Applied Numerical Mathematics 118 (2017) 238–248

Contents lists available at ScienceDirect

Applied Numerical Mathematics

www.elsevier.com/locate/apnum

Analysis of a group finite element formulation

Gabriel R. Barrenechea [a], Petr Knobloch [b,*]

[a] Department of Mathematics and Statistics, University of Strathclyde, 26 Richmond Street, Glasgow G1 1XH, Scotland, United Kingdom
[b] Department of Numerical Mathematics, Faculty of Mathematics and Physics, Charles University, Sokolovská 83, 18675 Praha 8, Czech Republic

ARTICLE INFO

Article history:
Received 30 December 2016
Received in revised form 9 March 2017
Accepted 17 March 2017
Available online 22 March 2017

Keywords:
Group finite element formulation
Existence of solutions
Stability
Error estimates
Convection–diffusion–reaction equation

ABSTRACT

The group finite element formulation is a strategy aimed at speeding the assembly of finite element matrices for time-dependent problems. This process modifies the Galerkin matrix of the problem in a non-consistent way. This may cause a deterioration of both the stability and convergence of the method. In this paper we prove results for a group finite element formulation of a convection–diffusion–reaction equation showing that the stability of the original discrete problem remains unchanged under appropriate conditions on the data of the problem and on the discretization parameters. A violation of these conditions may lead to non-existence of solutions, as one of our main results shows. An analysis of the consistency error introduced by the group finite element formulation and its skew-symmetric variant is given.

1. Introduction

The numerical solution of convection-dominated transient problems is a topic that has received much attention over the last couple of decades. If the interest is to produce discretizations that preserve properties such as positivity, then the family of flux-corrected transport (FCT) schemes [3,16,13,14,12,10,8] has been actively used over the past years. These methods are related to the shock-capturing idea, and thus are nonlinear, but the main advantage is that they have provided some of the best results to date (see, e.g., [6,7] for computational surveys).

When dealing with the numerical solution of the transient transport problem (or any time-dependent problem with time-varying coefficients) by means of the finite element method (FEM), a very costly part of the computations is the assembly of the finite element matrix at every time step. This is due to the possible time dependence of the convective field. Then, in order to make the implementation more efficient, the group finite element formulation can be applied. This technique was introduced in [5,4] to simplify the implementation of nonlinear (convective) terms and to increase the efficiency of computations. Its main idea is to represent products (i.e., groups) of variables by single finite element functions. In this way, assembling the matrix corresponding to the convective term reduces to the multiplication of the nodal values of the convective field by a collection of matrix entries that are computed only once at the beginning of the computation. This formulation can be interpreted also as evaluating the convective term by using nodal quadrature. Over the years, the group finite element formulation has been frequently applied in the context of explicit piecewise linear finite element discretizations of compressible flow problems. However, the group formulation has been also used intensively in implicit FCT discretizations of conservation laws and transport and convection–diffusion problems with incompressible convective

* Corresponding author.
E-mail addresses: gabriel.barrenechea@strath.ac.uk (G.R. Barrenechea), knobloch@karlin.mff.cuni.cz (P. Knobloch).

http://dx.doi.org/10.1016/j.apnum.2017.03.008

fields (see, e.g., [10,11,2,9,6]), with very satisfactory numerical results. The main focus of this work is on implicit schemes for convection–diffusion–reaction equations with divergence-free convective fields.

There is, nevertheless, a lack of theoretical exploration on the limits of the group formulation. In particular, no results seem to be available on the impact that the lack of antisymmetry of the discrete convective term has in the formulation. One particular point that, in our opinion, deserves attention is the following. The FCT-like schemes can be reinterpreted as nonlinear stabilized finite element methods, where the stabilizing term is positive semidefinite and, in particular, may vanish for some meshes and discrete solutions. Consequently, the possible stability of the whole discretization relies on the stability of the group formulation of the underlying Galerkin scheme. Thus, the impact of the modification made by the group finite element method on the Galerkin scheme needs to be studied more in detail.

The purpose of this work is to fill the gap that was described in the last paragraph. To this end we consider the convection–diffusion–reaction equation as a model problem. Our main objective is to explore what is the impact of replacing the original convective term by its group formulation, both in terms of stability and lack of consistency. Concerning the stability of the method, the situation is as follows. For the steady-state case, the ellipticity of the approximate bilinear form can be proved by supposing that the convection is small enough or the mesh is sufficiently fine. For the time-dependent case, this requirement can be overcome by supposing, in turn, that the time step is small enough, which, in practice, reduces to imposing a CFL condition. On the other hand, if the assumptions that guarantee the stability are not fulfilled, the discrete problems based on the group formulation are not solvable in general, as we demonstrate by constructing a counterexample. We then move onto the analysis of the error introduced by the group finite element formulation. Our aim in this paper is not to perform a detailed error analysis of FCT schemes for time-dependent problems, and we will thus only present results estimating the consistency error induced by the group formulation. This will, in turn, give us an insight of what sort of convergence results can be expected for the considered schemes.

