An implementation detail about the scaling of monomial bases in polytopal finite element methods

Applied Mathematics Letters 159 (2025) 109281

Contents lists available at ScienceDirect

Applied Mathematics Letters

journal homepage: www.elsevier.com/locate/aml

Regular Article

An implementation detail about the scaling of monomial bases in polytopal finite element methods

Matteo Cicuttin

Dipartimento di Scienze Matematiche "G. L. Lagrange", Politecnico di Torino, Corso Duca degli Abruzzi, 24, Torino, 10129, Italy

ARTICLE INFO

Dataset link: https://github.com/wareHHOuse/diskpp

Keywords:
Polytopal FEM
Matrix conditioning
Polynomial bases

ABSTRACT

The usual definition of scaled monomials found in polytopal finite elements literature leads to elemental matrices with an unnecessarily high condition number. A trivial but apparently overlooked rescaling significantly improves the situation. The extent of the improvement is demonstrated numerically.

1. Problem statement

Let $T \subset \mathbb{R}^d$ be a d-dimensional polytope and $\mathrm{vtx}(T)$ the set of points of $\mathbb{R}^d$ which are its vertices. The diameter of T is $h_T := \max_{\boldsymbol{v}_m, \boldsymbol{v}_n \in \mathrm{vtx}(T)} \|\boldsymbol{v}_m - \boldsymbol{v}_n\|_2$, where $\|\cdot\|_2$ is the euclidean norm in $\mathbb{R}^d$; $\boldsymbol{x}_T \in T$ finally denotes the barycenter (center of mass) of T. $\mathbb{P}^k_d(T)$ denotes the set of d-variate polynomials of degree k with support on T and a generic polynomial $p \in \mathbb{P}^k_d(T)$ can be written as the linear combination

$$p(\boldsymbol{x}) = \sum_{i=1}^{N^k_d} p_i \phi_{T,i}(\boldsymbol{x}),$$

where $\boldsymbol{x} = (x_1, \ldots, x_d) \in T$ is the evaluation point, p_i denotes the ith coefficient of the linear combination and $\phi_{T,i}$ denotes the ith function of the chosen polynomial basis of size N^k_d.

In finite element methods one needs to construct different elemental matrices out of these basis functions, for example

$$M_{ij} = \int_T \phi_{T,i} \phi_{T,j} \mathrm{d}\boldsymbol{x}, \qquad K_{ij} = \int_T \nabla\phi_{T,i} \cdot \nabla\phi_{T,j} \mathrm{d}\boldsymbol{x},$$

where M is the *mass matrix* and K the *stiffness matrix*. It is well known that the choice of the basis functions greatly influences the condition number of such matrices [1, Chapter 3].

2. Scaled monomials

In polytopal discretizations for PDEs the *scaled monomials* are very frequently used as basis, as they are cheap to evaluate and, in contrast for example with Legendre basis, they can be constructed irrespectively of the shape of T (even if T has non-planar faces). On the other hand, scaled monomials can lead to ill-conditioned matrices, and literature on orthogonalization techniques to control this issue abounds, see for example [2]. Such techniques however are computationally expensive.

E-mail address: matteo.cicuttin@polito.it.

https://doi.org/10.1016/j.aml.2024.109281
Received 16 July 2024; Received in revised form 16 August 2024; Accepted 16 August 2024
Available online 22 August 2024

The usual definition of scaled monomials is as follows: let $\alpha \in \mathbb{N}^d$ be a multi-index with magnitude $|\alpha| := \sum_{1\le i\le d} \alpha_i$. By setting for all $\alpha \in \mathbb{N}^d$

$$\mu_{T,\alpha}(\boldsymbol{x}) := \prod_{1\le i\le d} \left(\frac{x_i - x_{T,i}}{h_T}\right)^{\alpha_i}, \tag{1}$$

the usual scaled monomial basis for $\mathbb{P}_d^k(T)$ is $\{\mu_{T,\alpha} : \alpha \in \mathbb{N}^d, |\alpha| \le k\}$.

Using the monomials (1) to form a polynomial basis is common practice in polytopal FEM literature, see for example [2–5], [6, Appendix B], [7]. However, as it turns out, (1) leads to unnecessarily ill-conditioned elemental matrices and it is *not* what should be implemented.

