Sparse and High-Dimensional Approximation

Lecture, SoSe 2018

based on manuscript by

G. Plonka, D. Potts, G. Steidl and M. Tasche

Chemnitz University of Technology, Germany

Outline

Multidimensional Fourier methods

We consider Fourier methods in more than one dimension d. We start with Fourier series of d-variate, 2π-periodic functions $f : \mathbb{T}^d \to \mathbb{C}$. In particular, we present basic properties of the Fourier coefficients and learn about their decay for smooth functions. Then we deal with Fourier transforms of functions on $\mathbb{R}^d$. We show that the Fourier transform is a linear, bijective operator on the Schwartz space $\mathcal{S}(\mathbb{R}^d)$ of rapidly decaying functions as well as on the space $\mathcal{S}'(\mathbb{R}^d)$ of tempered distributions. Using the density of $\mathcal{S}(\mathbb{R}^d)$ in $L_1(\mathbb{R}^d)$ and $L_2(\mathbb{R}^d)$, the Fourier transform on these spaces is discussed. The Poisson summation formula and the Fourier transforms of radial functions are also addressed. As in the univariate case, any numerical application of d-dimensional Fourier series or Fourier transforms leads to d-dimensional discrete Fourier transforms. We present the basic properties of the two-dimensional and higher dimensional DFT, including the convolution property and the aliasing formula.

Multidimensional Fourier series

We consider *d-variate*, 2π*-periodic functions* $f : \mathbb{R}^d \to \mathbb{C}$, i.e., functions fulfilling $f(\mathbf{x}) = f(\mathbf{x} + 2\pi\,\mathbf{k})$ for all $\mathbf{x} = (x_j)_{j=1}^d \in \mathbb{R}^d$ and all $\mathbf{k} = (k_j)_{j=1}^d \in \mathbb{R}^d$. Note that the function f is 2π-periodic in each variable x_j, $j = 1, \ldots, d$, and that f is uniquely determined by its restriction to the hypercube $[0,\, 2\pi)^d$. Hence f can be considered as a function defined on the d-dimensional torus $\mathbb{T}^d = \mathbb{R}^d/(2\pi\,\mathbb{Z}^d)$. For fixed $\mathbf{n} = (n_j)_{j=1}^d \in \mathbb{Z}^d$, the *d-variate complex exponential*

$$\mathrm{e}^{\mathrm{i}\,\mathbf{n}\cdot\mathbf{x}} = \prod_{j=1}^{d} \mathrm{e}^{\mathrm{i}\,n_j\,x_j}\,, \quad \mathbf{x} \in \mathbb{R}^d\,,$$

is 2π-periodic, where $\mathbf{n}\cdot\mathbf{x} := n_1\,x_1 + \ldots + n_d\,x_d$ is the inner product of $\mathbf{n} \in \mathbb{Z}^d$ and $\mathbf{x} \in \mathbb{R}^d$.

Further, we use the Euclidean norm $\|\mathbf{x}\|_2 := (\mathbf{x} \cdot \mathbf{x})^{1/2}$ of $\mathbf{x} \in \mathbb{R}^d$.
For a multi-index $\boldsymbol{\alpha} = (\alpha_k)_{k=1}^d \in \mathbb{N}_0^d$ with $|\boldsymbol{\alpha}| = \alpha_1 + \ldots + \alpha_d$, we use the notation

$$\mathbf{x}^{\boldsymbol{\alpha}} := \prod_{k=1}^{d} x_k^{\alpha_k}.$$

Let $C(\mathbb{T}^d)$ be the Banach space of continuous functions $f : \mathbb{T}^d \to \mathbb{C}$ equipped with the norm

$$\|f\|_{C(\mathbb{T}^d)} := \max_{\mathbf{x} \in \mathbb{T}^d} |f(\mathbf{x})| \,.$$

By $C^r(\mathbb{T}^d)$, $r \in \mathbb{N}$, we denote the Banach space of r-times continuously differentiable functions with the norm

$$\|f\|_{C^r(\mathbb{T}^d)} := \sum_{|\boldsymbol{\alpha}| \le r} \max_{\mathbf{x} \in \mathbb{T}^d} |D^{\boldsymbol{\alpha}} f(\mathbf{x})| \,,$$

where

$$D^{\boldsymbol{\alpha}} f(\mathbf{x}) := \frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}} \ldots \frac{\partial^{\alpha_d}}{\partial x_d^{\alpha_d}} \, f(\mathbf{x})$$

denotes the partial derivative with the multi-index $\boldsymbol{\alpha} = (\alpha_j)_{j=1}^d \in \mathbb{N}_0^d$ and $|\boldsymbol{\alpha}| \le r$.

For $1 \le p \le \infty$, let $L_p(\mathbb{T}^d)$ denote the Banach space of all measurable functions $f : \mathbb{T}^d \to \mathbb{C}$ with finite norm

$$\|f\|_{L_p(\mathbb{T}^d)} := \begin{cases} \left(\frac{1}{(2\pi)^d} \int_{[0,\,2\pi]^d} |f(\mathbf{x})|^p \,\mathrm{d}\mathbf{x}\right)^{1/p} & 1 \le p < \infty\,, \\ \operatorname{ess\,sup}\{|f(\mathbf{x})| : \mathbf{x} \in [0,\,2\pi]^d\} & p = \infty\,, \end{cases}$$

where almost everywhere equal functions are identified. The spaces $L_p(\mathbb{T}^d)$ with $1 < p < \infty$ are continuously embedded as

$$L_1(\mathbb{T}^d) \supset L_p(\mathbb{T}^d) \supset L_\infty(\mathbb{T}^d)\,.$$

By the periodicity of $f \in L_1(\mathbb{T}^d)$ we have

$$\int_{[0,\,2\pi]^d} f(\mathbf{x})\,\mathrm{d}\mathbf{x} = \int_{[-\pi,\,\pi]^d} f(\mathbf{x})\,\mathrm{d}\mathbf{x}\,.$$

For $p = 2$, we obtain the Hilbert space $L_2(\mathbb{T}^d)$ with the inner product and norm

$$\langle f, g \rangle_{L_2(\mathbb{T}^d)} := \frac{1}{(2\pi)^d} \int_{[0,\,2\pi]^d} f(\mathbf{x})\,\overline{g(\mathbf{x})}\,\mathrm{d}\mathbf{x}\,, \quad \|f\|_{L_2(\mathbb{T}^d)} := \sqrt{\langle f, f \rangle_{L_2(\mathbb{T}^d)}}$$

for arbitrary f, $g \in L_2(\mathbb{T}^d)$. For all f, $g \in L_2(\mathbb{T}^d)$ it holds the Cauchy–Schwarz inequality

$$|\langle f, g \rangle_{L_2(\mathbb{T}^d)}| \le \|f\|_{L_2(\mathbb{T}^d)}\, \|g\|_{L_2(\mathbb{T}^d)}\,.$$

The set of all complex exponentials $\left\{\mathrm{e}^{\mathrm{i}\,\mathbf{k}\cdot\mathbf{x}} : \mathbf{k} \in \mathbb{Z}^d\right\}$ forms an orthonormal basis of $L_2(\mathbb{T}^d)$. A linear combination of complex exponentials

$$p(\mathbf{x}) = \sum_{\mathbf{k}\in\mathbb{Z}^d} a_{\mathbf{k}}\,\mathrm{e}^{\mathrm{i}\,\mathbf{k}\cdot\mathbf{x}}$$

with only finitely many coefficients $a_{\mathbf{k}} \in \mathbb{C}\setminus\{0\}$ is called *d-variate*, 2π*-periodic trigonometric polynomial*. The *degree* of p is the largest number $\|\mathbf{k}\|_1 = |k_1| + \ldots + |k_d|$ such that $a_{\mathbf{k}} \neq 0$ with $\mathbf{k} = (k_j)_{j=1}^d \in \mathbb{Z}^d$. The set of all trigonometric polynomials is dense in $L_p(\mathbb{T}^d)$ for $1 \le p < \infty$ (see [16, p. 168]).
For $f \in L_1(\mathbb{T}^d)$ and arbitrary $\mathbf{k} \in \mathbb{Z}^d$, the $\mathbf{k}$th *Fourier coefficient* of f is defined as

$$c_{\mathbf{k}}(f) := \langle f(\mathbf{x}), \mathrm{e}^{\mathrm{i}\,\mathbf{k}\cdot\mathbf{x}}\rangle_{L_2(\mathbb{T}^d)} = \frac{1}{(2\pi)^d}\int_{[0,\,2\pi]^d} f(\mathbf{x})\,\mathrm{e}^{-\mathrm{i}\mathbf{k}\cdot\mathbf{x}}\,\mathrm{d}\mathbf{x}\,.$$

As in the univariate case, the $\mathbf{k}$th *modulus* and *phase* of f are defined by $|c_{\mathbf{k}}(f)|$ and $\arg c_{\mathbf{k}}(f)$, respectively. Obviously, we have

$$|c_{\mathbf{k}}(f)| \le \frac{1}{(2\pi)^d}\int_{[0,\,2\pi]^d} |f(\mathbf{x})|\,\mathrm{d}\mathbf{x} = \|f\|_{L_1(\mathbb{T}^d)}\,.$$

The Fourier coefficients possess similar properties as in the univariate setting.

Lemma 1

The Fourier coefficients of any functions $f,\, g \in L_1(\mathbb{T}^d)$ have the following properties for all $\mathbf{k} = (k_j)_{j=1}^d \in \mathbb{Z}^d$:

1. Uniqueness: *If $c_{\mathbf{k}}(f) = c_{\mathbf{k}}(g)$ for all $\mathbf{k} \in \mathbb{Z}^d$, then $f = g$ almost everywhere.*
2. Linearity: *For all $\alpha,\, \beta \in \mathbb{C}$,*

$$c_{\mathbf{k}}(\alpha\, f + \beta\, g) = \alpha\, c_{\mathbf{k}}(f) + \beta\, c_{\mathbf{k}}(g)\,.$$

3. Translation and modulation: *For all $\mathbf{x}_0 \in [0,\, 2\pi)^d$ and $\mathbf{k}_0 \in \mathbb{Z}^d$,*

$$c_{\mathbf{k}}\big(f(\mathbf{x} - \mathbf{x}_0)\big) = \mathrm{e}^{-\mathrm{i}\,\mathbf{k}\cdot\mathbf{x}_0}\, c_{\mathbf{k}}(f)\,,$$
$$c_{\mathbf{k}}\big(\mathrm{e}^{-\mathrm{i}\,\mathbf{k}_0\cdot\mathbf{x}}\, f(\mathbf{x})\big) = c_{\mathbf{k}+\mathbf{k}_0}(f)\,.$$

Lemma 1 (continue)

4 Differentiation: *For $f \in L_1(\mathbb{T}^d)$ with partial derivative $\frac{\partial f}{\partial x_j} \in L_1(\mathbb{T}^d)$,*

$$c_{\mathbf{k}}\Big(\frac{\partial f}{\partial x_j}\Big) = \mathrm{i}\, k_j\, c_{\mathbf{k}}(f)\,.$$

5 Convolution: *For $f, g \in L_1(\mathbb{T}^d)$, the d-variate convolution*

$$(f * g)(\mathbf{x}) := \frac{1}{(2\pi)^d} \int_{[0,\,2\pi]^d} f(\mathbf{y})\, g(\mathbf{x}-\mathbf{y})\, \mathrm{d}\mathbf{y}\,, \quad \mathbf{x} \in \mathbb{R}^d\,,$$

is contained in $L_1(\mathbb{T}^d)$ and we have

$$c_{\mathbf{k}}(f * g) = c_{\mathbf{k}}(f)\, c_{\mathbf{k}}(g)\,.$$

The proof of Lemma 1 can be given similarly as in the univariate case and is left to the reader.

Remark 2

The differentiation property 1 of Lemma 1 can be generalized. Assume that $f \in L_1(\mathbb{R}^d)$ *possesses partial derivatives* $D^{\boldsymbol{\alpha}} f \in L_1(\mathbb{T}^d)$ *for all multi-indices* $\boldsymbol{\alpha} \in \mathbb{N}_0^d$ *with* $|\boldsymbol{\alpha}| := \|\boldsymbol{\alpha}\|_1 \le r$*, where* $r \in \mathbb{N}$ *is fixed. Repeated application of the differentiation property 1 of Lemma 1 provides*

$$c_{\mathbf{k}}(D^{\boldsymbol{\alpha}} f) = (\mathrm{i}\,\mathbf{k})^{\boldsymbol{\alpha}}\, c_{\mathbf{k}}(f) \tag{1}$$

for all $\mathbf{k} \in \mathbb{Z}^d$*, where* $(\mathrm{i}\,\mathbf{k})^{\boldsymbol{\alpha}}$ *denotes the product* $(\mathrm{i}\,k_1)^{\alpha_1} \ldots (\mathrm{i}\,k_d)^{\alpha_d}$ *with the convention* $0^0 = 1$*.* □

Remark 3

If the 2π*-periodic function*

$$f(\mathbf{x}) = \prod_{j=1}^{d} f_j(x_j)$$

is the product of univariate functions $f_j \in L_1(\mathbb{T})$, $j = 1, \ldots, d$, *then we have for all* $\mathbf{k} = (k_j)_{j=1}^{d} \in \mathbb{Z}^d$

$$c_{\mathbf{k}}(f) = \prod_{j=1}^{d} c_{k_j}(f_j)\,. \quad \square$$

Example 4

Let $n \in \mathbb{N}_0$ be given. The nth *Dirichlet kernel* $D_n : \mathbb{T}^d \to \mathbb{C}$

$$D_n(\mathbf{x}) := \sum_{k_1=-n}^{n} \dots \sum_{k_d=-n}^{n} \mathrm{e}^{\mathrm{i}\,\mathbf{k}\cdot\mathbf{x}}$$

is a trigonometric polynomial of degree $d\,n$. It is the product of univariate nth Dirichlet kernels

$$D_n(\mathbf{x}) = \prod_{j=1}^{d} D_n(x_j)\,. \quad \square$$

For arbitrary $n \in \mathbb{N}_0$, the nth *Fourier partial sum* of $f \in L_1(\mathbb{T}^d)$ is defined by

$$(S_n f)(\mathbf{x}) := \sum_{k_1=-n}^{n} \ldots \sum_{k_d=-n}^{n} c_{\mathbf{k}}(f)\, \mathrm{e}^{\mathrm{i}\,\mathbf{k}\cdot\mathbf{x}}\,. \qquad (2)$$

Using the nth Dirichlet kernel D_n, the nth Fourier partial sum $S_n f$ can be represented as convolution $S_n f = f * D_n$.
For $f \in L_1(\mathbb{T}^d)$, the d-dimensional *Fourier series*

$$\sum_{\mathbf{k}\in\mathbb{Z}^d} c_{\mathbf{k}}(f)\, \mathrm{e}^{\mathrm{i}\,\mathbf{k}\cdot\mathbf{x}} \qquad (3)$$

is called *convergent* to f in $L_2(\mathbb{T}^d)$, if the sequence of Fourier partial sums (2) converges to f, i.e.,

$$\lim_{n\to\infty} \|f - S_n f\|_{L_2(\mathbb{T}^d)} = 0\,.$$

Then it holds the following result on convergence in $L_2(\mathbb{T}^d)$:

Theorem 5

Every function $f \in L_2(\mathbb{T}^d)$ can be expanded into the Fourier series (3) which converges to f in $L_2(\mathbb{T}^d)$. Further the Parseval equality

$$\|f\|^2_{L_2(\mathbb{T}^d)} = \frac{1}{(2\pi)^d} \int_{[0,\,2\pi]^d} |f(\mathbf{x})|^2 \,\mathrm{d}\mathbf{x} = \sum_{\mathbf{k}\in\mathbb{Z}^d} |c_{\mathbf{k}}(f)|^2 \qquad (4)$$

is fulfilled.

Now we investigate the relation between the smoothness of the function $f : \mathbb{T}^d \to \mathbb{C}$ and the decay of its Fourier coefficients $c_{\mathbf{k}}(f)$ as $\|\mathbf{k}\|_2 \to \infty$. We show that the smoother a function $f : \mathbb{T}^d \to \mathbb{C}$ is, the faster its Fourier coefficients $c_{\mathbf{k}}(f)$ tend to zero as $\|\mathbf{k}\|_2 \to \infty$ (cf. Lemma of Riemann-Lebesgue and Theorem of Bernstein for $d = 1$).

Lemma 6

1. For $f \in L_1(\mathbb{T}^d)$ we have

$$\lim_{\|\mathbf{k}\|_2 \to \infty} c_{\mathbf{k}}(f) = 0\,. \tag{5}$$

2. Let $r \in \mathbb{N}$ be given. If f and its partial derivatives $D^{\boldsymbol{\alpha}} f$ are contained in $L_1(\mathbb{T}^d)$ for all multi-indices $\boldsymbol{\alpha} \in \mathbb{N}_0^d$ with $|\boldsymbol{\alpha}| \le r$, then

$$\lim_{\|\mathbf{k}\|_2 \to \infty} \left(1 + \|\mathbf{k}\|_2^r\right) c_{\mathbf{k}}(f) = 0\,. \tag{6}$$

Proof:

1. If $f \in L_2(\mathbb{T}^d)$, then (5) is a consequence of the Parseval equality (4). For all $\varepsilon > 0$, any function $f \in L_1(\mathbb{T}^d)$ can be approximated by a trigonometric polynomial p of degree n such that $\|f - p\|_{L_1(\mathbb{T}^d)} < \varepsilon$. Then the Fourier coefficients of $r := f - p \in L_1(\mathbb{T}^d)$ fulfill $|c_{\mathbf{k}}(r)| \le \|r\|_{L_1(\mathbb{T}^d)} < \varepsilon$ for all $\mathbf{k} \in \mathbb{Z}^d$. Further we have $c_{\mathbf{k}}(p) = 0$ for all $\mathbf{k} \in \mathbb{Z}^d$ with $\|\mathbf{k}\|_1 > n$, since the trigonometric polynomial p has the degree n. By the linearity of the Fourier coefficients and by $\|\mathbf{k}\|_1 \ge \|\mathbf{k}\|_2$, we obtain for all $\mathbf{k} \in \mathbb{Z}^d$ with $\|\mathbf{k}\|_2 > n$ that

$$|c_{\mathbf{k}}(f)| = |c_{\mathbf{k}}(p) + c_{\mathbf{k}}(r)| = |c_{\mathbf{k}}(r)| < \varepsilon\,.$$

2. We consider a fixed multi-index $\mathbf{k} \in \mathbb{Z}^d \setminus \{\mathbf{0}\}$ with $|k_\ell| = \max_{j=1,\ldots,d} |k_j| > 0$. From (1) it follows that

$$(\mathrm{i}\, k_\ell)^r\, c_{\mathbf{k}}(f) = c_{\mathbf{k}}\big(\frac{\partial^r f}{\partial x_\ell^r}\big)\,.$$

Using $\|\mathbf{k}\|_2 \le \sqrt{d}\, |k_\ell|$, we obtain the estimate

$$\|\mathbf{k}\|_2^r\, |c_{\mathbf{k}}(f)| \le d^{r/2}\, |c_{\mathbf{k}}\big(\frac{\partial^r f}{\partial x_\ell^r}\big)| \le d^{r/2} \max_{|\boldsymbol{\alpha}|=r} |c_{\mathbf{k}}(D^{\boldsymbol{\alpha}} f)|\,.$$

Then from (5) it follows the assertion (6). ■

Now we consider the uniform convergence of d-dimensional Fourier series.

Theorem 7

If $f \in C(\mathbb{T}^d)$ has the property

$$\sum_{\mathbf{k}\in\mathbb{Z}^d} |c_{\mathbf{k}}(f)| < \infty\,, \tag{7}$$

then the d-dimensional Fourier series (3) *converges uniformly to f on $\mathbb{T}^d$, i.e.,*

$$\lim_{n\to\infty} \|f - S_n f\|_{C(\mathbb{T}^d)} = 0\,.$$

Proof: By (7), the Weierstrass criterion ensures that the Fourier series (3) converges uniformly to a continuous function

$$g(\mathbf{x}) := \sum_{\mathbf{k}\in\mathbb{Z}^d} c_{\mathbf{k}}(f)\,\mathrm{e}^{\mathrm{i}\,\mathbf{k}\cdot\mathbf{x}}.$$

Since f and g have the same Fourier coefficients, the uniqueness property in Lemma 1 gives $f = g$ on $\mathbb{T}^d$. ■

Now we want to show that a sufficiently smooth function $f : \mathbb{T}^d \to \mathbb{C}$ fulfills condition (7). We need the following result:

Lemma 8

If $2\,r > d$, then

$$\sum_{\mathbf{k} \in \mathbb{Z}^d \setminus \{\mathbf{0}\}} \|\mathbf{k}\|_2^{-2\,r} < \infty\,. \tag{8}$$

Proof: For all $\mathbf{k} = (k_j)_{j=1}^d \in \mathbb{Z}^d \setminus \{\mathbf{0}\}$ we have $\|\mathbf{k}\|_2 \ge 1$. Using the inequality of arithmetic and geometric means, it follows

$$(d+1)\,\|\mathbf{k}\|_2^2 \ge d + \|\mathbf{k}\|_2^2 = \sum_{j=1}^d (1+k_j^2) \ge d\,\Big(\prod_{j=1}^d (1+k_j^2)\Big)^{1/d}$$

and hence

$$\|\mathbf{k}\|_2^{-2\,r} \le \Big(\frac{d+1}{d}\Big)^r \prod_{j=1}^d (1+k_j^2)^{-r/d}\,.$$

Consequently, we obtain

$$\sum_{\mathbf{k}\in\mathbb{Z}^d\setminus\{\mathbf{0}\}} \|\mathbf{k}\|_2^{-2\,r} \le \Big(\frac{d+1}{d}\Big)^r \sum_{k_1\in\mathbb{Z}} (1+k_1^2)^{-r/d} \,\ldots \sum_{k_d\in\mathbb{Z}} (1+k_d^2)^{-r/d}$$

$$= \Big(\frac{d+1}{d}\Big)^r \Big(\sum_{k\in\mathbb{Z}} (1+k^2)^{-r/d}\Big)^d < \Big(\frac{d+1}{d}\Big)^r \Big(1 + 2\sum_{k=1}^\infty k^{-2\,r/d}\Big)^d < \infty.$$

■

Theorem 9

If $f \in C^r(\mathbb{T}^d)$ with $2\,r > d$, then the condition (7) is fulfilled and the d-dimensional Fourier series (3) converges uniformly to f on $\mathbb{T}^d$.

Proof: By assumption, each partial derivative $D^{\boldsymbol{\alpha}} f$ with $|\boldsymbol{\alpha}| \le r$ is continuous on $\mathbb{T}^d$. Hence we have $D^{\boldsymbol{\alpha}} f \in L_2(\mathbb{T}^d)$ such that by (1) and the Parseval equality (4),

$$\sum_{|\boldsymbol{\alpha}|=r} \sum_{\mathbf{k}\in\mathbb{Z}^d} |c_{\mathbf{k}}(f)|^2\, \mathbf{k}^{2\boldsymbol{\alpha}} < \infty\,,$$

where $\mathbf{k}^{2\boldsymbol{\alpha}}$ denotes the product $k_1^{2\alpha_1} \ldots k_d^{2\alpha_d}$ with $0^0 := 1$. Then there exists a positive constant c, depending only on the dimension d and on r, such that

$$\sum_{|\boldsymbol{\alpha}|=r} \mathbf{k}^{2\boldsymbol{\alpha}} \ge c\, \|\mathbf{k}\|_2^{2\,r}\,.$$

By the Cauchy–Schwarz inequality in $\ell_2(\mathbb{Z}^d)$ and by Lemma 8 we obtain

$$\sum_{\mathbf{k}\in\mathbb{Z}^d\setminus\{\mathbf{0}\}} |c_{\mathbf{k}}(f)| \le \sum_{\mathbf{k}\in\mathbb{Z}^d\setminus\{\mathbf{0}\}} |c_{\mathbf{k}}(f)| \Big(\sum_{|\boldsymbol{\alpha}|=r} \mathbf{k}^{2\boldsymbol{\alpha}}\Big)^{1/2} c^{-1/2}\, \|\mathbf{k}\|_2^{-r}$$

$$\le \Big(\sum_{|\boldsymbol{\alpha}|=r}\sum_{\mathbf{k}\in\mathbb{Z}^d} |c_{\mathbf{k}}(f)|^2\, \mathbf{k}^{2\boldsymbol{\alpha}}\Big)^{1/2} \Big(\sum_{\mathbf{k}\in\mathbb{Z}^d\setminus\{\mathbf{0}\}} \|\mathbf{k}\|_2^{-2r}\Big)^{1/2} c^{-1/2} < \infty\,. \qquad \blacksquare$$

Multidimensional Fourier transform

Let $C_0(\mathbb{R}^d)$ be the Banach space of all functions $f : \mathbb{R}^d \to \mathbb{C}$, which are continuous on $\mathbb{R}^d$ and vanish as $\|\mathbf{x}\|_2 \to \infty$, with norm

$$\|f\|_{C_0(\mathbb{R}^d)} := \max_{\mathbf{x} \in \mathbb{R}^d} |f(\mathbf{x})| .$$

Let $C_c(\mathbb{R}^d)$ be the subspace of all continuous functions with compact supports. By $C^r(\mathbb{R}^d)$, $r \in \mathbb{N} \cup \{\infty\}$, we denote the set of r-times continuously differentiable functions and by $C_c^r(\mathbb{R}^d)$ the set of r-times continuously differentiable functions with compact supports.

For $1 \le p \le \infty$, let $L_p(\mathbb{R}^d)$ be the Banach space of all measurable functions $f : \mathbb{R}^d \to \mathbb{C}$ with finite norm

$$\|f\|_{L_p(\mathbb{R}^d)} := \begin{cases} \left(\int_{\mathbb{R}^d} |f(\mathbf{x})|^p \, \mathrm{d}\mathbf{x} \right)^{1/p} & 1 \le p < \infty\,, \\ \operatorname{ess\,sup} \{|f(\mathbf{x})| : \mathbf{x} \in \mathbb{R}^d\} & p = \infty\,, \end{cases}$$

where almost everywhere equal functions are identified.
In particular, we are interested in the Hilbert space $L_2(\mathbb{R}^d)$ with inner product and norm

$$\langle f, g \rangle_{L_2(\mathbb{R}^d)} := \int_{\mathbb{R}^d} f(\mathbf{x}) \, \overline{g(\mathbf{x})} \, \mathrm{d}\mathbf{x}\,, \quad \|f\|_{L_2(\mathbb{R}^d)} := \left(\int_{\mathbb{R}^d} |f(\mathbf{x})|^2 \, \mathrm{d}\mathbf{x} \right)^{1/2}.$$

Fourier transform on $\mathcal{S}(\mathbb{R}^d)$

By $\mathcal{S}(\mathbb{R}^d)$, we denote the set of all functions $\varphi \in C^\infty(\mathbb{R}^d)$ with the property $\mathbf{x}^{\boldsymbol{\alpha}} D^{\boldsymbol{\beta}} \varphi(\mathbf{x}) \in C_0(\mathbb{R}^d)$ for all multi-indices $\boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathbb{N}_0^d$.

We define the *convergence in* $\mathcal{S}(\mathbb{R}^d)$ as follows:

A sequence $(\varphi_k)_{k\in\mathbb{N}}$ of functions $\varphi_k \in \mathcal{S}(\mathbb{R}^d)$ converges to $\varphi \in \mathcal{S}(\mathbb{R}^d)$, if for all multi-indices $\boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathbb{N}_0^d$, the sequences $\left(\mathbf{x}^{\boldsymbol{\alpha}} D^{\boldsymbol{\beta}} \varphi_k\right)_{k\in\mathbb{N}}$ converge uniformly to $\mathbf{x}^{\boldsymbol{\alpha}} D^{\boldsymbol{\beta}} \varphi$ on $\mathbb{R}^d$.
We will write $\varphi_k \underset{\mathcal{S}}{\to} \varphi$ as $k \to \infty$.

Then the linear space $\mathcal{S}(\mathbb{R}^d)$ with this convergence is called *Schwartz space* or *space of rapidly decreasing functions*. The name is in honor of the French mathematician L. Schwartz (1915 – 2002).

Any function $\varphi \in \mathcal{S}(\mathbb{R}^d)$ is *rapidly decreasing* in the sense that for all multi-indices $\boldsymbol{\alpha}$, $\boldsymbol{\beta} \in \mathbb{N}_0^d$,

$$\lim_{\|\mathbf{x}\|_2 \to \infty} \mathbf{x}^{\boldsymbol{\alpha}} \, D^{\boldsymbol{\beta}} \varphi(\mathbf{x}) = 0 \, .$$

Introducing the seminorms

$$\|\varphi\|_m := \max_{|\boldsymbol{\beta}| \le m} \|(1 + \|\mathbf{x}\|_2)^m \, D^{\boldsymbol{\beta}} \varphi(\mathbf{x})\|_{C_0(\mathbb{R}^d)} \, , \quad m \in \mathbb{N}_0 \, , \qquad (9)$$

we see that $\|\varphi\|_0 \le \|\varphi\|_1 \le \|\varphi\|_2 \le \ldots$ for $\varphi \in \mathcal{S}(\mathbb{R}^d)$. Then we can describe the convergence in the Schwartz space by means of the seminorms (9):

Lemma 10

For $\varphi_k, \varphi \in \mathcal{S}(\mathbb{R}^d)$, we have $\varphi_k \underset{\mathcal{S}}{\longrightarrow} \varphi$ as $k \to \infty$ if and only if for all $m \in \mathbb{N}_0$,

$$\lim_{k\to\infty} \|\varphi_k - \varphi\|_m = 0\,. \tag{10}$$

Proof:
1. Let (10) be fulfilled for all $m \in \mathbb{N}_0$. Then for all $\boldsymbol{\alpha} = (\alpha_j)_{j=1}^d \in \mathbb{N}_0^d \setminus \{\mathbf{0}\}$ with $|\boldsymbol{\alpha}| \le m$, we get by the relation between geometric and quadratic means that

$$|\mathbf{x}^{\boldsymbol{\alpha}}| \le \left(\frac{\alpha_1 x_1^2 + \ldots + \alpha_d x_d^2}{|\boldsymbol{\alpha}|}\right)^{|\boldsymbol{\alpha}|/2} \le \left(x_1^2 + \ldots + x_d^2\right)^{|\boldsymbol{\alpha}|/2} \le (1+\|\mathbf{x}\|_2)^m$$

so that

$$|\mathbf{x}^{\boldsymbol{\alpha}} D^{\boldsymbol{\beta}}(\varphi_k - \varphi)(\mathbf{x})| \le (1 + \|\mathbf{x}\|_2)^m \, |D^{\boldsymbol{\beta}}(\varphi_k - \varphi)(\mathbf{x})|.$$

Hence, for all $\boldsymbol{\beta} \in \mathbb{N}_0^d$ with $|\boldsymbol{\beta}| \le m$, it holds

$$\|\mathbf{x}^{\boldsymbol{\alpha}} D^{\boldsymbol{\beta}}(\varphi_k - \varphi)(\mathbf{x})\|_{C_0(\mathbb{R}^d)} \le \sup_{\mathbf{x}\in\mathbb{R}^d} (1+\|\mathbf{x}\|_2)^m |D^{\boldsymbol{\beta}}(\varphi_k - \varphi)(\mathbf{x})| \le \|\varphi_k - \varphi\|_m.$$

2. Assume that $\varphi_k \underset{\mathcal{S}}{\rightarrow} \varphi$ as $k \to \infty$, i.e., for all $\boldsymbol{\alpha}$, $\boldsymbol{\beta} \in \mathbb{N}_0^d$ we have

$$\lim_{k\to\infty} \|\mathbf{x}^{\boldsymbol{\alpha}} D^{\boldsymbol{\beta}}(\varphi_k - \varphi)(\mathbf{x})\|_{C_0(\mathbb{R}^d)} = 0\,.$$

We consider multi-indices $\boldsymbol{\alpha}$, $\boldsymbol{\beta} \in \mathbb{N}_0^d$ with $|\boldsymbol{\alpha}| \le m$ and $|\boldsymbol{\beta}| \le m$ for $m \in \mathbb{N}$. Since x^m is convex, we use

$$\Big(\frac{1+x}{2}\Big)^m \le 1 + x^m$$

and obtain for $x = \|\mathbf{x}\|_2$, with $\mathbf{x} \in \mathbb{R}^d$,

$$(1 + \|\mathbf{x}\|_2)^m \le 2^m(1 + \|\mathbf{x}\|_2^m)\,.$$

Since $\sum_{j=1}^d |x_j| \ge \|\mathbf{x}\|_2$ and

$$d^{(1/2-1/m)}\,\Big(\sum_{j=1}^d |x_j|^m\Big)^{1/m} \ge \|\mathbf{x}\|_2$$

for $m \ge 2$ (see e.g. [19, formula (6.4)]), we see that $\sum_{j=1}^d |x_j|^m \ge c\,\|\mathbf{x}\|_2^m$ for all $\mathbf{x} \in \mathbb{R}^d$ with some positive constant $c \le 1$.

Hence we obtain

$$(1+\|\mathbf{x}\|_2)^m \leq 2^m(1+\|\mathbf{x}\|_2^m) \leq 2^m\Big(1+\frac{1}{c}\sum_{j=1}^{d}|x_j|^m\Big) \leq \frac{2^m}{c}\sum_{|\boldsymbol{\alpha}|\leq m}|\mathbf{x}^{\boldsymbol{\alpha}}|\,. \tag{11}$$

This implies that

$$\|(1+\|\mathbf{x}\|_2)^m\, D^{\boldsymbol{\beta}}(\varphi_k-\varphi)(\mathbf{x})\|_{C_0(\mathbb{R}^d)} \leq \frac{2^m}{c}\sum_{|\boldsymbol{\alpha}|\leq m}\|\mathbf{x}^{\boldsymbol{\alpha}}\, D^{\boldsymbol{\beta}}(\varphi_k-\varphi)(\mathbf{x})\|_{C_0(\mathbb{R}^d)}$$

and hence

$$\|\varphi_k-\varphi\|_m \leq \frac{2^m}{c}\max_{|\boldsymbol{\beta}|\leq m}\sum_{|\boldsymbol{\alpha}|\leq m}\|\mathbf{x}^{\boldsymbol{\alpha}}\, D^{\boldsymbol{\beta}}(\varphi_k-\varphi)(\mathbf{x})\|_{C_0(\mathbb{R}^d)}$$

such that $\lim_{k\to\infty}\|\varphi_k-\varphi\|_m = 0$. ■

Remark 11

The Schwartz space $\mathcal{S}(\mathbb{R}^d)$ *is a complete metric space with the metric*

$$\rho(\varphi,\psi) := \sum_{m=0}^{\infty} \frac{1}{2^m} \frac{\|\varphi-\psi\|_m}{1+\|\varphi-\psi\|_m}\,, \quad \varphi,\,\psi \in \mathcal{S}(\mathbb{R}^d)\,,$$

since by Lemma 10 the convergence $\varphi_k \underset{\mathcal{S}}{\to} \varphi$ *as* $k \to \infty$ *is equivalent to*

$$\lim_{k\to\infty} \rho(\varphi_k,\,\varphi) = 0\,.$$

This metric space is complete by the following reason: Let $(\varphi_k)_{k\in\mathbb{N}}$ *be a Cauchy sequence with respect to* ρ. *Then, for every* $\boldsymbol{\alpha}$, $\boldsymbol{\beta} \in \mathbb{N}_0^d$, $\left(x^{\boldsymbol{\alpha}}\, D^{\boldsymbol{\beta}}\varphi_k\right)_{k\in\mathbb{N}}$ *is a Cauchy sequence in Banach space* $C_0(\mathbb{R}^d)$ *and converges uniformly to a function* $\psi_{\boldsymbol{\alpha},\boldsymbol{\beta}}$. *Then, by definition of* $\mathcal{S}(\mathbb{R}^d)$, *it follows* $\psi_{\boldsymbol{\alpha},\boldsymbol{\beta}}(\mathbf{x}) = \mathbf{x}^{\boldsymbol{\alpha}}\, D^{\boldsymbol{\beta}}\psi_{\mathbf{0},\mathbf{0}}(\mathbf{x})$ *with* $\psi_{\mathbf{0},\mathbf{0}} \in \mathcal{S}(\mathbb{R}^d)$ *and hence* $\varphi_k \underset{\mathcal{S}}{\to} \psi_{\mathbf{0},\mathbf{0}}$ *as* $k \to \infty$. *Note that the metric* ρ *is not generated by a norm, since* $\rho(c\,\varphi,\,0) \neq |c|\,\rho(\varphi,\,0)$ *for all* $c \in \mathbb{C}\setminus\{0\}$ *with* $|c| \neq 1$ *and non-vanishing* $\varphi \in \mathcal{S}(\mathbb{R}^d)$. $\square$

Clearly, it holds $\mathcal{S}(\mathbb{R}^d) \subset C_0(\mathbb{R}^d)$ and $\mathcal{S}(\mathbb{R}^d) \subset L_p(\mathbb{R}^d)$, $1 \le p < \infty$, by the following argument: For each $\varphi \in \mathcal{S}(\mathbb{R}^d)$ we have by (9)

$$|\varphi(\mathbf{x})| \le \|\varphi\|_{d+1} \, (1 + \|\mathbf{x}\|_2)^{-d-1}$$

for all $x \in \mathbb{R}^d$. Then, using polar coordinates with $r = \|\mathbf{x}\|_2$, we obtain

$$\int_{\mathbb{R}^d} |\varphi(\mathbf{x})|^p \, \mathrm{d}\mathbf{x} \le \|\varphi\|_{d+1}^p \int_{\mathbb{R}^d} (1 + \|\mathbf{x}\|_2)^{-p(d+1)} \, \mathrm{d}\mathbf{x}$$
$$\le C \int_0^\infty \frac{r^{d-1}}{(1+r)^{p(d+1)}} \, \mathrm{d}r \le C \int_0^\infty \frac{1}{(1+r)^2} \, \mathrm{d}r < \infty$$

with some constant $C > 0$. Hence the Schwartz space $\mathcal{S}(\mathbb{R}^d)$ is contained in $L_1(\mathbb{R}^d) \cap L_2(\mathbb{R}^d)$.

Obviously, $C_c^\infty(\mathbb{R}^d) \subset \mathcal{S}(\mathbb{R}^d)$. Since $C_c^\infty(\mathbb{R}^d)$ is dense in $L_p(\mathbb{R}^d)$, $p \in [1, \infty)$, see e.g. [64, Satz 3.6], we also have that $\mathcal{S}(\mathbb{R}^d)$ is dense in $L_p(\mathbb{R}^d)$, $p \in [1, \infty)$. Summarizing we find that

$$C_c^\infty(\mathbb{R}^d) \subset \mathcal{S}(\mathbb{R}^d) \subset C_0^\infty(\mathbb{R}^d) \subset C^\infty(\mathbb{R}^d) \,. \tag{12}$$

Example 12

A typical function in $C_c^\infty(\mathbb{R}^d) \subset \mathcal{S}(\mathbb{R}^d)$ is the test function

$$\varphi(\mathbf{x}) := \begin{cases} \exp\left(-\frac{1}{1-\|\mathbf{x}\|_2^2}\right) & \|\mathbf{x}\|_2 < 1\,, \\ 0 & \|\mathbf{x}\|_2 \geq 1\,. \end{cases} \tag{13}$$

The compact support of φ is the unit ball $\{\mathbf{x} \in \mathbb{R}^d : \|\mathbf{x}\|_2 \leq 1\}$.
Any Gaussian function $\mathrm{e}^{-a\,\|\mathbf{x}\|_2^2}$ with $a > 0$ is contained in $\mathcal{S}(\mathbb{R}^d)$, but it is not in $C_c^\infty(\mathbb{R}^d)$.
For any $n \in \mathbb{N}$, the function

$$f(\mathbf{x}) := (1 + \|\mathbf{x}\|_2^2)^{-n} \in C_0^\infty(\mathbb{R}^d)$$

does not belong to $\mathcal{S}(\mathbb{R}^d)$, since $\|\mathbf{x}\|_2^{2n}\, f(\mathbf{x})$ does not tend to zero as $\|\mathbf{x}\|_2 \to \infty$. $\square$

Example 13

In the univariate case, each product of a polynomial and the Gaussian function $\mathrm{e}^{-x^2/2}$ is a rapidly decreasing function. The Hermite functions $h_n(x) = H_n(x)\,\mathrm{e}^{-x^2/2}$, $n \in \mathbb{N}_0$, are contained in $\mathcal{S}(\mathbb{R})$ and form an orthogonal basis of $L_2(\mathbb{R})$ (see lecture last year). Here H_n denotes the nth Hermite polynomial. Thus $\mathcal{S}(\mathbb{R})$ is dense in $L_2(\mathbb{R})$. For each multi-index $\mathbf{n} = (n_j)_{j=1}^d \in \mathbb{N}_0^d$, the function $\mathbf{x}^{\mathbf{n}}\,\mathrm{e}^{-\|\mathbf{x}\|_2^2/2}$, $\mathbf{x} = (x_j)_{j=1}^d \in \mathbb{R}^d$, is a rapidly decreasing function. The set of all functions

$$h_{\mathbf{n}}(\mathbf{x}) := \mathrm{e}^{-\|\mathbf{x}\|_2^2/2} \prod_{j=1}^{d} H_{n_j}(x_j) \in \mathcal{S}(\mathbb{R}^d)\,, \quad \mathbf{n} \in \mathbb{N}_0^d\,,$$

is an orthogonal basis of $L_2(\mathbb{R}^d)$. Further $\mathcal{S}(\mathbb{R}^d)$ is dense in $L_2(\mathbb{R}^d)$. $\square$

For $f \in L_1(\mathbb{R}^d)$ we define its *Fourier transform* at $\boldsymbol{\omega} \in \mathbb{R}^d$ by

$$\mathcal{F}f(\boldsymbol{\omega}) = \hat{f}(\boldsymbol{\omega}) := \int_{\mathbb{R}^d} f(\mathbf{x})\, \mathrm{e}^{-\mathrm{i}\,\mathbf{x}\cdot\boldsymbol{\omega}}\, \mathrm{d}\mathbf{x}\,. \tag{14}$$

Since

$$|\hat{f}(\boldsymbol{\omega})| \le \int_{\mathbb{R}^d} |f(\mathbf{x})|\, \mathrm{d}\mathbf{x} = \|f\|_{L_1(\mathbb{R}^d)}\,,$$

the Fourier transform (14) exists for all $\boldsymbol{\omega} \in \mathbb{R}^d$ and is bounded on $\mathbb{R}^d$.

Example 14

Let $L > 0$ be given. The characteristic function $f(\mathbf{x})$ of the hypercube $[-L,\, L]^d \subset \mathbb{R}^d$ is the product $\prod_{j=1}^d \chi_{[-L,\,L]}(x_j)$ of univariate characteristic functions. The related Fourier transform reads as follows

$$\hat{f}(\boldsymbol{\omega}) = (2L)^d \prod_{j=1}^{d} \operatorname{sinc}(L\omega_j)\,. \quad \square$$

Example 15

The Gaussian function $f(\mathbf{x}) := \mathrm{e}^{-\|\sigma\mathbf{x}\|_2^2/2}$ with fixed $\sigma > 0$ is the product of the univariate functions $f(x_j) = \mathrm{e}^{-\sigma^2 x_j^2/2}$ such that

$$\hat{f}(\boldsymbol{\omega}) = \left(\frac{2\pi}{\sigma^2}\right)^{d/2} \mathrm{e}^{-\|\boldsymbol{\omega}\|_2^2/(2\sigma^2)} . \quad \square$$

By the following theorem the Fourier transform maps the Schwartz space $\mathcal{S}(\mathbb{R}^d)$ into itself.

Theorem 16

For every $\varphi \in \mathcal{S}(\mathbb{R}^d)$, it holds $\mathcal{F}\varphi \in \mathcal{S}(\mathbb{R}^d)$, i.e., $\mathcal{F} : \mathcal{S}(\mathbb{R}^d) \to \mathcal{S}(\mathbb{R}^d)$. Furthermore, $D^{\boldsymbol{\alpha}}(\mathcal{F}\varphi) \in \mathcal{S}(\mathbb{R}^d)$ and $\mathcal{F}(D^{\boldsymbol{\alpha}}\varphi) \in \mathcal{S}(\mathbb{R}^d)$ for all $\boldsymbol{\alpha} \in \mathbb{N}_0^d$, and we have

$$D^{\boldsymbol{\alpha}}(\mathcal{F}\varphi) = (-\mathrm{i})^{|\boldsymbol{\alpha}|}\, \mathcal{F}(\mathbf{x}^{\boldsymbol{\alpha}}\, \varphi)\,, \tag{15}$$

$$\boldsymbol{\omega}^{\boldsymbol{\alpha}}\, (\mathcal{F}\varphi) = (-\mathrm{i})^{|\boldsymbol{\alpha}|}\, \mathcal{F}(D^{\boldsymbol{\alpha}}\, \varphi)\,. \tag{16}$$

where the partial derivative $D^{\boldsymbol{\alpha}}$ in (15) acts on $\boldsymbol{\omega}$ and in (16) on $\mathbf{x}$.

Proof: 1. Let $\boldsymbol{\alpha} \in \mathbb{N}_0^d$ be an arbitrary multi-index with $|\boldsymbol{\alpha}| \le m$. By definition, each rapidly decreasing function $\varphi \in \mathcal{S}(\mathbb{R}^d)$ has the property

$$\lim_{\|\mathbf{x}\|_2 \to \infty} \varphi(\mathbf{x})\,(1 + \|\mathbf{x}\|_2)^{m+d+1} = 0\,.$$

Therefore we can change the order of differentiation and integration in $D^{\boldsymbol{\alpha}}(\mathcal{F}\varphi)$ such that

$$D^{\boldsymbol{\alpha}}(\mathcal{F}\varphi)(\boldsymbol{\omega}) = \int_{\mathbb{R}^d} (-\mathrm{i}\,\mathbf{x})^{\boldsymbol{\alpha}}\,\varphi(\mathbf{x})\,\mathrm{e}^{-\mathrm{i}\,\mathbf{x}\cdot\boldsymbol{\omega}}\,\mathrm{d}\mathbf{x} = (-\mathrm{i})^{|\boldsymbol{\alpha}|}\,\mathcal{F}(\mathbf{x}^{\boldsymbol{\alpha}}\,\varphi)(\boldsymbol{\omega})\,.$$

Note that $\mathbf{x}^{\boldsymbol{\alpha}}\,\varphi \in \mathcal{S}(\mathbb{R}^d)$. Thus $\mathcal{F}\varphi$ belongs to $C^\infty(\mathbb{R}^d)$.

2. For simplicity, we show (16) only for $\boldsymbol{\alpha} = \mathbf{e}_1 = (\delta_{j-1})_{j=1}^d$. From the theorem of Fubini it follows that

$$\omega_1\,(\mathcal{F}\varphi)(\boldsymbol{\omega}) = \int_{\mathbb{R}^d} \omega_1\,\mathrm{e}^{-\mathrm{i}\,\mathbf{x}\cdot\boldsymbol{\omega}}\,\varphi(\mathbf{x})\,\mathrm{d}\mathbf{x}$$

$$= \int_{\mathbb{R}^{d-1}} \exp\Big(-\mathrm{i}\sum_{j=2}^{d} x_j\omega_j\Big)\Big(\int_{\mathbb{R}} \omega_1\,\mathrm{e}^{-\mathrm{i}\,x_1\omega_1}\,\varphi(\mathbf{x})\,\mathrm{d}x_1\Big)\,\mathrm{d}x_2\ldots\mathrm{d}x_d\,.$$

For the inner integral, integration by parts yields

$$\int_{\mathbb{R}} \omega_1\,\mathrm{e}^{-\mathrm{i}\,x_1\omega_1}\,\varphi(\mathbf{x})\,\mathrm{d}x_1 = \lim_{r\to\infty}\int_{-r}^{r} \mathrm{i}\,\frac{\mathrm{d}}{\mathrm{d}x_1}\big(\mathrm{e}^{-\mathrm{i}x_1\omega_1}\big)\,\varphi(\mathbf{x})\,\mathrm{d}x_1$$

$$= \lim_{r\to\infty}\Big(\mathrm{i}\,\mathrm{e}^{-\mathrm{i}\,x_1\omega_1}\,\varphi(\mathbf{x})\Big|_{x_1=-r}^{x_1=r} - \mathrm{i}\int_{-r}^{r} \mathrm{e}^{-\mathrm{i}\,x_1\omega_1}\,D^{\mathbf{e}_1}\varphi(\mathbf{x})\,\mathrm{d}x_1\Big)$$

$$= 0 - \mathrm{i}\int_{\mathbb{R}} \mathrm{e}^{-\mathrm{i}\,x_1\omega_1}\,D^{\mathbf{e}_1}\varphi(\mathbf{x})\,\mathrm{d}x_1\,.$$

Thus we obtain

$$\omega_1\,(\mathcal{F}\varphi)(\boldsymbol{\omega}) = -\mathrm{i}\,\mathcal{F}(D^{\mathbf{e}_1}\varphi)(\boldsymbol{\omega})\,.$$

For an arbitrary multi-index $\boldsymbol{\alpha} \in \mathbb{N}_0^d$, the formula (16) follows by induction.