The plan of this work is as follows. In Section 2 we summarize the FCT methodology and we motivate then why we require the group formulation of the Galerkin part to be stable. Then, in Section 3 we present the problem of interest, namely the transient convection–diffusion–reaction equation, and the basic formulation of the group finite element strategy. The main result of that section is the aforementioned negative result in Theorem 3.1, where we show that, without further assumptions, the discrete problem may not have a solution. Next, in Section 4 we lie down conditions on the data, the mesh and the time step to make sure that the bilinear form associated to the discrete problem is elliptic and hence that the discrete problem is solvable. Moreover, we present an alternative skew-symmetric group formulation that is stable without any additional assumptions on the data and discretization parameters. Finally, in Section 5 we estimate the consistency errors caused by the two group formulations.

2. A flux-corrected transport scheme

Consider a linear initial-boundary value problem and let us discretize it in space by the finite element method. Then, at a time instant $t \in [0, T]$, the approximate solution can be represented by a vector $\mathrm{U}(t) \in \mathbb{R}^N$ of its coefficients with respect to a basis of the respective finite element space. Let us assume that the last $N - M$ components of $\mathrm{U}(t)$ $(0 < M < N)$ correspond to nodes where Dirichlet boundary conditions are prescribed whereas the first M components of $\mathrm{U}(t)$ are computed using the semidiscretization of the underlying partial differential equation. Then $\mathrm{U}(t) \equiv (u_1(t), \ldots, u_N(t))$ satisfies a system of linear ordinary differential equations equipped with boundary and initial conditions of the form

$$\mathbb{M}\,\frac{\mathrm{dU}}{\mathrm{d}t}(t) + \mathbb{A}(t)\,\mathrm{U}(t) = \mathrm{F}(t)\,, \qquad t \in (0, T]\,, \tag{2.1}$$

$$u_i(t) = u_i^b(t)\,, \qquad i = M+1, \ldots, N\,, \quad t \in (0, T]\,, \tag{2.2}$$

$$\mathrm{U}(0) = \mathrm{U}_0\,, \tag{2.3}$$

where $\mathbb{M} = (m_{ij})_{j=1,\ldots,N}^{i=1,\ldots,M}$ is the mass matrix and $\mathbb{A}(t) = (a_{ij}(t))_{j=1,\ldots,N}^{i=1,\ldots,M}$ is the stiffness matrix. It is assumed that the entries of the mass matrix are nonnegative. Introducing discrete time instants $0 = t_0 < t_1 < \cdots < t_K = T$ and approximating the time derivative by a difference formula, one obtains a discrete scheme for the approximations $U^n \in \mathbb{R}^N$ of $U(t_n)$. For example, the Crank–Nicolson method leads to

$$\mathbb{M}\,\frac{\mathrm{U}^n - \mathrm{U}^{n-1}}{\Delta t_n} + \frac{1}{2}\left(\mathbb{A}^n\,\mathrm{U}^n + \mathbb{A}^{n-1}\,\mathrm{U}^{n-1}\right) = \frac{1}{2}\left(\mathrm{F}^n + \mathrm{F}^{n-1}\right), \quad n = 1, \ldots, K\,, \tag{2.4}$$

$$u_i^n = u_i^b(t_n)\,, \qquad i = M+1, \ldots, N\,, \quad n = 1, \ldots, K\,, \tag{2.5}$$

$$\mathrm{U}^0 = \mathrm{U}_0\,, \tag{2.6}$$

where $\Delta t_n = t_n - t_{n-1}$, $\mathbb{A}^n = \mathbb{A}(t_n)$, and $\mathrm{F}^n = \mathrm{F}(t_n)$.

In this work we are mainly interested in solving convection-dominated problems. Then, if the semidiscrete equation (2.1) corresponds to a standard (conforming) finite element method, an additional stabilization has to be considered, see, e.g., [15]. One possibility is to apply a flux-corrected transport scheme, see, e.g., [12,10,8]. To formulate it, one first extends the matrices $\mathbb{A}^n$ to $(a_{ij}^n)_{i,j=1,\ldots,N}$. A common way is to use the stiffness matrices corresponding to the above-mentioned finite