EFFICIENT SOLUTION OF HETEROGENEOUS ANISOTROPIC DIFFUSION PROBLEMS

Costanza Aricò and Tullio Tucciarelli
Università di Palermo
Dipartimento di Ingegneria Civile, Ambientale e Aerospaziale
Viale delle Scienze, 90128, Palermo, Italy
e-mail: arico@idra.unipa.it

Key words: anisotropic diffusion, heterogeneous medium, $M$-matrix, Delaunay mesh, affine transformation, edge swap

1 INTRODUCTION

Diffusion equation with anisotropic coefficients arise in many environmental topics, heat transfer, groundwater flow and transport problems, petroleum reservoir simulations, hydrodynamic simulations, ... These problems are characterized by a full rank diffusion tensor, that is diagonal only if the reference system is aligned with the principal direction of anisotropy (Bear). Numerical difficulties in solving this type of equation could arise when anisotropy, the tendency of a phenomenon to progressively align along a preferential direction, becomes strong.

A new methodology for the solution of the anisotropic heterogeneous diffusion problem is presented in this paper. The governing equations are discretized over a basic unstructured triangular mesh satisfying the so called Generalized Delaunay condition, further specified. The resulting spatial discretization of the fluxes across the control volume is similar to the one occurring in the standard linear ($P_1$) Galerkin Finite Element scheme.

2 GOVERNING EQUATIONS AND FLUX SPATIAL DISCRETIZATION

$\Omega$ is a 2D domain, $\Gamma$ its boundary and $x = [x_1, x_2]^T$ the spatial co-ordinate vector. $H_1(\Omega)$ is the Sobolev space of square-integrable functions with square-integrable 1st order derivatives over $\Omega$. Assume the following diffusive problem in the unknown variable $u(x, t) \in H_1(\Omega)$:

$$\begin{cases}
\nabla \cdot q = f & \text{in } \Omega \times [0, T] \\
u = u_D(x), x \in \Gamma_D \\
q(x) \cdot n = g_N(x), x \in \Gamma_N
\end{cases}$$

(1),

where $q$ is the unitary diffusive flux vector, $D(x)$ the (2x2) diffusion matrix, symmetric and positive definite, $\Gamma_D$ and $\Gamma_N$ the portions of $\Gamma$ where Dirichlet and Neumann conditions respectively hold, $u_D$ a fixed Dirichlet value on $\Gamma_D$, $g_N$ the assigned Neumann flux ($n$ the unit outward normal to the boundary) and $f = f(x, t) \in L_2(\Omega)$ a given function (source term).

$\Omega_h$ is a polygonal approximation of $\Omega$ and $T_h$ (basic mesh) an unstructured triangulation of $\Omega_h$. $N_T$ is the number of triangle $T$ of $T_h$, $P_i, i = 1, \ldots, N$ is the set of all vertices (nodes) of all $T \in T_h$ and $N$ the node number. A dual mesh is constructed over $T_h$ and the dual finite control volume
(or cell) associated with node \( i \) is the closed polygon given by the union of the midpoint of each side with the “anisotropic” circumcentre of each triangle \( T \) sharing \( i \), further defined.

### 2.1 The anisotropic circumcentre

Assume an homogeneous medium with \( \mathbf{D} \) its full diffusion tensor. Nodes of triangle \( T \) are \( i, ip \) and \( im \), in counterclockwise direction. Edge vector \( \mathbf{r}_{i,ip} (\mathbf{r}_{i,im}) \) connects \( i \) and \( ip (im) \), oriented from \( i \) to \( ip (im) \). Call \( c_T \) the anisotropic circumcentre of triangle \( T \) (\( x_T^i \) its co-ordinate vector).

Pi,ip and Pi,im are the midpoints of \( r_{i,ip} \) and \( r_{i,im} \) (\( x_{i,ip} \) and \( x_{i,im} \) the corresponding co-ordinate vectors). \( c_T \) is computed in order to set to zero the flux across segments \( P_{i,ip}c_T \) and \( P_{i,im}c_T \) due to the component of the \( u \) gradient orthogonal to \( r_{i,ip} \) and \( r_{i,im} \). This is equivalent to set fluxes in triangle \( T \) equal to:

\[
F_{n_{i,ip}} = -\mathbf{D} n_{i,ip} \cdot (x_T^i - x_{i,ip}) = 0 \quad \text{and} \quad F_{n_{i,im}} = -\mathbf{D} n_{i,im} \cdot (x_T^i - x_{i,im}) = 0
\]

(2),

where \( n_{i,ip} \) and \( n_{i,im} \) are the inward unit vectors orthogonal to the edges \( r_{i,ip} \) and \( r_{i,im} \) respectively.

