ground water

The Nested Superblock Approach for Regional-Scale Analytic Element Models

by James R. Craig[1], Igor Janković[2], and Randal Barnes[3]

Abstract

A new approach is presented for improving the computational efficiency of regional-scale ground water models based on the analytic element method (AEM). The algorithm is an extension of the existing "superblock" algorithm, which combines the effects of multiple analytic elements into Laurent series and Taylor series (superblock expansions). With the new "nested superblock" formulation, Laurent series are nested in a hierarchical (quad-tree) data structure with direct mathematical relationships between parent and child superblock coefficients. Nested superblocks significantly accelerate the evaluation of the complex potential and discharge function in models that contain a large number of analytic elements at multiple scales. This evaluation process, the primary computational cost of AEM models, is required to determine the element coefficients, generate contour plots, and trace pathlines. The performance of the nested superblocks is demonstrated with a simplified model based on the Lake Ontario watershed geometry comprising thousands of hydrogeologic features at multiple geographic scales.

Introduction

The analytic element method (AEM), discussed in Strack (1989), Haitjema (1995), and Fitts (2002), is an alternative to finite-difference and finite-element methods for the modeling of regional (two-dimensional [2D]) ground water flow. The AEM may also be used to simulate other physical phenomena (e.g., electric, magnetic, and thermal conduction) and three-dimensional (3D) flow. This paper deals with regional ground water flow only.

The irrotational and divergence-free portion of the AEM solution is expressed here in terms of complex potential $\Omega(z) = \Phi(z) + i\Psi(z)$ and the discharge function $W(z) = Q_x(z) - iQ_y(z)$ where Φ is the discharge potential, Ψ is the stream function, Q_x and Q_y are components of the discharge vector, and $z = x + iy$ is a location in the complex plane. The method is superposition based: Ω (or W) is expressed as a sum of complex potentials (or discharge functions) due to distinct hydrologic features (elements). Q_x and Q_y components that account for divergent and rotational flow may be added to this solution. The computational expense of the method is proportional to the number of elements but not their size. Hence, the method may be considered as scale independent.

Analytic elements that correspond to hydrologic features contain coefficients that are computed from internal and external boundary conditions. For example, coefficients of an analytic element included to model a section of a river are computed by requiring a certain head distribution along the element. The coefficients in analytic element models have traditionally been calculated using explicit algorithms in which the coefficients of all elements are computed simultaneously, by solving a linear system of equations that contains all unknown coefficients. An alternative, proposed by Janković and Barnes (1999), is to compute coefficients iteratively, for one element at a time. Element coefficients are computed simultaneously (by solving a linear system of equations that contains only element coefficients) before moving to the

[1]Corresponding author: Department of Civil Engineering, University of Waterloo, 200 University Ave West, Waterloo, Ontario, Canada, N2L3G1; (519) 888-4567 x7554; fax (519) 888-4349; jrcraig@uwaterloo.ca

[2]Department of Civil, Structural and Environmental Engineering, State University of New York at Buffalo, 231 Jarvis Hall, Buffalo, NY 14260-4400; (716) 645-2839 x2328; fax (716) 645-3667; ijankovi@eng.buffalo.edu

[3]Department of Civil Engineering, University of Minnesota, CivE 234, Minneapolis, MN 55455-0116; barne003@umn.edu

Received June 2003, accepted August 2004.

Copyright © 2005 The Author(s)
Journal compilation © 2006 National Ground Water Association

doi: 10.1111/j.1745-6584.2005.00081.x

next element. As soon as element coefficients are computed, they are used (within the same iteration) for other elements. The iterations between the elements are equivalent to iterations of a Gauss-Seidel algorithm.

The procedure of computing element coefficients requires evaluations of Ω and W along element boundaries at many control points. This is the primary computational expense associated with implementation of large regional-scale models. The computational expense, per element, is proportional to the number of control points on the element and to the total number of elements. The purpose of this paper is to develop an efficient algorithm that significantly reduces the computational expense in evaluation of Ω and W, by eliminating the dependency of the per-element computational cost on the total number of elements. These improvements are particular to our globally iterative solution method; the developments presented here are not applicable to conventional globally explicit algorithms.