3. From (15) and (16) it follows for all multi-indices $\boldsymbol{\alpha}$, $\boldsymbol{\beta} \in \mathbb{N}_0^d$ and each $\varphi \in \mathcal{S}(\mathbb{R}^d)$,

$$\boldsymbol{\omega}^{\boldsymbol{\alpha}}\,[D^{\boldsymbol{\beta}}(\mathcal{F}\varphi)] = (-\mathrm{i})^{|\boldsymbol{\beta}|}\,\boldsymbol{\omega}^{\boldsymbol{\alpha}}\mathcal{F}(\mathbf{x}^{\boldsymbol{\beta}}\varphi) = (-\mathrm{i})^{|\boldsymbol{\alpha}|+|\boldsymbol{\beta}|}\,\mathcal{F}[D^{\boldsymbol{\alpha}}(\mathbf{x}^{\boldsymbol{\beta}}\varphi)]\,. \tag{17}$$

Hence $\boldsymbol{\omega}^{\boldsymbol{\alpha}}\,[D^{\boldsymbol{\beta}}(\mathcal{F}\varphi)](\boldsymbol{\omega})$ is uniformly bounded on $\mathbb{R}^d$, since

$$|\boldsymbol{\omega}^{\boldsymbol{\alpha}}\,[D^{\boldsymbol{\beta}}(\mathcal{F}\varphi)](\boldsymbol{\omega})| = |\mathcal{F}[D^{\boldsymbol{\alpha}}(\mathbf{x}^{\boldsymbol{\beta}}\varphi)](\boldsymbol{\omega})| \le \int_{\mathbb{R}^d} |D^{\boldsymbol{\alpha}}(\mathbf{x}^{\boldsymbol{\beta}}\varphi)|\,\mathrm{d}x < \infty\,.$$

Thus we see that $\mathcal{F}\varphi \in \mathcal{S}(\mathbb{R}^d)$. ■

Based on the above theorem we can show that the Fourier transform is indeed a bijection on $\mathcal{S}(\mathbb{R}^d)$.

Theorem 17

The Fourier transform $\mathcal{F} : \mathcal{S}(\mathbb{R}^d) \to \mathcal{S}(\mathbb{R}^d)$ *is a linear, bijective mapping. Further the Fourier transform is continuous with respect to the convergence in* $\mathcal{S}(\mathbb{R}^d)$*, i.e., for* $\varphi_k,\, \varphi \in \mathcal{S}(\mathbb{R}^d)$*,* $\varphi_k \underset{\mathcal{S}}{\longrightarrow} \varphi$ *as* $k \to \infty$ *implies* $\mathcal{F}\varphi_k \underset{\mathcal{S}}{\longrightarrow} \mathcal{F}\varphi$ *as* $k \to \infty$*. For all* $\varphi \in \mathcal{S}(\mathbb{R}^d)$ *and all* $\mathbf{x} \in \mathbb{R}^d$*, the inverse Fourier transform* $\mathcal{F}^{-1} : \mathcal{S}(\mathbb{R}^d) \to \mathcal{S}(\mathbb{R}^d)$ *is given by*

$$(\mathcal{F}^{-1}\varphi)(\mathbf{x}) := \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \varphi(\boldsymbol{\omega})\, \mathrm{e}^{\mathrm{i}\,\mathbf{x}\cdot\boldsymbol{\omega}}\, \mathrm{d}\boldsymbol{\omega}\,. \tag{18}$$

The inverse Fourier transform is also a linear, bijective mapping on $\mathcal{S}(\mathbb{R}^d)$ *which is continuous with respect to the convergence in* $\mathcal{S}(\mathbb{R}^d)$*. Further for all* $\varphi \in \mathcal{S}(\mathbb{R}^d)$ *and all* $\mathbf{x} \in \mathbb{R}^d$ *it holds the Fourier inversion formula*

$$\varphi(\mathbf{x}) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} (\mathcal{F}\varphi)(\boldsymbol{\omega})\, \mathrm{e}^{\mathrm{i}\,\mathbf{x}\cdot\boldsymbol{\omega}}\, \mathrm{d}\boldsymbol{\omega}\,.$$

Proof: 1. By Theorem 16 the Fourier transform $\mathcal{F}$ maps the Schwartz space $\mathcal{S}(\mathbb{R}^d)$ into itself. The linearity of the Fourier transform $\mathcal{F}$ follows from those of the integral operator (14). For arbitrary $\varphi \in \mathcal{S}(\mathbb{R}^d)$, for all $\boldsymbol{\alpha}$, $\boldsymbol{\beta} \in \mathbb{N}_0^d$ with $|\boldsymbol{\alpha}| \le m$ and $|\boldsymbol{\beta}| \le m$, and for all $\boldsymbol{\omega} \in \mathbb{R}^d$ we obtain by (17)

$$|\boldsymbol{\omega}^{\boldsymbol{\beta}}\, D^{\boldsymbol{\alpha}}(\mathcal{F}\varphi)(\boldsymbol{\omega})| = |\mathcal{F}\big(D^{\boldsymbol{\beta}}(\mathbf{x}^{\boldsymbol{\alpha}}\,\varphi(\mathbf{x}))\big)(\boldsymbol{\omega})| \le \int_{\mathbb{R}^d} |D^{\boldsymbol{\beta}}(\mathbf{x}^{\boldsymbol{\alpha}}\,\varphi(\mathbf{x}))|\,\mathrm{d}\mathbf{x}$$

$$\le C \int_{\mathbb{R}^d} (1 + \|\mathbf{x}\|_2)^m \sum_{|\boldsymbol{\gamma}| \le m} |D^{\boldsymbol{\gamma}}\varphi(\mathbf{x})|\,\mathrm{d}\mathbf{x}$$

$$\le C \int_{\mathbb{R}^d} \frac{(1 + \|\mathbf{x}\|_2)^{m+d+1}}{(1 + \|\mathbf{x}\|_2)^{d+1}} \sum_{|\boldsymbol{\gamma}| \le m} |D^{\boldsymbol{\gamma}}\varphi(\mathbf{x})|\,\mathrm{d}\mathbf{x}$$

$$\le C \int_{\mathbb{R}^d} \frac{\mathrm{d}\mathbf{x}}{(1 + \|\mathbf{x}\|_2)^{d+1}}\, \|\varphi\|_{m+d+1}\,.$$

By

$$\|\mathcal{F}\varphi\|_m = \max_{|\gamma|\le m} \|(1+\|\boldsymbol{\omega}\|_2)^m\, D^{\gamma}\mathcal{F}\varphi(\boldsymbol{\omega})\|_{C_0(\mathbb{R}^d)}$$

we see that

$$\|\mathcal{F}\varphi\|_m \le C'\,\|\varphi\|_{m+d+1} \tag{19}$$

for all $\varphi \in \mathcal{S}(\mathbb{R}^d)$ and each $m \in \mathbb{N}_0$, where $C' > 0$ is a constant. Now we show the continuity of the Fourier transform. Assume that $\varphi_k \underset{\mathcal{S}}{\longrightarrow} \varphi$ as $k \to \infty$ for $\varphi_k,\, \varphi \in \mathcal{S}(\mathbb{R}^d)$. Applying the inequality (19) to $\varphi_k - \varphi$, we obtain for all $m \in \mathbb{N}_0$

$$\|\mathcal{F}\varphi_k - \mathcal{F}\varphi\|_m \le C'\,\|\varphi_k - \varphi\|_{m+d+1}\,.$$

From Lemma 10 it follows that $\mathcal{F}\varphi_k \underset{\mathcal{S}}{\longrightarrow} \mathcal{F}\varphi$ as $k \to \infty$.

2. The mapping

$$(\tilde{\mathcal{F}}\varphi)(\mathbf{x}) := \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \varphi(\boldsymbol{\omega})\, \mathrm{e}^{\mathrm{i}\,\mathbf{x}\cdot\boldsymbol{\omega}}\, \mathrm{d}\boldsymbol{\omega}\,, \quad \varphi \in \mathcal{S}(\mathbb{R}^d)\,,$$

is a linear continuous mapping on $\mathcal{S}(\mathbb{R}^d)$ into itself by the first step of this proof, since $(\tilde{\mathcal{F}}\varphi)(\mathbf{x}) = \frac{1}{(2\pi)^d}\,(\mathcal{F}\varphi)(-\mathbf{x})$.

Now we demonstrate that $\tilde{\mathcal{F}}$ is the inverse mapping of $\mathcal{F}$ For arbitrary φ, $\psi \in \mathcal{S}(\mathbb{R}^d)$ it holds by Fubini's theorem

$$\int_{\mathbb{R}^d} (\mathcal{F}\varphi)(\boldsymbol{\omega})\, \psi(\boldsymbol{\omega})\, \mathrm{e}^{\mathrm{i}\,\boldsymbol{\omega}\cdot\mathbf{x}}\, \mathrm{d}\boldsymbol{\omega} = \int_{\mathbb{R}^d} \Big(\int_{\mathbb{R}^d} \varphi(\mathbf{y})\, \mathrm{e}^{-\mathrm{i}\,\boldsymbol{\omega}\cdot\mathbf{y}}\, \mathrm{d}\mathbf{y}\Big)\, \psi(\boldsymbol{\omega})\, \mathrm{e}^{\mathrm{i}\,\boldsymbol{\omega}\cdot\mathbf{x}}\, \mathrm{d}\boldsymbol{\omega}$$
$$= \int_{\mathbb{R}^d} \varphi(\mathbf{y}) \Big(\int_{\mathbb{R}^d} \psi(\boldsymbol{\omega})\, \mathrm{e}^{\mathrm{i}\,(\mathbf{x}-\mathbf{y})\cdot\boldsymbol{\omega}}\, \mathrm{d}\boldsymbol{\omega}\Big) \mathrm{d}\mathbf{y}$$
$$= \int_{\mathbb{R}^d} \varphi(\mathbf{y})\, (\mathcal{F}\psi)(\mathbf{y}-\mathbf{x})\, \mathrm{d}\mathbf{y} = \int_{\mathbb{R}^d} \varphi(\mathbf{z}+\mathbf{x})\, (\mathcal{F}\psi)(\mathbf{z})\, \mathrm{d}\mathbf{z}\,.$$

For the Gaussian function $\psi(\mathbf{x}) := \mathrm{e}^{-\|\varepsilon\mathbf{x}\|_2^2/2}$ with $\varepsilon > 0$, we have by Example 15 that $(\mathcal{F}\psi)(\boldsymbol{\omega}) = \left(\frac{2\pi}{\varepsilon^2}\right)^{d/2} \mathrm{e}^{-\|\boldsymbol{\omega}\|_2^2/(2\varepsilon^2)}$ and consequently

$$\int_{\mathbb{R}^d} (\mathcal{F}\varphi)(\boldsymbol{\omega})\; \mathrm{e}^{-\|\varepsilon\boldsymbol{\omega}\|_2^2/2} \mathrm{e}^{\mathrm{i}\,\boldsymbol{\omega}\cdot\mathbf{x}}\, \mathrm{d}\boldsymbol{\omega} = \Big(\frac{2\pi}{\varepsilon^2}\Big)^{d/2} \int_{\mathbb{R}^d} \varphi(\mathbf{z}+\mathbf{x})\, \mathrm{e}^{-\|\mathbf{z}\|_2^2/(2\varepsilon^2)}\, \mathrm{d}\mathbf{z}$$
$$= (2\pi)^{d/2} \int_{\mathbb{R}^d} \varphi(\varepsilon\, \mathbf{y}+\mathbf{x})\, \mathrm{e}^{-\|\mathbf{y}\|_2^2/2}\, \mathrm{d}\mathbf{y}\,.$$

Since $|(\mathcal{F}\varphi)(\boldsymbol{\omega})\,\mathrm{e}^{-\|\varepsilon\boldsymbol{\omega}\|_2^2/2}| \le |\mathcal{F}\varphi(\boldsymbol{\omega})|$ for all $\boldsymbol{\omega} \in \mathbb{R}^d$ and $\mathcal{F}\varphi \in \mathcal{S}(\mathbb{R}^d) \subset L_1(\mathbb{R}^d)$, we obtain by Lebesgue's dominated convergence theorem

$$\begin{aligned}\big(\tilde{\mathcal{F}}(\mathcal{F}\varphi)\big)(\mathbf{x}) &= \frac{1}{(2\pi)^d}\,\lim_{\varepsilon\to 0}\int_{\mathbb{R}^d}(\mathcal{F}\varphi)(\boldsymbol{\omega})\,\mathrm{e}^{-\|\varepsilon\mathbf{y}\|_2^2/2}\;\mathrm{e}^{\mathrm{i}\,\boldsymbol{\omega}\cdot\mathbf{x}}\,\mathrm{d}\boldsymbol{\omega} \\ &= (2\pi)^{-d/2}\lim_{\varepsilon\to 0}\int_{\mathbb{R}^d}\varphi(\mathbf{x}+\varepsilon\mathbf{y})\,\mathrm{e}^{-\|\mathbf{y}\|_2^2/2}\,\mathrm{d}\mathbf{y} \\ &= (2\pi)^{-d/2}\varphi(\mathbf{x})\int_{\mathbb{R}^d}\mathrm{e}^{-\|\mathbf{y}\|_2^2/2}\,\mathrm{d}\mathbf{y} = \varphi(\mathbf{x})\,,\end{aligned}$$

since by the Fourier transform of the Gaussian function

$$\int_{\mathbb{R}^d}\mathrm{e}^{-\|\mathbf{y}\|_2^2/2}\,\mathrm{d}\mathbf{y} = \Big(\int_{\mathbb{R}}\mathrm{e}^{y^2/2}\,\mathrm{d}y\Big)^d = (2\pi)^{d/2}\,.$$

From $\tilde{\mathcal{F}}(\mathcal{F}\varphi) = \varphi$ it follows immediately that $\mathcal{F}(\tilde{\mathcal{F}}\varphi) = \varphi$ for all $\varphi \in \mathcal{S}(\mathbb{R}^d)$. Hence, $\tilde{\mathcal{F}} = \mathcal{F}^{-1}$ and $\mathcal{F}$ is bijective. ■

The *convolution* $f * g$ of two d-variate functions $f, g \in L_1(\mathbb{R}^d)$ is defined by

$$(f * g)(\mathbf{x}) := \int_{\mathbb{R}^d} f(\mathbf{y})\, g(\mathbf{x} - \mathbf{y})\, \mathrm{d}\mathbf{y}\,.$$

The convolution theorem of Young carries over to the multivariate setting. Moreover, by the following lemma the product and the convolution of two rapidly decreasing functions are again rapidly decreasing.

Lemma 18

For arbitrary $\varphi, \psi \in \mathcal{S}(\mathbb{R}^d)$, the product $\varphi\,\psi$ and the convolution $\varphi * \psi$ are in $\mathcal{S}(\mathbb{R}^d)$ too and it holds $\mathcal{F}(\varphi * \psi) = \hat{\varphi}\,\hat{\psi}$.

Proof: 1. By the Leibniz' formula

$$D^{\boldsymbol{\alpha}}(\varphi\,\psi) = \sum_{\boldsymbol{\beta}\le\boldsymbol{\alpha}} \binom{\boldsymbol{\alpha}}{\boldsymbol{\beta}} (D^{\boldsymbol{\beta}}\varphi)\,(D^{\boldsymbol{\alpha}-\boldsymbol{\beta}}\psi)$$

with $\boldsymbol{\alpha} = (\alpha_j)_{j=1}^d \in \mathbb{N}_0^d$, where the sum runs over all $\boldsymbol{\beta} = (\beta_j)_{j=1}^d \in \mathbb{N}_0^d$ with $\beta_j \le \alpha_j$ for $j = 1, \ldots, d$, and where

$$\binom{\boldsymbol{\alpha}}{\boldsymbol{\beta}} := \frac{\alpha_1! \ldots \alpha_d!}{\beta_1! \ldots \beta_d!\,(\alpha_1-\beta_1)! \ldots (\alpha_d-\beta_d)!}\,,$$

we obtain that $\mathbf{x}^{\boldsymbol{\gamma}}\,D^{\boldsymbol{\alpha}}\big(\varphi(\mathbf{x})\,\psi(\mathbf{x})\big) \in C_0(\mathbb{R}^d)$ for all $\boldsymbol{\alpha}$, $\boldsymbol{\gamma} \in \mathbb{N}_0^d$, i.e., $\varphi\,\psi \in \mathcal{S}(\mathbb{R}^d)$.
2. By Theorem 17, we know that $\hat\varphi$, $\hat\psi \in \mathcal{S}(\mathbb{R}^d)$ and hence $\hat\varphi\,\hat\psi \in \mathcal{S}(\mathbb{R}^d)$ by the first step. Using Theorem 17, we obtain that $\mathcal{F}(\hat\varphi\,\hat\psi) \in \mathcal{S}(\mathbb{R}^d)$.

Otherwise we receive by Fubini's theorem

$$\begin{aligned}\mathcal{F}(\varphi * \psi)(\boldsymbol{\omega}) &= \int_{\mathbb{R}^d} \Big(\int_{\mathbb{R}^d} \varphi(\mathbf{y})\, \psi(\mathbf{x}-\mathbf{y})\, \mathrm{d}\mathbf{y}\Big)\, \mathrm{e}^{-\mathrm{i}\,\mathbf{x}\cdot\boldsymbol{\omega}}\, \mathrm{d}\mathbf{x} \\ &= \int_{\mathbb{R}^d} \varphi(\mathbf{y})\, \mathrm{e}^{-\mathrm{i}\,\mathbf{y}\cdot\boldsymbol{\omega}} \Big(\int_{\mathbb{R}^d} \psi(\mathbf{x}-\mathbf{y})\, \mathrm{e}^{-\mathrm{i}\,(\mathbf{x}-\mathbf{y})\cdot\boldsymbol{\omega}}\, \mathrm{d}\mathbf{x}\Big)\, \mathrm{d}\mathbf{y} \\ &= \Big(\int_{\mathbb{R}^d} \varphi(\mathbf{y})\, \mathrm{e}^{-\mathrm{i}\,\mathbf{y}\cdot\boldsymbol{\omega}}\, \mathrm{d}\mathbf{y}\Big)\, \hat{\psi}(\boldsymbol{\omega}) = \hat{\varphi}(\boldsymbol{\omega})\, \hat{\psi}(\boldsymbol{\omega})\,.\end{aligned}$$

Therefore $\varphi * \psi = \mathcal{F}^{-1}(\hat{\varphi}\,\hat{\psi}) \in \mathcal{S}(\mathbb{R}^d)$. ■

The basic properties of the d-variate Fourier transform on $\mathcal{S}(\mathbb{R}^d)$ can be proved similarly as in the univariate case. The following properties 1, 3, and 4 hold also true for functions in $L_1(\mathbb{R}^d)$, whereas property 2 holds only under additional smoothness assumptions.

Properties of the Fourier transform on $\mathcal{S}(\mathbb{R}^d)$

Theorem 19

The Fourier transform of a function $\varphi \in \mathcal{S}(\mathbb{R}^d)$ has the following properties:

1. Translation and modulation: *For fixed $\mathbf{x}_0, \boldsymbol{\omega}_0 \in \mathbb{R}^d$,*

$$\big(\varphi(\mathbf{x}-\mathbf{x}_0)\big)\hat{}(\boldsymbol{\omega}) = \mathrm{e}^{-\mathrm{i}\,\mathbf{x}_0\cdot\boldsymbol{\omega}}\,\hat{\varphi}(\boldsymbol{\omega})\,,$$
$$\big(\mathrm{e}^{-\mathrm{i}\,\boldsymbol{\omega}_0\cdot\mathbf{x}}\,\varphi(\mathbf{x})\big)\hat{}(\boldsymbol{\omega}) = \hat{\varphi}(\boldsymbol{\omega}+\boldsymbol{\omega}_0)\,.$$

2. Differentiation and multiplication: *For $\boldsymbol{\alpha} \in \mathbb{N}_0^d$,*

$$\big(D^{\boldsymbol{\alpha}}\varphi(\mathbf{x})\big)\hat{}(\boldsymbol{\omega}) = \mathrm{i}^{|\boldsymbol{\alpha}|}\,\boldsymbol{\omega}^{\boldsymbol{\alpha}}\,\hat{\varphi}(\boldsymbol{\omega})\,,$$
$$\big(\mathbf{x}^{\boldsymbol{\alpha}}\varphi(\mathbf{x})\big)\hat{}(\boldsymbol{\omega}) = \mathrm{i}^{|\boldsymbol{\alpha}|}\,(D^{\boldsymbol{\alpha}}\hat{\varphi})(\boldsymbol{\omega})\,.$$

Theorem 19 (continue)

3. Scaling: *For* $c \in \mathbb{R} \setminus \{0\}$,

$$\big(\varphi(c\,\mathbf{x})\big)\hat{}(\boldsymbol{\omega}) = \frac{1}{|c|^d}\,\hat{\varphi}(c^{-1}\,\boldsymbol{\omega})\,.$$

4. Convolution: *For* φ, $\psi \in \mathcal{S}(\mathbb{R}^d)$,

$$(\varphi * \psi)\hat{}(\boldsymbol{\omega}) = \hat{\varphi}(\boldsymbol{\omega})\,\hat{\psi}(\boldsymbol{\omega})\,.$$

Fourier transform on $L_1(\mathbb{R}^d)$ and $L_2(\mathbb{R}^d)$

Similar to the univariate case, we obtain the following theorem for the Fourier transform on $L_1(\mathbb{R}^d)$.

Theorem 20

The Fourier transform $\mathcal{F}$ defined by (14) is a linear continuous operator from $L_1(\mathbb{R}^d)$ into $C_0(\mathbb{R}^d)$ with the operator norm $\|\mathcal{F}\|_{L_1(\mathbb{R}^d)\to C_0(\mathbb{R}^d)} = 1$.

Proof: By (12) there exists for any $f \in L_1(\mathbb{R}^d)$ a sequence $(\varphi_k)_{k\in\mathbb{N}}$ with $\varphi_k \in \mathcal{S}(\mathbb{R}^d)$ such that $\lim_{k\to\infty} \|f - \varphi_k\|_{L_1(\mathbb{R}^d)} = 0$. Then the $C_0(\mathbb{R}^d)$ norm of $\mathcal{F}f - \mathcal{F}\varphi_k$ can be estimated by

$$\|\mathcal{F}f - \mathcal{F}\varphi_k\|_{C_0(\mathbb{R}^d)} = \max_{\boldsymbol{\omega}\in\mathbb{R}^d} |\mathcal{F}(f - \varphi_k)(\boldsymbol{\omega})| \leq \|f - \varphi_k\|_{L_1(\mathbb{R}^d)},$$

i.e., $\lim_{k\to\infty} \mathcal{F}\varphi_k = \mathcal{F}f$ in the norm of $C_0(\mathbb{R}^d)$. By $\mathcal{S}(\mathbb{R}^d) \subset C_0(\mathbb{R}^d)$ and the completeness of $C_0(\mathbb{R}^d)$ we conclude that $\mathcal{F}f \in C_0(\mathbb{R}^d)$. The operator norm of $\mathcal{F} : L_1(\mathbb{R}^d) \to C_0(\mathbb{R}^d)$ can be deduced as in the univariate case, where we have just to use the d-variate Gaussian function. ■

Theorem 21 (Fourier inversion formula for $L_1(\mathbb{R}^d)$ functions)

Let $f \in L_1(\mathbb{R}^d)$ and $\hat{f} \in L_1(\mathbb{R}^d)$. Then the Fourier inversion formula

$$f(\mathbf{x}) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \hat{f}(\boldsymbol{\omega})\, \mathrm{e}^{\mathrm{i}\boldsymbol{\omega}\cdot\mathbf{x}}\, \mathrm{d}\boldsymbol{\omega} \tag{20}$$

holds true for almost all $\mathbf{x} \in \mathbb{R}^d$.

The proof follows similar lines as those of Theorem in the univariate case. Another proof of Theorem 21 is sketched in Remark 42.

The following lemma is related to the more general Lemma proved in the univariate case.

Lemma 22

For arbitrary $\varphi,\ \psi \in \mathcal{S}(\mathbb{R}^d)$*, the following* Parseval equality *is valid:*

$$(2\pi)^d \, \langle \varphi, \psi \rangle_{L_2(\mathbb{R}^d)} = \langle \mathcal{F}\varphi, \mathcal{F}\psi \rangle_{L_2(\mathbb{R}^d)} \, .$$

In particular, we have $(2\pi)^{d/2} \, \|\varphi\|_{L_2(\mathbb{R}^d)} = \|\mathcal{F}\varphi\|_{L_2(\mathbb{R}^d)}$.

Proof: By Theorem 17 we have $\varphi = \mathcal{F}^{-1}(\mathcal{F}\varphi)$ for $\varphi \in \mathcal{S}(\mathbb{R}^d)$. Then Fubini's theorem yields

$$\begin{aligned}
(2\pi)^d \, \langle \varphi,\, \psi \rangle_{L_2(\mathbb{R}^d)} &= (2\pi)^d \int_{\mathbb{R}^d} \varphi(\mathbf{x}) \, \overline{\psi(\mathbf{x})} \, \mathrm{d}\mathbf{x} \\
&= \int_{\mathbb{R}^d} \overline{\psi(\mathbf{x})} \Big(\int_{\mathbb{R}^d} (\mathcal{F}\varphi)(\boldsymbol{\omega}) \, \mathrm{e}^{\mathrm{i}\,\mathbf{x}\cdot\boldsymbol{\omega}} \, \mathrm{d}\boldsymbol{\omega} \Big) \, \mathrm{d}\mathbf{x} \\
&= \int_{\mathbb{R}^d} (\mathcal{F}\varphi)(\boldsymbol{\omega}) \overline{\int_{\mathbb{R}^d} \psi(\mathbf{x}) \, \mathrm{e}^{-\mathrm{i}\,\mathbf{x}\cdot\boldsymbol{\omega}} \, \mathrm{d}\mathbf{x}} \, \mathrm{d}\boldsymbol{\omega} \\
&= \int_{\mathbb{R}^d} \mathcal{F}\varphi(\boldsymbol{\omega}) \, \overline{\mathcal{F}\psi(\boldsymbol{\omega})} \, \mathrm{d}\boldsymbol{\omega} = \langle \mathcal{F}\varphi,\, \mathcal{F}\psi \rangle_{L_2(\mathbb{R}^d)} \, . \qquad \blacksquare
\end{aligned}$$

We will use the following extension theorem of bounded linear operator, see e.g. [1, Theorem 2.4.1], to extend the Fourier transform from $\mathcal{S}(\mathbb{R}^d)$ to $L_2(\mathbb{R}^d)$.

Theorem 23 (Extension of a bounded linear operator)

Let H be a Hilbert space and let $D \subset H$ be a linear subset which is dense in H. Further let $F : D \to H$ be a linear bounded operator. Then F admits a unique extension to a bounded linear operator $\tilde{F} : H \to H$ with equal operator norms

$$\|F\|_{D\to H} = \|\tilde{F}\|_{H\to H}\,.$$

For each $f \in H$ with $f = \lim_{k\to\infty} f_k$, where $f_k \in D$, it holds $\tilde{F}f = \lim_{k\to\infty} Ff_k$.

Theorem 24 (Plancherel)

The Fourier transform $\mathcal{F} : \mathcal{S}(\mathbb{R}^d) \to \mathcal{S}(\mathbb{R}^d)$ *can be uniquely extended to a linear continuous bijective transform* $\mathcal{F} : L_2(\mathbb{R}^d) \to L_2(\mathbb{R}^d)$, *which fulfills the Parseval equality*

$$(2\pi)^d \langle f, g \rangle_{L_2(\mathbb{R}^d)} = \langle \mathcal{F}f, \mathcal{F}g \rangle_{L_2(\mathbb{R}^d)} \tag{21}$$

for all $f, g \in L_2(\mathbb{R}^d)$. *In particular, it holds* $(2\pi)^{d/2} \|f\|_{L_2(\mathbb{R}^d)} = \|\mathcal{F}f\|_{L_2(\mathbb{R}^d)}$.

The above extension is also called *Fourier transform* on $L_2(\mathbb{R}^d)$ or sometimes *Fourier–Plancherel transform*.

Proof: We consider $D = \mathcal{S}(\mathbb{R}^d)$ as linear, dense subspace of the Hilbert space $H = L_2(\mathbb{R}^d)$. By Lemma 22 we know that $\mathcal{F}$ as well as $\mathcal{F}^{-1}$ are bounded linear operators from D to H with the operator norms $(2\pi)^{d/2}$ and $(2\pi)^{-d/2}$. Therefore both operators admit a unique extensions $\mathcal{F} : L_2(\mathbb{R}^d) \to L_2(\mathbb{R}^d)$ and $\mathcal{F}^{-1} : L_2(\mathbb{R}^d) \to L_2(\mathbb{R}^d)$ and (21) is fulfilled. ■

Fourier transforms of radial functions

A function $f : \mathbb{R}^d \to \mathbb{C}$ is called a *radial function*, if $f(\mathbf{x}) = f(\mathbf{y})$ for all $\mathbf{x}$, $\mathbf{y} \in \mathbb{R}^d$ with $\|\mathbf{x}\|_2 = \|\mathbf{y}\|_2$. Thus a radial function f can be written in the form $f(\mathbf{x}) = F(\|\mathbf{x}\|_2)$ with certain univariate function $F : [0, \infty) \to \mathbb{C}$. A radial function f is characterized by the property $f(\mathbf{A}\,\mathbf{x}) = f(\mathbf{x})$ for all orthogonal matrices $\mathbf{A} \in \mathbb{R}^{d \times d}$. The Gaussian function in Example 15 is a typical example of a radial function.

Lemma 25

Let $\mathbf{A} \in \mathbb{R}^{d\times d}$ *be invertible and let* $f \in L_1(\mathbb{R}^d)$*. Then we have*

$$\big(f(\mathbf{A}\,\mathbf{x})\big)\hat{}(\boldsymbol{\omega}) = \frac{1}{|\det \mathbf{A}|}\,\hat{f}(\mathbf{A}^{-\top}\boldsymbol{\omega})\,.$$

In particular, for an orthogonal matrix $\mathbf{A} \in \mathbb{R}^{d\times d}$ *we have the relation*

$$\big(f(\mathbf{A}\,\mathbf{x})\big)\hat{}(\boldsymbol{\omega}) = \hat{f}(\mathbf{A}\boldsymbol{\omega})\,.$$

Proof: Substituting $\mathbf{y} := \mathbf{A}\,\mathbf{x}$, it follows

$$\begin{aligned}
\big(f(\mathbf{A}\,\mathbf{x})\big)\hat{}(\boldsymbol{\omega}) &= \int_{\mathbb{R}^d} f(\mathbf{A}\mathbf{x})\,\mathrm{e}^{-\mathrm{i}\,\boldsymbol{\omega}\cdot\mathbf{x}}\,\mathrm{d}\mathbf{x} \\
&= \frac{1}{|\det \mathbf{A}|}\int_{\mathbb{R}^d} f(\mathbf{y})\,\mathrm{e}^{-\mathrm{i}\,(\mathbf{A}^{-\top}\boldsymbol{\omega})\cdot\mathbf{y}}\,\mathrm{d}\mathbf{y} = \frac{1}{|\det \mathbf{A}|}\,\hat{f}(\mathbf{A}^{-\top}\boldsymbol{\omega})\,.
\end{aligned}$$

If $\mathbf{A}$ is orthogonal, then $\mathbf{A}^{-\top} = \mathbf{A}$ and $|\det \mathbf{A}| = 1$. ■

Corollary 26

If $f \in L_1(\mathbb{R}^d)$ *is a radial function of the form* $f(\mathbf{x}) = F(r)$ *with* $r := \|\mathbf{x}\|_2$*, then its Fourier transform* $\hat{f}$ *is a radial function too. In the case* $d = 2$*, we have*

$$\hat{f}(\boldsymbol{\omega}) = 2\pi \int_0^\infty F(r)\, J_0(r\, \|\boldsymbol{\omega}\|_2)\, r \,\mathrm{d}r\,, \tag{22}$$

where J_0 *denotes the* Bessel function of order zero

$$J_0(x) := \sum_{k=0}^{\infty} \frac{(-1)^k}{(k!)^2} \left(\frac{x}{2}\right)^{2k}\,.$$

Proof: The first assertion is an immediate consequence of Lemma 25. Let $d = 2$. Using polar coordinates (r, φ) and (ρ, ψ) with $r = \|\mathbf{x}\|_2$, $\rho = \|\boldsymbol{\omega}\|_2$ and $\varphi, \psi \in [0, 2\pi)$ such that

$$\mathbf{x} = (r \cos\varphi, r \sin\varphi)^\top, \quad \boldsymbol{\omega} = (\rho \cos\psi, \rho \sin\psi)^\top,$$

we obtain

$$\begin{aligned}\hat{f}(\boldsymbol{\omega}) &= \int_{\mathbb{R}^2} f(\mathbf{x})\, \mathrm{e}^{-\mathrm{i}\, \mathbf{x}\cdot\boldsymbol{\omega}}\, \mathrm{d}\mathbf{x} \\ &= \int_0^\infty \int_0^{2\pi} F(r)\, \mathrm{e}^{-\mathrm{i}\, r\rho \cos(\varphi-\psi)}\, r\, \mathrm{d}\varphi\, \mathrm{d}r\,.\end{aligned}$$

The inner integral with respect to φ is independent of ψ, since the integrand is 2π-periodic. For $\psi = -\frac{\pi}{2}$ we conclude by Bessel's integral formula

$$\int_0^{2\pi} \mathrm{e}^{-\mathrm{i}\, r\rho \cos(\varphi+\pi/2)}\, \mathrm{d}\varphi = \int_0^{2\pi} \mathrm{e}^{\mathrm{i}\, r\rho \sin\varphi}\, \mathrm{d}\varphi = 2\pi\, J_0(r\rho)\,.$$

This yields the integral representation (22) which is called *Hankel transform of order zero* of F. ■

Remark 27

In the case $d = 3$, *we can use spherical coordinates for the computation of the Fourier transform of a radial function* $f \in L_1(\mathbb{R}^3)$, *where* $f(\mathbf{x}) = F(\|\mathbf{x}\|_2)$. *This results in*

$$\hat{f}(\boldsymbol{\omega}) = \frac{4\pi}{\|\boldsymbol{\omega}\|_2} \int_0^\infty F(r)\, r \sin(r\,\|\boldsymbol{\omega}\|_2)\,\mathrm{d}r\,, \quad \boldsymbol{\omega} \in \mathbb{R}^3 \setminus \{\mathbf{0}\}\,. \tag{23}$$

For an arbitrary dimension $d \in \mathbb{N} \setminus \{1\}$, *we obtain for* $\boldsymbol{\omega} \in \mathbb{R}^d \setminus \{\mathbf{0}\}$

$$\hat{f}(\boldsymbol{\omega}) = (2\pi)^{d/2}\, \|\boldsymbol{\omega}\|_2{}^{1-d/2} \int_0^\infty F(r)\, r^{d/2-1}\, J_{d/2-1}(r\,\|\boldsymbol{\omega}\|_2)\,\mathrm{d}r\,,$$

where

$$J_\nu(x) := \sum_{k=0}^{\infty} \frac{(-1)^k}{k!\,\Gamma(k+\nu+1)} \left(\frac{x}{2}\right)^{2k+\nu}$$

denotes the Bessel function of order $\nu \geq 0$, *see [61, p. 155].* □

Example 28

Let $f : \mathbb{R}^2 \to \mathbb{R}$ be the characteristic function of the unit disk, i.e., $f(\mathbf{x}) := 1$ for $\|\mathbf{x}\|_2 \le 1$ and $f(\mathbf{x}) := 0$ for $\|\mathbf{x}\|_2 > 1$. By (22) it follows for $\boldsymbol{\omega} \in \mathbb{R}^2 \setminus \{\mathbf{0}\}$ that

$$\hat{f}(\boldsymbol{\omega}) = 2\pi \int_0^1 J_0(r\,\|\boldsymbol{\omega}\|_2)\, r\, \mathrm{d}r = \frac{2\pi}{\|\boldsymbol{\omega}\|_2}\, J_1(\|\boldsymbol{\omega}\|_2)$$

and $\hat{f}(\mathbf{0}) = \pi$.

Let $f : \mathbb{R}^3 \to \mathbb{R}$ be the characteristic function of the unit ball. Then from (23) it follows for $\boldsymbol{\omega} \in \mathbb{R}^3 \setminus \{\mathbf{0}\}$ that

$$\hat{f}(\boldsymbol{\omega}) = \frac{4\pi}{\|\boldsymbol{\omega}\|_2^3} \left(\sin \|\boldsymbol{\omega}\|_2 - \|\boldsymbol{\omega}\|_2 \cos \|\boldsymbol{\omega}\|_2 \right),$$

and in particular $\hat{f}(\mathbf{0}) = \frac{4\pi}{3}$. □

Poisson summation formula

Now we generalize the one-dimensional Poisson summation formula. For $f \in L_1(\mathbb{R}^d)$ we introduce its 2π–*periodization* by

$$\tilde{f}(\mathbf{x}) := \sum_{\mathbf{k} \in \mathbb{Z}^d} f(\mathbf{x} + 2\pi\mathbf{k}), \quad \mathbf{x} \in \mathbb{R}^d. \tag{24}$$

First we prove the existence of the 2π–periodization $\tilde{f} \in L_1(\mathbb{T}^d)$ of $f \in L_1(\mathbb{R}^d)$.

Lemma 29

For given $f \in L_1(\mathbb{R}^d)$, the series in (24) converges absolutely for almost all $\mathbf{x} \in \mathbb{R}^d$ and $\tilde{f}$ is contained in $L_1(\mathbb{T}^d)$.

Proof: At first we show that the 2π-periodization φ of $|f|$ belongs to $L_1(\mathbb{T}^d)$, i.e.

$$\varphi(\mathbf{x}) := \sum_{\mathbf{k}\in\mathbb{Z}^d} |f(\mathbf{x} + 2\pi\,\mathbf{k})|\,.$$

For each $n \in \mathbb{N}$, we form the nonnegative function

$$\varphi_n(\mathbf{x}) := \sum_{k_1=-n}^{n-1} \ldots \sum_{k_d=-n}^{n-1} |f(\mathbf{x} + 2\pi\,\mathbf{k})|\,.$$

Then we obtain

$$\begin{aligned}
\int_{[0,\,2\pi]^d} \varphi_n(\mathbf{x})\,\mathrm{d}\mathbf{x} &= \sum_{k_1=-n}^{n-1} \ldots \sum_{k_d=-n}^{n-1} \int_{[0,\,2\pi]^d} |f(\mathbf{x} + 2\pi\,\mathbf{k})|\,\mathrm{d}\mathbf{x} \\
&= \sum_{k_1=-n}^{n-1} \ldots \sum_{k_d=-n}^{n-1} \int_{2\pi\mathbf{k}+[0,\,2\pi]^d} |f(\mathbf{x})|\,\mathrm{d}\mathbf{x} \\
&= \int_{[-2\pi n,\,2\pi n]^d} |f(\mathbf{x})|\,\mathrm{d}\mathbf{x}
\end{aligned}$$

and hence

$$\lim_{n\to\infty} \int_{[0,\,2\pi]^d} \varphi_n(\mathbf{x})\,\mathrm{d}\mathbf{x} = \int_{\mathbb{R}^d} |f(\mathbf{x})|\,\mathrm{d}\mathbf{x} = \|f\|_{L_1(\mathbb{R}^d)} < \infty\,. \qquad (25)$$

Since $(\varphi_n)_{n\in\mathbb{N}}$ is a monotone increasing sequence of nonnegative integrable functions with the property (25), we receive by the monotone convergence theorem of B. Levi that $\lim_{n\to\infty} \varphi(\mathbf{x}) = \varphi_n(\mathbf{x})$ for almost all $\mathbf{x} \in \mathbb{R}^d$ and $\varphi \in L_1(\mathbb{T}^d)$, where it holds

$$\int_{[0,\,2\pi]^d} \varphi(\mathbf{x})\,\mathrm{d}\mathbf{x} = \lim_{n\to\infty} \int_{[0,\,2\pi]^d} \varphi_n(\mathbf{x})\,\mathrm{d}\mathbf{x} = \|f\|_{L_1(\mathbb{R}^d)}\,.$$

In other words, the series in (24) converges absolutely for almost all $\mathbf{x} \in \mathbb{R}^d$. From

$$|\tilde{f}(\mathbf{x})| = \Big|\sum_{\mathbf{k}\in\mathbb{Z}^d} f(\mathbf{x} + 2\pi\mathbf{k})\Big| \le \sum_{\mathbf{k}\in\mathbb{Z}^d} |f(\mathbf{x} + 2\pi\mathbf{k})| = \varphi(\mathbf{x})\,,$$

it follows that $\tilde{f} \in L_1(\mathbb{T}^d)$ with

$$\|\tilde{f}\|_{L_1(\mathbb{T}^d)} = \int_{[0,\,2\pi]^d} |\tilde{f}(\mathbf{x})|\,\mathrm{d}\mathbf{x} \le \int_{[0,\,2\pi]^d} \varphi(\mathbf{x})\,\mathrm{d}\mathbf{x} = \|f\|_{L_1(\mathbb{R})^d}\,. \qquad \blacksquare$$

The d-dimensional Poisson summation formula describes an interesting connection between the values $\hat{f}(\mathbf{n})$, $\mathbf{n} \in \mathbb{Z}^d$, of the Fourier transform $\hat{f}$ of a given function $f \in L_1(\mathbb{R}^d) \cap C_0(\mathbb{R}^d)$ and the Fourier series of the 2π-periodization $\tilde{f}$.

Theorem 30

Let $f \in C_0(\mathbb{R}^d)$ be a given function which fulfills the decay conditions

$$|f(\mathbf{x})| \le \frac{c}{1 + \|\mathbf{x}\|_2^{d+\varepsilon}}\,, \quad |\hat{f}(\boldsymbol{\omega})| \le \frac{c}{1 + \|\boldsymbol{\omega}\|_2^{d+\varepsilon}} \tag{26}$$

for all $\mathbf{x}$, $\boldsymbol{\omega} \in \mathbb{R}^d$ with some constants $\varepsilon > 0$ and $c > 0$.
Then for all $\mathbf{x} \in \mathbb{R}^d$, it holds the Poisson summation formula

$$(2\pi)^d\, \tilde{f}(\mathbf{x}) = (2\pi)^d \sum_{\mathbf{k} \in \mathbb{Z}^d} f(\mathbf{x} + 2\pi\,\mathbf{k}) = \sum_{\mathbf{n} \in \mathbb{Z}^d} \hat{f}(\mathbf{n})\, \mathrm{e}^{\mathrm{i}\,\mathbf{n}\cdot\mathbf{x}}\,, \tag{27}$$

where both series in (27) converge absolutely for all $\mathbf{x} \in \mathbb{R}^d$. In particular, for $\mathbf{x} = \mathbf{0}$ it holds

$$(2\pi)^d \sum_{\mathbf{k} \in \mathbb{Z}^d} f(2\pi\,\mathbf{k}) = \sum_{\mathbf{n} \in \mathbb{Z}^d} \hat{f}(\mathbf{n})\,.$$

Proof: From the decay conditions (26) it follows that f, $\hat{f} \in L_1(\mathbb{R}^d)$ such that $\tilde{f} \in L_1(\mathbb{T}^d)$ by Lemma 29. Then we obtain

$$\begin{aligned} c_{\mathbf{n}}(\tilde{f}) &= \frac{1}{(2\pi)^d} \int_{[0,\,2\pi]^d} \tilde{f}(\mathbf{x})\, \mathrm{e}^{-\mathrm{i}\,\mathbf{n}\cdot\mathbf{x}}\, \mathrm{d}\mathbf{x} \\ &= \frac{1}{(2\pi)^d} \int_{[0,\,2\pi]^d} \Big(\sum_{\mathbf{k}\in\mathbb{Z}^d} f(\mathbf{x} + 2\pi\,\mathbf{k})\, \mathrm{e}^{-\mathrm{i}\,\mathbf{n}\cdot(\mathbf{x}+2\pi\mathbf{k})} \Big)\, \mathrm{d}\mathbf{x} \\ &= \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} f(\mathbf{x})\, \mathrm{e}^{-\mathrm{i}\,\mathbf{n}\cdot\mathbf{x}}\, \mathrm{d}\mathbf{x} = \frac{1}{(2\pi)^d}\, \hat{f}(\mathbf{n})\,. \end{aligned}$$

From the second decay condition and Lemma 8 it follows that $\sum_{\mathbf{n}\in\mathbb{Z}^d} |\hat{f}(\mathbf{n})| < \infty$. Thus, by Theorem 7, the 2π-periodization $\tilde{f} \in C(\mathbb{T}^d)$ possesses the uniformly convergent Fourier series

$$\tilde{f}(\mathbf{x}) = \frac{1}{(2\pi)^d} \sum_{\mathbf{n}\in\mathbb{Z}^d} \hat{f}(\mathbf{n})\, \mathrm{e}^{\mathrm{i}\,\mathbf{n}\cdot\mathbf{x}}\,.$$

Further we have $\tilde{f} \in C(\mathbb{T}^d)$ such that (27) is valid for all $\mathbf{x} \in \mathbb{R}^d$.

■

Remark 31

The decay conditions (26) *on f and $\hat{f}$ are needed only for the absolute convergence of both series and the pointwise validity of* (27)*. Note that the Poisson summation formula* (27) *holds pointwise or almost everywhere under much weaker conditions on f and $\hat{f}$, see [17].* $\square$

Finally, we will see that the Fourier transform can be generalized to so-called tempered distributions which are linear continuous functionals on the Schwartz space. The simplest tempered distribution which cannot be described just by integrating the product of some function with those from $\mathcal{S}(\mathbb{R}^d)$, is the Dirac distribution δ defined by $\langle \delta, \varphi \rangle := \varphi(\mathbf{0})$ for all $\varphi \in \mathcal{S}(\mathbb{R}^d)$.

A *tempered distribution* T is a continuous linear functional on $\mathcal{S}(\mathbb{R}^d)$. In other words, a tempered distribution $T : \mathcal{S}(\mathbb{R}^d) \to \mathbb{C}$ fulfills the following conditions:

(i) Linearity: For all $\alpha_1,\ \alpha_2 \in \mathbb{C}$ and all $\varphi_1,\ \varphi_2 \in \mathcal{S}(\mathbb{R}^d)$,

$$\langle T, \alpha_1 \varphi_1 + \alpha_2 \varphi_2 \rangle = \alpha_1 \langle T, \varphi_1 \rangle + \alpha_2 \langle T, \varphi_2 \rangle \,.$$

(ii) Continuity: If $\varphi_j \underset{\mathcal{S}}{\longrightarrow} \varphi$ as $j \to \infty$ with $\varphi_j,\ \varphi \in \mathcal{S}(\mathbb{R}^d)$, then

$$\lim_{j \to \infty} \langle T,\, \varphi_j \rangle = \langle T,\, \varphi \rangle \,.$$

The set of tempered distributions is denoted by $\mathcal{S}'(\mathbb{R}^d)$. Defining for $T_1,\ T_2 \in \mathcal{S}'(\mathbb{R}^d)$ and all $\varphi \in \mathcal{S}(\mathbb{R}^d)$ the operation

$$\langle \alpha_1\, T_1 + \alpha_2\, T_2, \varphi \rangle := \alpha_1 \, \langle T_1, \varphi \rangle + \alpha_2 \, \langle T_2,\, \varphi \rangle,$$

the set $\mathcal{S}'(\mathbb{R}^d)$ becomes a linear space. We say that a sequence $(T_k)_{k \in \mathbb{N}}$ of tempered distributions $T_k \in \mathcal{S}'(\mathbb{R}^d)$ *converges in* $\mathcal{S}'(\mathbb{R}^d)$ to $T \in \mathcal{S}'(\mathbb{R}^d)$, if for all $\varphi \in \mathcal{S}(\mathbb{R}^d)$,

$$\lim_{k \to \infty} \langle T_k,\, \varphi \rangle = \langle T,\, \varphi \rangle \,.$$

We will use the notation $T_k \underset{\mathcal{S}'}{\longrightarrow} T$ as $k \to \infty$.

Lemma 32 (Schwartz)

A linear functional $T : \mathcal{S}(\mathbb{R}^d) \to \mathbb{C}$ is a tempered distribution if and only if there exist constants $m \in \mathbb{N}_0$ and $C \geq 0$ such that for all $\varphi \in \mathcal{S}(\mathbb{R}^d)$,

$$|\langle T, \varphi \rangle| \leq C\, \|\varphi\|_m \,. \tag{28}$$

Proof: 1. Assume that (28) holds true. Let $\varphi_j \underset{\mathcal{S}}{\to} \varphi$ as $j \to \infty$, i.e., by Lemma 10, $\lim_{j\to\infty} \|\varphi_j - \varphi\|_m = 0$ for all $m \in \mathbb{N}_0$. From (28) it follows

$$|\langle T, \varphi_j - \varphi \rangle| \leq C\, \|\varphi_j - \varphi\|_m$$

for some $m \in \mathbb{N}_0$ and $C \geq 0$. Thus $\lim_{j\to\infty} \langle T, \varphi_j - \varphi \rangle = 0$ and hence $\lim_{j\to\infty} \langle T,\, \varphi_j \rangle = \langle T,\, \varphi \rangle$.

