Helsinki University of Technology, Institute of Mathematics, Research Reports

Teknillisen korkeakoulun matematiikan laitoksen tutkimusraporttisarja

Espoo 2007 A525

The discrete maximum principle for linear simplicial finite element approximations of a reaction-diffusion problem

Jan Brandts Sergey Korotov Michal Křížek

TEKNILLINEN KORKEAKOULU
TEKNISKA HÖGSKOLAN
HELSINKI UNIVERSITY OF TECHNOLOGY
TECHNISCHE UNIVERSITÄT HELSINKI
UNIVERSITE DE TECHNOLOGIE D'HELSINKI

Helsinki University of Technology, Institute of Mathematics, Research Reports
Teknillisen korkeakoulun matematiikan laitoksen tutkimusraporttisarja
Espoo 2007 A525

The discrete maximum principle for linear simplicial finite element approximations of a reaction-diffusion problem

Jan Brandts Sergey Korotov Michal Křížek

Helsinki University of Technology
Department of Engineering Physics and Mathematics
Institute of Mathematics

Jan Brandts, Sergey Korotov, Michal Křížek: *The discrete maximum principle for linear simplicial finite element approximations of a reaction-diffusion problem* ; Helsinki University of Technology, Institute of Mathematics, Research Reports A525 (2007).

Abstract: *This paper provides a sufficient condition for the discrete maximum principle for a fully discrete linear simplicial finite element discretization of a reaction-diffusion problem to hold. It explicitly bounds the dihedral angles and heights of simplices in the finite element partition in terms of the magnitude of the reaction coefficient and the spatial dimension. As a result, it can be computed how small the acute simplices should be for the discrete maximum principle to be valid. Numerical experiments suggests that the bound, which considerably improves a similar bound in* [6]*, is in fact sharp.*

AMS subject classifications: 65N30, 65N50

Keywords: reaction-diffusion problem, maximum principle, finite element method, discrete maximum principle, simplicial partition, angle condition

Correspondence

brandts@science.uva.nl, sergey.korotov@hut.fi, krizek@math.cas.cz

ISBN - 978-951-22-8781-9
ISSN 0784-3143

Helsinki University of Technology
Department of Engineering Physics and Mathematics
Institute of Mathematics
P.O. Box 1100, FI-02015 TKK, Finland
email:math@tkk.fi http://math.tkk.fi/

1 Introduction

Given $d \geq 1$, let $\Omega \subset \mathbf{R}^{\mathbf{d}}$ be a bounded polytopic domain with Lipschitz boundary $\partial\Omega$, and let $f \in C(\overline{\Omega})$. Write

$$\mathcal{C} = \{g \in C(\overline{\Omega}) \mid 0 \leq g\}, \tag{1}$$

and consider for given $g \in \mathcal{C}$ the reaction-diffusion problem to find $u^g = u(g) \in C^2(\overline{\Omega})$ for which

$$-\Delta u^g + g u^g = f \quad \text{in} \quad \Omega, \quad \text{and} \quad u^g = 0 \quad \text{on} \quad \partial\Omega. \tag{2}$$

We assume that for each $g \in \mathcal{C}$ a solution u^g of (2) exists. Notice that u^0 corresponds to $g = 0$, the pure diffusion problem.

1.1 Maximum principle and comparison principle

It is well-known that each u^g satisfies the maximum principle [14, 16, 17], which is the implication

$$f \leq 0 \quad \Rightarrow \quad u^g \leq 0. \tag{3}$$

The maximum principle induces a comparison principle: if $f \leq 0$ and $g, h \in C(\overline{\Omega})$ then

$$0 \leq h \leq g \quad \Rightarrow \quad u^0 \leq u^h \leq u^g \leq 0. \tag{4}$$

Indeed, the middle inequality in the right-hand side follows from the fact that

$$-\Delta(u^h - u^g) + h(u^h - u^g) = (g - h)u^h \quad \text{in} \quad \Omega, \quad \text{and} \quad u^h - u^g = 0 \quad \text{on} \quad \partial\Omega \tag{5}$$

and the observation that $(g - h)u^h \leq 0$, which implies $u^h - u^g \leq 0$ according to (3). The first inequality follows similarly.

1.2 History and relevance of discrete maximum principles

Already during the early development of numerical methods for problems like (2), it was realized that if a numerical approximation U^g of u^g satisfies the corresponding discrete maximum principle,

$$f \leq 0 \quad \Rightarrow \quad U^g \leq 0, \tag{6}$$

uniform error bounds for the method could be derived. For the finite difference method, we refer to [2, 4, 7, 9, 15, 19]. Later, similar discrete maximum principles were proved for finite volume and finite element approximations of elliptic and parabolic problems: see [11] and the references therein. In particular, conditions on the simplicial finite element partitions of $\overline{\Omega}$ were given in order for discrete maximum principles to hold. For linear and nonlinear diffusion problems this led to the condition that all dihedral angles between facets of simplices in the finite element partition should be non-obtuse, whereas for reaction-diffusion problems (2), the dihedral angles were even supposed to be acute [6, 10]. With the goal to derive uniform error bounds for the finite element method, (6) was proved in [6] provided that all dihedral angles of the simplices in the finite element partition are acute, and their diameters small enough.

