SOLVING PDEs USING WAVELETS

Oleg V. Vasilyev, David A. Yuen, and Samuel Paolucci

Department Editors:
Ralf Gruber
gruber@sic.epfl.ch
Jacques Rappaz
rappaz@dma.epfl.ch

Wavelet theory, developed mostly over the last 15 years, has generated a tremendous interest in many areas of research in mathematics, physics, computer science, and engineering. However, most applications of wavelets have focused on analyzing data and using wavelets as a tool for data compression.

The application of wavelets to the solution of difficult partial differential equations (PDEs) arising in various areas of physics and engineering has been very limited. However, most applications of wavelets have focused on analyzing data and using wavelets as a tool for data compression. The application of wavelets to the solution of difficult partial differential equations (PDEs) arising in various areas of physics and engineering has been very limited. This is mostly because wavelet-based numerical algorithms are in their infancy.

A large communication gap must be diminished before researchers can take advantage of this new mathematical tool. In this article we hope to convince the reader that wavelets are a useful tool in large-scale numerical modeling of physical problems and to bridge the communication gap by providing a short description of wavelet-based numerical algorithms and their advantages over conventional numerical methods.

Common numerical techniques for the solution of physical problems mostly fall into three classes: finite-difference and finite-volume methods, finite-element methods, and spectral methods. Sometimes the latter two methods are considered as subsets of the method of weighted residuals.

Briefly, the finite-difference method consists in defining the different unknowns by their values on a discrete (finite) grid and in replacing differential operators by difference operators using neighboring points. In the method of weighted residuals, the unknown solution is approximated by a linear combination of a set of linearly independent trial or basis functions. In the finite-element methods, the trial functions are only piecewise continuous and nonvanishing on certain elements of the domain, whereas the spectral method utilizes basis functions that are infinitely differentiable and nonvanishing on the whole domain.

The method of weighted residuals proceeds by substituting the approximation into a differential equation and imposing the condition that the integral over the residual, weighted by some weighting function, is zero. Different choices of the weighting function give rise to the different methods, which collectively are known as methods of weighted residuals. Two well-known examples are the collocation method and the Galerkin method. In the collocation method, the weighting functions are taken to be Dirac delta functions. In the Galerkin method, the weighting functions are chosen from the same family as the test functions.

If the solution of a physical problem has regular features, any of these numerical techniques can be applied. However, in many physical problems there exists a multiplicity of very different spatial and temporal scales in the solution. This situation arises in such cases as strongly time-dependent non-Newtonian convection, the formation of shock waves in compressible gas flow, pattern formation in hydrodynamic systems, turbulent flow around bluff bodies, and dendritic crystal growth.

The particular attribute of multiple spatial scales, which possibly change over time, puts great strain on conventional numerical methods. Spectral methods have problems in capturing large irregularities of the solutions. The main difficulty of existing adaptive finite-difference or finite-element methods is in developing a computationally efficient and robust procedure that dynamically adapts the computational grid to local structures of the solution.

The basic idea behind wavelet decomposition is to represent a function in terms of building blocks, called wavelets, that are localized in both position and scale. Good wavelet-localization properties in physical and wavenumber spaces are to be contrasted with the spectral approach, which employs infinitely differentiable functions but with global support and small discrete changes in the resolution. On the other hand, finite-difference, finite-volume, and finite-element methods use bases with small compact support but poor continuity properties. Wavelets appear to combine the advantages of both spectral and finite-difference bases. We can expect that numerical methods based on wavelets will attain both good spatial and spectral resolution.

Methodologies

The most important property of wavelet analysis is that a function is decomposed in terms of basis functions having
various discrete scales and locations. These basis functions are constructed by the discrete (typically dyadic) dilation and translation of a single function that has good localization properties in both physical as well as wavenumber spaces. In other words, wavelet analysis can be viewed as a multilevel or multiresolution representation of a function, in which each level of resolution consists of basis functions having the same scale but located at different positions.

