Computing 78, 195–209 (2006)
Digital Object Identifier (DOI) 10.1007/s00607-006-0175-1

Computing
Printed in Austria

Sparse Grids, Adaptivity, and Symmetry

H. Yserentant, Berlin

Received May 22, 2006; revised August 21, 2006
Published online: November 24, 2006

Abstract

Sparse grid methods represent a powerful and efficient technique for the representation and approximation of functions and particularly the solutions of partial differential equations in moderately high space dimensions. To extend the approach to truly high-dimensional problems as they arise in quantum chemistry, an additional property has to be brought into play, the symmetry or antisymmetry of the functions sought there.

In the present article, an adaptive sparse grid refinement scheme is developed that takes full advantage of such symmetry properties and for which the amount of work and storage remains strictly proportional to the number of degrees of freedom. To overcome the problems with the approximation of the inherently complex antisymmetric functions, augmented sparse grid spaces are proposed.

AMS Subject Classifications: 65D15, 65N50, 65T65.

Keywords: Sparse grids, adaptivity, symmetry, antisymmetry.

1. Introduction

Since the work [13] of Zenger, sparse grid techniques have developed into a very powerful tool in numerical analysis and scientific computing. The dimension of sparse grid spaces grows substantially slower with the spatial dimension than with the usual full grid spaces, essentially without detracting from the approximation properties of those. The only condition is that the functions to be approximated possess correspondingly bounded mixed derivatives. For a comprehensive survey and an extensive bibliography on sparse grid methods, see the article [2] of Bungartz and Griebel.

Sparse grid methods are very well suited to the approximation of functions depending on a moderately high number of variables, less than twenty. This does not suffice for applications in quantum chemistry. The wavefunctions that describe a system consisting of N electrons depend on $3N$ variables, three for the coordinates of each electron. These wavefunctions possess the required type of smoothness [10], [12], but to break the curse of dimensionality, an additional property has to be brought into play, their antisymmetry enforced by the Pauli principle. See [3], [4], and [11] for recent attempts to solve problems of quantum chemistry by sparse grid methods.

The purpose of the present article is to develop a framework for adaptive sparse grid methods that takes full advantage of symmetry properties as they arise in such fields of application. We develop a refinement scheme that works directly on the equivalence classes of the grid points that transfer to each other under the given symmetry operations. The approach is fairly general and applies to arbitrary symmetry groups that are composed of permutations and coordinate reflections.

The paper is organized as follows. In Sect. 1, the anisotropic tensor product hierarchical bases of which the approximation spaces are built up are introduced. In Sect. 2, the refinement scheme is presented. The basic data structure are the points around which the basis functions are centered and by which they are completely determined. In the transition from one refinement level to the next, only a single component of a grid point is exchanged for one of its two neighbors of the next finer level. The number of the potential direct descendants of a grid point grows therefore only proportional to the space dimension. In Sect. 3, the basic symmetry operations are introduced and the refinement scheme is generalized to symmetry classes of grid points. It is described how symmetric functions are represented in this framework.

Section 4 deals with antisymmetric functions. Formally, the number of degrees of freedom of antisymmetric functions is still further reduced by that compared to symmetric sparse grid functions, but an additional complication arises there. A large number of refinement steps is needed in higher space dimensions until a non-trivial contribution to the function to be approximated is found. This difficulty reflects the inherent complexity of antisymmetric functions in higher-space dimensions.

A solution to this problem is proposed in Sect. 6. The idea is to assign to the each single grid point first a whole family of ansatz functions with good local approximation properties. A particular linear combination of these functions is then selected and instead of the old hierarchical basis function attached to the grid point. This linear combination should on one hand be chosen from the approximation point of view, but, on the other hand, be sufficiently simple and still well manageable. In some cases as for the full permutation group there is an obvious choice for a linear combination of this kind whereas other cases like those arising from quantum physics or chemistry require much more involved strategies. The linear combinations can also be given implicitly, for example as antisymmetrized products of lower-dimensional functions, and can itself depend on the function to be approximated. In the extreme, such procedures are vaguely reminiscent of the Hartree-Fock method [5], [6] from quantum physics. But in contrast to true Hartree-Fock methods, these procedures serve in the given context only to augment the sparse grid spaces and to shorten and bridge the pre-asymptotic phase. The basic approximation properties of the underlying sparse grid spaces are in no way affected by such constructions and are fully maintained.

In Sect. 7, finally the conversion of the given concepts into computational procedures is discussed for two examples including the one that arises from the approximation of electronic wavefunctions and that motivated the present work.

2. Tensor Product Hierarchical Bases

Starting point of the whole construction is a one-dimensional hierarchical basis. The hierarchical basis functions of level j are centered around the grid points of this level in the interior of a given interval $[-L, L]$. The only grid point q of the initial level $j = 0$ is the point $q = 0$. The grid points q of level $\ell(q) \leq j$ subdivide the given interval into intervals of equal length. The midpoints q of these intervals are the grid points of the next finer level $\ell(q) = j + 1$. The level of the hierarchical basis function

$$\phi(x; q) = \phi\Big(\frac{x - q}{h(q)}\Big) \tag{2.1}$$

is by definition the level j of the point q around which it is centered. Its width $h(q) = L/2^j$ can be reconstructed from the position of q. The function ϕ is an even shape function that fixes the type of the approximation spaces. We suppose that the resulting hierarchical basis functions are linearly independent of each other. An example of such a shape function is the piecewise linear hat function

$$\phi(\xi) = \max(1 - |\xi|,\ 0) \tag{2.2}$$

that leads to a piecewise linear hierarchical basis [8], [9] as known from finite elements. Some of the corresponding basis functions are depicted in Fig. 1. It is, however, important to note that this choice is by no means unique. There is a vast number of other possibilities. Particularly, the presented concept can be generalized to wavelets, in which a different shape function is used on the initial level than on the other levels. Wavelets are superior to piecewise linear hat functions in many respects.