1.3 Motivation and outline of this paper

Our main contribution is in Section 3. We will make the conditions in [6] explicit and verifiable in terms of dihedral angles and heights of simplices on the one hand, and the magnitude of the reaction coefficient $\|g\|_\infty$ and the spatial dimension d on the other. Moreover, in Section 4 we discuss their concrete realization: as a matter of fact, it turns out that the conditions can never be satisfied for $d \geq 5$. Before that, in Section 2 we discuss a particular type of numerical integration that leads to a fully discrete finite element method. This is necessary because the conditions for (6) turn out to depend on g, whereas g needs to be integrated in the finite element formulation. This can, in general, not be done exactly.

1.4 Why the discrete maximum principle can fail

First however, we will show that the complications with the discrete maximum principle for $g \neq 0$ are already present for $d = 1$. For $j \in \{0, \dots, 15\}$, we apply the method of Section 2.2 to problem (2) on the unit interval with choices for f_j and g_j for f and g, defined by

$$f_j(x) = 2^j f(x) \quad \text{with} \quad f(x) = -(2x-1)^2, \quad \text{and} \quad g_j(x) = 2^j. \qquad (7)$$

Due to their simplicity, the computations could be performed in exact arithmetic. Left in Figure 1, all the finite element approximations U^{g_j} are shown in the same picture, together with f, marked by circles ('o'). Each U^{g_j} is continuous on $[0, 1]$ and linear on each sub-interval $I_k = [(k-1)/4, k/4]$, where $k \in \{1, \dots, 4\}$. For $j \to 15$, the values $U^{g_j}(\frac{1}{2})$ tend to a substantially positive value of about 0.2, while the graphs of U^{g_j} seem to converge to the W-shape with vertical coordinates $0, -0.54, 0.2, -0.54, 0$. This phenomenon is not hard to understand. The scaling of f in (7) with a factor $2^j = g$ yields, by linearity, a scaling of U^{g_j} by 2^j as well. In particular, it does not influence positivity and negativity. But it does turn the problem into an equivalent family of singularly perturbed problems

$$-\varepsilon \Delta u_\varepsilon + u_\varepsilon = f \quad \text{in} \quad \Omega, \quad \text{and} \quad u_\varepsilon = 0 \quad \text{on} \quad \partial\Omega, \quad \text{with} \quad \varepsilon = 2^{-j}, \qquad (8)$$

of which the solution u_ε in this simple one-dimensional example can be given exactly as

$$u_\varepsilon(x) = (1+8\varepsilon)\left[\frac{\mathrm{e}^{\varepsilon^{-\frac{1}{2}}}}{1+\mathrm{e}^{\varepsilon^{-\frac{1}{2}}}}\mathrm{e}^{-x\varepsilon^{-\frac{1}{2}}} + \frac{1}{1+\mathrm{e}^{\varepsilon^{-\frac{1}{2}}}}\mathrm{e}^{x\varepsilon^{-\frac{1}{2}}}\right] + f(x) - 8\varepsilon. \qquad (9)$$

The graphs of the functions u_ε with the values of $\varepsilon = 2^{-j}$ for $j \in \{0, \ldots, 15\}$ are shown in the right picture of Figure 1. Clearly, for $x \in (0,1)$, $u_\varepsilon(x)$ tends to $f(x)$ for $\varepsilon \to 0$.

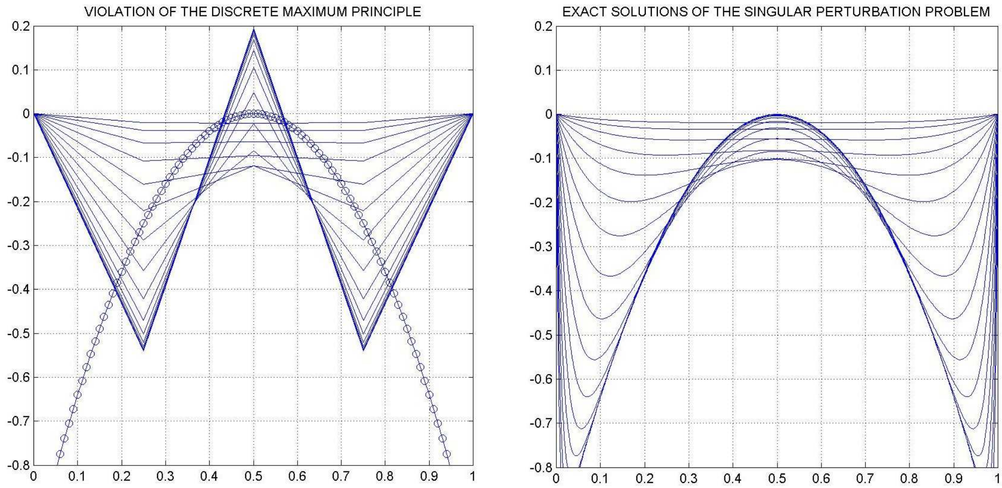


Figure 1. Violation of the discrete maximum principle for a 1d reaction-diffusion problem.

