Numerical Approximation of Generalized Functions: Aliasing, the Gibbs Phenomenon and a Numerical Uncertainty Principle

Patrick Guidotti

To the memory of Günter Lumer

Abstract. A general recipe for high order approximation of generalized functions is introduced which is based on the use of L_2-orthonormal bases consisting of C^∞-functions and the appropriate choice of a discrete quadrature rule. Particular attention is paid to maintaining the distinction between pointwise functions (that is, which can be evaluated pointwise) and linear functionals defined on spaces of smooth functions (that is, distributions). It turns out that "best" pointwise approximation and "best" distributional approximation cannot be achieved simultaneously. This entails the validity of a kind of "numerical uncertainty principle": The local value of a function and its action as a linear functional on test functions cannot be known at the same time with high accuracy, in general.

In spite of this, high order accurate pointwise approximations can be obtained in special cases from a high accuracy distributional approximation when more information is available concerning the function which is to be approximated. A few special cases with application to PDEs are considered in detail.

Mathematics Subject Classification (2000). Primary 99Z99; Secondary 00A00.

Keywords. Generalized functions, approximation, Gibbs phenomenon.

1. Introduction

Let $\Omega \subset \mathbb{R}^n$ be open. Starting with the spaces

$$\mathcal{D}(\Omega) = C_c^\infty(\Omega) = \{\varphi \in C^\infty(\Omega) \,|\, \operatorname{supp}(\varphi) \text{ is compact}\} \text{ and } C^\infty(\Omega)$$

of test functions, generalized functions are introduced as

$$\mathcal{D}'(\Omega) = \mathcal{L}\big(\mathcal{D}(\Omega), \mathbb{K}\big) = \{u : \mathcal{D}(\Omega) \to \mathbb{K} \,|\, u \text{ is linear and continuous}\}$$
$$\mathcal{E}'(\Omega) = \mathcal{L}\big(\mathrm{C}^\infty(\Omega), \mathbb{K}\big) = \{u : \mathrm{C}^\infty(\Omega) \to \mathbb{K} \,|\, u \text{ is linear and continuous}\}$$

continuous linear functionals on the former. If $(e_n)_{n \in \mathbb{N}}$ is an orthonormal basis of $\mathrm{L}_2(\Omega)$, then one has

$$\big[\varphi = \sum_{n \in \mathbb{N}} \varphi_n e_n \mapsto (\varphi_n)_{n \in \mathbb{N}}\big] \in \mathcal{GL}\big(\mathrm{L}_2(\Omega), l_2(\mathbb{N})\big)$$

is an isometry. In particular one has Bessel's equality

$$\|\varphi\|_2 = \Big(\sum_{n \in \mathbb{N}} \varphi_n^2\Big)^{1/2} \tag{1.1}$$

and Parseval's identity

$$\int_\Omega \varphi\psi \, dx = \sum_{n \in \mathbb{N}} \varphi_n \psi_n. \tag{1.2}$$

Unfortunately, even for $\varphi \in \mathcal{D}(\Omega)$ the "Fourier series"

$$\sum_{n=1}^m \varphi_n e_n \xrightarrow[m \to \infty]{} \varphi \tag{1.3}$$

merely converges in the $\mathrm{L}_2(\Omega)$-topology, in general. Whenever it converges in the topology of $\mathrm{C}^\infty(\Omega)$ for all $\varphi \in \mathcal{D}(\Omega)$, then one also has Parseval's identity

$$\langle u, \varphi \rangle = \langle u, \sum_{n=1}^m \varphi_n e_n \rangle = \sum_{n=1}^m \varphi_n \langle u, e_n \rangle =: \sum_{n=1}^m \varphi_n u_n \tag{1.4}$$

for pairs $(u, \varphi) \in \mathcal{E}'(\Omega) \times \mathcal{D}(\Omega)$.

Remarks 1.1. **(a)** The fact that the basis functions are in most cases fully supported in the domain Ω entails that the convergence in (1.3) can only occur in $\mathrm{C}^\infty(\Omega)$ even though $\varphi \in \mathcal{D}(\Omega)$. For the same reason only compactly supported distributions $u \in \mathcal{E}'(\Omega)$ can be approximated by by their coefficient series

$$\sum_{n=1}^m u_n e_n.$$

(b) In numerical analysis one is often confronted with the fact the approximations which are high order in the interior of the domain deteriorate as the boundary is approached. This is related to the previous remark.

The approximation procedure introduced in Section 3 for $u \in \mathcal{E}'(\Omega)$ is obtained bearing (1.4) in mind and is based on the choice of a convergent quadrature rule (x^m, q^m) given by

$$x^m = (x_j^m)_{j=1,\,\ldots,g(m)}, \ q^m = (q_j^m)_{j=1,\,\ldots,g(m)}.$$

In particular

$$q^m \cdot \varphi^m = \sum_{j=1}^{g(m)} \varphi_j^m q_j^m \xrightarrow[m \to \infty]{} \int_\Omega \varphi(x)\,dx \,,\ \varphi \in \mathcal{D}(\Omega)$$

for $\varphi^m := \big(\varphi(x_j^m)\big)_{j=1,\,\dots,g(m)}$. Given $(e_n)_{n \in \mathbb{N}}$ it is desirable to choose the quadrature rule in such a way that the underlying orthogonality structure is preserved, that is,

$$e_j^m \cdot e_k^m = \delta_{jk} \,,\ 1 \le j, k \le g(m)$$

and $g(m) = m$. In the case of classical Jacobi polynomials this would correspond to working with Gauß quadrature.

Consider now the problem of approximating a general element $u \in \mathcal{E}'(\Omega)$. Only the information loss associated to analytically exact projections is considered here. For smooth or piecewise smooth functions one can choose to project either in *physical space* with $P^{\mathcal{P}}$ or *Fourier space* with $P^{\mathcal{F}}$. The projections are defined through

$$P^{\mathcal{P}} u := u^m = \big(u(x_j^m)\big)_{j=1,\,\dots,m} \tag{1.5}$$

$$P^{\mathcal{F}} u := \sum_{n=1}^m u_n e_n^m = \Big(\sum_{n=1}^m u_n e_n(x_j^m)\Big)_{j=1,\,\dots,m}. \tag{1.6}$$

The two projections have ranges of the same dimension. In the case of general distributions only the second projection can be used. A particularly interesting situation is that of piecewise smooth functions where one has the choice of using either projection. If one chooses $P^{\mathcal{P}}$ then one is confronted with the problem of aliasing, whereas the Gibbs phenomenon imposes limitations on the use of $P^{\mathcal{F}}$. The two well-known effects are dual to each other. The Gibbs phenomenon, or the appearance of oscillations at points of non-smoothness, is in reality the manifestation of the fact that the approximation defined by $P^{\mathcal{F}}$ is spectrally accurate in the sense of distributions (as will be shown in Section 4).

In any case it cannot be expected to obtain approximations which are of a high pointwise and distributional degree of accuracy simultaneously. The dual effects of aliasing and the Gibbs phenomenon make this an impossible goal to achieve. This is what is labeled *numerical uncertainty principle* in this paper.

The rest of the paper is organized as follows. In the section 2 a brief review of the basic concepts and facts about Schwartz' theory of distributions is given. Section 3 is devoted to introducing the fundamental concept of approximation family for a distribution. It will play a central role in the rest of the paper. A few simple concrete examples will be given. In section 4 the important question of convergence of general approximation families is addressed. The rest of the paper is devoted to applications and examples. Section 5.2 deals with the Gibbs phenomenon and its resolution. In Section 6 a simple illustrative PDE example is considered.

2. Basics of the theory of distributions

A very brief summary of the main basic concepts and results of Schwartz' theory of distributions is needed to set the stage. The reader familiar with the theory can, however, skip this section and move on to the its numerical implications presented in the rest of the paper. A central role in the theory of generalized functions is played by the underlying duality structure. One of its advantages is that it allows to define and carry out many important (linear) operations for distributions at the level of test functions. At the discrete level it might seems pedantic and superfluous to want to keep a distinction between classical functions and distributions. It, however, turns out that many seemingly bad pointwise approximations are indeed good distributional ones. Here only a sketchy and incomplete overview of the theory of distributions is given. The interested reader is therefore referred to the literature [8, 7, 3, 11] for more in depth treatment. For Ω be a bounded open subset of $\mathbb{R}^n$ with $n \in \mathbb{N}$, the space of infinitely many times differentiable functions of compact support $\mathcal{D}(\Omega)$ is defined as

$$\mathcal{D}(\Omega) = \left\{ u \in C^\infty(\Omega) \,\middle|\, \mathrm{supp}(u) = \overline{[u \neq 0]} \subset\subset \Omega \right\}$$

where "$\subset\subset$" means "compactly contained". To be able to generate the appropriate associated class of distributions this space needs to be endowed with the inductive limit topology obtained by means of the following family of locally convex spaces

$$\mathcal{D}_{\Omega'}(\Omega) = \left(\left\{ u \in \mathcal{D}(\Omega) \,\middle|\, \mathrm{supp}(u) \subset \Omega' \subset\subset \Omega \right\}, \, \mathcal{P} = \left\{ p_{\Omega',m} \,\middle|\, m \in \mathbb{N} \right\} \right)$$

where $\mathcal{P}$ is the separating family of semi-norms defined through

$$p_{\Omega',m}(u) = \sup_{x \in \Omega', \, |\alpha| \leq m} |\partial^\alpha u(x)|, \ m \in \mathbb{N}$$

and choosing the smallest topology which makes the following inclusions continuous

$$\mathcal{D}_{\Omega'}(\Omega) \hookrightarrow \mathcal{D}(\Omega), \ \Omega' \subset\subset \Omega.$$