A grid point $\mathbf{q} = (q_1, q_2, \ldots, q_d)$ in d space dimensions is composed of one-dimensional grid points q_i of the described kind. Its level

$$\ell(\mathbf{q}) = \ell(q_1) + \ldots + \ell(q_d) \tag{2.3}$$

is the sum of the levels of its components. The d-dimensional basis function

$$\psi(\mathbf{x}; \mathbf{q}) = \prod_{i=1}^{d} \phi(x_i; q_i) \tag{2.4}$$

is again uniquely determined by the d-dimensional grid point $\mathbf{q}$ around which it is centered. The hierarchical basis functions (2.4) are linearly independent of each other, as can be shown by induction over the space dimension.

The full grid space of level j is spanned by the basis functions associated with the grid points $\mathbf{q}$ with components q_i of levels $\ell(q_i) \leq j$. The number of these basis

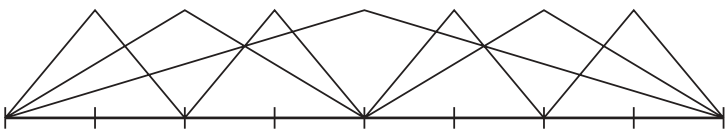

Fig. 1. The hierarchical basis functions of levels $j = 0, 1, 2$

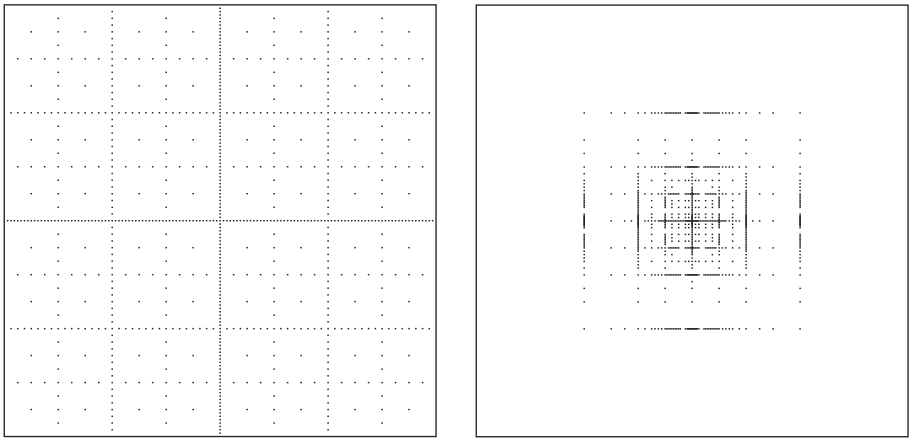

Fig. 2. A regular and an adaptively refined sparse grid

functions increases like $\sim (2^j)^d$, which is outside any reach for larger d. In contrast, the dimension of the sparse grid space spanned by the functions

$$\psi(\mathbf{x};\mathbf{q}), \quad \ell(\mathbf{q}) \le j, \tag{2.5}$$

increases only like $\sim j^{d-1}2^j$, that is, much slower with the space dimension d. It is a surprising fact that has been noted at many places since the work of Babenko, Korobov, or Smolyak [1], [7] in the early sixties of the last century that the approximation properties of these sparse grid spaces are not much worse for sufficiently smooth functions than the approximation properties of the corresponding full grid spaces. For some cases, so-called energy based sparse grid spaces [2], [4] even represent a further improvement. These considerations demonstrate that already linear combinations of comparatively few anisotropic tensor product hierarchical basis functions can have astonishingly good approximation properties. Figure 2 shows the regular sparse grid consisting of the grid points of levels $j \le 6$ together with an adaptively refined sparse grid in two space dimensions. The hope is that under corresponding additional symmetry assumptions the logarithmic term disappears and the curse of dimensionality can be broken. Adaptivity will help to exploit the approximation properties fully.

3. The Refinement Scheme

The given grid points can be generated recursively. The only grid point of level zero is the point $\mathbf{q} = \mathbf{0}$. The grid points of level $j+1$ can be represented in the form

$$\mathbf{q} \pm \frac{1}{2}\,\mathbf{H}(\mathbf{q})\mathbf{e}_i, \quad i = 1, \ldots, d, \tag{3.1}$$

with the points $\mathbf{q}$ the grid points of the preceding level j and the $\mathbf{e}_i$ the unit vectors pointing into the direction of the coordinate axes. The diagonal matrices

$$\mathbf{H}(\mathbf{q}) = \operatorname{diag}\big(h(q_1), \ldots, h(q_d)\big) \tag{3.2}$$

depend on the components q_i of $\mathbf{q}$ and are composed of the widths of the one-dimensional functions of which the hierarchical basis functions (2.4) are built up. That is, a grid point of level $j+1$ is generated from a grid point of level j exchanging one of its components for one of its two neighbors on the next finer level. We will denote the grid points (3.1) as the direct descendants of the grid point $\mathbf{q}$.

This observation opens the door to adaptive refinement schemes. Starting from a space spanned by finitely many hierarchical basis functions represented by a finite set of grid points, first the direct descendants of these points are determined. Which of these points is then added to the current grid is decided by means of some error criterion. The procedure is then repeated until the desired accuracy is reached.

Grid points of higher levels can be reached on many different refinement paths. Without corresponding countermeasures, many identical copies of the grid points will therefore be generated. The amount of work and storage explodes then very rapidly. The most obvious method to avoid this is to check in every refinement step whether the given grid point already exists. Strategies like this one can be most easily transfered to complicated situations as they will be considered in the following sections.

