Chapter 6
Satisfaction of DMPs for a Finite Element Discretization of the Poisson Problem

In this chapter, necessary and sufficient conditions for the satisfaction of DMPs for a finite element discretization of the Poisson problem with P_1 finite elements will be discussed. The proofs of the DMP consists of checking the hypotheses of Theorem 5.3 and 5.5.

Consider the Poisson problem

$$\begin{aligned} -\Delta u &= f \qquad \text{in } \Omega, \\ u &= u_{\mathrm{D}} \quad \text{on } \partial\Omega. \end{aligned} \tag{6.1}$$

A weak formulation is derived as presented in Section 1.3 and Theorem 1.11 shows the existence and uniqueness of a weak solution. Then, a finite element discretization on a simplicial grid $\mathcal{T}_h$ with P_1 finite elements is applied, leading to an algebraic system of form (5.1)-(5.2). The entries of the system matrix are given by

$$a_{ij} = (\nabla\phi_j, \nabla\phi_i)\,, \quad i = 1, \ldots, m,\ j = 1, \ldots .n, \tag{6.2}$$

where $\{\phi_i\}_{i=1}^n$ are the standard basis functions (hat functions).

The analysis requires a formula for the entries (6.2) of A. To this end, a formula relating the gradient of the basic functions and the normal outward vector to the mesh cell K is utilized. Since the basis function $\phi_i|_K$ vanishes on the facet $F_i^K \subset \partial K$, its derivative in any direction tangent to F_i^K vanishes. Hence, $\nabla\phi_i|_K$ is proportional to the unit normal $\boldsymbol{n}_i^K$. Consider the height vector $\boldsymbol{h}_i$ from F_i^K to $\boldsymbol{x}_i$, compare Figure 6.1. This vector is parallel to $\boldsymbol{n}_i^K$, pointing in the opposite direction, and the derivative of $\phi_i|_K$ in the direction of $\boldsymbol{h}_i$ is the constant $1/\left\|\boldsymbol{h}_i\right\|_2$. Altogether, using the formula for the volume of the simplex K leads to

$$\nabla\phi_i|_K = -\frac{1}{\left\|\boldsymbol{h}_i\right\|_2}\boldsymbol{n}_i^K = -\frac{|F_i^K|}{d|K|}\boldsymbol{n}_i^K\,. \tag{6.3}$$

So, the local matrix entry is given by

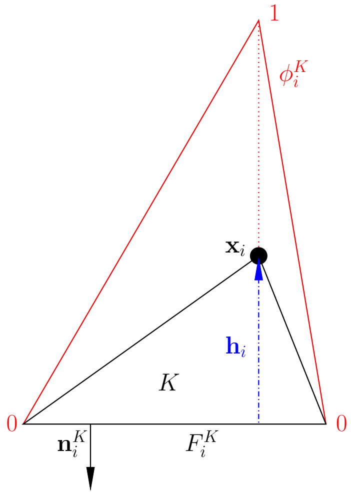

Fig. 6.1 Illustration for the derivation of (6.3).

$$a_{ij}^K = (\nabla\phi_j, \nabla\phi_i)_K = |K|\frac{|F_j^K|\,|F_i^K|}{d^2|K|^2}\boldsymbol{n}_j^K \cdot \boldsymbol{n}_i^K = -\frac{|F_j^K|\,|F_i^K|}{d^2|K|}\cos\theta_E^K. \quad (6.4)$$

Here, θ_E^K is the angle formed by F_i^K and F_j^K, or, more precisely, θ_E^K is the dihedral angle given by

$$\cos\theta_E^K = -\boldsymbol{n}_i^K \cdot \boldsymbol{n}_j^K.$$

A careful inspection of the statements of the results from Section 5.1.1 reveals that one only needs to show properties for the first m rows of the coefficient matrix of the system (5.1)-(5.2), that is, one only needs to worry about the equations associated to nodes interior to Ω. This observation motivates to define, for a matrix $A \in \mathbb{R}^{n\times n}$, the matrix $(A)^m \in \mathbb{R}^{m\times n}$ as the matrix containing only the first m rows of A. In fact, showing that $(A)^m$ is of non-negative type is what is needed to use Theorems 5.3 and 5.5 due to the expression (5.3) for the matrix associated to the system (5.1)-(5.2).