The reader can find the details of this topological construction in [3]. The corresponding space of distributions (generalized functions) is then obtained as the topological dual of $\mathcal{D}(\Omega)$, that is,

$$\mathcal{D}'(\Omega) = \mathcal{D}(\Omega)'.$$

It is endowed with its weak* topology. Since the main focus of this paper is on finite discrete approximations particular interest lies in the convergence of sequences and series. In the chosen topologies their convergence is equivalent to "pointwise" convergence. Indeed, let $(u_n)_{n \in \mathbb{N}}$ and $(\varphi_n)_{n \in \mathbb{N}}$ be sequences in $\mathcal{D}'(\Omega)$ and $\mathcal{D}(\Omega)$, respectively. Then the first converges to a limit $u_\infty \in \mathcal{D}'(\Omega)$ if and only if

$$\langle u_n, \varphi \rangle := u_n(\varphi) \to u_\infty(\varphi) = \langle u_\infty, \varphi \rangle \text{ for each } \varphi \in \mathcal{D}(\Omega)$$

whereas the second converges towards a limit $\varphi_\infty \in \mathcal{D}(\Omega)$ if and only if there exists $\Omega' \subset\subset \Omega$ such that $\mathrm{supp}(\varphi_n)$, $\mathrm{supp}(\varphi_\infty) \subset \Omega'$ for $n \in \mathbb{N}$ and

$$\langle u, \varphi_n \rangle \to \langle u, \varphi_\infty \rangle \text{ for each } u \in \mathcal{D}'(\Omega)$$

Given any $f \in \mathrm{L}_{1,loc}(\Omega)$ in the space of locally integrable functions, a distribution u_f can be defined by

$$\langle u_f, \varphi \rangle := \int_\Omega f(x)\varphi(x)\,dx\,, \ \varphi \in \mathcal{D}(\Omega)\,.$$

In other words, locally integrable functions naturally act on test functions. Any distribution of this form is called *regular*. The duality between distributions and test functions and, in particular, the *integral duality* between regular distributions and test functions is exploited in the definition of all standard operations for distributions and will soon play an essential role in understanding *"discrete distributions"*. Another space of test functions, $\mathcal{E}(\Omega) = \mathrm{C}^\infty(\Omega)$ is endowed with its natural locally convex topology generated by the separating family

$$\left\{ p_{\Omega',m} \,\middle|\, \Omega' \subset\subset \Omega\,, \ m \in \mathbb{N} \right\}.$$

Its dual $\mathcal{E}'(\Omega)$ of $\mathcal{E}(\Omega)$ can also naturally be viewed as a space of distributions. It actually is the space of distributions of compact support, where the support of a distribution u is defined as the closure of the complement of

$$\left\{ x \in \Omega \,\middle|\, \text{there exists a neighborhood } \Omega' \text{ of } x \text{ s.t. } \langle u, \varphi \rangle = 0\,, \ \varphi \in \mathcal{D}(\Omega') \right\}$$

It turns out that all compactly supported distributions are of finite order. The order of a distribution is defined as follows. Any $u \in \mathcal{D}'(\Omega)$ is said to be of order $m \in \mathbb{N}$ if and only if for any given $\Omega' \subset\subset \Omega$ there is a constant $C = C(\Omega') > 0$ such that

$$|\langle u, \varphi \rangle| \le C\, p_{m,\Omega'}(\varphi)\,, \ \varphi \in \mathcal{D}_{\Omega'}(\Omega)\,.$$

In particular, the compactly supported distributions $\partial^\alpha \delta_y \in \mathcal{E}'(\Omega)$, $\alpha \in \mathbb{N}^n$ and $y \in \Omega$, defined by

$$\langle \partial^\alpha \delta_y, \varphi \rangle = (-1)^{|\alpha|}(\partial^\alpha \varphi)(y)\,, \ \varphi \in \mathcal{E}(\Omega) \tag{2.1}$$

are of finite order $|\alpha|$. The space of finite regularity test functions is given by

$$\mathcal{D}^m(\Omega) = \left\{ \varphi \in \mathrm{C}^m(\Omega) \,\middle|\, \mathrm{supp}(\varphi) \subset\subset \Omega \right\} \tag{2.2}$$

in a way similar to $\mathcal{D}(\Omega)$, by duality one obtains the space $\mathcal{D}^m(\Omega)'$ of distributions of order at most $m \in \mathbb{N}$. Correspondingly, $\mathcal{E}^m(\Omega) = \mathrm{C}^m(\Omega)$ gives rise to the space $\mathcal{E}^m(\Omega)'$ of compactly supported distribution of order at most $m \in \mathbb{N}$.
It is important to point out that

$$\mathcal{D}(\Omega) \xrightarrow{d} \mathcal{E}(\Omega) \xrightarrow{d} \mathrm{L}_{1,loc}(\Omega) \xrightarrow{d} \mathcal{D}'(\Omega)\,, \tag{2.3}$$

where the "d" indicates density of the inclusion. The Hilbert space $\mathrm{L}_2(\Omega)$ and orthonormal bases on it are also basic ingredients of the approach presented here. It will be convenient to work with bases of smooth functions. They usually do not

have compact support, however. Just think of all eigenfunction bases associated to boundary value problems. Given an orthonormal basis $(e_n)_{n\in\mathbb{N}}$ for $L_2(\Omega)$ with

$$e_n \in C^\infty(\overline{\Omega})\,,$$

$(u_k)_{k\in\mathbb{N}} \in \mathbb{R}^\mathbb{N}$ and a distribution $u \in \mathcal{E}'(\Omega)$ of compact support, it follows that

$$\sum_{k=1}^m u_k e_k \xrightarrow[m\to\infty]{} u \text{ in } \mathcal{D}'(\Omega)$$

if and only if

$$\langle \sum_{k=1}^m u_k e_k, \varphi \rangle = \sum_{k=1}^m u_k \varphi_k \xrightarrow[m\to\infty]{} \langle u, \varphi \rangle\,,\ \varphi \in \mathcal{D}(\Omega)\,,$$

where

$$\varphi_k = \int_\Omega e_k(x)\varphi(x)\,dx\,.$$

The following simple but important relation

$$\langle u, \varphi \rangle = \big\langle u, \sum_{k=1}^\infty \langle e_k, \varphi \rangle\, e_k \big\rangle = \sum_{k=1}^\infty \langle u, e_k \rangle \langle e_k, \varphi \rangle = \big\langle \sum_{k=1}^\infty \langle u, e_k \rangle\, e_k\,, \varphi \big\rangle\,,$$

$$u \in \mathcal{E}'(\Omega)\,,\ \varphi \in \mathcal{D}(\Omega) \quad (2.4)$$

shows that the basis coefficients of u can be computed by

$$u_k = \langle u, e_k \rangle\,,\ k \in \mathbb{N} \tag{2.5}$$

whenever convergence takes place. Unfortunately nothing can be said about the latter in general. In view of the duality construction used to introduce distributions, however, it is always the case that if either the series

$$\sum_{k=1}^\infty \langle \varphi, e_k \rangle\, e_k \tag{2.6}$$

converges in $\mathcal{E}(\Omega)$ or

$$\sum_{k=1}^\infty \langle u, e_k \rangle\, e_k \tag{2.7}$$

converges in $\mathcal{D}'(\Omega)$ then the validity of (2.4) is assured.

Remark 2.1. Formula (2.4) shows in particular that

$$\langle u, \varphi \rangle = \sum_{k=1}^\infty \langle u, e_k \rangle \langle e_k, \varphi \rangle = \sum_{k=1}^\infty u_k \varphi_k\,,\ u \in \mathcal{E}'(\Omega)\,,\ \varphi \in \mathcal{D}(\Omega)\,, \tag{2.8}$$

whenever there is convergence. Assuming that all functions and distributions considered be real-valued and switching to the case where $u, \varphi \in L_2(\Omega)$, (2.8) becomes Parseval's identity. It is therefore legitimate to call it *generalized Parseval's identity*.

Remarks 2.2. (a) The requirement that (2.6)-(2.7) be convergent in the given topologies is essentially a "smoothness and boundary behavior assumption" on u and φ with respect to the basis $(e_n)_{n \in \mathbb{N}}$. If no further information about properties of the basis are available, it is, however, not possible to measure this regularity in classical function spaces. Fortunately, many basis of practical interest consist of eigenfunctions of some operator. In that case more can be said about the convergence of (2.6)-(2.7) and consequently of (2.8). More details are found in Subsection 4.