Since \( \mathbf{D} \) is a full tensor, \( r_{i,ip} \) and \( r_{i,im} \) are generally not orthogonal to vectors – \( \mathbf{D} n_{i,ip} \) and – \( \mathbf{D} n_{i,im} \).

### 2.2. Computation of the diffusion coefficients

According to the computation of \( c_T \), fluxes across \( P_{i,ip}c_T \) and \( P_{i,im}c_T \) are entirely due to the component of the \( u \) gradient along \( r_{i,ip} \) and \( r_{i,im} \). A computational structure of the flux across the segment from midpoint edge to triangle circumcentre similar to the one of the conforming linear \((P1)\) Galerkin Finite Element (FE) scheme (Putti and Cordes\(^2\)) can be derived.

We compute the flux across \( P_{i,ip}c_T \) and \( P_{i,im}c_T \) due to a unitary difference between \( u \) values in \( i \) and \( ip \) and \( i \) and \( im \) (specific fluxes) as (see Aricò and Tucciarelli\(^3\)):

\[
F_{d_{i,ip}} = \mathbf{D}(x_{i,ip} - x_i) \cdot \frac{1}{|r_{i,ip}|} \quad \text{and} \quad F_{d_{i,im}} = \mathbf{D}(x_{i,im} - x_i) \cdot \frac{1}{|r_{i,im}|}
\]

(3),

where the Euclidean norm of \( r_{i,ip} (r_{i,im}) \) is also the distance between nodes \( i \) and \( ip (im) \). Assume now \( T_1 \) and \( T_2 \) sharing \( r_{i,im} \), with anisotropic circumcentres \( c_{T_1} \) and \( c_{T_2} \). In the stiffness matrix of the resolving system, the extra-diagonal coefficient \( F_{di,m} \) corresponding to the connected nodes \( i \) and \( im \) can be viewed as the sum of the unitary fluxes across \( P_{i,ip}c_T \) and \( P_{i,im}c_T \), equal to:

\[
F_{d_{i,im}} = -F_{d_{i,ip}} - F_{d_{i,im}}
\]

(4),

with coefficients \( F_{d_{i,ip}} \) and \( F_{d_{i,im}} \) defined in Eq. (3). Eq. (4) can be written as:

\[
F_{d_{i,im}} = -\mathbf{D}^1(x_{i,im} - x_i) \cdot \frac{1}{|r_{i,im}|} - \mathbf{D}^2(x_{i,ip} - x_i) \cdot \frac{1}{|r_{i,ip}|}
\]

(5),

with \( \mathbf{D}^{(1,2)} \) diffusive tensor in \( T^{(1,2)} \). Eq. (5) can be written as:
Costanza Aricò and Tullio Tucciarelli.

\[ F_{d_{im}} = -\frac{(d_{pm}^2 c_{im}^r \sin \theta_i + d_{pm}^2 c_{im}^r \sin \theta_j)}{|r_{im}|^2}, \quad d_{pm}^r = |D^q(x_p - x_{pm})| \quad \text{with} \quad q = \frac{1}{2}, \quad p = \frac{i}{im}, \quad pm = \frac{jm}{i} \quad (6), \]

\[ c_{pm}^r \] is the distance between \( c_{T_q} \) and midpoint \( P_{i,im} \) in triangle \( T_q \) and \( \theta_q \) is the angle between vectors \( D^q(x_p - x_{pm}) \) and \( (x_p - x_{i,im}) \) in \( T_q \). \( F_{d_{im}} \) has to be always negative to guarantee the \( M \)-property of the system matrix.