The algorithm developed here is an extension of the existing superblock algorithm developed by Strack et al. (1999). This paper is organized as follows: the key features of the existing formulation are reviewed first, followed by presentation of the new algorithm and an example application. For brevity, only expressions for Ω are presented. Similar expressions for the discharge function may be obtained from the definition of the discharge function, $W = -\mathrm{d}\Omega/\mathrm{d}z$.

Existing Superblock Formulation

The simulation domain is covered with a square grid. Hydrologic features (and corresponding elements) are assigned to individual grid cells based on geometry: an element is considered a member of a grid cell if the cell contains the centroid of the element. These cells, referred to as superblocks, may contain elements that extend beyond cell boundaries (partially contained members) and those that do not (fully contained members).

The complex potential due to all fully contained superblock members may be represented, at a distance R and beyond from the superblock center, as a single asymptotic expansion (Laurent series of negative terms only) plus a logarithmic term

$$\Omega_{\mathrm{B}} = \sum_{n=0}^{\infty} a_n Z^{-n} + \frac{Q_{\mathrm{B}}}{2\pi}\ln(Z) \quad \text{where} \quad Z = \frac{z - z_c}{R} \tag{1}$$

where z_c is the complex coordinate of the superblock center and Q_{B} is the combined extraction rate of all fully contained members. The circle of radius R, centered at z_c, must contain the entire superblock. The coefficients a_n are computed by matching the combined discharge potential due to all fully contained members with the real part of the Laurent series and the logarithmic term. This matching is performed at control points, $Z_m = (z_m - z_c)/R = e^{i\theta_m}$, placed along a circle of radius R that is slightly larger than the cell (e.g., 10%). The control points Z_m are selected at Fourier frequencies, $\theta_m = 2\pi(m\text{-}1)/M$, $m = 1\ldots M$, where M is the number of control points. The sum of the squared matching errors over M control points is minimized; this yields (Strack et al. 1999)

$$a_0 = \frac{1}{M}\sum_{m=1}^{M} \Phi_B(Z_m) \tag{2}$$

$$a_n = \frac{2}{M}\sum_{m=1}^{M} \Phi_B(Z_m)[\cos(n\theta_m) + i\sin(n\theta_m)] \tag{3}$$

where Φ_B is the discharge potential from all fully contained members. Using this approach, Strack et al. (1999) reported matching errors in the 12th (or higher) significant digits for Laurent series with 60 terms. The number of control points (M) should be set to about three times the number of degrees of freedom, $2N + 1$, where N is the number of terms retained in the Laurent series Equation 1 (Barnes and Janković 1999).

The coefficients a_n may alternatively be obtained by matching the combined stream function due to all fully contained members with the imaginary part of the Laurent series and the logarithmic term along the circular boundary. This approach is somewhat more complicated because of the presence of branch cuts. However, the difference between the two methods was found to be practically negligible if enough terms in Equation 1 are retained (e.g., $N = 60$). The former approach is hence adopted here because of simplicity.

As many as 250,000 (2D) and 100,000 (3D) elements with $>10^7$ element coefficients have been successfully simulated using the existing formulation (Janković et al. 2003a). In contrast, explicit formulations are limited to $\sim 10^4$ element coefficients.

The existing algorithm works well if all elements are of similar size and most of them are fully contained by the superblock, as was the case in Janković et al. (2003a, 2003b, 2003c). If the variation in size of superblock members is large, the formulation is less efficient since more partially contained elements are present.

To reduce the number of partially contained elements, the restriction that an element must be fully contained by a superblock may be somewhat relaxed. Elements that are fully contained by the circle (where matching with Laurent series is performed) can be included in the equivalent Laurent series (e.g., element A in Figure 1). Many larger elements (e.g., element B in Figure 1) still extend beyond the circle boundary. These elements are not represented by Laurent series and must be accounted for individually.