On a fixed partition, as ε tends to zero, the finite element approximations U^{g_j} will tend to the L^2-orthogonal projection U^∞ of u_0 onto the space of continuous piecewise linear functions that vanish at the boundary. This is so because the discretized diffusion disappears for $\varepsilon \to 0$ and the reaction term remains. To minimize the L^2-distance between U^∞ and u_0, a logical overshoot must take place at the midpoint, violating the discrete maximum principle.

2 Preliminaries

We will use the standard notation $H^k(\Omega)$ for the Sobolev space of order k, with norm and semi-norm $\|\cdot\|_k$ and $|\cdot|_k$, respectively. Moreover, we write $H^{-1}(\Omega)$ for the topological dual of $H_0^1(\Omega)$ with norm

$$\|w\|_{-1} = \sup_{0 \neq v \in H_0^1(\Omega)} \frac{\langle w, v\rangle}{\|v\|_1}, \qquad (10)$$

where $\langle \cdot, \cdot \rangle$ is the dual pairing between $H^{-1}(\Omega)$ and $H_0^1(\Omega)$.

2.1 Weak formulation

Consider the weak formulation of (2) aiming to find $u \in H_0^1(\Omega)$ such that for all $v \in H_0^1(\Omega)$,

$$a(g; u^g, v) = (f, v), \tag{11}$$

where the bilinear form

$$a(g; \cdot, \cdot) : H_0^1(\Omega) \times H_0^1(\Omega) \to \mathbf{R} : \quad \mathrm{a(g; w, v) = (\nabla w, \nabla v) + (gw, v)} \tag{12}$$

is easily verified to be continuous. The Poincaré inequality guarantees that there exists a constant $\alpha > 0$ such that for all $g \in \mathcal{C}$ and all $v \in H_0^1(\Omega)$,

$$\alpha \|v\|_1^2 \leq a(g; v, v), \tag{13}$$

hence $a(g; \cdot, \cdot)$ is also coercive. Consequently, the Lax-Milgram lemma provides unique weak solutions u^g of (11) that coincide with the classical solutions u^g of (2).

2.2 Finite element discretization

Let $\mathcal{T}$ be a face-to-face simplicial finite element partition of $\overline{\Omega}$. Denote the vertices in $\mathcal{T}$ by $v_1, \ldots, v_{n+m}$, and such that

$$v_j \in \partial\Omega \quad \Leftrightarrow \quad n + 1 \leq j \leq n + m. \tag{14}$$

Let $V \subset H^1(\Omega)$ be the space of continuous piecewise linear functions relative to $\mathcal{T}$ with the usual nodal basis $\phi_1, \ldots, \phi_{n+m}$ and set

$$V_0 = V \cap H_0^1(\Omega). \tag{15}$$

Then $\phi_1, \ldots, \phi_n$ is the nodal basis for V_0, and the finite element approximation U^g of u^g is the unique function from V_0 such that for all $v \in V_0$,

$$a(g; U^g, v) = (f, v). \tag{16}$$

Notice that if $g \in \mathcal{C}$, due to (13) and (16), we have that

$$\alpha \|U^g\|_1^2 \leq a(g; U^g, U^g) = (f, U^g) \leq \|f\|_{-1} \|U^g\|_1, \tag{17}$$

and hence, irrespective of g,

$$\|U^g\|_1 \leq \frac{1}{\alpha} \|f\|_{-1}. \tag{18}$$

In the next section, this bound will be used to control the error introduced by the following fully discrete formulation, which includes a convenient type of quadrature.

2.3 Fully discrete finite element method

Define the usual nodal interpolation operator

$$\Pi : C(\overline{\Omega}) \to V : w \mapsto \Pi w = \sum_{j=1}^{n+m} w(v_j)\phi_j. \tag{19}$$

Clearly, if $g \in \mathcal{C}$ then $\Pi g \in \mathcal{C}$. Thus, if we consider the problem to find $U^{\Pi g} \in V_0$ such that for all $v \in V_0$,

$$a(\Pi g; U^{\Pi g}, v) = (\Pi f, v), \tag{20}$$

then due to (18) we have that

$$\|U^{\Pi g}\|_1 \leq \frac{1}{\alpha}\|\Pi f\|_{-1}. \tag{21}$$

We can now provide a bound for the difference between U^g and the actually computed $U^{\Pi g}$.

Proposition 2.1 *Let U^g solve (16) and $U^{\Pi g}$ solve (20). Then,*

$$\|U^g - U^{\Pi g}\|_1 \leq \frac{1}{\alpha}\|f - \Pi f\|_{-1} + \frac{1}{\alpha^2}\|g - \Pi g\|_\infty \|\Pi f\|_{-1}. \tag{22}$$

Proof. From (20) we observe that for all $v \in V_0$,

$$a(g; U^{\Pi g}, v) = (\Pi f, v) + ((g - \Pi g)U^{\Pi g}, v). \tag{23}$$

Write $Z^g = U^g - U^{\Pi g}$. Then together with (13) and (16), equality (23) gives that

$$\alpha\|Z^g\|_1^2 \leq a(g; Z^g, Z^g) = (f - \Pi f, Z^g) - ((g - \Pi g)U^{\Pi g}, Z^g). \tag{24}$$

Using the rather crude bound

$$|((g - \Pi g)U^{\Pi g}, Z^g)| \leq \|g - \Pi g\|_\infty \|U^{\Pi g}\|_1 \|Z^g\|_1, \tag{25}$$

completes, after applying (21), the proof. □

This result shows that for f and g smooth enough, the proposed fully discrete scheme (20) results in an approximation $U^{\Pi g}$ of U^g with similar approximation quality.

