

# Chapter 1

## An Adapted Waveform Functional Calculus \*

Mladen Victor Wickerhauser<sup>†</sup>

### Abstract

We briefly survey how to use libraries of (orthonormal) bases of well-behaved waveforms, including wavelets and lapped orthogonal transforms, so as to obtain fast numerical algorithms for the expansion of functions and operators in these bases. The most important applications are fast approximate matrix multiplication, and application of matrices to vectors.

### 1 Nonstandard matrix form

By decomposing a matrix into its 2-dimensional best basis [6], we reduce the number of nonnegligible coefficients and thus reduce the computational complexity of matrix application.

For simplicity we will consider the periodic case only. Write  $\mathcal{W}$  for the collection of one-dimensional  $r$ -periodic wavelet packets [3] with  $r = 2^L$ . Let  $\psi_{sfp}$  be a representative wavelet packet of frequency  $f$  at scale  $s$  and position  $p$ , where  $0 \leq s \leq L$ ,  $0 \leq f < 2^s$ , and  $0 \leq p < 2^{L-s}$ . We may call this index set  $Z$ .

Let  $\mathcal{W}^1$  denote the space of wavelet packet sequences. This is an  $r \log_2 r$ -dimensional space, and there is a natural injection  $J^1 : \mathbf{R}^r \hookrightarrow \mathcal{W}^1$  given by  $J^1 x = \{\lambda_{sfp} = \langle x, \psi_{sfp} \rangle\}$  for  $x \in \mathbf{R}^r$ , the sequence of inner products with functions in  $\mathcal{W}$ . If  $B \subset Z$  indexes a basis, then  $J_B^1$ , which is  $J^1$  restricted to the  $B$  indices, is an isomorphism. By orthogonality, its inverse is a map  $R_B^1 : \mathcal{W}^1 \rightarrow \mathbf{R}^r$  defined at the left in 1. This map  $R_B^1$  can also be extended to the range of  $J^1$ , which is all of  $\mathcal{W}^1$ . We may call this extension  $R^1 : \mathcal{W}^1 \rightarrow \mathbf{R}^r$ ; it is defined at the right of 1.

$$(1) \quad R_B^1 \lambda = \sum_{(s,f,p) \in B} \lambda_{sfp} \psi_{sfp}; \quad R^1 \lambda = \sum_{(s,f,p) \in Z} \lambda_{sfp} \psi_{sfp}.$$

For any basis subset  $B$ ,  $R^1 J_B^1$  is the identity on  $\mathbf{R}^r$ , and  $J^1 R_B^1$  gives an orthogonal projection of  $\mathcal{W}^1$  onto the range of  $J^1$ .

By using tensor products we get another space  $\mathcal{W}^2 = \mathcal{W}^1 \otimes \mathcal{W}^1$  of sequences  $\{\mu_{sfp s' f' p'}\}$ . Objects in the space  $\mathbf{R}^{2r}$ , *i.e.*,  $r \times r$  matrices  $M$ , inject into this space  $\mathcal{W}^2$  in the obvious way:  $J^2 M = \{\mu_{sfp s' f' p'} = \langle M, \psi_{sfp} \otimes \psi_{s' f' p'} \rangle : \psi \in \mathcal{W}, (s, f, p), (s', f', p') \in Z, \}$ , where  $J^2 : \mathbf{R}^{2r} \hookrightarrow \mathcal{W}^2$  denotes the injection. If  $B$  is a basis subset of  $Z \times Z$ , then the restriction  $J_B^2$  of  $J^2$  to the  $B$ -indexed components is also injective, and  $J_B^2$  is an isomorphism.

The *nonstandard form of a matrix*  $M$  is the organized array of coefficients  $J_B^2 M$ . We can perform linear algebra in the nonstandard form just as well as in the standard form. The advantage is that  $B$  may be chosen to maximize the number of negligible coefficients

---

\*Supported in part by AFOSR award F49620-92-J-0106 and NSF grant DMS-9302828.