2. Conversely, let $T \in \mathcal{S}'(\mathbb{R}^d)$. Then $\varphi_j \underset{\mathcal{S}}{\to} \varphi$ as $j \to \infty$ implies $\lim_{j\to\infty}\langle T, \varphi_j\rangle = \langle T, \varphi\rangle$.
Assume that for all $m \in \mathbb{N}$ and $C > 0$ there exists $\varphi_{m,C} \in \mathcal{S}(\mathbb{R}^d)$ such that

$$|\langle T,\, \varphi_{m,C}\rangle| > C\,\|\varphi_{m,C}\|_m.$$

Choose $C = m$ and set $\varphi_m := \varphi_{m,m}$. Then it follows $|\langle T,\, \varphi_m\rangle| > m\|\varphi_m\|_m$ and hence

$$1 = |\langle T, \frac{\varphi_m}{\langle T, \varphi_m\rangle}\rangle| > m\,\|\frac{\varphi_m}{\langle T, \varphi_m\rangle}\|_m\,.$$

We introduce the function

$$\psi_m := \frac{\varphi_m}{\langle T, \varphi_m\rangle} \in \mathcal{S}(\mathbb{R}^d)$$

which has the properties $\langle T, \psi_m\rangle = 1$ and $\|\psi_m\|_m < \frac{1}{m}$. Thus, $\psi_m \underset{\mathcal{S}}{\to} 0$ as $m \to \infty$. On the other hand, we have by assumption $T \in \mathcal{S}'(\mathbb{R}^d)$ that $\lim_{m\to\infty}\langle T, \psi_m\rangle = 0$. This contradicts $\langle T, \psi_m\rangle = 1$. ■

A measurable function $f : \mathbb{R}^d \to \mathbb{C}$ is called *slowly increasing*, if there exist $C > 0$ and $N \in \mathbb{N}_0$ such that for all $\mathbf{x} \in \mathbb{R}^d$,

$$|f(\mathbf{x})| \le C\,(1 + \|\mathbf{x}\|_2)^N\,. \tag{29}$$

These functions grow at most polynomial as $\|\mathbf{x}\|_2 \to \infty$. In particular, polynomials and complex exponential functions $\mathrm{e}^{\mathrm{i}\,\boldsymbol{\omega}\cdot\mathbf{x}}$ are slowly increasing functions. But the reciprocal Gaussian function $f(\mathbf{x}) := \mathrm{e}^{\|\mathbf{x}\|_2^2}$ is not a slowly increasing function.
For each slowly increasing function f, we can form the linear functional $T_f : \mathcal{S}(\mathbb{R}^d) \to \mathbb{C}$,

$$\langle T_f,\, \varphi\rangle := \int_{\mathbb{R}^d} f(\mathbf{x})\,\varphi(\mathbf{x})\,\mathrm{d}\mathbf{x}\,, \quad \varphi \in \mathcal{S}(\mathbb{R}^d)\,. \tag{30}$$

By Lemma 32 we obtain $T_f \in \mathcal{S}'(\mathbb{R}^d)$, because for every $\varphi \in \mathcal{S}(\mathbb{R}^d)$,

$$\begin{aligned}
|\langle T_f, \varphi\rangle| &\le \int_{\mathbb{R}^d} \frac{|f(\mathbf{x})|}{(1+\|\mathbf{x}\|_2)^{N+d+1}} \, (1+\|\mathbf{x}\|_2)^{N+d+1} \, |\varphi(\mathbf{x})| \, \mathrm{d}\mathbf{x} \\
&\le C \int_{\mathbb{R}^d} \frac{\mathrm{d}\mathbf{x}}{(1+\|\mathbf{x}\|_2)^{d+1}} \sup_{\mathbf{x}\in\mathbb{R}^d} \left((1+\|\mathbf{x}\|_2)^{N+d+1} \, |\varphi(\mathbf{x})|\right) \\
&\le C \int_{\mathbb{R}^d} \frac{\mathrm{d}\mathbf{x}}{(1+\|\mathbf{x}\|_2)^{d+1}} \, \|\varphi\|_{N+d+1} \, .
\end{aligned}$$

In the following, we identify a slowly increasing function f and the corresponding functional $T_f \in \mathcal{S}'(\mathbb{R}^d)$. Then T_f is called a *regular tempered distribution*. In this case we also say that $f \in \mathcal{S}'(\mathbb{R}^d)$. A tempered distribution, which is not a regular tempered distribution, is called a *singular tempered distribution*. The constant function 1 and any polynomial are in $\mathcal{S}'(\mathbb{R}^d)$, but the function $\mathrm{e}^{\|\mathbf{x}\|_2^2}$ is not in $\mathcal{S}'(\mathbb{R}^d)$.

Example 33

Every function $f \in L_p(\mathbb{R}^d)$, $1 \le p \le \infty$, is in $\mathcal{S}'(\mathbb{R}^d)$ by Lemma 32. For $p = 1$ we have

$$|\langle T_f, \varphi \rangle| \le \int_{\mathbb{R}^d} |f(\mathbf{x})||\varphi(\mathbf{x})| \,\mathrm{d}\mathbf{x} \le \|f\|_{L_1(\mathbb{R}^d)} \|\varphi\|_0 < \infty \,.$$

For $1 < p \le \infty$, let q be given by $\frac{1}{p} + \frac{1}{q} = 1$, where $q = 1$ if $p = \infty$. Then we obtain for $m \in \mathbb{N}_0$ with $m\,q \ge d + 1$ by Hölder's inequality

$$\begin{aligned} |\langle T_f, \varphi \rangle| &\le \int_{\mathbb{R}^d} |f(\mathbf{x})|(1 + \|\mathbf{x}\|_2)^{-m}(1 + \|\mathbf{x}\|_2)^m \,|\varphi(\mathbf{x})| \,\mathrm{d}\mathbf{x} \\ &\le \|\varphi\|_m \int_{\mathbb{R}^d} |f(\mathbf{x})|(1 + \|\mathbf{x}\|_2)^{-m} \,\mathrm{d}\mathbf{x} \\ &\le \|\varphi\|_m \, \|f\|_{L_p(\mathbb{R}^d)} \Big(\int_{\mathbb{R}^d} (1 + \|\mathbf{x}\|_2)^{-qm} \,\mathrm{d}\mathbf{x} \Big)^{1/q} . \square \end{aligned}$$

Example 34

The *Dirac distribution* δ is defined by

$$\langle \delta, \varphi \rangle := \varphi(\mathbf{0})$$

for all $\varphi \in \mathcal{S}(\mathbb{R}^d)$. Clearly, the Dirac distribution δ is a continuous linear functional with $|\langle \delta, \varphi \rangle| \le \|\varphi\|_0$ for all $\varphi \in \mathcal{S}(\mathbb{R}^d)$ so that $\delta \in \mathcal{S}'(\mathbb{R}^d)$. By the following argument the Dirac distribution is a singular tempered distribution: Assume that δ is a regular tempered distribution. Then there exists a slowly increasing function f with

$$\varphi(\mathbf{0}) = \int_{\mathbb{R}^d} f(\mathbf{x})\, \varphi(\mathbf{x})\, \mathrm{d}\mathbf{x}$$

for all $\varphi \in \mathcal{S}(\mathbb{R}^d)$. By (29) this function f is integrable over the unit ball. Let φ be the compactly supported test function (13) and $\varphi_n(\mathbf{x}) := \varphi(n\,\mathbf{x})$ for $n \in \mathbb{N}$.

Example 34 (continue)

Then we obtain the contradiction

$$\mathrm{e}^{-1} = |\varphi_n(\mathbf{0})| = \Big| \int_{\mathbb{R}^d} f(\mathbf{x})\, \varphi_n(\mathbf{x})\, \mathrm{d}\mathbf{x} \Big| \le \int_{B_{1/n}(\mathbf{0})} |f(\mathbf{x})|\, |\varphi(n\mathbf{x})|\, \mathrm{d}\mathbf{x}$$

$$\le \mathrm{e}^{-1} \int_{B_{1/n}(\mathbf{0})} |f(\mathbf{x})|\, \mathrm{d}\mathbf{x} \to 0 \quad \text{as} \quad n \to \infty\,,$$

where $B_{1/n}(\mathbf{0}) = \{\mathbf{x} \in \mathbb{R}^d : \|\mathbf{x}\|_2 \le 1/n\}$. $\square$

Important operations on tempered distributions are translations, dilations, multiplications with smooth, sufficiently fast decaying functions and derivations. In the following, we consider these operations.

The *translation* by $\mathbf{x}_0 \in \mathbb{R}^d$ of a tempered distribution $T \in \mathcal{S}'(\mathbb{R}^d)$ is the tempered distribution $T(\cdot - \mathbf{x}_0)$ defined for all $\varphi \in \mathcal{S}(\mathbb{R}^d)$ by

$$\langle T(\cdot - \mathbf{x}_0), \varphi \rangle := \langle T, \varphi(\cdot + \mathbf{x}_0) \rangle \,.$$

The *scaling* with $c \in \mathbb{R} \setminus \{0\}$ of $T \in \mathcal{S}'(\mathbb{R}^d)$ is the tempered distribution $T(c\,\cdot)$ given for all $\varphi \in \mathcal{S}(\mathbb{R}^d)$ by

$$\langle T(c\,\cdot), \varphi \rangle := \frac{1}{|c|^d} \langle T, \varphi(c^{-1}\,\cdot) \rangle \,.$$

In particular for $c = -1$, we obtain the *reflection* of $T \in \mathcal{S}'(\mathbb{R}^d)$, namely

$$\langle T(-\,\cdot), \varphi \rangle := \langle T, \tilde{\varphi} \rangle$$

for all $\varphi \in \mathcal{S}(\mathbb{R}^d)$, where $\tilde{\varphi}(\mathbf{x}) := \varphi(-\mathbf{x})$ denotes the reflection of $\varphi \in \mathcal{S}(\mathbb{R}^d)$.

Assume that $\psi \in C^{\infty}(\mathbb{R}^d)$ fulfills

$$|D^{\boldsymbol{\alpha}}\psi(\mathbf{x})| \le C_{\boldsymbol{\alpha}}\,(1 + \|\mathbf{x}\|_2)^{N_{\boldsymbol{\alpha}}}$$

for all $\boldsymbol{\alpha} \in \mathbb{N}_0^d$ and positive constants $C_{\boldsymbol{\alpha}}$ and $N_{\boldsymbol{\alpha}}$, i.e., $D^{\boldsymbol{\alpha}}\psi$ has at most polynomial growth at infinity for all $\boldsymbol{\alpha} \in \mathbb{N}_0^d$. Then the *product* of ψ with a tempered distribution $T \in \mathcal{S}'(\mathbb{R}^d)$ with is the tempered distribution $\psi\,T$ defined as

$$\langle \psi\,T, \varphi\rangle := \langle T, \psi\,\varphi\rangle\,, \quad \varphi \in \mathcal{S}(\mathbb{R}^d)\,.$$

Note that the product of an arbitrary $C^{\infty}(\mathbb{R}^d)$ function with a tempered distribution is not defined.

Example 35

For a regular distribution $T_f \in \mathcal{S}'(\mathbb{R}^d)$ with a slowly increasing function f, we obtain

$$T_f(\cdot - \mathbf{x}_0) = T_{f(\cdot - \mathbf{x}_0)}\,, \quad T_f(\varepsilon \cdot) = T_{f(\varepsilon \cdot)}\,, \quad \psi\, T_f = T_{\psi f}\,.$$

For the Dirac distribution δ, we have

$$\begin{aligned}
\langle \delta(\cdot - \mathbf{x}_0), \varphi \rangle &= \langle \delta, \varphi(\cdot + \mathbf{x}_0) \rangle = \varphi(\mathbf{x}_0)\,, \\
\langle \delta(\varepsilon \cdot), \varphi \rangle &= \frac{1}{|\varepsilon|^d}\, \langle \delta,\, \varphi\big(\frac{\cdot}{\varepsilon}\big) \rangle = \frac{1}{|\varepsilon|^d}\, \varphi(\mathbf{0})\,, \\
\langle \psi\, \delta, \varphi \rangle &= \langle \delta,\, \psi\, \varphi \rangle = \psi(\mathbf{0})\, \varphi(\mathbf{0})
\end{aligned}$$

for all $\varphi \in \mathcal{S}(\mathbb{R}^d)$. $\square$

Another important operation on tempered distributions is the differentiation. For $\boldsymbol{\alpha} \in \mathbb{N}_0^d$, the *derivative* $D^{\boldsymbol{\alpha}}\, T$ of a distribution $T \in \mathcal{S}'(\mathbb{R}^d)$ is defined for all $\varphi \in \mathcal{S}(\mathbb{R}^d)$ by

$$\langle D^{\boldsymbol{\alpha}}\, T,\, \varphi \rangle := (-1)^{|\boldsymbol{\alpha}|} \, \langle T,\, D^{\boldsymbol{\alpha}} \varphi \rangle \,. \tag{31}$$

Assume that $f \in C^r(\mathbb{R}^d)$ with $r \in \mathbb{N}$ possesses slowly increasing partial derivatives $D^{\boldsymbol{\alpha}} f$ for all $|\boldsymbol{\alpha}| \le r$. Thus $T_{D^{\boldsymbol{\alpha}} f} \in \mathcal{S}'(\mathbb{R}^d)$. Then we see by integration by parts that $T_{D^{\boldsymbol{\alpha}} f} = D^{\boldsymbol{\alpha}}\, T_f$ for all $\boldsymbol{\alpha} \in \mathbb{N}_0^d$ with $|\boldsymbol{\alpha}| \le r$, i.e., the distributional derivatives and the classical derivatives coincide.

Lemma 36

Let T, $T_k \in \mathcal{S}'(\mathbb{R}^d)$ with $k \in \mathbb{N}$ be given. For λ_1, $\lambda_2 \in \mathbb{R}$ and $\boldsymbol{\alpha}$, $\boldsymbol{\beta} \in \mathbb{N}_0^d$, the following relations hold true:

1. $D^{\boldsymbol{\alpha}}\, T \in \mathcal{S}'(\mathbb{R}^d)\,,$
2. $D^{\boldsymbol{\alpha}}\,(\lambda_1\, T_1 + \lambda_2\, T_2) = \lambda_1\, D^{\boldsymbol{\alpha}}\, T_1 + \lambda_2\, D^{\boldsymbol{\alpha}}\, T_2\,,$
3. $D^{\boldsymbol{\alpha}}\,(D^{\boldsymbol{\beta}}\, T) = D^{\boldsymbol{\beta}}\,(D^{\boldsymbol{\alpha}}\, T) = D^{\boldsymbol{\alpha}+\boldsymbol{\beta}}\, T\,.$
4. $T_k \underset{\mathcal{S}'}{\longrightarrow} T$ *as* $k \to \infty$ *implies* $D^{\boldsymbol{\alpha}}\, T_k \underset{\mathcal{S}'}{\longrightarrow} D^{\boldsymbol{\alpha}}\, T$ *as* $k \to \infty\,.$

Proof: The properties 1 – 3 follow directly from the definition of the derivative of tempered distributions. Property 4 can be derived by

$$\begin{aligned}\lim_{k\to\infty}\langle D^{\boldsymbol{\alpha}}\, T_k,\, \varphi\rangle &= \lim_{k\to\infty}(-1)^{|\boldsymbol{\alpha}|}\,\langle T_k,\, D^{\boldsymbol{\alpha}}\,\varphi\rangle = (-1)^{|\boldsymbol{\alpha}|}\,\langle T,\, D^{\boldsymbol{\alpha}}\,\varphi\rangle \\ &= \langle D^{\boldsymbol{\alpha}}\, T,\, \varphi\rangle\end{aligned}$$

for all $\varphi \in \mathcal{S}(\mathbb{R}^d)$. ∎

Example 37

For the slowly increasing univariate function

$$f(x) := \begin{cases} 0 & x \leq 0\,, \\ x & x > 0\,, \end{cases}$$

we obtain

$$\begin{aligned} \langle D\,T_f,\, \varphi\rangle &= -\langle f,\, \varphi'\rangle = -\int_{\mathbb{R}} f(x)\,\varphi'(x)\,\mathrm{d}x \\ &= -\int_0^\infty x\,\varphi'(x)\,\mathrm{d}x = -x\,\varphi(x)\big|_0^\infty + \int_0^\infty \varphi(x)\,\mathrm{d}x \\ &= \int_0^\infty \varphi(x)\,\mathrm{d}x \end{aligned}$$

so that

$$D\,T_f(x) = H(x) := \begin{cases} 0 & x \leq 0\,, \\ 1 & x > 0\,. \end{cases}$$

Example 37 (continue)

The function H is called *Heaviside function*. Further we get

$$\begin{aligned} \langle D^2\, T_f,\, \varphi\rangle &= -\langle D\, T_f,\, \varphi'\rangle = -\int_0^\infty \varphi'(x)\,\mathrm{d}x \\ &= -\varphi(x)\big|_0^\infty = \varphi(0) = \langle \delta,\, \varphi\rangle \end{aligned}$$

so that $D^2\, T_f = D\, T_H = \delta$. Thus the distributional derivative of the Heaviside function is equal to the Dirac distribution. □

For arbitrary $\psi \in \mathcal{S}(\mathbb{R}^d)$ and $T \in \mathcal{S}'(\mathbb{R}^d)$, the *convolution* $\psi * T$ is defined as

$$\langle \psi * T, \varphi\rangle := \langle T,\, \tilde{\psi} * \varphi\rangle\,, \quad \varphi \in \mathcal{S}(\mathbb{R}^d)\,, \tag{32}$$

where $\tilde{\psi}$ denotes the reflection of ψ.

Example 38

Let f be a slowly increasing function. For the regular tempered distribution $T_f \in \mathcal{S}'(\mathbb{R}^d)$ and $\psi \in \mathcal{S}(\mathbb{R}^d)$ we have by Fubini's theorem for all $\varphi \in \mathcal{S}(\mathbb{R}^d)$,

$$\langle \psi * T_f, \varphi \rangle = \langle T_f, \tilde{\psi} * \varphi \rangle = \int_{\mathbb{R}^d} f(\mathbf{y})\, (\tilde{\psi} * \varphi)(\mathbf{y})\, \mathrm{d}\mathbf{y}$$

$$= \int_{\mathbb{R}^d} f(\mathbf{y}) \Big(\int_{\mathbb{R}^d} \psi(\mathbf{x} - \mathbf{y})\, \varphi(\mathbf{x})\, \mathrm{d}\mathbf{x} \Big)\, \mathrm{d}\mathbf{y} = \int_{\mathbb{R}^d} (\psi * f)(\mathbf{x})\, \varphi(\mathbf{x})\, \mathrm{d}\mathbf{x}\,,$$

i.e., $\psi * T_f = T_{\psi * f}$ is a regular tempered distribution generated by the $C^\infty(\mathbb{R}^d)$ function

$$\int_{\mathbb{R}^d} \psi(\mathbf{x} - \mathbf{y})\, f(\mathbf{y})\, \mathrm{d}\mathbf{y} = \langle T_f, \psi(\mathbf{x} - \cdot) \rangle\,.$$

Example 38 (continue)

For the Dirac distribution δ and $\psi \in \mathcal{S}(\mathbb{R}^d)$, we get for all $\varphi \in \mathcal{S}(\mathbb{R}^d)$

$$\langle \psi * \delta, \varphi \rangle = \langle \delta, \tilde{\psi} * \varphi \rangle = (\tilde{\psi} * \varphi)(\mathbf{0}) = \int_{\mathbb{R}^d} \psi(\mathbf{x}) \, \varphi(\mathbf{x}) \, \mathrm{d}\mathbf{x}$$

i.e., $\psi * \delta = \psi$. $\square$

The convolution $\psi * T$ of $\psi \in \mathcal{S}(\mathbb{R}^d)$ and $T \in \mathcal{S}'(\mathbb{R}^d)$ possesses the following properties:

Theorem 39

For all $\psi \in \mathcal{S}(\mathbb{R}^d)$ and $T \in \mathcal{S}'(\mathbb{R}^d)$, the convolution $\psi * T$ is a regular tempered distribution generated by the slowly increasing $C^\infty(\mathbb{R}^d)$ function $\langle T, \psi(\mathbf{x} - \cdot)\rangle$, $\mathbf{x} \in \mathbb{R}^d$. For all $\boldsymbol{\alpha} \in \mathbb{N}_0^d$ it holds

$$D^{\boldsymbol{\alpha}}(\psi * T) = (D^{\boldsymbol{\alpha}}\psi) * T = \psi * (D^{\boldsymbol{\alpha}} T)\,. \tag{33}$$

Proof: 1. For arbitrary $\varphi \in \mathcal{S}(\mathbb{R}^d)$, $T \in \mathcal{S}'(\mathbb{R}^d)$, and $\boldsymbol{\alpha} \in \mathbb{N}_0^d$, we obtain by (31) and (32)

$$\langle D^{\boldsymbol{\alpha}}(\psi * T), \varphi\rangle = (-1)^{|\boldsymbol{\alpha}|} \langle \psi * T, D^{\boldsymbol{\alpha}}\varphi\rangle = (-1)^{|\boldsymbol{\alpha}|} \langle T, \tilde{\psi} * D^{\boldsymbol{\alpha}}\varphi\rangle ,$$

where $\tilde{\psi}(\mathbf{x}) = \psi(-\mathbf{x})$ and

$$(\tilde{\psi} * D^{\boldsymbol{\alpha}}\varphi)(\mathbf{x}) = \int_{\mathbb{R}^d} \tilde{\psi}(\mathbf{y})\, D^{\boldsymbol{\alpha}}\varphi(\mathbf{x} - \mathbf{y})\, \mathrm{d}\mathbf{y}\,.$$

Now we have

$$\begin{aligned}(\tilde{\psi} * D^{\boldsymbol{\alpha}}\varphi)(\mathbf{x}) &= \int_{\mathbb{R}^d} \tilde{\psi}(\mathbf{y})\, D^{\boldsymbol{\alpha}}\varphi(\mathbf{x} - \mathbf{y})\, \mathrm{d}\mathbf{y} = D^{\boldsymbol{\alpha}}\,(\tilde{\psi} * \varphi)(\mathbf{x})\\ &= D^{\boldsymbol{\alpha}} \int_{\mathbb{R}^d} \tilde{\psi}(\mathbf{x} - \mathbf{y})\, \varphi(\mathbf{y})\, \mathrm{d}\mathbf{y} = \int_{\mathbb{R}^d} D^{\boldsymbol{\alpha}}\, \tilde{\psi}(\mathbf{x} - \mathbf{y})\, \varphi(\mathbf{y})\, \mathrm{d}\mathbf{y}\\ &= (D^{\boldsymbol{\alpha}}\, \tilde{\psi} * \varphi)(\mathbf{x})\,,\end{aligned}$$

since the interchange of differentiation and integration in above integrals is justified, because $\tilde{\psi}$ and φ belong to $\mathcal{S}(\mathbb{R}^d)$.

From

$$D^{\boldsymbol{\alpha}}\,\tilde{\psi} = (-1)^{|\boldsymbol{\alpha}|}\,\widetilde{D^{\boldsymbol{\alpha}}\,\psi}$$

it follows that

$$\begin{aligned}
&\langle D^{\boldsymbol{\alpha}}(\psi * T),\, \varphi\rangle = (-1)^{|\boldsymbol{\alpha}|}\,\langle \psi * T,\, D^{\boldsymbol{\alpha}}\varphi\rangle = \langle D^{\boldsymbol{\alpha}}\,T,\, \tilde{\psi} * \varphi\rangle \\
&= \langle \psi * D^{\boldsymbol{\alpha}}\,T,\, \varphi\rangle \\
&= (-1)^{|\boldsymbol{\alpha}|}\,\langle T,\, D^{\boldsymbol{\alpha}}\tilde{\psi} * \varphi\rangle = \langle T,\, \widetilde{D^{\boldsymbol{\alpha}}\psi} * \varphi\rangle = \langle (D^{\boldsymbol{\alpha}}\psi) * T,\, \varphi\rangle\,.
\end{aligned}$$

Thus we have shown (33).
2. Now we prove that the convolution $\psi * T$ is a regular tempered distribution generated by the complex-valued function $\langle T,\, \psi(\mathbf{x} - \,\cdot)\rangle$ for $\mathbf{x} \in \mathbb{R}^d$. In Example 38 we have seen that this is true for each regular tempered distribution.
Let $\psi,\, \varphi \in \mathcal{S}(\mathbb{R}^d)$ and $T \in \mathcal{S}'(\mathbb{R}^d)$ be given. By Lemma 18 we know that $\tilde{\psi} * \varphi \in \mathcal{S}(\mathbb{R}^d)$.

We represent $(\tilde{\psi} * \varphi)(\mathbf{y})$ for arbitrary $\mathbf{y} \in \mathbb{R}^d$ a limit of Riemann sums

$$(\tilde{\psi} * \varphi)(\mathbf{y}) = \int_{\mathbb{R}^d} \psi(\mathbf{x} - \mathbf{y})\, \varphi(\mathbf{x})\, \mathrm{d}\mathbf{x} = \lim_{j \to \infty} \sum_{\mathbf{k} \in \mathbb{Z}^d} \psi(\mathbf{x_k} - \mathbf{y}) \psi(\mathbf{x_k})\, \frac{1}{j^d}\,,$$

where $\mathbf{x_k} := \frac{\mathbf{k}}{j}$, $\mathbf{k} \in \mathbb{Z}^d$, is the midpoint of a hypercube with side length $\frac{1}{j}$. Indeed, since $\tilde{\psi} * \varphi \in \mathcal{S}(\mathbb{R}^d)$, it is not hard to check that the above Riemann sums converge in $\mathcal{S}(\mathbb{R}^d)$. Since T is a continuous linear functional, we get

$$\begin{aligned} \langle T, \tilde{\psi} * \varphi \rangle &= \lim_{j \to \infty} \langle T, \sum_{\mathbf{k} \in \mathbb{Z}^d} \varphi(\mathbf{x_k} - \cdot)\, \psi(\mathbf{x_k})\, \frac{1}{j^d} \rangle \\ &= \lim_{j \to \infty} \sum_{\mathbf{k} \in \mathbb{Z}^d} \varphi(\mathbf{x_k})\, \frac{1}{j^d}\, \langle T,\, \psi(\mathbf{x_k} - \cdot) \rangle \\ &= \int_{\mathbb{R}^d} \langle T,\, \psi(\mathbf{x} - \cdot) \rangle\, \varphi(\mathbf{x})\, \mathrm{d}\mathbf{x}\,, \end{aligned}$$

i.e., the convolution $\psi * T$ is a regular tempered distribution generated by the function $\langle T,\, \psi(\mathbf{x} - \cdot) \rangle$ which belongs to $C^\infty(\mathbb{R}^d)$ by (33).

3. Finally, we show that the $C^\infty(\mathbb{R}^d)$ function $\langle T, \psi(\mathbf{x} - \cdot)\rangle$ is slowly increasing. Here we use the simple estimate

$$1 + \|\mathbf{x} - \mathbf{y}\|_2 \le 1 + \|\mathbf{x}\|_2 + \|\mathbf{y}\|_2 \le (1 + \|\mathbf{x}\|_2)\,(1 + \|\mathbf{y}\|_2)$$

for all $\mathbf{x}$, $\mathbf{y} \in \mathbb{R}^d$.
For arbitrary fixed $\mathbf{x}_0 \in \mathbb{R}^d$ and every $m \in \mathbb{N}_0$, we obtain for $\psi \in \mathcal{S}(\mathbb{R}^d)$,

$$\begin{aligned}
\|\psi(\mathbf{x}_0 - \cdot)\|_m &= \max_{|\boldsymbol{\beta}| \le m} \|(1 + \|\mathbf{x}\|_2)^m\, D^{\boldsymbol{\beta}} \psi(\mathbf{x}_0 - \mathbf{x})\|_{C_0(\mathbb{R}^d)} \\
&= \max_{|\boldsymbol{\beta}| \le m} \max_{\mathbf{x} \in \mathbb{R}^d} (1 + \|\mathbf{x}\|_2)^m\, |D^{\boldsymbol{\beta}}\, \psi(\mathbf{x}_0 - \mathbf{x})| \\
&= \max_{|\boldsymbol{\beta}| \le m} \max_{\mathbf{y} \in \mathbb{R}^d} (1 + \|\mathbf{x}_0 - \mathbf{y}\|_2)^m\, |D^{\boldsymbol{\beta}} \psi(\mathbf{y})| \\
&\le (1 + \|\mathbf{x}_0\|_2)^m \sup_{|\boldsymbol{\beta}| \le m} \sup_{\mathbf{y} \in \mathbb{R}^d} (1 + \|\mathbf{y}\|_2)^m\, |D^{\boldsymbol{\beta}}\, \psi(\mathbf{y})| \\
&= (1 + \|\mathbf{x}_0\|_2)^m\, \|\psi\|_m\,.
\end{aligned}$$

Since $T \in \mathcal{S}'(\mathbb{R}^d)$, by Lemma 32 of Schwartz there exist constants $m \in \mathbb{N}_0$ and $C > 0$, so that $|\langle T, \varphi \rangle| \le C \, \|\varphi\|_m$ for all $\varphi \in \mathcal{S}(\mathbb{R}^d)$. Then we conclude

$$|\langle T, \psi(\mathbf{x} - \cdot) \rangle| \le C \, \|\psi(\mathbf{x} - \cdot)\|_m \le C \, (1 + \|\mathbf{x}\|_2)^m \, \|\psi\|_m \, .$$

Hence $\langle T, \psi(\mathbf{x} - \cdot) \rangle$ is a slowly increasing function. ■

The *Fourier transform* $\mathcal{F}T = \hat{T}$ of a tempered distribution $T \in \mathcal{S}'(\mathbb{R}^d)$ is defined by

$$\langle \mathcal{F}T, \varphi \rangle = \langle \hat{T}, \varphi \rangle := \langle T, \mathcal{F}\varphi \rangle = \langle T, \hat{\varphi} \rangle \tag{34}$$

for all $\varphi \in \mathcal{S}(\mathbb{R}^d)$. Indeed $\hat{T}$ is again a continuous linear functional on $\mathcal{S}(\mathbb{R}^d)$, since by Theorem 17, the expression $\langle T, \mathcal{F}\varphi \rangle$ defines a linear functional on $\mathcal{S}(\mathbb{R}^d)$. Further, $\varphi_k \underset{\mathcal{S}}{\longrightarrow} \varphi$ as $k \to \infty$, implies $\mathcal{F}\,\varphi_k \underset{\mathcal{S}}{\longrightarrow} \mathcal{F}\,\varphi$ as $k \to \infty$ so that for $T \in \mathcal{S}'(\mathbb{R}^d)$, it follows

$$\lim_{k\to\infty} \langle \hat{T}, \varphi_k \rangle = \lim_{k\to\infty} \langle T, \mathcal{F}\,\varphi_k \rangle = \langle T, \mathcal{F}\,\varphi \rangle = \langle \hat{T}, \varphi \rangle\,.$$

Example 40

Let $f \in L_1(\mathbb{R}^d)$. Then we obtain for an arbitrary $\varphi \in \mathcal{S}(\mathbb{R}^d)$ by Fubini's theorem

$$\begin{aligned} \langle \mathcal{F} T_f, \varphi \rangle = \langle T_f, \hat{\varphi} \rangle &= \int_{\mathbb{R}^d} \Big(\int_{\mathbb{R}^d} \varphi(\mathbf{x})\, \mathrm{e}^{-\mathrm{i}\,\mathbf{x}\cdot\boldsymbol{\omega}}\, \mathrm{d}\mathbf{x} \Big)\, f(\boldsymbol{\omega})\, \mathrm{d}\boldsymbol{\omega} \\ &= \int_{\mathbb{R}^d} \hat{f}(\mathbf{x})\, \varphi(\mathbf{x})\, \mathrm{d}\mathbf{x} = \langle T_{\hat{f}}, \varphi \rangle, \end{aligned}$$

i.e., $\mathcal{F}\, T_f = T_{\mathcal{F}f}$.
Let $\mathbf{x}_0 \in \mathbb{R}^d$ be fixed. For the shifted Dirac distribution $\delta_{\mathbf{x}_0} := \delta(\cdot - \mathbf{x}_0)$, we have

$$\begin{aligned} \langle \mathcal{F}\delta_{\mathbf{x}_0}, \varphi \rangle = \langle \delta_{\mathbf{x}_0}, \hat{\varphi} \rangle &= \langle \delta_{\mathbf{x}_0}, \int_{\mathbb{R}^d} \varphi(\boldsymbol{\omega}) \mathrm{e}^{-\mathrm{i}\,\boldsymbol{\omega}\cdot\mathbf{x}}\, \mathrm{d}\boldsymbol{\omega} \rangle \\ &= \int_{\mathbb{R}^d} \varphi(\boldsymbol{\omega})\, \mathrm{e}^{-\mathrm{i}\,\boldsymbol{\omega}\cdot\mathbf{x}_0}\, \mathrm{d}\boldsymbol{\omega} = \langle \mathrm{e}^{-\mathrm{i}\,\boldsymbol{\omega}\cdot\mathbf{x}_0}, \varphi(\boldsymbol{\omega}) \rangle, \end{aligned}$$

so that $\mathcal{F}\delta_{\mathbf{x}_0} = \mathrm{e}^{-\mathrm{i}\,\boldsymbol{\omega}\cdot\mathbf{x}_0}$ and in particular, for $\mathbf{x}_0 = \mathbf{0}$ we obtain $\mathcal{F}\delta = 1$. $\square$

Theorem 41

The Fourier transform on $\mathcal{S}'(\mathbb{R}^d)$ is a linear, bijective operator $\mathcal{F} : \mathcal{S}'(\mathbb{R}^d) \to \mathcal{S}'(\mathbb{R}^d)$. The Fourier transform on $\mathcal{S}'(\mathbb{R}^d)$ is continuous in the sense that for $T_k,\ T \in \mathcal{S}'(\mathbb{R}^d)$ the convergence $T_k \underset{\mathcal{S}'}{\longrightarrow} T$ as $k \to \infty$ implies $\mathcal{F}\, T_k \underset{\mathcal{S}'}{\longrightarrow} \mathcal{F}\, T$ as $k \to \infty$. The inverse Fourier transform is given by

$$\langle \mathcal{F}^{-1} T, \varphi \rangle = \langle T, \mathcal{F}^{-1} \varphi \rangle \tag{35}$$

for all $\varphi \in \mathcal{S}(\mathbb{R}^d)$ which means

$$\mathcal{F}^{-1} T := \frac{1}{(2\pi)^d} \mathcal{F}\, T(-\,\cdot)\,.$$

For all $T \in \mathcal{S}'(\mathbb{R}^d)$ it holds the Fourier inversion formula

$$\mathcal{F}^{-1}(\mathcal{F}\, T) = \mathcal{F}\,(\mathcal{F}^{-1} T) = T\,.$$

Proof: By definition (34), the Fourier transform $\mathcal{F}$ maps $\mathcal{S}'(\mathbb{R}^d)$ into itself. Obviously, $\mathcal{F}$ is a linear operator. We show that $\mathcal{F}$ is a continuous linear operator of $\mathcal{S}'(\mathbb{R}^d)$ onto $\mathcal{S}'(\mathbb{R}^d)$. Assume that $T_k \underset{\mathcal{S}'}{\longrightarrow} T$ as $k \to \infty$. Then, we get by (34),

$$\lim_{k\to\infty} \langle \mathcal{F}\, T_k, \varphi\rangle = \lim_{k\to\infty} \langle T_k, \mathcal{F}\,\varphi\rangle = \langle T, \mathcal{F}\,\varphi\rangle = \langle \mathcal{F}\, T, \varphi\rangle$$

for all $\varphi \in \mathcal{S}(\mathbb{R}^d)$. This means that $\mathcal{F}\, T_k \underset{\mathcal{S}'}{\longrightarrow} \mathcal{F}\, T$ as $k \to \infty$, i.e., the operator $\mathcal{F} : \mathcal{S}'(\mathbb{R}^d) \to \mathcal{S}'(\mathbb{R}^d)$ is continuous.
Next we show that (35) is the inverse Fourier transform, i.e.,

$$\mathcal{F}^{-1}\,(\mathcal{F}\, T) = T\,, \quad \mathcal{F}\,(\mathcal{F}^{-1}\, T) = T \tag{36}$$

for all $T \in \mathcal{S}'(\mathbb{R}^d)$.

By Theorem 17, we find that for all $\varphi \in \mathcal{S}(\mathbb{R}^d)$,

$$\begin{aligned}
\langle \mathcal{F}^{-1}(\mathcal{F}\,T),\,\varphi\rangle &= \frac{1}{(2\pi)^d}\,\langle \mathcal{F}\big(\mathcal{F}\,T(-\,\cdot)\big),\,\varphi\rangle \\
&= \frac{1}{(2\pi)^d}\,\langle \mathcal{F}\,T(-\,\cdot),\,\mathcal{F}\,\varphi\rangle = \frac{1}{(2\pi)^d}\,\langle \mathcal{F}\,T,\,(\mathcal{F}\,\varphi)(-\,\cdot)\rangle \\
&= \langle \mathcal{F}\,T,\,\mathcal{F}^{-1}\,\varphi\rangle = \langle T,\,\mathcal{F}(\mathcal{F}^{-1}\,\varphi)\rangle = \langle T,\,\varphi\rangle\,.
\end{aligned}$$

By (36), each $T \in \mathcal{S}'(\mathbb{R}^d)$ is the Fourier transform of the tempered distribution $S = \mathcal{F}^{-1}\,T$, i.e., $T = \mathcal{F}\,S$. Thus both $\mathcal{F}$ and $\mathcal{F}^{-1}$ map $\mathcal{S}'(\mathbb{R}^d)$ one-to-one onto $\mathcal{S}'(\mathbb{R}^d)$. ■

Remark 42

From Theorem 41 it follows immediately Theorem 21. If $f \in L_1(\mathbb{R}^d)$ with $\hat{f} \in L_1(\mathbb{R}^d)$ is given, then T_f and $T_{\hat{f}}$ are regular tempered distributions by Example 33. By Theorem 41 and Example 40 we have

$$T_{\mathcal{F}^{-1}\hat{f}} = \mathcal{F}^{-1}\, T_{\hat{f}} = \mathcal{F}^{-1}(\mathcal{F}\, T_f) = T_f$$

so that the functions f and

$$(\mathcal{F}^{-1}\hat{f})(\mathbf{x}) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \hat{f}(\boldsymbol{\omega})\, \mathrm{e}^{\mathrm{i}\,\mathbf{x}\cdot\boldsymbol{\omega}}\, \mathrm{d}\boldsymbol{\omega}$$

are equal almost everywhere. □

The following theorem summarizes properties of Fourier transform on $\mathcal{S}'(\mathbb{R}^d)$.

Theorem 43 (Properties of the Fourier transform on $\mathcal{S}'(\mathbb{R}^d)$)

The Fourier transform of a tempered distribution $T \in \mathcal{S}'(\mathbb{R}^d)$ has the following properties:

1. Translation and modulation: *For fixed $\mathbf{x}_0$, $\boldsymbol{\omega}_0 \in \mathbb{R}^d$,*

$$\mathcal{F}T(\cdot - \mathbf{x}_0) = \mathrm{e}^{-\mathrm{i}\,\boldsymbol{\omega}\cdot\mathbf{x}_0}\,\mathcal{F}T\,,$$
$$\mathcal{F}\big(\mathrm{e}^{-\mathrm{i}\,\boldsymbol{\omega}_0\cdot\mathbf{x}}\,T\big) = \mathcal{F}T(\cdot + \boldsymbol{\omega}_0)\,.$$

2. Differentiation and multiplication: *For $\boldsymbol{\alpha} \in \mathbb{N}_0^d$,*

$$\mathcal{F}(D^{\boldsymbol{\alpha}}T) = \mathrm{i}^{|\boldsymbol{\alpha}|}\,\boldsymbol{\omega}^{\boldsymbol{\alpha}}\,\mathcal{F}T\,,$$
$$\mathcal{F}(\mathbf{x}^{\boldsymbol{\alpha}}T) = \mathrm{i}^{|\boldsymbol{\alpha}|}\,D^{\boldsymbol{\alpha}}\mathcal{F}T\,.$$

Theorem 43 (continue)

3. Scaling: *For* $c \in \mathbb{R} \setminus \{0\}$,

$$\mathcal{F}T(c\,\cdot) = \frac{1}{|c|^d}\,\mathcal{F}T(c^{-1}\,\cdot)\,.$$

4. Convolution: *For* $\varphi \in \mathcal{S}(\mathbb{R}^d)$,

$$\mathcal{F}(T * \varphi) = (\mathcal{F}T)\,(\mathcal{F}\varphi)\,.$$

The proof follows in a straightforward way from the definitions of corresponding operators, in particular the Fourier transform (34) on $\mathcal{S}'(\mathbb{R}^d)$ and Theorem 19.

Finally, we present some additional examples of Fourier transforms of tempered distributions.

Example 44

In Example 40 we have seen that for fixed $\mathbf{x}_0 \in \mathbb{R}^d$,

$$\mathcal{F}\,\delta_{\mathbf{x}_0} = \mathrm{e}^{-\mathrm{i}\,\boldsymbol{\omega}\cdot\mathbf{x}_0}\,, \quad \mathcal{F}\,\delta = 1\,.$$

Now we determine $\mathcal{F}^{-1}\,1$. By Theorem 41, we obtain

$$\mathcal{F}^{-1}\,1 = \frac{1}{(2\pi)^d}\,\mathcal{F}\,1(-\,\cdot) = \frac{1}{(2\pi)^d}\,\mathcal{F}\,1,$$

since the reflection $1(-\,\cdot)$ is equal to 1. Thus, we have $\mathcal{F}\,1 = (2\pi)^d\,\delta$. From Theorem 43, it follows for any $\boldsymbol{\alpha} \in \mathbb{N}_0^d$,

$$\begin{aligned} \mathcal{F}\,(D^{\boldsymbol{\alpha}}\,\delta) &= (\mathrm{i}\,\boldsymbol{\omega})^{\boldsymbol{\alpha}}\,\mathcal{F}\,\delta = (\mathrm{i}\,\boldsymbol{\omega})^{\boldsymbol{\alpha}}\,1 = (\mathrm{i}\,\boldsymbol{\omega})^{\boldsymbol{\alpha}}\,, \\ \mathcal{F}\,(\mathbf{x}^{\boldsymbol{\alpha}}) &= \mathcal{F}\,(\mathbf{x}^{\boldsymbol{\alpha}}\,1) = \mathrm{i}^{|\boldsymbol{\alpha}|}\,D^{\boldsymbol{\alpha}}\,\mathcal{F}\,1 = (2\pi)^d\,\mathrm{i}^{|\boldsymbol{\alpha}|}\,D^{\boldsymbol{\alpha}}\,\delta\,. \quad \square \end{aligned}$$

The spaces $\mathcal{S}(\mathbb{R}^d)$, $L_2(\mathbb{R}^d)$ and $\mathcal{S}'(\mathbb{R}^d)$ are a typical example of a so-called *Gelfand triple* named after the mathematician I.M. Gelfand (1913 – 2009). To obtain a Gelfand triple (B, H, B'), we equip a Hilbert space H with a dense topological vector subspace B of test functions carrying a finer topology than H such that the natural inclusion $B \subset H$ is continuous. We consider the inclusion of the dual space H' in B', where B' is the dual space of all linear continuous functionals on B with its topology. Applying the Riesz representation theorem, we can identify H with H' leading to the Gelfand triple

$$B \subset H \cong H' \subset B'.$$

We are interested in

$$\mathcal{S}(\mathbb{R}^d) \subset L_2(\mathbb{R}^d) \cong L_2(\mathbb{R}^d)' \subset \mathcal{S}'(\mathbb{R}^d). \tag{37}$$

We already know that $\mathcal{S}(\mathbb{R}^d)$ is dense in $L_2(\mathbb{R}^d)$. Moreover, the first natural embedding is continuous, since $\varphi_k \underset{\mathcal{S}}{\to} \varphi$ as $k \to \infty$ implies

$$\begin{aligned} \|\varphi_k - \varphi\|^2_{L_2(\mathbb{R}^d)} &= \int_{\mathbb{R}^d} (1 + \|\mathbf{x}\|_2)^{-d-1} \, (1 + \|\mathbf{x}\|_2)^{d+1} \, |\varphi_k(\mathbf{x}) - \varphi(\mathbf{x})|^2 \, \mathrm{d}\mathbf{x} \\ &\le \sup_{\mathbf{x} \in \mathbb{R}^d} (1 + \|\mathbf{x}\|_2)^{d+1} |\varphi_k(\mathbf{x}) - \varphi(\mathbf{x})|^2 \int_{\mathbb{R}^d} \frac{\mathrm{d}\mathbf{y}}{(1 + \|\mathbf{y}\|_2)^{d+1}} \\ &\le C \sup_{\mathbf{x} \in \mathbb{R}^d} (1 + \|\mathbf{x}\|_2)^{d+1} \, |\varphi_k(\mathbf{x}) - \varphi(\mathbf{x})|^2 \ \to \ 0 \end{aligned}$$

as $k \to \infty$.

Let F be a continuous linear functional on $L_2(\mathbb{R}^d)$. Then we can identify F with the unique function $f \in L_2(\mathbb{R}^d)$ fulfilling $F = \langle \cdot, \bar{f} \rangle_{L_2(\mathbb{R}^d)}$ and consider the mapping $\iota : L_2(\mathbb{R}^d)' \to \mathcal{S}'(\mathbb{R}^d)$ defined by $\iota F := T_f$, see Example 33. Indeed ι is injective by the following argument: Assume that $F_n = \langle \cdot, \bar{f}_n \rangle_{L_2(\mathbb{R}^d)}$, $n = 1, 2$, are different continuous linear functionals on $L_2(\mathbb{R}^d)$, but $T_{f_1} = T_{f_2}$. Then we get $\langle T_{f_1}, \varphi \rangle = \langle T_{f_2}, \varphi \rangle$ for all $\varphi \in \mathcal{S}(\mathbb{R}^d)$, i.e.,

$$\langle \varphi, \bar{f}_1 \rangle_{L_2(\mathbb{R}^d)} = \langle \varphi, \bar{f}_2 \rangle_{L_2(\mathbb{R}^d)}$$

which is impossible, since $\mathcal{S}(\mathbb{R}^d)$ is dense in $L_2(\mathbb{R}^d)$.

Corollary 45

If we identify $f \in L_2(\mathbb{R}^d)$ with $T_f \in \mathcal{S}'(\mathbb{R}^d)$, then the Fourier transforms on $L_2(\mathbb{R}^d)$ and $\mathcal{S}'(\mathbb{R}^d)$ coincide in the sense $\mathcal{F}T_f = T_{\mathcal{F}f}$.

Proof: For any sequence $(f_k)_{k\in\mathbb{N}}$ of functions $f_k \in \mathcal{S}(\mathbb{R}^d)$ converging to f in $L_2(\mathbb{R}^d)$, we obtain

$$\lim_{k\to\infty} \langle \mathcal{F}\varphi, \bar{f}_k \rangle_{L_2(\mathbb{R}^d)} = \langle \mathcal{F}\varphi, \bar{f} \rangle_{L_2(\mathbb{R}^d)} = \langle T_f, \mathcal{F}\varphi \rangle = \langle \mathcal{F}T_f, \varphi \rangle$$

for all $\varphi \in \mathcal{S}(\mathbb{R}^d)$. On the other hand, we conclude by continuity of $\mathcal{F}$ that

$$\lim_{k\to\infty} \langle \mathcal{F}\varphi, \bar{f}_k \rangle_{L_2(\mathbb{R}^d)} = \lim_{k\to\infty} \langle \varphi, \overline{\mathcal{F}f_k} \rangle_{L_2(\mathbb{R}^d)} = \langle \varphi, \overline{\mathcal{F}f} \rangle_{L_2(\mathbb{R}^d)} = \langle T_{\mathcal{F}f}, \varphi \rangle$$

for all $\varphi \in \mathcal{S}(\mathbb{R}^d)$. Thus, $\mathcal{F}T_f = T_{\mathcal{F}f}$ and we are done. ■

Multidimensional discrete Fourier transforms

The multidimensional DFT is necessary for the computation of Fourier coefficients of a function $f \in C(\mathbb{T}^d)$ as well as for the calculation of the Fourier transform of a function $f \in L_1(\mathbb{R}^d) \cap C(\mathbb{R}^d)$. Further the two-dimensional DFT finds numerous applications in image processing. The properties of the one-dimensional DFT can be extended to the multidimensional DFT in a straightforward way.

Computation of multivariate Fourier coefficients

We describe the computation of Fourier coefficients $c_{\mathbf{k}}(f)$, $\mathbf{k} = (k_j)_{j=1}^d \in \mathbb{Z}^d$, of a given function $f \in C(\mathbb{T}^d)$, where f is sampled on the uniform grid $\{\frac{2\pi}{N}\,\mathbf{n} :\, \mathbf{n} \in I_N^d\}$, where $N \in \mathbb{N}$ is even, $I_N := \{0,\, \ldots, N-1\}$, and $I_N^d := \{\mathbf{n} = (n_j)_{j=1}^d : n_j \in I_N,\, j = 1, \ldots, d\}$. Using the rectangle rule of numerical integration, we can compute $c_{\mathbf{k}}(f)$ for $\mathbf{k} \in \mathbb{Z}^d$ approximately. Since $[0,\, 2\pi]^d$ is equal to the union of the N^d hypercubes $\frac{2\pi}{N}\,\mathbf{n} + [0,\, \frac{2\pi}{N}]^d$, $\mathbf{n} \in I_N^d$, we obtain

$$c_{\mathbf{k}}(f) = \frac{1}{(2\pi)^d} \int_{[0,\, 2\pi]^d} f(\mathbf{x})\, \mathrm{e}^{-\mathrm{i}\,\mathbf{k}\cdot\mathbf{x}}\, \mathrm{d}\mathbf{x} \approx \frac{1}{N^d} \sum_{\mathbf{n} \in I_N^d} f\big(\frac{2\pi}{N}\,\mathbf{n}\big)\, \mathrm{e}^{-2\pi\mathrm{i}\,(\mathbf{k}\cdot\mathbf{n})/N}$$

$$= \frac{1}{N^d} \sum_{\mathbf{n} \in I_N^d} f\big(\frac{2\pi}{N}\,\mathbf{n}\big)\, w_N^{\mathbf{k}\cdot\mathbf{n}}$$

with $w_N = \mathrm{e}^{-2\pi\mathrm{i}/N}$.