In Section 3 we will show that if $f \leq 0$, and the elements of the triangulation $\mathcal{T}$ satisfy certain angle properties, then $U^g \leq 0$. Since $f \leq 0$ immediately implies that $\Pi f \leq 0$, and similarly that $g \geq 0$ implies that $\Pi g \geq 0$, we will also have that $U^{\Pi g} \leq 0$.

3 Conditions for the discrete maximum principle

With respect to the nodal basis, U^g can be written as

$$U^g = \sum_{j=1}^{n} u_j^g \phi_j, \quad \text{with} \quad u_j^g = U^g(v_j) \ \text{ for all } j \in \{1, \ldots, n\}. \tag{26}$$

Define for $i, j \in \{1, \ldots, n\}$ the matrices A and M^g by

$$A = (a_{ij})_{i,j=1}^n, \ M^g = (m_{ij}^g)_{i,j=1}^n, \quad \text{with} \quad a_{ij} = (\nabla\phi_i, \nabla\phi_j), \ m_{ij}^g = (g\phi_i, \phi_j).$$

Then the vector $U_g = (u_1^g, \ldots, u_n^g)^*$ of coordinates u_j^g of U^g defined in (26) satisfies

$$B^g U_g = F, \quad \text{where} \quad B^g = A + M^g, \tag{27}$$

and $F = (f_1, \ldots, f_n)^*$ with $f_j = (f, \phi_j)$ for each $j \in \{1, \ldots, n\}$. Since $\phi_j \geq 0$ for all $j \in \{1, \ldots, n\}$, the inequality $f \leq 0$ in (28) implies that $F \leq 0$ in (27), where here and further on, a matrix- or vector inequality is meant to be taken entry-wise. Moreover, because $U^g \in V_0$, its extrema are taken at certain vertices of $\mathcal{T}$. Thus, in view of (26), the discrete maximum principle (6) can be rephrased linear algebraically as

$$F \leq 0 \quad \Rightarrow \quad U_g \leq 0. \tag{28}$$

In the following, we will study the discrete maximum principle in terms of linear algebra.

3.1 The discrete maximum principle in terms of linear algebra

A sufficient condition for (28) to hold is obviously that

$$(B^g)^{-1} \geq 0, \tag{29}$$

because U_g is then a linear combination of columns of $(B^g)^{-1}$ with non-positive coefficients.

Remark 3.1 It is not clear if condition (29) is necessary: if f is non-zero on $T \in \mathcal{T}$ with $T \cap \partial\Omega = \emptyset$, then F will have at least $d + 1$ non-zero entries, in general. In that case, the product $(B^g)^{-1}F$ will not be a single column of B^g but a non-trivial linear combination of $d + 1$ columns.

Condition (29) is satisfied if B^g is a so-called Stieltjes matrix (see Varga [18, p. 85]). We will work with this concept because it avoids irreducibility of B^g, which does not always hold [8].

Definition 3.2 (Stieltjes Matrix) A matrix is called a *Stieltjes matrix* if it is symmetric positive definite and has non-positive off-diagonal entries.

Notice that B^g is symmetric positive definite due to (12) and (13). Hence, it remains to prove that it has non-positive off-diagonal entries. First, we introduce some additional notations.

Definition 3.3 Let $d \geq 1$. For a given d-simplex T with facets F_i and F_j, denote their proper volumes by $|F_i|$, $|F_j|$, and $|T|$, where we use the convention that $|F_i| = |F_j| = 1$ if $d = 1$. For $d > 1$ the interior dihedral angle α_{ij} between F_i and F_j is defined as

$$\alpha_{ij} = \pi - \gamma_{ij}, \tag{30}$$

where $\gamma_{ij} \in [0, \pi]$ is the angle between outward normals q_i and q_j to F_i and F_j, respectively. To stress the dependence on the facets, we will write $\cos(F_i, F_j)$ for $\cos(\alpha_{ij})$. Finally, we write h_j for the (positive) height of T above F_j, which satisfies

$$h_j = \frac{d|T|}{|F_j|}, \tag{31}$$

relating the volume of T to that of its facets.

3.2 The pure diffusion problem

First we recall the case $g = 0$, that corresponds to the pure diffusion problem. The results for $d \leq 3$ are well-known [13]. For arbitrary d we refer to [3, 20].