(b) So far only the real-valued case is considered. The complex-valued case can clearly be covered with only minor modifications caused by the incongruence of the sesquilinearity of the scalar product and the bilinearity of the duality pairing.

Turning to a simple concrete example, let $y \in \Omega$ and $\alpha \in \mathbb{N}^n$ and consider the series

$$\partial^\alpha \delta_y = (-1)^{|\alpha|} \sum_{k \in \mathbb{Z}} (\partial^\alpha \bar{e}_k)(y) \, e_k \, . \tag{2.9}$$

It will provide with a smooth approximation

$$\Big[(-1)^{|\alpha|} \sum_{|k| \le m} (\partial^\alpha \bar{e}_k)(y) \, e_k \Big]_{m \in \mathbb{N}} \tag{2.10}$$

to $\partial^\alpha \delta$ whenever the series converges in the sense of distributions. This is a very natural way of taking advantage of the duality built-in in the distributional framework without completely giving up the benefits of orthogonal expansions which are restricted to Hilbert spaces. A particular case of the above is given by Fourier series of periodic functions. Any test function of the periodicity cube $B_n = [-\pi, \pi]^n$ can be viewed as a periodic function. It can therefore be developed in a Fourier series

$$\varphi = \frac{1}{(2\pi)^{n/2}} \sum_{k \in \mathbb{Z}^n} \hat{\varphi}_k \exp(ik \cdot x) \tag{2.11}$$

with Fourier coefficients given by

$$\hat{\varphi}_k = \frac{1}{(2\pi)^{n/2}} \int_{B_n} \varphi(x) \exp(-ik \cdot x) \, dx \, . \tag{2.12}$$

In this case one has

Lemma 2.3. *Let* $\varphi \in \mathcal{D}(B_n)$. *Then its Fourier series expansion* (2.11) *converges in the topology of* $\mathcal{E}(B_n)$. *For any distribution* $u \in \mathcal{E}'(B_n)$ *its Fourier coefficient* u_k *is defined to be*

$$u_k = \frac{1}{(2\pi)^{n/2}} \langle u, \exp(-ik \cdot x) \rangle \, . \tag{2.13}$$

Then

$$\frac{1}{(2\pi)^n} \sum_{k=-m}^{m} u_k \exp(ik \cdot x) \to u \ (m \to \infty) \ in \ \mathcal{D}'(B_n) \, . \tag{2.14}$$

Proof. The proof follows from the general convergence result of Theorem 4.1. $\square$

Example 1. Consider the distribution $\partial^\alpha \delta$ for $\alpha \in \mathbb{N}^n$. Then (2.13) gives

$$(\partial^\alpha \delta)_k = (-ik)^\alpha / (2\pi)^{n/2}$$

and

$$\frac{1}{(2\pi)^n} \sum_{k=-m}^{m} (-ik)^\alpha \exp(ik \cdot x) \xrightarrow[m \to \infty]{} \partial^\alpha \delta \text{ in } \mathcal{D}'(B_n). \qquad (2.15)$$

Remarks 2.4. (a) More in general if a distribution is of finite order, then convergence in a weaker topology than that of $\mathcal{E}(\Omega)$ suffices to obtain convergence of its basis development in a stronger topology than that of the space $\mathcal{E}'(\Omega)$. This point will be raised again in section 4.
(b) Specializing to $u = \delta$ and $n = 1$, the one dimensional Dirac distribution supported in the origin, one sees that its approximating Fourier sum coincides with the classical Dirichlet kernel

$$D_m(x) = \frac{1}{2\pi} \sum_{k=-m}^{m} \exp(ik \cdot x). \qquad (2.16)$$

These very simple facts and examples turn out to be very important when approaching the problem of discretization. The problem of discretizing a (generalized) function is in fact not independent of the problem of discretizing the underlying duality structure between test functions and distributions. For this reason the discretization of functions and that of the duality pairing have to be related to each other in order to produce optimal results.

3. Approximating families

Next the discrete version of the above concepts is considered. Vectors will appear instead of distributions but the distinction between function and distribution in the duality sense will not be given up. This might seem a minor point since every finite dimensional space is naturally isomorphic to its dual but it is not. It determines, among other things, the way in which the information content of a given vector of finite length has to be read. The main concept introduced in this section is that of *approximation family* for a distribution. It is meant to faithfully reproduce the continuous analytical structure at the finite dimensional, discrete level. To realize a discrete duality pairing quadrature rules are used. Assume that a family of discretization points has been chosen

$$x^m = (x_1^m, \ldots, x_{g(m)}^m), \ m \in \mathbb{N} \qquad (3.1)$$

where the strictly increasing function $g : \mathbb{N} \to \mathbb{N}$ with $g(m) \geq m$ counts the total number of grid points. A *discrete quadrature rule* for Ω with respect to a family of discretization points $(x^m)_{m \in \mathbb{N}}$ is a family of vectors

$$q^m = (q_1^m, \ldots, q_{g(m)}^m) \qquad (3.2)$$

such that

$$\sum_{k=1}^{g(m)} q_k^m \varphi(x_k^m) \xrightarrow[m\to\infty]{} \int_\Omega \varphi(x)\, dx\,, \ \varphi \in \mathcal{E}(\Omega)\,. \tag{3.3}$$

The quadrature rules for Ω shall be denoted by $\left(x^m, q^m\right)_{m\in\mathbb{N}}$ or simply by $(q^m)_{m\in\mathbb{N}}$ if it is clear which discretization points have been fixed. The following definition plays a central and fundamental role.

Definition 3.1. Let $u \in \mathcal{D}'(\Omega) \left[\mathcal{E}'(\Omega)\right]$ be a given distribution. Then

$$[u^m, x^m, q^m]_{m\in\mathbb{N}}$$

is called a *discretization family* (in the sense of distributions) for u iff

(i) $u^m \in \mathbb{R}^{g(m)}$, $m \in \mathbb{N}$. $\hspace{5cm}$ (3.4)

(ii) $\left(x^m,\, q^m\right)_{m\in\mathbb{N}}$ is quadrature rule for Ω. $\hspace{3cm}$ (3.5)

(iii) $u^m \cdot \varphi^m := \displaystyle\sum_{k=1}^{g(m)} u_k^m \varphi(x_k^m) q_k^m \xrightarrow[m\to\infty]{} \langle u, \varphi\rangle$ for each $\varphi \in \mathcal{D}(\Omega) \left[\mathcal{E}(\Omega)\right]$. $\hspace{0.3cm}$ (3.6)

If it is assumed that the approximated distribution is regular, that is, if $u \in \mathrm{L}_{1,loc}(\Omega)$, then the above definition entails that

$$u^m \cdot \varphi^m \to \int_\Omega u(x)\varphi(x)\, dx \text{ for each } \varphi \in \mathcal{D}(\Omega)\,. \tag{3.7}$$

It follows that the definition is the discrete version of the classical concept of weak convergence for sequences of functions.

Remarks 3.2. (a) It should be pointed out that by modifying the definition of the approximating sequence as follows

$$\tilde{u}^m = (q_1^m u_1^m, \ldots, q_{g(m)}^m u_{g(m)}^m)$$

the discrete duality pairing could be normalized to be the Euclidean scalar product. This will always be done whenever dealing with concrete examples. For abstract calculations, however, it is preferable to have the quadrature rule appear explicitly in the formulæ.

(b) Choosing $u^m = (1, \ldots, 1)$, $m \in \mathbb{N}$, one can think of some distributions (read measures) as being approximated by quadrature rules. Or better still, one could view quadrature rules as discretizations families of measures in the sense of distributions. In the above definition quadrature rules are obviously encoded in the choice of duality pairing.

A few prototypical examples are considered next.

Example 2. Let $y \in (0,1)$ and $\mathbf{1}^m = \underbrace{(1, \ldots, 1)}_{m\text{times}}$ and consider the family

$$[\delta_y^m, x^m, q^m] = \Big[2\sum_{k=0}^m \sin(k\pi y)\sin(k\pi x^m)\,,\ x^m = (k/m)_{k=1,\ldots,m-1}\,,\ \frac{1}{m}\mathbf{1}^{m-1}\Big]_{m\in\mathbb{N}}$$

It approximates the Dirac distribution δ_y supported at $y \in (0,1)$ with respect to the trapezoidal rule of quadrature.