Nested Superblocks

A new superblock formulation, developed in this paper, is designed to remove the limitation posed by partially contained members and to further enhance the efficiency of the superblock approach. The idea is to group adjacent superblocks into a new, larger superblock that fits some partially contained members of the smaller blocks. In addition, this high-level superblock contains all fully contained members of smaller superblocks (e.g., the four superblocks in Figure 1). The equivalent Laurent series of the high-level superblock (referred to as a parent superblock)

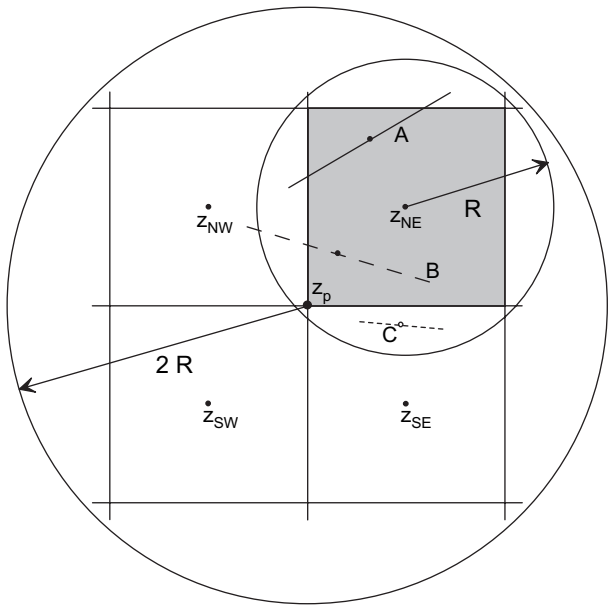

Figure 1. Schematic of a superblock configuration.

accounts for the influence of all lower-level (child) superblocks. Furthermore, elements that extend beyond the boundaries of child blocks, yet fit inside the parent block, are included in the Laurent series of the parent block.

Any number and arrangement of child blocks can be included in a parent block using the matching algorithm described in the previous section. In the following analysis, a parent block is composed of four adjacent child blocks as shown in Figure 1. The Laurent series of the parent block represents both the complex potential of all four child blocks and the complex potential of the partially contained members of child blocks that are fully contained by the parent block. This simple geometric arrangement results in significant savings of computational effort, as shown subsequently.

The following analysis relates to the child block labeled NE in Figure 1. For brevity, similar analyses for the other three child blocks are omitted. The Laurent series of the NE child block can be represented as:

$$\Omega_{\mathrm{NE}}(z) = \frac{Q_{\mathrm{NE}}}{2\pi}\ln\left(\frac{z - z_{\mathrm{NE}}}{R}\right) + \sum_{n=0}^{\infty} a_n^{\mathrm{NE}}\left(\frac{z - z_{\mathrm{NE}}}{R}\right)^{-n} \quad (4)$$

Since the NE child block is contained by the parent block, Equation 4 can also be written as an expansion about z_p, the center of the parent block

$$\Omega_{\mathrm{NE}}(z) = \frac{Q_{\mathrm{NE}}}{2\pi}\ln\left(\frac{z - z_p}{2R}\right) + \sum_{n=0}^{\infty} b_n^{\mathrm{NE}}\left(\frac{z - z_p}{2R}\right)^{-n} \quad (5)$$

Equating Equations 4 and 5 and rearranging yields

$$\sum_{n=0}^{\infty} b_n^{\mathrm{NE}}\left(\frac{z-z_p}{2R}\right)^{-n} = \frac{Q_{\mathrm{NE}}}{2\pi}\ln\left(2\frac{z-z_{\mathrm{NE}}}{z-z_p}\right) + \sum_{n=0}^{\infty} a_n^{\mathrm{NE}}\left(\frac{z-z_{\mathrm{NE}}}{R}\right)^{-n} \quad (6)$$

The matching of the left-hand and right-hand side is done on a truncated form of Equation 6 using M control points $Z_m = (z_m - z_p)/(2R) = e^{i\theta_m}$; $\theta_m = 2\pi(m-1)/M$, $m = 1 \ldots M$. Both series are truncated at the same order, N,

$$\sum_{n=0}^{N} b_n^{\mathrm{NE}} e^{-in\theta_m} = \frac{Q_{\mathrm{NE}}}{2\pi}\ln\left(2\frac{z_m - z_{\mathrm{NE}}}{z_m - z_p}\right) + \sum_{n=0}^{N} a_n^{\mathrm{NE}}\left(\frac{z - z_{\mathrm{NE}}}{R}\right)^{-n} \quad (7)$$