Proposition 3.4 *Let $i, j \in \{1, \dots, n\}$ be distinct, and choose a d-simplex $T \in \mathcal{T}$ with*

$$T \subset \operatorname{supp}(\phi_i) \cap \operatorname{supp}(\phi_j). \tag{32}$$

Write F_i and F_j for the facets of T opposite v_i and v_j, respectively. Then

$$(\nabla\phi_i, \nabla\phi_j)_T = -\frac{\cos(F_i, F_j)}{h_i h_j}|T|. \tag{33}$$

Corollary 3.5 *If $\mathcal{T}$ contains no simplices with obtuse dihedral angles, then B^0 has non-positive off-diagonal entries. Hence, B^0 is a Stieltjes matrix and (6) holds.*

Proof. Follows immediately from (33) and the fact that

$$a_{ij} = \sum_{T \in \mathcal{T}} (\nabla\phi_i, \nabla\phi_j)_T. \tag{34}$$

The non-obtuseness condition on $\mathcal{T}$ guarantees that each term in the sum is non-positive. □

Remark 3.6 The outward normals to an interval make an angle of $\gamma_{ij} = \pi$. Therefore, using (30), we find that if $d = 1$,

$$\cos(F_i, F_j) = 1, \tag{35}$$

showing that (6) holds for any partition $\mathcal{T}$. In fact, the finite element approximation U^0 is then equal to Πu^0, which proves the discrete maximum principle in an alternative way.

3.3 The reaction-diffusion problem

Now we will continue with the general case $g \neq 0$ and consider B^g. The complication is that the off-diagonal entries of m^g_{ij} are positive. Indeed, [5, p. 201] yields that for $i \neq j$,

$$(\phi_i, \phi_j)_T = \frac{|T|}{(d+1)(d+2)}. \tag{36}$$

The requirement $a_{ij} + m^g_{ij} \leq 0$ results in the following restriction on the shape of the simplices.

Theorem 3.7 *If for each pair of distinct facets F_i and F_j of any simplex $T \in \mathcal{T}$ we have that*

$$\frac{\cos(F_i, F_j)}{h_i h_j} \geq \frac{\|g\|_\infty}{(d+1)(d+2)}, \tag{37}$$

then B^g has non-positive off-diagonal entries $a_{ij} + m^g_{ij}$ and is therefore a Stieltjes matrix.

Proof. Let $i, j \in \{1, \ldots, n\}$ be distinct. Due to $\phi_i \geq 0, \phi_j \geq 0$, and $g \geq 0$ we infer that

$$a_{ij} + m^g_{ij} = \sum_{T \in \mathcal{T}} (\nabla \phi_i, \nabla \phi_j)_T + (g\phi_i, \phi_j)_T \leq \sum_{T \in \mathcal{T}} (\nabla \phi_i, \nabla \phi_j)_T + \|g\|_\infty (\phi_i, \phi_j)_T. \tag{38}$$

The statement follows from combining (33) with (36), which shows that for a given $T \in \mathcal{T}$,

$$(\nabla \phi_i, \nabla \phi_j)_T + \|g\|_\infty (\phi_i, \phi_j)_T = -|T| \left(\frac{\cos(F_i, F_j)}{h_i h_j} - \frac{\|g\|_\infty}{(d+1)(d+2)} \right), \tag{39}$$

where F_i and F_j are the facets of T opposite v_i and v_j, respectively. □

Remark 3.8 In [6] the authors derived the similar, though less sharp condition

$$\frac{\cos(F_i, F_j)}{h^2} \geq \|g\|_\infty, \tag{40}$$

where h is the maximum diameter of all simplices in $\mathcal{T}$. For instance, for a planar triangulation into equilateral triangles (see also Section 4.1) this forces h to be four times smaller as required in (37). Solving the corresponding finite element system would then cost at least sixteen times more.

Remark 3.9 If $g > 0$ is constant, the inequality in (38) becomes an equality. Nevertheless, (37) may not be necessary for non-positivity of $a_{ij} + m^g_{ij}$, because a positive term (39) in the sum in (38) may be compensated for by the other terms. Moreover, B^g does not need to be a Stieltjes matrix for (29) to hold, and even (29) may not be a necessary condition (see Remark 3.1). Still, (37) seems to be necessary for $d = 1$ and $d = 2$ in the experiments of Section 4.

3.4 Discrete comparison principle

Let $g, h \in C(\overline{\Omega})$ with $0 \leq h \leq g$. Consider the finite element problems to find $U^h, U^g \in V_0$ such that for all $v \in V_0$,

$$a(h; U^h, v) = (f, v) \quad \text{and} \quad a(g; U^g, v) = (f, v). \tag{41}$$

Similarly as in Section 1.1, we are now able to derive a discrete comparison principle.

Theorem 3.10 *Let $g, h \in C(\overline{\Omega})$. Assume that $\mathcal{T}$ is a finite element partition satisfying (37). Then the solutions U^g and U^h of (41) satisfy*

$$(f \leq 0 \quad \text{and} \quad 0 \leq h \leq g) \quad \Rightarrow \quad U^h \leq U^g \leq 0. \tag{42}$$