Proof. For any given test function $\varphi \in \mathcal{D}(0,1)$ the duality pairing is given by

$$\delta_y^m \cdot \varphi^m = \sum_{j=1}^{m} \Big(2 \sum_{k=1}^{m} \sin(k\pi y) \sin(k\pi x_j^m) \Big) \varphi(x_j^m) \frac{1}{m}$$

$$= \sum_{k=1}^{m} \sqrt{2} \sin(k\pi y) \Big(\frac{1}{m} \sum_{j=1}^{m} \sqrt{2} \sin(k\pi x_j^m) \varphi(x_j^m) \Big).$$

The inner sum in the second line appears to be a trapezoidal rule discretization of

$$\sqrt{2} \int_0^1 \sin(k\pi x) \varphi(x) \, dx = \langle \varphi, \sqrt{2} \sin(k\pi \cdot) \rangle = \varphi_k$$

and it is therefore obtained that

$$\delta^m \cdot \varphi^m \longrightarrow \sum_{k=1}^{\infty} \varphi_k \sin(k\pi y) = \varphi(y) \; (m \to \infty) \tag{3.8}$$

which proves the claim. The convergence of the trapezoidal rule, which has implicitly been used, is elementary and omitted. $\square$

Remark 3.3. In this case

$$e_k(x) := \sqrt{2} \sin(k\pi x) \,, \; x \in (0,1) \,, \; k \in \mathbb{N} \tag{3.9}$$

build an orthonormal basis of $L_2(0,1)$ with $e_k \in C^{\infty}[0,1]$, $k \in \mathbb{N}$. It is very important that this basis consists of solutions (eigenfunctions) of

$$-\partial_{xx} u = \lambda u \,, \; u(0) = u(1) = 0 \,.$$

In more general situations, like in Example 5, this will provide the means for proving convergence.

This simple but important example can easily be extended to any dimension.

Example 3. Let $y \in \Omega = (0,1)^n$, $\alpha \in \mathbb{N}^n$ and

$$\mathbf{x}^m = \otimes_{k=1}^{n} x^m \,, \; (\delta_y^{\alpha})^m = \otimes_{k=1}^{n} (\delta_{y_k}^{\alpha})^m \tag{3.10}$$

for

$$(\delta_{y_k}^{\alpha})^m = 2 \sum_{j=0}^{m} (-j\pi)^{|\alpha_k|} \big(\partial^{\alpha_k} \sin \big) (j\pi y_k) \sin(j\pi x_k^m) \,, \; k = 1, \ldots, n \,.$$

Then $(\delta_y^{\alpha})^m$ is a discretization family for $\delta_y^{\alpha} = \partial^{\alpha} \delta_y \in \mathcal{E}'(0,1)$.

In a periodic context this construction essentially produces the Dirichlet kernel D_m evaluated at the grid points.

Example 4. Let $\Omega = (-\pi, \pi)$ and let

$$x^m = \frac{\pi}{m}(-m, -m+1, \ldots, m-1, m)$$

be the tuple of $2m+1$ equidistant discretization points. Fix the trapezoidal rule

$$q^m = (\frac{1}{2}, \underbrace{1, \ldots, 1}_{2m-1\text{-times}}, \frac{1}{2})$$

as the associated quadrature rule. Then the following modified Dirichlet kernel

$$\delta^m = \frac{1}{2\pi} + \frac{1}{\pi}\sum_{k=1}^{m-1}\cos(kx^m) + \frac{1}{2\pi}\cos(m\,x^m) =$$

$$\frac{1}{2\pi}\frac{\sin\big((m-\frac{1}{2})x^m\big)}{\sin(\frac{x^m}{2})} + \frac{1}{2\pi}\cos(m\,x^m) = D_{m-1}(x^m) + \frac{1}{2\pi}\cos(m\,x^m) \quad (3.11)$$

defines a discretization family for $\delta \in \mathcal{E}'(-\pi, \pi)$.

Proof. The proof is identical to that of example 2.   □

Remarks 3.4. (a) In example 4 the so-called alternating point trapezoidal rule given by the family of weights

$$q^m = (1, 0, 2, 0, \ldots, 2, 0, 1)$$

with $2m+1$ equidistant discretization points as above could have been chosen.
(b) The approximating family obtained in example 2 is closely related to that of example 4. In fact

$$2\sum_{k=0}^{m}\sin(k\pi y)\sin(k\pi x^m) = \pi D_m(\pi(x-y)) - \pi D_m(\pi(x+y)). \quad (3.12)$$

(c) Disregarding the role played by the quadrature rule in the construction of the a discrete approximation for the Dirac distribution in the above example, it would seem natural to use the Dirichlet kernel (2.16) to produce a discretization family for δ. Even though this would be a viable discretization, it would unfortunately converge more slowly than (3.11).

Lastly a genuinely higher dimensional example is considered.

Example 5. Let $\Omega = \mathbb{B}^2$ be the open unit circle parameterized by spherical coordinates and

$$e_{m,n}(r, \theta) = \frac{1}{c_{mn}}J_m(r\sqrt{\lambda_{mn}})e^{im\,\theta}, \ r \in [0, 1], \ \theta \in [-\pi, \pi)$$

be the orthonormal basis of $L_2(\Omega)$ given by the eigenfunctions of the Dirichlet problem

$$-\triangle u = \lambda u \text{ in } \Omega, \ u = 0 \text{ on } \partial\Omega.$$

to eigenvalues λ_{mn}, $n \in \mathbb{N}$ given by the positive zeros of the Bessel function J_m, $m \in \mathbb{N}$. The constants c_{mn} are chosen so that $\|e_{mn}\|_{L_2(\Omega)} = 1$.
Discretization points are

$$r_j^M = j/M\,,\ 0 \le j \le M \text{ and } \theta_j^N = \pi(-1 + j/N)\,,\ 0 \le j \le 2N\,,\ M, N \in \mathbb{N}\,.$$

For quadrature, the trapezoidal rule is used in both variables, that is,

$$q_r^M = \frac{1}{M}(r_0^M, r_1^M, \dots, r_M^M)\,,\ q_\theta^N = \frac{\pi}{N}\mathbf{1}^{2N+1}\,,\ M, N \in \mathbb{N}\,.$$

Then the family

$$\sum_{m=0}^{M} \sum_{n=-N}^{N} \frac{1}{c_{mn}^2} J_m(r_0\sqrt{\lambda_{mn}}) J_m(r^M\sqrt{\lambda_{mn}}) e^{im(\theta^N - \theta_0)} \tag{3.13}$$

defines a discretization family for $\delta_{(r_0, \theta_0)}$ for any $(r_0, \theta_0) \in \Omega$. Let now Γ be a closed smooth curve completely contained in Ω. Denote by δ_Γ the line integral distribution defined through

$$\langle \delta_\Gamma, \varphi \rangle = \int_\Gamma \varphi(x)\, d\sigma_\Gamma(x)\,,\ \varphi \in \mathrm{C}^\infty(\Omega)\,.$$

where σ_Γ is the surface measure. Then

$$\sum_{m=0}^{M} \sum_{n=-N}^{N} \frac{I_{mn}}{c_{mn}} J_m(r^M\sqrt{\lambda_{mn}}) e^{im\,\theta^N} \tag{3.14}$$

with

$$I_{mn} = \frac{1}{c_{mn}} \int_\Gamma J_m(r\sqrt{\lambda_{mn}}) e^{-im\,\theta}\, d\sigma_\Gamma(r, \theta)\,. \tag{3.15}$$

defines an approximation family for δ_Γ.

Proof. As for the Dirac distribution the claim would be a direct consequence of (2.4)-(2.5) and (2.9) combined with the know fact that $\big(e_{m,n}\big)_{m,n \in \mathbb{N}}$ is indeed an orthonormal basis of $L_2(\Omega)$ and with the convergence of the chosen quadrature rule if

$$\sum_{m=0}^{M} \sum_{n=-N}^{N} \varphi_{mn} e_{m,n}$$

converged at least in $\mathrm{C}(\Omega)$ as $M, N \to \infty$ for any $\varphi \in \mathcal{E}(\Omega)$. This will be considered in a more general setting in section 4. In the case of the line integral distribution the proof is similar and uses the complex version of (2.4)-(2.5). $\qquad \square$

Remarks 3.5. (a) In the last example (3.15) can not be evaluated analytically. It can, however, be easily approximated numerically with high order of accuracy.
(b) Wavelet bases can of course also be used. One of their main purpose is, however, to provide localized basis functions. They also usually have finite degree of smoothness. Unless C^∞ wavelets are used, these two facts concur in making it impossible to achieve spectrality (in the sense of distributions) of any approximation

method based on their use. Their main advantage lies of course in their multi-scale resolution properties.

4. Convergence

Definition 3.1 of approximation family given in the previous section leads to useful discrete approximations provided (2.8) does indeed converge for the chosen basis $(e_n)_{n\in\mathbb{N}}$. A criterion for their convergence is therefore derived which entails convergence of general discretization families. The standard basic tool is to trade smoothness for convergence.