A more elegant way to overcome this difficulty is to attach $2d$ flags associated with its possible direct descendants to every grid point. These flags indicate whether the corresponding new points may be generated. Assume that a new point has been generated exchanging the component q_i of a grid point $\mathbf{q}$ for one of its two neighbors on the next finer level. The new point then inherits all flags from its ancestor except for those associated with the direction i. These are set to "allowed". In return, the flag of the original point associated with the new point is set to "forbidden". This point will then have no other descendant on the corresponding side of the hyperplane $x_i = q_i$ than the just generated one. The forbidden descendants can be reached from the new point. Of course, all flags of the initial point $\mathbf{q} = \mathbf{0}$ have to be set to "allowed" in the beginning of the refinement process.

4. Symmetry

We start in this section from a group $\mathbf{G}$ of $(d \times d)$-matrices that are products of transpositions and coordinate reflections. A transposition is a matrix that interchanges two components of a vector and leaves the other components invariant. A coordinate reflection exchanges a single component of a vector for its negative value. The order $|\mathbf{G}|$ of such a group, that is, the number of its elements is finite. The entries of a matrix $\mathbf{P} \in \mathbf{G}$ are ± 1 or 0. Such groups can nevertheless have a rather complicated structure since, as is well-known and one can easily think about, every finite group of order d is isomorphic to a subgroup of the $(d \times d)$-permutation matrices.

The matrices $\mathbf{P} \in \mathbf{G}$ map the set of the grid points of a given level bijectively onto itself. A function u is called symmetric with respect to the group $\mathbf{G}$ if

$$u(\mathbf{P}\mathbf{x}) = u(\mathbf{x}), \quad \mathbf{P} \in \mathbf{G}. \tag{4.1}$$

The aim of this section is to derive a refinement scheme from the given one that takes full advantage of this kind of symmetry. We begin our considerations with the property by which the ansatz functions and the symmetry group interact.

Lemma 1: *For all grid points* $\mathbf{q}$ *and all matrices* $\mathbf{P}$ *in the symmetry group* $\mathbf{G}$,

$$\psi(\mathbf{Px};\mathbf{q}) = \psi(\mathbf{x};\mathbf{P}^{-1}\mathbf{q}). \tag{4.2}$$

Proof: Let $\mathbf{P}$ be the permutation matrix that exchanges the component i of a vector for the component $\pi(i)$. As the order of the factors in a product is arbitrary, then

$$\psi(\mathbf{Px};\mathbf{q}) = \prod_{i=1}^{d} \phi\big(x_{\pi(i)};q_i\big) = \prod_{i=1}^{d} \phi\big(x_i;q_{\pi^{-1}(i)}\big) = \psi(\mathbf{x};\mathbf{P}^{-1}\mathbf{q}) \tag{4.3}$$

so that (4.2) particularly holds for all transpositions. Since

$$\phi(-x;q) = \phi(x;-q) \tag{4.4}$$

for the univariate functions (2.1) of which the hierarchical basis functions (2.4) are composed, (4.2) holds also for coordinate reflections and therefore for all $\mathbf{P}\in\mathbf{G}$. □

We emphasize once more that the structure of the ansatz functions enters only via the relation (4.2) into the further considerations that therefore generalize to all functions for which this relation holds. As long as $\mathbf{G}$ is a sheer permutation group, one can even dispense with the condition (4.4), that is, in the case of the hierarchical basis functions introduced in Sect. 2, with the fact that univariate shape function is even.

Lemma 2: *For all matrices* $\mathbf{P}$ *in the symmetry group* $\mathbf{G}$,

$$\sum_{\mathbf{q}} \alpha(\mathbf{q})\,\psi(\mathbf{Px};\mathbf{q}) = \sum_{\mathbf{q}} \alpha(\mathbf{Pq})\,\psi(\mathbf{x};\mathbf{q}). \tag{4.5}$$

Proof: Lemma 1 yields

$$\sum_{\mathbf{q}} \alpha(\mathbf{q})\,\psi(\mathbf{Px};\mathbf{q}) = \sum_{\mathbf{q}} \alpha(\mathbf{PP}^{-1}\mathbf{q})\,\psi(\mathbf{x};\mathbf{P}^{-1}\mathbf{q}).$$

Since the matrix $\mathbf{P}^{-1}$ maps the set of the grid points of every given level bijectively onto itself, this already proves the proposition. □

We are now in the position to study the linear combinations of the hierarchical basis functions that are symmetric in the sense of Eq. (4.1).

Lemma 3: *A finite linear combination*

$$u(\mathbf{x}) = \sum_{\mathbf{q}} \alpha(\mathbf{q})\,\psi(\mathbf{x};\mathbf{q}) \tag{4.6}$$

of hierarchical basis functions is symmetric if and only if for all $\mathbf{P}\in\mathbf{G}$

$$\alpha(\mathbf{Pq}) = \alpha(\mathbf{q}). \tag{4.7}$$

Proof: By Lemma 2, the linear combination (4.6) is symmetric if and only if

$$\sum_{\mathbf{q}} \alpha(\mathbf{Pq})\, \psi(\mathbf{x}; \mathbf{q}) = \sum_{\mathbf{q}} \alpha(\mathbf{q})\, \psi(\mathbf{x}; \mathbf{q})$$

for all matrices $\mathbf{P} \in \mathbf{G}$. The condition (4.7) hence implies the symmetry. The other direction follows from the linear independence of the basis functions. □