Wavelet theory is often discussed in light of the multiresolution analysis introduced by Meyer\(^1\) and Mallat.\(^5\) We choose to introduce a wavelet decomposition in a different way that is easier to understand in the context of partial differential equations. Here we describe only the essential points necessary for introducing the wavelet concept and wavelet-based numerical algorithms.

In one-dimensional space, the wavelet basis consists of a double-indexed function set \( \{ \psi^j_k(x); j, k \in \mathbb{Z}, x \in R \} \) given by

\[
\psi^j_k(x) = a_j^{-1/2} \psi \left( \frac{x - b_j^k}{a_j} \right),
\]

where \( \psi(x) \) is a one-dimensional “mother” wavelet and \( \psi^j_k(x) \) is a wavelet of scale \( a_j = a_0 2^j \), located at position \( b_j^k(x) = a_j k \). We use the superscript to denote the level of resolution and the subscript to denote the location in physical space (with the exception of \( a_0 \)). Examples of wavelets are shown in Fig. 1 (a, b).

In multiple dimensions, wavelet bases can be constructed as a tensor product of one-dimensional bases. Thus an \( n \)-dimensional wavelet basis is given by

\[
\psi_j^k(x) = \prod_{i=1}^n a_{x_i}^{-1/2} \psi \left( \frac{x_i - b_{x_i}^k}{a_{x_i}} \right),
\]

where \( a_{x_i} \) denote \( n \)-dimensional vectors, e.g. \( x = (x_1, \ldots, x_n) \), \( k = (k_1, \ldots, k_n) \), \( b_x = (b_{x_1}, \ldots, b_{x_n}) \), and \( a_{x_i}, b_{x_i} \) and \( b_{x_i}^k \) (i = 1, ..., n, j, k \( \in \mathbb{Z} \)) are wavelet scales and locations at the \( j \) level of resolution.

Let us consider a function \( u(x) \) defined on a closed \( n \)-dimensional domain \( \Omega \). Let \( j = 0 \) be the coarsest level of resolution with scales defined by the corresponding dimensions of the domain \( \Omega \). The numerical resolution is constrained by memory limits and computational times. Thus \( u(x) \) can be approximated only up to a certain level of resolution. Let \( j = J \) be the finest level of resolution for the approximation of a function. Due to the compact or effectively compact wavelet support, at each level of resolution \( j = 0, \ldots, J \) there exists a finite \( n \)-dimensional integer set \( \mathbb{Z}_Q \) such that function \( u(x) \) can be approximated as

\[
u_j^j(x) = \sum_{j=0}^{J} \sum_{k \in \mathbb{Z}_Q} c_j^k \psi_j^k(x).
\]

The absolute value of the wavelet coefficient \( c_j^k \) appearing in the approximation \( (2) \) depends upon the local regularity of \( u(x) \) in the neighborhood of the wavelet location. A good approximation is maintained even when wavelets whose coefficients are below a certain threshold \( \varepsilon \) are omitted, that is, \( |c_j^k| < \varepsilon \), and only those wavelets whose coefficients are above the threshold, that is, \( |c_j^k| \geq \varepsilon \), are kept.

In Fig. 2 we sketch the locations of wavelets used in a four-level approximation of a function. Figure 2 also shows schematically the locations of wavelets whose coefficients are above a given threshold parameter for a function that has a sharp transition in the middle of the domain. This schematic example illustrates the considerable savings we can achieve by keeping in the approximation only significant wavelets and omitting those that play an insignificant role in the approximation.