Looking at the examples 2-5 given in section 3 it is recognized that they all share a specific structure which yields the desired convergence. One starts with some operator pair $(\mathcal{A}, \mathcal{B})$ where

$$\mathcal{A} = \sum_{|\alpha|\leq m} a_\alpha \partial^\alpha \,,\ a_\alpha \in \mathbb{R}^M \text{ and } \mathcal{B} = \sum_{|\alpha|\leq k} b_\alpha \partial^\alpha \,,\ b_\alpha \in \mathbb{R}^M \tag{4.1}$$

are a vector of differential operators on Ω of order $m \geq 1$ and a vector of boundary differential operators on $\partial\Omega$ of order $k < m$, respectively. It is always possible to introduce

$$L : \operatorname{dom} L \subset \mathrm{L}_2(\Omega) \to \mathrm{L}_2(\Omega)^M \tag{4.2}$$

where

$$\operatorname{dom}(L) = \mathrm{H}^m_\mathcal{B}(\Omega) := \left\{ u \in \mathrm{H}^m(\Omega) \,\middle|\, \mathcal{B}u = 0 \right\} \text{ and } Lu = \mathcal{A}u \,,\ u \in \operatorname{dom}(L)\,, \tag{4.3}$$

with the understanding that the homogeneous boundary condition has to be imposed in the sense of traces whenever it makes sense and has to considered empty otherwise. Then, if L is closed and densely defined, the self-adjoint operator

$$A = L'L : \operatorname{dom}(A) \subset \mathrm{L}_2(\Omega) \to \mathrm{L}_2(\Omega) \tag{4.4}$$

can be defined. If the latter turns out to be invertible, it has a compact resolvent by the compact embedding

$$\operatorname{dom}(A) \subset \mathrm{H}^{2m}(\Omega) \hookrightarrow \mathrm{L}_2(\Omega)\,.$$

It is therefore possible to introduce an orthonormal basis $(e_n)_{n\in\mathbb{N}}$ on $\mathrm{L}_2(\Omega)$ consisting of the necessarily smooth eigenfunctions of A to the ordered family of positive eigenvalues $(\lambda_n)_n \in \mathbb{N}$. Then this basis has all the needed properties.

Theorem 4.1. *Assume that $(\mathcal{A}, \mathcal{B})$ is such that the operator L defined through (4.2)-(4.3) be closed and densely defined. Let the self-adjoint operator A given in (4.4) have finite dimensional kernel, then*

$$\sum_{k\in\mathbb{N}} \langle e_k, \varphi\rangle e_k \overset{k\to\infty}{\longrightarrow} \varphi \ \text{in } \mathcal{E}(\Omega) \tag{4.5}$$

for any $\varphi \in \mathcal{D}(\Omega)$ and

$$\sum_{k\in\mathbb{N}} \langle u, e_k\rangle e_k \overset{k\to\infty}{\longrightarrow} u \ \text{in } \mathcal{D}'(\Omega) \tag{4.6}$$

for any $u \in \mathcal{E}'(\Omega)$.

Proof. Factoring out the finite dimensional kernel, if necessary, it can be assumed that A is invertible and therefore has compact inverse by the Rellich-Kondratev embedding theorem, [10, Proposition 4.4]. The spectral theory of compact operators implies that $(e_k)_{k \in \mathbb{N}}$ is indeed a orthonormal basis of $\mathrm{L}_2(\Omega)$. Series (4.5) therefore converges in $\mathrm{L}_2(\Omega)$ at least. Now

$$\varphi \in \mathrm{dom}(A^m) \, , \ m \in \mathbb{N} \, ,$$

since $\mathcal{D}(\Omega) \subset \mathrm{dom}(A)$ and $A\varphi \in \mathcal{D}(\Omega)$ for any $\varphi \in \mathcal{D}(\Omega)$. The closure of A implies that

$$A^m \sum_{k \in \mathbb{N}} \langle e_k, \varphi \rangle e_k = \sum_{k \in \mathbb{N}} \langle e_k, \varphi \rangle A^m e_k = \sum_{k \in \mathbb{N}} \langle \lambda_k^m e_k, \varphi \rangle e_k = \sum_{k \in \mathbb{N}} \langle e_k, A^m \varphi \rangle e_k$$

which yields $\mathrm{H}^m(\Omega)$ convergence for the series since $\|A^m \cdot\|_{\mathrm{L}_2(\Omega)}$ is an equivalent norm on $\mathrm{H}^m(\Omega)$ and $A^m \varphi \in \mathrm{L}_2(\Omega)$. Notice that factoring the kernel has no impact on the convergence of the series in any way, since it is assumed to be finite dimensional. Sobolev embedding theorem (cf. [10, Prop. 1.5])

$$\mathrm{H}^s(\Omega) \hookrightarrow \mathrm{C}^{s-n/2}(\overline{\Omega}) \, , \ s > 0 \text{ with } s - n/2 > 0$$

then implies convergence in $\mathrm{C}^s(\Omega)$ for every $s > 0$ which implies the first assertion. The second convergence claim follows from (2.4). $\qquad\square$

Example 6. In examples 2-3 and 5 the basis functions are eigenfunctions of the Laplacian on $\mathrm{L}_2(\Omega)$ with Dirichlet boundary conditions for $\Omega = (0,1) \, , \ (0,1)^n$ and $\mathbb{B}^2$, respectively.

Example 7. Taking the operator $L = \nabla$ with domain of definition

$$\mathrm{dom}(L) = \mathrm{H}_p^1\big((-\pi, \pi)^n\big)$$

where the subscript "p" stands for periodic on the periodicity box $(-\pi, \pi)^n$, one has that

$$L' = -\,\mathrm{div}$$

with same domain of definition, and therefore

$$A = -\,\mathrm{div}(\nabla \cdot)$$

with domain $\mathrm{dom}(A) = H_p^2\big((-\pi, \pi)^n\big)$. This is the case of Fourier series, example 4.

This, together with the proofs given there, shows that the families of examples 2-4 are indeed discretization families in the sense of definition 3.1. As to the general situation, assume that an orthonormal basis $(e_k)_{k \in \mathbb{N}}$ of $\mathrm{L}_2(\Omega)$ is given. Let $u \in \mathcal{E}'(\Omega)$ and $\varphi \in \mathcal{D}(\Omega)$ be given. Choose a family of discretization points $(x^m)_{m \in \mathbb{N}}$ and an associated convergent quadrature rule $(q^m)_{m \in \mathbb{N}}$. Let $\bar{u}^m$ and $\bar{\varphi}^m$

be approximations to the exact coefficient vector u^m and to the exact grid point values φ^m, respectively. Then

$$\langle u, \varphi \rangle - \bar{u}^m \cdot \bar{\varphi}^m = \sum_{k=m+1}^{\infty} u_k \varphi_k + \sum_{k=1}^{m} u_k \varphi_k - u^m \cdot \varphi^m$$
$$+ (u^m - \bar{u}^m) \cdot \bar{\varphi}^m + u^m \cdot (\varphi^m - \bar{\varphi}^m)$$
$$= \sum_{k=m+1}^{\infty} u_k \varphi_k + \sum_{k=1}^{m} u_k (\varphi_k - \hat{\varphi}_k^m) + (u^m - \bar{u}^m) \cdot \bar{\varphi}^m + u^m \cdot (\varphi^m - \bar{\varphi}^m) \quad (4.7)$$

where

$$\hat{\varphi}_k^m = e_k^m \cdot \varphi^m = \sum_{k=1}^{m} e_k(x_k^m) \varphi(x_k^m) q_k^m \,.$$

Disregarding the numerical error in the approximation of the coefficients, the error depends on the order of the distribution u and on the difference between the exact and discrete basis coefficients of the test function φ. The first determines the polynomial growth rate of the u_k's, the latter is super-algebraically convergent for any test function φ. More precisely, since the distribution u is of finite order, one has that

$$|u_k| \le c(u) k^p \,, \ k \in \mathbb{N}$$

for some positive constants $c(u)$ and some order $p \in \mathbb{N}$. As for the test function, for every $P \ge 0$ a constant $c(P, \varphi) \ge 0$ can be found such that

$$|\varphi_k| \le c(P, \varphi) k^{-P} \,, \ k \in \mathbb{N} \,.$$

It follows that

$$| \sum_{k=m+1}^{\infty} u_k \varphi_k | \le c(P', u, \varphi) \frac{1}{m^{P'}}$$

for any $P' \ge 0$. As for the error $|\varphi_k - \hat{\varphi}_k|$ incurred in the computation of the discrete basis coefficient, it is determined by the effect of aliasing and can be bound as follows. It can be easily seen from

$$\varphi = \sum_{k=1}^{\infty} \varphi_k e_k$$

that the discrete basis coefficient differs from the exact one in the amount

$$|\varphi_j - \hat{\varphi}_j| = | \sum_{k=m+1}^{\infty} \varphi_k e_k^m \cdot e_j^m | \,.$$

This follows from the choice of the discrete duality pairing which preserves orthogonality of the first m basis functions. Again, since φ is a test function and since $e_k^m \cdot e_j^m = O(1)$, it follows

$$|\varphi_j - \hat{\varphi}_j| \le c(P, \varphi) \frac{1}{m^P}$$

for any $P \ge 0$ and suitable constant $c(P, \varphi) \ge 0$. Therefore one obtains

Theorem 4.2. *Let $u \in \mathcal{E}'(\Omega)$ and $\varphi \in \mathcal{D}(\Omega)$. Then for any $P \geq 0$, a constant $c = c(P, u, \varphi)$ can be found such that*

$$|\langle u, \varphi \rangle - u^m \cdot \varphi^m| \leq c(P, u, \varphi) \frac{1}{m^P} \qquad (4.8)$$

Remark 4.3. Approximations for test functions are best generated and handled in physical space. This is due to the fact that there is no easy way to tell when a rapidly convergent basis expansion will have compact support. On the other hand, special distributions often allow for an easy computation of their exact coefficients in the basis expansion. This implies that, in many cases of interest, the errors neglected in (4.7) can be made very small as well.