The expression

$$\sum_{\mathbf{n}\in I_N^d} f\big(\frac{2\pi}{N}\,\mathbf{n}\big)\,w_N^{\mathbf{k}\cdot\mathbf{n}}$$

is called the *d-dimensional discrete Fourier transform of size* $N_1 \times \ldots \times N_d$ of the d-dimensional array $\big(f\big(\frac{2\pi}{N}\,\mathbf{n}\big)\big)_{\mathbf{n}\in I_N^d}$, where $N_1 = \ldots = N_d := N$. Thus we obtain the approximate Fourier coefficients

$$\hat{f}_{\mathbf{k}} := \frac{1}{N^d} \sum_{\mathbf{n}\in I_N^d} f\big(\frac{2\pi}{N}\,\mathbf{n}\big)\,w_N^{\mathbf{k}\cdot\mathbf{n}}\,. \tag{38}$$

Obviously, the values $\hat{f}_{\mathbf{k}}$ are N-periodic, i.e., for all $\mathbf{k}$, $\mathbf{m} \in \mathbb{Z}^d$ we have

$$\hat{f}_{\mathbf{k}+N\,\mathbf{m}} = \hat{f}_{\mathbf{k}}\,.$$

But by Lemma 6 we know that $\lim_{\|\mathbf{k}\|_2 \to \infty} c_{\mathbf{k}}(f) = 0$. Therefore we can only expect that

$$\hat{f}_{\mathbf{k}} \approx c_{\mathbf{k}}(f)\,, \quad k_j = -\frac{N}{2}, \ldots, \frac{N}{2} - 1\,;\; j = 1, \ldots, d\,.$$

To see this effect more clearly, we will derive a multidimensional aliasing formula. By $\delta_{\mathbf{m}}$, $\mathbf{m} \in \mathbb{Z}^d$, we denote the *d-dimensional Kronecker symbol*

$$\delta_{\mathbf{m}} := \begin{cases} 1 & \mathbf{m} = \mathbf{0}\,, \\ 0 & \mathbf{m} \in \mathbb{Z}^d \setminus \{\mathbf{0}\}\,. \end{cases}$$

Lemma 46

Let $N_j \in \mathbb{N} \setminus \{1\}$, $j = 1, \ldots, d$, *be given. Then for each* $\mathbf{m} = (m_j)_{j=1}^d \in \mathbb{Z}^d$, *we have*

$$\sum_{k_1=0}^{N_1-1} \ldots \sum_{k_d=0}^{N_d-1} w_{N_1}^{m_1 k_1} \ldots w_{N_d}^{m_d k_d} = \prod_{j=1}^{d} (N_j \, \delta_{m_j \bmod N})$$

$$= \begin{cases} \prod_{j=1}^d N_j & \mathbf{m} \in N_1 \mathbb{Z} \times \ldots \times N_d \mathbb{Z}\,, \\ 0 & \mathbf{m} \in \mathbb{Z}^d \setminus (N_1 \mathbb{Z} \times \ldots \times N_d \mathbb{Z})\,. \end{cases}$$

If $N_1 = \ldots = N_d = N$, *then for each* $\mathbf{m} \in \mathbb{Z}^d$,

$$\sum_{\mathbf{k} \in I_N^d} w_N^{\mathbf{m}\cdot\mathbf{k}} = N^d \, \delta_{\mathbf{m} \bmod N} = \begin{cases} N^d & \mathbf{m} \in N\,\mathbb{Z}^d\,, \\ 0 & \mathbf{m} \in \mathbb{Z}^d \setminus (N\,\mathbb{Z}^d)\,, \end{cases}$$

where the vector $\mathbf{m} \bmod N := (m_j \bmod N)_{j=1}^d$ *denotes the nonnegative residue of* $\mathbf{m} \in \mathbb{Z}^d$ *modulo* N, *and* $\delta_{\mathbf{m} \bmod N} = \prod_{j=1}^d \delta_{m_j \bmod N}$.

Proof: This result is an immediate consequence

$$\begin{aligned}\sum_{k_1=0}^{N_1-1} \dots \sum_{k_d=0}^{N_d-1} w_{N_1}^{m_1 k_1} \dots w_{N_d}^{m_d k_d} &= \prod_{j=1}^{d} \Big(\sum_{k_j=0}^{N_j-1} w_{N_j}^{m_j k_j} \Big) \\ &= \prod_{j=1}^{d} \big(N_j \, \delta_{m_j \bmod N_j} \big) \,. \qquad \blacksquare\end{aligned}$$

The following aliasing formula describes a close relation between the Fourier coefficients $c_{\mathbf{k}}(f)$ and the approximate values $\hat{f}_{\mathbf{k}}$.

Theorem 47

Let $N \in \mathbb{N}$ be even and let $f \in C(\mathbb{T}^d)$ be given. Assume that the Fourier coefficients $c_{\mathbf{k}}(f)$ satisfy the condition $\sum_{\mathbf{k}\in\mathbb{Z}^d} |c_{\mathbf{k}}(f)| < \infty$.

Then we have the aliasing formula

$$\hat{f}_{\mathbf{k}} = \sum_{\mathbf{m}\in\mathbb{Z}^d} c_{\mathbf{k}+N\,\mathbf{m}}(f)\,. \tag{39}$$

Thus for $k_j = -\frac{N}{2}, \ldots, \frac{N}{2} - 1$ and $j = 1, \ldots, d$, we have the error estimate

$$|\hat{f}_{\mathbf{k}} - c_{\mathbf{k}}(f)| \leq \sum_{\mathbf{m}\in\mathbb{Z}^d\setminus\{\mathbf{0}\}} |c_{\mathbf{k}+N\,\mathbf{m}}(f)|\,.$$

Proof: By Theorem 7, the d-dimensional Fourier series of f converges uniformly to f. Hence for all $\mathbf{x} \in \mathbb{T}^d$, we have

$$f(\mathbf{x}) = \sum_{\mathbf{m} \in \mathbb{Z}^d} c_{\mathbf{m}}(f)\, \mathrm{e}^{\mathrm{i}\, \mathbf{m} \cdot \mathbf{x}} .$$

In particular for $\mathbf{x} = \frac{2\pi}{N}\, \mathbf{n}$, $\mathbf{n} \in I_N^d$, we obtain

$$f\big(\frac{2\pi}{N}\, \mathbf{n}\big) = \sum_{\mathbf{m} \in \mathbb{Z}^d} c_{\mathbf{m}}(f)\, \mathrm{e}^{2\pi\, \mathrm{i}\, (\mathbf{m} \cdot \mathbf{n})/N} = \sum_{\mathbf{m} \in \mathbb{Z}^d} c_{\mathbf{m}}(f)\, w_N^{-\mathbf{m} \cdot \mathbf{n}} .$$

Hence due to (38) and the pointwise convergence of the Fourier series,

$$\begin{aligned} \hat{f}_{\mathbf{k}} &= \frac{1}{N^d} \sum_{\mathbf{n} \in \mathbf{I}_N} \Big(\sum_{\mathbf{m} \in \mathbb{Z}^d} c_{\mathbf{m}}(f)\, w_N^{-\mathbf{m} \cdot \mathbf{n}} \Big)\, w_N^{\mathbf{k} \cdot \mathbf{n}} \\ &= \frac{1}{N^d} \sum_{\mathbf{m} \in \mathbb{Z}^d} c_{\mathbf{m}}(f) \sum_{\mathbf{n} \in I_N^d} w_N^{(\mathbf{k} - \mathbf{m}) \cdot \mathbf{n}} , \end{aligned}$$

which yields the aliasing formula (39) by Lemma 46. ■

Now we sketch the computation of the Fourier transform $\hat{f}$ of a given function $f \in L_1(\mathbb{R}^d) \cap C_0(\mathbb{R}^d)$. Since $f(\mathbf{x}) \to 0$ as $\|\mathbf{x}\|_2 \to \infty$, we obtain for sufficiently large $n \in \mathbb{N}$ that

$$\hat{f}(\boldsymbol{\omega}) = \int_{\mathbb{R}^d} f(\mathbf{x})\, \mathrm{e}^{-\mathrm{i}\,\mathbf{x}\cdot\boldsymbol{\omega}}\, \mathrm{d}\mathbf{x} \approx \int_{[-n\pi,\, n\pi]^d} f(\mathbf{x})\, \mathrm{e}^{-\mathrm{i}\,\mathbf{x}\cdot\boldsymbol{\omega}}\, \mathrm{d}\mathbf{x}\,, \quad \boldsymbol{\omega} \in \mathbb{R}^d\,.$$

Using the uniform grid $\{\frac{2\pi}{N}\,\mathbf{k} :\ k_j = -\frac{nN}{2}, \ldots, \frac{nN}{2} - 1;\ j = 1, \ldots, d\}$ of the hypercube $[-n\pi,\ n\pi)^d$ for even $N \in \mathbb{N}$, we receive by the rectangle rule of numerical integration

$$\begin{aligned} &\int_{[-n\pi,\, n\pi]^d} f(\mathbf{x})\, \mathrm{e}^{-\mathrm{i}\,\mathbf{x}\cdot\boldsymbol{\omega}}\, \mathrm{d}\mathbf{x} \\ \approx\ & \big(\frac{2\pi}{N}\big)^d \sum_{k_1=-nN/2}^{nN/2-1} \ldots \sum_{k_d=-nN/2}^{nN/2-1} f\big(\frac{2\pi}{N}\,\mathbf{k}\big)\, \mathrm{e}^{-2\pi\mathrm{i}\,(\mathbf{k}\cdot\boldsymbol{\omega})/N}\,. \end{aligned}$$

For $\boldsymbol{\omega} = \frac{1}{n}\,\mathbf{m}$ with $m_j = -\frac{nN}{2}, \ldots, \frac{nN}{2} - 1$ and $j = 1, \ldots, d$, we obtain the following values

$$\Big(\frac{2\pi}{N}\Big)^d \sum_{k_1=-nN/2}^{nN/2-1} \ldots \sum_{k_d=-nN/2}^{nN/2-1} f\Big(\frac{2\pi}{N}\,\mathbf{k}\Big)\, w_{nN}^{\mathbf{k}\cdot\boldsymbol{\omega}} \approx \hat{f}\Big(\frac{1}{n}\,\mathbf{m}\Big)\,,$$

which can be considered as d-dimensional DFT$(N_1 \times \ldots \times N_d)$ with $N_1 = \ldots = N_d = n\,N$.

Two-dimensional discrete Fourier transforms

Let $N_1,\ N_2 \in \mathbb{N} \setminus \{1\}$ be given, and let $I_{N_j} := \{0, \ldots, N_j - 1\}$ for $j = 1,\ 2$ be the corresponding index sets. The linear map from $\mathbb{C}^{N_1 \times N_2}$ into itself which maps any matrix $\mathbf{A} = (a_{k_1,k_2})_{k_1,k_2=0}^{N_1-1,N_2-1} \in \mathbb{C}^{N_1 \times N_2}$ to the matrix

$$\hat{\mathbf{A}} = (\hat{a}_{n_1,n_2})_{n_1,n_2=0}^{N_1-1,N_2-1} := \mathbf{F}_{N_1}\,\mathbf{A}\,\mathbf{F}_{N_2}\,,$$

is called *two-dimensional discrete Fourier transform of size $N_1 \times N_2$* and abbreviated by $\mathrm{DFT}(N_1 \times N_2)$. The entries of the transformed matrix $\hat{\mathbf{A}}$ read as follows

$$\hat{a}_{n_1,n_2} = \sum_{k_1=0}^{N_1-1} \sum_{k_2=0}^{N_2-1} a_{k_1,k_2}\, w_{N_1}^{k_1 n_1}\, w_{N_2}^{k_2 n_2}\,, \quad n_j \in I_{N_j};\ j = 1,\ 2\,. \quad (40)$$

If we form the entries (40) for all $n_1,\ n_2 \in \mathbb{Z}$, then we observe the *periodicity* of $\mathrm{DFT}(N_1 \times N_2)$, i.e., for all $\ell_1,\ \ell_2 \in \mathbb{Z}$, one has

$$\hat{a}_{n_1,n_2} = \hat{a}_{n_1+\ell_1 N_1, n_2+\ell_2 N_2}\,, \quad n_j \in I_{N_j}\,,\ j = 1,\ 2\,.$$

Remark 48

The two-dimensional DFT is of great importance for digital image processing. The light intensity measured by a camera is generally sampled over a rectangular array of pictures elements, so-called pixels. Thus a digital grayscale image is a matrix $\mathbf{A} = (a_{k_1,k_2})_{k_1,k_2=0}^{N_1-1,N_2-1}$ *of* $N_1\,N_2$ *pixels* $(k_1, k_2) \in I_{N_1} \times I_{N_2}$ *and corresponding grayscale values* $a_{k_1,k_2} \in \{0, 1, \ldots, 255\}$*, where zero means black and 255 is white. Typically,* N_1*,* $N_2 \in \mathbb{N}$ *are relatively large, for instance* $N_1 = N_2 = 512$*.*
The modulus of the transformed matrix $\hat{\mathbf{A}}$ *is given by* $|\hat{\mathbf{A}}| := (|\hat{a}_{n_1,n_2}|)_{n_1,n_2=0}^{N_1-1,N_2-1}$ *and its phase by*

$$\mathrm{atan2}\,(\mathrm{Im}\,\hat{\mathbf{A}}, \mathrm{Re}\,\hat{\mathbf{A}}) := \big(\mathrm{atan2}\,(\mathrm{Im}\,\hat{a}_{n_1,n_2}, \mathrm{Re}\,\hat{a}_{n_1,n_2})\big)_{n_1,n_2=0}^{N_1-1,N_2-1},$$

where atan2 *is defined in Matlab. In natural images the phase contains important structure information as illustrated in Figure 1.*
□

Figure 1: Top: Images *Barbara* (left) and *Lena* (right). Bottom: Images reconstructed with modulus of *Barbara* and phase of *Lena* (left) and conversely, with modulus of *Lena* and phase of *Barbara* (right). The phase appears to be dominant with respect to structures.

For the computation of DFT($N_1 \times N_2$) the following simple relation to one-dimensional DFT's is very useful. If the data a_{k_1,k_2} can be factorized as

$$a_{k_1,k_2} = b_{k_1}\, c_{k_2}\,, \quad k_j \in I_{N_j};\ j = 1,\, 2\,,$$

then the DFT($N_1 \times N_2$) of $\mathbf{A} = (a_{k_1,k_2})_{k_1,k_2=0}^{N_1-1,N_2-1} = \mathbf{b}\,\mathbf{c}^\top$ reads as follows

$$\hat{\mathbf{A}} = \mathbf{F}_{N_1}\mathbf{b}\mathbf{c}^\top\mathbf{F}_{N_2}^\top = (\hat{b}_{n_1}\ \hat{c}_{n_2})_{n_1,n_2=0}^{N_1-1,N_2-1}\,, \tag{41}$$

where $(\hat{b}_{n_1})_{n_1=0}^{N_1-1}$ is the one-dimensional DFT(N_1) of $\mathbf{b} = (b_{k_1})_{k_1=0}^{N_1-1}$ and $(\hat{c}_{n_2})_{n_2=0}^{N_2-1}$ is the one-dimensional DFT(N_2) of $\mathbf{c} = (c_{k_2})_{k_2=0}^{N-1}$.

Example 49

For fixed $s_j \in I_{N_j}$, $j = 1, 2$, the sparse matrix

$$\mathbf{A} := \big(\delta_{(k_1 - s_1) \bmod N_1}\, \delta_{(k_2 - s_2) \bmod N_2}\big)_{k_1, k_2 = 0}^{N_1 - 1, N_2 - 1}$$

is transformed to $\hat{\mathbf{A}} = \big(w_{N_1}^{n_1 s_1}\, w_{N_2}^{n_2 s_2}\big)_{n_1, n_2 = 0}^{N_1 - 1, N_2 - 1}$. Thus we see that a sparse matrix (i.e., a matrix with few nonzero entries) is not transformed to a sparse matrix.
Conversely, the matrix $\mathbf{B} = \big(w_{N_1}^{-s_1 k_1}\, w_{N_2}^{-s_2 k_2}\big)_{k_1, k_2 = 0}^{N_1 - 1, N_2 - 1}$ is mapped to

$$\hat{\mathbf{B}} := N_1\, N_2\, \big(\delta_{(n_1 - s_1) \bmod N_1}\, \delta_{(n_2 - s_2) \bmod N_2}\big)_{n_1, n_2 = 0}^{N_1 - 1, N_2 - 1}. \quad \square$$

Example 50

Let $N_1 = N_2 = N \in \mathbb{N} \setminus \{1\}$. We consider the matrix $\mathbf{A} = \big(a_{k_1}\, a_{k_2}\big)_{k_1,k_2=0}^{N-1}$, where a_{k_j} is defined by

$$a_{k_j} := \begin{cases} \frac{1}{2} & k_j = 0\,, \\ \frac{k_j}{N} & k_j = 1, \ldots, N-1\,. \end{cases}$$

Thus by (41) we obtain the entries of the transformed matrix $\hat{\mathbf{A}}$ by

$$\hat{a}_{n_1,n_2} = \hat{a}_{n_1}\, \hat{a}_{n_2} = -\frac{1}{4}\, \cot\frac{\pi n_1}{N}\, \cot\frac{\pi n_2}{N}\,, \quad n_j \in I_{N_j};\ j = 1,\, 2\,. \quad \square$$

The DFT($N_1 \times N_2$) maps $\mathbb{C}^{N_1 \times N_2}$ one-to-one onto itself. The *inverse* DFT($N_1 \times N_2$) *of size* $N_1 \times N_2$ is given by

$$\mathbf{A} = \mathbf{F}_{N_1}^{-1}\, \hat{\mathbf{A}}\, \mathbf{F}_{N_2}^{-1} = \frac{1}{N_1 N_2}\, \mathbf{J}'_{N_1}\, \mathbf{F}_{N_1}\, \hat{\mathbf{A}}\, \mathbf{F}_{N_2}\, \mathbf{J}'_{N_2}$$

such that

$$a_{k_1,k_2} = \frac{1}{N_1 N_2} \sum_{n_1=0}^{N_1-1} \sum_{n_2=0}^{N_2-1} \hat{a}_{n_1,n_2}\, w_{N_1}^{-k_1 n_1}\, w_{N_2}^{-k_2 n_2}\,, \quad k_j \in I_{N_j};\ j = 1, 2\,.$$

In practice, one says that the DFT($N_1 \times N_2$) is defined on the *time domain* or *space domain* $\mathbb{C}^{N_1 \times N_2}$. The range of the DFT($N_1 \times N_2$) is called *frequency domain* which is $\mathbb{C}^{N_1 \times N_2}$ too.

In the linear space $\mathbb{C}^{N_1 \times N_2}$ we introduce the *inner product* of two complex matrices $\mathbf{A} = \big(a_{k_1,k_2}\big)_{k_1,k_2=0}^{N_1-1,N_2-1}$, $\mathbf{B} = \big(b_{k_1,k_2}\big)_{k_1,k_2=0}^{N_1-1,N_2-1}$,

$$\langle \mathbf{A},\, \mathbf{B} \rangle := \sum_{k_1=0}^{N_1-1} \sum_{k_2=0}^{N_2-1} a_{k_1,k_2}\, \overline{b}_{k_1,k_2}$$

and the *Frobenius norm*

$$\|\mathbf{A}\|_{\mathrm{F}} := \langle \mathbf{A},\, \mathbf{A} \rangle^{1/2} = \Big(\sum_{k_1=0}^{N_1-1} \sum_{k_2=0}^{N_2-1} |a_{k_1,k_2}|^2 \Big)^{1/2} .$$

Lemma 51

For given $N_1,\, N_2 \in \mathbb{N} \setminus \{1\}$, the set of exponential matrices

$$\mathbf{E}_{m_1,m_2} := \big(w_{N_1}^{-k_1 m_1}\, w_{N_2}^{-k_2 m_2}\big)_{k_1,k_2=0}^{N_1-1,N_2-1}$$

forms an orthogonal basis of $\mathbb{C}^{N_1 \times N_2}$, where $\|\mathbf{E}_{m_1,m_2}\|_{\mathrm{F}} = \sqrt{N_1\, N_2}$ for all $m_j \in I_{N_j}$ and $j = 1,\, 2$. Any matrix $\mathbf{A} \in \mathbb{C}^{N_1 \times N_2}$ can be represented in the form

$$\mathbf{A} = \frac{1}{N_1\, N_2} \sum_{m_1=0}^{N_1-1} \sum_{m_2=0}^{N_2-1} \langle \mathbf{A},\, \mathbf{E}_{m_1,m_2} \rangle\, \mathbf{E}_{m_1,m_2}\,,$$

and we have

$$\hat{\mathbf{A}} = \big(\langle \mathbf{A},\, \mathbf{E}_{m_1,m_2} \rangle\big)_{m_1,m_2=0}^{N_1-1,N_2-1}\,.$$

Proof: From Lemma 46 it follows that for $p_j \in I_{N_j}$, $j = 1, 2$,

$$\langle \mathbf{E}_{m_1,m_2},\, \mathbf{E}_{p_1,p_2} \rangle = \sum_{k_1=0}^{N_1-1} \sum_{k_2=0}^{N_2-1} w_{N_1}^{k_1\,(p_1-m_1)}\, w_{N_2}^{k_2\,(p_2-m_2)}$$

$$= N_1 N_2\, \delta_{(m_1-p_1)\bmod N_1}\, \delta_{(m_2-p_2)\bmod N_2}$$

$$= \begin{cases} N_1\, N_2 & (m_1,\, m_2) = (p_1,\, p_2)\,, \\ 0 & (m_1,\, m_2) \neq (p_1,\, p_2)\,. \end{cases}$$

Further we see that $\|\mathbf{E}_{m_1,m_2}\|_{\mathrm{F}} = \sqrt{N_1\, N_2}$. Since $\dim\, \mathbb{C}^{N_1 \times N_2} = N_1\, N_2$, the set of the $N_1\, N_2$ exponential matrices forms an orthogonal basis of $\mathbb{C}^{N_1 \times N_2}$. ■

In addition, we introduce the *cyclic convolution*

$$\mathbf{A} * \mathbf{B} := \Big(\sum_{k_1=0}^{N_1-1} \sum_{k_2=0}^{N_2-1} a_{k_1,k_2}\, b_{(m_1-k_1) \bmod N_1,\, (m_2-k_2) \bmod N_2} \Big)_{m_1,m_2=0}^{N_1-1,N_2-1}$$

and the *entrywise product*

$$\mathbf{A} \circ \mathbf{B} := \big(a_{k_1,k_2}\, b_{k_1,k_2} \big)_{k_1,k_2=0}^{N_1-1,N_2-1} .$$

In the case $N_1 = N_2 = N$, the cyclic convolution in $\mathbb{C}^{N \times N}$ is a commutative, associative, and distributive operation with the unity $(\delta_{k_1 \bmod N}\, \delta_{k_2 \bmod N})_{k_1,k_2=0}^{N-1}$.

High dimensional FFT

In this chapter, we discuss methods for the approximation of d-variate functions in high dimension $d \in \mathbb{N}$ based on sampling along rank-1 lattices and we derive the corresponding fast algorithms. In contrast to Chapter 3, our approach to compute the Fourier coefficients of d-variate functions is no longer based on tensor product methods. We introduce weighted subspaces of $L_1(\mathbb{T}^d)$ which are characterized by the decay properties of the Fourier coefficients. We show that functions in these spaces can be already well approximated by d-variate trigonometric polynomials on special frequency index sets. We study the fast evaluation of d-variate trigonometric polynomials on finite frequency index sets. We introduce so-called rank-1 lattices and derive an algorithm for the fast evaluation of these trigonometric polynomials at the lattice points. The special structure of the rank-1 lattice enables us to perform this computation using only a one-dimensional FFT.

In order to reconstruct the Fourier coefficients of the d-variate trigonometric polynomials from the polynomial values at the lattice points exactly, the used rank-1 lattice needs to satisfy a special condition. Using so-called reconstructing rank-1 lattices, the stable computation of the Fourier coefficients of a d-variate trigonometric polynomial can be again performed by employing only a one-dimensional FFT, where the numerical effort depends on the lattice size. In Section 6, we come back to the approximation of periodic functions in weighted subspaces of $L_1(\mathbb{T}^d)$ on rank-1 lattices. Section 7 considers the construction of rank-1 lattices. We present a constructive component-by-component algorithm with less than $|I|^2$ lattice points, where I denotes the finite index set of nonzero Fourier coefficients that have to be computed. In particular, this means that the computational effort to reconstruct the Fourier coefficients depends only linearly on the dimension and mainly on the size of the frequency index sets of the considered trigonometric polynomials. In order to overcome the limitations of the single rank-1 lattice approach, we generalize the proposed methods to multiple rank-1 lattices.

Fourier partial sums of smooth multivariate functions

In order to ensure a good quality of the obtained approximations of d-variate periodic functions, we need to assume that these functions satisfy certain smoothness conditions, which are closely related to the decay properties of their Fourier coefficients. As we have already seen for $d = 1$, the smoothness properties of a function strongly influence the quality of a specific approximation method, for example see Theorem of Bernstein.
We consider a d-variate periodic function $f\colon \mathbb{T}^d \to \mathbb{C}$ with the Fourier series

$$f(\mathbf{x}) = \sum_{\mathbf{k}\in\mathbb{Z}^d} c_{\mathbf{k}}(f)\, \mathrm{e}^{\mathrm{i}\,\mathbf{k}\cdot\mathbf{x}}\,. \tag{42}$$

We will always assume that $f \in L_1(\mathbb{T}^d)$ in order to guarantee the existence of all Fourier coefficients $c_{\mathbf{k}}(f)$, $\mathbf{k} \in \mathbb{Z}^d$. For the definition of function spaces $L_p(\mathbb{T}^d)$, $1 \le p < \infty$, we refer to Section 4.

The decay properties of Fourier coefficients can also be used to characterize the smoothness of the function f, see Theorem 9 for $d > 1$. For a detailed characterization of periodic functions and suitable function spaces, in particular with respect to the decay properties of the Fourier coefficients, we refer to [56, Chapter 3].

In this section, we consider the approximation of a d-variate periodic function $f \in L_1(\mathbb{T}^d)$ using Fourier partial sums $S_I f$,

$$(S_I f)(\mathbf{x}) := \sum_{\mathbf{k} \in I} c_{\mathbf{k}}(f)\, \mathrm{e}^{\mathrm{i}\,\mathbf{k}\cdot\mathbf{x}}\,, \tag{43}$$

where the finite index set $I \subset \mathbb{Z}^d$ needs to be carefully chosen with respect to the properties of the sequence of the Fourier coefficients $\big(c_{\mathbf{k}}(f)\big)_{\mathbf{k}\in\mathbb{Z}^d}$. The set I is called *frequency index set* of the Fourier partial sum. The operator $S_I : L_1(\mathbb{T}^d) \to C(\mathbb{T}^d)$ maps f to a trigonometric polynomial with frequencies supported on the finite index set I.

We call

$$\Pi_I := \mathrm{span}\,\{\mathrm{e}^{\mathrm{i}\,\mathbf{k}\cdot\mathbf{x}} :\, \mathbf{k} \in I\}$$

the *space of trigonometric polynomials supported on I*. We will be interested in frequency index sets of type

$$I = I_{p,N}^d := \{\mathbf{k} = (k_s)_{s=1}^d \in \mathbb{Z}^d :\, \|\mathbf{k}\|_p \le N\}\,, \tag{44}$$

where $\|\mathbf{k}\|_p$ is the usual p-(quasi-)norm

$$\|\mathbf{k}\|_p := \begin{cases} \Big(\sum\limits_{s=1}^{d} |k_s|^p\Big)^{1/p} & 0 < p < \infty\,, \\ \max\limits_{s=1,\ldots,d} |k_s| & p = \infty\,. \end{cases}$$

The Figure 2 illustrates the two-dimensional frequency index sets $I_{p,16}^2$ for $p \in \{\frac{1}{2},\, 1,\, 2, \infty\}$, see also [69, 68, 25].

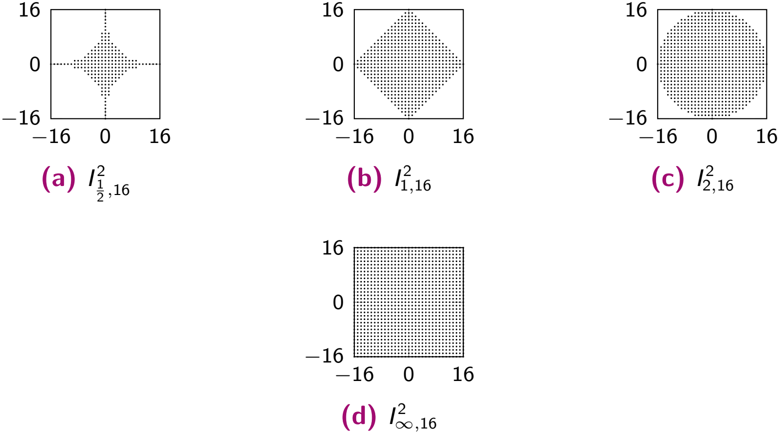


Figure 2: Two-dimensional frequency index sets $I^2_{p,16}$ for $p \in \{\frac{1}{2}, 1, 2, \infty\}$.

If the absolute values of the Fourier coefficients decrease sufficiently fast for growing frequency index $\mathbf{k}$, we can very well approximate the function f using only a few terms $c_{\mathbf{k}}(f)\,\mathrm{e}^{\mathrm{i}\,\mathbf{k}\cdot\mathbf{x}}$, $\mathbf{k} \in I \subset \mathbb{Z}^d$ with $|I| < \infty$. In particular, we will consider a periodic function $f \in L_1(\mathbb{T}^d)$ whose sequence of Fourier coefficients is absolutely summable. This implies by Theorem 9 that f has a continuous representative within $L_1(\mathbb{T}^d)$. We introduce the weighted subspace $\mathcal{A}_\omega(\mathbb{T}^d)$ of $L_1(\mathbb{T}^d)$ of functions $f : \mathbb{T}^d \to \mathbb{C}$ equipped with the norm

$$\|f\|_{\mathcal{A}_\omega(\mathbb{T}^d)} := \sum_{\mathbf{k}\in\mathbb{Z}^d} \omega(\mathbf{k})|c_{\mathbf{k}}(f)|\,, \tag{45}$$

if f has the Fourier expansion (42). Here $\omega : \mathbb{Z}^d \to [1, \infty)$ is called *weight function* and characterizes the decay of the Fourier coefficients. If ω is increasing for $\|\mathbf{k}\|_p \to \infty$, then the Fourier coefficients $c_{\mathbf{k}}(f)$ of $f \in \mathcal{A}_\omega(\mathbb{T}^d)$ have to decrease faster than the weight function ω increases with respect to $\mathbf{k} = (k_s)_{s=1}^d \in \mathbb{Z}^d$.

Example 52

Important examples for a weight function ω are

$$\omega(\mathbf{k}) = \omega_p^d(\mathbf{k}) := \max\{1, \|\mathbf{k}\|_p\}$$

for $0 < p \leq \infty$. Instead of the p-norm, one can also consider a weighted p-norm. To characterize function spaces with dominating smoothness, also weight functions of the form

$$\omega(\mathbf{k}) = \prod_{s=1}^{d} \max\{1, |k_s|\}$$

have been considered, see e.g. [63, 10, 24]. $\square$

Observe that $\omega(\mathbf{k}) \geq 1$ for all $\mathbf{k} \in \mathbb{Z}^d$. Let ω_1 be the special weight function with $\omega_1(\mathbf{k}) = 1$ for all $\mathbf{k} \in \mathbb{Z}^d$ and $\mathcal{A}(\mathbb{T}^d) := \mathcal{A}_{\omega_1}(\mathbb{T}^d)$. The space $\mathcal{A}(\mathbb{T}^d)$ is called *Wiener algebra*. Further, we recall that $C(\mathbb{T}^d)$ denotes the Banach space of continuous d-variate 2π-periodic functions. The norm of $C(\mathbb{T}^d)$ coincides with the norm of $L_\infty(\mathbb{T}^d)$. The next lemma, see [24, Lemma 2.1], states that the embeddings $\mathcal{A}_\omega(\mathbb{T}^d) \subset \mathcal{A}(\mathbb{T}^d) \subset C(\mathbb{T}^d)$ are true.

Lemma 53

Each function $f \in \mathcal{A}(\mathbb{T}^d)$ has a continuous representative. In particular, we obtain $\mathcal{A}_\omega(\mathbb{T}^d) \subset \mathcal{A}(\mathbb{T}^d) \subset C(\mathbb{T}^d)$ with the usual interpretation.

Proof: Let $f \in \mathcal{A}_\omega(\mathbb{T}^d)$ be given. Then the function f belongs to $\mathcal{A}(\mathbb{T}^d)$, since the following estimate holds

$$\sum_{\mathbf{k}\in\mathbb{Z}^d} |c_{\mathbf{k}}(f)| \le \sum_{\mathbf{k}\in\mathbb{Z}^d} \omega(\mathbf{k})\,|c_{\mathbf{k}}(f)| < \infty\,.$$

Now let $f \in \mathcal{A}(\mathbb{T}^d)$ be given. The summability of the sequence $\big(|c_{\mathbf{k}}(f)|\big)_{\mathbf{k}\in\mathbb{Z}^d}$ of the absolute values of the Fourier coefficients implies the summability of the sequence $\big(|c_{\mathbf{k}}(f)|^2\big)_{\mathbf{k}\in\mathbb{Z}^d}$ of the squared absolute values of the Fourier coefficients and, thus, the embedding $\mathcal{A}(\mathbb{T}^d) \subset L_2(\mathbb{T}^d)$ is proved using Parseval equation (4). Clearly, the function $g(\mathbf{x}) = \sum_{\mathbf{k}\in\mathbb{Z}^d} c_{\mathbf{k}}(f)\,\mathrm{e}^{\mathrm{i}\,\mathbf{k}\cdot\mathbf{x}}$ is a representative of f in $L_2(\mathbb{T}^d)$ and also in $\mathcal{A}(\mathbb{T}^d)$. We show that g is the continuous representative of f. The absolute values of the Fourier coefficients of $f \in \mathcal{A}(\mathbb{T}^d)$ are summable. So, for each $\varepsilon > 0$ there exists a finite index set $I \subset \mathbb{Z}^d$ with $\sum_{\mathbf{k}\in\mathbb{Z}^d\setminus I} |c_{\mathbf{k}}(f)| < \frac{\varepsilon}{4}$.

For a fixed $\mathbf{x}_0 \in \mathbb{T}^d$, we estimate

$$\begin{aligned} |g(\mathbf{x}_0) - g(\mathbf{x})| &= \Big| \sum_{\mathbf{k}\in\mathbb{Z}^d} c_{\mathbf{k}}(f)\,\mathrm{e}^{\mathrm{i}\,\mathbf{k}\cdot\mathbf{x}_0} - \sum_{\mathbf{k}\in\mathbb{Z}^d} c_{\mathbf{k}}(f)\,\mathrm{e}^{\mathrm{i}\,\mathbf{k}\cdot\mathbf{x}} \Big| \\ &\le \Big| \sum_{\mathbf{k}\in I} c_{\mathbf{k}}(f)\,\mathrm{e}^{\mathrm{i}\,\mathbf{k}\cdot\mathbf{x}_0} - \sum_{\mathbf{k}\in I} c_{\mathbf{k}}(f)\,\mathrm{e}^{\mathrm{i}\,\mathbf{k}\cdot\mathbf{x}} \Big| + \frac{\varepsilon}{2}\,. \end{aligned}$$

The trigonometric polynomial $(S_I f)(\mathbf{x}) = \sum_{\mathbf{k}\in I} c_{\mathbf{k}}\,\mathrm{e}^{\mathrm{i}\,\mathbf{k}\cdot\mathbf{x}}$ is a continuous function. Accordingly, for $\varepsilon > 0$ and $\mathbf{x}_0 \in \mathbb{T}^d$ there exists a $\delta_0 > 0$ such that $\|\mathbf{x}_0 - \mathbf{x}\|_1 < \delta_0$ implies $|(S_I f)(\mathbf{x}_0) - (S_I f)(\mathbf{x})| < \frac{\varepsilon}{2}$. Then we obtain $|g(\mathbf{x}_0) - g(\mathbf{x})| < \varepsilon$ for all $\mathbf{x}$ with $\|\mathbf{x}_0 - \mathbf{x}\|_1 < \delta_0$. ■

In particular for our further considerations on sampling methods, it is essential that we identify each function $f \in \mathcal{A}(\mathbb{T}^d)$ with its continuous representative in the following. Note that the definition of $\mathcal{A}_\omega(\mathbb{T}^d)$ in (45) using the Fourier series representation of f already comprises the continuity of the contained functions. Considering Fourier partial sums, we will always call them exact Fourier partial sums in contrast to approximate partial Fourier sums that will be introduced later.

Lemma 54

Let $I_N = \{\mathbf{k} \in \mathbb{Z}^d : \omega(\mathbf{k}) \le N\}$, $N \in \mathbb{R}$, *be a frequency index set being defined by the weight function* ω. *Assume that the cardinality* $|I_N|$ *is finite.*
Then the exact Fourier partial sum

$$(S_{I_N} f)(\mathbf{x}) := \sum_{\mathbf{k} \in I_N} c_{\mathbf{k}}(f)\, \mathrm{e}^{\mathrm{i}\, \mathbf{k} \cdot \mathbf{x}} \tag{46}$$

approximates the function $f \in \mathcal{A}_\omega(\mathbb{T}^d)$ *and we have*

$$\|f - S_{I_N} f\|_{L_\infty(\mathbb{T}^d)} \le N^{-1}\, \|f\|_{\mathcal{A}_\omega(\mathbb{T}^d)}\,.$$

Proof: We follow the ideas of [24, Lemma 2.2]. Let $f \in \mathcal{A}_\omega(\mathbb{T}^d)$. Obviously, $S_{I_N} f \in \mathcal{A}_\omega(\mathbb{T}^d) \subset C(\mathbb{T}^d)$ and we obtain

$$\begin{aligned}
\|f - S_{I_N} f\|_{L_\infty(\mathbb{T}^d)} &= \operatorname*{ess\,sup}_{\mathbf{x}\in\mathbb{T}^d} |(f - S_{I_N} f)(\mathbf{x})| \\
&= \operatorname*{ess\,sup}_{\mathbf{x}\in\mathbb{T}^d} \Big| \sum_{\mathbf{k}\in\mathbb{Z}^d\setminus I_N} c_{\mathbf{k}}(f)\, \mathrm{e}^{\mathrm{i}\,\mathbf{k}\cdot\mathbf{x}} \Big| \\
&\le \sum_{\mathbf{k}\in\mathbb{Z}^d\setminus I_N} |c_{\mathbf{k}}(f)| \le \frac{1}{\inf\limits_{\mathbf{k}\in\mathbb{Z}^d\setminus I_N} \omega(\mathbf{k})} \sum_{\mathbf{k}\in\mathbb{Z}^d\setminus I_N} \omega(\mathbf{k}) |\, c_{\mathbf{k}}(f)| \\
&\le \frac{1}{N} \sum_{\mathbf{k}\in\mathbb{Z}^d} \omega(\mathbf{k})\, |c_{\mathbf{k}}(f)| = N^{-1} \|f\|_{\mathcal{A}_\omega(\mathbb{T}^d)}. \quad \blacksquare
\end{aligned}$$

Remark 55

For the weight function $\omega(\mathbf{k}) = \big(\max\{1, \|\mathbf{k}\|_p\}\big)^{\alpha/2}$ *with* $0 < p \leq \infty$ *and* $\alpha > 0$ *we similarly obtain for the index set* $I_N = I^d_{p,N}$ *given in* (44)

$$\begin{aligned} \|f - S_{I^d_{p,N}} f\|_{L_\infty(\mathbb{T}^d)} &\leq N^{-\alpha/2} \sum_{\mathbf{k}\in\mathbb{Z}^d\setminus I^d_{p,N}} \big(\max\{1, \|\mathbf{k}\|_p\}\big)^{\alpha/2} |c_{\mathbf{k}}(f)| \\ &\leq N^{-\alpha/2} \, \|f\|_{\mathcal{A}_\omega(\mathbb{T}^d)} \,. \end{aligned}$$

The error estimates can be also transferred to other norms. Let $H^{\alpha,p}(\mathbb{T}^d)$ *denote the periodic Sobolev space of isotropic smoothness consisting of all* $f \in L_2(\mathbb{T}^d)$ *with finite norm*

$$\|f\|_{H^{\alpha,p}(\mathbb{T}^d)} := \sum_{\mathbf{k}\in\mathbb{Z}^d} \big(\max\{1, \|\mathbf{k}\|_p\}\big)^{\alpha} \, |c_{\mathbf{k}}(f)|^2 \,, \qquad (47)$$

where f *possesses the Fourier expansion* (42) *and where* $\alpha > 0$ *is the smoothness parameter.*

Remark 55 (continue)

Using the Cauchy–Schwarz inequality, we obtain here

$$\begin{aligned}\|f - S_{I^d_{p,N}} f\|_{L_\infty(\mathbb{T}^d)} &\leq \sum_{\mathbf{k}\in\mathbb{Z}^d\setminus I^d_{p,N}} |c_{\mathbf{k}}(f)| \\ &\leq \Big(\sum_{\mathbf{k}\in\mathbb{Z}^d\setminus I^d_{p,N}} \|\mathbf{k}\|_p^{-\alpha}\Big)^{1/2} \Big(\sum_{\mathbf{k}\in\mathbb{Z}^d\setminus I^d_{p,N}} \|\mathbf{k}\|_p^{\alpha}\, |c_{\mathbf{k}}(f)|^2\Big)^{1/2} \\ &\leq \Big(\sum_{\mathbf{k}\in\mathbb{Z}^d\setminus I^d_{p,N}} \|\mathbf{k}\|_p^{-\alpha}\Big)^{1/2} \|f\|_{H^{\alpha,p}(\mathbb{T}^d)}\,.\end{aligned}$$

Note that this estimate is related to the estimates on the decay of Fourier coefficients for functions $f \in C^r(\mathbb{T}^d)$ in (1) and Theorem 9. For detailed estimates of the approximation error of Fourier partial sums in these spaces, we refer to [36].
As we will see later, for efficient approximation, other frequency index sets, as e.g. frequency index sets related to the hyperbolic crosses, are of special interest. The corresponding approximation errors have been studied in [30, 31, 8]. □

Lemma 56

Let $N \in \mathbb{N}$ and the frequency index set $I_N := \{\mathbf{k} \in \mathbb{Z}^d : 1 \leq \omega(\mathbf{k}) \leq N\}$ with the cardinality $0 < |I_N| < \infty$ be given.
Then the norm of the operator S_{I_N} that maps $f \in \mathcal{A}_\omega(\mathbb{T}^d)$ to its Fourier partial sum $S_{I_N} f$ on the index set I_N is bounded by

$$\frac{1}{\min_{\mathbf{k}\in\mathbb{Z}^d} \omega(\mathbf{k})} \leq \|S_{I_N}\|_{\mathcal{A}_\omega(\mathbb{T}^d)\to C(\mathbb{T}^d)} \leq \frac{1}{\min_{\mathbf{k}\in\mathbb{Z}^d} \omega(\mathbf{k})} + \frac{1}{N}.$$

Proof: 1. Since $|I_N|$ is finite, there exists $\min_{\mathbf{k}\in I_N}\omega(\mathbf{k})$. The definition of I_N implies that $\min_{\mathbf{k}\in\mathbb{Z}^d}\omega(\mathbf{k}) = \min_{\mathbf{k}\in I_N}\omega(\mathbf{k})$. To obtain the upper bound for the operator norm we apply the triangle inequality and Lemma 54,

$$\|S_{I_N}\|_{\mathcal{A}_\omega(\mathbb{T}^d)\to C(\mathbb{T}^d)} = \sup_{\substack{f\in\mathcal{A}_\omega(\mathbb{T}^d)\\ \|f\|_{\mathcal{A}_\omega(\mathbb{T}^d)}=1}} \|S_{I_N}f\|_{C(\mathbb{T}^d)}$$

$$\leq \sup_{\substack{f\in\mathcal{A}_\omega(\mathbb{T}^d)\\ \|f\|_{\mathcal{A}_\omega(\mathbb{T}^d)}=1}} \|S_{I_N}f - f\|_{C(\mathbb{T}^d)} + \sup_{\substack{f\in\mathcal{A}_\omega(\mathbb{T}^d)\\ \|f\|_{\mathcal{A}_\omega(\mathbb{T}^d)}=1}} \|f\|_{C(\mathbb{T}^d)}$$

$$\leq \sup_{\substack{f\in\mathcal{A}_\omega(\mathbb{T}^d)\\ \|f\|_{\mathcal{A}_\omega(\mathbb{T}^d)}=1}} \left(\sum_{\mathbf{k}\in\mathbb{Z}^d} |c_{\mathbf{k}}(f)| + N^{-1}\|f\|_{\mathcal{A}_\omega(\mathbb{T}^d)}\right)$$

$$\leq \sup_{\substack{f\in\mathcal{A}_\omega(\mathbb{T}^d)\\ \|f\|_{\mathcal{A}_\omega(\mathbb{T}^d)}=1}} \left(\sum_{\mathbf{k}\in\mathbb{Z}^d} \frac{\omega(\mathbf{k})}{\min_{\tilde{\mathbf{k}}\in\mathbb{Z}^d}\omega(\tilde{\mathbf{k}})} |c_{\mathbf{k}}(f)| + N^{-1}\|f\|_{\mathcal{A}_\omega(\mathbb{T}^d)}\right)$$

$$\leq \frac{1}{\min_{\mathbf{k}\in\mathbb{Z}^d}\omega(\mathbf{k})} + \frac{1}{N}.$$

2. To prove the lower bound we construct a suitable example. Let $\mathbf{k}' \in I_N$ be a frequency index with $\omega(\mathbf{k}') = \min_{\mathbf{k}\in\mathbb{Z}^d} \omega(\mathbf{k})$. We choose the trigonometric polynomial $g(\mathbf{x}) = \frac{1}{\omega(\mathbf{k}')}\,\mathrm{e}^{\mathrm{i}\,\mathbf{k}'\cdot\mathbf{x}}$ which is an element of $\mathcal{A}_\omega(\mathbb{T}^d)$ with $\|g\|_{\mathcal{A}_\omega(\mathbb{T}^d)} = 1$. Since $S_{I_N} g = g$, we find

$$\begin{aligned}\|S_{I_N}\|_{\mathcal{A}_\omega(\mathbb{T}^d)\to C(\mathbb{T}^d)} &\geq \|S_{I_N} g\|_{C(\mathbb{T}^d)} \\ &= \|g\|_{C(\mathbb{T}^d)} = g(\mathbf{0}) \\ &= \frac{1}{\omega(\mathbf{k}')} = \frac{1}{\min\limits_{\mathbf{k}\in I_N} \omega(\mathbf{k})}. \qquad \blacksquare\end{aligned}$$

Our observations in this section imply that smooth functions with special decay of their Fourier coefficients can be well approximated by d-variate trigonometric polynomials on special index sets. In the next section we will therefore study the efficient evaluation of d-variate trigonometric polynomials on special grids, as well as the corresponding efficient computation of their Fourier coefficients.

Fast evaluation of multivariate trigonometric polynomials

As we have seen in the last section, smooth functions in $\mathcal{A}_\omega(\mathbb{T}^d)$ can be already well approximated by d-variate trigonometric polynomials on index sets $I_N = \{\mathbf{k} \in \mathbb{Z}^d : \omega(\mathbf{k}) \leq N\}$. In Figure 2, we have seen possible two-dimensional index sets, where $\omega(\mathbf{k}) = \max\{1, \|\mathbf{k}\|_p\}$. Therefore we study trigonometric polynomials $p \in \Pi_I$ on the d-dimensional torus $\mathbb{T}^d \cong [0, 2\pi)^d$ of the form

$$p(\mathbf{x}) = \sum_{\mathbf{k} \in I} \hat{p}_{\mathbf{k}} \, \mathrm{e}^{\mathrm{i}\, \mathbf{k} \cdot \mathbf{x}} \tag{48}$$

with Fourier coefficients $\hat{p}_{\mathbf{k}} \in \mathbb{C}$ and with a fixed finite frequency index set $I \subset \mathbb{Z}^d$ of cardinality $|I|$.

Let $X \subset [0, 2\pi)^d$ be a finite set of sampling points with $|X|$ elements. Now we are interested in solving the following two problems:

(i) *Evaluation of trigonometric polynomials*. For given Fourier coefficients $\hat{p}_{\mathbf{k}}$, $\mathbf{k} \in I$, how to compute the polynomial values $p(\mathbf{x})$ for all $\mathbf{x} \in X$ efficiently?

(ii) *Evaluation of the Fourier coefficients*. For given polynomial values $p(\mathbf{x})$, $\mathbf{x} \in X$, how to compute $\hat{p}_{\mathbf{k}}$ for all $\mathbf{k} \in I$ efficiently?

The second problem also involves the question, how the sampling set X has to be chosen such that $\hat{p}_{\mathbf{k}}$ for all $\mathbf{k} \in I$ can be uniquely computed in a stable way.

Let us consider the $|X|$-by-$|I|$ *Fourier matrix* $\mathbf{A} = \mathbf{A}(X, I)$ defined by

$$\mathbf{A} = \mathbf{A}(X, I) := \left(\mathrm{e}^{\mathrm{i}\,\mathbf{k}\cdot\mathbf{x}}\right)_{\mathbf{x}\in X,\,\mathbf{k}\in I} \in \mathbb{C}^{|X|\times|I|}\,,$$

as well as the two vectors $\mathbf{p} := (p(\mathbf{x}))_{\mathbf{x}\in X} \in \mathbb{C}^{|X|}$ and $\hat{\mathbf{p}} := (\hat{p}(\mathbf{k}))_{\mathbf{k}\in I} \in \mathbb{C}^{|I|}$. To solve problem (i), we need to perform the matrix-vector multiplication

$$\mathbf{p} = \mathbf{A}\,\hat{\mathbf{p}}\,. \tag{49}$$

To compute $\hat{\mathbf{p}}$ from $\mathbf{p}$, we have to solve the inverse problem. For arbitrary polynomial $p \in \Pi_I$ this problem is only uniquely solvable, if $|X| \geq |I|$ and if $\mathbf{A}$ possesses full rank $|I|$. In other words, the sampling set X needs to be large enough and the obtained samples need to contain "enough information" about p. Then $\mathbf{A}^{\mathrm{H}}\,\mathbf{A} \in \mathbb{C}^{|I|\times|I|}$ is invertible, and we have

$$\hat{\mathbf{p}} = (\mathbf{A}^{\mathrm{H}}\,\mathbf{A})^{-1}\,\mathbf{A}^{\mathrm{H}}\,\mathbf{p}\,. \tag{50}$$

In order to ensure stability of this procedure, we want to assume that the columns of $\mathbf{A}$ are orthogonal, i.e., $\mathbf{A}^{\mathrm{H}}\mathbf{A} = M\,\mathbf{I}_{|I|}$, where $\mathbf{I}_{|I|}$ is the $|I|$-by-$|I|$ unit matrix and $M = |X|$. Then (50) simplifies to

$$\hat{\mathbf{p}} = \frac{1}{M}\,\mathbf{A}^{\mathrm{H}}\,\mathbf{p}\,.$$

In the following, we will consider very special sampling sets X, so-called rank-1 lattices.