To formally illustrate the issue the 1D situation is analyzed; the same reasoning can be readily extended to any dimension. Let $T = [a, b]$: according to the definitions given in Section 1, $h_T = b - a$ is the diameter of T and $x_T = (a + b)/2$ is its barycenter. The mass matrix of T therefore is

$$M_{ij}^{(\eta)} = \int_a^b \left(\eta\frac{x - x_T}{h_T}\right)^i \left(\eta\frac{x - x_T}{h_T}\right)^j dx,$$

where an additional real scaling factor η was introduced in the basis functions if compared with (1) above. The entries of the mass matrix are then computed explicitly, obtaining

$$M_{ij}^{(\eta)} = \int_a^b \left(\eta\frac{x - x_T}{h_T}\right)^{i+j} dx = \left(\frac{\eta}{2}\right)^{i+j} \frac{h_T}{2} \frac{1 - (-1)^{i+j+1}}{i + j + 1}. \tag{2}$$

Inspecting (2) it can be noticed that any $\eta \neq 2$ introduces a scaling dependent on i and j that deteriorates the condition number: indeed, by analyzing the Gershgorin circles of $M_{ij}^{(\eta)}$, for $\eta = 2$ they are centered on $M_{ii}^{(2)} = \frac{2h_T}{2i+1}$, whereas for $\eta = 1$ their center is $M_{ii}^{(1)} = \frac{2h_T}{2i+1} 2^{-2i}$. Therefore, in the case $\eta = 1$, the smallest eigenvalues are pushed towards zero in an exponential fashion.

A very practical rule of thumb deriving from the result above is that optimally scaled monomials on T are a stretched and translated version of the *unscaled* monomials in $[-1, 1]$. It is readily seen that (1) violates this rule: since $|x - x_T| \le h_T/2$, the monomials defined by (1) are a stretched and translated version of the unscaled monomials in $[-1/2, 1/2]$.

Extending the reasoning above (and, consequently, the given rule of thumb) to the multi-dimensional case it can be also deduced that, when $d \ge 2$, in anisotropic elements a single value of h_T for all the directions is not appropriate. To rectify this, the half-sizes of the rectangular bounding box around T are collected in the vector $\mathcal{H}_T$ with components

$$\mathcal{H}_{T,i} := \max_{\boldsymbol{v}_m, \boldsymbol{v}_n \in \mathrm{vtx}(T)} \frac{|v_{m,i} - v_{n,i}|}{2}, \qquad i \in \{1, \dots, d\}, \tag{3}$$

which is subsequently used to give the improved definition of scaled monomials

$$\hat{\mu}_{T,\alpha}(\boldsymbol{x}) := \prod_{1\le i\le d} \left(\frac{x_i - x_{T,i}}{\mathcal{H}_{T,i}}\right)^{\alpha_i}. \tag{4}$$

As a last remark, it is noted that the ideas discussed above are applicable without modification to skeletal methods [3,7] employing $(d - 1)$-variate polynomials attached to the *faces* of T: it can be readily shown that, as long as the rule of thumb given above is fulfilled, the resulting bases provide optimal matrix conditioning.

3. Numerical experiments

The mass and stiffness matrices for some elements are computed using differently scaled bases, subsequently their condition numbers are computed using a singular value decomposition (on stiffness matrices, the row and the column corresponding to the constant were removed before computing the SVD). The experiments were implemented in the `DiSk++` open-source library (https://github.com/wareHHOuse/diskpp).

As first numerical test the element $[0, 1]$ is considered and matrices are constructed using $\eta = 1$ and $\eta = 2$. In Table 1 it can be observed that already at $k = 4$ the mass matrix computed with $\eta = 1$ has a conditioning 5 orders of magnitude worse than the matrix computed with $\eta = 2$; on the stiffness matrix a gap of 3 orders of magnitude is observed.

For the 2D tests $K_\rho^n := \{(\rho\cos(2\pi i/n), \rho\sin(2\pi i/n)) : 0 \le i < n\}$ denotes the set of n points lying on a circumference of radius ρ and, clearly, the convex hull of K_ρ^n yields a regular polygon of n edges. The standard rotation matrix $\mathbf{R}_\theta$ and the matrix $\mathbf{T} = \mathrm{diag}(1, 0.1)$ are then introduced to generate from K_ρ^n some anisotropic test elements as described in the following.

For the second numerical test, the elements identified by the sets $\hat{K}_{0.4}^3$ and $\hat{K}_{0.4}^5$ obtained by applying $\mathbf{T}$ to the points of $K_{0.4}^3$ and $K_{0.4}^5$ are considered. Those elements are an highly stretched triangle and an highly stretched pentagon respectively, both aligned along the x axis. Mass and stiffness matrices are then formed and their condition number is computed under three settings: (i) using monomials (1), (ii) using monomials (4) where $\mathcal{H}_{T,i} = h_T/2$ for all i and (iii) using monomials (4) with $\mathcal{H}_{T,i}$ set as defined by (3). As reported in Tables 2 and 3, also in this case the use of the scaled monomials defined by (4) yields significant improvements in the condition numbers.