5. Interpolation

The general procedure described above applied to special distributions leads to recovering many interesting applications. One of them is interpolation. It can be naturally viewed as the approximation of (derivatives of the) Dirac distributions. In this case the above procedure delivers an "optimal" interpolation functional. It also turns out that a high order (in the sense of distributions) approximation of a (possibly only piecewise) smooth function has to carry non vanishing oscillations at the right scale. Put differently, even if singularities occur at non-grid points, their locations and intensity are captured via oscillations which degrade the pointwise quality of the approximation. However, if the oscillations are not viewed pointwise but rather in the sense of distributions, they are seen contain high order information concerning the approximated object. In the case of piecewise regular functions this allows for developing "recovery" methods which produce genuinely pointwise approximation. Even though it is not formulated in the general abstract terms of this paper, this is precisely what a series of methods proposed in the literature implicitly exploit (see for instance [5, 4]).

In general, however, the oscillatory behavior of high order distributional approximation causes a limitation in the pointwise accuracy that can be achieved without giving up non-local out-of-grid information concerning the limit. Even worse, the high order implicit information encoded in the oscillations can, in general, not be made explicit. This is precisely what has earlier been labeled as *numerical uncertainty principle*.

5.1. Location and intensity of singularities

So far it has been observed that distributions give rise to oscillations when they are approximated. Next the question is investigated of how to find out which distribution (if any) is represented by some given set of data with oscillatory behavior. To answer this question, the concept of φ-sweep, or test function sweep, is introduced. Consider the test functions given by

$$\varphi_{x_0, \alpha, \beta}(x) = \begin{cases} \exp\left(-\alpha \frac{\psi^2(x, x_0, \beta)}{1 - \psi^2(x, x_0, \beta)}\right), & |x - x_0| \leq \beta \\ 0, & |x - x_0| > \beta \end{cases} \qquad (5.1)$$

for

$$\psi(x, x_0, \beta) = \frac{x - x_0}{\beta}. \qquad (5.2)$$

and $\alpha, \beta > 0$. To distinguish the continuous test function from its discrete counterpart, the latter is denoted by $\varphi^m_{x_0, \alpha, \beta}$.

Definition 5.1. If u^m is a vector of discrete values approximating some distribution $u \in \mathcal{D}'(\Omega)$ with $\mathrm{supp}(u) \subset\subset \Omega$, the function defined by

$$S\big(u^m, \varphi^m_{\cdot, \alpha, \beta}\big)(x) = u^m \cdot \varphi^m_{x, \alpha, \beta}, \ x \in \Omega, \qquad (5.3)$$

is called a $\varphi^m_{\alpha, \beta}$-sweep of u^m.

Assume u^m is the manifestation of a Dirac distribution $c_0 \delta_{x_0}$ supported at $x_0 \in \Omega$ and with intensity $c_0 \in \mathbb{R}$, for instance. By applying a φ-sweep to it, an approximation to $\varphi_{x_0, \alpha, \beta}$ would be produced since

$$\langle c_0 \delta_{x_0}, \varphi_{x, \alpha, \beta} \rangle = c_0 \, \varphi_{\cdot, \alpha, \beta}(x_0) = c_0 \, \varphi_{x_0, \alpha, \beta}(x), \ x \in \Omega,$$

by definition of φ-sweep. Consequently the location of the peak would point to the location of the singularity and the height of the peak to its intensity. For a one dimensional illustration, take

$$m = 64, \ u^m = \delta^m_{\pi/6} - 0.5\delta^m_{1/\sqrt{2}} \qquad (\pi/6 \approx 0.5236, \ 1/\sqrt{2} \approx 0.7071)$$

and compute its φ-sweep

$$S\big(u^m, \varphi^m_{\cdot, \alpha, \beta}\big)$$

with $\alpha = 10$, $\beta = 0.1$ at $N = 256$ equidistant grid points. Figure 1 shows the result. The peaks are located at

$$x_1 = 0.5273 \text{ and } x_2 = 0.7109$$

and their heights are

$$p_1 \approx 0.9996 \text{ and } p_2 \approx 0.4999,$$

respectively. The example shows that, given oscillations due to Dirac singularities on a certain grid, it is possible to approximately locate their support with the help of test function sweeps, even if it is located outside the grid.

Remarks 5.2. (a) The above simple example has illustrative purposes only. The approximate location of the singularities is taken as the grid point at which the maximum is reached. By interpolation or, equivalently, by using finer sweeps better results can be achieved (up to a limit depending on m).
(b) If one is given some random oscillatory data without further information it wont be possible to use sweeps to find out anything, in general. Fortunately, in many cases, the information available is not confined to the data set itself.
(c) In general, how would can it be checked if the singularities observed are of δ-type, and not worse: say of δ'-type, for instance? A rough diagnosis can be given by looking at order of the discretization family, that is, at the power p for which $u^m = O(m^p)$. It roughly points to the order of the distributions.
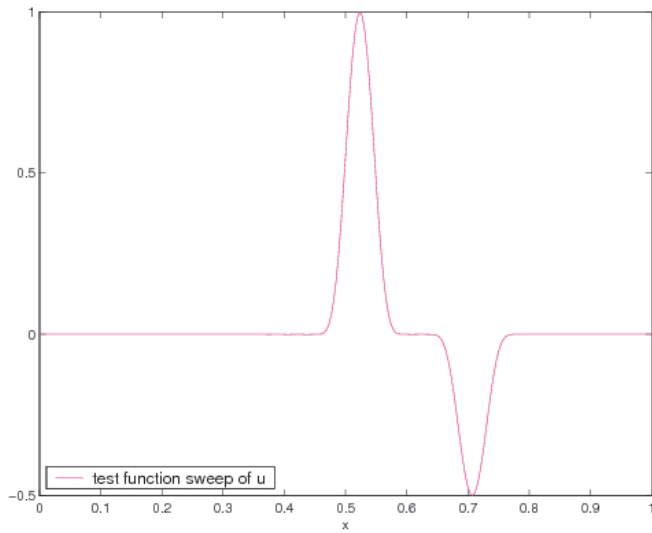
FIGURE 1. Locating singularities with a φ-sweep.

It is now possible to go back to the interpolation problem for discontinuous functions.

5.2. A remark about Gibbs phenomenon

Gibbs phenomenon is a widely observed phenomenon with which everybody working in numerical analysis or applied mathematics is very familiar with. It is observed whenever a discontinuous function is approximated by its Fourier series. The approximating series converges (albeit poorly) away from the jump discontinuity but keeps oscillating around it and does therefore not converge pointwise to its limit (due to persistent overshooting in the standard example). An historical perspective on the topic which raises and answers questions about the correct attribution of the discovery of this phenomenon can be found in [5]. This paper also contains a selection of many useful references from the vast literature about this problem.
Consider the simple sawtooth function

$$s_0(x) = \begin{cases} -(x/\pi + 1)/2\,, & x \in [-\pi, 0] \\ (1 - x/\pi)/2\,, & x \in (0, \pi] \end{cases}$$

defined on the interval $[-\pi, \pi]$. This is one of the standard one dimensional example used to illustrate Gibbs phenomenon. The function s_0 is to be considered as a 2π-periodic function. It obviously has a jump discontinuity at $x = 0$. If one tries to

approximate s_0 by its Fourier series

$$\sum_{n \in \mathbb{Z}^*} \frac{1}{\pi n} \exp(inx) \,,\; x \in [-\pi, \pi) \,,$$

then one can observe "experimentally" that the oscillations around the discontinuity point will concentrate but never disappear or attenuate no matter how many terms in the series are taken into the approximating sum. The series of course converges in $L_2(-\pi, \pi)$ but this is not quite good enough for many purposes. It is actually more advantageous to think of s_0 as a distribution than it is to consider it a function. It can in fact be argued that the most interesting convergence which is taking place is the one in the sense of distributions.

Now, if the $L_{2,p}(0,1)$ convergence is not good enough, why would one be interested in an even weaker convergence?

Well, this should be clear in light of the analysis presented in the previous sections.