Rank-1 lattices

Initially, rank-1 lattices were introduced as sampling schemes for (equally weighted) cubature formulas in the late 1950's and 1960's, see [35]. A summary of the early work on cubature rules based on rank-1 lattice sampling can be found in [46]. The recent increased interest in rank-1 lattices is particularly caused by new approaches to describe lattice rules that allow optimal theoretical error estimates for cubature formulas for specific function classes, see e.g. [60]. We also refer to [58] for a survey on lattice methods for numerical integration.
In contrast to general lattices which are spanned by several vectors, we consider only sampling on so-called rank-1 lattices. This simplifies the evaluation of trigonometric polynomials essentially and allows to derive necessary and sufficient conditions for unique or stable reconstruction.

For a given vector $\mathbf{z} \in \mathbb{Z}^d$ and a positive integer $M \in \mathbb{N}$ we define the *rank*-1 *lattice*

$$X = \Lambda(\mathbf{z}, M) := \left\{\mathbf{x}_j := \frac{2\pi}{M}(j\,\mathbf{z} \bmod M\,\mathbf{1}) \in [0, 2\pi)^d : j = 0, \ldots, M-1\right\} \tag{51}$$

as spatial discretization in $[0,\, 2\pi)^d$. Here, $\mathbf{1} := (1)_{s=1}^d \in \mathbb{Z}^d$ and for $\mathbf{z} = (z_s)_{s=1}^d \in \mathbb{Z}^d$ the term $j\,\mathbf{z} \bmod M\,\mathbf{1}$ denotes the vector $(j\,z_s \bmod M)_{s=1}^d$. We call $\mathbf{z}$ the *generating vector* and M the *lattice size* of the rank-1 lattice $\Lambda(\mathbf{z}, M)$. To ensure that $\Lambda(\mathbf{z}, M)$ has exactly M distinct elements, we assume that M is coprime with at least one component of $\mathbf{z}$. Further, for a given rank-1 lattice $\Lambda(\mathbf{z}, M)$ with generating vector $\mathbf{z} \in \mathbb{Z}^d$ we call the set

$$\Lambda^{\perp}(\mathbf{z}, M) := \{\mathbf{k} \in \mathbb{Z}^d : \mathbf{k} \cdot \mathbf{z} \equiv 0 \bmod M\} \tag{52}$$

the *integer dual lattice* of $\Lambda(\mathbf{z}, M)$. The integer dual lattice $\Lambda^{\perp}(\mathbf{z}, M)$ will play an important role, when we approximate the Fourier coefficients of a function f using only samples of f on the rank-1 lattice $\Lambda(\mathbf{z}, M)$.

Example 57

Let $d = 2$, $\mathbf{z} = (1,\, 3)^\top$ and $M = 11$, then we obtain

$$\begin{aligned}\Lambda(\mathbf{z}, M) &= \frac{2\pi}{11}\Big\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \end{pmatrix}, \begin{pmatrix} 2 \\ 6 \end{pmatrix}, \begin{pmatrix} 3 \\ 9 \end{pmatrix}, \begin{pmatrix} 4 \\ 1 \end{pmatrix}, \begin{pmatrix} 5 \\ 4 \end{pmatrix}, \\ &\quad \begin{pmatrix} 6 \\ 7 \end{pmatrix}, \begin{pmatrix} 7 \\ 10 \end{pmatrix}, \begin{pmatrix} 8 \\ 2 \end{pmatrix}, \begin{pmatrix} 9 \\ 5 \end{pmatrix}, \begin{pmatrix} 10 \\ 8 \end{pmatrix} \Big\},\end{aligned}$$

and $\Lambda^\perp(\mathbf{z}, M)$ contains all vectors $\mathbf{k} = (k_1,\, k_2)^\top \in \mathbb{Z}^2$ with $k_1 + 3k_2 \equiv 0 \bmod 11$. Figure 3 illustrates the construction of this two-dimensional rank-1 lattice.

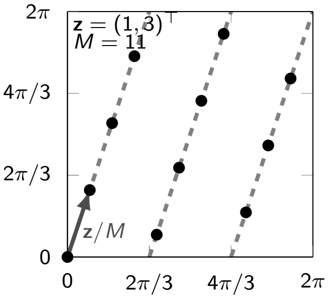


Figure 3: Rank-1 lattice $\Lambda(\mathbf{z}, M)$ of Example 57.

A rank-1 lattice possesses the following important property:

Lemma 58

Let a frequency index set $I \subset \mathbb{Z}^d$ of finite cardinality and a rank-1 lattice $X = \Lambda(\mathbf{z}, M)$ be given.
Then two distinct columns of the corresponding M-by-$|I|$ Fourier matrix $\mathbf{A}$ are either orthogonal or equal, i.e., the $(\mathbf{h}, \mathbf{k})$th entry $(\mathbf{A}^{\mathrm{H}}\,\mathbf{A})_{\mathbf{h},\mathbf{k}} \in \{0, M\}$ for all $\mathbf{h}$, $\mathbf{k} \in I$.

Proof: The matrix $\mathbf{A}^{\mathrm{H}}\,\mathbf{A}$ contains all inner products of two columns of the Fourier matrix $\mathbf{A}$, i.e., the $(\mathbf{h}, \mathbf{k})$th entry $(\mathbf{A}^{\mathrm{H}}\,\mathbf{A})_{\mathbf{h},\mathbf{k}}$ is equal to the inner product of the $\mathbf{k}$th column and the $\mathbf{h}$th column of $\mathbf{A}$. For $\mathbf{k}\cdot\mathbf{z} \not\equiv \mathbf{h}\cdot\mathbf{z} \mod M$ we obtain

$$(\mathbf{A}^{\mathrm{H}}\,\mathbf{A})_{\mathbf{h},\mathbf{k}} = \sum_{j=0}^{M-1} \left(\mathrm{e}^{2\pi\mathrm{i}\,[(\mathbf{k}-\mathbf{h})\cdot\mathbf{z}]/M}\right)^j = \frac{\mathrm{e}^{2\pi\mathrm{i}\,(\mathbf{k}-\mathbf{h})\cdot\mathbf{z}} - 1}{\mathrm{e}^{2\pi\mathrm{i}\,[(\mathbf{k}-\mathbf{h})\cdot\mathbf{z}]/M} - 1} = 0\,,$$

since $\mathbf{k} - \mathbf{h} \in \mathbb{Z}^d$.
For $\mathbf{k}\cdot\mathbf{z} \equiv \mathbf{h}\cdot\mathbf{z} \mod M$ it follows immediately that the $\mathbf{k}$th and $\mathbf{h}$th column of $\mathbf{A}$ are equal and that $(\mathbf{A}^{\mathrm{H}}\,\mathbf{A})_{\mathbf{h},\mathbf{k}} = M$. ■

Evaluation of trigonometric polynomials on rank-1 lattice

Let us now consider the efficient evaluation of a d-variate trigonometric polynomial p supported on I on the sampling set X being a rank-1 lattice $X = \Lambda(\mathbf{z}, M)$. We have to compute $p(\mathbf{x}_j)$ for all M nodes $\mathbf{x}_j \in \Lambda(\mathbf{z}, M)$, i.e.,

$$p(\mathbf{x}_j) = \sum_{\mathbf{k}\in I} \hat{p}_{\mathbf{k}} \, \mathrm{e}^{\mathrm{i}\, \mathbf{k}\cdot\mathbf{x}_j} = \sum_{\mathbf{k}\in I} \hat{p}_{\mathbf{k}} \, \mathrm{e}^{2\pi\mathrm{i}\, j\, (\mathbf{k}\cdot\mathbf{z})/M}, \quad j = 0, \ldots, M-1.$$

We observe that $\{\mathbf{k}\cdot\mathbf{z} \bmod M : \mathbf{k}\in I\} \subset \{0, \ldots, M-1\}$ and consider the values

$$\hat{g}_\ell = \sum_{\substack{\mathbf{k}\in I \\ \ell \equiv \mathbf{k}\cdot\mathbf{z} \bmod M}} \hat{p}_{\mathbf{k}}\,, \quad \ell = 0, \ldots, M-1\,. \tag{53}$$

Then, we can write

$$
\begin{aligned}
p(\mathbf{x}_j) &= \sum_{\mathbf{k}\in I} \hat{p}_{\mathbf{k}}\, \mathrm{e}^{2\pi \mathrm{i}\, j\,(\mathbf{k}\cdot\mathbf{z})/M} = \sum_{\ell=0}^{M-1} \sum_{\substack{\mathbf{k}\in I \\ \ell \equiv \mathbf{k}\cdot\mathbf{z} \bmod M}} \hat{p}_{\mathbf{k}}\, \mathrm{e}^{2\pi \mathrm{i}\, j\ell/M} \\
&= \sum_{\ell=0}^{M-1} \hat{g}_{\ell}\, \mathrm{e}^{2\pi \mathrm{i}\, j\ell/M} \qquad (54)
\end{aligned}
$$

for $j = 0, \ldots, M-1$. Therefore, the right-hand side of (54) can be evaluated using a one-dimensional FFT of length M with at most $C \cdot (M \log M + d\,|I|)$ arithmetic operations, where the constant C does not depend on the dimension d. Here we assume that $\hat{g}_{\ell}$, $\ell = 0, \ldots, M$, can be computed with $C\, d\, |I|$ arithmetic operations. The fast realization of the matrix-vector product in (49) or equivalently of (54) is presented in the following

Algorithm: (Lattice based FFT (LFFT))

Input: $M \in \mathbb{N}$ lattice size of rank-1 lattice $\Lambda(\mathbf{z}, M)$,
$\mathbf{z} \in \mathbb{Z}^d$ generating vector of $\Lambda(\mathbf{z}, M)$,
$I \subset \mathbb{Z}^d$ finite frequency index set,
$\hat{\mathbf{p}} = (\hat{p}_{\mathbf{k}})_{\mathbf{k}\in I}$ Fourier coefficients of $p \in \Pi_I$.

1. Set $\hat{\mathbf{g}} := (0)_{\ell=0}^{M-1}$.
2. For each $\mathbf{k} \in I$ do $\hat{g}_{\mathbf{k}\cdot\mathbf{z} \bmod M} := \hat{g}_{\mathbf{k}\cdot\mathbf{z} \bmod M} + \hat{p}_{\mathbf{k}}$ endfor.
3. Apply a one-dimensional FFT of length M in order to compute $\mathbf{p} := \mathbf{F}_M^{-1}\left((\hat{g}_\ell)_{\ell=0}^{M-1}\right)$.
4. Compute $\mathbf{p} := M\,\mathbf{p}$.

Output: $\mathbf{p} = \mathbf{A}\,\hat{\mathbf{p}}$ vector of values of the trigonometric polynomial $p \in \Pi_I$.
Arithmetic cost: $\mathcal{O}(M \log M + d|I|)$.

Adjoint single lattice based FFT (aLFFT)

We immediately obtain also a fast algorithm for the matrix-vector multiplication with the adjoint Fourier matrix $\mathbf{A}^{\mathrm{H}}$.

Input: $M \in \mathbb{N}$ lattice size of rank-1 lattice $\Lambda(\mathbf{z}, M)$,
$\mathbf{z} \in \mathbb{Z}^d$ generating vector of $\Lambda(\mathbf{z}, M)$,
$I \subset \mathbb{Z}^d$ finite frequency index set,
$\mathbf{p} = \left(p(\frac{j}{M}\,\mathbf{z})\right)_{j=0}^{M-1}$ values of the trigonometric polynomial $p \in \Pi_I$.

1. Apply a one-dimensional FFT of length M in order to compute $\hat{\mathbf{g}} := \mathbf{F}_M\,\mathbf{p}$.
2. Set $\hat{\mathbf{a}} := (0)_{\mathbf{k} \in I}$.
3. For each $\mathbf{k} \in I$ do $\hat{a}_{\mathbf{k}} := \hat{a}_{\mathbf{k}} + \hat{g}_{\mathbf{k}\cdot\mathbf{z} \bmod M}$ endfor.

Output: $\hat{\mathbf{a}} = \mathbf{A}^{\mathrm{H}}\,\mathbf{p}$ with the adjoint Fourier matrix $\mathbf{A}^{\mathrm{H}}$.
Arithmetic cost: $\mathcal{O}(M \log M + d\,|I|)$.

Evaluation of the Fourier coefficients

Our considerations of the Fourier matrix $\mathbf{A} = \mathbf{A}(X, I)$ in (49) and (50) show that a unique evaluation of all Fourier coefficients of an arbitrary d-variate trigonometric polynomial $p \in \Pi_I$ is only possible, if the $|X|$-by-$|I|$ matrix $\mathbf{A}$ has full rank $|I|$. By Lemma 58 we have seen that for a given frequency index set I and a rank-1 lattice $\Lambda(\mathbf{z}, M)$, two distinct columns of $\mathbf{A}$ are either orthogonal or equal. Therefore, $\mathbf{A}$ has full rank if and only if for all distinct $\mathbf{k}$, $\mathbf{h} \in I$,

$$\mathbf{k} \cdot \mathbf{z} \not\equiv \mathbf{h} \cdot \mathbf{z} \mod M\,. \tag{55}$$

If (55) holds, then the sums determining $\hat{g}_\ell$ in (53) contain only one term for each ℓ and no aliasing occurs. We define the *difference set of the frequency index set I* as

$$\mathcal{D}(I) := \{\mathbf{k} - \mathbf{l} \colon \mathbf{k}, \mathbf{l} \in I\}\,. \tag{56}$$

Then the condition (55) is equivalent to

$$\mathbf{k} \cdot \mathbf{z} \not\equiv 0 \mod M \quad \text{for all } \mathbf{k} \in \mathcal{D}(I) \setminus \{\mathbf{0}\}. \tag{57}$$

Therefore, we define a *reconstructing rank*-1 *lattice* to a given frequency index set I as a rank-1 lattice satisfying (55) or equivalently (57) and denote it by

$$\Lambda(\mathbf{z}, M, I) := \left\{ \mathbf{x} \in \Lambda(\mathbf{z}, M) : \mathbf{k} \in \mathcal{D}(I) \setminus \{\mathbf{0}\} \,\mathrm{with}\, \mathbf{k} \cdot \mathbf{z} \not\equiv 0 \mod M \right\}.$$

The condition (57) ensures that the mapping of $\mathbf{k} \in I$ to $\mathbf{k} \cdot \mathbf{z} \mod M \in \{0, \ldots, M-1\}$ is injective. Assuming that we have a reconstructing rank-1 lattice, we will be able to evaluate the Fourier coefficients of $p \in \Pi_I$ uniquely.
If condition (57) is satisfied, then Lemma 58 implies $\mathbf{A}^{\mathrm{H}} \mathbf{A} = M \, \mathbf{I}_M$ for the Fourier matrix $\mathbf{A}$ such that $\hat{\mathbf{p}} = (\hat{p}_{\mathbf{k}})_{\mathbf{k} \in I} = \frac{1}{M} \mathbf{A}^{\mathrm{H}} \mathbf{p}$.

Equivalently, for each Fourier coefficient we have

$$\hat{p}_{\mathbf{k}} = \frac{1}{M} \sum_{j=0}^{M-1} p(\mathbf{x}_j)\, \mathrm{e}^{-2\pi \mathrm{i}\, j\, (\mathbf{k}\cdot\mathbf{z})/M} = \frac{1}{M} \sum_{j=0}^{M-1} p(\mathbf{x}_j)\, \mathrm{e}^{-2\pi \mathrm{i}\, j\ell/M}$$

for all $\mathbf{k} \in I$ and $\ell = \mathbf{k} \cdot \mathbf{z} \bmod M$. Algorithm 164 computes all Fourier coefficients $\hat{f}_{\mathbf{k}}$ using only a one-dimensional FFT of length M and the inverse mapping of $\mathbf{k} \mapsto \mathbf{k} \cdot \mathbf{z} \bmod M$, see also [24, Algorithm 3.2].

Reconstruction via reconstructing rank-1 lattice

Input: $I \subset \mathbb{Z}^d$ finite frequency index set,
$M \in \mathbb{N}$ lattice size of reconstructing rank-1 lattice $\Lambda(\mathbf{z}, M, I)$,
$\mathbf{z} \in \mathbb{Z}^d$ generating vector of reconstructing rank-1 lattice $\Lambda(\mathbf{z}, M, I)$,
$\mathbf{p} = \left(p\left(\frac{2\pi}{M}\,(j\,\mathbf{z} \bmod M\,\mathbf{1})\right)\right)_{j=0}^{M-1}$ values of $p \in \Pi_I$.

1. Compute $\hat{\mathbf{a}} := \mathbf{A}^{\mathrm{H}}\,\mathbf{p}$ using Algorithm above.
2. Set $\hat{\mathbf{p}} := M^{-1}\,\hat{\mathbf{a}}$.

Output: $\hat{\mathbf{p}} = M^{-1}\mathbf{A}^{\mathrm{H}}\,\mathbf{p} = (\hat{p}_{\mathbf{k}})_{\mathbf{k}\in I}$ Fourier coefficients supported on I.

Arithmetic cost: $\mathcal{O}(M\,\log M + d\,|I|)$.

Example 59

Let $I^d_{\infty,N}$ be the full grid defined by (44). Then straightforward calculation shows that the rank-1 lattice $\Lambda(\mathbf{z}, M)$ with the generating vector $\mathbf{z} = (1, 2N+2, \ldots, (2N+2)^{d-1})^\top$ and the lattice size $M = (2N+2)^d$ is a reconstructing rank-1 lattice to the full grid $I^d_{\infty,N}$. It provides a perfectly stable spatial discretization. The resulting reconstruction algorithm is based on a one-dimensional FFT of size $(2N+2)^d$, and has similar arithmetic costs as the usual d-dimensional tensor-product FFT. Our goal is to construct smaller reconstructing rank-1 lattices for special index sets, such that the arithmetic cost for the reconstruction of Fourier coefficients can be significantly reduced. $\square$

As a corollary of the observations above we show that a reconstructing rank-1 lattice implies the following important quadrature rule, see [59].

Theorem 60

For a given finite frequency index set I and a corresponding reconstructing rank-1 lattice $\Lambda(\mathbf{z}, M, I)$ we have

$$\int_{[0,2\pi]^d} p(\mathbf{x})\,\mathrm{d}\mathbf{x} = \frac{1}{M} \sum_{j=0}^{M-1} p(\mathbf{x}_j)$$

for all trigonometric polynomials $p \in \Pi_{\mathcal{D}(I)}$, where $\mathcal{D}(I)$ is defined by (56).

Proof: For $\mathbf{x}_j = \frac{2\pi}{M}\,(j\mathbf{z} \mod M\,\mathbf{1}) \in \Lambda(\mathbf{z}, M, I)$ it follows that

$$\begin{aligned}\sum_{j=0}^{M-1} p(\mathbf{x}_j) &= \sum_{j=0}^{M-1} \Big(\sum_{\mathbf{k}\in\mathcal{D}(I)} \hat{p}_{\mathbf{k}}\, \mathrm{e}^{2\pi \mathrm{i}\, j\, (\mathbf{k}\cdot\mathbf{z})/M} \Big) \\ &= \sum_{\mathbf{k}\in\mathcal{D}(I)} \hat{p}_{\mathbf{k}} \Big(\sum_{j=0}^{M-1} \mathrm{e}^{2\pi \mathrm{i}\, j\, (\mathbf{k}\cdot\mathbf{z})/M} \Big)\,.\end{aligned}$$

According to (57) we have $\mathbf{k}\cdot\mathbf{z} \not\equiv 0 \mod M$ for all $\mathbf{k} \in \mathcal{D}(I) \setminus \{\mathbf{0}\}$. Therefore

$$\sum_{j=0}^{M-1} \mathrm{e}^{2\pi \mathrm{i}\, j\, (\mathbf{k}\cdot\mathbf{z})/M} = \begin{cases} 0 & \mathbf{k} \in \mathcal{D}(I) \setminus \{\mathbf{0}\}\,, \\ M & \mathbf{k} = \mathbf{0}\,, \end{cases}$$

and the equation above simplifies to

$$\sum_{j=0}^{M-1} p(\mathbf{x}_j) = M\,\hat{p}(\mathbf{0}) = M \int_{[0,\,2\pi]^d} p(\mathbf{x})\,\mathrm{d}\mathbf{x}. \quad \blacksquare$$

Efficient function approximation on rank-1 lattices

Now we come back to the problem of approximation of a smooth d-variate periodic function f by a Fourier series (42) or by a Fourier partial sum (43). Let f be an arbitrary continuous function in $\mathcal{A}(\mathbb{T}^d) \cap C(\mathbb{T}^d)$. Then we determine approximate values $\hat{f}_{\mathbf{k}}$ of the Fourier coefficients $c_{\mathbf{k}}(f)$ using only the sampling values on a rank-1 lattice $\Lambda(\mathbf{z}, M)$ as given in (51) and obtain

$$\begin{aligned}\hat{f}_{\mathbf{k}} &:= \frac{1}{M} \sum_{j=0}^{M-1} f\Big(\frac{2\pi}{M}\,(j\,\mathbf{z} \bmod M\,\mathbf{1})\Big)\mathrm{e}^{-2\pi\mathrm{i}\,j\,(\mathbf{k}\cdot\mathbf{z})/M} \qquad (58)\\ &= \frac{1}{M} \sum_{j=0}^{M-1} \sum_{\mathbf{h}\in\mathbb{Z}^d} c_{\mathbf{h}}(f)\,\mathrm{e}^{2\pi\mathrm{i}\,j\,[(\mathbf{h}-\mathbf{k})\cdot\mathbf{z}]/M}\\ &= \sum_{\mathbf{h}\in\mathbb{Z}^d} c_{\mathbf{k}+\mathbf{h}}(f)\,\frac{1}{M}\sum_{j=0}^{M-1} \mathrm{e}^{2\pi\mathrm{i}\,j\,(\mathbf{h}\cdot\mathbf{z})/M} = \sum_{\mathbf{h}\in\Lambda^{\perp}(\mathbf{z},M)} c_{\mathbf{k}+\mathbf{h}}(f)\,,\end{aligned}$$

where the integer dual lattice $\Lambda^{\perp}(\mathbf{z}, M)$ is defined by (52).

Obviously we have $\mathbf{0} \in \Lambda^\perp(\mathbf{z}, M)$ and hence

$$\hat{f}_{\mathbf{k}} = c_{\mathbf{k}}(f) + \sum_{\mathbf{h} \in \Lambda^\perp(\mathbf{z}, M) \setminus \{\mathbf{0}\}} c_{\mathbf{k}+\mathbf{h}}(f). \tag{59}$$

The absolute convergence of the series of the Fourier coefficients of f ensures that all terms in the calculation above are well-defined. We call $\hat{f}_{\mathbf{k}}$ the *approximate Fourier coefficients* of f. The formula (59) can be understood as an *aliasing formula for the rank-*1 *lattice* $\Lambda(\mathbf{z}, M)$. If the sum

$$\sum_{\mathbf{h} \in \Lambda^\perp(\mathbf{z}, M) \setminus \{\mathbf{0}\}} |c_{\mathbf{k}+\mathbf{h}}(f)|$$

is sufficiently small, then $\hat{f}_{\mathbf{k}}$ is a convenient approximate value of $c_{\mathbf{k}}(f)$.

Assume that f can be already well approximated by a trigonometric polynomial p on a frequency index set I. Further, assume that we have a corresponding reconstructing rank-1 lattice $X = \Lambda(\mathbf{z}, M, I)$. Then we can compute the approximative Fourier coefficients $\hat{f}_{\mathbf{k}}$ with $\mathbf{k} \in I$ using Algorithm 164 by employing M sample values $f\big(\frac{2\pi}{M}(j\,\mathbf{z} \bmod M\,\mathbf{1})\big)$ instead of the corresponding polynomial values. In this way, we obtain $\hat{f}_{\mathbf{k}}$, $\mathbf{k} \in I$, with arithmetical costs of $\mathcal{O}(M \log M + d\,|I|)$.

Now we want to study the approximation error that occurs if the exact Fourier coefficients $c_{\mathbf{k}}(f)$ are replaced by the approximate Fourier coefficients $\hat{f}_{\mathbf{k}}$ in (59). We consider the corresponding approximate Fourier partial sum on the frequency index set $I_N = \{\mathbf{k} \in \mathbb{Z}^d : \omega(\mathbf{k}) \le N\}$. Let $\Lambda(\mathbf{z}, M, I_N)$ be a reconstructing rank-1 lattice for I_N and $\Lambda^{\perp}(\mathbf{z}, M, I_N)$ the corresponding integer dual lattice (52). By definition of the reconstructing rank-1 lattice it follows that $I_N \cap \Lambda^{\perp}(\mathbf{z}, M, I_N) = \{\mathbf{0}\}$. Generally we can show the following result:

Lemma 61

Let $I \subset \mathbb{Z}^d$ be an arbitrary finite frequency index set and let $\Lambda(\mathbf{z}, M, I)$ be a reconstructing rank-1 lattice with the integer dual lattice $\Lambda^{\perp}(\mathbf{z}, M, I)$.
Then we have

$$\left\{\mathbf{k} + \mathbf{h} : \mathbf{k} \in I, \mathbf{h} \in \Lambda^{\perp}(\mathbf{z}, M, I) \setminus \{\mathbf{0}\}\right\} \subset \mathbb{Z}^d \setminus I \,.$$

Proof: Assume to the contrary that there exist $\mathbf{k} \in I$ and $\mathbf{h} \in \Lambda^{\perp}(\mathbf{z}, M, I) \setminus \{\mathbf{0}\}$ such that $\mathbf{k} + \mathbf{h} \in I$. Since $\Lambda(\mathbf{z}, M, I)$ is a reconstructing rank-1 lattice for I, it follows that $\mathbf{0} \neq \mathbf{h} = (\mathbf{k} + \mathbf{h}) - \mathbf{k} \in \mathcal{D}(I)$. Thus, $\mathbf{h} \in \mathcal{D}(I) \cap \Lambda^{\perp}(\mathbf{z}, M, I) \setminus \{\mathbf{0}\}$. But this is a contradiction, since on the one hand (57) implies that $\mathbf{h} \cdot \mathbf{z} \not\equiv 0 \mod M$, and on the other hand $\mathbf{h} \cdot \mathbf{z} \equiv 0 \mod M$ by definition of $\Lambda^{\perp}(\mathbf{z}, M, I)$. ■

Theorem 62

Let $f \in \mathcal{A}_\omega(\mathbb{T}^d)$ *and let a frequency index set* $I_N = \{\mathbf{k} \in \mathbb{Z}^d : \omega(\mathbf{k}) \le N\}$ *of finite cardinality be given. Further, let* $\Lambda(\mathbf{z}, M, I_N)$ *be a reconstructing rank-1 lattice for* I_N*. Moreover, let the approximate Fourier partial sum*

$$(S_{I_N}^{\Lambda} f)(\mathbf{x}) := \sum_{\mathbf{k} \in I_N} \hat{f}_{\mathbf{k}} \, \mathrm{e}^{\mathrm{i}\,\mathbf{k} \cdot \mathbf{x}} \tag{60}$$

of f *be determined by*

$$\hat{f}_{\mathbf{k}} := \frac{1}{M} \sum_{j=0}^{M-1} f\Big(\frac{2\pi}{M}\,(j\,\mathbf{z} \mod M\,\mathbf{1})\Big)\, \mathrm{e}^{-2\pi\mathrm{i}\,j\,(\mathbf{k}\cdot\mathbf{z})/M}\,, \quad \mathbf{k} \in I_N\,, \tag{61}$$

that are computed using the values on the rank-1 lattice $\Lambda(\mathbf{z}, M, I_N)$*.*
Then we have

$$\| f - S_{I_N}^{\Lambda} f \|_{L_\infty(\mathbb{T}^d)} \le 2\, N^{-1}\, \| f \|_{\mathcal{A}_\omega(\mathbb{T}^d)}\,. \tag{62}$$

Proof: Using the triangle inequality, we find

$$\|f - S_{I_N}^{\Lambda} f\|_{L_\infty(\mathbb{T}^d)} \le \|f - S_{I_N} f\|_{L_\infty(\mathbb{T}^d)} + \|S_{I_N}^{\Lambda} f - S_{I_N} f\|_{L_\infty(\mathbb{T}^d)} .$$

For the first term, Lemma 54 yields

$$\|f - S_{I_N} f\|_{L_\infty(\mathbb{T}^d)} \le N^{-1} \|f\|_{\mathcal{A}_\omega(\mathbb{T}^d)} .$$

For the second term we obtain by using (59)

$$\begin{aligned}
\|S_{I_N}^{\Lambda} f - S_{I_N} f\|_{L_\infty(\mathbb{T}^d)} &= \operatorname*{ess\,sup}_{\mathbf{x} \in \mathbb{T}^d} \Big| \sum_{\mathbf{k} \in I_N} \big(\hat{f}_{\mathbf{k}} - c_{\mathbf{k}}(f) \big) \, \mathrm{e}^{\mathrm{i}\, \mathbf{k} \cdot \mathbf{x}} \Big| \\
&\le \sum_{\mathbf{k} \in I_N} \Big| \sum_{\mathbf{h} \in \Lambda^{\perp}(\mathbf{z}, M) \setminus \{\mathbf{0}\}} c_{\mathbf{k}+\mathbf{h}}(f) \Big| \\
&\le \sum_{\mathbf{k} \in I_N} \sum_{\mathbf{h} \in \Lambda^{\perp}(\mathbf{z}, M) \setminus \{\mathbf{0}\}} |c_{\mathbf{k}+\mathbf{h}}(f)| .
\end{aligned}$$

By Lemma 61 it follows that

$$\begin{aligned}\|S_{I_N}^{\Lambda} f - S_{I_N} f\|_{L_\infty(\mathbb{T}^d)} &\leq \sum_{\mathbf{k}\in\mathbb{Z}^d\setminus I_N} |c_{\mathbf{k}}(f)| \leq \frac{1}{\inf_{\mathbf{h}\in\mathbb{Z}^d\setminus I_N}\omega(\mathbf{h})} \sum_{\mathbf{k}\in\mathbb{Z}^d} \omega(\mathbf{k})\,|c_{\mathbf{k}}(f)| \\ &\leq N^{-1}\|f\|_{\mathcal{A}_\omega(\mathbb{T}^d)}\end{aligned}$$

and hence the assertion. ■

Theorem 62 states that the worst case error of the approximation $S_{I_N}^{\Lambda} f$ in (60) given by the approximate Fourier coefficients computed from samples on the reconstructing rank-1 lattice $\Lambda(\mathbf{z}, M, I_N)$, is qualitatively as good as the worst case error of the approximation $S_{I_N} f$, see (46). Improved error estimates for the approximation of functions in $\mathcal{A}_\omega(\mathbb{T}^d)$ with a special weight function ω as in Remark 55 can be similarly derived. The approximation error essentially depends on the considered norms. In particular, we have focussed on the $L_\infty(\mathbb{T}^d)$-norm on the left-hand side and the weighted $\ell_1(\mathbb{Z}^d)$-norm of the Fourier coefficients on the right-hand side. Further results with different norms are given in [24, 66].

Remark 63

The idea to use special rank-1 lattices $\Lambda(\mathbf{z}, M)$ *of Korobov type as sampling schemes to approximate functions by trigonometric polynomials has been already considered by V.N. Temlyakov [62]. Later, D. Li and F.J. Hickernell studied a more general setting in [39]. They presented an approximation error using an aliasing formula as* (59) *for the given rank-1 lattice* $\Lambda(\mathbf{z}, M)$*. But both approaches did not lead to a constructive way to determine rank-1 lattices of high quality. In contrast to their approach, we have constructed the frequency index set* $I_N := \{\mathbf{k} \in \mathbb{Z}^d : \omega(\mathbf{k}) \leq N\}$ *with* $|I_N| < \infty$ *depending on the arbitrary weight function* ω*. The problem to find a reconstructing rank-1 lattice* $\Lambda(\mathbf{z}, M, I_N)$ *which is well adapted to the frequency index set* I_N *will be studied in the next section. Approximation properties of rank-1 lattices have been also investigated in information based complexity and applied analysis, see e.g. [72, 38, 45].* $\square$

Reconstructing rank-1 lattices

As shown in the two last sections, we can use so-called reconstructing rank-1 lattices in order to compute the Fourier coefficients of a d-variate trigonometric polynomial in Π_I in a stable way by applying a one-dimensional FFT. The reconstructing rank-1 lattice $\Lambda(\mathbf{z}, M, I)$ for a frequency index set I is determined as a rank-1 lattice $\Lambda(\mathbf{z}, M)$ in (51) satisfying the condition (55). The arithmetic costs to reconstruct the Fourier coefficients of the d-variate trigonometric polynomial p from its sampling values of the given rank-1 lattice mainly depend on the number M of needed sampling values. In this section we will present a deterministic procedure to obtain reconstructing rank-1 lattices using a component-by-component approach. We start with considering the problem, how large the number M of sampling values in $\Lambda(\mathbf{z}, M, I)$ needs to be, see also [25, 28]. For simplicity, we consider only a *symmetric frequency index set* $I \subset \mathbb{Z}^d$ satisfying the condition that for each $\mathbf{k} \in I$ also $-\mathbf{k} \in I$. For example, all frequency index sets in Example 57 and Figure 2 are symmetric.

Theorem 64

Let I be a symmetric frequency index set with finite cardinality $|I|$ such that $I \subset [-\frac{|I|}{2}, \frac{|I|}{2}]^d \cap \mathbb{Z}^d$.
Then there exists a reconstructing rank-1 lattice $X = \Lambda(\mathbf{z}, M, I)$ with prime cardinality M, such that

$$|I| \le M \le |\mathcal{D}(I)| \le |I|^2 - |I| + 1, \qquad (63)$$

where $\mathcal{D}(I)$ denotes the difference set (56).

Proof: 1. The lower bound $|I| \le M$ is obvious, since we need a Fourier matrix $\mathbf{A} = \mathbf{A}(X, I) \in \mathbb{C}^{|X| \times |I|}$ of full rank $|I|$ in (50) to reconstruct $\hat{\mathbf{p}}$, and this property follows from (55).
Recall that $|\mathcal{D}(I)|$ is the number of all pairwise distinct vectors $\mathbf{k} - \mathbf{l}$ with $\mathbf{k}, \mathbf{l} \in I$. We can form at most $|I|\,(|I| - 1) + 1$ pairwise distinct vectors in $\mathcal{D}(I)$. Therefore we obtain the bound $|\mathcal{D}(I)| \le |I|^2 - |I| + 1$.

2. In order to show that there exists a reconstructing rank-1 lattice with $M \le |\mathcal{D}(I)|$, we choose M as a prime number satisfying $|\mathcal{D}(I)|/2 < M \le |\mathcal{D}(I)|$ and show that there exists a generating vector $\mathbf{z}$ such that the condition (57) is satisfied for $X = \Lambda(\mathbf{z}, M, I)$. The prime number M can be always chosen in $(|\mathcal{D}(I)|/2, \, |\mathcal{D}(I)|]$ by Bertrand's postulate.
For the special case $d = 1$ we have $I \subset [-\frac{|I|}{2}, \frac{|I|}{2}] \cap \mathbb{Z}$. Taking $z = z_1 = 1$, each $M \ge |I| + 1$ satisfies the assumption $k \cdot z = k \not\equiv 0 \bmod M$ for $k \in \mathcal{D}(I) \subset [-|I|, |I|]$. In particular, we can take M as a prime number in $(|\mathcal{D}(I)|/2, \, |\mathcal{D}(I)|]$, since we have $|\mathcal{D}(I)| \ge 2\,|I|$ in this case.
Let us now assume that $d \ge 2$. We need to show that there exists a generating vector $\mathbf{z}$ such that

$$\mathbf{k} \cdot \mathbf{z} \not\equiv 0 \bmod M \quad \text{for all } \mathbf{k} \in \mathcal{D}(I) \setminus \{\mathbf{0}\},$$

and want to use an induction argument with respect to the dimension d.

We consider the projection of $\mathcal{D}(I)$ on the index set

$$\mathcal{D}(I_{d-1}) := \{\tilde{\mathbf{k}} = (k_j)_{j=1}^{d-1} : \mathbf{k} = (k_j)_{j=1}^{d} \in \mathcal{D}(I)\},$$

such that each $\mathbf{k} \in \mathcal{D}(I)$ can be written as $(\tilde{\mathbf{k}}^\top, k_d)^\top$ with $\tilde{\mathbf{k}} \in \mathcal{D}(I_{d-1})$. Assume that we have found already a vector $\tilde{\mathbf{z}} \in \mathbb{Z}^{d-1}$ such that the condition

$$\tilde{\mathbf{k}} \cdot \tilde{\mathbf{z}} \not\equiv 0 \bmod M \quad \text{for all } \tilde{\mathbf{k}} \in \mathcal{D}(I_{d-1}) \setminus \{\mathbf{0}\} \tag{64}$$

is satisfied. We show now that there exists a vector $\mathbf{z} = (\tilde{\mathbf{z}}^\top, z_d)^\top$ with $z_d \in \{1, \ldots, M-1\}$ such that

$$\mathbf{k} \cdot \mathbf{z} = \tilde{\mathbf{k}} \cdot \tilde{\mathbf{z}} + k_d z_d \not\equiv 0 \bmod M \quad \text{for all } \mathbf{k} \in \mathcal{D}(I) \setminus \{\mathbf{0}\}. \tag{65}$$

For that purpose we will use a counting argument. We show that there are at most $(|\mathcal{D}(I_{d-1})| - 1)/2$ integers $z_d \in \{1, \ldots, M-1\}$ with the property

$$\mathbf{k} \cdot \mathbf{z} = \tilde{\mathbf{k}} \cdot \tilde{\mathbf{z}} + k_d z_d \equiv 0 \bmod M \qquad \text{for at least one } \mathbf{k} \in \mathcal{D}(I) \setminus \{\mathbf{0}\}. \tag{66}$$

Since $(|\mathcal{D}(I_{d-1})| - 1)/2 \le (|\mathcal{D}(I)| - 1)/2 < M - 1$, we always find a z_d satisfying the desired condition (65).

3. We show now that for each pair of elements $\mathbf{k}$, $-\mathbf{k}$ with $\mathbf{k} = (\tilde{\mathbf{k}}^\top, k_d)^\top \in \mathcal{D}(I) \setminus \{\mathbf{0}\}$ and given $\tilde{\mathbf{z}}$ satisfying (64), there is at most one z_d such that (66) is satisfied.
If $k_d = 0$, then (66) yields $\tilde{\mathbf{k}} \cdot \tilde{\mathbf{z}} \equiv 0 \bmod M$ contradicting (64). Thus in this case no z_d is found to satisfy (66).
If $\tilde{\mathbf{k}} = \mathbf{0}$ and $k_d \neq 0$, then (66) yields $k_d z_d \equiv 0 \bmod M$. Since $|k_d| \leq |I| < M$ and $z_d \in \{1, \ldots, M-1\}$, it follows that $k_d z_d$ and M are coprime such that no z_d is found to satisfy (66).
If $\tilde{\mathbf{k}} \neq \mathbf{0}$ and $k_d \neq 0$, then (66) yields $\tilde{\mathbf{k}} \cdot \tilde{\mathbf{z}} \equiv -k_d z_d \bmod M$. Since $\tilde{\mathbf{k}} \cdot \tilde{\mathbf{z}} \neq 0$ by assumption (64) and k_d and M are coprime, there exists one unique solution z_d of this equation. The same unique solution z_d is found, if we replace $\mathbf{k} = (\tilde{\mathbf{k}}^\top, k_d)^\top$ by $-\mathbf{k} = (-\tilde{\mathbf{k}}^\top, -k_d)^\top$ in (66).
Taking into account that $\mathcal{D}(I_{d-1})$ and $\mathcal{D}(I)$ always contain the corresponding zero vector, it follows that at most $(|\mathcal{D}(I_{d-1})| - 1)/2$ integers satisfy (66). Thus the assertion is proved. ■
The idea of the proof of Theorem 64 leads us also to an algorithm, the so-called component-by-component Algorithm 182. This algorithm computes for a known lattice size M the generating vector $\mathbf{z}$ of the reconstructing rank-1 lattice, see also [25].

Component-by-component lattice search

Input: $M \in \mathbb{N}$ prime, cardinality of rank-1 lattice,
$I \subset \mathbb{Z}^d$ finite frequency index set.

1. Set $z_1 := 1$.
2. For $s = 2, \ldots, d$ do
   form the set $I_s := \left\{ (k_j)_{j=1}^{s} \colon \mathbf{k} = (k_j)_{j=1}^{d} \in I \right\}$
   search for one $z_s \in [1,\ M-1] \cap \mathbb{Z}$ with

$$\left|\left\{ (z_1, \ldots, z_s)^\top \cdot \mathbf{k} \bmod M : \ \mathbf{k} \in I_s \right\}\right| = |I_s| \, .$$

   end for.

Output: $\mathbf{z} = (z_j)_{j=1}^{d} \in \mathbb{N}^d$ generating vector.

The construction of the generating vector $\mathbf{z} \in \mathbb{N}^d$ in Algorithm 182 requires at most $2d\,|I|\,M \leq 2d\,|I|^3$ arithmetic operations. For each component z_s, $s \in \{2, \ldots, d\}$, of the generating vector $\mathbf{z}$ in the component-by-component step s, the tests for the reconstruction property (54) for a given component z_s, in step 2 of Algorithm 182 require at most $s\,|I|$ multiplications, $(s-1)\,|I|$ additions and $|I|$ modulo operations. Since each component z_s, $s \in \{2, \ldots, d\}$, of the generating vector $\mathbf{z}$ can only take $M-1$ possible values, the construction requires at most $d|I|\,(M-1) \leq 2\,d\,|I|\,M$ arithmetic operations in total.

Remark 65

The lower bound for the number M in Theorem 64 can be improved for arbitrary frequency index sets, if we employ the exact cardinalities of the projected index sets $I_s := \{(k_j)_{j=1}^{s} : \mathbf{k} = (k_j)_{j=1}^{d} \in I\}$, see also [25].
The assumption on the index set can be also relaxed. In particular, the complete index set can be shifted in $\mathbb{Z}^d$ without changing the results. $\square$

A drawback of Algorithm 182 is that the cardinality M needs to be known in advance. As we have shown in Theorem 64, M can be always taken as a prime number satisfying $|\mathcal{D}(I)|/2 < M \leq |\mathcal{D}(I)|$. But this may be far away from an optimal choice. Once we have discovered a reconstructing rank-1 lattice $\Lambda(\mathbf{z}, M, I)$ satisfying for all distinct $\mathbf{k}$, $\mathbf{h} \in I$,

$$\mathbf{k} \cdot \mathbf{z} \not\equiv \mathbf{h} \cdot \mathbf{z} \bmod M\,,$$

we can ask for $M' < M$ such that for all distinct $\mathbf{k}$, $\mathbf{h} \in I$,

$$\mathbf{k} \cdot \mathbf{z} \not\equiv \mathbf{h} \cdot \mathbf{z} \bmod M'$$

is still true for the computed generating vector $\mathbf{z}$. This leads to the following simple algorithm for lattice size decreasing, see also [25].

Lattice size decreasing

Input: $M \in \mathbb{N}$ cardinality of rank-1 lattice,
$I \subset \mathbb{Z}^d$ finite frequency index set,
$\mathbf{z} \in \mathbb{N}^d$ generating vector of reconstructing rank-1 lattice $\Lambda(\mathbf{z}, M, I)$.

1. For $j = |I|, \ldots, M$ do
   if $|\{\mathbf{k} \cdot \mathbf{z} \bmod j : \mathbf{k} \in I\}| = |I|$ then
   $M_{\text{min}} := j$, stop
   end if
   end for.

Output: M_{min} reduced lattice size.

There exist also other strategies to determine reconstructing rank-1 lattices for given frequency index sets, where the lattice size M needs not to be known a priori, see e.g. [25, Algorithms 4 and 5]. These algorithms are also component-by-component algorithms and compute complete reconstructing rank-1 lattices, i.e., the generating vectors $\mathbf{z} \in \mathbb{N}^d$ and the lattice sizes $M \in \mathbb{N}$, for a given frequency index set I. The algorithms are applicable for arbitrary frequency index sets of finite cardinality $|I|$.
As we have seen in Theorem 64 the sampling size M can be bounded by the cardinality of the difference set $\mathcal{D}(I)$. Interestingly, this cardinality strongly depends on the structure of I.

Example 66

Let $I = I^d_{p,N} := \{\mathbf{k} \in \mathbb{Z}^d \colon \|\mathbf{k}\|_p \leq N\}$, $N \in \mathbb{N}$, be the $\ell_p(\mathbb{Z}^d)$-ball with $0 < p \leq \infty$ and the size $N \in \mathbb{N}$, see Figure 2. The cardinality of the frequency index set $I^d_{p,N}$ is bounded by $c_{p,d}\, N^d \leq |I^d_{p,N}| \leq C_{d,p}\, N^d$, while the cardinality of the difference set satisfies $c_{p,d}\, N^d \leq |\mathcal{D}(I^d_{p,N})| \leq C_{d,p}\, 2^d\, N^d$ with the some constants $0 < c_{p,d} \leq C_{d,p}$. Consequently, we can find a reconstructing rank-1 lattice of size $M \leq \tilde{C}_{p,d}\, |I^d_{p,N}|$ using a component-by-component strategy, where the constant $\tilde{C}_{p,d} > 0$ only depends on p and d.

On the other hand, we obtain for $p \to 0$ the frequency index set $I := \{\mathbf{k} \in \mathbb{Z}^d \colon \|\mathbf{k}\|_1 = \|\mathbf{k}\|_\infty \leq N\}$ with $N \in \mathbb{N}$, which is supported on the coordinate axes. In this case we have $|I| = 2d\, N + 1$, while we obtain $(2N+1)^2 \leq |\mathcal{D}(I)| \leq d\,(2N+1)^2$. Hence, there exists a positive constant $\tilde{c}_d \in \mathbb{R}$ with $\tilde{c}_d\, |I|^2 \leq |\mathcal{D}(I)|$ and the theoretical upper bound on M is quadratic in $|I|$ for each fixed dimension d. In fact, reconstructing rank-1 lattices for these specific frequency index sets need at least $\mathcal{O}(N^2)$ nodes, see [29, Theorem 3.5] and [28]. $\square$

Example 67

Important frequency index sets in higher dimensions $d > 2$ are so-called (energy-norm based) hyperbolic crosses, see e.g. [3, 6, 7, 71]. In particular, we can consider a frequency index set of the form

$$I_N^{d,T} := \left\{\mathbf{k} \in \mathbb{Z}^d \colon (\max\{1, \|\mathbf{k}\|_1\})^{T/(T-1)} \prod_{s=1}^{d} (\max\{1, |k_s|\})^{1/(1-T)} \leq N\right\}$$

with parameters $T \in [0, 1)$ and $N \in \mathbb{N}$, see Figure 4 for illustration. The frequency index set $I_N^{d,0}$ for $T = 0$ is a *symmetric hyperbolic cross*, and the frequency index set $I_N^{d,T}$, $T \in (0, 1)$, is called *energy-norm based hyperbolic cross*.

Example 67 (continue)

The cardinality of $I_N^{d,T}$ can be estimated by

$$\begin{aligned} c_{d,0}\, N\, (\log N)^{d-1} &\leq |I_N^{d,T}| \leq C_{d,0}\, N\, (\log N)^{d-1}\,, \quad \text{for } T = 0\,, \\ c_{d,T}\, N &\leq |I_N^{d,T}| \leq C_{d,T}\, N\,, \quad \text{for } T \in (0,1) \end{aligned}$$

with some constants $0 < c_{d,T} \leq C_{d,T}$, depending only on d and T, see [30, Lemma 2.6]. Since the axis cross is a subset of the considered frequency index sets, i.e., $\{\mathbf{k} \in \mathbb{Z}^d \colon \|\mathbf{k}\|_1 = \|\mathbf{k}\|_\infty \leq N\} \subset I_N^{d,T}$ for $T \in [0,1)$, it follows that $(2N+1)^2 \leq |\mathcal{D}(I_N^{d,T})|$. On the other hand, we obtain upper bounds of the cardinality of the difference set $\mathcal{D}(I_N^{d,T})$ of the form

$$\begin{aligned} |\mathcal{D}(I_N^{d,T})| &\leq \tilde{C}_{d,0}\, N^2\, (\log N)^{d-2}\,, \quad \text{for } T = 0\,, \\ |\mathcal{D}(I_N^{d,T})| &\leq |I_N^{d,T}|^2 \leq C_{d,T}^2\, N^2\,, \quad \text{for } T \in (0,1)\,, \end{aligned}$$

see e.g. [23, Theorem 4.8].

Example 67 (continue)

Theorem 64 offers a constructive strategy to find reconstructing rank-1 lattices for $I_N^{d,T}$ of cardinality $M \leq |\mathcal{D}(I_N^{d,T})|$. For $T \in (0,1)$, these rank-1 lattices are of optimal order in N, see [23, Lemmata 2.1 and 2.3, and Corollary 2.4] and [24]. Reconstructing rank-1 lattices for these frequency index sets are discussed in more detail in [24]. □

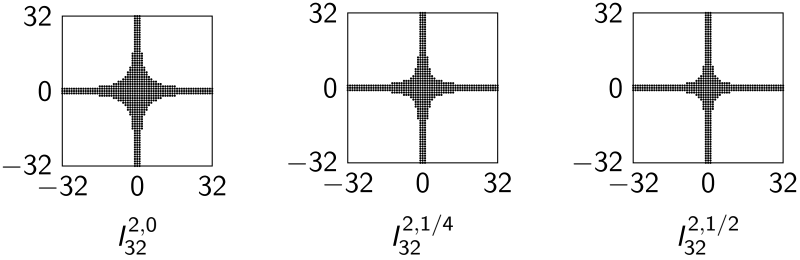


Figure 4: Two-dimensional frequency index sets $I_{32}^{2,T}$ for $T \in \{0, \frac{1}{4}, \frac{1}{2}\}$.