Next we introduce the equivalence relation

$$\mathbf{q} \sim \mathbf{q}' \quad :\Leftrightarrow \quad \mathbf{q}' = \mathbf{Pq} \qquad \text{for a } \mathbf{P} \in \mathbf{G} \tag{4.8}$$

on the set of all grid points $\mathbf{q}$ and fix a set S of representatives of the corresponding equivalence classes $[\,\mathbf{q}\,]$, that is, a set of grid points that contains exactly one element from each of these equivalence classes. Furthermore, let

$$\widetilde{\psi}(\mathbf{x}; \mathbf{q}) \;=\; \frac{1}{|\mathbf{G}|} \sum_{\mathbf{P} \in \mathbf{G}} \psi(\mathbf{Px}; \mathbf{q}). \tag{4.9}$$

By (4.2), these functions are also symmetric with respect to $\mathbf{q}$ in the sense that

$$\widetilde{\psi}(\mathbf{x}; \mathbf{Pq}) = \widetilde{\psi}(\mathbf{x}; \mathbf{q}), \quad \mathbf{P} \in \mathbf{G}. \tag{4.10}$$

Lemma 4: *Every symmetric finite linear combination of hierarchical basis functions possesses the representation*

$$\sum_{\mathbf{q}} \alpha(\mathbf{q})\, \psi(\mathbf{x}; \mathbf{q}) \;=\; \sum_{\mathbf{q} \in S} |[\,\mathbf{q}\,]|\, \alpha(\mathbf{q})\, \widetilde{\psi}(\mathbf{x}; \mathbf{q}), \tag{4.11}$$

where $|[\,\mathbf{q}\,]|$ *denotes the number of elements in the equivalence class* $[\,\mathbf{q}\,]$.

Proof: A symmetric function satisfies the relation

$$u(\mathbf{x}) \;=\; \frac{1}{|\mathbf{G}|} \sum_{\mathbf{P} \in \mathbf{G}} u(\mathbf{Px}).$$

From that, the representation

$$\sum_{\mathbf{q}} \alpha(\mathbf{q})\, \psi(\mathbf{x}; \mathbf{q}) = \sum_{\mathbf{q}} \alpha(\mathbf{q})\, \widetilde{\psi}(\mathbf{x}; \mathbf{q})$$

follows. The equivalence classes form a partition of the sets of the grid points into disjoint subsets. Hence, the sum on the right-hand side splits into

$$\sum_{\mathbf{q}} \alpha(\mathbf{q})\, \widetilde{\psi}(\mathbf{x}; \mathbf{q}) = \sum_{\mathbf{q} \in S} \sum_{\mathbf{q}' \in [\,\mathbf{q}\,]} \alpha(\mathbf{q}')\, \widetilde{\psi}(\mathbf{x}; \mathbf{q}').$$

Every $\mathbf{q}' \in [\,\mathbf{q}\,]$ can by definition be written in the form $\mathbf{q}' = \mathbf{Pq}$ with a matrix $\mathbf{P}$ in the symmetry group $\mathbf{G}$. By (4.7) and (4.10) therefore

$$\alpha(\mathbf{q}')\, \widetilde{\psi}(\mathbf{x}; \mathbf{q}') = \alpha(\mathbf{q})\, \widetilde{\psi}(\mathbf{x}; \mathbf{q})$$

for all these $\mathbf{q}'$, which proves the proposition. □

A symmetric linear combination of hierarchical basis functions is therefore uniquely determined by its coefficients $\alpha(\mathbf{q})$ for the grid points $\mathbf{q}$ in the given set of representatives. These coefficients can be arbitrarily given, the other coefficients are then fixed by the condition (4.7). Depending on the symmetry group, the number of degrees of freedom is substantially reduced by that. This suggests to work with the equivalence classes of the grid points, respectively their representatives, and with the symmetrized basis functions (4.9) instead of the original ones.

The next step is to incorporate the given kind of symmetry into the refinement scheme from Sect. 3 and to generalize it to the given equivalence classes. We begin with an observation on the action of the symmetry operations onto the matrices (3.2).

Lemma 5: *For all grid points* $\mathbf{q}$ *and all matrices* $\mathbf{P}$ *in the symmetry group* $\mathbf{G}$*,*

$$\mathbf{P}\mathbf{H}(\mathbf{q}) = \mathbf{H}(\mathbf{P}\mathbf{q})\mathbf{P}. \tag{4.12}$$

Proof: The proposition is more or less obvious for permutation matrices. Since by assumption $h(-q_i) = h(q_i)$, it also remains true for coordinate reflections and therefore for all products of permutations matrices and such coordinate reflections. □

Equivalent grid points are of the same level. Therefore we can define the equivalence classes of level j as the equivalence classes of the grid points of level j. Recall that the grid points of level $j+1$ are the descendants (3.1) of the grid points of the preceding level j, and that their equivalence classes are therefore the equivalence classes of these descendants. The next lemma shows that these equivalence classes in fact only depend on the equivalence class of the ancestor.

Lemma 6: *For all grid points* $\mathbf{q}$*, the set of the equivalence classes*

$$\big[\,\mathbf{q} \pm \frac{1}{2}\,\mathbf{H}(\mathbf{q})\mathbf{e}_i\,\big], \quad i = 1, \ldots, d, \tag{4.13}$$

depends only on the equivalence class of $\mathbf{q}$*.*

Proof: By Lemma 5, for all $\mathbf{P}$ in the group $\mathbf{G}$,

$$\mathbf{P}\mathbf{q} \pm \frac{1}{2}\,\mathbf{H}(\mathbf{P}\mathbf{q})\mathbf{e}_i = \mathbf{P}\big(\mathbf{q} \pm \frac{1}{2}\,\mathbf{H}(\mathbf{q})\mathbf{P}^{-1}\mathbf{e}_i\big).$$

Since the set of the vectors $\pm\mathbf{P}^{-1}\mathbf{e}_i$ coincides with the set of the vectors $\pm\mathbf{e}_i$ for all permutations and coordinate reflections $\mathbf{P}$ and therefore for all matrices $\mathbf{P}$ in the symmetry group, this proves the proposition. □