Proof. Subtracting the second equality from the first shows that for all $v \in V_0$,

$$a(h; U^h - U^g, v) = a(g; U^g, v) - a(h; U^g, v) = ((g - h)U^g, v). \tag{43}$$

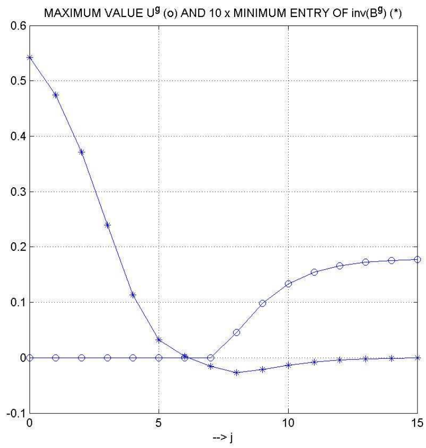
Since $\|g\|_\infty \geq \|h\|_\infty$, the partition $\mathcal{T}$ also satisfies (37) with g replaced by h. Thus, both problems in (41) satisfy the discrete maximum principle. Therefore, $f \leq 0$ implies that $U^g \leq 0$. From this we get that $(g - h)U^g \leq 0$, which in turn implies that $U^h \leq U^g$. $\square$

4 Numerical experiments

For $d = 1$, all dihedral angles are zero, and the condition of Theorem 3.7 reduces to the requirement that

$$h^2 \leq \frac{6}{\|g\|_\infty}, \tag{44}$$

where h is the length of the largest sub-interval in the partition. Notice that the bound on h^2 resulting from (40) is six times smaller. Returning to the experiments in Section 1.4, where we fixed $h = 1/4$, condition (44) is violated if $\|g\|_\infty > 96$. In the picture in Figure 2 below, we plotted the maximum value of U^g against j in the graph with circles ('o'), and (for clarity 16 times) the minimal entry of $(B^g)^{-1}$ in asterisks ('*'). At the right, the minimal entries of $(B^g)^{-1}$ are given around the critical value 96. Even though the discrete maximum principle is not violated immediately, the non-negativity of $(B^g)^{-1}$ is lost straight away.



g	$\min\,(B^g)^{-1}$
91	$3.4914e-006$
92	$2.1868e-006$
93	$1.2040e-006$
94	$5.2389e-007$
95	$1.2824e-007$
96	0
97	$-7.1344e-005$
98	$-1.4074e-004$
99	$-2.0826e-004$

Figure 2. Verifying if $\|g\|_\infty > 96$ violates the non-negativity of $(B^g)^{-1}$.

Remark 4.1 Without giving numerical evidence, we note that by taking for f the function

$$f(x) = -(2x-1)^{10}, \tag{45}$$

instead of (7), and using the fully discrete method of Section 2.3, the discrete maximum principle was violated already for $g = 97$. We suspect that raising the power in (45) further will show that (44) is indeed necessary for the discrete maximum principle to hold, though round-off may obscure the results.

4.1 An experiment with equilateral triangles

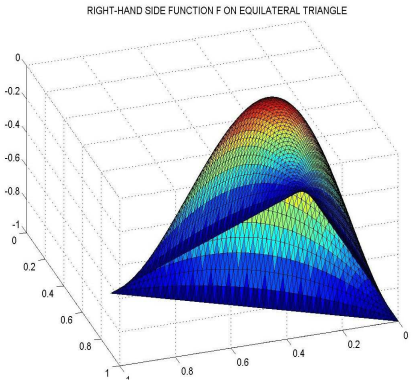
It is easily verified that (37) results in a similar requirement as in (44) for planar partitions into equilateral triangles, namely,

$$h^2 \leq \frac{8}{\|g\|_\infty}. \tag{46}$$

Again, h stands for the edge length in the partition. As already mentioned in Remark 3.8, the bound from [6] would force h to be four times smaller. We test (46) by taking for Ω the equilateral triangle with vertices $(0,0),(1,0)$, and $(\frac{1}{2}, \frac{1}{2}\sqrt{3})$. As in the one-dimensional case, we take g to be constant, and scale the right-hand side with g,

$$f_g(x,y) = -f(x,y)^{10}g, \quad \text{where} \quad f(x,y) = -1+6\sqrt{3}y(y-x\sqrt{3})(y-(1-x)\sqrt{3}). \tag{47}$$

The function f, which is depicted in Figure 3, is the natural generalization of (45). We use the method of Section 2.3, which includes numerical integration of the right-hand side.



g	min $(B^g)^{-1}$	max U^g
508	$1.7453e-017$	0
509	$4.1074e-018$	0
510	$5.3640e-019$	0
511	$1.6624e-020$	0
512	$-1.3878e-017$	0
513	$-2.3443e-005$	$4.6807e-005$
514	$-4.6787e-005$	$1.1783e-004$
515	$-7.0032e-005$	$2.2662e-004$
516	$-9.3180e-005$	$3.3527e-004$

Figure 3. Verifying if $\|g\|_\infty > 512$ violates the non-negativity of $(B^g)^{-1}$.

Subdividing Ω into 64 equilateral triangles by three consecutive uniform refinements, gives that $h = 1/8$. Thus, condition (46) becomes

$$\|g\|_\infty \leq 512. \tag{48}$$

As is clear from the tabular in Figure 3, this can indeed be confirmed, if we consider the small negative value for $g = 512$ an effect of rounding errors. Also, similarly as for (45), the discrete maximum principle is violated already at $g = 513$. Again without presenting evidence, we note that without the exponent 10 in (47), this took much larger values of g.

4.2 An experiment with right triangles

Even though the previous experiment shows, that there exist triangulations for which (37) is a necessary condition, we will conclude our investigations with showing that for a triangulation into right triangles as in the left of Figure 4, the discrete maximum principle may still hold.

