UNIVERSITY OF WISCONSIN-MADISON

CENTER FOR THE MATHEMATICAL SCIENCES

Asymptotically Optimal Approximation and Numerical Solutions of Differential Equations

Martin D. Buhmann
Mathematik Departement
ETH Zentrum,
8092 Zürich
Switzerland
mdb@math.ethz.ch

Charles A. Micchelli
Department of Mathematical Sciences
IBM T.J. Watson Research Center
P.O. Box 218
Yorktown Heights NY 10598, USA
cam@watson.ibm.com

Amos Ron
Computer Science Department
University of Wisconsin-Madison
1210 West Dayton Street
Madison, Wisconsin 53706, USA
amos@cs.wisc.edu

Technical Summary Report # 97–03
October 1996

ABSTRACT

Given finite subset $J \subset \mathbb{R}^n$, and a point $\lambda \in \mathbb{R}^n$, we study in this paper the possible convergence, as $h \to 0$, of the coefficients in least-squares approximation to $f(\cdot + h\lambda)$ from the space spanned by $(f(\cdot + hj)_{j\in J}$. We invoke the 'least solution of the polynomial interpolation problem' to show that the coefficient do converge for a generic J and λ, provided that the underlying function f is sufficiently smooth. Moreover, in certain cases (e.g., in the case $J \cup \lambda$ for a rectangular mesh), the limit of the least-squares coefficients are shown to be independent of f, and are characterized by their polynomial accuracy. Finally, we employ a different argument to show that convergence of the least squares coefficients occurs also for a certain class of functions which are not "sufficiently smooth".

The above study is relevant to the problem of selecting an optimal differencing scheme for solving PDE's, a connection that is briefly discussed as well.

This work was supported by the National Science Foundation under Grants DMS-9224748 and DMS-9626319, and by the U.S. Army Research Office under Contract DAAH04-95-1-0089.

Asymptotically Optimal Approximation and Numerical Solutions of Differential Equations

Martin D. Buhmann
Charles A. Micchelli
Amos Ron

1. Introduction

An optimal finite difference method for the numerical solution of the Cauchy problem for a given partial differential equation is, by definition, the scheme that minimises the local truncation error after one step. In this paper we conduct a study of certain extremal problems that are closely related to optimal finite difference schemes for finding numerical solutions of such problems. For relevant general information on difference methods connected with the Cauchy problem cf., e.g., (Iserles and Strang, 1983).

Consider now the concrete problem of finding an optimal finite difference method for approximating the solution of a well-posed Cauchy problem using the linear constant coefficient differential equation

$$\begin{aligned} u_t(x,t) &= p(D_x)u(x,t), \quad t \geq 0, \\ u(x,0) &= f(x), \qquad x \in \mathbb{R}, \end{aligned}$$

where p is a polynomial and $D_x = \frac{\partial}{\partial x}$. It is shown in (Micchelli and Miranker, 1973a, 1973b) that the optimal method is computable in terms of the optimal finite difference method for the advection partial differential equation

$$u_t = u_x, \qquad t \geq 0, \quad x \in \mathbb{R},$$

whose solution is $u(x,t) = f(x+t)$ with the appropriate ranges of x and t. If h denotes the mesh increment in the x-direction and λh is the mesh increment in the t direction, λ being the Courant number, then

$$f(x+\lambda h) - \sum_{j\in J} c_j f(x+jh), \qquad x \in \mathbb{R}, \tag{1.1}$$

is the local truncation error after one time step corresponding to a generic finite difference scheme that uses the pointset J at the backward level of time. This means that our goal can be accomplished by solving a corresponding least-squares minimisation problem which is the univariate case ($n = 1$) of Problem A below. In the statement of the problem, we use the notation E^t for the shift operator, i.e.

$$E^t : f \mapsto f(\cdot + t). \tag{1.2}$$

Problem A. *Given $f \in L^2(\mathbb{R}^n)$, a finite set $J \subset \mathbb{R}^n$, a point $\lambda \in \mathbb{R}^n$, and $h > 0$, minimise the error function*

$$\|E^{h\lambda} f - \sum_{j \in J} c_j E^{hj} f\|_{L^2(\mathbb{R}^n)} \tag{1.3}$$

over all $c := (c_j)_{j \in J} \in \mathbb{C}^J$. The optimal coefficients are denoted here and hereafter by $c(\lambda, h) := (c_j(\lambda, h))_{j \in J}$.

We can therefore view the stated problem as the problem of approximating the $h\lambda$-translate of some given function f from the finite dimensional space

$$S_h^J(f) := \operatorname{span}\{E^{jh} f \mid \ j \in J\},$$

J being a given finite subset of $\mathbb{R}^n$. We think of the function f as given to us (the initial value of the above partial differential equation); on the other hand, the choice of the set J and the scaling parameter h are within our discretion.

Since the space $S_h^J(f)$ is finite dimensional, it is seemingly feasible to find the exact solution to Problem A, i.e., the least squares approximation to $E^{h\lambda} f$ from $S_h^J(f)$. This, however, requires the recomputation of the approximant whenever f or h are changed. Furthermore, since the least squares solution is expressed via the basis $\{E^{jh} f \mid j \in J\}$ of $S_h^J(f)$ (i.e. we compute the coefficients with respect to that basis), the deteriorating condition number of that basis, as $h \to 0$, may lead to loss of significance through numerical instability of the process. In addition, knowing the theoretical background of such an approximation should help us in determining the size of h to be used, and the configuration and cardinality of the set J to be chosen. A large set J will make the computation of the least square solution computationally prohibitive, while a small set J may force us to select a scaling parameter h that is too small.

This problem has already been dealt with in the literature if $f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ and for a positive α

$$\lim_{x \to \pm\infty} |\widehat{f}(x)||x|^{\alpha + \frac{1}{2}} = \text{const} \neq 0 \tag{1.4}$$

in the above-mentioned two papers by Micchelli and Miranker. Specifically, it is proved there that, upon assuming (1.4), the optimal coefficients $c_j(\lambda, h)$ converge, for every $j \in J$, to a limit $c_j(\lambda) := \lim_{h \to 0} c_j(\lambda, h)$; that limit is referred to as the *principal part* of the intermediate coefficient $c_j(\lambda, h)$. This admits employing the h-independent principal parts as a suitable alternative to the optimal coefficients in (1.1). Moreover, it is the principal part of the optimal scheme which determines its stability.

Although then the local truncation error is not necessarily minimal, we will show that it is "near optimal" in a certain sense, and it allows for a simpler, easier to compute finite difference scheme. In other words, while the coefficients of the best approximation to $E^{h\lambda} f$ from $S_h^J(f)$ may exhibit undesired numerical behaviour as $h \to 0$, one explicitly identifies cases where the coefficients of the best approximation each converge to a limit (defined as the principal part of the coefficient), and considers the approximation scheme when the $c_j(\lambda, h)$ are replaced by their principal parts.

In extending these univariate results we have two approaches in this paper, one involving a multivariate generalisation of (1.4) which is also more general even in one dimension. It leads to so-called radial basis function interpolation, and the principal parts turn out to be Lagrange functions for this interpolation with respect to the set J. Another approach involves multivariate polynomial interpolation on the set J.

Indeed, when we use that second *ansatz*, the message of this paper concerning Problem A becomes strikingly simple, and so we are going to describe it first. The statements below require J to be in a *total degree* configuration. This notion is defined near the end of Section 2. Here, we mention that any pointset J in $\mathbb{R}$ satisfies the total degree assumption, and that, in two dimensions, 3-sets J whose points are not colinear, as well as 6-sets J whose points do not lie on a conic, are of total degree. The *degree* $m(J)$ of a total degree configuration J is also defined in Section 2. We remark here that $m(J) = |J| - 1$ in case $J \subset \mathbb{R}$, and that the degrees of the two-dimensional 3-sets and 6-sets discussed above are 1 and 2 respectively.

One of the main results of this paper, that invokes the polynomial interpolation approach, is as follows. In its statement, and elsewhere in the paper, we use the notation e_j, $j \in \mathbb{R}^n$, for the exponential with frequency j,

$$e_j : t \mapsto e^{ij\cdot t},$$

and

$$\Pi_m$$

for the space of all n-variate polynomials of degree $\le m$.

Theorem 1. *Let J be a total degree subset of $\mathbb{R}^n$ of degree $m(J)$. Then, for every $\lambda \in \mathbb{R}^n$, there is a sequence $c(\lambda) \in \mathbb{C}^J$, that depends only on λ and J, with the following properties:*

1. *For every $\alpha > m(J)$, and for every function in the Sobolev space $W_2^\alpha(\mathbb{R}^n)$, the optimal coefficients $c_j(\lambda, h)$, $j \in J$, of Problem A converge each to $c_j(\lambda)$.*
2. *For every $\alpha \le m(J) + 1$, and for every $f \in W_2^\alpha(\mathbb{R}^n)$, the $L^2(\mathbb{R}^n)$-error in the approximation scheme*
$$E^{h\lambda} f \approx \sum_{j\in J} c_j(\lambda) E^{hj} f$$
is $O(h^\alpha)$ as $h \to 0$.