3 discretization of the diffusive system (1)

3.1 Isotropic case and Generalized Delaunay (GD) mesh condition

From Eq. (2), the dual finite volume \( e_i \) associated with node \( P_i \) in isotropic media is the polygon given by the union of the midpoint of each side with the circumcentre of each triangle \( T \), equal to the control volume of the mass balance of the standard Galerkin FE stiffness equation for node \( P_i \) (Putti and Cordes). A Delaunay triangulation in \( \mathbb{R}^2 \) is defined by the condition that all the nodes in the mesh are not interior to the circles defined by the three nodes of each triangle (Joe). A Delaunay triangulation satisfies the following condition (Putti and Cordes):

\[ c_{im}^r + c_{im}^c \geq 0 \quad (7), \]

for each interior edge \( r_{i,im} \). \( D^q = D^q I (q = 1, 2, I \) is the identity matrix and \( D^q \) is a positive scalar value) are the diffusive tensors of triangles \( T_1 \) and \( T_2 \) sharing \( r_{i,im} \). Given a Delaunay mesh, starting from Eq. (6) we propose to compute the extra-diagonal coefficient \( F_{d_{im}} \) as:

\[ F_{d_{im}} = -\frac{(d_{pm}^2 c_{im}^r + d_{pm}^2 c_{im}^r)}{|r_{im}|^2}, \quad \left[ \begin{array}{c} c_1 = c_{i,im}^r, \quad c_2 = c_{i,im}^r + c_{i,im}^r \quad \text{if} \quad c_{i,im}^r \geq 0 \quad \text{and} \quad c_{i,im}^r \geq 0 \\
\quad c_1 = c_{im}^r, \quad c_2 = c_{im}^r \quad \text{if} \quad c_{i,im}^r < 0 \quad \text{and} \quad c_{i,im}^r < 0 \\
\quad c_1 = 0, \quad c_2 = c_{i,im}^r + c_{i,im}^r \quad \text{if} \quad c_{i,im}^r < 0 \quad \text{and} \quad c_{i,im}^r \geq 0 \quad \text{and} \quad c_{i,im}^r > c_{i,im}^r \end{array} \right] \quad (8). \]

According to Eq. (8), \( c_q \) is never smaller than zero. In the standard Galerkin FE discretization, if the two element fluxes \( F_{d_{i,im}} \) and \( F_{d_{j,im}} \) are computed with different elemental parameters \( D^q \), the sign of the total flux from node \( i \) to node \( im \) can lose consistency with the \( u \) difference, even if the mesh satisfies the Delaunay property. On the opposite, the formulation provided by Eq. (8) guarantees the negative sign of \( F_{d_{i,im}} \) defined by Eq. (8). If Delaunay property is not satisfied, it is possible to obtain a new mesh that satisfies Eq. (7) for all the internal edges starting from the original one, without changing the location of the original nodes. This can be done by a series of local edge swaps (Forsyth).

If element \( T_1 \) is a boundary element with boundary side \( r_{i,im} \) opposite to an obtuse angle, specific flux is negative, even in a Delaunay mesh, since the distance of the circumcentre from the boundary edge is negative and \( F_{d_{i,im}} \) is positive. We define Generalized Delaunay (GD) mesh a Delaunay mesh where the following condition holds for all the boundary edges:

\[ c_{pm}^r \]
If Eq. (9) does not hold for one or more boundary edges, and/or common edges are fixed as internal boundaries, it is still possible to obtain a GD mesh, also saving internal boundaries, by adding new nodes along the original boundary edges (Aricò and Tucciarelli3).

3.2 Directionally homogeneous anisotropic case and Generalized Anisotropic Delaunay mesh

Diffusion tensor \( D \) can be split into a scalar \( d_0 \) and a directional \( \mathbf{D}' \) component, according to:

\[
D = d_0 \mathbf{D}' = d_0 \begin{pmatrix} D'_{11} & D'_{12} \\ D'_{21} & D'_{22} \end{pmatrix}
\]

with \( d_0 = D_{11} + D_{22} \) and \( D'_{rs} = \frac{D_{rs}}{d_0}, \ r, s = 1, 2 \). \( \text{Eq. (10)} \)