The next issue is how to compute wavelet coefficients for a given function \( u(x) \). With orthonormal wavelet bases, the natural choice is to use the wavelet-Galerkin projection to obtain a set of wavelet coefficients. This approach is easy to use in the case of a periodic function defined in a rectangular domain. Nevertheless, the efficient implementation of the wavelet-Galerkin projection in the case of general geometry is still an open issue, even though different possibilities have been studied.\(^6\)\(^-\)\(^8\)

An alternative to the wavelet-Galerkin projection is to use a collocation method analogous to the spectral-collocation method. In this approach wavelet coefficients are found based on the values of a function at certain locations, called collocation points. In a wavelet-collocation algorithm a set of...
Collocation Transform (FWCT), consists in finding wavelet coefficients based on values of a function at collocation points. The important feature of the FWCT and FIWCT is that once wavelet coefficients are known, derivatives of the approximate function at the collocation points are found by differentiating (2) and evaluating the resulting equation at collocation points. This collocation method therefore formalizes the procedure of finding derivatives of a function in a two-stage process. The first stage, which we call the Fast Wavelet Collocation Transform (FWCT), consists in finding wavelet coefficients based on values of a function at collocation points. The second stage, which we call the Fast Inverse Wavelet Collocation Transform (FIWCT), consists in mapping the values of wavelet coefficients into the values of the derivatives of the approximate function at the collocation points. The important feature of the FWCT and FIWCT is that they do not change even in the case in which some wavelets are omitted from the approximation. Furthermore, the computational cost of both FWCT and FIWCT is \( O((j+1) M_W N_f) \), where \( N_f \) is the total number of wavelets kept in the approximation, \( M_W \) is a parameter characterizing the wavelet support, and \( n \) is the dimensionality of the wavelet.

Now we have all the necessary components for constructing a dynamically adaptive wavelet-collocation numerical method for solving PDEs. Many systems of PDEs that arise in physics and engineering can be written in the following form:

\[
\frac{\partial u}{\partial t} = F(x,t,u,\nabla u), \quad \text{Eq. (3a)}
\]

\[
0 = \Phi(x,t,u,\nabla u) \quad \text{Eq. (3b)}
\]

where Eq. (3a) describes the time evolution of a vector function \( u \) and Eq. (3b) represents boundary conditions and possibly algebraic/differential constraints.

There are two different ways to use the wavelet decomposition (2) for the numerical solution of (3). The first approach, commonly known as the wavelet-Galerkin (WG) method, is to assume that wavelet coefficients are functions of time, then substitute this decomposition for each component of the solution into (3) and use a Galerkin projection to derive the nonlinear system of ordinary differential-algebraic equations, which describe the time evolution of wavelet coefficients. The second approach is to evaluate (3) at collocation points and obtain a system of nonlinear ordinary differential-algebraic equations describing the evolution of the solution at these collocation points. This approach is called the wavelet-collocation (WC) method.

Although WC algorithms are much more effective in the treatment of general boundary conditions and nonlinearities,\(^9\)-\(^12\) they have not enjoyed as much popularity as WG methods.\(^13\)-\(^17\) One reason is that WC algorithms do not fall exactly into the theory of wavelet multiresolution approximation. Often they even utilize scaling functions (sometimes called interpolating wavelets)\(^12\) instead of wavelets\(^11\),\(^12\) or frames instead of orthonormal bases.\(^9\)-\(^11\) Examples of a scaling function and interpolating “wavelet” are shown in Fig. 1 (c, d).

The reasons why WG and WC methods can be combined into the same class of numerical methods are the multilevelness of the approximation, good spatial and spectral localization properties of basis functions, and the existence of fast algorithms for obtaining wavelet coefficients. In distinguishing WC from WG methods, we shall call the former multilevel WC algorithms and the latter multiresolution WG methods.

In order for the numerical algorithm to be able to resolve all structures appearing in the solution and yet be efficient in terms of minimizing the number of unknowns, the basis of active wavelets and, consequently, the computational grid for the WC algorithm should adapt dynamically in time to reflect local changes in the solution. Such adaptation of the wavelet basis or computational grid is based on the analysis of wavelet coefficients. The contribution of a wavelet to the approximation is considerable if and only if the nearby structures of the solution have a size comparable with the wavelet scale. Thus, we may drop the large number of fine-scale wavelets with small coefficients in the regions where the solution is smooth.