Furthermore, the sequence $c(\lambda)$ is characterised by its polynomial accuracy: it is the unique sequence supported on J that satisfies

$$p(\lambda) = \sum_{j\in J} c_j(\lambda) p(j), \quad \forall p \in \Pi_{m(J)}.$$

The analysis that leads to Theorem 1 is given in Sections 3 and 4. We mention that the convergence asserted in (1.) is valid for configurations J that are more general than the "total degree" mentioned here; however, in these more general setups, either the limit coefficients are not universal, i.e., they may depend on f, and/or the convergence is proved for only certain, specific λ. For example, convergence to universal limits occurs if $J \cup \lambda$ consists of the vertices of a rectangular mesh.

Thus, for a large collection of functions and for quite general configurations J, the optimal coefficients of Problem A converge to an f-independent sequence $c(\lambda)$. Further, that sequence "works well" for other functions too (even though for these other functions the optimal coefficients may not converge at all), in the sense that the error $\|E^{h\lambda} - \sum_{j\in J} c_j(\lambda)e_j\|_{L^2(\mathbb{R}^n)}$ decays at rates that are related to the smoothness of f. This leads us naturally to examining also f-independent h-independent schemes of the form

$$E^{h\lambda} \approx \sum_{j\in J} c_j E^{hj} f.$$

We call schemes such as the above "near optimal" if they provide approximation rates like those in (2.) of Theorem 1.

The paper is laid out as follows. In Section 2, we review relevant facts concerning the *least solution of the polynomial interpolation problem* (de Boor, Ron, 1990). That least solution turns out to be the main tool in our study of Problem A for "sufficiently smooth" functions f that is carried out in Section 4. In that section, Theorem 1 is proved, together with some more general results. Before, in Section 3, we study the problem of near optimal schemes. In Section 5, we consider the convergence of the optimal coefficients in case the underlying function f is not sufficiently smooth for the application of Theorem 1. We identify situations when the Fourier transform of f is asymptotically homogeneous at ∞ (similarly to (1.4)) and prove that the optimal coefficients converge, regardless of the configuration of J and/or the choice of λ. As stated before, the limit coefficients are identified as the Lagrange functions of certain interpolation problems that involve radial basis functions (Buhmann, 1993, Dyn, 1989, Micchelli, 1986).

2. The least solution of the polynomial interpolation problem

We briefly discuss here the least solution of the polynomial interpolation problem of (de Boor, Ron, 1990), a tool that we require in our analysis in Section 4, and to a lesser extent in our analysis of Section 3.

Given a finite $J \subset \mathbb{R}^n$, and a polynomial space P in n variables, we say that P **is correct** for interpolation on J if every data given on J can be interpolated by a unique $p \in P$. For each J, there are of course many polynomial spaces P that interpolate correctly on J. In (de Boor, Ron, 1990, 1992a, 1992b) a correct polynomial space Π_J, the least solution, of least possible degree and various other desired properties was constructed. That construction applies to any J, and is fit to the problem we discuss in this paper: the limit of the optimal coefficients (the principal parts) for sufficiently smooth functions f will be expressed in terms of this least solution. We review below some of its basic features and refer the reader to the above-mentioned references for further discussions.

Let Exp_J be the exponential space

$$\mathrm{Exp}_J := \mathrm{span}\{e_j \mid j \in J\}.$$

Since each $g \in \mathrm{Exp}_J$ is entire, it can be written uniquely as a convergent sum

$$g = g_0 + g_1 + g_2 + \dots,$$

with each g_m being a *homogeneous* polynomial of degree m. One sets

$$g_\downarrow$$

(read "g least") to be the non-zero polynomial g_m of *least* degree in the above expansion. The **least solution** of the polynomial interpolation problem is then the *homogeneous* polynomial space

$$\Pi_J := \mathrm{span}\{g_\downarrow \mid g \in \mathrm{Exp}_J\}.$$

The space Π_J is correct for interpolation on J. In fact, it is of minimal degree among all such polynomial spaces: if there exists any polynomial space P that is correct for J and contains the space Π_α (of all polynomials of degree $\leq \alpha$), then Π_J contains Π_α, as well. Also, Π_J coincides with standard choices of correct spaces, in cases such choices exist, for example when J forms a rectangular grid.

We associate now with the set J two important parameters: its *minimal degree* (or *accuracy*) $m(J)$, and its *maximal degree* $M(J)$; both notions are essential in our analysis of Problem A.

Definition 1. *Let J be a finite subset of $\mathbb{R}^n$, Π_J its associated least polynomial space.*

1. *$m(J)$ is the maximal integer m for which $\Pi_m \subset \Pi_J$.*
2. *$M(J)$ is the minimal m for which $\Pi_J \subset \Pi_m$.*

Example. In one dimension, the least solution of any m-set J is the space Π_{m-1}. Thus, we always have $m(J) = M(J) = |J| - 1$. In more than one dimension, an equality $m(J) = M(J) =: m$ implies that $\Pi_J = \Pi_m$, and is possible only in case the cardinality $|J|$ matches the dimension of Π_α for some α. However, with $k := \dim \Pi_m$, for some m, the equality $m(J) = M(J) = m$ holds for a generic k-set in $\mathbb{R}^n$.

The above example motivates the following definition:

Definition 2. *We say that $J \subset \mathbb{R}^n$ is*

1. *in* **general position** *if $M(J) \leq m(J) + 1$ and*
2. *of* **total degree** *if $m(J) = M(J)$.*

We note that "general position" is the generic case. However, various important configurations for J, such as the vertices of a rectangular grid, are not in general position. "Total degree" occurs when J is in general position, and further its cardinality matches the dimension of some Π_m; for example, in two dimensions, three points that are not colinear, and six points that do not lie on a conic are of total degree. In contrast, a set consisting of four or five points (still in $\mathbb{R}^2$) cannot be of total degree. Such a set will be in general position, though, unless all its points lie on one line.

3. Near optimality

We say that "$S_h^J()$ provides approximation order α to $f \in L^2(\mathbb{R}^n)$", if $\mathrm{dist}(E^{h\lambda} f, S_h^J(f)) = O(h^\alpha)$, for every/some $\lambda \in \mathbb{R}^n$. The range of relevant λ will be clear in each context. Of course, the optimal coefficients of Problem A realize any approximation order that can be provided (after all, they are the best). However, our objective in this paper is to replace the h-dependent optimal coefficients by h-independent ones. Therefore, it seems useful to consider the following problem:

Problem B. *Assume that $S_h^J()$ provides, for some fixed J and λ, approximation order $\alpha > 0$ to all functions in some smoothness class L. Are there h-independent sequences $c \in \mathbb{C}^J$ that realize that order, i.e., that*

$$\|E^{h\lambda} f - \sum_{j \in J} c_j E^{hj} f\|_{L^2(\mathbb{R}^n)} = O(h^\alpha), \quad \forall f \in L \ ?$$

Schemes that use such sequences c as described above are called **near optimal** because they realise the approximation rate α of best approximation, but not necessarily with the same constant factor.

One of the fundamental principles of Approximation Theory is the intimate relation between the decay rate of the error in approximation schemes and the smoothness class of the function being approximated. Roughly speaking, one expects that "functions with only α derivatives" will be approximated at rates no better than α. This also explains the custom of studying rates of convergence simultaneously for all functions in the same smoothness class. For our particular problem, the close relation between the decay of the error with $h \to 0$ and the smoothness class of f is even more basic: the mere definition of "smoothness" via the moduli of smoothness notion (cf., e.g., DeVore, Lorentz, 1993, p. 44) shows that smoothness can be defined, hence interpreted, as the ability to approximate a function well by close-by translates of it.