The third and last numerical test explores a setting where the proposed rescaling should fail to provide improved condition numbers. The transformation $\mathbf{R}_{\pi/4}\mathbf{T}$ is applied to $K_{0.4}^5$, yielding a very stretched pentagon $\bar{K}_{0.4}^5$ aligned with neither axis. Indeed,

Table 1
Condition number of the 1D mass matrix for varying polynomial degree k.

k	Mass		Stiffness	
k	$\eta = 1$	$\eta = 2$	$\eta = 1$	$\eta = 2$
1	4.80e+01	3.00e+00	1.00e+00	1.00e+00
2	2.88e+03	1.41e+01	1.20e+01	1.33e+00
3	1.80e+05	6.76e+01	3.23e+02	7.67e+00
4	1.13e+07	3.58e+02	1.13e+04	2.48e+01
5	7.19e+08	1.87e+03	4.57e+05	1.18e+02
6	4.58e+10	1.02e+04	2.02e+07	5.60e+02
8	1.87e+14	3.07e+05	4.62e+10	1.47e+04
10	7.69e+17	9.44e+06	1.21e+14	4.15e+05

Table 2
Condition number of the 2D matrices for the triangle $\hat{K}^3_{0.4}$.

k	Mass			Stiffness		
	(i)	(ii)	(iii)	(i)	(ii)	(iii)
1	1.81e+03	4.51e+02	6.00e+00	1.00e+00	1.00e+00	7.50e+01
2	3.07e+06	2.01e+06	6.56e+01	4.52e+02	1.13e+02	8.44e+01
3	4.79e+09	7.90e+07	7.25e+02	3.48e+05	3.74e+04	3.96e+02
4	7.09e+12	2.99e+10	6.77e+03	3.06e+08	1.34e+07	1.24e+03
5	1.01e+16	1.08e+13	8.64e+04	2.91e+11	5.65e+09	5.76e+03
6	8.63e+18	3.85e+15	9.75e+05	2.89e+14	2.50e+12	2.47e+04
7	2.77e+20	1.17e+18	1.07e+07	1.82e+17	1.13e+15	1.32e+05
8	2.85e+22	6.80e+18	1.43e+08	7.16e+18	4.45e+17	1.07e+06

Table 3
Condition number of the 2D matrices for the pentagon $\hat{K}^5_{0.4}$.

k	Mass			Stiffness		
	(i)	(ii)	(iii)	(i)	(ii)	(iii)
1	1.70e+03	4.26e+02	4.70e+00	1.00e+00	1.00e+00	9.05e+01
2	3.11e+06	2.05e+05	3.45e+01	4.25e+02	1.06e+02	9.61e+01
3	5.26e+09	8.67e+07	2.59e+02	3.54e+05	3.67e+04	4.81e+02
4	8.59e+12	3.59e+10	2.30e+03	3.38e+08	9.84e+06	1.31e+03
5	1.38e+16	1.44e+13	1.85e+04	3.54e+11	3.54e+09	6.16e+03
6	8.99e+17	5.73e+15	1.61e+05	3.94e+14	1.19e+12	2.39e+04
7	6.49e+20	1.73e+18	1.42e+06	3.31e+17	4.47e+14	1.18e+05
8	8.55e+22	1.32e+19	1.28e+07	1.29e+19	1.30e+17	5.31e+05

already at $k = 4$ a huge ill-conditioning is observed (Table 4, columns (iii)). This situation is cured by aligning the monomial basis with the element's inertia axes [4], but applying the appropriate scaling. The standard inertia matrix is therefore introduced

$$\int_T (\boldsymbol{x} - \boldsymbol{x}_T)(\boldsymbol{x} - \boldsymbol{x}_T)^T \mathrm{d}\boldsymbol{x} = \mathbf{Q}\boldsymbol{\Lambda}\mathbf{Q}^T \in \mathbb{R}^{d\times d}, \qquad \boldsymbol{x} \in T,$$

and factorized in a diagonal $\boldsymbol{\Lambda}$ containing its eigenvalues and an orthonormal $\mathbf{Q}$ with its eigenvectors as columns. Letting $\lambda_{\max} = \max(\boldsymbol{\Lambda})$, the linear transformation defined by the matrix $\mathbf{B}^{(\eta)} := \eta h_T^{-1} \sqrt{\lambda_{\max}} \sqrt{\boldsymbol{\Lambda}^{-1}} \mathbf{Q}^T$ allows to evaluate the monomials in a reference frame aligned with the principal axes of T. It can be easily seen that also in this case the correct choice is $\eta = 2$, and the transformation just introduced allows to write the scaled and principal-axis-aligned monomial basis as

$$\bar{\mu}_{T,\boldsymbol{\alpha}}(\boldsymbol{x}) := \prod_{i=1}^{d} (\mathbf{B}^{(2)}(\boldsymbol{x} - \boldsymbol{x}_T))_i^{\alpha_i}. \tag{5}$$