Equation 7 may be written in a matrix form as

$$\mathbf{Cb} = \mathbf{d}\cdot Q + \mathbf{Ea} \quad (8)$$

where the complex matrix elements for the NE block are defined as

$$b_n = b_n^{\mathrm{NE}},\ a_n = a_n^{\mathrm{NE}},\ C_{mn} = e^{-in\theta_m},$$
$$d_m = \frac{1}{2\pi}\ln\left(2\frac{z_m - z_{\mathrm{NE}}}{z_m - z_p}\right),\ \text{and}\ E_{mn} = \left(\frac{z_m - z_{\mathrm{NE}}}{R}\right)^{-n} \quad (9)$$

The matrix elements for the other child blocks are obtained by replacing z_{NE} with the center of respective blocks. Equations 8 and 9 define the influence of the NE child coefficients on the parent block coefficients. Equation 8 is solved in the least-squares sense (Golub and Van-Loan 1996) by multiplying it by the transposed conjugate of $\mathbf{C}$, $\overline{\mathbf{C}}^T$,

$$\overline{\mathbf{C}}^T\mathbf{Cb} = \overline{\mathbf{C}}^T\mathbf{d}\cdot Q + \overline{\mathbf{C}}^T\mathbf{Ea} \quad (10)$$

With the selection of control points adopted in this paper, $\overline{\mathbf{C}}^T\mathbf{C} = M\cdot\mathbf{I}$, where $\mathbf{I}$ is the identity matrix. Hence

$$\mathbf{b} = \mathbf{f}\cdot Q + \mathbf{Ga} \quad (11)$$

where

$$\mathbf{f} = \frac{1}{M}\overline{\mathbf{C}}^T\mathbf{d}\ \text{and}\ \mathbf{G} = \frac{1}{M}\overline{\mathbf{C}}^T\mathbf{E} \quad (12)$$

Equations 11 and 12 contain a portion of the coefficients b_n due to the NE block. Similar analysis must be performed for the other three child blocks, and the coefficients b_n are computed as

$$b_n^p = b_n^{\mathrm{NE}} + b_n^{\mathrm{NW}} + b_n^{\mathrm{SE}} + b_n^{\mathrm{SW}} + b_n^C \quad (13)$$

where b_n^C are the Laurent series coefficients due to partially contained members of the child blocks that are fully contained by the parent block. These coefficients are computed using the matching algorithm described in the previous section.

The complex potential due to all child blocks and their partially contained members that fit in the parent block is

$$\Omega_P(z) = \frac{Q_p}{2\pi}\ln\left(\frac{z - z_p}{2R}\right) + \sum_{n=0}^{N} b_n^p\left(\frac{z - z_p}{2R}\right)^{-n} \quad (14)$$

where

$$Q_p = Q_{\mathrm{NE}} + Q_{\mathrm{NW}} + Q_{\mathrm{SW}} + Q_{\mathrm{SE}} + Q_C \quad (15)$$

Q_C is the combined extraction rate of partially contained members of child blocks that fit in the parent block.

The process of combining child blocks into a parent block may be repeated at the level of parent blocks. Four

adjacent parent blocks are considered as a new parent block. This process can be carried out for any number of levels, resulting in a quad-tree (Samet 1989) superblock structure. At the highest level, the simulation domain is covered by four blocks. Each of these blocks is subdivided into four child blocks. The process goes down the quad-tree structure until the smallest blocks are too small to fully contain member elements. The complex potential at a given location is evaluated using an algorithm that progresses from the highest level (largest blocks) to the lowest level (smallest blocks). If a high-level block is used, the contributions from all its members are accounted for; there is no need to use lower-level blocks. The example considered in the next section contains up to five nesting levels.

The quad-tree structure, originally developed for rapid storage and querying of spatially distributed data (Samet 1989), provides additional advantages in ground water flow and transport modeling. For example, the structure may be used to rapidly access values of hydraulic conductivity in models of highly heterogeneous formations.

Regional Modeling Example

A test case was developed to examine the performance of the new formulation. The test case is based on the geometry and topography of the Lake Ontario–St. Lawrence River surface watershed and surrounding area. The case is not considered as a representative model of this watershed; the only purpose is to illustrate the efficiency of the nested superblock approach in modeling regional-scale ground water flow. The aquifer is represented as a homogeneous unconfined formation with conductivity of 3 m/d and uniform recharge of 12.5 cm/year, with a horizontal base.