Theorem 1 provides simultaneous answers to both Problems A and B: it shows that for smooth functions the optimal coefficients of Problem A converge to their universal limits and that for functions of lesser smoothness these universal coefficients provide the expected approximation orders. This result is stated with respect to Sobolev spaces. However, it is still valid if we replace these spaces by the larger *Besov spaces* that we now define. Let

$$\Omega_j := \{t \in \mathbb{R}^n \mid 2^{j-1} \leq \|t\| < 2^j\}$$

and

$$\Omega_0 := \{t \in \mathbb{R}^n \mid \|t\| < 1\},$$

where $\|\cdot\|$ denotes the Euclidean norm on $\mathbb{R}^n$. For any $f \in L^2(\mathbb{R}^n)$ we let $a_j(f) := \|\|\cdot\|^\alpha \widehat{f}\|_{L^2(\Omega_j)}$. The **Besov space** $B_\infty^\alpha(L^2(\mathbb{R}^n))$ is the space that contains all f with uniformly bounded $a_j(f)$:

$$\|f\|_{B_\infty^\alpha(L^2(\mathbb{R}^n))} := \sup_{j \geq 0} \|\|\cdot\|^\alpha \widehat{f}\|_{L^2(\Omega_j)}.$$

Theorem 2. *Let $J \subset \mathbb{R}^n$ be given, let $\lambda \in \mathbb{R}^n$, and $c = (c_j)_{j \in J} \in \mathbb{C}^J$. Suppose that, with*

$$E_c := e_\lambda - \sum_{j \in J} c_j e_j,$$

$|E_c(t)| = O(\|t\|^m)$ as $\|t\| \to 0$ for some positive integer m. Then, with

$$A_h f := \sum_{j \in J} c_j E^{hj} f,$$

1. *for every $f \in W_2^m(\mathbb{R}^n)$, and $h \to 0$,*

$$\|E^{h\lambda} f - A_h f\|_{L^2(\mathbb{R}^n)} = O(h^m).$$

2. *for every $\alpha < m$, every $f \in B_\infty^\alpha(L^2(\mathbb{R}^n))$, and $h \to 0$,*

$$\|E^{h\lambda} f - A_h f\|_{L^2(\mathbb{R}^n)} = O(h^\alpha).$$

Proof: Measuring the error in the Fourier domain, one has

$$\|E^{h\lambda} f - A_h f\|_{L^2(\mathbb{R}^n)} = \tfrac{1}{(2\pi)^n} \|\widehat{f} E_c(h\cdot)\|_{L^2(\mathbb{R}^n)} \leq$$

$$\tfrac{1}{(2\pi)^n} \|(h\|\cdot\|)^m \widehat{f}\|_{L^2(\mathbb{R}^n)} = \|f\|_{W_2^m(\mathbb{R}^n)} h^m,$$

which proves (1.). For the proof of (2.), we assume without loss of generality that $h = 2^{-k}$ for some k. We estimate

$$\int |E_c(ht)|^2 |\widehat{f}(t)|^2 \, dt. \tag{3.1}$$

We divide the integral into its range of integration over $\|t\| > 2^k = h^{-1}$ and the remaining part. We deal with the former part first. Here, since $f \in B_\infty^\alpha(L^2(\mathbb{R}^n))$, we may write, for a generic positive constant C,

$$\int_{\Omega_j} |\widehat{f}(t)|^2 \, dt \leq 2^{-2(j-1)\alpha} \int_{\Omega_j} \|t\|^{2\alpha} |\widehat{f}(t)|^2 \, dt \leq C 2^{-2j\alpha}.$$

Summing over $j \geq k+1$, we obtain that

$$\int_{\|t\| > 2^k} |E_c(t)|^2 |\widehat{f}(t)|^2 \, dt \leq C \int_{\|t\| > 2^k} |\widehat{f}(t)|^2 \, dt \leq C 2^{-2k\alpha} = O(h^{2\alpha}).$$

For $j \leq k$, we have

$$\begin{aligned} \int_{\Omega_j} |E_c(t/2^k)|^2 |\widehat{f}(t)|^2 \, dt &\leq C 2^{-2km} \int_{\Omega_j} \|t\|^{2m} |\widehat{f}(t)|^2 \, dt \\ &\leq C 2^{-2km} 2^{2j(m-\alpha)} \int_{\Omega_j} \|t\|^{2\alpha} |\widehat{f}(t)|^2 \, dt. \end{aligned}$$

Invoking the fact that $f \in B_\infty^\alpha(L^2(\mathbb{R}^n))$, we can sum over $j = 0, 1, \ldots, k$, to obtain the required $O(h^{2\alpha})$ bound. $\square$

Discussion. Expressed with the least term of a smooth function (as defined in Section 2), the requirement in Theorem 2 concerning E_c is that $\deg E_{c\downarrow} \ge m$. If $m \le m(J)+1$, with $m(J)$ the accuracy of J, then, for each $\lambda \in \mathbb{R}^J$, this requirement can be fulfilled by a suitable choice of c: for any $J \subset \mathbb{R}^n$, and any polynomial p of degree at most $m(J)$, there exists $g \in \mathrm{Exp}_J$ whose Taylor expansion up to degree $m(J)$ yields p. For higher values of m, such a condition may be satisfied only for very special values of λ that lie in the zero sets of certain polynomials. In any event, if $m > M(J)+1$, then the aforementioned condition cannot hold for any $\lambda \notin J$.

Proposition 1. *Given $J \subset \mathbb{R}^n$, the spaces $(S_h^J())_h$ provide approximation order*

1. *$m(J)+1$ to the Sobolev space $W_2^{m(J)+1}(\mathbb{R}^n)$ and*
2. *for $\alpha < m(J)+1$, α to the Besov space $B_\infty^\alpha(L^2(\mathbb{R}^n))$.*

The fact that we do not get in the above proposition approximation order $m(J)+1$ for the entire Besov class is expected. Indeed, the definition of Besov spaces, say in the univariate case, in terms of divided differences (that is intimately related to the approximation problem we are considering here) involve, for an integer smoothness parameter k, a difference operator that is supported on $k+2$ points. Difference operators that involve only $k+1$ points can be used, however, to define the smaller smoothness space $\mathrm{Lip}_k(\mathbb{R})$. One can therefore expect the space $\mathrm{Lip}_{m(J)+1}(\mathbb{R})$ to be the *saturation class* for our problem (in the univariate case).

We show finally that a better approximation rate cannot be obtained for all functions from our Besov space of order α. We define a specific f in that space via its Fourier transform, namely

$$\widehat{f}(t) = (1-\chi(t))\|t\|^{-\alpha-n/2}, \qquad t \in \mathbb{R}^n.$$

Here, χ is the characteristic function of the unit ball. This is in the Besov space because for positive j

$$\int_{\Omega_j} \|t\|^{2\alpha}|\widehat{f}(t)|^2\,dt = \int_{\Omega_j} \|t\|^{-n}\,dt = \text{const}.$$

Note that f is not contained in Sobolev space W_2^α.

A simple change of variables shows that α is the best order one can achieve:

$$\mathrm{dist}(E^{h\lambda}f, S_h^J(f)) = h^{n/2}\,\mathrm{dist}(E^\lambda f(h\cdot), S_1^J(f(h\cdot))) \ge h^\alpha\,\mathrm{dist}(E^\lambda f, S_1^J(f)).$$

Thus, the rate of approximation is α, unless $E^\lambda f$ happened to lie in $S_1^J(f)$, i.e. unless $\lambda \in J$.

4. Optimal Approximation and multivariate polynomial interpolation

Given a finite $J \subset \mathbb{R}^n$, we consider Problem A under the assumption that the underlying function f is "sufficiently smooth", a notion that we make precise soon. Given $\lambda \in \mathbb{R}^n$, we show that for each $j \in J$, there exists an integer $k_j := k_j(J,\lambda)$, so that, for all sufficiently smooth f, the sequence $h^{-k_j}(c_j(\lambda,h))_h$ converges to a finite limit as h tends to zero. Furthermore, under certain assumptions on J and λ, the limit is shown to be f-independent.

Given a finite set $J \subset \mathbb{R}^n$, we assume throughout this section (with the exception of Theorem 4) that our function f of Problem A lies in the Sobolev space $W_2^{M(J)+\varepsilon}(\mathbb{R}^n)$, for some $\varepsilon > 0$. Here, $M(J)$ is the maximal degree of J as defined in Section 2.

EXAMPLE. $n = 1$: If $J \subset \mathbb{R}$, then $\Pi_J = \Pi_{m-1}$, with m the cardinality of J. We conclude that our smoothness class in one variable is slightly smaller than $W_2^{m-1}(\mathbb{R})$.

As we will see momentarily, for a given smooth f, the convergence of the optimal coefficient $c_j(\lambda, h)$ depends critically on a certain connection among the three least spaces Π_J, $\Pi_{(J\backslash j)\cup\lambda}$ and $\Pi_{J\cup\lambda}$. All these spaces are homogeneous, each one of them is a superspace of $\Pi_{J\backslash j}$. Thus, we can think of each as constructed from $\Pi_{J\backslash j}$ by appending first to that space one (for J and $(J\backslash j) \cup \lambda$) or two (for $J \cup \lambda$) *homogeneous* polynomials, and then taking the span of the so-obtained polynomial set. The critical information in this regard is the *degree* of the polynomials appended in such a procedure. This motivates the

Definition 3. *Let K be a finite subset of $\mathbb{R}^n$ and $k \in \mathbb{R}^n\backslash K$. We denote by*

$$d(K, k)$$

the degree of any homogeneous polynomial p that satisfies

$$\Pi_{K\cup k} = \operatorname{span}\{\Pi_K \cup p\}.$$

In general, the value of $d(K, k)$ depends on subtle relations between K and k. However, the following estimates (albeit crude ones) are valid:

$$m(K) + 1 \le d(K, k) \le M(K) + 1.$$

Lemma 1. *Let J be a finite subset of $\mathbb{R}^n$, $\lambda \in \mathbb{R}^n$, and $f \in W_2^{M(J)+\varepsilon}(\mathbb{R}^n)$, for some $\varepsilon > 0$. Let $(c_j(\lambda, h))_{j\in J}$ be the optimal coefficients of Problem A. Fix $j \in J$, and let k_j be the f-independent integer*

$$k_j = d(J\backslash j, \lambda) - d(J\backslash j, j).$$

Then the following holds:

1. *If $k_j > 0$, $c_j(\lambda, h)$ converges to 0.*
2. *If $k_j \le 0$, there exists a number $c_j(\lambda)$, which is independent of h, such that*

$$c_j(\lambda, h) = h^{k_j}(c_j(\lambda) + o(1)),$$

for $h \to 0$.