Tensor \( \mathbf{D}' \), similarly to \( D \), is symmetric and positive definite. \( D^1 \) and \( D^2 \) are the diffusive tensors in triangles \( T_1 \) and \( T_2 \) sharing side \( r_{im} \) with the same \( \mathbf{D}' \) component. The extra-diagonal coefficient \( F_{d_{im}} \) can be computed by Eq. (6), which becomes:

\[
F_{d_{im}} = -\left( \frac{d'_{im}c_i + d'_{jm}c_j}{\rho_{r_{im}}} \right) \sin \theta \geq 0 \quad \text{(9)}
\]

and \( \theta = \theta_1 = \theta_2 \) (medium directionally homogeneous). Even if \( d'_{im} \) and \( d'_{jm} \) have different weight, coefficient \( F_{d_{im}} \) computed by Eq. (11) is always negative and the \( M \)-property is preserved if \( \theta \leq \pi \). We say a mesh to satisfy the Generalized Anisotropic Delaunay (GAD) property if \( F_{d_{im}} \) computed for each side of the mesh by Eq. (11) is always negative. The orthogonality in \( T_1 \) between \( r_{ip} \) (with \( q = im \) or \( ip \)) and \( n_{iq} = \mathbf{D}'^1 \mathbf{s}, \ \mathbf{s} \) is the unit vector going from the edge midpoint to \( c_{r_{ip}} \), implies:

\[
\left( x_q^i - 1/2(x_q + x_i) \right) \left( x_q - x_i \right) = 0
\]

Changing \( x_q \) with the co-ordinate vector \( \mathbf{x} \) of a generic point, Eq. (12) is:

\[
\left( x_q^i - 1/2(x_q + x_i) \right) \left( x_q - x_i \right) = 0
\]

which is the equation of an ellipse, with centre \( c_{T_1} \), passing through the three \( T_1 \) nodes (Aricò and Tucciarelli3). Eq. (13) can be written as (Aricò and Tucciarelli3):

\[
\left( x_q^i - 1/2(x_q + x_i) \right)^T \mathbf{C} C^T \left( x_q - x_i \right) = 0 \quad \text{with} \quad \mathbf{C} = \mathbf{H} \Lambda^{1/2}
\]

where columns of matrix \( \mathbf{H} \) are the \( \mathbf{D}'^1 \)-eigenvectors, \( \Lambda \) is the diagonal matrix with diagonal elements equal to the \( \mathbf{D}'^1 \)-eigenvalues. Matrix \( \mathbf{C} \) acts on \( \mathbf{x} \) with a rotation and a distortion: \( \mathbf{H} \) accounts for rotation, \( \Lambda^{1/2} \) for distortion. We define a new co-ordinate vector:

\[
\xi = \mathbf{F} \mathbf{x} \quad \text{with} \quad \mathbf{F} = \mathbf{C}^T
\]

\( \text{Eq. (15)} \)
x and ξ represent the co-ordinate vectors respectively in the physical and in the computational space. According to Eqs. (15), Eq. (14) in the computational space becomes a circle equation (Aricò and Tucciarelli3):

$$\left(\xi_n - 1/2 (\xi_n + \xi_s)\right)^T (\xi - \xi_s) = 0$$

(16).

The original physical anisotropic problem becomes in the computational space an isotropic one. Matrix $F = A^{1/2} H^T$ provides an affine transformation between the physical and the computational space. $H^T$ can be viewed as an intermediate transformation, from $x$ to $\xi' = H^T x$. The original figure is simply rotated in the new space $\xi'$ and unitary flux in the original $x$ space and in the new $\xi'$ space is the same. $A^{1/2}$ provides a second transformation from $\xi'$ to $\xi$, that is a simple contraction along the $\xi'$ principal axis. Since eigenvalues are both positives, the sign of two coefficients in the physical and the computational space is the same (Aricò and Tucciarelli3).

3.3 Flux computation in the heterogeneous anisotropic case

In the most general heterogeneous anisotropic case, it is not possible to guarantee the M-property, unless some “smoothing” is applied to the $D$ dispersion tensor. To do that, we treat differently the scalar $d_0$ and the directional $D'$ components of tensor $D$. The scalar component is assigned to each node and in the flux computation its original average value in the triangle is always saved and assumed as constant inside the element. The directional component $D'$ is assigned to each triangle of the initial mesh, that satisfies the GD condition. When an internal edge is swapped, the average value of the $D'$ directional components in the two old triangles is assumed to hold for both the new triangles (further details in Aricò and Tucciarelli3).