Summarizing, we can construct a reconstructing rank-1 lattice $\Lambda(\mathbf{z}, M, I)$ for arbitrary finite frequency index set I. The choice of the frequency index set I always depends on the approximation properties of the considered function space. The positive statement is that the size M of the reconstructing rank-1 lattice can be always bounded by $|I|^2$ being independent of the dimension d. However for important index sets, such as the hyperbolic cross or thinner index sets, the lattice size M is bounded from below by $M \geq C\, N^2$. We overcome this disadvantage in the following by considering the union of several rank-1 lattices.

Multiple rank-1 lattices

To overcome the limitations of the single rank-1 lattice approach, we consider now multiple rank-1 lattices which are obtained by taking a union of rank-1 lattices. For s rank-1 lattices $\Lambda(\mathbf{z}_r, M_r)$, $r = 1, \ldots, s$ as given in (51) we call the union

$$X = \Lambda(\mathbf{z}_1, M_1, \mathbf{z}_2, M_2, \ldots, \mathbf{z}_s, M_s) := \bigcup_{r=1}^{s} \Lambda(\mathbf{z}_r, M_r)$$

*multiple rank-*1 *lattice*. In order to work with this multiple rank-1 lattices, we need to consider the question, how many distinct points are contained in X. Assuming that for each r the lattice size M_r is coprime with at least one component of $\mathbf{z}_r$, the single rank-1 lattice $\Lambda(\mathbf{z}_r, M_r)$ possesses exactly M_r distinct points in $[0, 2\pi)^d$ including $\mathbf{0}$. Consequently, the number of distinct points in $\Lambda(\mathbf{z}_1, M_1, \mathbf{z}_2, M_2, \ldots, \mathbf{z}_s, M_s)$ is bounded from above by

$$|\Lambda(\mathbf{z}_1, M_1, \mathbf{z}_2, M_2, \ldots, \mathbf{z}_s, M_s)| \leq 1 - s + \sum_{r=1}^{s} M_r \,.$$

In the special case $s = 2$, we obtain the following result, see also [27, Lemma 2.1].

Lemma 68

Let $\Lambda(\mathbf{z}_1, M_1)$ and $\Lambda(\mathbf{z}_2, M_2)$ be two rank-1 lattices with coprime lattice sizes M_1 and M_2.
Then the multiple rank-1 lattice $\Lambda(\mathbf{z}_1, M_1, \mathbf{z}_2, M_2)$ is a subset of the rank-1 lattice $\Lambda(M_2\mathbf{z}_1 + M_1\mathbf{z}_2, M_1 M_2)$. Furthermore, if the cardinalities of the single rank-1 lattices $\Lambda(\mathbf{z}_1, M_1)$ and $\Lambda(\mathbf{z}_2, M_2)$ are M_1 and M_2, then

$$|\Lambda(\mathbf{z}_1, M_1, \mathbf{z}_2, M_2)| = M_1 + M_2 - 1\,.$$

Proof: 1. We show that $\Lambda(\mathbf{z}_1, M_1)$ is a subset of $\Lambda(M_2\mathbf{z}_1 + M_1\mathbf{z}_2, M_1 M_2)$. Let

$$\mathbf{x}_j := \frac{2\pi}{M_1}(j\,\mathbf{z}_1 \bmod M_1\mathbf{1})$$

be an arbitrary point of $\Lambda(\mathbf{z}_1, M_1)$. Since M_1 and M_2 are coprime, there exists a $k \in \{0, \ldots, M_1 - 1\}$ such that $k\,M_2 \equiv j \bmod M_1$. Choose now $\ell = kM_2$, then

$$\mathbf{y}_\ell := \frac{2\pi}{M_1 M_2}\big(\ell(M_2\mathbf{z}_1 + M_1\mathbf{z}_2) \bmod M_1 M_2\mathbf{1}\big)$$

is a point of $\Lambda(M_2\mathbf{z}_1 + M_1\mathbf{z}_2, M_1 M_2)$. Further we find by

$$\ell\,(M_2\,\mathbf{z}_1 + M_1\,\mathbf{z}_2) \bmod M_1 M_2\mathbf{1} = k\,(M_2^2\,\mathbf{z}_1 + M_1 M_2\,\mathbf{z}_2) \bmod M_1 M_2\mathbf{1}$$
$$= k\,M_2^2\,\mathbf{z}_1 \bmod M_1 M_2\mathbf{1} = k\,M_2\,\mathbf{z}_1 \bmod M_1\mathbf{1} = j\,\mathbf{z}_1 \bmod M_1\mathbf{1}$$

that $\mathbf{x}_j = \mathbf{y}_\ell$. Analogously, we conclude that $\Lambda(\mathbf{z}_2, M_2) \subset \Lambda(M_2\mathbf{z}_1 + M_1\mathbf{z}_2, M_1 M_2)$.

2. Now we prove that $\Lambda(\mathbf{z}_1, M_1) \cap \Lambda(\mathbf{z}_2, M_2) = \{\mathbf{0}\}$. For this purpose it is sufficient to show that the $M_1 M_2$ points of $\Lambda(M_2\mathbf{z}_1 + M_1\mathbf{z}_2, M_1 M_2)$ are distinct. Suppose that there is an $\ell \in \{0, \ldots, M_1 M_2 - 1\}$ such that

$$\ell\,(M_2\mathbf{z}_1 + M_1\mathbf{z}_2) \equiv \mathbf{0} \bmod M_1 M_2 \mathbf{1}\,.$$

Then there exist j_1, $k_1 \in \{0, \ldots, M_1 - 1\}$ and j_2, $k_2 \in \{0, \ldots, M_2 - 1\}$ with $\ell = j_2\,M_1 + j_1 = k_1\,M_2 + k_2$, and we find

$$\ell\,(M_2\mathbf{z}_1 + M_1\mathbf{z}_2) \bmod M_1 M_2 \mathbf{1} = j_1\,M_2\mathbf{z}_1 + k_2\,M_1\mathbf{z}_2 \bmod M_1 M_2 \mathbf{1}\,.$$

Thus, we arrive at

$$j_1\,M_2\,\mathbf{z}_1 \equiv -k_2\,M_1\,\mathbf{z}_2 \bmod M_1 M_2 \mathbf{1}\,.$$

Since M_1 and M_2 are coprime, it follows that M_1 is a divisor of each component of $j_1\,\mathbf{z}_1$, and that M_2 is a divisor of each component of $-k_2\,\mathbf{z}_2$. But this can be only true for $j_1 = k_2 = 0$, since we had assumed that $\Lambda(\mathbf{z}_1, M_1)$ and $\Lambda(\mathbf{z}_2, M_2)$ have the cardinalities M_1 and M_2. This observation implies now $\ell = j_2\,M_1 = k_1\,M_2$ which is only possible for $j_2 = k_1 = 0$, since M_1 and M_2 are coprime. Thus $\ell = 0$, and the assertion is proven. ■

Lemma 68 can be simply generalized.

Corollary 69

Let the multiple rank-1 lattice $\Lambda(\mathbf{z}_1, M_1, \ldots, \mathbf{z}_s, M_s)$ *with pairwise coprime lattice sizes* $M_1, \ldots, M_s$ *be given. Assume that* $|\Lambda(\mathbf{z}_r, M_r)| = M_r$ *for each* $r = 1, \ldots, s$.
Then we have

$$|\Lambda(\mathbf{z}_1, M_1, \ldots, \mathbf{z}_s, M_s)| = 1 - s + \sum_{r=1}^{s} M_r \,.$$

Further, let $\Lambda(\mathbf{z}, M)$ *be the rank-1 lattice with the generating vector* $\mathbf{z}$ *and lattice size* M *given by*

$$\mathbf{z} := \sum_{r=1}^{s} \Big(\prod_{\substack{\ell=1\\ \ell\neq r}}^{s} M_\ell \Big) \, \mathbf{z}_r \,, \quad M := \prod_{r=1}^{s} M_r \,.$$

Then

$$\Lambda(\mathbf{z}_1, M_1, \ldots, \mathbf{z}_s, M_s) \subset \Lambda(\mathbf{z}, M).$$

Proof: The proof follows similarly as for Lemma 68. ■

As in Section 5 we define now the *Fourier matrix* for the sampling set $X = \Lambda(\mathbf{z}_1, M_1, \mathbf{z}_2, M_2, \ldots, \mathbf{z}_s, M_s)$ and the frequency index set I,

$$\mathbf{A} = \mathbf{A}(\Lambda(\mathbf{z}_1, M_1, \mathbf{z}_2, M_2, \ldots, \mathbf{z}_s, M_s), I)$$
$$:= \begin{pmatrix} \left(\mathrm{e}^{2\pi\mathrm{i}\, j\,(\mathbf{k}\cdot\mathbf{z}_1)/M_1}\right)_{j=0,\ldots,M_1-1,\,\mathbf{k}\in I} \\ \left(\mathrm{e}^{2\pi\mathrm{i}\, j\,(\mathbf{k}\cdot\mathbf{z}_2)/M_2}\right)_{j=0,\ldots,M_2-1,\,\mathbf{k}\in I} \\ \vdots \\ \left(\mathrm{e}^{2\pi\mathrm{i}\, j\,(\mathbf{k}\cdot\mathbf{z}_s)/M_s}\right)_{j=0,\ldots,M_s-1,\,\mathbf{k}\in I} \end{pmatrix}, \qquad (67)$$

where we assume that the frequency indices $\mathbf{k} \in I$ are arranged in a fixed order. Thus $\mathbf{A}$ has $\sum_{r=1}^{s} M_r$ rows and $|I|$ columns, where the first rows of the s partial Fourier matrices coincide.

We also introduce the reduced Fourier matrix

$$\widetilde{\mathbf{A}} := \begin{pmatrix} \left(\mathrm{e}^{2\pi\mathrm{i}\, j\,(\mathbf{k}\cdot\mathbf{z}_1)/M_1}\right)_{j=0,\ldots,M_1-1,\,\mathbf{k}\in I} \\ \left(\mathrm{e}^{2\pi\mathrm{i}\, j\,(\mathbf{k}\cdot\mathbf{z}_2)/M_2}\right)_{j=1,\ldots,M_2-1,\,\mathbf{k}\in I} \\ \vdots \\ \left(\mathrm{e}^{2\pi\mathrm{i}\, j\,(\mathbf{k}\cdot\mathbf{z}_s)/M_s}\right)_{j=1,\ldots,M_s-1,\,\mathbf{k}\in I} \end{pmatrix},$$

where we use beside $\left(\mathrm{e}^{2\pi\mathrm{i}\, j\,(\mathbf{k}\cdot\mathbf{z}_1)/M_1}\right)_{j=0,\ldots,M_1-1,\,\mathbf{k}\in I}$ only the partial matrices

$$\left(\mathrm{e}^{2\pi\mathrm{i}\, j\,(\mathbf{k}\cdot\mathbf{z}_r)/M_r}\right)_{j=1,\ldots,M_r-1,\,\mathbf{k}\in I}, \quad r = 2, \ldots, s,$$

such that $\widetilde{\mathbf{A}}$ has $\sum_{r=1}^{s} M_r - s + 1$ rows and $|I|$ columns. Obviously, $\mathbf{A}$ and $\widetilde{\mathbf{A}}$ have the same rank, since we have only removed redundant rows.

As in Section 5, we consider the fast evaluation of trigonometric polynomials on multiple rank-1 lattices on the one hand and the evaluation of their Fourier coefficients from samples on multiple rank-1 lattices on the other hand.

(i) *Evaluation of trigonometric polynomials*. To evaluate a trigonometric polynomial at all nodes of a multiple rank-1 lattice $\Lambda(\mathbf{z}_1, M_1, \ldots, \mathbf{z}_s, M_s)$, we can apply the ideas from Section 5 and compute the trigonometric polynomial on s different rank-1 lattices $\Lambda(\mathbf{z}_1, M_1)$, ..., $\Lambda(\mathbf{z}_s, M_s)$ separately. The corresponding Algorithm 201 applies the known rank-1 Algorithm s-times, once for each single rank-1 lattice. The arithmetic costs of the fast evaluation at all nodes of the whole multiple rank-1 lattice $\Lambda(\mathbf{z}_1, M_1, \ldots, \mathbf{z}_s, M_s)$ is therefore $\mathcal{O}(\sum_{r=1}^{s} M_r \log M_r + s\, d\, |I|)$.

Evaluation at multiple rank-1 lattices

Input: $M_1, \ldots, M_s \in \mathbb{N}$ lattice sizes of rank-1 lattices $\Lambda(\mathbf{z}_\ell, M_\ell)$, $\ell = 1, \ldots, s$,

$\mathbf{z}_1, \ldots, \mathbf{z}_s \in \mathbb{Z}^d$ generating vectors of $\Lambda(\mathbf{z}_\ell, M_\ell)$, $\ell = 1, \ldots, s$,

$I \subset \mathbb{Z}^d$ finite frequency index set,

$\hat{\mathbf{p}} = (\hat{p}_{\mathbf{k}})_{\mathbf{k} \in I}$ Fourier coefficients of $p \in \Pi_I$ in (48).

1. For $\ell = 1, \ldots, s$ do by rank-1 Algorithm

$$\mathbf{p}_\ell := \mathrm{LFFT}(M_\ell, \mathbf{z}_\ell, I, \hat{\mathbf{p}})$$

   end for

2. Set $\mathbf{p} := \big(\mathbf{p}_1(1), \ldots, \mathbf{p}_1(M_1), \mathbf{p}_2(2), \ldots, \mathbf{p}_2(M_2), \ldots, \mathbf{p}_s(2), \ldots \mathbf{p}_s(M_s)\big)^\top$.

Output: $\mathbf{p} = \widetilde{\mathbf{A}}\,\hat{\mathbf{p}}$ polynomial values of $p \in \Pi_I$.
Arithmetic costs: $\mathcal{O}(\sum_{\ell=1}^{s} M_\ell \log M_\ell + s\, d\, |I|)$.

The algorithm is a fast realization of the matrix-vector product with the Fourier matrix $\widetilde{\mathbf{A}}$ in (67). The fast computation of the matrix-vector product with the adjoint Fourier matrix $\mathbf{A}^{\mathrm{H}}$ can be realized by employing rank-1 Algorithm separately to each rank-1 lattice with a numerical effort of $\mathcal{O}(\sum_{\ell=1}^{s} M_\ell \log M_\ell + s\, d\, |I|)$.

(ii) *Evaluation of the Fourier coefficients*. To solve the inverse problem, i.e., to compute the Fourier coefficients of an arbitrary trigonometric polynomial $p \in \Pi_I$ as given in (48), we need to ensure that our Fourier matrix $\mathbf{A}$ in (67)) has full rank $|I|$. This means that p needs to be already completely determined by the sampling set $\Lambda(\mathbf{z}_1, M_1, \ldots, \mathbf{z}_s, M_s)$. Then we can apply formula (50) for reconstruction. We are especially interested in a fast and stable reconstruction method.

We define a *reconstructing multiple rank*-1 *lattice* to a given frequency index set I as a multiple rank-1 lattice satisfying that

$$\mathbf{A}^{\mathrm{H}}\mathbf{A}$$

with $\mathbf{A} = \mathbf{A}(\Lambda(\mathbf{z}_1, M_1, \mathbf{z}_2, M_2, \ldots, \mathbf{z}_s, M_s), I)$ has full rank $|I|$.
In order to keep the needed number of sampling points $|\Lambda(\mathbf{z}_1, M_1, \ldots, \mathbf{z}_s, M_s)| = \sum_{r=1}^{s} M_r - s + 1$ small, we do not longer assume that each single rank-1 lattice is a reconstructing rank-1 lattice. But still, we can use Lemma 58 in order to compute the matrix $\mathbf{A}^{\mathrm{H}}\mathbf{A}$ in an efficient way.

Lemma 70

Let $\mathbf{A}$ *be the* $(\sum_{r=1}^{s} M_r)$*-by-*$|I|$ *Fourier matrix* (67) *for a frequency index set* $|I|$ *and a multiple rank-*1 *lattice* $\Lambda(\mathbf{z}_1, M_1, \mathbf{z}_2, M_2, \ldots, \mathbf{z}_s, M_s)$ *with cardinality* $1 - s + \sum_{r=1}^{s} M_r$*. Then the entries of* $\mathbf{A}^{\mathrm{H}}\mathbf{A} \in \mathbb{C}^{|I| \times |I|}$ *have the form*

$$(\mathbf{A}^{\mathrm{H}}\mathbf{A})_{\mathbf{h},\mathbf{k}} = \sum_{r=1}^{s} M_r \delta_{(\mathbf{k}-\mathbf{h})\cdot\mathbf{z}_r \bmod M_r} \,,$$

where

$$\delta_{(\mathbf{k}-\mathbf{h})\cdot\mathbf{z}_r \bmod M_r} := \begin{cases} 1 & \mathbf{k} \cdot \mathbf{z}_r \equiv \mathbf{h} \cdot \mathbf{z}_r \bmod M_r \,, \\ 0 & \mathbf{k} \cdot \mathbf{z}_r \not\equiv \mathbf{h} \cdot \mathbf{z}_r \bmod M_r \,. \end{cases}$$

Proof: The assertion follows directly from Lemma 58. The entry $(\mathbf{A}^{\mathrm{H}}\mathbf{A})_{\mathbf{h},\mathbf{k}}$ is the inner product of the **k**th and the **h**th column of **A**. Thus we find

$$(\mathbf{A}^{\mathrm{H}}\mathbf{A})_{\mathbf{h},\mathbf{k}} = \sum_{r=1}^{s} \sum_{j=0}^{M_r-1} \left(\mathrm{e}^{2\pi\mathrm{i}\,[(\mathbf{k}-\mathbf{h})\cdot\mathbf{z}_r]/M_r}\right)^j,$$

where the sums

$$\sum_{j=0}^{M_r-1} \left(\mathrm{e}^{2\pi\mathrm{i}\,[(\mathbf{k}-\mathbf{h})\cdot\mathbf{z}_r]/M_r}\right)^j$$

can be simply computed as in Lemma 58. ■

Lemma 70 also shows that $\mathbf{A}^{\mathrm{H}}\mathbf{A}$ can be sparse for suitably chosen rank-1 lattices. If the single rank-1 lattices are already reconstructing rank-1 lattices, then it directly follows that $\mathbf{A}^{\mathrm{H}}\mathbf{A}$ is a multiple of the identity matrix.

Now the question remains, how to choose the parameters s as well as $\mathbf{z}_r$ and M_r, $r = 1, \ldots s$, to ensure that $\mathbf{A}^{\mathrm{H}}\mathbf{A}$ indeed possesses full rank $|I|$. The following strategy given in Algorithm 207, see [26, Algorithm 1], yields with high probability such a multiple rank-1 lattice. Here we take the lattice sizes $M_r := M$ for all $r = 1, \ldots, s$ as a prime number and choose the generating vectors $\mathbf{z}_r$ randomly in the set $[0,\, M-1]^d \cap \mathbb{Z}^d$. In order to determine the lattice size M large enough for the index set I, we define the *expansion of the frequency set I* by

$$N_I := \max_{j=1,\ldots,d} \{\max_{\mathbf{k}\in I} k_j - \min_{\mathbf{l}\in I} \ell_j\}\,, \tag{68}$$

where $\mathbf{k} = (k_j)_{j=1}^d$ and $\mathbf{l} = (\ell_j)_{j=1}^d$ belong to I. The expansion N_I can be interpreted as the size of a d-dimensional cube we need to cover the index set I.

Determining reconstructing multiple rank-1 lattices

Input: $T \in \mathbb{N}$ upper bound of the cardinality of a frequency set I,
$d \in \mathbb{N}$ dimension of the frequency set I,
$N \in \mathbb{N}$ upper bound of the expansion N_I,
$\delta \in (0,1)$ upper bound of failure probability,
$c > 1$ minimal oversampling factor.

1. Set $c := \max\left\{c, \frac{N}{T-1}\right\}$ and $\lambda := c\,(T-1)$.
2. Set $s := \lceil \left(\frac{c}{c-1}\right)^2 \frac{\ln T - \ln \delta}{2} \rceil$.
3. Set $M = \operatorname{argmin}\{p > \lambda : p \in \mathbb{N} \text{ prime}\}$.
4. For $r = 1$ to s do
   choose $\mathbf{z}_r$ from $[0, M-1]^d \cap \mathbb{Z}^d$ uniformly at random
   endfor

Output: M lattice size of all rank-1 lattices,
$\mathbf{z}_1, \ldots, \mathbf{z}_s$ generating vectors of rank-1 lattices such that
$\Lambda(\mathbf{z}_1, M, \ldots, \mathbf{z}_s, M)$ is a reconstructing multiple rank-1
with probability at least $1 - \delta$.

Arithmetical cost: $\mathcal{O}(\lambda \ln\ln\lambda + ds)$ for $c > 1$, $\lambda \sim \max\{T, N\}$,
and $s \sim \ln T - \ln \delta$.

Due to [26, Theorem 3.4] the Algorithm 207 determines a reconstructing sampling set for trigonometric polynomials supported on the given frequency set I with probability at least $1 - \delta_s$, where

$$\delta_s = T\,\mathrm{e}^{-2s\,(c-1)^2/c^2} \tag{69}$$

is an upper bound on the probability that the approach fails. There are several other strategies in the literature to find appropriate reconstructing multiple rank-1 lattices, see [27, 26, 32]. Finally, if a reconstructing multiple rank-1 lattice is found, then the Fourier coefficients of the trigonometric polynomial $p \in \Pi_I$ in (48) can be efficiently computed by solving the system

$$\mathbf{A}^{\mathrm{H}}\mathbf{A}\,\hat{\mathbf{p}} = \mathbf{A}^{\mathrm{H}}\mathbf{p}\,,$$

where $\mathbf{p} := \big(p(\mathbf{x}_j)_{\mathbf{x}_j\in\Lambda(\mathbf{z}_1,M_1)}, \ldots, p(\mathbf{x}_j)_{\mathbf{x}_j\in\Lambda(\mathbf{z}_s,M_s)}\big)^\top$.

Unknown frequency index set I

until now:

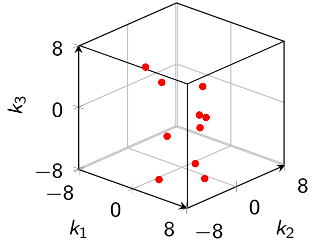
- fast reconstruction / approximation from samples for arbitrary **given** frequency index set $I \subset \mathbb{Z}^d$, $|I| < \infty$

next: unknown frequency index set $I \Rightarrow$ multi-dim. sparse FFT

- task: Determine frequency index set I from samples belonging to $\approx$largest Fourier coefficients $\hat{f}_{\boldsymbol{k}}$ or to $\hat{f}_{\boldsymbol{k}} \neq 0$
- search domain $\Gamma \subset \mathbb{Z}^d$, e.g. full grid $\hat{G}_N^d := \{-N, -N+1, \ldots, N\}^d$, $N \in \mathbb{N}$
- various existing methods, e.g., based on
  - filters [Indyk, Kapralov '14]
  - Chinese Remainder Theorem [Cuyt, Lee '08] [Iwen '13]
  - Prony's method [Tasche, P. '13] [Peter, Plonka, Schaback '15] [Kunis, Peter, Römer, von der Ohe '15]
- **problems:** non-sparsity, implementations?, stability, many frequencies

$\Rightarrow$ dimension-incremental sparse FFT based on rank-1 lattices [P., Volkmer. '15] [Volkmer '17]; (similar basic idea without rank-1 lattices: [Zippel '79] [Kaltofen, Lee '03] [Javadi Monagan '10] [P., Tasche '13])

Dimension incremental reconstruction

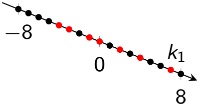


supp $\hat{p} \subset \Gamma = \hat{G}_8^3$

$$\tilde{\hat{p}}_{k_1} := \frac{1}{17}\sum_{\ell=0}^{16} p\left(\begin{pmatrix}\ell/17\\ x_2'\\ x_3'\end{pmatrix}\right) \mathrm{e}^{-2\pi\mathrm{i}\frac{\ell k_1}{17}}$$

$$= \sum_{\substack{(h_2,h_3)\in\{-8,\dots,8\}^2\\ (k_1,h_2,h_3)^\top\in\mathrm{supp}\,\hat{p}}} \hat{p}_{\begin{pmatrix}k_1\\ h_2\\ h_3\end{pmatrix}} \mathrm{e}^{2\pi\mathrm{i}(h_2x_2'+h_3x_3')},$$

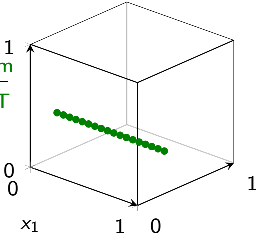
$k_1 = -8, \dots, 8$



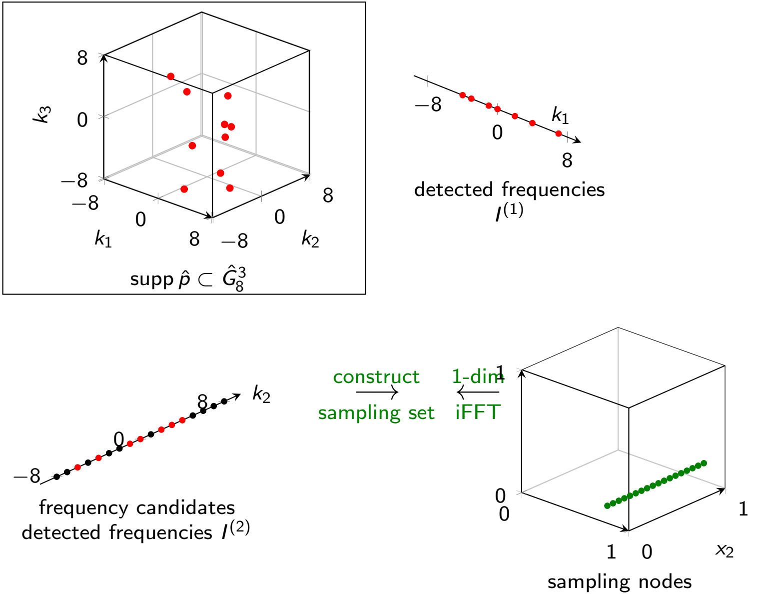
frequency candidates
detected frequencies $I^{(1)}$

construct $\longrightarrow$ sampling set

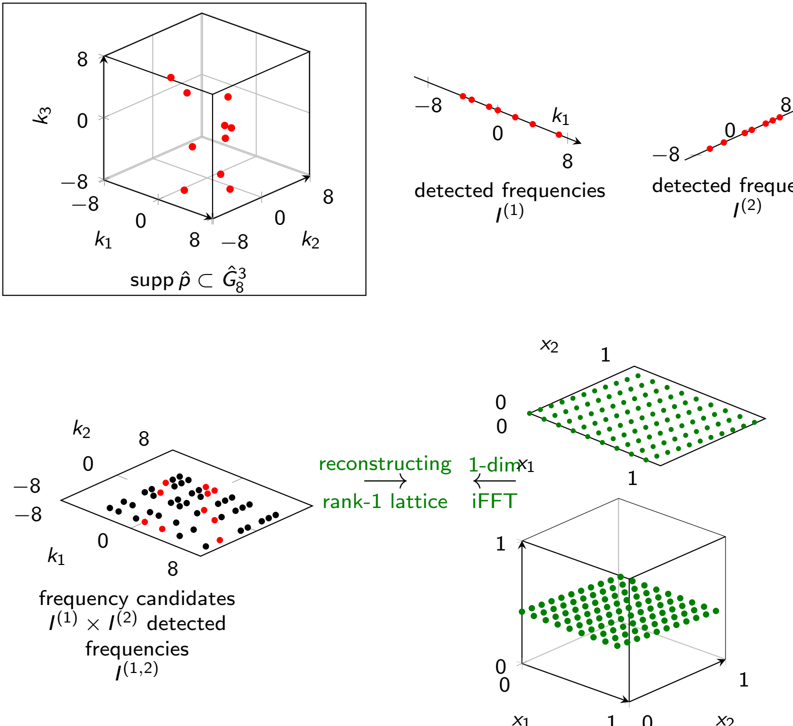
1-dim $\longleftarrow$ iFFT



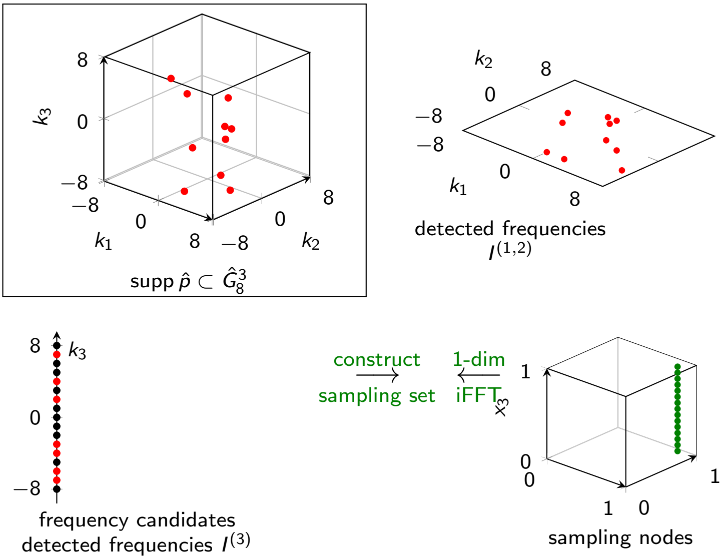
sampling nodes
$(\frac{\ell}{17}, x_2', x_3')^\top$, $\ell = 0, \dots, 16$

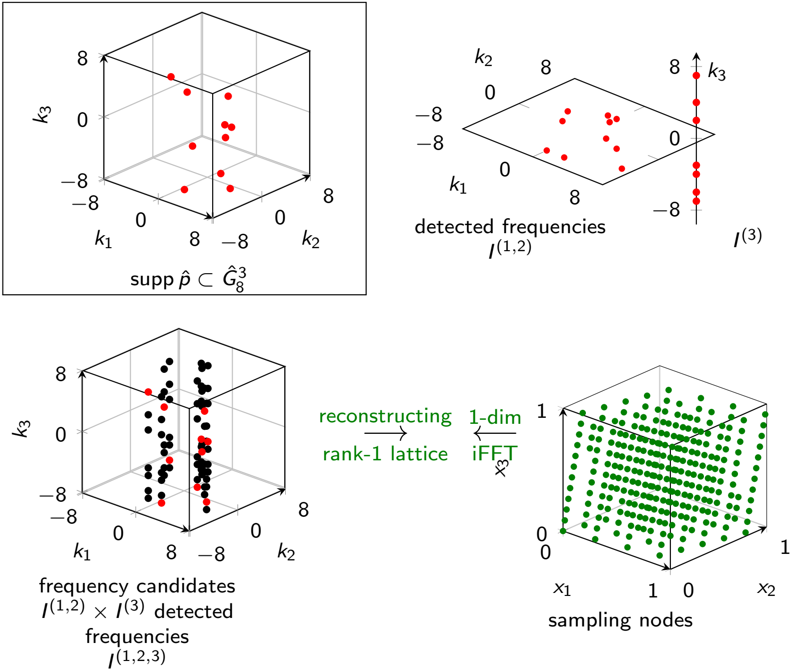


8
0
−8
k_3
−8
0
8
k_1
−8
0
8
k_2
supp $\hat{p} \subset \hat{G}_8^3$
−8
0
8
k_1
detected frequencies
$I^{(1)}$
−8
0
8
k_2
frequency candidates
detected frequencies $I^{(2)}$
construct
sampling set
1-dim
iFFT
1
0
0
1
1
0
x_2
sampling nodes



supp $\hat{p} \subset \hat{G}_8^3$
detected frequencies $I^{(1)}$
detected frequencies $I^{(2)}$
frequency candidates $I^{(1)} \times I^{(2)}$ detected frequencies $I^{(1,2)}$
reconstructing rank-1 lattice
1-dim iFFT
k_1
k_2
k_3
x_1
x_2


supp $\hat{p} \subset \hat{G}_8^3$
detected frequencies $I^{(1,2)}$
frequency candidates
detected frequencies $I^{(3)}$
construct
sampling set
1-dim
iFFT
sampling nodes
k_1
k_2
k_3
x_3



supp $\hat{p} \subset \hat{G}_8^3$
detected frequencies $I^{(1,2)}$
$I^{(3)}$
frequency candidates $I^{(1,2)} \times I^{(3)}$ detected frequencies $I^{(1,2,3)}$
reconstructing rank-1 lattice
1-dim iFFT
sampling nodes
k_1
k_2
k_3
x_1
x_2
x_3

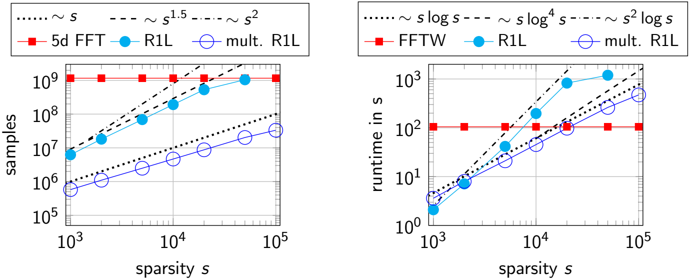
- B-spline $N_m(x) := \sum_{k\in\mathbb{Z}} C_m \,\mathrm{sinc}\left(\frac{\pi}{m}k\right)^m \cos(\pi k)\,\mathrm{e}^{2\pi\mathrm{i}kx}$, $\|N_m|L^2(\mathbb{T})\| = 1$, $|\hat{N}_m(k)| \sim |k|^{-m}$
- $f(\boldsymbol{x}) := \prod_{t\in\{1,3,8\}} N_2(x_t) + \prod_{t\in\{2,5,6,10\}} N_4(x_t) + \prod_{t\in\{4,7,9\}} N_6(x_t)$
- full grid for $N = 64$, $d = 10$: $|\hat{G}_{64}^{10}| = 129^{10} \approx 1.28 \cdot 10^{21}$
- symmetric hyperbolic cross: $|I_{64}^{10}| = 696\,036\,321$
  relative $L^2(\mathbb{T}^d)$-error (best case) 4.1e-04
- results for dimension incremental algorithm with $\Gamma = \hat{G}_{64}^{10}$ (tests repeated 10 times):

threshold	#samples	$\lvert I \rvert$	rel. L_2-error
1.0e-02	254 530	491	1.4e-01
1.0e-03	2 789 050	1 121	1.1e-02
1.0e-04	17 836 042	3 013	1.7e-03
1.0e-05	82 222 438	7 163	4.7e-04

complexity of dimension-incremental sparse FFT using multiple rank-1 lattices:

- sparsity $s = |I|$, search domain $\Gamma = \hat{G}_N^d := \{-N, \ldots, N\}^d \supset I$, number of detection iterations r
- samples: $\mathcal{O}(d\, r s N \log^2(rsN))$ (w.h.p.)
  instead of $\mathcal{O}(d\, r^2 s^2 N)$
- arithmetic operations: $\mathcal{O}\big(d^2 r s N \log^4(rsN)\big)$ (w.h.p.)
  instead of $\mathcal{O}(d\, r^3 s^3 + d\, r^2 s^2 N \log(rsN))$

Example: $p_I(\boldsymbol{x}) = \sum_{\boldsymbol{k}\in I} \hat{p}_{\boldsymbol{k}}\, \mathrm{e}^{2\pi\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}}$, $I \subset \Gamma = \hat{G}_{32}^5 := \{-32,\dots,32\}^5$, $|\Gamma| \approx 1.16\cdot 10^9$



Example:

- B-spline $N_m(x) := \sum_{k\in\mathbb{Z}} C_m \operatorname{sinc}\left(\frac{\pi}{m}k\right)^m (-1)^k \mathrm{e}^{2\pi\mathrm{i}kx}$
- $f(\boldsymbol{x}) :=$
  $\prod_{t\in\{1,3,8\}} N_2(x_t) + \prod_{t\in\{2,5,6,10\}} N_4(x_t) + \prod_{t\in\{4,7,9\}} N_6(x_t)$
- dimension-incremental method for $\Gamma = \hat{G}_{64}^{10}$:
  $(|\hat{G}_{64}^{10}| \approx 1.28 \cdot 10^{21})$

threshold	single rank-1 lattices #samples	$\lvert I\rvert$	rel. L_2 error	multiple rank-1 lattices #samples	$\lvert I\rvert$	rel. L_2 error
1.0e-02	327 689	493	1.3e-01	246 681	501	5.5e-01
1.0e-03	2 551 143	1 109	1.1e-02	1 441 455	1 205	1.1e-02
1.0e-04	17 198 228	3 009	2.0e-03	7 473 447	3 463	2.1e-03
1.0e-05	132 285 922	7 435	4.8e-04	37 056 491	11 053	4.9e-04

- multivariate periodic functions and rank-1 lattices
  - fast reconstruction of multivariate trigonometric polynomials p_I for arbitrary frequency index sets I [Kämmerer '14]
  - fast approximation [Kämmerer '14], error estimates in [Kämmerer '14], [Volkmer '17]
- similar results for multivariate non-periodic functions and rank-1 Chebyshev lattices (not in this talk)
- high-dimensional dimension-incremental sparse FFT and rank-1 lattices [P., Volkmer '16] [Volkmer '17]
  - determination of unknown frequency index set I
  - very good numerical results for high-dimensional sparse trigonometric polynomials and for high-dimensional functions (non-sparse in frequency domain)
- high-dimensional dim.-incremental sparse FFT and multiple rank-1 lattices
  - based on multiple reconstructing rank-1 lattices [Kämmerer '16] [Kämmerer '17]
  - distinct reduction of number of samples and arithmetic operations

Prony's method for reconstruction of structured functions

The recovery of a structured function from noisy sampled data is a fundamental problem in applied mathematics and signal processing. In Section 222, we consider the frequency analysis problem, where the classical Prony method and its relatives are described. Section 240 describes frequently used methods for solving the frequency analysis problem, namely MUSIC, the approximate Prony method, and ESPRIT. The algorithms for recovery of exponential sums will be mainly derived for noiseless data. Fortunately, these methods work also for noisy data. This important property is based on the stability of exponentials which will be handled in Section 269.

The reconstruction of a compactly supported function of special structure from given Fourier data is a common problem in scientific computing. In Section **??** we present an algorithm for recovery of a spline function from given samples of its Fourier transform. In Section **??** we study a phase retrieval problem, i.e., we investigate the question whether a complex-valued function f can be reconstructed from the modulus $|\hat{f}|$ of its Fourier transform.

Prony method

The following problem arises quite often in electrical engineering, signal processing, and mathematical physics and is known as *frequency analysis problem* (see [47] or [41, Chapter 9]):

Recover the positive integer M, distinct frequencies $\varphi_j \in [-\pi, \pi)$, complex coefficients $c_j \neq 0$, $j = 1, \ldots, M$, in the *exponential sum of order M*

$$h(x) := \sum_{j=1}^{M} c_j \, \mathrm{e}^{\mathrm{i}\varphi_j x}\,, \quad x \geq 0\,, \tag{70}$$

if noisy sampled data $h_k := h(k) + e_k$, $k = 0, \ldots, N-1$, with $N \geq 2\,M$ are given, where $e_k \in \mathbb{C}$ are small error terms.

In this frequency analysis problem, we have to detect the significant exponentials of the signal h.

The classical Prony method works for noiseless sampled data of the exponential sum (70) in the case of known order M. Following an idea of G.R. de Prony from 1795 (see [53]), we recover all parameters of the exponential sum (70), if sampled data

$$h(k) := \sum_{j=1}^{M} c_j \, \mathrm{e}^{\mathrm{i}\,\varphi_j\, k} = \sum_{j=1}^{M} c_j \, z_j^k \,, \quad k = 0, \ldots, 2M-1 \qquad (71)$$

are given, where $z_j := \mathrm{e}^{\mathrm{i}\,\varphi_j}$ are distinct points on the unit circle. We introduce the *Prony polynomial*

$$p(z) := \prod_{j=1}^{M} (z - z_j) = \sum_{k=0}^{M-1} p_k \, z^k + z^M \,, \quad z \in \mathbb{C} \,, \qquad (72)$$

with corresponding coefficients $p_k \in \mathbb{C}$.

Further we define the *companion matrix* $\mathbf{C}_M(p) \in \mathbb{C}^{M\times M}$ of the Prony polynomial (72) by

$$\mathbf{C}_M(p) := \begin{pmatrix} 0 & 0 & \dots & 0 & -p_0 \\ 1 & 0 & \dots & 0 & -p_1 \\ 0 & 1 & \dots & 0 & -p_2 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \dots & 1 & -p_{M-1} \end{pmatrix} . \tag{73}$$

It is known that the companion matrix $\mathbf{C}_M(p)$ has the property

$$\det\big(z\,\mathbf{I}_M - \mathbf{C}_M(p)\big) = p(z)\,,$$

where $\mathbf{I}_M \in \mathbb{C}^{M\times M}$ denotes the identity matrix. Hence the zeros of the Prony polynomial (72) coincide with the eigenvalues of the companion matrix $\mathbf{C}_M(p)$.

Setting $p_M := 1$, we observe the following relation for all $m \in \mathbb{N}_0$,

$$\begin{aligned}
& \sum_{k=0}^{M} p_k\, h(k+m) \\
= & \sum_{k=0}^{M} p_k \Big(\sum_{j=1}^{M} c_j\, z_j^{k+m} \Big) \\
= & \sum_{j=1}^{M} c_j\, z_j^{m} \Big(\sum_{k=0}^{M} p_k\, z_j^{k} \Big) = \sum_{j=1}^{M} c_j\, z_j^{m}\, p(z_j) = 0\,. \qquad (74)
\end{aligned}$$

Using the known values $h(k)$, $k = 0, \ldots, 2M-1$, the formula (74) implies that the homogeneous linear difference equation

$$\sum_{k=0}^{M-1} p_k\, h(k+m) = -h(M+m)\,, \quad m = 0, \ldots, M-1 \qquad (75)$$

is fulfilled.

In matrix-vector notation, we obtain the linear system

$$\mathbf{H}_M(0)\,\big(p_k\big)_{k=0}^{M-1} = -\big(h(M+m)\big)_{m=0}^{M-1} \tag{76}$$

with the square *Hankel matrix*

$$\mathbf{H}_M(0) := \begin{pmatrix} h(0) & h(1) & \dots & h(M-1) \\ h(1) & h(2) & \dots & h(M) \\ \vdots & \vdots & & \vdots \\ h(M-1) & h(M) & \dots & h(2M-2) \end{pmatrix} = \big(h(k+m)\big)_{k,m=0}^{M-1}. \tag{77}$$

The matrix $\mathbf{H}_M(0)$ is invertible, since by the special structure (71) of the values $h(k)$ we have the factorization

$$\mathbf{H}_M(0) = \mathbf{V}_M(\mathbf{z})\,\big(\mathrm{diag}\,\mathbf{c}\big)\,\mathbf{V}_M(\mathbf{z})^\top,$$

where the diagonal matrix $\mathrm{diag}\,\mathbf{c}$ with $\mathbf{c} := (c_j)_{j=1}^{M}$ contains the nonzero coefficients of (70) in the main diagonal, and where

$$\mathbf{V}_M(\mathbf{z}) := \left(z_k^{j-1}\right)_{j,k=1}^{M} = \begin{pmatrix} 1 & 1 & \dots & 1 \\ z_1 & z_2 & \dots & z_M \\ \vdots & \vdots & & \vdots \\ z_1^{M-1} & z_2^{M-1} & \dots & z_M^{M-1} \end{pmatrix}$$

denotes the *Vandermonde matrix* generated by the vector $\mathbf{z} := (z_j)_{j=1}^{M}$. Since all z_j, $j = 1, \dots, M$, are distinct, the Vandermonde matrix $\mathbf{V}_M(\mathbf{z})$ is invertible. Note that by (71) we have

$$\mathbf{V}_M(\mathbf{z})\,\mathbf{c} = \big(h(k)\big)_{k=0}^{M-1}\,. \tag{78}$$

We summarize:

Classical Prony method

Input: $M \in \mathbb{N}$, sampled values $h(k)$, $k = 0, \ldots, 2M-1$.

1. Solve the linear system (76).
2. Compute all zeros $\tilde{z}_j \in \mathbb{C}$, $j = 1, \ldots, M$, of the Prony polynomial (72), i.e., calculate all eigenvalues $\tilde{z}_j$ of the associated companion matrix (73). For $z_j := \tilde{z}_j / |\tilde{z}_j|$ and form $\varphi_j := \mathrm{Im}(\log z_j) \in [-\pi, \pi)$, $j = 1, \ldots, M$, where log is the principal value of the complex logarithm.
3. Solve the Vandermonde system (78).

Output: $\varphi_j \in [-\pi, \pi)$, $c_j \in \mathbb{C}$, $j = 1, \ldots, M$.

As shown, Prony's idea is mainly based on the separation of the unknown frequencies φ_j from the unknown coefficients c_j. But the main problem is the determination of φ_j, since the coefficients c_j are uniquely determined by the linear system (78).

Remark 71

For simplicity, we consider only an undamped exponential sum (70)*. Analogously, one can handle a damped exponential sum*

$$h(x) := \sum_{j=1}^{M} c_j \, \mathrm{e}^{f_j x} \,, \quad x \geq 0 \,,$$

where $c_j \neq 0$ and $f_j \in [-\alpha, 0] + \mathrm{i}\,[-\pi, \pi)$ are distinct numbers with small $\alpha > 0$ (see [49]). Then the negative real part of f_j is the damping factor and the imaginary part of f_j is the angular frequency of the exponential $\mathrm{e}^{f_j x}$.

Remark 72

The Prony method can be also applied to the recovery of an extended exponential sum

$$h(x) := \sum_{j=1}^{m} c_j(x)\, \mathrm{e}^{\mathrm{i}\varphi_j x}\,, \quad x \geq 0\,,$$

where c_j are polynomials of low degree. For simplicity, we sketch only the case of linear polynomials $c_j(x) = c_{j,0} + c_{j,1}x$. With distinct $z_j = \mathrm{e}^{\mathrm{i}\varphi_j}$, $j = 1, \ldots, M$, the corresponding Prony polynomial reads as follows

$$p(z) := \prod_{j=1}^{M} (z - z_j)^2 = \sum_{k=0}^{2M-1} p_k\, z^k + z^{2M}\,. \tag{79}$$

Remark 72 (continue)

Assuming that the sampled values $h(k)$, $k = 0, \ldots, 4M-1$, *are given, one has to solve the linear system*

$$\sum_{k=0}^{2M-1} p_k \, h(k+\ell) = -h(2M+\ell)\,, \quad \ell = 0, \ldots, 2M-1\,,$$

and to compute all double zeros z_j *of corresponding Prony polynomial* (79), *i.e., all double eigenvalues of the related companion matrix.*

Remark 72 (continue)

Introducing the confluent Vandermonde matrix

$$\mathbf{V}^c_{2M}(\mathbf{z}) := \begin{pmatrix} 1 & 0 & \dots & 1 & 0 \\ z_1 & 1 & \dots & z_M & 1 \\ z_1^2 & 2z_1 & \dots & z_M^2 & 2z_M \\ \vdots & \vdots & & \vdots & \vdots \\ z_1^{2M-1} & (2M-1)z_1^{2M-2} & \dots & z_M^{2M-1} & (2M-1)z_M^{2M-2} \end{pmatrix}$$

finally one has to solve the confluent Vandermonde system

$$\mathbf{V}^c_{2M}(\mathbf{z})\left(c_{0,1},\, z_1 c_{1,1},\, \dots,\, c_{M,0},\, z_1 c_{M,1}\right)^\top = \left(h(k)\right)_{k=0}^{2M-1}. \qquad \square$$

Remark 73

The Prony method is closely related to Padé approximation (see [67]). Let $(f_k)_{k\in\mathbb{N}_0}$ *be a complex sequence with* $\rho := \limsup_{k\to\infty} |f_k|^{1/k} < \infty$. *The z-transform of such a sequence is the Laurent series* $\sum_{k=0}^{\infty} f_k\, z^{-k}$ *which converges in the neighborhood* $\{z\in\mathbb{C}: |z|>\rho\}$ *of* $z=\infty$. *Thus the z-transform of each sequence* $(z_j^k)_{k\in\mathbb{N}_0}$ *with* $z_j = \mathrm{e}^{\mathrm{i}\,\varphi_j}$ *is equal to* $\frac{z}{z-z_j}$, $j=1,\ldots,M$. *Since the z-transform is linear, the z-transform maps the data sequence* $\big(h(k)\big)_{k\in\mathbb{N}_0}$ *with* (71) *for all* $k\in\mathbb{N}_0$ *into the rational function*

$$\sum_{k=0}^{\infty} h(k)\, z^{-k} = \sum_{j=1}^{M} c_j\, \frac{z}{z-z_j} = \frac{a(z)}{p(z)}\,, \qquad (80)$$

where p is the Prony polynomial (72) *and* $a(z) := a_M\, z^M + \ldots + a_1\, z$. *Now we substitute z for* z^{-1} *in* (80) *and form the reverse Prony polynomial* $\operatorname{rev} p(z) := z^M\, p(z^{-1})$ *of degree M with* $\operatorname{rev} p(0) = 1$ *as well as the reverse polynomial* $\operatorname{rev} a(z) := z^M\, a(z^{-1})$ *of degree at least* $M-1$.