Table 4 reports the comparison of the condition numbers obtained using the setting (iii) described above and using the mapping $\mathbf{B}^{(\eta)}$ as in (5) but with $\eta \in \{1, 2\}$ to evaluate the monomials. It can be noticed how the choice $\eta = 2$ provides significant improvement over (iii) and over the setting $\eta = 1$ already used in [4].

The monomials defined by (5), underlining the choice $\eta = 2$, are therefore the "optimal" monomial basis in terms of finite element matrix conditioning, and it is what should be implemented if monomial bases are to be used. We remark that the computation of $\mathbf{B}^{(2)}$ is inexpensive, as it requires to solve, once per element, only a $d \times d$ eigenvalue problem.

4. Conclusions

Scaled monomials are largely used in polygonal FEM, but at the same time dismissed as ill-conditioned. Unnecessarily high ill-conditioning however frequently results from a wrong scaling of the basis functions. As demonstrated algebraically and with

Table 4
Condition number of the 2D matrices for the pentagon $\bar{K}^5_{0.4}$.

k	Mass			Stiffness		
	(iii)	$\mathbf{B}^{(1)}$	$\mathbf{B}^{(2)}$	(iii)	$\mathbf{B}^{(1)}$	$\mathbf{B}^{(2)}$
1	2.26e+02	1.70e+01	4.26e+00	1.00e+00	1.00e+02	1.00e+02
2	9.60e+04	4.30e+02	3.15e+01	1.13e+02	4.26e+02	1.06e+02
3	3.60e+07	1.17e+04	2.26e+02	3.66e+04	3.32e+03	5.82e+02
4	1.47e+10	3.52e+05	1.94e+03	1.15e+07	3.09e+04	1.76e+03
5	5.65e+12	1.00e+07	1.48e+04	4.16e+09	3.37e+05	8.40e+03
6	2.33e+15	3.03e+08	1.27e+05	1.56e+12	3.94e+06	3.61e+04
7	7.01e+15	9.05e+09	1.08e+06	6.11e+14	4.97e+07	1.83e+05
8	5.42e+16	2.76e+11	9.41e+06	6.98e+15	7.22e+08	9.00e+05

the numerical tests, the ill-conditioning can be strongly mitigated by using an easy to implement and computationally inexpensive variation of the definition normally found in literature.

Data availability

The numerical results for the article were obtained using the DiSk++ open source library, available at https://github.com/wareHHOuse/diskpp.

Acknowledgments

The author is member of the Gruppo Nazionale di Calcolo Scientifico at Istituto Nazionale di Alta Matematica (INdAM-GNCS) and kindly acknowledges partial financial support from 2024 project CUP: E53C23001670001.

References

[1] J.S. Hesthaven, T. Warburton, Nodal Discontinuous Galerkin Methods: Algorithms, Analysis, and Applications, in: Texts in Applied Mathematics, Springer, 2008.
[2] S. Berrone, A. Borio, Orthogonal polynomials in badly shaped polygonal elements for the virtual element method, Finite Elem. Anal. Des. 129 (2017) 14–31.
[3] L. Beirão da Veiga, F. Brezzi, L.D. Marini, A. Russo, The Hitchhiker's guide to the virtual element method, Math. Models Methods Appl. Sci. 24 (08) (2014) 1541–1573.
[4] S. Berrone, G. Teora, F. Vicini, Improving high-order VEM stability on badly-shaped elements, Math. Comput. Simulation 216 (2024) 367–385, URL https://www.sciencedirect.com/science/article/pii/S0378475423004287.
[5] F. Dassi, S. Scacchi, Parallel solvers for virtual element discretizations of elliptic equations in mixed form, Comput. Math. Appl. 79 (7) (2020) 1972–1989, URL https://www.sciencedirect.com/science/article/pii/S0898122119303797.
[6] D.A. Di Pietro, J. Droniou, The Hybrid High-Order Method for Polytopal Meshes: Design, Analysis, and Applications, in: Modeling, Simulation and Applications, Springer, 2020.
[7] D.A. Di Pietro, A. Ern, A hybrid high-order locking-free method for linear elasticity on general meshes, Comput. Methods Appl. Mech. Engrg. 283 (2015) 1–21, URL https://www.sciencedirect.com/science/article/pii/S0045782514003181.