ANALYSIS. Already at this point, we can use the lemma's statement to classify the different cases of Problem A as follows:

1. The parameter k_j in the lemma is positive. Then, the coefficient $c_j(\lambda, h)$ converges to 0 for every smooth f as h tends to zero.

2. The parameter k_j in the lemma is negative. Then, the coefficient $c_j(\lambda, h)$ becomes unbounded for a generic smooth f and $h \to 0$.

3. The parameter k_j is 0, and the principal part $c_j(\lambda)$ is actually independent of f. Then, the optimal coefficients converge to *universal limits* $c_j(\lambda)$ that depend on J, λ and j, but not on f.

4. The parameter k_j is zero, but the principal parts depend, in general, on f. The optimal coefficients in Problem A converge for every smooth f to the f-dependent limit $c_j(\lambda)$.

In case (3.) above, a characterisation of the universal coefficients is desired.

In any event, the sign of k_j is important. Therefore, it is useful to observe that the space $\Pi_{J\cup\lambda}$ is obtained from $\Pi_{J\backslash j}$ by adding to the latter two homogeneous polynomials: one of degree $d(J\backslash j, j)$ and one of degree $d(J, \lambda)$. Thus, unless $k_j = 0$, it must have the value

$$k_j = d(J, \lambda) - d(J\backslash j, j).$$

While the value of $d(J\backslash j, j)$ may depend on $j \in J$, we have the obvious bound $d(J\backslash j, j) \leq M(J)$. Thus

$$k_j \geq \min\{0, d(J, \lambda)\} - M(J)\}. \tag{4.1}$$

Proof of Lemma 1. We first state and prove another lemma. All inner products appearing here and elsewhere in this section are standard L^2-inner products.

Lemma 2. *Let $f \in L^2(\mathbb{R}^n)$ and $(g_k)_{k\in J}$ be a finite collection of linearly independent real-analytic functions. Assume that $g_k\widehat{f} \in L^2$, for every $k \in J$. Then the matrix whose entries are*

$$\langle g_j\widehat{f}, g_k\widehat{f}\rangle, \quad (j,k) \in J \times J,$$

is non-singular.

Proof of Lemma 2. The non-singularity of the matrix is equivalent to the linear independence of $(g_k\widehat{f})_{k\in J}$, the latter is argued as follows. If

$$\sum_{k\in J} c_k g_k \widehat{f} = 0,$$

then $g\widehat{f} = 0$, with $g = \sum_{k\in J} c_k g_k$. This implies that g must vanish on a non-null subset of $\mathbb{R}^n$ (viz., the support of $\widehat{f}$). Hence $g = 0$ everywhere (g being real analytic). This forces $c_k \equiv 0$ since $(g_k)_k$ are linearly independent by assumption. □

We now prove Lemma 1 as follows. First, we find the coefficient sequence $c(\lambda, h)$ of the best approximation $\sum_{j\in J} c_j(\lambda, h)E^{hj}f$ to $E^{h\lambda}f$ by solving the normal equations. The corresponding Gram matrix is

$$D_h := (\langle E^{hj}f, E^{hk}f\rangle)_{j,k\in J} = (2\pi)^{-n}(\langle e_{hj}\widehat{f}, e_{hk}\widehat{f}\rangle), \tag{4.2}$$

where the right-most equality follows Parseval's formula. This Gram matrix is not singular, as an application of Lemma 2 with $g_k := e_{hk}$ shows.

Now, we fix $j \in J$, and construct a basis for Π_J as follows. First, we choose a basis $B_0 = (b_k)_{k\in(J\backslash j)}$ for $\mathrm{Exp}_{J\backslash j}$ such that (a) the transformation matrix T_0 that maps $(e_k)_{k\in(J\backslash j)}$ to B_0 is unit lower triangular, and (b) $B_\downarrow := (b_{k\downarrow})_{k\in J\backslash j}$ is a basis for $\Pi_{J\backslash j}$ (such a basis always exists and can be constructed inductively; cf. (de Boor, Ron, 1992a)). Then we extend this basis to a basis B of Exp_J by adding an additional function b, such that $b_\downarrow$ completes $B_{0\downarrow}$ to a basis $B_\downarrow$ of Π_J. Let T be the matrix that converts the basis $(e_k)_{k\in J}$ to the basis B; by normalising B if necessary, we may assume, as we do, that $\det T = 1$. Then, using (4.2), we obtain that

$$G_h := TD_hT^T = (2\pi)^{-n}(\langle b_i(h\cdot)\widehat{f}, b_k(h\cdot)\widehat{f}\rangle).$$

Due to our smoothness assumption on f, it is easy to prove that for small h

$$\langle b_i(h\cdot)\widehat{f}, b_k(h\cdot)\widehat{f}\rangle = h^{\deg b_{i\downarrow}+\deg b_{k\downarrow}}(\langle b_{i\downarrow}\widehat{f}, b_{k\downarrow}\widehat{f}\rangle + o(1)).$$

Thus,

$$\det D_h = \det G_h = h^\ell(\det G_\downarrow + o(1)),$$

where $G_\downarrow$ has entries

$$(2\pi)^{-n}\langle b_{i\downarrow}\widehat{f}, b_{k\downarrow}\widehat{f}\rangle, \quad (i,k) \in J\times J,$$

and where

$$\ell := 2\sum_{k\in J\backslash j} \deg b_{k\downarrow} + 2\deg b_\downarrow.$$

The matrix $G_\downarrow$ is non-singular, as follows from Lemma 2, when choosing (g_k) there to be $B_\downarrow$. Thus $h^{-\ell}\det D_h = \det G_\downarrow + o(1)$, $\det G_\downarrow \neq 0$.

We proceed by computing $c_j(\lambda, h)$ via Cramer's rule. The denominator is $\det D_h$. The numerator is the determinant of $D_{h,j}$ which is obtained from D_h when replacing the jth column there by

$$(2\pi)^{-n}\langle e_{hi}\widehat{f}, e_{h\lambda}\widehat{f}\rangle,\ i \in J.$$

We apply to $D_{h,j}$ the same row-operations as before, i.e., multiply that matrix by T from the left. As to column operations, we need to modify the underlying least space, since the columns now relate to the pointset $J' := (J\backslash j)\cup\lambda$, hence to the least space $\Pi_{J'}$. We construct a basis for $\Pi_{J'}$ by adding to the previous basis B_0 of $\mathrm{Exp}_{J\backslash j}$ an additional function b', such that $B' := B_0 \cup b'$ is a basis for $\mathrm{Exp}_{J'}$, while $B'_\downarrow$ is a basis for $\Pi_{J'}$. We let T_j be the matrix that transform $(e_k)_{k\in J'}$ to B'. Then,

$$G_{h,j} := TD_{h,j}T_j^T$$

has entries

$$(2\pi)^{-n}\langle b_i(h\cdot)\widehat{f}, b'_k(h\cdot)\widehat{f}\rangle, \quad (i,k)\in J\times J'.$$

As before, we thus obtain that, with

$$G_{j\downarrow} := (2\pi)^{-n}(\langle b_{i\downarrow}\widehat{f}, b'_{k\downarrow}\widehat{f}\rangle),$$

one has for small h

$$h^{-\ell_j}\det G_{h,j} = \det G_{j\downarrow} + o(1),$$

where

$$\ell_j := \deg b_\downarrow + \deg b'_\downarrow + 2 \sum_{k \in J \backslash j} \deg b_{k\downarrow}.$$

Collecting terms, we finally arrive at

$$c_j(\lambda, h) = h^{\deg b'_\downarrow - \deg b_\downarrow} a_j \left(\frac{\det G_{j\downarrow}}{\det G_\downarrow} + o(1) \right),$$

with $a_j := 1/\det T_j$. Since, by the above construction details, $\deg b_\downarrow = d(J\backslash j, j)$ while $\deg b'_\downarrow = d(J\backslash j, \lambda)$, and in view of the discussion in the next paragraph, we obtain the desired result.