Lemma 6 enables it to define the equivalence classes (4.13) as the direct descendants of the equivalence class $[\,\mathbf{q}\,]$. The set of the equivalence classes of level $j+1$ consists then of these descendants of the equivalence classes of level j. Every equivalence class can be reached from the equivalence class of the initial point $\mathbf{q} = \mathbf{0}$ stepping

forward from one descendant to the next. If the grid point $\mathbf{q} = \mathbf{q}^{(j)}$ is reached from the initial point $\mathbf{q}^{(0)} = \mathbf{0}$ along the refinement path

$$\mathbf{q}^{(0)} \to \mathbf{q}^{(1)} \to \ \dots \ \to \mathbf{q}^{(j)}, \tag{4.14}$$

the equivalence class of $\mathbf{q}$ is reached on the corresponding path

$$[\mathbf{q}^{(0)}] \to \ \dots \ \to [\mathbf{q}^{(j)}] \tag{4.15}$$

of equivalence classes. With this observation, we have generalized the refinement scheme of the previous section to the case of symmetric functions.

Double occurrences of equivalence classes again have to be eliminated in the course of the refinement process. An actual computation would start from a subset of points $\mathbf{q} \in S$ representing a symmetric grid or, more abstractly, a set of equivalence classes. For all these points $\mathbf{q}$ then the points $\mathbf{q}' \in S$ representing the equivalence classes (4.13) of their direct descendants are calculated, where copies have to be singled out in a suitable way, for example by means of balanced search trees. A two-stage process in which first the descendants of each single grid point are considered separately from the other points may be advantageous here. Which of the remaining points $\mathbf{q}'$ are then actually added to the present grid is decided with the help of problem dependent refinement criterions that are not subject of the present study.

The presented technique can be easily transfered to more general refinement patterns. Let $\mathbf{a}_1, \dots, \mathbf{a}_r$ be a set of grid points that fulfills the following two conditions. First, it must contain the vectors $\pm \mathbf{e}_i/2$ to guarantee that every grid point can be reached from the initial point. Secondly, it has to be invariant under the given symmetry operations to allow for an argumentation as in the proof of Lemma 6. The grid points

$$\mathbf{q} + \mathbf{H}(\mathbf{q})\mathbf{a}_i, \quad i = 1, \dots, r, \tag{4.16}$$

can then replace the grid points (3.1) as the direct descendants of the grid point $\mathbf{q}$ and their equivalence classes correspondingly the equivalence classes (4.13).

5. Antisymmetric Functions

Let $\mathrm{sign}: \mathbf{G} \to \{-1, 1\}$ be a group homomorphism, that is, a mapping with

$$\mathrm{sign}(\mathbf{PQ}) = \mathrm{sign}(\mathbf{P})\mathrm{sign}(\mathbf{Q}) \tag{5.1}$$

for all matrices $\mathbf{P}$ and $\mathbf{Q}$ in the group $\mathbf{G}$. A function u is called antisymmetric with respect to the symmetry group $\mathbf{G}$ and this sign function, if

$$u(\mathbf{Px}) = \mathrm{sign}(\mathbf{P})u(\mathbf{x}), \quad \mathbf{P} \in \mathbf{G}. \tag{5.2}$$

The aim of this section is to study antisymmetric linear combinations

$$u(\mathbf{x}) = \sum_{\mathbf{q}} \alpha(\mathbf{q})\, \psi(\mathbf{x}; \mathbf{q}) \tag{5.3}$$

of hierarchical basis functions. The first observation is that such a linear combination is antisymmetric if and only if the coefficients satisfy the relation

$$\alpha(\mathbf{Pq}) = \mathrm{sign}(\mathbf{P})\alpha(\mathbf{q}), \quad \mathbf{P} \in \mathbf{G}, \tag{5.4}$$

that is, are antisymmetric in $\mathbf{q}$. The proof is the same as for Lemma 3. Let

$$\psi^*(\mathbf{x}; \mathbf{q}) = \frac{1}{|\mathbf{G}|} \sum_{\mathbf{P} \in \mathbf{G}} \mathrm{sign}(\mathbf{P}) \psi(\mathbf{Px}; \mathbf{q}) \tag{5.5}$$

be the antisymmetrized counterpart of the hierarchical basis function (2.4). Corresponding to the property (4.10) of the symmetrized basis functions, then

$$\psi^*(\mathbf{x}; \mathbf{Pq}) = \mathrm{sign}(\mathbf{P})\psi^*(\mathbf{x}; \mathbf{q}), \quad \mathbf{P} \in \mathbf{G}. \tag{5.6}$$

Antisymmetric functions can be represented as

$$u(\mathbf{x}) = \frac{1}{|\mathbf{G}|} \sum_{\mathbf{P} \in \mathbf{G}} \mathrm{sign}(\mathbf{P}) u(\mathbf{Px}). \tag{5.7}$$

As in the proof of Lemma 4, from that and (5.4) and (5.6) the representation

$$\sum_{\mathbf{q}} \alpha(\mathbf{q})\, \psi(\mathbf{x}; \mathbf{q}) = \sum_{\mathbf{q} \in S} |[\,\mathbf{q}\,]|\, \alpha(\mathbf{q})\, \psi^*(\mathbf{x}; \mathbf{q}) \tag{5.8}$$

of antisymmetric linear combinations of hierarchical basis functions can be deduced. This corresponds to the representation (4.11) of symmetric functions. Antisymmetric sparse grid functions can therefore be adaptively generated and administrated in the same way as symmetric sparse grid functions.