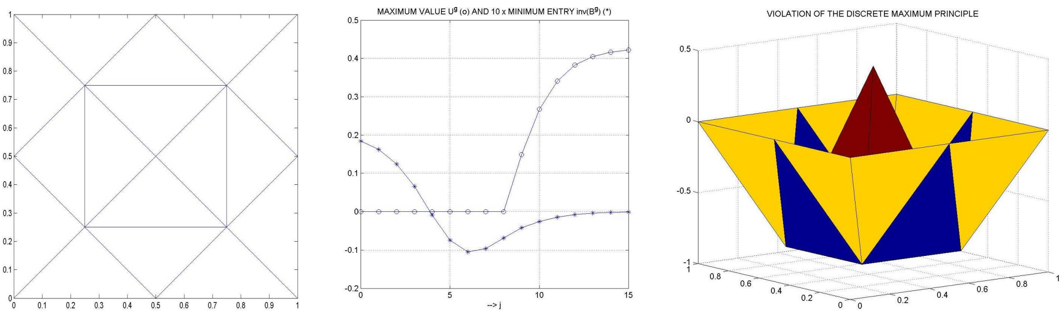


Figure 4. Right triangles do not necessarily violate the discrete maximum principle.

The right-hand side functions for this experiment were again scaled with g, where $g = 2^j$ for $j \in \{0, \ldots, 15\}$, and

$$f_g(x, y) = f(x, y)g, \quad \text{where} \quad f(x, y) = -1+16x(1-x)y(1-y) \quad \text{on } [0, 1]\times[0, 1]. \tag{49}$$

We use the method of Section 2.3. The middle picture in Figure 4 shows that the minimum entry of $(B^g)^{-1}$ becomes negative around $g = 14$, whereas the discrete maximum principle is lost around $g = 256$. In fact, $g = 264$ is the smallest integer for which U^g has a positive value. In the right picture we see the discrete solution for $j = 15$, which is a two-dimensional version of the discrete solution for $j = 15$ in Figure 1.

The fact that the discrete maximum principle holds for moderate reaction coefficients g even though h and $\|g\|_\infty$ do not satisfy (37) may be explained by observing that acute simplices are not needed for convergence of the finite element method. Thus, for h tending to zero, U^g converges to u^g, and u^g satisfies the maximum principle. Complication in this argument is that convergence takes place only in $H^1(\Omega)$ and not in $L^\infty(\Omega)$. In fact, for the latter, the discrete maximum principle was used [6].

Remark 4.2 The recent paper [1], which came to our attention while finishing this paper, may explain the above situation in an alternative way. Here, it is studied linear algebraically which perturbations of A keep the property $A^{-1} \geq 0$ intact. A moderate reaction term gM^g may be such a perturbation.

Remark 4.3 For $d = 3$ and for regular tetrahedra, it can also be explicitly computed which relation h and $\|g\|_\infty$ should satisfy in order for the discrete maximum principle to hold. However, space cannot be filled with regular tetrahedra.

5 Conclusions and final remarks

In the implementation of the finite element method it can be verified if B^g will be a Stieltjes matrix by checking if for each $T \in \mathcal{T}$, the $(d+1) \times (d+1)$ element matrix E_T^g happens to have a positive off-diagonal entry. Those matrices E_T^g are explicitly and easily computed to form B^g from the affine invertible transforms $F_T : \hat{T} \to T, x \mapsto p_0 + Px$, where p_0 together with p_0 plus each of the columns $p_1, \ldots, p_d$ of P are the vertices of T, and $\hat{T}$ is the reference simplex with as vertices the canonical basis vectors of $\mathbf{R^d}$ together with the origin. The computational costs of this verification is only of order d^2t, where t is the number of simplices $T \in \mathcal{T}$, which is modest in comparison to solving the linear system $B^gU^g = F$ in optimal complexity. This approach is however rather naive and only provides an answer to the question if the discrete maximum principle is satisfied or not. In particular, it does not tell, given the partition, which reaction terms could be allowed. Ranging over the partition and computing the quotients in the left-hand side of (37) does.

Condition (37) can only be satisfied if all dihedral angles in the partition are acute, and then only if all products of distinct pairs of heights are small enough. This shows for instance that uniform refinement of a planar triangulation satisfying (37) results in a triangulation that can cope with a reaction term that is even four times larger. In higher dimensions, the situation is less clear. In [12], it was proved that there are no partitions of $\mathbf{R}^5$ into acute simplices, and in $\mathbf{R}^3$ and $\mathbf{R}^4$ there are no algorithms yet known to decompose even a simple polyhedron or polytope like a simplex or a (hyper)cube into acute simplices. This shows that much research remains to be done, both in the area of finding weaker, or alternative, conditions on the partition and in the area of mesh generation and refinement.

References

[1] Bouchon, F., Monotonicity of some perturbations of irreducibly diagonally dominant M-matrices, *Numer. Math.* 105 (2007), 591–601.

[2] Bramble, J. H., Hubbard, B. E., and Thomée, V, Convergence estimates for essentially positive type discrete problems, *Math. Comp.* 23 (1969), 695–710.