In one case, the above argument may fail to go through: this is the case when $\deg b'_\downarrow > M(J)$, a case in which inner products of the form $\langle p\widehat{f}, b'_\downarrow \widehat{f} \rangle$, $\deg p = M(J)$, that necessarily appear in $G_{j\downarrow}$, may not make sense without a stronger smoothness assumption on f. However, in such a case $k_j = \deg b'_\downarrow - \deg b_\downarrow \geq \deg b'_\downarrow - M(J) > 0$, and we can modify the previous proof as follows: we extend B_0 to a basis B'' of $\mathrm{Exp}_{J'}$ such that, with b'' the function added to B_0, $\deg b''_\downarrow = M(J)$. With this degree reduction, the smoothness assumption on f allows us to invoke the previous argument (with B'' replacing B'). The new k_j is now $\deg b''_\downarrow - \deg b_\downarrow \geq M(J) - M(J) = 0$. At the same time, $B''_\downarrow$ is now a subset of $\Pi_{J\backslash j}$, hence is dependent, which forces the matrix G_j to be singular. Our previous argument thus yields for that case that $c_j(\lambda, h) = o(1)$, and the proof is now complete. □

We recall the definition of *general position* from Section 2, and remind the reader that this is the generic situation: it is proved in (de Boor, Ron, 1990) that for any integer ℓ, the sets $J \subset \mathbb{R}^n$ with cardinality ℓ that are in general position form an open and dense subset (in $\mathbb{R}^{\ell n}$). Also, note that the degrees $m(J)$ and $M(J)$ of a set J in general position are determined by its cardinality; for example $M(J)$ is the least integer m for which $|J| \leq \dim \Pi_m$. Though general position is the generic case, there are important configurations for J that are not in general position: the most notable case is that of a grid.

Corollary 1. *Let J be a finite subset of $\mathbb{R}^n$ in general position, $\lambda \in \mathbb{R}^n$, and $f \in W_2^{M(J)+\varepsilon}(\mathbb{R}^n)$, $\varepsilon > 0$. Let $(c_j(\lambda, h))_{j \in J}$ be the optimal coefficients of Problem A. Then, for each $j \in J$, the sequence*

$$h \mapsto c_j(\lambda, h)$$

converges to a finite limit $c_j(\lambda)$.

Proof: In view of Lemma 1, it suffices to show that the integer k_j there is non-negative, for each j. Since we have the estimate (4.1), we need only to prove that $d(J, \lambda) \geq M(J)$. However, the general position assumption grants us that $\Pi_{M(J)-1} \subset \Pi_J$. Since $d(J, \lambda)$ is the degree of some homogeneous polynomial which is not in Π_J, this degree is trivially $> M(J)-1$. □

Example. Let J consist of three points in $\mathbb{R}^2$ which are not colinear. Then J is in general position, hence the corollary applies. The corollary further tells us that, for some λ and j, the optimal coefficient $c_j(\lambda, h)$ tends to 0. Precisely, the penultimate proof shows that this happens if $\Pi_{(J\backslash j)\cup\lambda}$ contains a quadratic polynomial (note that $\Pi_J = \Pi_1$). However, the least space of a 3-set contains a quadratic if and only if the three points are colinear. We conclude that $c_j(\lambda, h)$ tends to 0 if λ lies on the line that goes through the two other points of J. This observation holds for *every* smooth f; here, "smooth" means lying in $W_2^\alpha(\mathbb{R}^2)$, for some $\alpha > 1$.

Suppose now that J is not in general position. First, whether or not J is in general position, it follows from (de Boor, Ron, 1990) that one can always find $j \in J$ such that $d(J\backslash j, j) = M(J)$. On the other hand, for a generic λ, the value of $d(J\backslash j, \lambda)$ will be the smallest possible, i.e., $m(J\backslash j) + 1$. For such λ and j, with k_j as in Lemma 1,

$$k_j = m(J\backslash j) + 1 - M(J) \leq m(J) + 1 - M(J).$$

This number is non-negative if and *only if* J is in general position. Thus, for J not in general position, the optimal coefficients diverge to ∞ for almost all λ and f. However, one should not mistakenly write off sets J that are not in general position. Such sets can be useful for particular values of λ. Since we consider λ as given, while J is for us to choose, we may adjust J to the value of λ. Specific examples to that extent are given in the sequel.

We now turn our attention to the case when the optimal coefficients of Problem A converge to limits that are f-independent. Sufficient conditions for that to happen are given in the next two results. The condition in the first result can be seen to be necessary, too, for the universality of the limit $c_j(\lambda)$. The second result is the main result of this section. Its condition can be shown to be necessary for the existence and universality of *all* coefficient limits.

Corollary 2. *In Lemma 1, if, for some $j \in J$, the least space $\Pi_{(J\backslash j)\cup\lambda}$ equals Π_J, then the limit $c_j(\lambda)$ of the optimal coefficients $c_j(\lambda, h)$ exists and is independent of f. Moreover, there exists a function $F \in \mathrm{Exp}_{J\backslash j}$ such that $(e_\lambda - (c_j(\lambda)e_j + F))_\downarrow \notin \Pi_J$.*

Proof: Set $J' := (J\backslash j) \cup \lambda$. Since we assume $\Pi_J = \Pi_{J'}$, we may choose in the proof of Lemma 1 the same homogeneous basis $B_\downarrow$ and $B'_\downarrow$ for this space. This shows that k_j in the lemma is 0 and the two matrices $G_\downarrow$ and $G_{j\downarrow}$ are identical. The lemma thus provides the estimate for small h

$$c_j(\lambda, h) = a_j + o(1).$$

Thus the optimal coefficient converges to an f-independent limit.

We investigate the nature of that limit with the aid of the notations and details introduced in the proof of our lemma above. We write $b = r_1 e_j + F_1$, and $b' = r_2 e_\lambda + F_2$, $F_1, F_2 \in \mathrm{Exp}_{J\backslash j}$, and assume $b_\downarrow = b'_\downarrow$. If either F_1 or F_2 is not unique, we choose them to maximise $\deg(b - b')_\downarrow$. We note that $(b - b')_\downarrow \notin \Pi_J$. Indeed, if $g = se_j + F$, $F \in \mathrm{Exp}_{J\backslash j}$, and $g_\downarrow = (b - b')_\downarrow$ then one of the following must happen: (i) $s = 0$. That contradicts the maximal choice of F_1, F_2, since $\deg(b - g - b')_\downarrow > \deg(b - b')_\downarrow$. (ii) $s \neq 0$. Then $sb - r_1 g \in \mathrm{Exp}_{J\backslash j}$, and $(sb - r_1 g)_\downarrow = sb_\downarrow = sp$ since $\deg(b - b')_\downarrow > \deg b_\downarrow$. Hence $b_\downarrow \in \Pi_{J\backslash j}$, contrary to our assumptions.

Now, for the matrices T and T_j in Lemma 1, we have that $\det T = r_1$ and $\det T_j = r_2$. Since we assume in the lemma that $\det T = 1$, we have that $r_1 = 1$. Thus, with $a_j = (\det T_j)^{-1} = 1/r_2$, the function $-a_j(b - b')$ is of the form $e_\lambda - (a_j e_j + F_3)$, with $F_3 \in \mathrm{Exp}_{J\backslash j}$, while $(b - b')_\downarrow \notin \Pi_J$. $\square$

Theorem 3. *Let $J \subset \mathbb{R}^n$ and $\lambda \in \mathbb{R}^n$ be given. If $M(J \cup \lambda) > M(J)$, then*

1. *The optimal coefficients in Problem A (with the current J and λ) converge for every $f \in W_2^{M(J)+\varepsilon}$, $\varepsilon > 0$, to an f-independent limit $c(\lambda)$.*
2. *The function $F := \sum_{j \in J} c_j(\lambda) e_j$ is the only function in Exp_J that satisfies $\deg(e_\lambda - F)_\downarrow > M(J)$, or equivalently, the sequence $j \mapsto c_j(\lambda)$ is the only sequence supported on J for which*
$$p(\lambda) = \sum_{j \in J} c_j(\lambda) p(j), \quad \forall p \in \Pi_{M(J)}.$$

Proof: Since $M(J \cup \lambda) > M(J)$, we must have $d(J, \lambda) = M(J \cup \lambda) > M(J)$. Now, fix $j \in J$. The space $\Pi_{J \cup \lambda}$ is obtained from $\Pi_{J \backslash j}$ by adding two homogeneous polynomials p, q of degrees $d(J \backslash j, j) \leq M(J)$ and $d(J, \lambda) > M(J)$ respectively. Since these two polynomials were just shown to have different degrees, the space $\Pi_{(J \backslash j) \cup \lambda}$ is obtained from $\Pi_{J \backslash j}$ by appending to the latter either p or q (but not a linear combination of them). We denote by $J_0 \subset J$ those $j \in J$ for which p is the appended polynomial. If $j \in J_0$, then $\Pi_{(J \backslash j) \cup \lambda} = \Pi_J$ (since Π_J is also obtained from $\Pi_{J \backslash j}$ by appending either p or q; however, q is ruled out because of its higher degree). By Corollary 2, $c_j(\lambda, h)$ converges for every smooth f to an f-independent limit. In the opposing case when the appended polynomial is q, then $d(J \backslash j, \lambda) = d(J, \lambda) = \deg q > M(J)$, and hence k_j of Lemma 1 is *positive*. In this case $c_j(\lambda, h) \to 0$.