A grid point $\mathbf{q}$ is called a diagonal point if there is a matrix $\mathbf{P}$ in $\mathbf{G}$ of negative sign with $\mathbf{q} = \mathbf{Pq}$. The grid points in an equivalence class are either all diagonal or all non-diagonal so that one can correspondingly define diagonal and non-diagonal equivalence classes. The set S' of all those points in the set of representatives S that are not diagonal points therefore forms a set of representatives of the non-diagonal equivalence classes. By (5.6), the antisymmetrized counterparts (5.5) of the basis functions associated with the diagonal points vanish. Antisymmetric linear combinations of hierarchical basis functions can therefore be written in the form

$$\sum_{\mathbf{q}} \alpha(\mathbf{q})\, \psi(\mathbf{x}; \mathbf{q}) = \sum_{\mathbf{q} \in S'} |[\,\mathbf{q}\,]|\, \alpha(\mathbf{q})\, \psi^*(\mathbf{x}; \mathbf{q}). \tag{5.9}$$

They are determined by the coefficients $\alpha(\mathbf{q})$, $\mathbf{q} \in S'$, that can be arbitrarily given. The other coefficients are fixed by the symmetry condition (5.4).

Although this still further reduces the number of degrees of freedom, this fact more complicates rather than simplifies the situation. In higher dimensions, a huge number of diagonal points is normally generated in the course of the refinement process until the first non-diagonal point is reached. An enormous amount of overhead is produced in such cases, which also reflects the inherent complexity of antisymmetric functions.

6. Augmented Function Spaces

The way out of this dilemma is to augment the function spaces and to assign modified functions to the diagonal grid points. We will restrict ourselves for the sake of simplicity in the sequel to groups **G** consisting only of permutations matrices and first replace the old hierarchical basis functions (2.4) by functions

$$\psi(\mathbf{x}; \mathbf{q}, \boldsymbol{\nu}) = \prod_{i=1}^{d} \phi(x_i; q_i, \nu_i) \tag{6.1}$$

that additionally depend on a second d-dimensional vector $\boldsymbol{\nu}$ with components ν_i in a given set Γ of real numbers. The univariate functions

$$\phi(x; q, \nu) = \phi\Big(\frac{x - q}{h(q)}; \nu\Big) \tag{6.2}$$

are hereby centered around the same grid points q as before and have the same associated widths as the one-dimensional hierarchical basis functions (2.1). We assume that $\nu = 0$ is contained in Γ and that the function $\phi(\xi) = \phi(\xi, 0)$ is the original shape function of which the hierarchical basis functions (2.4) are composed.

The following examples illustrate the span of possibilities. In the first one, Γ is the set of the nonnegative integers and the underlying shape functions are of the form

$$\phi(\xi; \nu) = \phi(\xi) L_\nu(\xi), \quad \nu = 0, 1, 2, \dots, \tag{6.3}$$

where the function ϕ on the right-hand side is the old shape function from Sect. 2 and the L_ν are, for example, orthogonal polynomials of degree ν. In the second example, Γ is the set of the one-dimensional grid points and

$$\phi(\xi; \nu) = \phi\Big(\frac{\xi - \nu}{h(\nu)}\Big). \tag{6.4}$$

This means that the basis functions assigned to all its possible descendants are once more virtually attached to a grid point $\mathbf{q}$ and shows at the same time that linear independence of the functions (6.1) can no longer be assumed. Sparse grid combination techniques (see [2], and [11] for applications in quantum theory) fall in the given category, too. Here, Γ is again the set of the nonnegative integers and

$$\phi(\xi; \nu) = \phi(2^\nu \xi). \tag{6.5}$$

In this case, correspondingly shrunken copies of the original basis function are additionally attached to the grid points. The approximation properties of such function spaces can be most easily understood if one switches to the Fourier transform and starts from an anisotropic Littlewood-Paley decomposition [11].

For all permutation matrices $\mathbf{P}$ and therefore by assumption particularly for all matrices $\mathbf{P}$ in the given symmetry group **G**,

$$\psi(\mathbf{P}\mathbf{x}; \mathbf{q}, \boldsymbol{\nu}) = \psi(\mathbf{x}; \mathbf{P}^{-1}\mathbf{q}, \mathbf{P}^{-1}\boldsymbol{\nu}). \tag{6.6}$$

This property generalizes (4.2) and is central for our further considerations. Equation (6.6) is at the same time the only place where the structure of the functions (6.1) interacts with the structure of the symmetry group. By (6.6), the linear combinations

$$\sum_{\mathbf{q},\boldsymbol{\nu}} \alpha(\mathbf{q},\boldsymbol{\nu})\,\psi(\mathbf{x};\mathbf{q},\boldsymbol{\nu}) \tag{6.7}$$

of the functions (6.1) satisfy the relation

$$\sum_{\mathbf{q},\boldsymbol{\nu}} \alpha(\mathbf{q},\boldsymbol{\nu})\,\psi(\mathbf{P}\mathbf{x};\mathbf{q},\boldsymbol{\nu}) \;=\; \sum_{\mathbf{q},\boldsymbol{\nu}} \alpha(\mathbf{P}\mathbf{q},\mathbf{P}\boldsymbol{\nu})\,\psi(\mathbf{x};\mathbf{q},\boldsymbol{\nu}) \tag{6.8}$$

for all permutation matrices $\mathbf{P}$. Because of the possible lack of linear independence we now enforce the antisymmetry of the linear combinations (6.7) by the condition

$$\alpha(\mathbf{P}\mathbf{q},\mathbf{P}\boldsymbol{\nu}) = \operatorname{sign}(\mathbf{P})\alpha(\mathbf{q},\boldsymbol{\nu}), \quad \mathbf{P}\in\mathbf{G}, \tag{6.9}$$

to the coefficients. The next step is to rewrite the corresponding linear combinations in terms of the antisymmetrized basis functions

$$\psi^*(\mathbf{x};\mathbf{q},\boldsymbol{\nu}) \;=\; \frac{1}{|\mathbf{G}|}\sum_{\mathbf{P}\in\mathbf{G}} \operatorname{sign}(\mathbf{P})\psi(\mathbf{P}\mathbf{x};\mathbf{q},\boldsymbol{\nu}). \tag{6.10}$$