<sup>†</sup>Department of Mathematics, Washington University in St. Louis, Missouri 63130 USA

and thus reduce the cost of approximate computations.  $B$  can be a complicated mixture of pieces from the two  $Z$  factors.

## 2 Applying a matrix to a vector

Let  $M$  be an  $r \times r$  matrix and suppose  $B \subset Z$  is a basis subset. Then we have:

$$(2) \quad \langle Mx, \psi_{sfp} \rangle = \sum_{(s',f',p') \in B} \langle M, \psi_{sfp} \otimes \psi_{s'f'p'} \rangle \langle x, \psi_{s'f'p'} \rangle \stackrel{\text{def}}{=} \sum_{(s',f',p') \in B} \mu_{sfp s'f'p'} \lambda_{s'f'p'}.$$

This generalizes to a linear action  $\lambda \mapsto \sum_{(s',f',p') \in Z} \mu_{sfp s'f'p'} \lambda_{s'f'p'}$  by  $\mu \in \mathcal{W}^2$  on  $\mathcal{W}^1$ . Now, images of matrices form a proper submanifold of  $\mathcal{W}^2$ . Likewise, images of vectors form a proper submanifold of  $\mathcal{W}^1$ . We can lift the action of  $M$  on  $x$  to these larger spaces via the commutative diagram at the left of 3.

$$(3) \quad \begin{array}{ccc} \mathcal{W}^1 & \xrightarrow{J_B^2 M} & \mathcal{W}^1 \\ J^1 \uparrow & & \downarrow R^1 \\ \mathbf{R}^r & \xrightarrow{M} & \mathbf{R}^r \end{array} \qquad \begin{array}{ccc} \mathcal{W}^2 & \xrightarrow{J_B^2 N} & \mathcal{W}^2 \\ J^2 \uparrow & & \downarrow R^2 \\ \mathbf{R}^{2r} & \xrightarrow{N} & \mathbf{R}^{2r} \end{array}$$

## 3 Composing operators

Suppose that  $M$  and  $N$  are  $r \times r$  matrices, and  $B$  is a basis subset. We have the identity

$$(4) \quad \langle NM, \psi_{sfp} \otimes \psi_{s'f'p'} \rangle = \sum_{(s'',f'',p'') \in B} \langle N, \psi_{s''f''p''} \otimes \psi_{s'f'p'} \rangle \langle M, \psi_{sfp} \otimes \psi_{s''f''p''} \rangle.$$

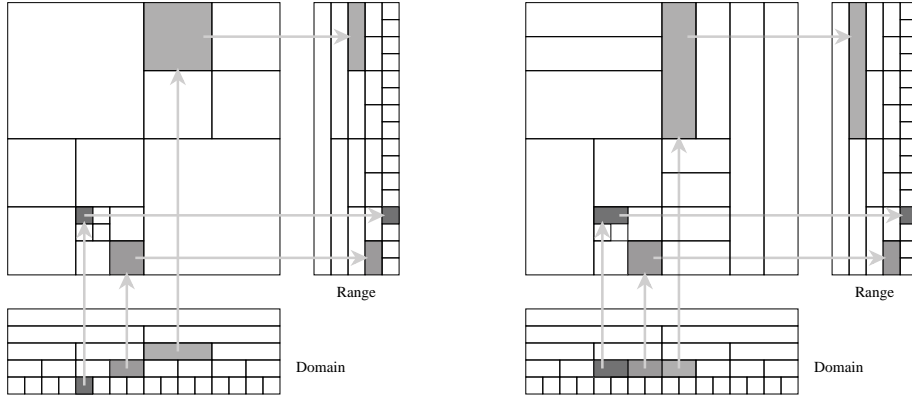
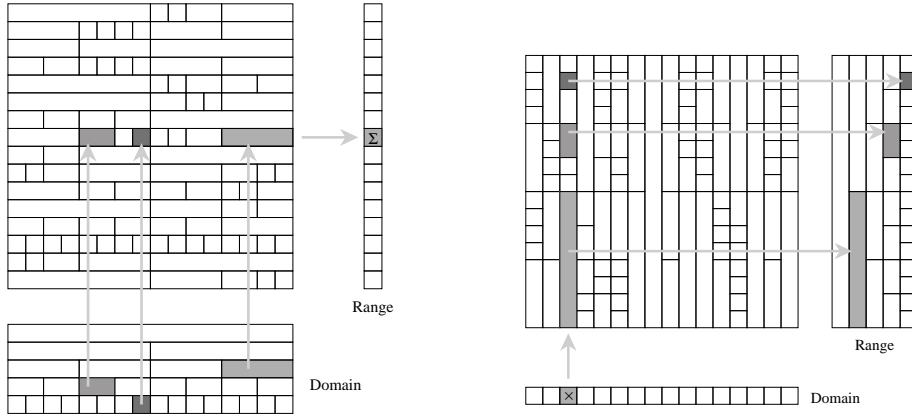
This generalizes to an action  $\nu \mapsto \sum_{(s'',f'',p'') \in Z} \mu_{sfp s''f''p''} \nu_{s''f''p'' s'f'p'}$  of  $\mathcal{W}^2$  on  $\mathcal{W}^2$ . Using  $J^2$ , we can lift multiplication by  $N$  to an action on these larger spaces via the commutative diagram at the right in 3.