Remark 73 (continue)

Then we obtain that

$$\sum_{k=0}^{\infty} h(k)\, z^k = \frac{\text{rev } a(z)}{\text{rev } p(z)} \tag{81}$$

in a certain neighborhood of $z = 0$*. In other words, the rational function on the right side of* (81) *is an* $(M-1, M)$ *Padé approximant of the power series* $\sum_{k=0}^{\infty} h(k)\, z^k$ *with vanishing* $\mathcal{O}(z^M)$ *term and it holds*

$$\left(\sum_{k=0}^{\infty} h(k)\, z^k\right) \text{rev } p(z) = \text{rev } a(z)$$

in a neighborhood of $z = 0$*.*

Remark 73 (continue)

Equating the coefficients of like powers of z yields

$$\begin{aligned} \sum_{k=M-m}^{M} p_k\, h(k+m-M) &= a_{M-m}\,, \quad m = 0, \ldots, M-1\,, \\ \sum_{k=0}^{M} p_k\, h(k+m) &= 0\,, \quad m \in \mathbb{N}_0\,. \end{aligned} \tag{82}$$

Now the equations (82) *for* $m = 0, \ldots, M-1$ *coincide with* (75)*. Hence the Prony method may also be regarded as a Padé approximation.* □

Remark 74

In signal processing, the Prony method is also known as the annihilating filter method, see e.g. [65]. For distinct $z_j = \mathrm{e}^{\mathrm{i}\varphi_j}$ *and complex coefficients* $c_j \neq 0$, $j = 1, \ldots, M$, *we consider the discrete signal* $\mathbf{h} = \big(h_n\big)_{n\in\mathbb{Z}}$ *with*

$$h_n := \sum_{j=1}^{M} c_j\, z_j^n, \quad n \in \mathbb{Z}\,. \tag{83}$$

For simplicity, we assume that M *is known. Then a discrete signal* $\mathbf{a} = (a_n)_{n\in\mathbb{Z}}$ *is called an annihilating filter of the signal* $\mathbf{h}$, *if the discrete convolution of the signals* $\mathbf{a}$ *and* $\mathbf{h}$ *vanishes, i.e.*

$$(\mathbf{a} * \mathbf{h})_n := \sum_{\ell\in\mathbb{Z}} a_\ell\, h_{n-\ell} = 0\,, \quad n \in \mathbb{Z}\,.$$

Remark 74 (continue)

For the construction of an annihilating filter **a** *we consider*

$$a(z) := \prod_{j=1}^{M} \left(1 - z_j \, z^{-1}\right) = \sum_{n=0}^{M} a_n \, z^{-n} , \quad z \in \mathbb{C} \setminus \{0\} ,$$

then $\mathbf{a} = (a_n)_{n \in \mathbb{Z}}$ *with* $a_n = 0$, $n \in \mathbb{Z} \setminus \{0, \ldots, M\}$ *is an annihilating filter of* **h** *in* (83). *Note that* $a(z)$ *is the z-transform of the annihilating filter* **a**. *Furthermore,* $a(z)$ *and the Prony polynomial* (72) *have the same zeros* $z_j \in \mathbb{D}$, $j = 1, \ldots, M$, *since* $z^M \, a(z) = p(z)$ *for all* $z \in \mathbb{C} \setminus \{0\}$. *Hence the Prony method and the method of annihilating filters are equivalent. For details see e.g. [65].* $\square$

Remark 75

Prony methods arise also from problems of science and engineering, where one is interested in predicting future information from previous ones using a linear model. Let $\mathbf{h} = (h_n)_{n\in\mathbb{N}_0}$ *be a discrete signal. The linear prediction method, see e.g. [5], aims at finding suitable predictor parameters* $p_j \in \mathbb{C}$ *such that the signal value* $h_{\ell+M}$ *can be expressed as a linear combination of the previous signal values* h_j, $j = \ell, \ldots, \ell + M - 1$, *i.e.*

$$h_{\ell+M} = \sum_{j=0}^{M-1} (-p_j)\, h_{\ell+j}\,, \quad \ell \in \mathbb{N}_0\,.$$

Therefore these equations are also called linear prediction equations. Setting $p_M := 1$, *we observe that this representation is equivalent to the homogeneous linear difference equation* (75).

Remark 75 (continue)

Assuming that

$$h_k = \sum_{j=1}^{M} c_j \, z_j^k, \quad k \in \mathbb{N}_0 \,,$$

we obtain the frequency analysis problem, i.e., the Prony polynomial (72) *coincides with the negative value of the forward predictor polynomial. The associated companion matrix* $\mathbf{C}_M(p)$ *in* (73) *is hence equal to the forward predictor matrix. Thus the linear prediction method can also be considered as a Prony method.* □

Unfortunately, the classical Prony method has some numerical drawbacks. Often the order M of the exponential sum (70) is unknown. Further the classical Prony method is known to perform poorly when noisy sampled data are given, since the Hankel matrix $\mathbf{H}_M(0)$ as well as the Vandermonde matrix $\mathbf{V}_M(\mathbf{z})$ are usually badly conditioned. We will see that one can attenuate these problems by using more sampled data. But then one has to deal with rectangular matrices.

Recovery of exponential sums

In this section, we present three efficient algorithms for solving the frequency analysis problem. Let $N \in \mathbb{N}$ with $N \geq 2M$ be given, where $M \in \mathbb{N}$ denotes the (unknown) order of the exponential sum (70). We introduce the *nonequispaced Fourier matrix*, see Chapter **??**,

$$\mathbf{A}_{N,M}^{\top} := \left(\mathrm{e}^{\mathrm{i}\varphi_j(k-1)}\right)_{k,j=1}^{N,\,M}.$$

Note that $\mathbf{A}_{N,M}^{\top}$ coincides with the *rectangular Vandermonde matrix*

$$\mathbf{V}_{N,M}(\mathbf{z}) := \left(z_j^{k-1}\right)_{k,j=1}^{N,\,M}$$

with the vector $\mathbf{z} := (z_j)_{j=1}^{M}$, where $z_j = \mathrm{e}^{\mathrm{i}\varphi_j}$, $j = 1, \ldots, M$, are distinct nodes on the unit circle. Then the frequency analysis problem can be formulated in following matrix-vector form

$$\mathbf{V}_{N,M}(\mathbf{z})\,\mathbf{c} = \left(h_k\right)_{k=0}^{N-1}, \tag{84}$$

where $\mathbf{c} = (c_j)_{j=1}^{M}$ is the vector of complex coefficients of (70).

In practice, the order M of the exponential sum (70) is often unknown. Assume $L \in \mathbb{N}$ is a convenient upper bound of M and $M \le L \le N - M + 1$. In applications, such an upper bound L of M is mostly known *a priori*. If this is not the case, then one can choose $L \approx \frac{N}{2}$. Later we will see that the choice $L \approx \frac{N}{2}$ is optimal in some sense. Often the sequence $\{h_0, h_1, \ldots, h_{N-1}\}$ of (noisy) sampled data is called a *time series of length N*. Then we form the *L-trajectory matrix* of this time series

$$\mathbf{H}_{L,N-L+1} := \big(h_{\ell+m}\big)_{\ell,m=0}^{L-1,\,N-L} \in \mathbb{C}^{L\times(N-L+1)} \tag{85}$$

with the *window length* $L \in \{M, \ldots, N - M + 1\}$. Obviously $\mathbf{H}_{L,N-L+1}$ is a rectangular *Hankel matrix*.
For simplicity, we consider mainly noiseless data $h_k = h(k)$, $k = 0, \ldots, N - 1$, i.e.

$$\mathbf{H}_{L,N-L+1} = \big(h(\ell + m)\big)_{\ell,m=0}^{L-1,\,N-L} \in \mathbb{C}^{L\times(N-L+1)}\,. \tag{86}$$

The main step in the solution of the frequency analysis problem is the determination of the order M and the computation of the frequencies φ_j or alternatively of the nodes $z_j = \mathrm{e}^{\mathrm{i}\,\varphi_j}$, $j = 1, \ldots, M$. Afterwards one can calculate the coefficient vector $\mathbf{c} \in \mathbb{C}^M$ as least squares solution of the overdetermined linear system (84), i.e., the coefficient vector $\mathbf{c}$ is the solution of the least squares problem

$$\min_{\mathbf{c} \in \mathbb{C}^M} \|\mathbf{V}_{N,M}(\mathbf{z})\,\mathbf{c} - \big(h_k\big)_{k=0}^{N-1}\|_2\,.$$

By (71) the L-trajectory matrix (86) can be factorized in the following form

$$\mathbf{H}_{L,N-L+1} = \mathbf{V}_{L,M}(\mathbf{z})\,(\mathrm{diag}\,\mathbf{c})\,\mathbf{V}_{N-L+1,M}(\mathbf{z})^\top\,. \tag{87}$$

We denote square matrices with only one index.

Additionally we introduce the rectangular Hankel matrices

$$\begin{aligned} \mathbf{H}_{L,N-L}(s) &:= \mathbf{H}_{L,N-L+1}(1:L,\,1+s:N-L+s) \\ &= \left(h_{s+\ell+m}\right)_{\ell,m=0}^{L-1,\,N-L-1}, \quad s \in \{0,\,1\}, \end{aligned} \tag{88}$$

for $L \in \{M, \ldots, N-M\}$. Here we use the known submatrix notation. For example, $\mathbf{H}_{L,N-L+1}(1:L,\,1:N-L)$ is the submatrix of $\mathbf{H}_{L,N-L+1}$ obtained by extracting rows 1 through L and columns 1 through $N-L$. Observe that the first row or column of a matrix can be indexed by zero.

Lemma 76

Let $N \geq 2M$ be given. For each window length $L \in \{M,\,\ldots,\,N-M+1\}$, the rank of the L-trajectory matrix (86) *of noiseless data is M. The related Hankel matrices $\mathbf{H}_{L,N-L}(s)$, $s \in \{0,\,1\}$, possess the same rank M for each window length $L \in \{M,\,\ldots,\,N-M\}$.*

Proof: 1. As known, the square Vandermonde matrix $\mathbf{V}_M(\mathbf{z})$ is invertible. Further we have

$$\mathrm{rank}\,\mathbf{V}_{L,M}(\mathbf{z}) = M\,, \quad L \in \{M,\,\ldots,\,N-M+1\}\,, \qquad (89)$$

since $\mathrm{rank}\,\mathbf{V}_{L,M}(\mathbf{z}) \le \min\{L,\,M\} = M$ and since the submatrix $\big(z_k^{j-1}\big)_{j,k=1}^{M}$ of $\mathbf{V}_{L,M}(\mathbf{z})$ is invertible.
For $L \in \{M,\,\ldots,\,N-M+1\}$, we see by (89) that

$$\mathrm{rank}\,\mathbf{V}_{L,M}(\mathbf{z}) = \mathrm{rank}\,\mathbf{V}_{N-L+1,M}(\mathbf{z}) = M\,.$$

Thus the rank of the matrix $(\mathrm{diag}\,\mathbf{c})\,\mathbf{V}_{N-L+1,M}(\mathbf{z})^\top$ is equal to M. Hence we conclude that

$$\begin{aligned}\mathrm{rank}\,\mathbf{H}_{L,N-L+1} &= \mathrm{rank}\Big(\mathbf{V}_{L,M}(\mathbf{z})\,\big((\mathrm{diag}\,\mathbf{c})\,\mathbf{V}_{N-L+1,M}(\mathbf{z})^\top\big)\Big)\\ &= \mathrm{rank}\,\mathbf{V}_{L,M}(\mathbf{z}) = M\,.\end{aligned}$$

Note that this proof is mainly based on the factorization (87).

2. From $\mathbf{H}_{L,N-L}(0) = \mathbf{H}_{L,N-L}$ it follows by step 1 that $\operatorname{rank} \mathbf{H}_{L,N-L}(0) = M$ for each $L \in \{M, \ldots, N-M\}$. The Hankel matrix $\mathbf{H}_{L,N-L}(1)$ has also the rank M for each $L \in \{M, \ldots, N-M\}$. This follows from the fact that $\mathbf{H}_{L,N-L}(1)$ can be factorized in a similar form as (87), namely

$$\mathbf{H}_{L,N-L}(1) = \mathbf{V}_{L,M}(\mathbf{z})\,(\operatorname{diag}\mathbf{c})\,(\operatorname{diag}\mathbf{z})\,\mathbf{V}_{N-L,M}(\mathbf{z})^{\top}\,. \qquad \blacksquare$$

Consequently, the order M of the exponential sum (70) coincides with the rank of the Hankel matrices (86) and (88).
The ranges of $\mathbf{H}_{L,N-L+1}$ and $\mathbf{V}_{L,M}(\mathbf{z})$ coincide in the noiseless case with $M \le L \le N - M + 1$ by (87). If $L > M$, then the range of $\mathbf{V}_{L,M}(\mathbf{z})$ is a proper subspace of $\mathbb{C}^L$. This subspace is called *signal space* $\mathcal{S}_L$. The signal space $\mathcal{S}_L$ is of dimension M and is generated by the M columns $\mathbf{e}_L(\varphi_j)$, $j = 1, \ldots, M$, where

$$\mathbf{e}_L(\varphi) := \left(\mathrm{e}^{\mathrm{i}\,\ell\varphi}\right)_{\ell=0}^{L-1}, \quad \varphi \in [-\pi,\,\pi)\,.$$

Note that $\|\mathbf{e}_L(\varphi)\|_2 = \sqrt{L}$ for each $\varphi \in [-\pi,\,\pi)$.

The *noise space* $\mathcal{N}_L$ is defined as the orthogonal complement of $\mathcal{S}_L$ in $\mathbb{C}^L$. The dimension of $\mathcal{N}_L$ is equal to $L - M$.
By $\mathbf{Q}_L$ we denote the orthogonal projection of $\mathbb{C}^L$ onto the left noise space $\mathcal{N}_L$. Since $\mathbf{e}_L(\varphi_j) \in \mathcal{S}_L$, $j = 1, \ldots, M$, and $\mathcal{N}_L \perp \mathcal{S}_L$, we obtain that

$$\mathbf{Q}_L \, \mathbf{e}_L(\varphi_j) = \mathbf{0} \,, \quad j = 1, \ldots, M \,.$$

For $\varphi \in [-\pi, \pi) \setminus \{\varphi_1, \ldots, \varphi_M\}$, the vectors $\mathbf{e}_L(\varphi_1), \ldots, \mathbf{e}_L(\varphi_M), \mathbf{e}_L(\varphi) \in \mathbb{C}^L$ are linearly independent, since the square Vandermonde matrix

$$\big(\mathbf{e}_L(\varphi_1) \,|\, \ldots \,|\, \mathbf{e}_L(\varphi_M) \,|\, \mathbf{e}_L(\varphi)\big)(1 : M+1, \, 1 : M+1)$$

is invertible for each $L \geq M + 1$. Hence $\mathbf{e}_L(\varphi) \notin \mathcal{S}_L = \mathrm{span}\,\{\mathbf{e}_L(\varphi_1), \ldots, \mathbf{e}_L(\varphi_M)\}$, i.e., $\mathbf{Q}_L \mathbf{e}_L(\varphi) \neq \mathbf{0}$.

Thus the frequencies φ_j can be determined via the zeros of the *noise-space correlation function*

$$N_L(\varphi) := \frac{1}{\sqrt{L}} \, \|\mathbf{Q}_L \, \mathbf{e}_L(\varphi)\|_2 \, , \quad \varphi \in [-\pi, \, \pi) \, ,$$

since $N_L(\varphi_j) = 0$ for each $j = 1, \ldots, M$ and $0 < N_L(\varphi) \leq 1$ for all $\varphi \in [-\pi, \, \pi) \setminus \{\varphi_1, \ldots, \varphi_M\}$, where $\mathbf{Q}_L \mathbf{e}_L(\varphi)$ can be computed on a fine equispaced grid of $[-\pi, \, \pi)$. Alternatively, one can seek the peaks of the *imaging function*

$$J_L(\varphi) := \sqrt{L} \, \|\mathbf{Q}_L \, \mathbf{e}_L(\varphi)\|_2^{-1} \, , \quad \varphi \in [-\pi, \, \pi) \, .$$

In this approach, we prefer the zeros or rather the lowest local minima of the noise-space correlation function $N_L(\varphi)$.
In the next step we determine the orthogonal projection $\mathbf{Q}_L$ of $\mathbb{C}^L$ onto the noise space $\mathcal{N}_L$.

Here we use the *singular value decomposition* (SVD) the L-trajectory matrix $\mathbf{H}_{L,N-L+1}$, i.e.,

$$\mathbf{H}_{L,N-L+1} = \mathbf{U}_L\, \mathbf{D}_{L,N-L+1}\, \mathbf{W}_{N-L+1}^{\mathrm{H}}\,, \tag{90}$$

where

$$\begin{aligned} \mathbf{U}_L &= \big(\mathbf{u}_1 \,|\, \ldots \,|\, \mathbf{u}_L\big) \in \mathbb{C}^{L\times L}\,, \\ \mathbf{W}_{N-L+1} &= \big(\mathbf{w}_1 \,|\, \ldots \,|\, \mathbf{w}_{N-L+1}\big) \in \mathbb{C}^{(N-L+1)\times(N-L+1)} \end{aligned}$$

are unitary and where

$$\mathbf{D}_{L,N-L+1} = \mathrm{diag}\,\big(\sigma_1, \ldots, \sigma_{\min\{L,\,N-L+1\}}\big) \in \mathbb{R}^{L\times(N-L+1)}$$

is a rectangular diagonal matrix. The diagonal entries of $\mathbf{D}_{L,N-L+1}$ are arranged in nonincreasing order

$$\sigma_1 \geq \ldots \geq \sigma_M > \sigma_{M+1} = \ldots = \sigma_{\min\{L,N-L+1\}} = 0\,.$$

The columns of $\mathbf{U}_L$ are the *left singular vectors* of $\mathbf{H}_{L,N-L+1}$, the columns of $\mathbf{W}_{N-L+1}$ are the *right singular vectors* of $\mathbf{H}_{L,N-L+1}$. The nonnegative numbers σ_k are called *singular values* of $\mathbf{H}_{L,N-L+1}$.

The rank of $\mathbf{H}_{L,N-L+1}$ is equal to the number of positive singular values. Thus we can determine the order M of the exponential sum (70) by the number of positive singular values σ_j.
From (90) it follows that

$$\mathbf{H}_{L,N-L+1}\,\mathbf{W}_{N-L+1} = \mathbf{U}_L\,\mathbf{D}_{L,N-L+1}\,,$$
$$\mathbf{H}^{\mathrm{H}}_{L,N-L+1}\,\mathbf{U}_L = \mathbf{W}_{N-L+1}\,\mathbf{D}^{\top}_{L,N-L+1}\,.$$

Comparing the columns in above equations, for each $k = 1, \ldots, \min\{L,\, N-L+1\}$ we obtain

$$\mathbf{H}_{L,N-L+1}\,\mathbf{w}_k = \sigma_k\,\mathbf{u}_k\,, \quad \mathbf{H}^{\mathrm{H}}_{L,N-L+1}\,\mathbf{u}_k = \sigma_k\,\mathbf{w}_k\,.$$

Introducing the matrices

$$\begin{aligned} \mathbf{U}^{(1)}_{L,M} &:= \mathbf{U}_L(1:L,\,1:M) = \big(\mathbf{u}_1\,|\,\ldots\,|\,\mathbf{u}_M\big) \in \mathbb{C}^{L\times M}\,, \\ \mathbf{U}^{(2)}_{L,L-M} &:= \mathbf{U}_L(1:L,\,M+1:L) = \big(\mathbf{u}_{M+1}\,|\,\ldots\,|\,\mathbf{u}_L\big) \in \mathbb{C}^{L\times (L-M)}\,, \end{aligned}$$

we see that the columns of $\mathbf{U}^{(1)}_{L,M}$ form an orthonormal basis of $\mathcal{S}_L$ and that the columns of $\mathbf{U}^{(2)}_{L,L-M}$ are an orthonormal basis of $\mathcal{N}_L$.

Hence the orthogonal projection onto the noise space $\mathcal{N}_L$ has the form

$$\mathbf{Q}_L = \mathbf{U}^{(2)}_{L,L-M}\,(\mathbf{U}^{(2)}_{L,L-M})^{\mathrm{H}}\,.$$

Consequently, we obtain that

$$\begin{aligned}
&\|\mathbf{Q}_L\,\mathbf{e}_L(\varphi)\|_2^2\\
=\ &\langle \mathbf{Q}_L\,\mathbf{e}_L(\varphi),\, \mathbf{Q}_L\,\mathbf{e}_L(\varphi)\rangle = \langle (\mathbf{Q}_L)^2\,\mathbf{e}_L(\varphi),\, \mathbf{e}_L(\varphi)\rangle\\
=\ &\langle \mathbf{Q}_L\,\mathbf{e}_L(\varphi),\, \mathbf{e}_L(\varphi)\rangle = \langle \mathbf{U}^{(2)}_{L,L-M}\,(\mathbf{U}^{(2)}_{L,L-M})^{\mathrm{H}}\,\mathbf{e}_L(\varphi),\, \mathbf{e}_L(\varphi)\rangle\\
=\ &\langle (\mathbf{U}^{(2)}_{L,L-M})^{\mathrm{H}}\,\mathbf{e}_L(\varphi),\, (\mathbf{U}^{(2)}_{L,L-M})^{\mathrm{H}}\,\mathbf{e}_L(\varphi)\rangle = \|(\mathbf{U}^{(2)}_{L,L-M})^{\mathrm{H}}\,\mathbf{e}_L(\varphi)\|_2^2\,.
\end{aligned}$$

Hence the noise-space correlation function can be represented by

$$\begin{aligned}
N_L(\varphi) &= \frac{1}{\sqrt{L}}\,\|(\mathbf{U}^{(2)}_{L,L-M})^{\mathrm{H}}\,\mathbf{e}_L(\varphi)\|_2\\
&= \frac{1}{\sqrt{L}}\Big(\sum_{k=M+1}^{L}|\mathbf{u}_k^{\mathrm{H}}\,\mathbf{e}_L(\varphi)|^2\Big)^{1/2}, \quad \varphi \in [-\pi,\, \pi)\,.
\end{aligned}$$

In MUSIC, one determines the lowest local minima of the left noise-space correlation function, see e.g. [57, 41, 11, 33].

MUSIC via SVD

Input: $N \in \mathbb{N}$ with $N \geq 2\,M$, $L \approx \frac{N}{2}$ window length, $\tilde{h}_k = h(k) + e_k \in \mathbb{C}$, $k = 0, \ldots, N-1$, noisy sampled values of (70), $0 < \varepsilon \ll 1$ tolerance.

1. Compute the singular value decomposition

$$\mathbf{H}_{L,N-L+1} = \tilde{\mathbf{U}}_L\, \tilde{\mathbf{D}}_{L,N-L+1}\, \tilde{\mathbf{W}}^{\mathrm{H}}_{N-L+1}$$

of the rectangular Hankel matrix (85), where the singular values $\tilde{\sigma}_\ell$ are arranged in nonincreasing order. Determine the numerical rank M of (85) such that $\tilde{\sigma}_M \geq \varepsilon\,\tilde{\sigma}_1$ and $\tilde{\sigma}_{M+1} < \varepsilon\tilde{\sigma}_1$. Form the matrix

$$\tilde{\mathbf{U}}^{(2)}_{L,L-M} = \tilde{\mathbf{U}}_L(1:L,\, M+1:L) = \big(\tilde{\mathbf{u}}_{M+1}\,|\,\ldots\,|\,\tilde{\mathbf{u}}_L\big)\,.$$

2. Calculate the squared noise-space correlation function

$$\tilde{N}_L(\varphi)^2 := \frac{1}{L}\sum_{k=M+1}^{L} |\tilde{\mathbf{u}}_k^{\mathrm{H}}\,\mathbf{e}_L(\varphi)|^2$$

on the equispaced grid $\{\frac{(2k-S)\pi}{S} :\, k = 0, \ldots, S-1\}$ for sufficiently large $S \in \mathbb{N}$ by fast Fourier transforms.

3. The M lowest local minima of $\tilde{N}_L(\frac{(2k-S)\pi}{S})$, $k = 0, \ldots, S-1$, form the frequencies $\tilde{\varphi}_1, \ldots, \tilde{\varphi}_M$. Set $\tilde{z}_j := \mathrm{e}^{\mathrm{i}\,\tilde{\varphi}_j}$, $j = 1, \ldots, M$.
4. Compute the coefficient vector $\tilde{\mathbf{c}} := (\tilde{c}_j)_{j=1}^M \in \mathbb{C}^M$ as solution of the least squares problem

$$\min_{\tilde{\mathbf{c}} \in \mathbb{C}^M} \|\mathbf{V}_{N,M}(\tilde{\mathbf{z}})\, \tilde{\mathbf{c}} - \big(\tilde{h}_k\big)_{k=0}^{N-1}\|_2 \,,$$

where $\tilde{\mathbf{z}} := \big(\tilde{z}_j\big)_{j=1}^M$ denotes the vector of computed nodes.
Output: $M \in \mathbb{N}$, $\tilde{\varphi}_j \in [-\pi,\, \pi)$, $\tilde{c}_j \in \mathbb{C}$, $j = 1, \ldots, M$.

The *approximate Prony method* can be immediately derived from the MUSIC method. We start with the squared noise-space correlation function

$$\begin{aligned} N_L(\varphi)^2 &= \frac{1}{L}\,\|(\mathbf{U}^{(2)}_{L,L-M})^{\mathrm{H}}\,\mathbf{e}_L(\varphi)\|_2^2 \\ &= \frac{1}{L}\sum_{k=M+1}^{L}\left|\mathbf{u}_k^{\mathrm{H}}\,\mathbf{e}_L(\varphi)\right|^2, \quad \varphi\in[-\pi,\,\pi)\,. \end{aligned}$$

For noiseless data, all frequencies φ_j, $j=1,\ldots,M$, are zeros of $N_L(\varphi)^2$ and hence especially zeros of

$$\left|\mathbf{u}_L^{\mathrm{H}}\,\mathbf{e}_L(\varphi)\right|^2.$$

Thus we obtain $\mathbf{u}_L^{\mathrm{H}}\,\mathbf{e}_L(\varphi_j)=0$ for $j=1,\ldots,M$. Note that $\mathbf{u}_L^{\mathrm{H}}\,\mathbf{e}_L(\varphi)$ can have additional zeros. For noisy data we observe small values $\left|\mathbf{u}_L^{\mathrm{H}}\,\mathbf{e}_L(\varphi)\right|^2$ near φ_j. Finally we determine the order M of the exponential sum (70) by the number of sufficiently large coefficients in the reconstructed exponential sum.

Approximate Prony method

Input: $N \in \mathbb{N}$ with $N \geq 2\,M$, $L \approx \frac{N}{2}$ window length,
$\tilde{h}_k = h(k) + e_k \in \mathbb{C}$, $k = 0, \ldots, N-1$, noisy sampled values of (70), $\varepsilon > 0$ lower bound with $|c_j| \geq 2\,\varepsilon$, $j = 1, \ldots, M$.

1. Compute the singular vector $\mathbf{u}_L = (u_\ell)_{\ell=0}^{L-1} \in \mathbb{C}^L$ of the rectangular Hankel matrix (85).

2. Calculate

$$\mathbf{u}_L^{\mathrm{H}}\,\mathbf{e}_L(\varphi) = \sum_{\ell=0}^{L-1} \bar{u}_\ell\,\mathrm{e}^{\mathrm{i}\,\ell\varphi}$$

on the equispaced grid $\{\frac{(2k-S)\pi}{S} :\ k = 0, \ldots, S-1\}$ for sufficiently large $S \in \mathbb{N}$ by FFT.

3. Determine the lowest local minima ψ_j, $j = 1, \ldots, \tilde{M}$, of $|\mathbf{u}_L^*\,\mathbf{e}_L(\frac{(2k-S)\pi}{S})|^2$, $k = 0, \ldots, S-1$. Set $\tilde{w}_j := \mathrm{e}^{\mathrm{i}\tilde{\psi}_j}$, $j = 1, \ldots, \tilde{M}$.

4. Compute the coefficients $\tilde{d}_j \in \mathbb{C}$ as least squares solution of the overdetermined linear system

$$\sum_{j=1}^{\tilde{M}} \tilde{d}_j \, \tilde{w}_j = h_k \,, \quad k = 0, \ldots, N-1 \,.$$

Delete all the $\tilde{w}_k$ with $|\tilde{d}_k| \le \varepsilon$ and denote the remaining nodes by $\tilde{z}_j$, $j = 1, \ldots, M$.
5. Compute the coefficients $\tilde{c}_j \in \mathbb{C}$ as least squares solution of the overdetermined linear system

$$\sum_{j=1}^{M} \tilde{c}_j \, \tilde{z}_j = h_k \,, \quad k = 0, \ldots, N-1 \,.$$

Output: $M \in \mathbb{N}$, $\tilde{\varphi}_j \in [-\pi, \pi)$, $\tilde{c}_j \in \mathbb{C}$, $j = 1, \ldots, M$.

Finally we sketch the frequently used ESPRIT method (see [54, 50]) which is based on singular value decomposition of the rectangular Hankel matrix. First we assume that noiseless data $\tilde{h}_k = h(k)$, $k = 0, \ldots, N-1$, of (70) are given. The set of all matrices of the form

$$z\,\mathbf{H}_{L,N-L}(0) - \mathbf{H}_{L,N-L}(1)\,, \quad z \in \mathbb{C}\,, \tag{91}$$

is called a rectangular *matrix pencil*. If a scalar $z_0 \in \mathbb{C}$ and a nonzero vector $\mathbf{v} \in \mathbb{C}^{N-L}$ satisfy

$$z_0\,\mathbf{H}_{L,N-L}(0)\,\mathbf{v} = \mathbf{H}_{L,N-L}(1)\,\mathbf{v}\,,$$

then z_0 is called an *eigenvalue* of the matrix pencil and $\mathbf{v}$ is called *eigenvector*. Note that a rectangular matrix pencil may not have eigenvalues in general. The ESPRIT method is based on following result:

Lemma 77

Assume that $N \in \mathbb{N}$ with $N \geq 2\,M$ and $L \in \{M, \ldots, N-M\}$ are given. In the case of noiseless data, the matrix pencil (91) has the nodes $z_j = \mathrm{e}^{\mathrm{i}\varphi_j}$, $j = 1, \ldots, M$, as eigenvalues. Further, zero is an eigenvalue of (91) with $N-L-M$ linearly independent eigenvectors.

Proof:

1. Let p denote the Prony polynomial (72) and let $q(z) := z^{N-L-M}\,p(z)$. Then the companion matrix of q reads as follows

$$\mathbf{C}_{N-L}(q) = \big(\mathbf{e}_1 \,|\, \mathbf{e}_2 \,|\, \ldots \,|\, \mathbf{e}_{N-L-1} \,|\, -\mathbf{q}\big)$$

with $\mathbf{q} := \big(0, \ldots, 0, p_0, p_1, \ldots, p_{M-1}\big)^\top$, where p_k are the coefficients of (72). Here $\mathbf{e}_k = (\delta_{k-\ell})_{\ell=0}^{N-L-1}$ denote the canonical basis vectors of $\mathbb{C}^{N-L}$. By (74) and (88) we obtain that

$$\mathbf{H}_{L,N-L}(0)\,\mathbf{q} = -\big(h(\ell)\big)_{\ell=N-L}^{N-1}$$

and hence

$$\mathbf{H}_{L,N-L}(0)\,\mathbf{C}_{N-L}(q) = \mathbf{H}_{L,N-L}(1)\,. \qquad (92)$$

2. Thus it follows by (92) that the rectangular matrix pencil (91) coincides with the square matrix pencil $z\,\mathbf{I}_{N-L} - \mathbf{C}_{N-L}(q)$ up to a matrix factor,

$$z\,\mathbf{H}_{L,N-L}(0) - \mathbf{H}_{L,N-L}(1) = \mathbf{H}_{L,N-L}(0)\,\big(z\,\mathbf{I}_{N-L} - \mathbf{C}_{N-L}(q)\big)\,.$$

Now we have to determine the eigenvalues of the companion matrix $\mathbf{C}_{N-L}(q)$. By

$$\det\,\big(z\,\mathbf{I}_{N-L} - \mathbf{C}_{N-L}(q)\big) = q(z) = z^{N-L-M}\prod_{j=1}^{M}(z - z_j)$$

the eigenvalues of $\mathbf{C}_{N-L}(q)$ are zero and z_j, $j = 1, \ldots, M$. Obviously, $z = 0$ is an eigenvalue of the rectangular matrix pencil (91), which has $L - M$ linearly independent eigenvectors, since $\operatorname{rank}\mathbf{H}_{L,N-L}(0) = M$ by Lemma 76. For each $z = z_j$, $j = 1, \ldots, M$, we can compute an eigenvector $\mathbf{v} = (v_k)_{k=0}^{N-L-1}$ of $\mathbf{C}_{N-L}(q)$, if we set $v_{N-L-1} = z_j$.

Thus we obtain

$$\big(z_j\, \mathbf{H}_{L,N-L}(0) - \mathbf{H}_{L,N-L}(1)\big)\, \mathbf{v} = \mathbf{0}\,.$$

We have shown that the generalized eigenvalue problem of the rectangular matrix pencil (91) can be reduced to the classical eigenvalue problem of the square matrix $\mathbf{C}_{N-L}(q)$. ■

We start the ESPRIT method by the singular value decomposition (90) of the L-trajectory matrix $\mathbf{H}_{L,N-L+1}$ with a window length $L \in \{M, \ldots, N-M\}$. Introducing the matrices

$$\mathbf{U}_{L,M} := \mathbf{U}_L(1:L,\, 1:M)\,, \quad \mathbf{W}_{N-L+1,M} := \mathbf{W}_{N-L+1}(1:N{-}L{+}1,\, 1:M)$$

with orthonormal columns as well as the diagonal matrix $\mathbf{D}_M := \mathrm{diag}\,(\sigma_j)_{j=1}^{M}$, we obtain the partial singular value decomposition of the matrix (86) with noiseless entries, i.e.,

$$\mathbf{H}_{L,N-L+1} = \mathbf{U}_{L,M}\, \mathbf{D}_M\, \mathbf{W}_{N-L+1,M}^{\mathrm{H}}\,.$$

Setting

$$\mathbf{W}_{N-L,M}(s) := \mathbf{W}_{N-L+1,M}(1+s : N-L+s, 1 : M)\,, \quad s \in \{0,\,1\}\,, \tag{93}$$

it follows by (93) and (88) that both Hankel matrices (88) can be simultaneously factorized in the form

$$\mathbf{H}_{L,N-L}(s) = \mathbf{U}_{L,M}\,\mathbf{D}_M\,\mathbf{W}_{N-L,M}(s)^{\mathrm{H}}\,, \quad s \in \{0,\,1\}\,. \tag{94}$$

Since $\mathbf{U}_{L,M}$ has orthonormal columns and since $\mathbf{D}_M$ is invertible, the generalized eigenvalue problem of the matrix pencil

$$z\,\mathbf{W}_{N-L,M}(0)^{\mathrm{H}} - \mathbf{W}_{N-L,M}(1)^{\mathrm{H}}\,, \quad z \in \mathbb{C}\,, \tag{95}$$

has the same non-zero eigenvalues z_j, $j = 1, \ldots, M$, as the matrix pencil (91) except for additional zero eigenvalues. Finally we determine the nodes z_j, $j = 1, \ldots, M$, as eigenvalues of the matrix

$$\mathbf{F}_M^{\mathrm{SVD}} := \mathbf{W}_{N-L,M}(1)^{\mathrm{H}}\left(\mathbf{W}_{N-L,M}(0)^{\mathrm{H}}\right)^{+} \in \mathbb{C}^{M\times M}\,, \tag{96}$$

where $\left(\mathbf{W}_{N-L,M}(0)^{\mathrm{H}}\right)^{+}$ denote the Moore–Penrose pseudoinverse.

Analogously, we can handle the general case of noisy data $\tilde{h}_k = h(k) + e_k \in \mathbb{C}$, $k = 0, \ldots, N-1$, with small error terms $e_k \in \mathbb{C}$, where $|e_k| \le \varepsilon_1$ and $0 < \varepsilon_1 \ll 1$. For the Hankel matrix (90) with the singular values $\tilde{\sigma}_1 \ge \ldots \ge \tilde{\sigma}_{\min\{L, N-L+1\}} \ge 0$, we can calculate the numerical rank M of (85) by the property $\tilde{\sigma}_M \ge \varepsilon\, \tilde{\sigma}_1$ and $\tilde{\sigma}_{M+1} < \varepsilon\, \tilde{\sigma}_1$ with convenient chosen tolerance ε. Using the IEEE double precision arithmetic, one can choose $\varepsilon = 10^{-10}$ for given noiseless data. In the case of noisy data, one has to use a larger tolerance $\varepsilon > 0$.

For the rectangular Hankel matrix (85) with noisy entries, we use its singular value decomposition

$$\mathbf{H}_{L,N-L+1} = \tilde{\mathbf{U}}_L \, \tilde{\mathbf{D}}_{L,N-L+1} \, \tilde{\mathbf{W}}_{N-L+1}^{\mathrm{H}}$$

and define as above the matrices $\tilde{\mathbf{U}}_{L,M}$, $\tilde{\mathbf{D}}_M := \mathrm{diag}\,\big(\tilde{\sigma}_j\big)_{j=1}^{M}$, and $\tilde{\mathbf{W}}_{N-L+1,M}$. Then

$$\tilde{\mathbf{U}}_{L,M} \, \tilde{\mathbf{D}}_M \, \tilde{\mathbf{W}}_{N-L+1,M}^{\mathrm{H}}$$

is a low-rank approximation of (85). Analogously to (93) and (96), we introduce corresponding matrices $\tilde{\mathbf{W}}_{N-L,M}(s)$, $s \in \{0, 1\}$ and $\tilde{\mathbf{F}}_M^{\mathrm{SVD}}$. Note that

$$\tilde{\mathbf{K}}_{L,N-L}(s) := \tilde{\mathbf{U}}_{L,M} \, \tilde{\mathbf{D}}_M \, \tilde{\mathbf{W}}_{N-L,M}(s)^* \,, \quad s \in \{0, 1\} \tag{97}$$

is a low-rank approximation of $\tilde{\mathbf{H}}_{L,N-L}(s)$. Thus the SVD-based ESPRIT algorithm reads as follows:

ESPRIT via SVD

Input: $N \in \mathbb{N}$ with $N \gg 1$, $M \le L \le N - M$, $L \approx \frac{N}{2}$, M unknown order of (70)), $h_k = h(k) + e_k \in \mathbb{C}$, $k = 0, \ldots, N-1$, noisy sampled values of (70), $0 < \varepsilon \ll 1$ tolerance.

1. Compute the singular value decomposition of the rectangular Hankel matrix (85). Determine the numerical rank M of (85) such that $\tilde{\sigma}_M \ge \varepsilon\,\tilde{\sigma}_1$ and $\tilde{\sigma}_{M+1} < \varepsilon\tilde{\sigma}_1$. Form the matrices $\tilde{\mathbf{W}}_{N-L,M}(s)$, $s \in \{0, 1\}$, as in (93).
2. Calculate the square matrix $\tilde{\mathbf{F}}_M^{\mathrm{SVD}}$ as in (96) and compute all eigenvalues $\tilde{z}_j$, $j = 1, \ldots, M$, of $\tilde{\mathbf{F}}_M^{\mathrm{SVD}}$. Replace $\tilde{z}_j$ by the corrected value $\frac{\tilde{z}_j}{|\tilde{z}_j|}$, $j = 1, \ldots, M$, and set $\tilde{\varphi}_j := \log \tilde{z}_j\,, \quad j = 1, \ldots, M$, where log denotes the principal value of the complex logarithm.
3. Compute the coefficient vector $\tilde{\mathbf{c}} := (\tilde{c}_j)_{j=1}^M \in \mathbb{C}^M$ as solution of the least squares problem

$$\|\mathbf{V}_{N,M}(\tilde{\mathbf{z}})\,\tilde{\mathbf{c}} - \big(\tilde{h}_k\big)_{k=0}^{N-1}\|_2 = \min\,,$$

where $\tilde{\mathbf{z}} := \big(\tilde{z}_j\big)_{j=1}^M$ denotes the vector of computed nodes.

Output: $M \in \mathbb{N}$, $\tilde{\varphi}_j \in [-\pi, \pi)$, $\tilde{c}_j \in \mathbb{C}$ for $j = 1, \ldots, M$.

Remark 78

One can pass on the computation of the Moore–Penrose pseudoinverse in (96)*. Then the second step of Algorithm 263 reads as follows (see [51, Algorithm 3.1]):*

$2'$*. Calculate the matrix products*

$$\tilde{\mathbf{A}}_M := \tilde{\mathbf{W}}_{N-L,M}(0)^{\mathrm{H}}\,\tilde{\mathbf{W}}_{N-L,M}(0)\,, \quad \tilde{\mathbf{B}}_M := \tilde{\mathbf{W}}_{N-L,M}(1)^{\mathrm{H}}\,\tilde{\mathbf{W}}_{N-L,M}(0)$$

and compute all eigenvalues $\tilde{z}_j$*,* $j = 1, \ldots, M$*, of the square matrix pencil* $z\,\tilde{\mathbf{A}}_M - \tilde{\mathbf{B}}_M$*,* $z \in \mathbb{C}$*, by the QZ–Algorithm (see [13, pp. 384–385]). Set* $\tilde{\varphi}_j := \log \tilde{z}_j$*,* $j = 1, \ldots, M$*.* $\square$

Remark 79

For various numerical examples as well as for a comparison between Algorithm 263 and another Prony–like method see [48]. The Algorithm 263 is very similar to the Algorithm 3.2 in [51]. Note that one can also use the QR decomposition of the rectangular Hankel matrix (85) *instead of the singular value decomposition. In that case one obtains an algorithm that is similar to the matrix pencil method [21, 55], see also Algorithm 3.1 in [51]. The matrix pencil method has been also applied to reconstruction of shapes from moments, see e.g. [12].*
In [4, 52], the condition number of a rectangular Vandermonde matrix is estimated. It is shown that this matrix is well conditioned, provided the nodes z_j are not extremely close to each other and provided N is large enough. □

Remark 80

The given data sequence $\{h_0, h_1, \ldots, h_{N-1}\}$ *can be also interpreted as time series. A powerful tool of time series analysis is the singular spectrum analysis (see [14, 15]). Similarly as step 1 of the Algorithm 263, this technique is based on the singular value decomposition of a rectangular Hankel matrix constructed upon the given time series* h_k*. By this method, the original time series can be decomposed into a sum of interpretable components such as trend, oscillatory components, and noise. For further details and numerous applications see [14, 15].* □

Remark 81

The considered Prony-like method can also be interpreted as a model reduction based on low-rank approximation of Hankel matrices, see [42]. The structured low-rank approximation problem reads as follows: For a given structure specification $\mathcal{S} : \mathbb{C}^K \to \mathbb{C}^{L\times N}$ *with* $L < N$, *a parameter vector* $\mathbf{h} \in \mathbb{C}^K$ *and an integer* M *with* $0 < M < L$, *find a vector*

$$\hat{\mathbf{h}}^* = \arg\min \{\|\mathbf{h} - \hat{\mathbf{h}}\| : \hat{\mathbf{h}} \in \mathbb{C}^K \text{ with } \operatorname{rank} \mathcal{S}(\hat{\mathbf{h}}) \le M\},$$

where $\|\cdot\|$ *denotes a suitable norm in* $\mathbb{C}^K$.

Remark 81 (continue)

In the special case of a Hankel matrix structure, the Hankel matrix $\mathcal{S}(\mathbf{h}) = (h_{\ell+k})_{\ell,\,k=0}^{L-1,N-1}$ *is rank-deficient of order* M *if there exists a nonzero vector* $\mathbf{p} = (p_k)_{k=0}^{M-1}$ *so that*

$$\sum_{k=0}^{M-1} p_k \, h(m+k) = -h(M+m)$$

for all $m = 0, \ldots, N+L-M-1$*. Equivalently, the values* $h(k)$ *can be interpreted as function values of an exponential sum of order* M *in* (70)*. The special kernel structure of rank-deficient Hankel matrices can already be found in [18].* □

Stability of exponentials

The three methods for recovery of exponential sums, namely MUSIC, approximate Prony method, and ESPRIT, were derived for noiseless data. Fortunately, these methods work also for noisy data $h_k = h(k) + e_k$, $k = 0, \ldots, N-1$, with error terms $e_k \in \mathbb{C}$ provided that the bound $\varepsilon_1 > 0$ of all $|e_k|$ is small enough. This property is based on the perturbation theory of a singular value decomposition of a rectangular Hankel matrix. Here we have to assume that the frequencies $\varphi_j \in [-\pi, \pi)$, $j = 1, \ldots, M$, are not too close to each other, that the number N of samples is sufficiently large with $N \geq 2\,M$, and that the window length $L \approx \frac{N}{2}$. We start with following stability result, see [22], [70, pp. 162–164] or [34, pp. 59 - 66].

Lemma 82

Let $M \in \mathbb{N}$ *and* $T > 0$ *be given. If the ordered frequencies* $\varphi_j \in \mathbb{R}$*,* $j = 1, \ldots, M$*, fulfill the gap condition*

$$\varphi_{j+1} - \varphi_j \geq q > \frac{\pi}{T}\,, \quad j = 1, \ldots, M-1\,, \tag{98}$$

then the exponentials $\mathrm{e}^{\mathrm{i}\,\varphi_j\,\cdot}$*,* $j = 1, \ldots, M$*, are Riesz stable in* $L_2[0,\,2T]$*, i.e., for all vectors* $\mathbf{c} = (c_j)_{j=1}^{M} \in \mathbb{C}^M$ *it holds the Ingham inequalities*

$$\alpha(T)\,\|\mathbf{c}\|_2^2 \leq \|\sum_{j=1}^{M} c_j\,\mathrm{e}^{\mathrm{i}\,\varphi_j\,\cdot}\|_{L_2[0,\,2T]}^2 \leq \beta(T)\,\|\mathbf{c}\|_2^2 \tag{99}$$

with positive constants
$\alpha(T) := \frac{2}{\pi}\left(1 - \frac{\pi^2}{T^2 q^2}\right), \quad \beta(T) := \frac{4\sqrt{2}}{\pi}\left(1 + \frac{\pi^2}{4T^2q^2}\right)$ *and the norm*
$\|f\|_{L_2[0,2T]} := \left(\frac{1}{2T}\int_0^{2T} |f(t)|^2\,\mathrm{d}t\right)^{1/2}, \quad f \in L^2[0,\,2T]\,.$

Proof: 1. Let

$$h(x) := \sum_{j=1}^{M} c_j \, \mathrm{e}^{\mathrm{i}\,\varphi_j x}\,, \quad x \in [0,\, 2T]\,. \tag{100}$$

Substituting $t = x - T \in [-T,\, T]$, we obtain

$$f(t) = \sum_{j=1}^{M} d_j \, \mathrm{e}^{\mathrm{i}\,\varphi_j t}\,, \quad t \in [-T,\, T]$$

with $d_j := c_j \, \mathrm{e}^{\mathrm{i}\,\varphi_j T}$, $j = 1, \ldots, M$. Note that $|d_j| = |c_j|$ and

$$\|f\|_{L_2[-T,\, T]} = \|h\|_{L_2[0,\, 2T]}\,.$$

For simplicity, we can assume that $T = \pi$. If $T \neq \pi$, then we substitute $s = \frac{\pi}{T}\, t \in [-\pi, \pi]$ for $t \in [-T, T]$ such that

$$f(t) = f(\frac{T}{\pi}\, s) = \sum_{j=1}^{M} d_j \, \mathrm{e}^{\mathrm{i}\,\psi_j s}\,, \quad s \in [-\pi,\, \pi]\,,$$

with $\psi_j := \frac{T}{\pi}\, \varphi_j$.