The statement given next was proved in (Xu & Zikatanov, 1999, Lemma 2.1). It presents a necessary and sufficient condition on the mesh to guarantee the satisfaction of the DMP.

Theorem 6.1 (Sufficient and necessary condition for $(A)^m$ to be of non-negative type, Xu & Zikatanov (1999)). *The matrix $(A)^m$ is of non-negative type if and only if the mesh $\mathcal{T}_h$ satisfies so-called the XZ-criterion: for every edge that is not contained on the boundary of Ω, it holds*

$$\sum_{K\subset\omega_E} \left|\kappa_E^K\right| \cot\theta_E^K \geq 0, \quad (6.5)$$

where for each edge E, it is

$$\omega_E = \cup\{K \in \mathcal{T}_h \ : \ E \subset K\},$$

$E = F_i^K \cap F_j^K$*, where* F_i^K *and* F_j^K *are facets of* K *with* $i \neq j$*, and*

$$\left|\kappa_E^K\right| = \begin{cases} 1 & \text{if } d = 2, \\ \left|F_i^K \cap F_j^K\right| & \text{if } d = 3, \ \text{i.e., the length of the edge.} \end{cases} \tag{6.6}$$

In addition, $(A)^m$ *satisfies* (5.13).

Proof. Let $\boldsymbol{x}_i, \boldsymbol{x}_j$ be two different nodes contained in the same mesh cell $K \in \mathcal{T}_h$, let E be the connecting edge, and assume that $i \in \{1, \ldots, m\}$. Recall the following formulas for the volume of a simplex

$$|K| = \frac{|F_i^K||F_j^K|}{2} \sin\theta_E^K \quad \text{if } d = 2, \qquad |K| = \frac{2|F_i^K||F_j^K|}{3\left|\kappa_E^K\right|} \sin\theta_E^K \quad \text{if } d = 3\,.$$

Inserting them in (6.4), and using (6.6) gives

$$a_{ij}^K = -\frac{1}{d(d-1)} \left|\kappa_E^K\right| \cot\theta_E^K\,. \tag{6.7}$$

Thus, it is

$$a_{ij} = \sum_{K \subset \omega_E} a_{ij}^K = -\sum_{K \subset \omega_E} \frac{|\kappa_E^K| \cot\theta_E^K}{d(d-1)}\,, \tag{6.8}$$

and then (5.5) is satisfied if and only if (6.5) holds. Finally, since the basis functions form a partition of unity, one has

$$\sum_{j=1}^N a_{ij} = \sum_{j=1}^N (\nabla\phi_j, \nabla\phi_i) = (\nabla 1, \nabla\phi_i) = 0\,. \tag{6.9}$$

So, (5.13) is satisfied, and in particular (5.6). ■

In Drăgănescu *et al.* (2005); Brandts *et al.* (2009), counterexamples are constructed showing that for meshes that violate the XZ-criterion the local DMP may not hold, thus confirming the optimality of the XZ-criterion. On the other hand, the condition on the matrix A being of non-negative type is not necessary for the validity, see Example 5.14.

Remark 6.2 (XZ-criterion). One can show that in two dimensions, the XZ-criterion (6.5) is equivalent to a Delaunay triangulation, e.g., see Xu & Zikatanov (1999). In two and three dimensions, so-called non-obtuse triangulations, i.e., triangulations where all dihedral angles $\theta_E^K \leq \pi/2$, satisfy the XZ-criterion. □

Remark 6.3 (Convection-diffusion-reaction equations). The construction of finite element discretizations for the convection-diffusion-reaction equation (2.3) that are on the one hand sufficiently accurate and on the other hand satisfy DMPs is quite complicated. There are only few such methods. All off them are nonlinear, i.e., there is a stabilization term where the parameter depends on the concrete numerical solution. In particular, such methods applying different strategies in a vicinity of layers and away from layers. The numerical analysis of those methods is an active topic of research. For some ideas concerning the construction of such methods and the corresponding analysis, it is referred to Barrenechea *et al.* (2018). □