4 NUMERICAL TESTS

4.1 Test 1. Analysis of the system condition number

A unitary square domain $\Omega = [0,1]^2$ and a diagonal homogeneous diffusion tensor $D$ are assumed. Coefficient $D_{11}$ is kept constant, equal to 1, $D_{22}$ ranges from 1 to 1.d-10 (see Mazzia et al.6). The basic mesh has 14 acute triangles and 12 nodes. Dirichlet conditions are imposed on the left and right boundary sides. Tables 1,a-1,b show the maximum and minimum stiffness matrix eigenvalues ($\lambda_{\text{max}}$ and $\lambda_{\text{min}}$) and their ratio, for different $D_{22}$ values, for the proposed and the $P1$ standard Galerkin schemes. Condition number of the stiffness matrix is the ratio $\lambda_{\text{max}}/\lambda_{\text{min}}$. In the proposed algorithm this ratio is closer to 1 than in the $P1$ Galerkin scheme, implying a better conditioning of the final system. Because of the anisotropy, some of the extra-diagonal coefficients of the stiffness matrix computed by the $P1$ Galerkin scheme are positive and $M$-property is lost. Results of the $P1$ Galerkin scheme are given by Mazzia et al.6

4.2. Test 2. Investigation of computation of unphysical oscillations in the solution

The following problem is solved over the square domain $\Omega = [0,1]^2$ (Edwards and Zengh7):
Costanza Aricò and Tullio Tucciarelli.

\[
\begin{align*}
\left\{ 
    u_D(0,0) &= 0, \\
    u_D(1,1) &= 200, \\
    u_D(0,x_2)|_{x_2\in[0,1]} &= 100
\right. 
\end{align*}
\]  

(17,a),

with discontinuous diffusion tensor: \( D_1 \) in sub-regions \((0 \leq x_1 \leq 1) \times (0 \leq x_2 \leq 1/3) \) and \( D_2 \) in sub-region \((0 \leq x_1 \leq 1) \times (1/3 < x_2 \leq 2/3) \):

\[
D_1 = \begin{pmatrix} 2464.36002 & 1148.683643 \\ 1148.683643 & 536.6399794 \end{pmatrix}, \quad D_2 = \begin{pmatrix} -2464.36002 & -1148.683643 \\ -1148.683643 & 536.6399794 \end{pmatrix}
\]  

(17,b).

Edwards and Zengh\(^7\) apply two Multi-Point Flux Approximation schemes: a linear Triangle Pressure Support and a Full-Pressure Support scheme. The first one produces unphysical oscillations in the solutions. The second one computes solutions almost free of oscillations. In the proposed approach the domain is discretized with a GD unstructured mesh (2624 triangles, 1377 nodes). Due to the strong contrast of tensor \( D \) in the three zones, two situations have been considered: the first one without internal boundaries, the second one, with assigned internal boundaries overlapping the jumps of \( D_1 \) and \( D_2 \). Results in the two cases are very similar and for brevity we show only the ones with internal boundaries. Figures 1,a-1,c show the computed contours of the iso-\( u \), the 3D \( u \) profile and the final mesh after the edge swaps. Computed solutions are in both cases free of spurious oscillations.

4.3. Test 3. Mesh locking investigation

The following problem is studied over the square domain \( \Omega = [0,1]^2 \) (Manzini and Putti\(^8\); Gao and Wu\(^9\)):

\[
\begin{align*}
u &= u_D, \quad x \in \Gamma_D \\
\mathbf{q} \cdot \mathbf{n} &= g_N, \quad x \in \Gamma_N
\end{align*}
\]  

(18),

with parameter \( \varepsilon \) a positive real number in the range \([10^{-6}, 1]\). The exact solution is given:

\[
u_e = \exp\left(-2\pi \sqrt{\varepsilon x_1}\right) \sin(2\pi x_2)
\]  

(19).