## 4 Examples

We illustrate the above ideas by considering nonstandard multiplication by matrices of order 16. Using wavelet packets of isotropic dilations, we obtain the *best isotropic basis* action depicted in the left half of Figure 1. The domain vector is first injected into  $\mathcal{W}^1 = \mathbf{R}^{16 \log 16}$  by a complete wavelet packet analysis. It then has its components multiplied by the best-basis coefficients in the matrix square at center, as indicated by the arrows. The products are summed into a wavelet packet tree at right, and then projected onto the range space  $\mathbf{R}^{16}$  by  $R^1$ , or wavelet packet synthesis.

By reading off the explicit frequency and position of wavelet packet components, we can see explicitly how frequencies and positions are mixed. It also is possible for each scale to affect any another, but since input and output blocks must have the same scale, this action is not immediately readable from the nonstandard matrix coefficients.

If we allow mixed  $x$  and  $y$  scales, we get the *best tensor basis* nonstandard representation. Finding the best basis becomes harder, but the multiplication algorithm is practically the same. Again  $J^1$  develops the complete wavelet packet analysis of the input. Blocks in the nonstandard matrix send input blocks to possibly larger or smaller output blocks, where they are synthesized by  $R^1$  into the output. This is depicted in the right half of Figure 1,


 FIG. 1.  $16 \times 16$  Matrix multiplication in isotropic and anisotropic best bases

 FIG. 2.  $16 \times 16$  Matrix multiplication in best bases row by row and column by column

for a matrix of order 16. Such an expansion explicitly shows how one scale affects another, as well as how position and frequency are mixed.

Rather than choose a basis set among the two-dimensional wavelet packets, we may also consider matrix multiplication to be a list of inner products and expressing each row of the matrix in its individual best basis. We expand each input vector in a complete wavelet packet analysis with  $J^1$ , then evaluate the inner products row by row, up to any desired accuracy. Schematically, this accelerated inner product may be depicted as in the left half of Figure 2. Then  $R^1$  is replaced by a simple sum.

We may also expand matrix columns in their individual best-bases and consider matrix multiplication to be the superposition of a weighted sum of these expansions. This *best column basis* algorithm is depicted schematically in the right half of Figure 2. The input value is sprayed into the output wavelet packet synthesis tree, then the output is reassembled from those components with  $R^1$ .

## 5 Operation counts

Suppose that  $M$  is a non-sparse  $r \times r$  matrix. Ordinary multiplication of a vector by  $M$  takes  $O(r^2)$  operations. On the other hand, the injection  $J^2$  will require  $O(r^2 \log r)$  operations, and each of  $J^1$  and  $R^1$  require  $O(r \log r)$  operations. For a fixed basis subset  $B$  of  $\mathcal{W}^2$ ,

the application of  $J_B^2 M$  to  $J^1 v$  requires at most  $\#|J_B^2 M| = O(r^2)$  operations, where  $\#|U|$  denotes the number of nonnegligible coefficients in  $U$ . Thus nonstandard multiplication costs an initial investment of  $O(r^2 \log r)$ , plus at most an additional  $O(r^2)$  per right-hand side; asymptotically, this is  $O(r^2)$  per vector just like the standard method.