In the WC method, every wavelet is uniquely associated with a collocation point. Consequently, the collocation point should be omitted from the computational grid if the associated wavelet is omitted from the approximation. This property of the multilevel wavelet approximation allows local grid refinement up to an arbitrarily small scale without a drastic increase in the number of collocation points.

To ensure accuracy, the basis should also consist of wavelets whose coefficients can possibly become significant during the period of time when the basis and consequently the computational grid remain unchanged. Thus, at any instant of time, the basis should include not only wavelets whose coefficients are above a prescribed threshold parameter \( \varepsilon \), but also the surrounding wavelets. In other words, at any instant of
time, the basis should include wavelets belonging to an adjacent zone of wavelets for which the magnitude of coefficients is within an a priori prescribed threshold. The concept of adjacent zone is illustrated in Fig. 3 by showing the adjacent zone of the one-dimensional wavelet $\psi_k$. The basic hypothesis motivating the dynamic adaptation procedure, first suggested by Liandrat and Tchiamichian for the WG method, is that during a time-integration step, the domain of wavelets with significant coefficients does not move in phase space beyond the adjacent zone. With such an algorithm, the grid of WC points is dynamically adapted in time and follows the local structures that appear in the solution.

By omitting wavelets with coefficients below a threshold parameter $\varepsilon$, we automatically control the error of the approximation. Thus the WC method has another important feature: active control of the local accuracy of the solution. The smaller $\varepsilon$ is chosen, the smaller the error of the solution is. In typical applications the value of $\varepsilon$ varies between $10^{-2}$ and $10^{-4}$, assuming that the unknown dependent variables have been properly normalized. We also note that the larger the set value of $\varepsilon$, the fewer the number of grid points.

The algorithm can utilize different criteria for adaptation of the collocation grid. For example, we can compose a computational grid based on the analysis of wavelet coefficients of both the function and its derivatives. If a system of equations is solved, the adaptation of the computational grid should be based on the analysis of wavelet coefficients associated with all dependent variables. The computational grid is constructed as a union of grids corresponding to each dependent variable. Note that the algorithm can be easily extended to the case in which each variable is treated on a separate computational grid. The mapping from one grid to another can be achieved via wavelet interpolation. This may be very important for problems where scales associated with different variables are considerably different, such as in double-diffusive convection with thermal and chemical fields.

**Applications and examples**

Most problems in physics and engineering requiring solution of PDEs can be divided into two general classes. The first class of problems is characterized by a fairly uniform distribution of spatial scales. Note that separation of scales can be large, as long as they are uniformly distributed: both small and large scales should be present at every location within the domain of interest. These problems generally do not require a time-adaptive and/or nonuniform computational grid and can be solved using a variety of conventional numerical algorithms with different degrees of computational efficiency and accuracy. The computational complexity of these problems increases with the growth of the separation of scales that need to be resolved.

The second class of problems is characterized by the presence of a wide range of spatial and temporal scales that are not distributed uniformly in space and time. In order to solve these problems in a computationally efficient way, the computational grid should adapt dynamically in time. For the numerical algorithm to be robust, grid adaptation—that is, local grid refinement and coarsening—should be based on the local demands of the solution and not on ad hoc assumptions. The multilevel structure of wavelet decomposition provides a simple and effective framework for spatially adaptive algorithms. The adaptation is achieved by retaining only those wavelets whose coefficients are greater than an a priori given threshold. Thus, high-resolution computations are carried out only in the regions where sharp transitions occur. This adaptive procedure is ideally suitable for handling problems with localized structures that might occur intermittently anywhere in the computational domain or change their locations and scales in space and time.

The implementation of conventional adaptive algorithms in such problems is costly because the grids vary drastically within short space or time intervals. In addition, the use of conventional algorithms on a uniform grid is impractical. Thus, the main advantage of the dynamically adaptive WC
algorithm is that it will employ far fewer grid points (degrees of freedom) than the other algorithms when applied to problems with a great diversity of spatial/temporal scales. In addition, the computational grid can be refined locally to an arbitrarily small size grid.