We now prove the characterisation of the limit coefficients. Let q be as in the previous paragraph. Let $F_0 \in \mathrm{Exp}_{J \cup \lambda}$ be such that $F_{0\downarrow} = q$. F_0 is unique: if $\mathcal{F}_{0\downarrow} = q$ as well, and $F_0 - \mathcal{F}_0 \neq 0$, then $(F_0 - \mathcal{F}_0)_\downarrow$ is a polynomial in $\Pi_{J \cup \lambda}$ of degree $> \deg q$, which is impossible. However, we have already proved that, for $j \in J \backslash J_0$, $q \in \Pi_{(J \backslash j) \cup \lambda}$; hence the uniqueness of F_0 implies that $F_0 \in \mathrm{Exp}_{(J \backslash j) \cup \lambda}$. Consequently, $F_0 \in \mathrm{Exp}_{J_0 \cup \lambda}$.

The coefficient of e_λ in F_0 is not zero, since otherwise $F_{0\downarrow} \in \Pi_J$, and Π_J cannot contain a polynomial whose degree exceeds its own maximal degree. Normalising if necessary, we may assume that $F_0 = e_\lambda - F$, for some $F \in \mathrm{Exp}_J$. Note that F_0 as above is the unique function in Exp_J for which $F_{0\downarrow} \notin \Pi_J$. Indeed, if $\mathcal{F}_0 = e_\lambda - \mathcal{F}$, for some $\mathcal{F} \in \mathrm{Exp}_J \backslash \{F\}$, then $\mathcal{F}_0 - F_0 = F - \mathcal{F} \in \mathrm{Exp}_J \backslash 0$, and hence $\deg(\mathcal{F}_0 - F_0)_\downarrow \leq M(J)$. This implies that $\deg \mathcal{F}_{0\downarrow} \leq M(J)$. Since $\deg F_{0\downarrow} > M(J)$, we had found in $\Pi_{J \cup \lambda} \backslash \Pi_J$ two homogeneous polynomials of different degrees, and that contradicts the fact that $\dim \Pi_{J \cup \lambda} = \dim \Pi_J + 1$.

We write $F = \sum_{j \in J} d_j e_j$, and will show that $c_j(\lambda) = d_j$ for every j. Since F lies in Exp_{J_0}, we have $d_j = 0$, for every $j \in J \backslash J_0$. This agrees with our previous observation that $c_j(\lambda) = 0$ for each such j. In the opposite case, $j \in J_0$. Corollary 2 and the uniqueness of F (proved in the previous paragraph) show $d_j = c_j(\lambda)$. □

Proof of Theorem 1. With $m := m(J) = M(J)$, we know that $\Pi_J = \Pi_m$. Since $\Pi_{J \cup \lambda}$ is a proper superspace of Π_J, it must contain polynomials of degree $> m$. Theorem 3 applies to yield the first part of Theorem 1.

From the corollary we also know that, with $c(\lambda)$ the limit coefficients, $\deg(e_\lambda - \sum_{j \in J} c_j(\lambda) e_j)_\downarrow > m$, and hence that $|E_{c(\lambda)}(t)| = O(\|t\|^{m+1})$ near the origin. Part (2.) follows by invoking Theorem 2. □

One may try to adapt J to the given λ and f in a way that admits the use of a small J on the one hand, as well as the relaxation of the smoothness assumption on f on the other hand. In the statement below, the symbol D_θ stands for the directional derivative in the θ-direction.

Theorem 4. *Given $\theta \in \mathbb{R}^n$, $m > 0$ and $f_0 \in L^2(\mathbb{R}^n)$, assume that $D_\theta^m f_0 \in L^2(\mathbb{R}^n)$. Given any $\lambda \in \mathbb{R}^n$, let J be a subset of $\mathbb{R}^n$ of cardinality m such that $J \cup \{\lambda\}$ lies on a line in the θ direction. Then, there exists a sequence $c(\lambda) \in \mathbb{C}^J$, that depends only on λ and J, with the following properties:*

1. *The optimal coefficients of Problem A w.r.t. f_0 converge to $c(\lambda)$.*
2. *For every $\alpha \leq m$, and every $f \in L^2(\mathbb{R}^n)$, the $L^2(\mathbb{R}^n)$ error in*

$$E^{h\lambda} f \approx \sum_{j \in J} c_j(\lambda) E^{hj} f$$

is $O(h^\alpha)$, with α the maximal integer for which $D_\theta^\alpha f \in L^2(\mathbb{R}^n)$.

We note that (de Boor, Ron, 1990), if a set J lies entirely on a line directed in the θ direction, then $\Pi_J = \operatorname{span}\{t \mapsto (\theta \cdot t)^\ell \mid 0 \leq \ell \leq |J| - 1\}$. In particular, for a colinear set J, $M(J) = |J| - 1$.

Proof of Theorem 4. Since we assume that $J \cup \lambda$ is colinear, so is J, and hence, by the remark preceding this proof, $M(J \cup \lambda) = |J| > |J| - 1 = M(J)$. Part (1.) of the theorem then follows from Theorem 3, as soon as we verify that the weaker smoothness assumption imposed on f is suitable.

The fact that f in Corollary 1 is required to be smooth is due to the estimation in its proof of expressions of the form $\int (b - b_\downarrow)|\widehat{f}|^2$, for certain $b \in \operatorname{Exp}_{J \cup \lambda}$. We used there the estimate $|(b - b_\downarrow)(t)| \leq \operatorname{const}\|t\|^{\deg b_\downarrow + 1}$. However, in the present case, $|b - b_\downarrow|$ is a function of the variable $t \mapsto \theta \cdot t$. Hence we can bound $|(b - b_\downarrow)(t)| \leq \operatorname{const}|\theta \cdot t|^{\deg b_\downarrow + 1}$, and that allows us to relax our smoothness condition on f.

In order to prove (2.), we invoke the characterisation of the principal parts of Theorem 3: the characterisation says that, with $c(\lambda)$ being the universal limits of the optimal coefficients, we have $\deg E_{c(\lambda)\downarrow} = M(J) + 1 = |J|$. This implies $|E_{c(\lambda)}(t)| = O((\theta \cdot t)^{|J|})$. By an argument analogous to that employed in the proof of Theorem 2, we obtain that the approximation scheme has the properties asserted in (2.). □

Example: Rectangular Grids. If $J \cup \lambda$ is the cartesian product of univariate sets $(J_1, \ldots, J_n)$ of cardinalities $(k_1, \ldots, k_n)$, then (de Boor, Ron, 1992a) $\Pi_{J \cup \lambda}$ is spanned by the monomials

$$t \mapsto t^\alpha, \quad \alpha \leq k := (k_1 - 1, \ldots, k_n - 1).$$

The same reference shows that, upon deleting the monomial $t \mapsto t^k$ from the above basis, we get a basis for $\Pi_{J'}$ with J' obtained from $J \cup \lambda$ by the deletion of *any* single point in $J \cup \lambda$. In particular, $M(J \cup \lambda) > M(J)$. Thus, Theorem 3 applies to show that the optimal coefficients converge here to universal limits.

We close this section with an example showing that the requirement $M(J \cup \lambda) > M(J)$ (cf. Theorem 3) is not *necessary* for the universality of the limits of *some* of the optimal coefficients.

EXAMPLE. Assume that J_i is a finite subset of the x_i axis, $i = 1, 2$, of $\mathbb{R}^2$, $J = J_1 \cup J_2$. Then Π_J is spanned by pure monomials (i.e., by functions of the form $t \mapsto t_i^m$). For λ on the x_1-axis, $\Pi_{J\cup\lambda}$ is obtained from Π_J by appending $t \mapsto t_1^{|J_1|}$ to Π_J, and thus $d(J, \lambda) = |J_1|$. However, if $|J_2| > |J_1|$, then $M(J) = |J_2| - 1 \geq |J_1|$, and hence $M(J \cup \lambda) = M(J)$. At the same time, for each $j \in J_1$, we have that $\Pi_{(J\backslash j)\cup\lambda} = \Pi_J$, hence that $c_j(\lambda, h)$ converges, for such j and for all smooth f, to an f-independent limit. In summary, in this example, despite the fact that the condition required in Theorem 3 fails to hold, Corollary 2 still applies to show that some of the coefficients converge to universal limits.

5. Optimal approximation and radial basis functions

In the previous section, a fairly thorough analysis of the convergence of the optimal coefficients in Problem A is provided when f is "sufficiently smooth". In many cases, we can easily circumvent the smoothness assumption on f by simply reducing the cardinality of J: the smoothness assumption is relaxed as we remove points from J. However, this clearly entails that, in such case, the optimal coefficients may converge, but at the expense of an increase in the error of best approximation. After all, the spaces $(S_h^J())_h$ ought to approximate worse as J is reduced.