The model consists of 3727 head-specified line segments that mimic significant surface water features. The geometry was digitized from a USGS 1:100,000 hydrologic basemap of the region (USGS 2003). Specified heads were interpolated from a 1:250,000 USGS digital elevation model (USGS 2003). The model covers ~200,000 km^2 with elements ranging in length from 60 to 30,000 m in length.

The algorithm was implemented in the object-oriented analytic element software Bluebird (Craig 2002). The solution was obtained using the iterative algorithm of Janković and Barnes (1999). Figure 2 shows head distribution and model configuration at three spatial scales obtained with 5 degrees of freedom per line segment after 5000 iterations.

The computational efficiency of the formulation was measured by solving the system at multiple levels of superblock nesting (from 0 to 5) and for varying number of degrees of freedom per segment (from 2 to 30). The number of superblocks ranged from 0 (level 0) to 1024 (level 5). Computations were performed on a desktop computer with a Pentium IV Xeon processor at 2.4 GHz. The results are shown in Figure 3. The nested superblock approach does not affect the iterative process (or the number of iterations) required for the model to converge (only the per-iteration cost is affected). Efficiency was hence measured based on a single iteration. The nested superblock speedup, defined as a ratio of computer time required to complete one iteration at level 0 (no blocks) to that at a given nesting level, was independent of the number of degrees of freedom. The speedup was 1.5, 3.3, 6.8, 14, and 19 for nesting levels 1, 2, 3, 4, and 5, respectively.

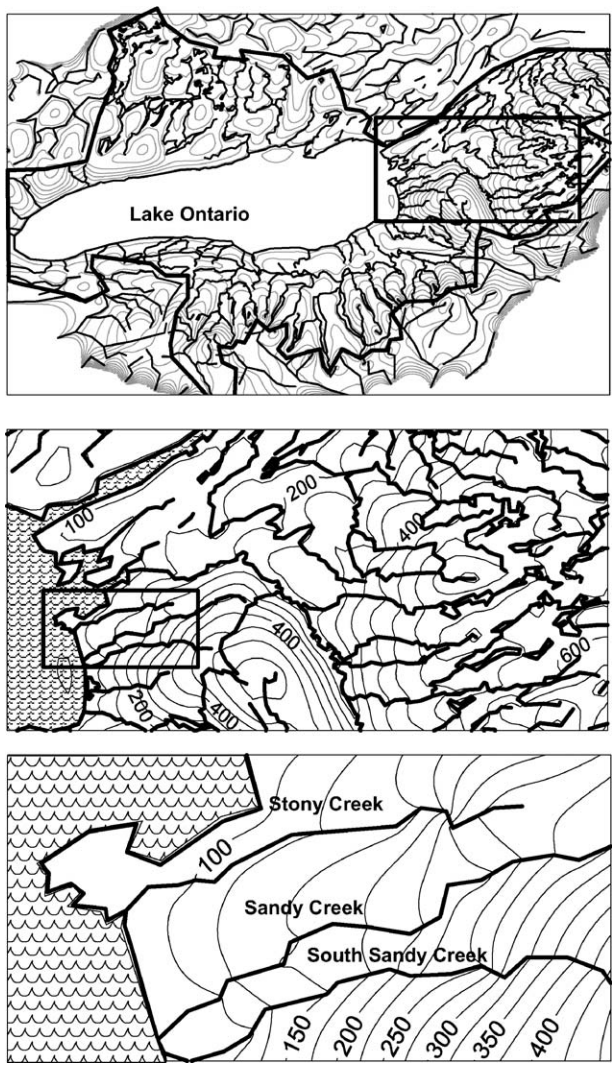

Figure 2. Head contours at three spatial scales for the test simulation.

The additional memory (random-access memory [RAM]) storage required for superblocks is small. Our highest precision test case (with 74,540 degrees of freedom) required only 16.4 MB of RAM with five nesting levels and 13.4 MB without superblocks (an increase of ~20%).

Conclusions and Discussion

A new superblock algorithm for improving the computational efficiency of analytic element models has been developed and tested. The algorithm reduces the per-iteration computational cost in models containing thousands of hydrogeologic features at multiple scales. The relative performance improves as the number of elements is increased.