We emphasize here that the adaptation of the computational grid does not require additional efforts and consists merely in turning on and off wavelets at different locations and scales. Furthermore, the compression of the solution is performed dynamically as opposed to a posteriori as is done in data analysis. These features make adaptive wavelet algorithms attractive in solving tough multiscale problems in physics and engineering.

Thermoacoustic waves. As a first example, we consider a thermoacoustic wave problem for illustrating the prowess of the adaptive WC algorithm in treating problems with localized structures, whose positions and sharpness change both spatially and temporally. Here we discuss only the evolution of the solution and the physical formulation of the problem. For the mathematical formulation we refer to Ref. 18 for the one-dimensional case and Ref. 19 for the two-dimensional problem.

Consider a compressible ideal gas in a rectangular cavity with rigid walls. The gas is initially quiescent at a uniform pressure and temperature. An abrupt temperature change in localized regions of the walls causes gas to expand in the immediate neighborhood of those regions. This expansion generates pressure waves, which propagate away from these regions. These thermally generated waves are referred to as thermoacoustic (TAC) waves. TAC waves, which decay on large time scales due to thermal and viscous diffusion, propagate, interact with each other, and reflect from the walls, creating complicated two-dimensional patterns. The processes of reflection and diffusion continue until the waves die out and a quiescent thermal-conduction state is achieved.

The efficiency of the numerical algorithm was illustrated in Ref. 9, in which results for a one-dimensional TAC wave obtained by the adaptive WC method were compared with the results of Huang and Bau. In their one-dimensional finite-difference numerical algorithm, Huang and Bau required 6000 evenly spaced grid points to achieve proper resolution of the physics. In the calculation using the adaptive WC method, the total number of collocation points did not exceed 195 at any time, and adaptive WC maintained proper resolution of all scales. In addition, small spurious oscillations had been observed in the solution of Huang and Bau at small times and ascribed to unresolved scales associated with the initial large gradients, and these oscillations were not present in the results obtained by the WC algorithm.

The two-dimensional TAC wave problem involves five unknowns, four partial differential equations (continuity, x- and y-momentum, and energy), and one algebraic equation (equation of state). The adaptation of the computational grid was based on the analysis of coefficients associated with all dependent variables. The irregular grid of WC points was constructed as a union of irregular grids corresponding to each dependent variable.

The solution of the problem (pressure field) corresponding to the case of four symmetrically placed localized temperature disturbances at the wall is shown in Fig. 4. These temperature sources generate four TAC waves, which initially propagate independently of each other until they start to interact. After interaction, the waves go through each other, creating complicated patterns.

The dynamical adaptation of the computational grid is shown in Fig. 5. As time progresses, this grid adapts to the multiple local structures appearing in the solution. Solved on a uniform nonadaptive computational grid, this problem would have required more than $10^6$ grid points, whereas in the present calculation, the total number of collocation points did not exceed $10^3$. This reduction in complexity illustrates the prowess and compression efficiency of the dynamically adaptive WC algorithm.

Viscoelastic deformation. As a second example, we consider a problem of geophysical relevance, viscoelastic deformation involving both steep viscosity stratification and density variations. This example illustrates the ability of the adaptive WC method to solve problems with localized structures and sharp gradients in physical properties. The viscoelastic-flow process involves a thin, highly viscous upper boundary layer (lithosphere) that interacts with a highly variable viscous interior (the mantle) associated with a rising diapir. This is modeled kinematically by a rising small low-density sphere with a viscosity considerably lower than the ambient mantle.