By (6.6), these functions are also antisymmetric with respect to simultaneous permutations of the grid points $\mathbf{q}$ and the vectors $\boldsymbol{\nu}$ in the sense that

$$\psi^*(\mathbf{x};\mathbf{P}\mathbf{q},\mathbf{P}\boldsymbol{\nu}) = \operatorname{sign}(\mathbf{P})\psi^*(\mathbf{x};\mathbf{q},\boldsymbol{\nu}), \quad \mathbf{P}\in\mathbf{G}. \tag{6.11}$$

Two pairs $(\mathbf{q},\boldsymbol{\nu})$ and $(\mathbf{q}',\boldsymbol{\nu}')$ are called equivalent if there is a matrix $\mathbf{P}$ in the symmetry group $\mathbf{G}$ with $\mathbf{q}' = \mathbf{P}\mathbf{q}$ and $\boldsymbol{\nu}' = \mathbf{P}\boldsymbol{\nu}$. A pair $(\mathbf{q},\boldsymbol{\nu})$ is a diagonal element if there is a $\mathbf{P}$ in $\mathbf{G}$ of negative sign with $\mathbf{q} = \mathbf{P}\mathbf{q}$ and $\boldsymbol{\nu} = \mathbf{P}\boldsymbol{\nu}$. The pairs in an equivalence class are then again either all diagonal or all non-diagonal. Let S_2 be a set of representatives of the corresponding equivalence classes and let S_2' consist of those elements in S_2 that are not diagonal elements. By (6.9) and (6.11), the given antisymmetric linear combination (6.7) can then again be represented in the form

$$\sum_{\mathbf{q},\boldsymbol{\nu}} \alpha(\mathbf{q},\boldsymbol{\nu})\,\psi(\mathbf{x};\mathbf{q},\boldsymbol{\nu}) \;=\; \sum_{(\mathbf{q},\boldsymbol{\nu})\in S_2'} |[(\mathbf{q},\boldsymbol{\nu})]|\,\alpha(\mathbf{q},\boldsymbol{\nu})\,\psi^*(\mathbf{x};\mathbf{q},\boldsymbol{\nu}). \tag{6.12}$$

The coefficients $\alpha(\mathbf{q},\boldsymbol{\nu})$ for the pairs $(\mathbf{q},\boldsymbol{\nu}) \in S_2'$ can be arbitrarily given. The other coefficients are then again fixed by the condition (6.9).

The set of representatives S_2 can be chosen such that, for every $(\mathbf{q},\boldsymbol{\nu}) \in S_2$, the grid point $\mathbf{q}$ belongs to the old set of representatives S introduced in Sect. 4. The refinement scheme can therefore be kept. The idea is to attach, to every $\mathbf{q} \in S$, a linear combination of the functions (6.10) assigned to the corresponding non-diagonal pairs $(\mathbf{q},\boldsymbol{\nu}) \in S_2'$. This linear combination can also be given in implicit form, for example as an antisymmetrized product of lower dimensional functions, and may itself depend on the function to be approximated. To reach our initial goal, pairs $(\mathbf{q},\boldsymbol{\nu})$ with nonvanishing $\boldsymbol{\nu}$ need only to be included for diagonal points $\mathbf{q}$.

7. Two Examples

We first consider the case that $\mathbf{G}$ is the full permutation group, the group that consists of all permutation matrices, and equip it with usual sign function

$$\operatorname{sign}(\mathbf{P}) = \det(\mathbf{P}). \tag{7.1}$$

The order of this group is $|\mathbf{G}| = d!$. Also for higher dimensions d, the antisymmetrized basis functions (6.10) can nevertheless be easily evaluated as normalized determinant

$$\psi^*(\mathbf{x}; \mathbf{q}, \boldsymbol{\nu}) = \frac{1}{d!} \det\big(\phi(x_i; q_j, \nu_j)\big) \tag{7.2}$$

of a $(d \times d)$-matrix. A similar rule holds for the L_2-inner product of two such functions.

A possible set of representatives S is the set of the grid points $\mathbf{q}$ with components

$$q_1 \le q_2 \le \ldots \le q_d. \tag{7.3}$$

The representative of a grid point $\mathbf{q}$ is found sorting the components of $\mathbf{q}$.

The diagonal points are the points with at least two coinciding components. For the non-diagonal points therefore $\mathbf{q} \neq \mathbf{P}\mathbf{q}$ for all non-trivial $\mathbf{P}$ in $\mathbf{G}$. The equivalence classes of these points hence contain exactly $d!$ elements so that the number of degrees of freedom diminishes by more than this factor in the case of antisymmetric functions.

To simplify matters, the pairs $(\mathbf{q}, \boldsymbol{\nu})$ of grid points $\mathbf{q}$ and vectors $\boldsymbol{\nu}$ with components in the set Γ from Sect. 6 are here represented as sequences

$$(q_i, \nu_i), \quad i = 1, \ldots, d, \tag{7.4}$$

of pairs of one-dimensional grid points q_i and elements ν_i of Γ. A pair $(\mathbf{q}, \boldsymbol{\nu})$ is diagonal if at least two of these pairs (q_i, ν_i) coincide. The set of representatives S_2 of the corresponding equivalence classes can then be chosen as the set of the lexicographically ordered sequences (7.4).