We can obtain lower complexity if we take into account the finite accuracy of our calculation. Given a fixed matrix  $M$ , write  $M_\delta$  for the same matrix with all coefficients set to 0 whose absolute values are less than  $\delta$ . By continuity, for every  $\epsilon > 0$  there is a  $\delta > 0$  such that  $\|M - M_\delta\| < \epsilon$ . Given  $M$  and  $\epsilon$  as well as a library of wavelet packets, we can choose a basis subset  $B \subset Z \times Z$  so as to minimize  $\#|(J_B^2 M)_\delta|$ . The choice algorithm has complexity  $O(r^2 \log r)$ , as shown in [4]. For certain classes of operators [1, 2], we can arrange that for every fixed  $\delta > 0$  we have  $\#|(J_B^2 M)_\delta| = O(r \log r)$ , with the constant depending, of course, on  $\delta$ . Call this class the *sparsifiable* matrices  $\mathcal{S}$ . Then finite-precision multiplication by sparsifiable matrices has asymptotic complexity  $O(r \log r)$ .

The best row basis algorithm also has complexity  $O(r \log r)$  for sparsifiable matrices, since it requires  $O(r \log r)$  operations to apply  $J^1$  to the input vector, then  $O(\log r)$  multiplies to evaluate each of the  $r$  inner products. Finding the best basis for each of the  $r$  rows of the matrix requires an initial investment of  $O(r^2 \log r)$  operations.

Standard multiplication of  $r \times r$  matrices  $N$  and  $M$  has complexity  $O(r^3)$ . For nonstandard composition, the complexity of injecting  $N$  and  $M$  into  $\mathcal{W}^2$  is  $O(r^2 \log r)$ . The action of  $J_B^2 N$  on  $J^2 M$  has complexity  $O(\#|J_B^2 N| r \log r)$ . Since the first factor is  $r^2$  in general, the complexity of the exact algorithm is  $O(r^3 \log r)$  for generic matrices, reflecting the extra cost of conjugating into the basis set  $B$ .

For approximate nonstandard composition, the complexity is  $O(\#|(J_B^2 N)_\delta| r \log r)$ . For the sparsifiable matrices, this can be reduced by a suitable choice of  $B$  to a complexity of  $O(r^2 \log r)$  for the complete algorithm. Since choosing  $B$  and evaluating  $J_B^2$  each have this complexity, it is not possible to do any better by this method.

## References

- [1] G. Beylkin, R. R. Coifman, and V. Rokhlin, *Fast wavelet transforms and numerical algorithms I*, Communications on Pure and Applied Math, XLIV (1991), pp. 141–183.
- [2] B. Bradie, R. R. Coifman, and A. Grossmann, *Fast numerical computations of oscillatory integrals related to acoustic scattering, I*, Applied and Computational Harmonic Analysis, 1 (1993), pp. 94–99.
- [3] R. R. Coifman, Y. Meyer, S. R. Quake, and M. V. Wickerhauser, *Signal processing and compression with wavelet packets*, in Meyer and Roques [5], pp. 77–93. ISBN 2-86332-130-7.
- [4] R. R. Coifman and M. V. Wickerhauser, *Entropy based algorithms for best basis selection*, IEEE Transactions on Information Theory, 32 (1992), pp. 712–718.
- [5] Y. MEYER AND S. ROQUES, eds., *Progress in Wavelet Analysis and Applications*, Proceedings of the International Conference “Wavelets and Applications”, Toulouse, France, 8–13 June 1992, Observatoire Midi-Pyrénées de l’Université Paul Sabatier, Editions Frontières. ISBN 2-86332-130-7.
- [6] M. V. Wickerhauser, *Computation with adapted time-frequency atoms*, in Meyer and Roques [5], pp. 175–184. ISBN 2-86332-130-7.