It can be argued that Gibbs phenomenon, at the discrete level, is essentially an interpolation phenomenon. Truncating a Fourier series one always obtains an approximation with finite dimensional information content. Now, if one plots the truncated Fourier series as a function of a continuous variable one is in a sense implicitly interpolating the value of the approximation at infinitely many points. In the presence of a singularity this automatically leads to oscillations. However, at the discrete level, it is more natural to evaluate the truncated Fourier series at a finite number of points only. If the points are chosen appropriately

$$\left[x_j^m \mapsto \delta_{m, x_j^m} \cdot z^m \,,\; j = 1, \ldots, m \right] ,$$

then one obtains an approximating sequence which converges (albeit poorly) to z. Oscillations are still observed but are now decaying. In fact even for the Dirac distribution itself, one has "pointwise" convergence in the sense that

$$\frac{1}{2\pi m} D_{m, x_{j^m}}(x_j) = \delta_{j j^m} + O(\frac{1}{m}) \,,\; j = -m, \ldots, m \,,$$

assuming that its support always sits on a grid point. This again indicates that persistent oscillations are manifestation of the fact that of out-of-grid information needs to be approximated or accounted for (read *interpolated*). But even when oscillations occur, their exact behavior carries the information needed to determine the location of the singularity. The information has, however, to be read in the sense of distributions. As for the rest, the oscillatory approximation is as good any pointwise converging approximation. In fact, since finite dimensional approximations are used, the space of test functions can be spanned by finitely many "functions" and the distributional information content of an oscillating approximation and that of a pointwise approximation do indeed have to coincide. It just needs to be extracted in the proper way. In many concrete applications it is not known where exactly a singularity might be located and the occurrence of persistent oscillations will be the rule rather than the exception. The pointwise convergence observed in that case, however, is so poor that the local method is always preferred.

Dirichlet's kernel (2.16) is nothing but the manifestation of a very general way of producing discretization families for the Dirac distribution. See (2.9) and Example 3, in particular. A δ^m-sweep (defined similarly to a test function sweep) by means of Dirichlet's kernel for a function ψ can be rephrased in terms of a convolution with it in view of the periodic structure. Or, in other words,

$$D_{m,x_0}(x) = D_m(x - x_0), \ x \in [-\pi, \pi].$$

In terms of the distributional framework, one sees that

$$S(D_{m,x_0}, \psi_m) = \sum_{k=-m}^{m} \hat{\psi}_m^k \exp(ikx_0), \ x_0 \in (-\pi, \pi)$$

for the approximate discrete Fourier coefficients (computed w.r.t. the trapezoidal rule)

$$\hat{\psi}_m^k = \psi_m \cdot \exp(-ikx^m)$$

in this particular case. This means that oscillations observed whenever using a δ-sweep to represent or interpolate a discontinuous function are in this particular case what is known as Gibbs phenomenon. By using localized biased δ-sweeps obtained by multiplication with both test functions like (5.1) and with one-sided functions like

$$h_{x_0,\gamma}^-(x) = \left(1 - \tanh(\gamma(x - x_0))\right),$$
$$h_{x_0,\gamma}^+(x) = 1 - h_{x_0,\gamma}(x), \ x \in [-\pi, \pi], \ \gamma > 0$$

it is possible to take advantage of the oscillations and get rid of them at the same time. To be able to effectively use one-sided localization one needs to estimate the location of the singularity. This can be done by using left- and right-sided (local) δ-sweeps and observing their difference. It is large where left and right limit do not coincide. The results obtained by approximating

$$z(x) := \begin{cases} \pi + x, & x \le \frac{1}{\sqrt{2}} \\ x - \pi, & x > \frac{1}{\sqrt{2}} \end{cases}$$

are depicted in Figure 2. $n = 129$ equidistant discretization points are used for $(-\pi, \pi)$ in order to perform sweeps at $N = 257$ points (midpoints of the original intervals are added). The choice of the discontinuity point makes sure it is never a grid point. As to the other parameters

$$\alpha = \frac{10}{\pi}, \ \beta = \frac{\pi}{5}, \ \gamma = 100.$$

A global δ-sweep produces z_{128}^b whereas a local δ-sweep with one-sided bias within distance $\pi/5$ of the singularity yields z_{128}^g from the discrete information contained in

$$z_{64} = \left(z(\frac{j}{64}\pi)\right)_{j=-64,\ldots,64}.$$

The definition of test function is adapted to the periodic case in that (5.1) is still used but (5.2) is modified to

$$\psi(x, x_0, \beta) =$$
$$\Big[\min\Big(\max(x, x_0) - \min(x, x_0), 2\pi - \max(x, x_0) + \min(x, x_0) \Big) \Big] / \beta$$
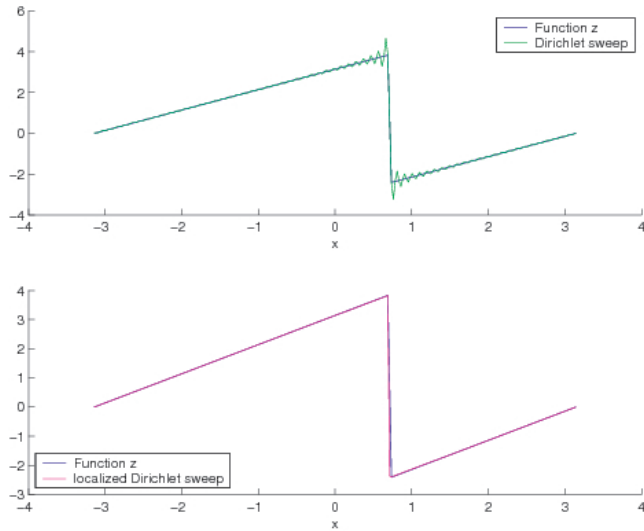
which is the periodic distance function.

FIGURE 2. "Good and bad" approximations z_{128}^g and z_{128}^b of z.

Both approximations are equivalent in the sense of distributions, that is, in the sense that z_{128}^b and z_{128}^g both determine perfectly valid and equivalent discretization families for the distribution z.

It is also observed in [5] that Gibbs phenomenon is observed when developing smooth functions with finitely many jump discontinuities in the more general framework of Gegenbauer polynomials. It follows from distributional perspective adopted here that oscillations are to be expected whenever approximating distributions by means of bases of smooth (oscillatory) functions and that oscillations, at the discrete level at least, are the nonlocal way in which local singular behavior manifests itself.

6. Oscillations in the wave equations

Finally an example is presented where a distribution or its approximation is not the starting point but rather where it appears as the solution of an equation. The example will demonstrate that the oscillations observed when solving the

wave equation for discontinuous data are an approximation to the exact solution provided they are interpreted in the sense of distributions. A standard test problem in this context is provided by the propagation of box waves. Here the natural problem of propagating Dirac impulses is chosen. In other words, the fundamental solution (Riemann function) of the wave operator will be computed numerically.

Consider the one dimensional wave equation

$$
\begin{cases}
u_{tt} - u_{xx} = f(t, x), & t \in \mathbb{R}, \; x \in (0, 1) \\
u(t, 0) = u(t, 1) = 0, & t \in \mathbb{R} \\
u(0, x) = u_0(x), \; u_t(0, x) = u_1(x), & x \in [0, 1].
\end{cases}
\tag{6.1}
$$

To compute the fundamental solution one needs to solve (6.1) for

$$
\bigl(f, u_0, u_1\bigr) = \bigl(0, 0, \delta_y\bigr), \; y \in (0, 1).
\tag{6.2}
$$

Taking the analytical view point first, let

$$
A : \mathrm{dom}(A) \subset \mathrm{L}_2(0, 1) \longrightarrow \mathrm{L}_2(0, 1)
$$

be the operator defined through

$$
\mathrm{dom}(A) = \mathrm{H}^2(0, 1) \cap \mathrm{H}_0^1(0, 1), \; Au = -u_{xx}, \; u \in \mathrm{dom}(A)
$$

A is a positive definite self-adjoint operator with compact resolvent and therefore admits a calculus (cf. [9]). With it one can define

$$
A^{-1/2} \sin(tA) \in \mathcal{L}\bigl(\mathrm{L}_2(\Omega), \mathrm{L}_2(\Omega)\bigr) \text{ for every } t \in \mathbb{R}
\tag{6.3}
$$

by

$$
A^{-1/2} \sin(tA)u = \sum_{k=1}^{\infty} \sin(kt)\langle u, e_k\rangle e_k
$$

for $e_k = \sin(k\pi\cdot)$. Unfortunately the Dirac distribution $\delta_y \notin \mathrm{L}_2(\Omega)$ and the functional calculus approach breaks down. The distributional framework, however, comes to rescue and one has

$$
A^{-1/2} \sin(tA)\delta_y = \sum_{k=1}^{\infty} \frac{1}{k} \sin(kt)\langle \delta_y, e_k\rangle e_k = \sum_{k=1}^{\infty} \frac{1}{k} \sin(kt) \sin(k\pi y) \sin(k\pi x) \tag{6.4}
$$

converges in the sense of distributions and represents the Riemann function. When numerically solving the above equation by any consistent scheme, one expects to obtain some discrete approximation to (6.4) and, consequently, oscillations. Finding a numerically acceptable solution therefore becomes the problem of filtering the information out in the spirit of Subsection 5.1. Since the propagation u_t of the initial Dirac impulse $u_t(0) = \delta_y$ is of interest, the problem is reformulated into the wave equation with initial condition

$$
u(0, x) = \delta_y, \; u_t(0, x) = 0.
\tag{6.5}
$$

The analytical expression for the solution then reads

$$\sum_{k=1}^{\infty} \cos(k\pi t)\sin(k\pi y)\sin(k\pi x)$$

$$= \sum_{k=1}^{\infty} \frac{1}{2}\big[\sin(k\pi(y-t)) + \sin(k\pi(y+t))\big]\sin(k\pi x)\,. \quad (6.6)$$