Thus we receive by the gap condition (98) that

$$\psi_{j+1} - \psi_j = \frac{T}{\pi}\left(\varphi_{j+1} - \varphi_j\right) \ge \frac{T}{\pi}\, q > 1\,.$$

2. For fixed function $k \in L_1(\mathbb{R})$ and its Fourier transform

$$\hat{k}(\omega) := \int_{\mathbb{R}} k(t)\, \mathrm{e}^{-\mathrm{i}\,\omega\, t}\, \mathrm{d}t\,, \quad \omega \in \mathbb{R}\,,$$

we see that

$$\begin{aligned} \int_{\mathbb{R}} k(t)\, |f(t)|^2\, \mathrm{d}t &= \sum_{j=1}^{M} \sum_{\ell=1}^{M} d_j\, \bar{d}_\ell \int_{\mathbb{R}} k(t)\, \mathrm{e}^{-\mathrm{i}\,(\psi_\ell - \psi_j)\, t}\, \mathrm{d}t \\ &= \sum_{j=1}^{M} \sum_{\ell=1}^{M} d_j\, \bar{d}_\ell\, \hat{k}(\psi_\ell - \psi_j)\,. \end{aligned}$$

If we choose

$$k(t) := \begin{cases} \cos \frac{t}{2} & t \in [-\pi,\, \pi]\,, \\ 0 & t \in \mathbb{R} \setminus [-\pi,\, \pi]\,, \end{cases}$$

then we obtain the Fourier transform

$$\hat{k}(\omega) = \frac{4\,\cos(\pi\omega)}{1 - 4\,\omega^2}\,, \quad \omega \in \mathbb{R} \setminus \{-\frac{1}{2},\, \frac{1}{2}\}\,, \tag{101}$$

with $\hat{k}(\pm\frac{1}{2}) = \pi$ and hence

$$\int_{-\pi}^{\pi} \cos\frac{t}{2}\,|f(t)|^2\,\mathrm{d}t = \sum_{j=1}^{M} \sum_{\ell=1}^{M} \hat{k}(\psi_\ell - \psi_j)\, d_j\, \bar{d}_\ell\,. \tag{102}$$

3. From (102) it follows immediately that

$$\int_{-\pi}^{\pi} |f(t)|^2 \,\mathrm{d}t \geq \sum_{j=1}^{M} \sum_{\ell=1}^{M} \hat{k}(\psi_\ell - \psi_j)\, d_j\, \bar{d}_\ell\,.$$

Let S_1 denote that part of the above double sum for which $j = \ell$ and let S_2 be the remaining part. Clearly, by $\hat{k}(0) = 4$ we get

$$S_1 = 4 \sum_{j=1}^{M} |d_j|^2\,. \tag{103}$$

Since $\hat{k}$ is even and since $2\,|d_j\,\bar{d}_\ell| \leq |d_j|^2 + |d_\ell|^2$, there are constants $\theta_{j,\ell} \in \mathbb{C}$ with $|\theta_{j,\ell}| \leq 1$ and $\theta_{j,\ell} = \bar{\theta}_{\ell,j}$ such that

$$\begin{aligned} S_2 &= \sum_{j=1}^{M} \sum_{\substack{\ell=1 \\ \ell \neq j}}^{M} \frac{|d_j|^2 + |d_\ell|^2}{2}\, \theta_{j,\ell}\, |\hat{k}(\psi_\ell - \psi_j)| \\ &= \sum_{j=1}^{M} |d_j|^2 \Big(\sum_{\substack{\ell=1 \\ \ell \neq j}}^{M} \mathrm{Re}\,\theta_{j,\ell}\, |\hat{k}(\psi_\ell - \psi_j)| \Big)\,. \end{aligned}$$

Consequently, there exists a constant $\theta \in [-1, 1]$ such that

$$S_2 = \theta \sum_{j=1}^{M} |d_j|^2 \Big(\sum_{\substack{\ell=1 \\ \ell \neq j}}^{M} |\hat{k}(\psi_\ell - \psi_j)| \Big) . \tag{104}$$

Since $|\psi_\ell - \psi_j| \geq |\ell - j|\, q > 1$ for $\ell \neq j$ by (98), we receive by 101 that

$$\begin{aligned} \sum_{\substack{\ell=1 \\ \ell \neq j}}^{M} |\hat{k}(\psi_\ell - \psi_j)| &\leq \sum_{\substack{\ell=1 \\ \ell \neq j}}^{M} \frac{4}{4\,(\ell - j)^2\, q^2 - 1} < \frac{8}{q^2} \sum_{n=1}^{\infty} \frac{1}{4n^2 - 1} \\ &= \frac{4}{q^2} \sum_{n=1}^{\infty} \Big(\frac{1}{2n-1} - \frac{1}{2n+1} \Big) = \frac{4}{q^2} . \end{aligned} \tag{105}$$

Hence from (103) – (105) it follows that

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} |f(t)|^2 \,\mathrm{d}t \geq \alpha(\pi) \sum_{j=1}^{M} |d_j|^2$$

with $\alpha(\pi) = \frac{2}{\pi} \big(1 - \frac{1}{q^2}\big)$.

In the case $T \neq \pi$, we obtain $\alpha(T) = \frac{2}{\pi}\left(1 - \frac{\pi^2}{T^2 q^2}\right)$ by the substitution in step 1 and hence

$$\|h\|^2_{L_2[0,\,2T]} \geq \alpha(T) \sum_{j=1}^{M} |c_j|^2 = \alpha(T)\, \|\mathbf{c}\|_2^2 \,.$$

4. From (102) – (105) it follows that

$$\int_{-\pi}^{\pi} \cos\frac{t}{2}\, |f(t)|^2 \,\mathrm{d}t \geq \int_{-\pi/2}^{\pi/2} \cos\frac{t}{2}\, |f(t)|^2 \,\mathrm{d}t \geq \frac{\sqrt{2}}{2} \int_{-\pi/2}^{\pi/2} |f(t)|^2 \,\mathrm{d}t$$

and further

$$\begin{aligned} & \int_{-\pi}^{\pi} \cos\frac{t}{2}\, |f(t)|^2 \,\mathrm{d}t \\ = \;& \sum_{j=1}^{M} \sum_{\ell=1}^{M} \hat{k}(\psi_\ell - \psi_j)\, d_j\, \bar{d}_\ell \\ \leq \;& 4 \sum_{j=1}^{M} |d_j|^2 + \frac{4}{q^2} \sum_{j=1}^{M} |d_j|^2 = \Big(1 + \frac{1}{q^2}\Big) \sum_{j=1}^{M} |d_j|^2 \,. \end{aligned}$$

Thus we obtain

$$\frac{1}{\pi}\int_{-\pi/2}^{\pi/2} |f(t)|^2\,\mathrm{d}t \le \frac{4\sqrt{2}}{\pi}\Big(1+\frac{1}{q^2}\Big)\sum_{j=1}^{M} |d_j|^2\,. \qquad (106)$$

5. Now we consider the function

$$g(t) := f(2t) = \sum_{j=1}^{M} d_j\,\mathrm{e}^{2\mathrm{i}\psi_j t}\,, \quad t \in \big[-\frac{\pi}{2}, \frac{\pi}{2}\big]\,,$$

where the ordered frequencies $2\,\psi_j$ fulfill the gap condition

$$2\,\psi_{j+1} - 2\,\psi_j \ge 2\,q\,, \quad j = 1, \ldots, M-1\,.$$

Applying (106) to the function g, we receive

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} |f(t)|^2\,\mathrm{d}t = \frac{1}{\pi}\int_{-\pi/2}^{\pi/2} |g(t)|^2\,\mathrm{d}t \le \frac{4\sqrt{2}}{\pi}\Big(1+\frac{1}{4q^2}\Big)\sum_{j=1}^{M} |d_j|^2\,.$$

Hence $\beta(\pi) = \frac{4\sqrt{2}}{\pi}\,(1 + \frac{1}{4q^2})$ and $\beta(T) = \frac{4\sqrt{2}}{\pi}\,(1 + \frac{\pi^2}{4\,T^2 q^2})$ by the substitution in step 1.

Thus we obtain

$$\|h\|^2_{L_2[0,\,2T]} \le \beta(T) \sum_{j=1}^{M} |d_j|^2 = \beta(T)\,\|\mathbf{c}\|_2^2 \,.$$

This completes the proof. ■

Remark 83

The Ingham inequalities (99) *can be considered as far-reaching generalization of the Parseval equa for Fourier series. The constants* $\alpha(T)$ *and* $\beta(T)$ *are not optimal in general. Note that these constants are independently of* M*. The assumption* $q > \frac{\pi}{T}$ *is necessary for the existence of positive* $\alpha(T)$*. Compare also with [9, Theorems 7.6.5 and 7.6.6] and [37].* □

In the following, we present a discrete version of the Ingham inequalities (99) (see [40, 43, 2]). For sufficiently large integer $P > M$, we consider the rectangular Vandermonde matrix

$$\mathbf{V}_{P,M}(\mathbf{z}) := \big(z_j^{k-1}\big)_{k,j=1}^{P,\,M} = \begin{pmatrix} 1 & 1 & \ldots & 1 \\ z_1 & z_2 & \ldots & z_M \\ \vdots & \vdots & & \vdots \\ z_1^{P-1} & z_2^{P-1} & \ldots & z_M^{P-1} \end{pmatrix}$$

with $\mathbf{z} = (z_j)_{j=1}^{M}$, where $z_j = \mathrm{e}^{\mathrm{i}\,\varphi_j}$, $j = 1, \ldots, M$, are distinct nodes on the unit circle. Setting $\varphi_j = 2\pi\,\psi_j$, $j = 1, \ldots, M$, we measure the distance between distinct frequencies ψ_j, ψ_ℓ by $d(\psi_j - \psi_\ell)$, where $d(x)$ denotes the *distance of $x \in \mathbb{R}$ to the nearest integer*, i.e.,

$$d(x) := \min_{n \in \mathbb{Z}} |x - n| \in [0,\, \frac{1}{2}]\,.$$

Our aim is a good estimation of the spectral condition number of $\mathbf{V}_{P,M}(\mathbf{z})$. Therefore we assume that ψ_j, $j = 1, \ldots, M$, satisfy the *gap condition*

$$\min\{d(\psi_j - \psi_\ell) : j, \ell = 1, \ldots, M, j \neq \ell\} \geq \Delta > 0 . \qquad (107)$$

The following discussion is mainly based on a generalization of Hilbert's inequality (see [43, 2]). Note that the Hilbert's inequality reads originally as follows:

Lemma 84

For all $\mathbf{x} = (x_j)_{j=1}^M \in \mathbb{C}^M$ *it holds Hilbert's inequality*

$$\Big|\sum_{\substack{j,\ell=1\\ j\neq\ell}}^{M} \frac{x_j\,\overline{x}_\ell}{j-\ell}\Big| \leq \pi\,\|\mathbf{x}\|_2^2 .$$

Proof: For arbitrary vector $\mathbf{x} = (x_j)_{j=1}^M \in \mathbb{C}^M$, we form the trigonometric polynomial

$$p(t) := \sum_{k=1}^{M} x_k \, \mathrm{e}^{\mathrm{i}\, k\, t}$$

such that

$$|p(t)|^2 = \sum_{k,\ell=1}^{M} x_k \, \overline{x}_\ell \, \mathrm{e}^{\mathrm{i}\,(k-\ell)\,t} \, .$$

Using

$$\frac{1}{2\pi\,\mathrm{i}} \int_0^{2\pi} (\pi - t)\, \mathrm{e}^{\mathrm{i}\,n\,t} \, \mathrm{d}t = \begin{cases} 0 & n = 0\,, \\ \frac{1}{n} & n \in \mathbb{Z} \setminus \{0\}\,, \end{cases}$$

we obtain

$$\frac{1}{2\pi\,\mathrm{i}} \int_0^{2\pi} (\pi - t)\, |p(t)|^2 \, \mathrm{d}t = \sum_{\substack{k,\ell=1 \\ k \neq \ell}}^{M} \frac{x_k \, \overline{x}_\ell}{k - \ell} \, .$$

Note that $|\pi - t| \le \pi$ for $t \in [0,\, 2\pi]$. From the triangle inequality and the Parseval equality in $L_2(\mathbb{T})$ it follows that

$$\frac{1}{2\pi}\,\Big|\int_0^{2\pi} (\pi - t)\,|p(t)|^2\,\mathrm{d}t\Big| \le \frac{1}{2}\,\int_0^{2\pi} |p(t)|^2\,\mathrm{d}t = \pi \sum_{j=1}^{M} |x_j|^2 = \pi\,\|\mathbf{x}\|_2^2\,.$$

The proof of generalized Hilbert's inequality applies the following result:

Lemma 85

For all $x \in \mathbb{R} \setminus \mathbb{Z}$ we have

$$\big(\sin(\pi x)\big)^{-2} + 2\,\Big|\frac{\cot(\pi x)}{\sin(\pi x)}\Big| \le \frac{3}{\pi^2\, d(x)^2}\,. \tag{108}$$

Proof: It suffices to show (108) for all $x \in (0, \frac{1}{2}]$. Substituting $t = \pi x \in (0, \frac{\pi}{2}]$, (108) means

$$3\,(\sin t)^2 \geq t^2\,(1 + 2\,\cos t)\,.$$

This inequality is equivalent to

$$3\,(\operatorname{sinc} t)^2 \geq 1 + 2\,\cos t\,, \quad t \in [0, \frac{\pi}{2}]\,,$$

which is true by the behaviors of the concave functions $3\,(\operatorname{sinc} t)^2$ and $1 + 2\,\cos t$ on the interval $[0, \frac{\pi}{2}]$. ■

Theorem 86

(see [44, Theorem 1]) Assume that the distinct values $\psi_j \in \mathbb{R}$, $j = 1, \ldots, M$, *fulfill the gap condition* (107) *with a constant* $\Delta > 0$.

Then generalized Hilbert's inequality

$$\Big| \sum_{\substack{j,\ell=1 \\ j \neq \ell}}^{M} \frac{x_j \, \overline{x}_\ell}{\sin\big(\pi\,(\psi_j - \psi_\ell)\big)} \Big| \le \frac{1}{\Delta}\, \|\mathbf{x}\|_2^2 \qquad (109)$$

holds for all $\mathbf{x} = (x_j)_{j=1}^{M} \in \mathbb{C}^M$.

Proof: 1. Setting

$$s_{j,\ell} := \begin{cases} \big[\sin\big(\pi\,(\psi_j - \psi_\ell)\big)\big]^{-1} & j \neq \ell\,, \\ 0 & j = \ell \end{cases}$$

for all $j,\ \ell = 1, \ldots, M$, we form the matrix $\mathbf{S} := -\mathrm{i}\,\big(s_{j,\ell}\big)_{j,\ell=1}^{M}$ which is Hermitian. Let the eigenvalues of $\mathbf{S}$ be arranged in increasing order $-\infty < \lambda_1 \leq \ldots \leq \lambda_M < \infty$. By the Rayleigh–Ritz theorem (see [20, pp. 234–235]) we have for all $\mathbf{x} \in \mathbb{C}^M$ with $\|\mathbf{x}\|_2 = 1$,

$$\lambda_1 \leq \mathbf{x}^{\mathrm{H}}\,\mathbf{S}\,\mathbf{x} \leq \lambda_M\,.$$

Suppose that $\lambda \in \mathbb{R}$ is such an eigenvalue of $\mathbf{S}$ with $|\lambda| = \max\{|\lambda_1|,\ |\lambda_M|\}$. Then we have the sharp inequality

$$|\mathbf{x}^{\mathrm{H}}\,\mathbf{S}\,\mathbf{x}| = \Big|\sum_{j,\ell=1}^{M} x_j\,\overline{x}_\ell\,s_{j,\ell}\Big| \leq |\lambda|$$

for all normed vectors $\mathbf{x} = (x_j)_{j=1}^{M} \in \mathbb{C}^M$. Now we show that $|\lambda| \leq \frac{1}{\Delta}$.

2. Related to the eigenvalue λ of $\mathbf{S}$, there exists a normed eigenvector $\mathbf{y} = (y_j)_{j=1}^M \in \mathbb{C}^M$ with $\mathbf{S}\,\mathbf{y} = \lambda\,\mathbf{y}$, i.e.,

$$\sum_{j=1}^{M} y_j\, s_{j,\ell} = \mathrm{i}\,\lambda\, y_\ell\,, \quad \ell = 1, \ldots, M\,. \tag{110}$$

Thus we have $\mathbf{y}^{\mathrm{H}}\,\mathbf{S}\,\mathbf{y} = \lambda\,\mathbf{y}^{\mathrm{H}}\,\mathbf{y} = \lambda$. Applying the Cauchy–Schwarz inequality, we estimate

$$\begin{aligned}
|\mathbf{y}^{\mathrm{H}}\,\mathbf{S}\,\mathbf{y}|^2 &= \Big|\sum_{j=1}^{M} y_j\,\Big(\sum_{\ell=1}^{M} \overline{y}_\ell\, s_{j,\ell}\Big|^2 \le \|\mathbf{y}\|_2^2\,\Big(\sum_{j=1}^{M} \Big|\sum_{\ell=1}^{M} \overline{y}_\ell\, s_{j,\ell}\Big|^2\Big) \\
&= \sum_{j=1}^{M} \Big|\sum_{\ell=1}^{M} \overline{y}_\ell\, s_{j,\ell}\Big|^2 = \sum_{j=1}^{M}\ \sum_{\ell,\,m=1}^{M} \overline{y}_\ell\, y_m\, s_{j,\ell}\, s_{j,m} \\
&= \sum_{\ell,\,m=1}^{M} \overline{y}_\ell\, y_m \sum_{j=1}^{M} s_{j,\ell}\, s_{j,m} = S_1 + S_2
\end{aligned}$$

with the partial sums

$$S_1 := \sum_{\ell=1}^{M} |y_\ell|^2 \sum_{j=1}^{M} s_{j,\ell}^2 \,, \quad S_2 := \sum_{\substack{\ell,\, m=1 \\ \ell \neq m}}^{M} \overline{y}_\ell \, y_m \sum_{j=1}^{M} s_{j,\ell} \, s_{j,m} \,.$$

3. For distinct α, $\beta \in \mathbb{R} \setminus (\pi \, \mathbb{Z})$ it holds

$$\frac{1}{(\sin \alpha)\,(\sin \beta)} = \frac{\cot \alpha - \cot \beta}{\sin(\beta - \alpha)}$$

such that for all indices with $j \neq \ell$, $j \neq m$, and $\ell \neq m$ we have

$$s_{j,\ell} \; s_{j,m} = s_{\ell,m} \left[\cot\big(\pi(\psi_j - \psi_\ell)\big) - \cot\big(\pi(\psi_j - \psi_m)\big) \right].$$

Now we split the sum S_2 in the following way

$$\begin{aligned} S_2 &= \sum_{\substack{\ell,\, m=1 \\ \ell \neq m}}^{M} \overline{y}_\ell \, y_m \sum_{\substack{j=1 \\ j \neq \ell\,,\, j \neq m}}^{M} s_{\ell,m} \left[\cot\big(\pi(\psi_j - \psi_\ell)\big) - \cot\big(\pi(\psi_j - \psi_m)\big) \right] \\ &= S_3 - S_4 + 2 \, \mathrm{Re}\, S_5 \end{aligned}$$

with following sums

$$S_3 := \sum_{\substack{\ell,\, m=1 \\ \ell \neq m}}^{M} \sum_{\substack{j=1 \\ j \neq \ell}}^{M} \overline{y}_\ell \, y_m \, s_{\ell,m} \cot\big(\pi(\psi_j - \psi_\ell)\big) \,,$$

$$S_4 := \sum_{\substack{\ell,\, m=1 \\ \ell \neq m}}^{M} \sum_{\substack{j=1 \\ j \neq m}}^{M} \overline{y}_\ell \, y_m \, s_{\ell,m} \cot\big(\pi(\psi_j - \psi_m)\big) \,,$$

$$S_5 := \sum_{\substack{j,\, \ell=1 \\ j \neq \ell}}^{M} \overline{y}_\ell \, y_j \, s_{j,\ell} \cot\big(\pi(\psi_j - \psi_\ell)\big) \,.$$

Note that $2\,\mathrm{Re}\, S_5$ is the correction sum, since S_3 contains the additional terms for $j = m$ and S_4 contains the additional terms for $j = \ell$.

4. First we show that $S_3 = S_4$. From (110) it follows that

$$S_3 = \sum_{\substack{\ell,\, j=1 \\ \ell \neq j}}^{M} \overline{y}_\ell \,\Big(\sum_{m=1}^{M} y_m\, s_{\ell,m} \Big)\, \cot\big(\pi(\psi_j - \psi_\ell)\big)$$

$$= -\mathrm{i}\, \lambda \sum_{\substack{\ell,\, j=1 \\ \ell \neq j}}^{M} |y_\ell|^2 \cot\big(\pi(\psi_j - \psi_\ell)\big)\,.$$

Analogously, we see that

$$S_4 = \sum_{\substack{j,\, m=1 \\ j \neq m}}^{M} y_m \,\Big(\sum_{\ell=1}^{M} \overline{y}_\ell\, s_{\ell,m} \Big)\, \cot\big(\pi(\psi_j - \psi_m)\big)$$

$$= -\mathrm{i}\, \lambda \sum_{\substack{j,\, m=1 \\ j \neq m}}^{M} |y_m|^2 \cot\big(\pi(\psi_j - \psi_m)\big)\,.$$

Hence we obtain the estimate

$$|\lambda|^2 = |\mathbf{y}^{\mathrm{H}}\, \mathbf{S}\, \mathbf{y}|^2 = S_1 + S_2 = S_1 + 2\, \mathrm{Re}\, S_5 \le S_1 + 2\, |S_5|\,.$$

Using $2\,|\overline{y}_\ell\, y_j| \le |y_\ell|^2 + |y_j|^2$, we estimate

$$2\,|S_5| \le \sum_{\substack{j,\,\ell=1\\ j\ne\ell}}^{M} 2\,|\overline{y}_\ell\, y_j|\,|s_{j,\ell}\,\cot\big(\pi(\psi_j-\psi_\ell)\big)|$$

$$\le 2\sum_{\substack{j,\,\ell=1\\ j\ne\ell}}^{M} |y_\ell|^2\,|s_{j,\ell}\,\cot\big(\pi(\psi_j-\psi_\ell)\big)|$$

such that

$$S_1 + 2\,|S_5| \le \sum_{\substack{j,\,\ell=1\\ j\ne\ell}}^{M} |y_\ell|^2\,\big[s_{j,\ell}^2 + 2\,|s_{j,\ell}\,\cot\big(\pi(\psi_j-\psi_{ell})\big)|\big]\,.$$

By Lemma 85 we obtain

$$S_1+2\,|S_5| \le \frac{3}{\pi^2}\sum_{\ell=1}^{M}|y_\ell|^2\sum_{\substack{j=1\\ j\ne\ell}}^{M} d(\psi_j-\psi_\ell)^{-2} = \frac{3}{\pi^2}\sum_{\substack{j,\,\ell=1\\ j\ne\ell}}^{M} d(\psi_j-\psi_\ell)^{-2}\,.$$

By assumption, the values ψ_j, $j = 1, \ldots, M$, are spaced from each other by at least Δ, so that

$$\sum_{\substack{j=1 \\ j \neq \ell}}^{M} d(\psi_j - \psi_\ell)^{-2} < 2 \sum_{k=1}^{\infty} (k\,\Delta)^{-2} = \frac{\pi^2}{3\,\Delta^2}$$

and hence

$$|\lambda|^2 = S_1 + S_2 \le S_1 + 2\,|S_5| < \frac{1}{\Delta^2}\,. \qquad \blacksquare$$

Under the natural assumption that the nodes $z_j = \mathrm{e}^{2\pi\,\mathrm{i}\,\psi_j}$, $j = 1, \ldots, M$, are well-separated on the unit circle, it can be shown that the rectangular Vandermonde matrix $\mathbf{V}_{P,M}(\mathbf{z})$ is well conditioned for sufficiently large $P > M$.

Theorem 87

(see [40, 43, 2]) *Let $P \in \mathbb{N}$ with $P > \max\{M, \frac{1}{\Delta}\}$ be given. Assume that the frequencies $\psi_j \in \mathbb{R}$, $j = 1, \ldots, M$, fulfill the gap condition* (107) *with a constant $\Delta > 0$.*
Then for all $\mathbf{c} \in \mathbb{C}^M$, the rectangular Vandermonde matrix $\mathbf{V}_{P,M}(\mathbf{z})$ with $\mathbf{z} = (z_j)_{j=1}^M$ fulfills the inequalities

$$\left(P - \frac{1}{\Delta}\right) \|\mathbf{c}\|_2^2 \le \|\mathbf{V}_{P,M}(\mathbf{z})\,\mathbf{c}\|_2^2 \le \left(P + \frac{1}{\Delta}\right) \|\mathbf{c}\|_2^2 . \tag{111}$$

Further the rectangular Vandermonde matrix $\mathbf{V}_{P,M}(\mathbf{z})$ has a uniformly bounded spectral norm condition number

$$\mathrm{cond}_2\, \mathbf{V}_{P,M}(\mathbf{z}) \le \sqrt{\frac{P\,\Delta + 1}{P\,\Delta - 1}} .$$

Proof: 1. Simple computation shows that

$$\|\mathbf{V}_{P,M}(\mathbf{z})\,\mathbf{c}\|_2^2 = \sum_{k=0}^{P-1} \Big| \sum_{j=1}^{M} c_j\, z_j^k \Big|^2 = \sum_{k=0}^{P-1} \sum_{j,\,\ell=1}^{M} c_j\, \overline{c}_\ell\, \mathrm{e}^{2\pi\,\mathrm{i}\,(\psi_j-\psi_\ell)\,k}$$

$$= \sum_{k=0}^{P-1} \Big(\sum_{j=1}^{M} |c_j|^2 + \sum_{\substack{j,\,\ell=1\\ j\neq\ell}}^{M} c_j\, \overline{c}_\ell\, \mathrm{e}^{2\pi\,\mathrm{i}\,(\psi_j-\psi_\ell)\,k} \Big)$$

$$= P\,\|\mathbf{c}\|_2^2 + \sum_{\substack{j,\,\ell=1\\ j\neq\ell}}^{M} c_j\, \overline{c}_\ell \Big(\sum_{k=0}^{P-1} \mathrm{e}^{2\pi\,\mathrm{i}\,(\psi_j-\psi_\ell)\,k} \Big)\,.$$

Determining the sum

$$\sum_{k=0}^{P-1} \mathrm{e}^{2\pi\,\mathrm{i}\,(\psi_j-\psi_\ell)\,k} = \frac{1-\mathrm{e}^{2\pi\,\mathrm{i}\,(\psi_j-\psi_\ell)\,P}}{1-\mathrm{e}^{2\pi\,\mathrm{i}\,(\psi_j-\psi_\ell)}}$$

$$= \frac{1-\mathrm{e}^{2\pi\,\mathrm{i}\,(\psi_j-\psi_\ell)\,P}}{2\mathrm{i}\,\mathrm{e}^{\pi\,\mathrm{i}\,(\psi_j-\psi_\ell)}\,\sin\big(\pi\,(\psi_j-\psi_\ell)\big)} = -\frac{\mathrm{e}^{-\pi\,\mathrm{i}\,(\psi_j-\psi_\ell)} - \mathrm{e}^{\pi\,\mathrm{i}\,(\psi_j-\psi_\ell)\,(2P-1)}}{2\mathrm{i}\,\sin\big(\pi\,(\psi_j-\psi_\ell)\big)}\,,$$

we obtain

$$\|\mathbf{V}_{P,M}(\mathbf{z})\,\mathbf{c}\|_2^2 = P\,\|\mathbf{c}\|_2^2 - \Sigma_1 + \Sigma_2 \tag{112}$$

with the sums

$$\Sigma_1 := \sum_{\substack{j,\,\ell=1\\ j\neq\ell}}^{M} \frac{c_j\,\overline{c}_\ell\,\mathrm{e}^{-\pi\,\mathrm{i}\,(\psi_j-\psi_\ell)}}{2\mathrm{i}\,\sin\big(\pi\,(\psi_j-\psi_\ell)\big)}\,, \quad \Sigma_2 := \sum_{\substack{j,\,\ell=1\\ j\neq\ell}}^{M} \frac{c_j\,\overline{c}_\ell\,\mathrm{e}^{\pi\,\mathrm{i}\,(\psi_j-\psi_\ell)\,(2P-1)}}{2\mathrm{i}\,\sin\big(\pi\,(\psi_j-\psi_\ell)\big)}\,.$$

The nodes $z_j = \mathrm{e}^{2\pi\,\mathrm{i}\,\psi_j}$, $j = 1, \ldots, M$, are distinct, since we have (107) by assumption. Applying generalized Hilbert's inequality (109) first with $x_k := c_k\,\mathrm{e}^{-\pi\mathrm{i}\,\psi_k}$, $k = 1, \ldots, M$, yields

$$|\Sigma_1| \le \frac{1}{2\,\Delta}\,\sum_{k=1}^{M} \big|c_k\,\mathrm{e}^{-\pi\mathrm{i}\,\psi_k}\big|^2 = \frac{1}{2\,\Delta}\,\|\mathbf{c}\|_2^2 \tag{113}$$

and then with $x_k := c_k\,\mathrm{e}^{\pi\mathrm{i}\,\psi_k\,(2P-1)}$, $k = 1, \ldots, M$, results in

$$|\Sigma_2| \le \frac{1}{2\,\Delta}\,\sum_{k=1}^{M} \big|c_k\,\mathrm{e}^{\pi\mathrm{i}\,\psi_k\,(2P-1)}\big|^2 = \frac{1}{2\,\Delta}\,\|\mathbf{c}\|_2^2 \tag{114}$$

From (112) – (114) it follows the assertion (111) by triangle inequality.

2. Let $\mu_1 \ge \ldots \ge \mu_M > 0$ be the ordered eigenvalues of $\mathbf{V}_{P,M}(\mathbf{z})^{\mathrm{H}}\,\mathbf{V}_{P,M}(\mathbf{z}) \in \mathbf{C}^{M\times M}$. Using the Raleigh–Ritz theorem (see [20, pp. 234–235]) and (111), we obtain that for all $\mathbf{c} \in \mathbb{C}^M$

$$\left(P-\frac{1}{\Delta}\right)\|\mathbf{c}\|_2^2 \le \mu_M\,\|\mathbf{c}\|_2^2 \le \|\mathbf{V}_{P,M}(\mathbf{z})\,\mathbf{c}\|_2^2 \le \mu_1\,\|\mathbf{c}\|_2^2 \le \left(P-\frac{1}{\Delta}\right)\|\mathbf{x}\|_2^2$$

and hence

$$0 < P - \frac{1}{\Delta} \le \lambda_M \le \lambda_1 \le P + \frac{1}{\Delta} < \infty\,. \qquad (115)$$

Thus $\mathbf{V}_{P,M}(\mathbf{z})^{\mathrm{H}}\,\mathbf{V}_{P,M}(\mathbf{z})$ is positive definite and

$$\mathrm{cond}_2\,\mathbf{V}_{P,M}(\mathbf{z}) = \sqrt{\frac{\mu_1}{\mu_M}} \le \sqrt{\frac{P\,\Delta+1}{P\,\Delta-1}}\,. \qquad \blacksquare$$

The inequalities (111) can be interpreted as discrete versions of the Ingham inequalities (99). Now the exponentials $\mathrm{e}^{2\pi\,\mathrm{i}\,\psi_j\,\cdot}$ are replaced by their discretizations

$$\mathbf{e}_P(\psi_j) = \big(\mathrm{e}^{2\pi\,\mathrm{i}\,\psi_j\,k}\big)_{k=0}^{P-1}\,, \quad j = 1, \ldots, M\,,$$

with sufficiently large integer $P > \max\{M, \frac{1}{\Delta}\}$. Thus the rectangular Vandermonde matrix can be written as

$$\mathbf{V}_{P,M}(\mathbf{z}) = \big(\mathbf{e}_P(\psi_1)\,|\,\mathbf{e}_P(\psi_2)\,|\,\ldots\,|\,\mathbf{e}_P(\psi_M)\big)$$

with $\mathbf{z} = (z_j)_{j=1}^M$, where $z_j = \mathrm{e}^{2\pi\,\mathrm{i}\,\psi_j}$, $j = 1, \ldots, M$, are distinct nodes on the unit circle. Then (111) provides the *discrete Ingham inequalities*

$$\big(P - \frac{1}{\Delta}\big)\,\|\mathbf{c}\|_2^2 \le \|\sum_{j=1}^{M} c_j\,\mathbf{e}_P(\varphi_j)\|_2^2 \le \big(P + \frac{1}{\Delta}\big)\,\|\mathbf{c}\|_2^2 \qquad (116)$$

for all $\mathbf{c} = (c_j)_{j=1}^M \in \mathbb{C}^M$. In other words, (116) means that the vectors $\mathbf{e}_P(\varphi_j)$, $j = 1, \ldots, M$, are Riesz stable too.

Corollary 88

Under the assumptions of Theorem 87, the inequalities

$$\big(P - \frac{1}{\Delta}\big)\,\|\mathbf{d}\|_2^2 \le \|\mathbf{V}_{P,M}(\mathbf{z})^\top\,\mathbf{d}\|_2^2 \le \big(P + \frac{1}{\Delta}\big)\,\|\mathbf{d}\|_2^2 \qquad (117)$$

hold for all $\mathbf{d} \in \mathbb{C}^P$*.*

Proof: The matrices $\mathbf{V}_{P,M}(\mathbf{z})$ and $\mathbf{V}_{P,M}(\mathbf{z})^\top$ possess the same singular values μ_j, $j = 1, \ldots, M$. By the Rayleigh–Ritz theorem we obtain that

$$\lambda_M\,\|\mathbf{d}\|_2^2 \le \|\mathbf{V}_{P,M}(\mathbf{z})^\top\,\mathbf{d}\|_2^2 \le \lambda_1\,\|\mathbf{d}\|_2^2$$

for all $\mathbf{d} \in \mathbb{C}^P$. Applying (115), we obtain the inequalities (117).
■

Remark 89

In [4, 2], the authors derive bounds on the extremal singular values and the condition number of the rectangular Vandermonde matrix $\mathbf{V}_{P,M}(\mathbf{z})$ *with* $P \geq M$ *and* $\mathbf{z} = (z_j)_{j=1}^{M} \in \mathbb{C}^M$*, where the nodes are in the unit disk, i.e.,* $|z_j| \leq 1$ *for* $j = 1, \ldots, M$. $\square$

By the Vandermonde decomposition of the Hankel matrix $\mathbf{H}_{L,N-L+1}$ we obtain that

$$\mathbf{H}_{L,N-L+1} = \mathbf{V}_{L,M}(\mathbf{z})\,(\mathrm{diag}\,\mathbf{c})\,(\mathbf{V}_{N-L+1,M}(\mathbf{z}))^{\top}\,. \tag{118}$$

Under mild conditions, the Hankel matrix $\mathbf{H}_{L,N-L+1}$ of noiseless data is well-conditioned too.

Theorem 90

Let $L,\, N \in \mathbb{N}$ *with* $M \le L \le N - M + 1$ *and* $\min\{L,\, N - L + 1\} > \frac{1}{\Delta}$ *be given. Assume that the frequencies* $\psi_j \in \mathbb{R}$, $j = 1, \ldots, M$, *are well-separated at least by a constant* $\Delta > 0$ *and that the nonzero coefficients* c_j, $j = 1, \ldots, M$, *of the exponential sum* (70) *fulfill the condition*

$$0 < \gamma_1 \le |c_j| \le \gamma_2 < \infty\,, \quad j = 1, \ldots, M\,. \tag{119}$$

Then for all $\mathbf{y} \in \mathbb{C}^{N-L+1}$

$$\gamma_1^2\, \alpha_1(L,\, N,\, \Delta)\, \|\mathbf{y}\|_2^2 \le \|\mathbf{H}_{L,N-L+1}\, \mathbf{y}\|_2^2 \le \gamma_2^2\, \alpha_2(L,\, N,\, \Delta)\, \|\mathbf{y}\|_2^2\,. \tag{120}$$

with

$$\alpha_1(L,\, N,\, \Delta) := \big(L - \frac{1}{\Delta}\big)\, \big(N - L + 1 - \frac{1}{\Delta}\big)\,,$$
$$\alpha_2(L,\, N,\, \Delta) := \big(L + \frac{1}{\Delta}\big)\, \big(N - L + 1 + \frac{1}{\Delta}\big)\,.$$

Theorem 90 (continue)

Further, the lowest (nonzero) respectively largest singular value of $\mathbf{H}_{L,N-L+1}$ can be estimated by

$$0 < \gamma_1 \sqrt{\alpha_1(L,\, N,\, \Delta)} \le \sigma_M \le \sigma_1 \le \gamma_2 \sqrt{\alpha_2(L,\, N,\, \Delta)}\,. \tag{121}$$

The spectral norm condition number of $\mathbf{H}_{L,N-L+1}$ is bounded by

$$\mathrm{cond}_2\, \mathbf{H}_{L,N-L+1} \le \frac{\gamma_2}{\gamma_1} \sqrt{\frac{\alpha_2(L,\, N,\, \Delta)}{\alpha_1(L,\, N,\, \Delta)}}\,. \tag{122}$$

Proof: By the Vandermonde decomposition (118) of the Hankel matrix $\mathbf{H}_{L,N-L+1}$, we obtain that for all $\mathbf{y} \in \mathbb{C}^{N-L+1}$

$$\|\mathbf{H}_{L,N-L+1}\,\mathbf{y}\|_2^2 = \|\mathbf{V}_{L,M}(\mathbf{z})\,(\mathrm{diag}\,\mathbf{c})\,\mathbf{V}_{N-L+1,M}(\mathbf{z})^\top\,\mathbf{y}\|_2^2\,.$$

By the estimates (111) and the assumption (119), it follows that

$$\begin{aligned}
&\gamma_1^2\,\big(L-\frac{1}{\Delta}\big)\,\|\mathbf{V}_{N-L+1,M}(\mathbf{z})^\top\,\mathbf{y}\|_2^2 \\
&\leq \|\mathbf{H}_{L,N-L+1}\,\mathbf{y}\|_2^2 \\
&\leq \gamma_2^2\,\big(L+\frac{1}{\Delta}\big)\,\|\mathbf{V}_{N-L+1,M}(\mathbf{z})^\top\,\mathbf{y}\|_2^2\,.
\end{aligned}$$

Using the inequalities (117), we obtain the estimates (120). Finally, the estimates of the extremal singular values and the spectral norm condition number of $\mathbf{H}_{L,N-L+1}$ arise from (120) and the Rayleigh–Ritz theorem. ■

Remark 91

For fixed N, the positive singular values as well as the spectral norm condition number of the Hankel matrix $\mathbf{H}_{L,N-L+1}$ depend strongly on $L \in \{M, \ldots, N-M+1\}$. A good criterion for the choice of optimal window length L is to maximize the lowest positive singular value σ_M of $\mathbf{H}_{L,N-L+1}$. It was shown in [52, Lemma 3.1 and Remark 3.3] that the squared singular values increase almost monotonously for $L = M, \ldots, \lceil \frac{N}{2} \rceil$ and decrease almost monotonously for $L = \lceil \frac{N}{2} \rceil, \ldots, N-M+1$. Note that the lower bound (121) of the lowest positive singular value σ_M is maximal for $L \approx \frac{N}{2}$. Further the upper bound (122) of the spectral norm condition number of the exact Hankel matrix $\mathbf{H}_{L,N-L+1}$ is minimal for $L \approx \frac{N}{2}$. Therefore we prefer to choose $L \approx \frac{N}{2}$ as optimal window length. Thus we can ensure that $\sigma_M > 0$ is not too small. This property is decisively for the correct detection of the order M in the first step of the MUSIC Algorithm and ESPRIT Algorithm. □

[1] K. Atkinson and W. Han.
Theoretical Numerical Analysis. A Functional Analysis Framework.
Springer-Verlag, Dordrecht, 2009.

[2] C. Aubel and H. Bölcskei.
Vandermonde matrices with nodes in the unit disk and the large sieve.
Appl. Comput. Harmon. Anal., 2017.

[3] G. Baszenski and F.-J. Delvos.
A discrete Fourier transform scheme for Boolean sums of trigonometric operators.
In *Multivariate Approximation Theory IV*, pages 15–24. Birkhäuser, Basel, 1989.

[4] F. S. V. Bazán.
Conditioning of rectangular Vandermonde matrices with nodes in the unit disk.
SIAM J. Matrix Anal. Appl., 21:679–693, 2000.

[5] F. S. V. Bazán and P. L. Toint.
Error analysis of signal zeros from a related companion matrix eigenvalue problem.
Appl. Math. Lett., 14(7):859–866, 2001.

[6] H.-J. Bungartz and M. Griebel.
A note on the complexity of solving Poisson's equation for spaces of bounded mixed derivatives.
J. Complexity, 15(2):167–199, 1999.

[7] H.-J. Bungartz and M. Griebel.
Sparse grids.
Acta Numer., 13:147–269, 2004.

[8] G. Byrenheid, L. Kämmerer, T. Ullrich, and T. Volkmer.
Tight error bounds for rank-1 lattice sampling in spaces of hybrid mixed smoothness.
Numer. Math., 136(4):993–1034, 2017.

[9] O. Christensen.
An Introduction to Frames and Riesz Bases.

Birkhäuser, Boston, 2003.

[10] D. Dung, V. N. Temlyakov, and T. Ullrich.
Hyperbolic Cross Approximation.
Advanced Courses in Mathematics. CRM Barcelona. Birkhäuser/Springer, 2017.

[11] A. C. Fannjiang.
The MUSIC algorithm for sparse objects: a compressed sensing analysis.
Inverse Problems, 27(3):035013, 2011.

[12] G. H. Golub, P. Milanfar, and J. Varah.
A stable numerical method for inverting shape from moments.
SIAM J. Sci. Comput., 21(4):1222–1243, 1999/2000.

[13] G. H. Golub and C. F. Van Loan.
Matrix Computations.
Johns Hopkins University Press, Baltimore, third edition, 1996.

[14] N. Golyandina, V. Nekrutkin, and A. Zhigljavsky.
Analysis of Time Series Structure. SSA and Related Techniques.
Chapman & Hall/CRC, Boca Raton, 2001.

[15] N. Golyandina and A. Zhigljavsky.
Singular Spectrum Analysis for Time Series.
Springer-Verlag, Heidelberg, 2013.

[16] L. Grafakos.
Classical Fourier Analysis.
Springer-Verlag, New York, second edition, 2008.

[17] K. Gröchenig.
An uncertainty principle related to the Poisson summation formula.
Studia Math., 121(1):87–104, 1996.

[18] G. Heinig and K. Rost.
Algebraic Methods for Toeplitz-like Matrices and Operators.
Akademie-Verlag, Berlin, 1984.

[19] N. J. Higham.
Accuracy and Stability of Numerical Algorithms.
SIAM, Philadelphia, second edition, 2002.

[20] R. A. Horn and C. R. Johnson.
Matrix Analysis.
Cambridge University Press, Cambridge, second edition, 2013.

[21] Y. Hua and T. K. Sarkar.
Matrix pencil method for estimating parameters of exponentially damped/undamped sinusoids in noise.
IEEE Trans. Acoust. Speech Signal Process., 38(5):814–824, 1990.

[22] A. E. Ingham.
Some trigonometrical inequalities with applications to the theory of series.
Math. Z., 41(1):367–379, 1936.

[23] L. Kämmerer.

Reconstructing hyperbolic cross trigonometric polynomials by sampling along rank-1 lattices.
SIAM J. Numer. Anal., 51(5):2773–2796, 2013.

[24] L. Kämmerer.
High Dimensional Fast Fourier Transform Based on Rank-1 Lattice Sampling.
Dissertation. Universitätsverlag Chemnitz, 2014.

[25] L. Kämmerer.
Reconstructing multivariate trigonometric polynomials from samples along rank-1 lattices.
In *Approximation Theory XIV: San Antonio 2013*, pages 255–271. Springer, Cham, 2014.

[26] L. Kämmerer.
Constructing spatial discretizations for sparse multivariate trigonometric polynomials that allow for a fast discrete Fourier transform.
ArXiv e-prints 1703.07230, 2017.

[27] L. Kämmerer.
Multiple rank-1 lattices as sampling schemes for multivariate trigonometric polynomials.
J. Fourier Anal. Appl., 24(1):17–44, 2018.

[28] L. Kämmerer, S. Kunis, I. Melzer, D. Potts, and T. Volkmer.
Computational methods for the Fourier analysis of sparse high-dimensional Functions.
In *Extraction of Quantifiable Information from Complex Systems*, pages 347–363. Springer, Cham, 2014.

[29] L. Kämmerer, S. Kunis, and D. Potts.
Interpolation lattices for hyperbolic cross trigonometric polynomials.
J. Complexity, 28(1):76–92, 2012.

[30] L. Kämmerer, D. Potts, and T. Volkmer.
Approximation of multivariate periodic functions by trigonometric polynomials based on rank-1 lattice sampling.
J. Complexity, 31(4):543–576, 2015.

[31] L. Kämmerer, D. Potts, and T. Volkmer.
Approximation of multivariate periodic functions by trigonometric polynomials based on sampling along rank-1 lattice with generating vector of Korobov form.
J. Complexity, 31(3):424–456, 2015.

[32] L. Kämmerer, D. Potts, and T. Volkmer.
High-dimensional sparse FFT based on sampling along multiple rank-1 lattices.
ArXiv e-prints 1711.05152, 2017.

[33] A. Kirsch.
The MUSIC algorithm and the factorization method in inverse scattering theory for inhomogeneous media.
Inverse Problems, 18(4):1025–1040, 2002.

[34] V. Komornik and P. Loreti.
Fourier Series in Control Theory.
Springer-Verlag, New York, 2005.

[35] N. M. Korobov.

Teoretiko-chislovye metody v priblizhënnom analize.
Moscow, second edition, 2004.

[36] T. Kühn, W. Sickel, and T. Ullrich.
Approximation numbers of Sobolev embeddings - sharp constants and tractability.
J. Complexity, 30(2):95–116, 2014.

[37] S. Kunis, H. M. Möller, T. Peter, and U. von der Ohe.
Prony's method under an almost sharp multivariate Ingham inequality.
J. Fourier Anal. Appl., 2017.

[38] F. Y. Kuo, I. H. Sloan, and H. Woźniakowski.
Lattice rule algorithms for multivariate approximation in the average case setting.
J. Complexity, 24(2):283–323, 2008.

[39] D. Li and F. J. Hickernell.
Trigonometric spectral collocation methods on lattices.

In *Recent Advances in Scientific Computing and Partial Differential Equations*, pages 121–132. Amer. Math. Soc., Providence, 2003.

[40] W. Liao and A. Fannjiang.
MUSIC for single-snapshot spectral estimation: Stability and super-resolution.
Appl. Comput. Harmon. Anal., 40(1):33–67, 2016.

[41] D. G. Manolakis, V. K. Ingle, and S. M. Kogon.
Statistical and Adaptive Signal Processing.
McGraw-Hill, Boston, 2005.

[42] I. Markovsky.
Structured low-rank approximation and its applications.
Automatica J. IFAC, 44(4):891–909, 2008.

[43] A. Moitra.
The threshold for super-resolution via extremal functions.
preprint, massachusetts institute of technology, cambridge.
2014.

[44] H. L. Montgomery and R. C. Vaughan.
Hilbert's inequality.
J. London Math. Soc., 8:73–82, 1974.

[45] H. Munthe-Kaas and T. Sørevik.
Multidimensional pseudo-spectral methods on lattice grids.
Appl. Numer. Math., 62(3):155–165, 2012.

[46] H. Niederreiter.
Quasi-Monte Carlo methods and pseudo-random numbers.
Bull. Amer. Math. Soc., 84(6):957–1041, 1978.

[47] V. Pereyra and G. Scherer.
Exponential Data Fitting and its Applications.
Bentham Sci. Publ., Sharjah, 2010.

[48] T. Peter, D. Potts, and M. Tasche.
Nonlinear approximation by sums of exponentials and translates.
SIAM J. Sci. Comput., 33:1920–1947, 2011.

[49] D. Potts and M. Tasche.
Nonlinear approximation by sums of nonincreasing exponentials.
Appl. Anal., 90(3–4):609–626, 2011.

[50] D. Potts and M. Tasche.
Parameter estimation for nonincreasing exponential sums by Prony-like methods.
Linear Algebra Appl., 439(4):1024–1039, 2013.

[51] D. Potts and M. Tasche.
Sparse polynomial interpolation in Chebyshev bases.
Linear Algebra Appl., 441:61–87, 2014.

[52] D. Potts and M. Tasche.
Error estimates for the ESPRIT algorithm.
In *Large Truncated Toeplitz Matrices, Toeplitz Operators, and Related Topics*, pages 621–648. Birkhäuser/Springer, Cham, 2017.

[53] R. Prony.

Essai éxperimental et analytique: sur les lois de la dilatabilité des fluides élastiques et sur celles de la force expansive de la vapeur de l'eau et de la vapeur de l'alkool, à différentes températures.
J. Ecole Polytechnique, 1:24–76, 1795.

[54] R. Roy and T. Kailath.
ESPRIT—estimation of signal parameters via rotational invariance techniques.
In *Signal Processing, Part II*, IMA Vol. Math. Appl. 23, pages 369–411. Springer, New York, 1990.

[55] T. K. Sarkar and O. Pereira.
Using the matrix pencil method to estimate the parameters of a sum of complex exponentials.
IEEE Antennas and Propag., 37:48–55, 1995.

[56] H.-J. Schmeisser and H. Triebel.
Topics in Fourier Analysis and Function Spaces.
Akademische Verlagsgesellschaft Geest & Portig, Leipzig, 1987.

[57] R. Schmidt.
Multiple emitter location and signal parameter estimation.
IEEE Trans. Antennas and Propagation, 34:276–280, 1986.

[58] I. H. Sloan and S. Joe.
Lattice Methods for Multiple Integration.
Clarendon Press, Oxford University Press, New York, 1994.

[59] I. H. Sloan and P. J. Kachoyan.
Lattice methods for multiple integration: theory, error analysis and examples.
SIAM J. Numer. Anal., 24(1):116–128, 1987.

[60] I. H. Sloan and A. V. Reztsov.
Component-by-component construction of good lattice rules.
Math. Comp., 71(237):263–273, 2002.

[61] E. M. Stein and G. Weiss.
Introduction to Fourier Analysis on Euclidean Spaces.
Princeton University Press, Princeton, 1971.

[62] V. N. Temlyakov.
Reconstruction of periodic functions of several variables from the values at the nodes of number-theoretic nets (in Russian).
Anal. Math., 12(4):287–305, 1986.

[63] V. N. Temlyakov.
Approximation of Periodic Functions.
Nova Science Publishers, Commack, 1993.

[64] H. Triebel.
Höhere Analysis.
Deutscher Verlag der Wissenschaften, Berlin, 1972.

[65] M. Vetterli, P. Marziliano, and T. Blu.
Sampling signals with finite rate of innovation.
IEEE Trans. Signal Process., 50(6):1417–1428, 2002.

[66] T. Volkmer.
Multivariate Approximation and High-Dimensional Sparse FFT Based on Rank-1 Lattice Sampling.
Dissertation. Universitätsverlag Chemnitz, 2017.

[67] L. Weiss and R. N. McDonough.
Prony's method, Z-transforms, and Padé approximation.
SIAM Rev., 5:145–149, 1963.

[68] F. Weisz.
Summability of multi-dimensional Fourier series and Hardy spaces.
Kluwer Academic, Dordrecht, 2002.

[69] F. Weisz.
Summability of multi-dimensional trigonometric Fourier series.
Surv. Approx. Theory, 7:1–179, 2012.

[70] R. M. Young.
An Introduction to Nonharmonic Fourier Series.
Academic Press, San Diego, revised first edition, 2001.

[71] H. Yserentant.
Regularity and Approximability of Electronic Wave Functions.
Springer-Verlag, Berlin, 2010.

[72] X. Zeng, K.-T. Leung, and F. J. Hickernell.
Error analysis of splines for periodic problems using lattice designs.
In *Monte Carlo and Quasi-Monte Carlo Methods 2004*, pages 501–514. Springer-Verlag, Berlin, 2006.