The second example is inspired by the Pauli principle from quantum physics. Here $d = 3N$, where N is the number of the spin 1/2-particles of which the considered system is composed. The vectors $\mathbf{x} \in \mathbb{R}^d$ are in this case partitioned into N vectors $\mathbf{x}_i \in \mathbb{R}^3$ that represent the positions of the particles. The grid points $\mathbf{q}$ and the vectors $\boldsymbol{\nu}$ split correspondingly. The group $\mathbf{G}$ consists in the case of particles of equal spin, to which we restrict ourselves here, of the permutation matrices that interchange the parts $\mathbf{x}_i \in \mathbb{R}^3$ of a vector $\mathbf{x}$, that is, the positions of the particles. The sign function is again the usual sign of these permutations. In the general case, only permutations that interchange particles with spin of the same sign are taken into account.

One can proceed in this case first precisely as in the previous example. The ansatz functions are evaluated as $(N \times N)$-determinants with the matrix entries now themselves products of three univariate functions (6.2). These determinants are denoted

as Slater determinants in quantum physics. The order relation between the three-dimensional vectors $\mathbf{q}_i$ is now the lexicographic ordering. With this convention, the set of representatives S now consists of the grid points $\mathbf{q} = (\mathbf{q}_1, \ldots, \mathbf{q}_N)$ with parts

$$\mathbf{q}_1 \le \mathbf{q}_2 \le \ldots \le \mathbf{q}_N. \tag{7.5}$$

The diagonal points are those for which at least two of the $\mathbf{q}_i$ coincide.

The pairs $(\mathbf{q}, \boldsymbol{\nu})$ of grid points $\mathbf{q}$ and vectors $\boldsymbol{\nu}$ are now represented as sequences

$$(\mathbf{q}_i, \boldsymbol{\nu}_i), \quad i = 1, \ldots, N, \tag{7.6}$$

of pairs of three-dimensional grid points $\mathbf{q}_i$ and vectors $\boldsymbol{\nu}_i$. A pair $(\mathbf{q}, \boldsymbol{\nu})$ is diagonal if two of these pairs $(\mathbf{q}_i, \boldsymbol{\nu}_i)$ coincide. The set of representatives S_2 of the corresponding equivalence classes consists of the lexicographically ordered sequences (7.6).

Up to this point, the two examples look quite similar. However, there is an essential difference. Assume that only vectors $\boldsymbol{\nu}$ with nonnegative integer components are admissible. In the first case, for every $\mathbf{q} \in S$ there exists then a unique such vector $\boldsymbol{\nu}$ of minimum Euclidian or 1-norm with the property that $(\mathbf{q}, \boldsymbol{\nu})$ is a non-diagonal pair in the given set of representatives. This vector can be found treating groups of equal components of the grid point $\mathbf{q}$ separately. Only the one-dimensional shape functions $\nu = 0, 1, \ldots, d-1$ enter into this construction. Remembering the example of the shape functions (6.3), it is suggestive to attach exactly the function (6.1) respectively (6.10) assigned to this pair to the given grid point. A similar approach is not possible in the second case. The number of the corresponding integer vectors $\boldsymbol{\nu}$ can be huge there and can become completely intractable in many cases. More subtle strategies are needed then that will surely be subject of further studies.

References

[1] Babenko, K. I.: Approximation by trigonometric polynomials in a certain class of periodic functions of several variables. Soviet Math. Dokl. *1*, 672–675 (1960).
[2] Bungartz, H. J., Griebel, M.: Sparse grids. In: Acta Numerica, pp. 1–123. Cambridge: Cambridge University Press 2004.
[3] Flad, H. J., Hackbusch, W., Schneider, R.: Best N-term approximation in electronic structure calculations. II. Jastrow factors. Max-Planck-Institut für Mathematik in den Naturwissenschaften, Leipzig, Germany, Preprint 80/2005.
[4] Griebel, M., Hamaekers, J.: Sparse grids for the Schrödinger equation. M2AN (submitted).
[5] Le Bris, C.: Computational chemistry from the perspective of numerical analysis. In: Acta Numerica, pp. 363–444. Cambridge: Cambridge University Press 2005.
[6] Pople, J.: Nobel lecture: Quantum chemical models. Rev. Mod. Phys. *71*, 1267–1274 (1999).
[7] Smolyak, S. A.: Quadrature and interpolation formulas for tensor products of certain classes of functions. Dokl. Akad. Nauk SSSR *4*, 240–243 (1963).
[8] Yserentant, H.: On the multi-level splitting of finite element spaces. Numer. Math. *49*, 379–412 (1986).
[9] Yserentant, H.: Hierarchical bases. In: Proc. ICIAM 91 (O'Malley, R. E. et al., eds.), pp. 256–276. Philadelphia: SIAM 1992.
[10] Yserentant, H.: On the regularity of the electronic Schrödinger equation in Hilbert spaces of mixed derivatives. Numer. Math. *98*, 731–759 (2004).
[11] Yserentant, H.: Sparse grid spaces for the numerical solution of the electronic Schrödinger equation. Numer. Math. *101*, 381–389 (2005).
[12] Yserentant, H.: The hyperbolic cross space approximation of electronic wavefunctions. Numer. Math. (in print).

[13] Zenger, C.: Sparse grids. In: Parallel Algorithms for Partial Differential Equations. Proc. (Hackbusch, W., ed.), Kiel 1990, pp. 241–251. Notes on Numerical Fluid Mechanics, Vol. 31. Braunschweig Wiesbaden: Vieweg 1991.

H. Yserentant
Technische Universität Berlin
Institut für Mathematik
Straße des 17. Juni 136
10623 Berlin
Germany
e-mail: yserentant@math.tu-berlin.de