The discretization is chosen as in Example 2, that is , a discretization with n equidistant points x_j^n, $j = 1,\ldots n$ is used. In addition the spatial differential operator is discretized by means of a centered second order finite difference approximation to $-\partial_{xx}$. The same is done for ∂_{tt}. Then the scheme consists in marching implicitly in time. More specifically, if

$$u^m = (0, u_2, \ldots, u_{n-1}, 0)$$

is the solution vector at time mdt, then

$$u^{m+1} = \left(1 + \frac{dt^2}{dx^2} A^n\right)^{-1}\left(2u^m - u^{m-1}\right) \quad (6.7)$$

where A^n is defined by

$$A(j,k) = \begin{cases} 2\,, & |j-k| = 0\,,\ 1 < j,k < n\,, \\ -1\,, & |j-k| = 1\,,\ 1 < j,k < n\,, \\ 1\,, & j = k = 1\,,\ j = k = n \\ 0\,, & \text{otherwise}\,. \end{cases}$$

As to the initial conditions it is chosen

$$u^0 = e_{j_0^m}/dx\,,\ u^1 = 0\,,\ j_0^m = 13, 33\,,\ m = 32, 64\,.$$

The solution is computed up to time $t = 0.13$ as to make sure that distributions which live on a grid point are avoided. To ensure convergence a small time step $dt = t/10^4$. Figure 3 show the solutions u_n^{10000} obtained for $n = 32, 64$. The functions $S_{n,100}$, $n = 32, 64$ obtained by using a $\varphi_{10,0.1}$-sweep at $N = 100$ equidistant points are depicted in Figure 4. The peaks are located at

$$x_{32}^- = 0.36\,,\ x_{32}^+ = 0.62 \text{ and } x_{64}^- = 0.37\,,\ x_{64}^+ = 0.63$$

and the maxima are

$$M_{32}^- = 0.4952\,,\ M_{32}^+ = 0.4954 \text{ and } M_{64}^- = 0.4911\,,\ M_{64}^+ = 0.4979\,,$$

respectively.

This shows that the location, type and intensity of the singularity can be computed approximately by means of the oscillatory finite difference approximations to (6.5) to obtain

$$u(t, \cdot) \approx M_n^- \delta_{x_n^-(t)} + M_n^+ \delta_{x_n^+(t)}\,. \quad (6.8)$$
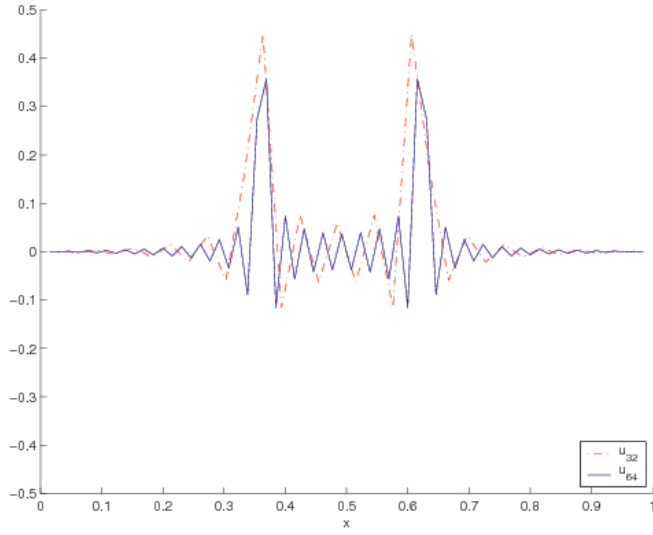
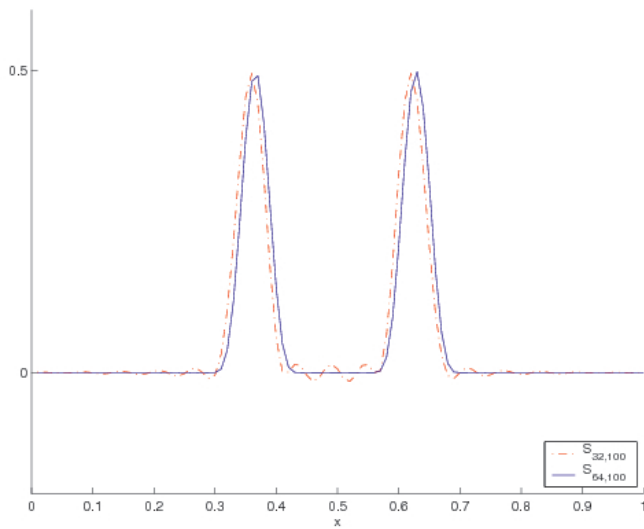FIGURE 3. The finite difference solutions u_n, $n = 32, 64$.

FIGURE 4. Test function sweeps $S_{n,100}$ of u_n^m, $n = 32, 64$, $m = 10^4$.

Remark 6.1. In this example it would have been possible also worked in the space
$$H^{-1}(0,1) = \left(H_0^1(0,1)\right)'$$

in order to look for the mild solution (cf. [6] for a definition) of the weak formulation of (6.1) with (6.5) as a function $u : \mathbb{R} \to \mathrm{H}^{-1}(0,1)$ satisfying

$$\begin{cases} \langle u_{tt}, \varphi \rangle + a(u, \varphi) = 0, & \varphi \in \mathrm{H}_0^1(0,1), \\ u(0) = \delta_y, \, u_t(0) = 0, & t = 0 \end{cases}$$

where $a : \mathrm{H}_0^1(0,1) \times \mathrm{H}_0^1(0,1) \to \mathbb{R}$ is defined through

$$a(u, \varphi) = \int_0^1 u_x(\xi) \varphi_x(\xi) \, d\xi \,.$$

The above discussion then shows that the numerical scheme automatically produces the mild weak solution of (6.5) and that it needs to be interpreted as such. A convergence analysis would reveal that the formal order of the scheme employed is indeed realized if understood in the sense of distributions.

It has been observed (cf. [1, 2] and the overview and references in the second paper) that the accuracy of "weak" solutions obtained by finite element methods can be improved by post-processing techniques. This is possible since the solution has a higher degree of accuracy in negative Sobolev norms. That is a further manifestation of the general point made in this paper albeit in the context of normed spaces and can be understood and explained by the framework proposed here.

7. Conclusion

In this paper it is shown how the theory of distribution can naturally be linked to and used in the discrete world. Actually, even more is true. The use of the natural duality structure existing between distributions and test functions in the discrete world paves the way to a systematic understanding of numerical techniques and the treatment of problems in which distributions are natural key players. It turns out that a central role in this approach is played by smooth orthonormal bases (with respect to $\mathrm{L}_2(\Omega)$) which can be used to effectively approximate any compactly supported distribution. The discretization obtained can be viewed as a kind of *abstract spectral methods*. Bases consisting of functions with only a finite degree of smoothness m only allow for approximating distributions of finite order at most m. In particular, bases consisting of merely $\mathrm{L}_2(\Omega)$ functions can only approximate $\mathrm{L}_2(\Omega)$ functions. Many methods have been developed which can be found in the literature (a small part of which is cite here) which take implicit advantage of the structure made explicit in this paper.

References

[1] J.H. Bramble and A.H. Schatz. Higher order local accuracy by averaging in the finite element method. *Math. Comp.*, 31:94–111, 1977.

[2] B. Cockburn, M. Luskin, Chi-Wang Shu, and E. Süli. Enhanced accuracy by post-processing for finite element methods for hyperbolic equations. *Math. Comp.*, 72, 2003.

[3] R. E. Edwards. *Functional Analysis.* Dover, New York, 1965.

[4] A. Gelb and E.Tadmor. Spectral Reconstruction of Piecewise Smooth Functions from their Discrete Data. *Mathematical Modelling and Numericical Analysis*, 36(2):155–175, 2002.

[5] D. Gottlieb and Chi-Wang Shu. On the Gibbs Phenomenon and its Resolution. *SIAM REV.*, 39(4):644–668, 1997.

[6] A. Pazy. *Semigroups of Linear Operators and Application to Partial Differential Equations.* Springer Verlag, New York, 1983.

[7] L. Schwartz. *Mathematics for the physical sciences.* Hermann, Paris, 1966.

[8] L. Schwartz. *Théorie des Distributions.* Hermann, Paris, 1966.

[9] V. S. Sunder. *Functional Analysis: Spectral Theory.* Birkhäuser, Basel, 1991.

[10] M. E. Taylor. *Partial Differential Equations. Basic Theory.* Springer-Verlag, New York, 1996.

[11] F. Treves. *Topological Vector Spaces, Distributions and Kernels.* Academic Press, New York, 1967.

Patrick Guidotti
103 Multipurpose Science and Technology Building
Department of Mathematics
University of California, Irvine
Irvine, CA 92697-3875
USA
e-mail: `gpatrick@math.uci.edu`