The purpose of this section is to give sufficient conditions for the existence of the principal parts of the optimal coefficient sequences for functions that are not covered by the results of the previous section.

To begin, we state a set of conditions on f, expressed in terms of its Fourier transform. They will apply to the next theorem, and they are related to, but more general than (1.4). To that end, we require that $f : \mathbb{R}^n \to \mathbb{R}$ be absolutely integrable and square-integrable. Its Fourier transform $\widehat{f}$ is therefore continuous. We require it to be *slowly varying*, i.e. there is a positive function G such that

$$\lim_{\|y\|\to\infty} \frac{\widehat{f}(t\|y\|)}{\widehat{f}(y)} = G(t), \qquad t \text{ a.e. in } \mathbb{R}^n. \tag{5.1}$$

Here "a.e." means almost everywhere, i.e. everywhere except perhaps on a set of measure zero.

In order that the convergence (5.1) is controlled, we require that there is a constant K so that

$$\left|\frac{\widehat{f}(t\|y\|)}{\widehat{f}(y)}\right| \leq KG(t) \tag{5.2}$$

for almost all $t \in B :=$ the unit ball. To explain the relevance of conditions (5.1)–(5.2), note that in the univariate case, they are a natural generalisation of (1.4).

We finally have a condition that fixes certain properties of G's behaviour at the origin and for large argument. Its purpose is to render the limiting minimisation problem well-defined. We state this condition by viewing G as the Fourier transform of a certain generalised function, and it will be convenient to deal with the square of G instead of G itself. Precisely, we require that the square of G is the distributional Fourier transform of a function ϕ in $\mathcal{P}$, a class with the following properties.

Every $\phi \in \mathcal{P}$ must have a distributional Fourier transform which agrees with a positive function $\widehat{\phi} \in C(\mathbb{R}^n \setminus 0) \cap L^1(\mathbb{R}^n \setminus B)$. Further, we require that there exists a sequence $c = (c_j)_{j\in J}$, so that for all $\lambda \in \mathbb{R}^n$

$$\int_{\mathbb{R}^n} |E_c(t)|^2 \widehat{\phi}(t)\, dt < \infty. \tag{5.3}$$

E_c still has the same meaning as in the previous sections.

Example. An example where all these assumptions are satisfied, is when $\widehat{f}$ is a negative power of the function $(1+\|\cdot\|)$ with a suitably large negative exponent $-\alpha - n/2$ to ensure $G^2 = \|\cdot\|^{-2\alpha-n} = \widehat{\phi} \in L^1(\mathbb{R}^n \setminus B)$. We note, however, that (5.1) really is an asymptotic condition and does not require $\widehat{f}$ to be of that form. If, as in this example, $\widehat{\phi}(t) = \|t\|^{-2\alpha-n}$, then the aforementioned $\phi(t)$ is a constant multiple of $\|t\|^{2\alpha}$ so long as α is *not* an integer, whereas in the opposing case $\phi(t)$ is a constant multiple of $\|t\|^{2\alpha}\log\|t\|$. Both of these ϕ are in the class of radial basis functions frequently considered in the literature (Buhmann, 1993, Dyn, 1989, Micchelli, 1986). In order to satisfy (5.3) for this example, we have to choose c such that

$$\sum_{j\in J} c_j j^\gamma = \lambda^\gamma, \qquad |\gamma| \le \alpha. \tag{5.4}$$

Then the condition (5.3) is true. It is always possible to achieve the above requirement by taking a large enough J.

The relevance of radial basis functions in general in this section will become clear in the following theorem, where we demonstrate that under our conditions, the limiting coefficients are Lagrange functions of certain radial basis function interpolation problems. We point out that J is still a set of scattered points in $\mathbb{R}^n$. It is a salient assumption that the only polynomial from the kernel $\mathcal{K}$ of the semi inner product associated with $\widehat{\phi}$, viz.,

$$\langle f, g\rangle = \int_{\mathbb{R}^n} \widehat{\phi}^{-1} \widehat{f}\overline{\widehat{g}},$$

that vanishes identically on J, is the zero polynomial. The above inner product is well-defined for all $f, g \in \mathcal{H}$, where

$$\mathcal{H} := \Big\{ f \in \mathcal{S}' \mid \int_{\mathbb{R}^n} \widehat{\phi}^{-1} |\widehat{f}|^2 < \infty \Big\}.$$

Here, $\mathcal{S}'$ is the space of distributions dual to the Schwartz space of rapidly decreasing smooth test functions on $\mathbb{R}^n$, cf., e.g., Jones (1982).

The above assumptions on the points we make from now on. We also find it convenient for the statement of the next theorem to define the space

$$\mathcal{J} := \Big\{ d = (d_j)_{j\in J} \in \mathbb{R}^J \mid \sum_{j\in J} d_j q(j) = 0 \,\forall\, q \in \mathcal{K} \Big\}.$$

Our last result now is as follows.

Theorem 5. *Let f be such that the above conditions* (5.1)–(5.2) *hold for $G = \sqrt{\widehat{\phi}}$, $\phi \in \mathcal{P}$. Then, with $(c_j(\lambda, h))_{j\in J}$ the optimal coefficients of Problem A, the following limits exist:*

$$\lim_{h\to 0} c_j(\lambda, h) = c_j(\lambda), \qquad j \in J, \qquad \lambda \in \mathbb{R}^n,$$

and $(c_j(\lambda))_{j\in J}$ are the coefficients that minimise

$$\min_{(c_j)_{j\in J}} \int_{\mathbb{R}^n} |E_c(t)|^2 \widehat{\phi}(t)\, dt \tag{5.5}$$

over all coefficients satisfying (5.3). *They are also the unique Lagrange functions in λ of the form*

$$c_j(\lambda) = \sum_{k\in J} d_{jk}\phi(\lambda - k) + p(\lambda), \qquad \lambda \in \mathbb{R}^n,\ j \in J,$$

that provide the interpolation conditions

$$c_j(k) = \delta_{jk}, \qquad j \in J,\ k \in J,$$

where p is a polynomial from $\mathcal{K}$ and $(d_{jk})_{k\in J} \in \mathcal{J}$.

Proof: First note that, due to Parseval's identity, the norm (1.3) that we need to minimise equals to the square root of

$$\frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} |E_c(th)|^2 |\widehat{f}(t)|^2\, dt, \tag{5.6}$$

and hence we are entitled to minimise, in lieu of (1.3), the above expression. To this end, we multiply (5.6) by

$$(2\pi h)^n |\widehat{f}(h^{-1}\mathbf{1}/\sqrt{n})|^{-2},$$

where $\mathbf{1} = (1, 1, \ldots, 1) \in \mathbb{R}^n$, and scale the argument in the integral to obtain

$$|\widehat{f}(h^{-1}\mathbf{1}/\sqrt{n})|^{-2} \int_{\mathbb{R}^n} |E_c(t)|^2 |\widehat{f}(h^{-1}t)|^2\, dt \tag{5.7}$$

instead of (5.6); thus, we can minimise (5.7), instead of (1.3). Now, let $\mathbf{c}$ be the minimising sequence of (5.5), and let $c(\lambda, h)$ be the minimising sequence of (5.7), the latter being the same as the minimising sequence of (1.3) (the uniqueness of these minimising sequences is granted by Lemma 2). Then, with c_0 any finite accumulation point of $(c(\lambda, h))_h$, and with $\mathcal{C}_k := c(\lambda, h_k)$ a subsequence that converges to c_0, we get from (5.1) that

$$|\widehat{f}(h_k^{-1}\mathbf{1}/\sqrt{n})|^{-2} |E_{\mathcal{C}_k}(t)|^2 |\widehat{f}(h_k^{-1}t)|^2$$

converges pointwise to $|E_{c_0}(t)|^2 \widehat{\phi}(t)$. Therefore, by Fatou's Lemma,

$$\int_{\mathbb{R}^n} |E_{c_0}(t)|^2 \widehat{\phi}(t)\, dt \le \liminf_{k\to\infty} |\widehat{f}(h_k^{-1}\mathbf{1}/\sqrt{n})|^{-2} \int_{\mathbb{R}^n} |E_{\mathcal{C}_k}(t)|^2 |\widehat{f}(h_k^{-1}t)|^2\, dt. \tag{5.8}$$

On the other hand, for any $h > 0$, since we are assuming $\mathbf{c}$ to satisfy (5.3), the following bound is valid:

$$|\widehat{f}(h^{-1}\mathbf{1}/\sqrt{n})|^{-2}|E_{\mathbf{c}}(t)|^2|\widehat{f}(h^{-1}t)|^2 \leq \begin{cases} |E_{\mathbf{c}}(t)|^2\widehat{\phi}(t) + \text{const}, & \text{if } \|t\| \leq 1, \\ \widehat{\phi}(t) \cdot \text{const}, & \text{if } \|t\| \geq 1. \end{cases}$$

By the dominated convergence theorem, we get thus that

$$\lim_{h\to 0}|\widehat{f}(h^{-1}\mathbf{1}/\sqrt{n})|^{-2}\int_{\mathbb{R}^n}|E_{\mathbf{c}}(t)|^2|\widehat{f}(h^{-1}t)|^2\,dt = \int_{\mathbb{R}^n}|E_{\mathbf{c}}(t)|^2\widehat{\phi}(t)\,dt.$$

Combining this with (5.8), we conclude that, in view of the optimality of $c(\lambda, h)$,

$$\int_{\mathbb{R}^n}|E_{c_0}(t)|^2\widehat{\phi}(t)\,dt \leq \int_{\mathbb{R}^n}|E_{\mathbf{c}}(t)|^2\widehat{\phi}(t)\,dt.$$

This, in turn, implies, since $\mathbf{c}$ is the unique solution of (5.5), that $c_0 = \mathbf{c}$; in other words, the sequence $(c(\lambda, h))_h$ has $\mathbf{c}$ as its unique finite accumulation point. This leaves us with showing just boundedness of the limiting coefficients to settle the lemma.