and source term \( f \) on the r.h.s. of eq. (19) is computed by space differentiating the same solution on the l.h.s. of Eq. (19). Three types of boundary condition are considered: Case A: fully Dirichlet conditions; Case B: mixed Dirichlet/Neumann conditions, with \( u = u_D \) on \( \Gamma_D = \{(x_1, x_2), \ x_1 = 0 \text{ or } x_2 = 0\} \) and \( \mathbf{q} \cdot \mathbf{n} = g_N \) on \( \Gamma_N = \{(x_1, x_2), \ x_1 = 1 \text{ or } x_2 = 1\} \). Case C: nearly pure Neumann conditions, with \( u = u_D \) on \( \Gamma_D = \{(x_1, x_2), \ 1-h \leq x_1 \leq 1 \text{ and } x_2 = 1 \text{ or } x_1 = 1 \text{ and } 1-h \leq x_2 \leq 1\} \) \((h \text{ is the characteristic linear mesh size}) \) and \( \mathbf{q} \cdot \mathbf{n} = g_N \) on \( \Gamma_N = \Gamma_D \). In Case C, a Dirichlet condition is imposed on the two boundary edges located at the upper right corner and the length of these edge goes to zero as \( h \). Locking may arise when solving parametric elliptic problems with FE methods and usually prohibits the convergence of low-order FE schemes when the parameter associated with the problem approaches an asymptotic value such as zero.

Domain is discretized with an unstructured GD mesh (272 acute elements, 159 nodes) and five refinements have been performed. Figure 2 shows the \( L_2 \) norms of relative errors versus \( h \).
for Cases A to C. No mesh locking is present and a second convergence order is obtained. In Case A, since the basic triangulation is acute, the most regular situation corresponds to $\varepsilon = 1$, where circumcentres of two triangles sharing an edge are on the opposite sides respect to that edge. Reducing $\varepsilon$, triangular elements tend to lengthen and to become parallel to the principal anisotropy direction and in some cases circumcentres of two triangles sharing the same edge fall on the same side respect to that edge and swapping could be necessary. This could introduce a small error in the boundary edges, where Dirichlet conditions are imposed. For Case B, errors seem to be quite independent on anisotropy ratio. This could be due to the error introduced by substituting a Dirichlet condition with a Neumann condition, greater than the error due to the progressive deformation of triangle mesh for decreasing $\varepsilon$. In Case C, computed solution is free of locking effects. Solutions computed by Manzini and Putti$^8$ and Gao and Wu$^9$ are affected by mesh locking especially for the smallest $\varepsilon$.

5 CONCLUSIONS

A new methodology for the solution of heterogeneous anisotropic diffusion problem is presented. Governing PDE is discretized over an unstructured triangular mesh satisfying the GD condition and a dual mesh is constructed over the basic mesh. The dual control volume is obtained connecting the midpoint of each triangle side with the anisotropic circumcentre of each triangle. The main advantages of the proposed procedure are: 1) the algorithm acts directly on the physical mesh, without dealing with the computational space; 2) the number and the location of nodes are not changed; 3) the stiffness matrix has the M property; 4) it can be easily applied to convection–diffusion transport problem, where diffusion matrix coefficients change in time. The proposed methodology does not compute unphysical oscillations, shows a second convergence order and no locking effects have been observed, also for strong anisotropy ratios.

REFERENCES


<table>
<thead>
<tr>
<th>proposed scheme</th>
<th>P1 Galerkin scheme</th>
</tr>
</thead>
<tbody>
<tr>
<td>$D_{22}$</td>
<td>$\lambda_{\text{min}}$</td>
</tr>
<tr>
<td>1.d+00</td>
<td>0.88030</td>
</tr>
<tr>
<td>1.d-01</td>
<td>0.72124</td>
</tr>
<tr>
<td>1.d-04</td>
<td>0.57819</td>
</tr>
<tr>
<td>1.d-06</td>
<td>0.57799</td>
</tr>
<tr>
<td>1.d-10</td>
<td>0.57799</td>
</tr>
</tbody>
</table>

Table 1. $\lambda_{\text{max}}$, $\lambda_{\text{min}}$ and spectral condition number of the proposed and P1 Galerkin scheme.

Figure 1: test 2. a) iso-$u$; b) 3D $u$ profiles; c) final computed mesh

Figure 2. $L_2$ norms of relative errors versus characteristic length size $h$; cases A) to C)