The approach is currently being extended to massively parallel computer architectures (to further reduce

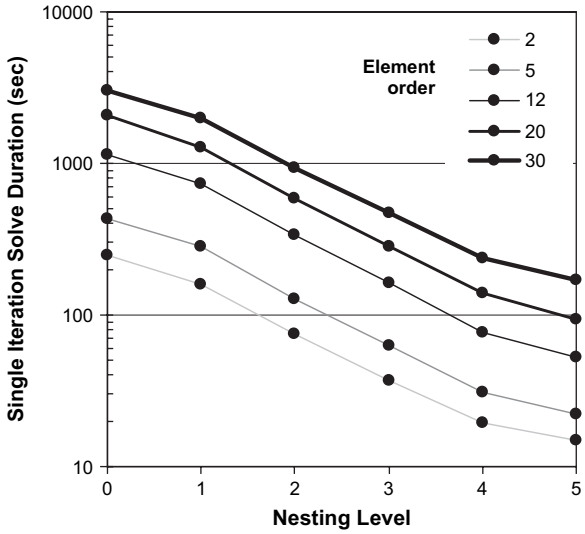

Figure 3. Nested superblock solution speedup.

per-iteration cost), permitting the solution of regional-scale models with tens of thousands of elements. The main difficulty in parallel implementation of the nested superblock algorithm is the amount of communications required to maintain our global Gauss-Seidel algorithm. In our ongoing research, we have replaced the Gauss-Seidel algorithm with a less communication-intensive Jacobi algorithm (where element coefficients computed during a given iteration are not used until the next iteration). This, however, significantly increases the total number of iterations. New algorithms are being developed that significantly reduce (by ~1 order of magnitude) the number of iterations required for both Gauss-Seidel and Jacobi methods. These algorithms are designed to be used in conjunction with the nested superblock method presented here. They will be reported on in the near future.

Acknowledgments

The authors wish to thank the National Science Foundation (project EAR-0218914; IGERT award DGE-987066) and the Environmental Protection Agency (NCER STAR grant R 82-7961) for providing partial funding for this research. The research described in the article has not been subjected to any EPA review and therefore does not necessarily reflect the views of the Agency, and no official endorsement should be inferred.

References

Barnes, R., and I. Janković. 1999. Two-dimensional flow through large numbers of circular inhomogeneities. *Journal of Hydrology* 226, no. 3–4: 204–210.

Craig, J.R. 2002. *Bluebird Developer's Manual*. http://www.groundwater.buffalo.edu/ (accessed July 2003).

Fitts, C.R. 2002. *Groundwater Science*. San Diego, California: Academic Press.

Golub, G.H., and C.F. Van-Loan. 1996. *Matrix Computations*. Baltimore, Maryland: The Johns Hopkins University Press.

Haitjema, H.M. 1995. *Analytic Element Modeling of Groundwater Flow.* San Diego, California: Academic Press.

Janković, I., and R. Barnes. 1999. High-order line elements in modeling two-dimensional groundwater flow. *Journal of Hydrology* 226, no. 3–4: 211–223.

Janković, I., A. Fiori, and G. Dagan. 2003a. Effective conductivity of an isotropic heterogeneous medium of lognormal conductivity distribution. *SIAM Multiscale Modeling, Analysis, and Simulation* 1, no. 1: 40–56.

Janković, I., A. Fiori, and G. Dagan. 2003b. Flow and transport through two-dimensional isotropic media of binary conductivity distribution, part 2: Numerical simulations and comparison with theoretical results. *Stochastic Environmental Research and Risk Assessment (SERRA)* 17, no. 6: 384–393.

Janković, I., A. Fiori, and G. Dagan. 2003c. Flow and transport in highly heterogeneous formations, part 3: Numerical simulations and comparison with theoretical results. *Water Resources Research* 39, 1270.

Samet, H. 1989. *The Design and Analysis of Spatial Data Structures*. Reading, Massachusetts: Addison-Wesley.

Strack, O.D.L. 1989. *Groundwater Mechanics*. New York: Prentice-Hall.

Strack, O.D.L., I. Janković, and R. Barnes. 1999. The superblock approach for the analytic element method. *Journal of Hydrology* 226, no. 3–4: 179–187.

USGS. 2003. *Earth Explorer.* http://earthexplorer.usgs.gov (accessed September 2002).