[3] Brandts, J., Korotov, S., Křížek, M., Dissection of the path-simplex in $\mathbf{R}^n$ into n path-subsimplices, *Linear Algebra Appl.* 421 (2007), 382–393.

[4] Ciarlet, P. G., Discrete maximum principle for finite-difference operators, *Aequationes Math.* 4 (1970), 338–352.

[5] Ciarlet, P. G., *The finite element method for elliptic problems,* North-Holland, Amsterdam, 1978.

[6] Ciarlet, P. G., Raviart, P.-A., Maximum principle and uniform convergence for the finite element method, *Comput. Methods Appl. Mech. Engrg.* 2 (1973), 17–31.

[7] Collatz, L., *Numerische Behandlung von Differentialgleichungen*, Springer-Verlag, Berlin-Göttingen-Heidelberg, 1951.

[8] Drăgănescu, A., Dupont, T.F., Scott, L.R. Failure of the discrete maximum principle for an elliptic finite element method. *Math. Comp.* 74 (2005), 1–23.

[9] Forsythe, G. E., Wasow, W. R., *Finite-difference methods for partial diffrential equations*, John Wiley & Sons, New York, London, 1960.

[10] Karátson, J., Korotov, S., Discrete maximum principles for finite element solutions of nonlinear elliptic problems with mixed boundary conditions, *Numer. Math.* 99 (2005), 669–698.

[11] Karátson, J., Korotov, S., Křížek, M., On discrete maximum principles for nonlinear elliptic problems, Technical Report **A504**, Helsinki University of Technology, accepted by *Math. Comput. Simulation* in 2005.

[12] Křížek, M., There is no face-to-face partition of $\mathbf{R}^5$ into acute simplices, *Discrete Comput. Geom.* 36 (2006), 381–390.

[13] Křížek, M., Lin Qun, On diagonal dominance of stiffness matrices in 3D, *East-West J. Numer. Math.* 3 (1995), 59–69.

[14] Ladyzhenskaya, O. A., Ural'tseva, N. N., *Linear and quasilinear elliptic equations*, Leon Ehrenpreis Academic Press, New York-London, 1968.

[15] Milne, W. E., *Numerical solution of differential equations*, John Wiley & Sons, New York, Chapman & Hall, London, 1953.

[16] Protter, M., Weinberger, H., *Maximum principles in differential equations*, Prentice-Hall, Englewood Cliffs, New Jersey, 1967.

[17] Sperb, R., *Maximum principles and their applications*, Academic Press, New York, 1981.

[18] Varga, R., *Matrix iterative analysis*, Prentice Hall, New Jersey, 1962.

[19] Varga, R., On a discrete maximum principle, *J. SIAM Numer. Anal.* 3 (1966), 355–359.

[20] Xu, J., Zikatanov, L., A monotone finite element scheme for convection-diffusion equations, *Math. Comp.* 68 (1999), 1429–1446.

(continued from the back cover)

A519 Teemu Lukkari
Elliptic equations with nonstandard growth involving measure data
February 2007

A518 Niko Marola
Regularity and convergence results in the calculus of variations on metric spaces
February 2007

A517 Jan Brandts , Sergey Korotov , Michal Krizek
Simplicial finite elements in higher dimensions
February 2007

A516 Sergey Repin , Rolf Stenberg
Two-sided a posteriori estimates for the generalized stokes problem
December 2006

A515 Sergey Korotov
Global a posteriori error estimates for convection-reaction-diffusion problems
December 2006

A514 Yulia Mishura , Esko Valkeila
An extension of the L'evy characterization to fractional Brownian motion
December 2006

A513 Wolfgang Desch , Stig-Olof Londen
On a Stochastic Parabolic Integral Equation
October 2006

A512 Joachim Schöberl , Rolf Stenberg
Multigrid methods for a stabilized Reissner-Mindlin plate formulation
October 2006

A511 Carlo Lovadina , Mikko Lyly , Rolf Stenberg
A posteriori estimates for the Stokes eigenvalue problem
February 2007

HELSINKI UNIVERSITY OF TECHNOLOGY INSTITUTE OF MATHEMATICS
RESEARCH REPORTS

The list of reports is continued inside. Electronical versions of the reports are available at *http://www.math.hut.fi/reports/* .

A526 Beirao da Veiga Lourenco , Jarkko Niiranen , Rolf Stenberg
A family of C^0 finite elements for Kirchhoff plates II: Numerical results
May 2007

A525 Jan Brandts , Sergey Korotov , Michal Krizek
The discrete maximum principle for linear simplicial finite element approximations of a reaction-diffusion problem
July 2007

A524 Dmitri Kuzmin , Antti Hannukainen , Sergey Korotov
A new a posteriori error estimate for convection-reaction-diffusion problems
May 2007

A522 Antti Hannukainen , Sergey Korotov , Marcus Ruter
A posteriori error estimates for some problems in linear elasticity
March 2007

A521 Sergey Korotov , Ales Kropac , Michal Krizek
Strong regularity of a family of face-to-face partitions generated by the longest-edge bisection algorithm
April 2007

ISBN - 978-951-22-8781-9
ISSN 0784-3143