Indeed, in the following fashion we can bound the optimal coefficients independently of h: Let, for the moment being, $g : \mathbb{R}^n \to \mathbb{R}$ be *any* square-integrable function that satisfies the following conditions for a fixed $k \in J$:

$$g(\lambda) = 0, \quad g(j) = \delta_{jk}, \qquad j \in J,$$

where we assume $\lambda \neq j$ for all $j \in J$ (otherwise the solution of the minimisation problem would be obvious). Such a function exists of course. Then, by Cauchy–Schwarz,

$$\begin{aligned}
|c_k(\lambda, h)|^2 &= \Big|g(\lambda) - \sum_{j\in J} c_j(\lambda, h)g(j)\Big|^2 \\
&\leq \frac{1}{(2\pi)^{2n}}\left[\int_{\mathbb{R}^n}|\widehat{g}(t)||E_{c(\lambda,h)}(t)|\,dt\right]^2 \\
&\leq \frac{1}{(2\pi)^{2n}}\int_{\mathbb{R}^n}\frac{|\widehat{g}(t)|^2}{|\widehat{f}(h^{-1}t)|^2}\,dt\int_{\mathbb{R}^n}|E_{c(\lambda,h)}(t)|^2|\widehat{f}(h^{-1}t)|^2\,dt \\
&= \frac{1}{(2\pi)^{2n}}\int_{\mathbb{R}^n}\frac{|\widehat{g}(t)|^2}{|\widehat{f}(h^{-1}\mathbf{1}/\sqrt{n})|^{-2}|\widehat{f}(h^{-1}t)|^2}dt\,\times \\
&\qquad |\widehat{f}(h^{-1}\mathbf{1}/\sqrt{n})|^{-2}\int_{\mathbb{R}^n}|E_{c(\lambda,h)}(t)|^2|\widehat{f}(h^{-1}t)|^2dt.
\end{aligned}$$

We take $g(t) = p(t)\psi(t-k)$, where p is a polynomial of suitable degree, $p(\lambda) = 0$, $p(j) = \delta_{jk}$, ψ entire and quickly decaying along $\mathbb{R}^n$, so that $\widehat{\psi}$ is sufficiently smooth, with $\widehat{\psi}$'s support in B, and $\psi(0) = 1$. Such a polynomial exists; we may form it, for instance, as a suitably scaled product of terms of the form $\|t-j\|^2$, $\|t-\lambda\|^2$. We get, using the properties of g and of $\widehat{f}$,

$$\begin{aligned}
\limsup_{h\to 0}|c_k(\lambda, h)|^2 \leq \text{const}\cdot\int_{\|t\|<1}\widehat{\phi}(t)^{-1}\,dt\,\times& \\
\limsup_{h\to 0}|\widehat{f}(h^{-1}\mathbf{1}/\sqrt{n})|^{-2}\int_{\mathbb{R}^n}|E_c(t)|^2|\widehat{f}(h^{-1}t)|^2\,dt&
\end{aligned}$$

which is uniformly finite.

Finally, we explain the representation of the principal parts as Lagrange functions. Indeed, it is clear from (1.3) that the principal parts $c_j(\lambda)$ as $h \to 0$ of the optimal coefficients $c_j(\lambda, h)$ are fundamental functions with respect to λ on the set J, i.e. they must yield the interpolation conditions

$$c_j(k) = \delta_{jk}, \qquad j \in J,\ k \in J.$$

Further, it follows from the fact that the principal parts solve (5.5) and from $\phi \in \mathcal{P}$ that they are of the form

$$c_j(\lambda) = \sum_{k \in J} d_{jk} \phi(\lambda - k) + p(\lambda), \qquad \lambda \in \mathbb{R}^n,$$

where p is an element of the kernel $\mathcal{K}$ of the aforementioned semi inner product $\langle \cdot, \cdot \rangle$ associated with $\widehat{\phi}$. Moreover, the d_{jk} have to satisfy the side conditions mentioned in the statement of the theorem. This fact is a consequence of standard Hilbert space theory for positive definite kernels on subspaces of $\mathbb{R}^J$, see (Dyn, 1989) and (Schaback, 1993), for good summaries of this issue. The subspace here is $\mathcal{J}$. □

We observe immediately that the coefficients of this theorem give a scheme to which Theorem 2 can be applied, for instance in the example given at the beginning of this section. In that case, i.e. when $\widehat{\phi}(t) = \|t\|^{-2\alpha-n}$, $\mathcal{K}$ contains $\Pi_{<\alpha+n/2}$ which means that all such polynomials are reproduced by the Lagrange interpolation of Theorem 5 (by uniqueness of interpolation). Therefore, (5.4) holds even for all $|\gamma| < \alpha + n/2$ using the $c_j(\lambda, h)$. Theorem 2 is thus applicable for $m \le \lceil \alpha + n/2 \rceil$, the f given in the example being in $B^\alpha_\infty(L^2(\mathbb{R}^n))$.

References

1. de Boor, C. and A. Ron (1990) "On multivariate polynomial interpolation", Constructive Approximation **6**, 287–302.

2. de Boor, C. and A. Ron (1992a) "Computational aspects of polynomial interpolation in several variables", Mathematics of Computations **58(198)**, 705–727.

3. de Boor, C. and A. Ron (1992b) "The least solution of the multivariate polynomial interpolation", Math. Z. **210**, 347–378.

4. Buhmann, M.D. (1993) "New developments in the theory of radial basis function interpolation", in Multivariate Approximation: From CAGD to Wavelets, K. Jetter and F.I. Utreras, eds., World Scientific (Singapore), 35–75.

5. Buhmann, M.D. and C.A. Micchelli (1992) "On radial basis approximation on periodic grids", Math. Proc. Cambridge Phil. Soc. **112**, 317–334.

6. Dyn, N. (1989) "Interpolation and approximation by radial and related functions", in Approximation Theory VI, C.K. Chui, L.L. Schumaker, and J.D. Ward, eds., Academic Press (New York), 211–234.

7. DeVore, R.A. and G.G. Lorentz (1993) Constructive Approximation, Springer-Verlag.

8. Iserles, A. and G. Strang (1983) "The optimal accuracy of difference schemes", Trans. Amer. Math. Soc. **277**, 779–803.

9. Jones, D.S. (1982) The Theory of Generalised Functions, Cambridge University Press (Cambridge).

10. Micchelli, C.A. (1986) "Interpolation of scattered data: distance matrices and conditionally positive definite functions", Constr. Approx. **1**, 11–22.

11. Micchelli, C.A. and W.L. Miranker (1973a) "Optimal difference schemes for linear initial value problems", SIAM J. Numer. Anal. **10**, 983–1009.

12. Micchelli, C.A. and W.L. Miranker (1973b) "Asymptotically optimal approximation in fractional Sobolev spaces and the numerical solution of differential equations", Numer. Math. **22**, 75–87.

13. Schaback, R. (1993) "Comparison of radial basis function interpolants", in Multivariate Approximation: From CAGD to Wavelets, K. Jetter and F.I. Utreras, eds., World Scientific (Singapore), 293–305.

Martin D. Buhmann
Mathematik Departement
ETH Zentrum
8092 Zürich
Switzerland
E-mail: mdb@math.ethz.ch

Charles A. Micchelli
IBM T.J. Watson Research Center
Department of Mathematical Sciences
P.O. Box 218
Yorktown Heights NY 10598
USA
E-mail: cam@watson.ibm.com

Amos Ron
Computer Sciences Department
1210 West Dayton Street
University of Wisconsin - Madison
Madison WI 53706
USA
E-mail: amos@cs.wisc.edu