The viscoelastic model problem involves six unknowns: two velocity components, pressure, and three components of the stress tensor. Six partial differential equations describe the temporal evolution of these variables. The computational grid is adapted in the same way as the previous example; that is, it is based on the analysis of wavelet coefficients associated with all dependent variables. The irregular grid of WC points is constructed as a union of irregular grids corresponding to...
The solution for the $\tau_{xx}$ component of the stress tensor and the corresponding computational grid is illustrated in Fig. 6 for two different times. The high-stress region is observed at the top 8% of the domain. Note that the computational grid is very fine only in regions where small-scale features or large gradients are present. The use of the dynamically adaptive WC method allows us to conduct the calculations for thinner layers and higher viscosity contrast (up to $10^5$), close enough for realistic mantle-lithospheric interaction. The multilevel wavelet approximation allows local grid refinement up to an arbitrary small scale without a drastic increase in the number of collocation points. The total number of collocation points at any instant of time does not exceed $10^4$. If we repeated the same calculation on a nonadaptive two-dimensional grid using conventional numerical algorithms, it would require more than $10^6$ grid points.

Discussion and future work

Efficient, dynamically adaptive wavelet-based numerical algorithms have recently been developed for the solution of nonlinear PDEs. These algorithms are competitive with classical methods and have a number of attractive special features. The computational grid or wavelet basis can efficiently adapt to local irregularities of the solution in order to resolve regions of large gradients. The algorithms can be applied to problems with wide ranges of temporal and spatial scales. Moreover, it is possible to actively control the relative error of the solution by prescribing a threshold parameter.

In spite of the progress thus far, the wavelets approach could still be improved. Development of an implicit time-integration algorithm would be helpful. After the time discretization is applied, problem (3) reduces to essentially an elliptic problem. The main difference between the algorithm for elliptic problems and that for time integration is that the computational grid is known for the latter.

In the development of an efficient adaptive elliptic solver, the multilevel structure of the wavelet approximation provides a natural way to obtain the solution on an optimal grid—namely, by starting calculations on a coarse grid and progressively adding wavelets at finer scales of resolution. This procedure results in the resolution of all the scales present in the solution and reduction of the total number of unknowns.

Due to the local character of the adjacent zone, between two successive grid changes the level of resolution increases at most by one. This local increase of the level of resolution guarantees that the solution on the new grid will not be too far off the solution on the old grid and that fewer iterations will be necessary to obtain the solution. Most of the iterations will be spent on the coarser levels. By the time the finest level of resolution is reached, a good initial guess of the solution will be available, so that only a few iterations will be needed.

In order to accelerate the convergence of the iterative solver at each computational grid, a procedure analogous to the multigrid algorithm can be utilized. Multilevel iterations can be considered as a preconditioner for the iterative algorithm. Due to the sparseness of the matrices involved in the multilevel wavelet approximation, it is natural to expect that an iterative algorithm will be much more efficient than direct solvers. Elliptic solvers for WG algorithms can be developed along the same lines.

Another area of possible improvement of existing wavelet-based numerical algorithms is their extension to complex geometries. This type of improvement should be based on the locality of the support of the basic wavelet. Each complex domain can be locally embedded within a regular domain. Wavelets can be located the same way as when dealing with simple geometries. There will be two classes of wavelets connected with the irregular domain: wavelets located within the domain and wavelets located outside of the domain. The collocation points for internal wavelets can be chosen to be the locations of the wavelets themselves, whereas the collocation points for external wavelets can be chosen on the boundaries of the domain. Thus, the problem reduces to a differential-algebraic system of equations, which can be solved by the methods described above depending on whether it is an elliptic or a time-evolution problem. The extension of WG algorithms to complex geometries is not as straightforward as in the case of the WC method, and at this time it is another potential weak point of WG algorithms.

Wavelets have opened up new horizons in numerical analysis. In this article, we have discussed the main advantages of adaptive wavelet-based algorithms and the benefits of using them in certain situations. Although we do not wish to suggest that wavelet-based approaches will replace all existing numerical algorithms, the application of wavelets to solving PDEs is a promising development. Researchers should not shy away from applying wavelets to complicated physics and engineering problems.

Acknowledgments

We thank Yuri Podladchikov for discussions and the editors Lewis Holmes and Ralf Gruber for their encourage-
References