25 **S**pringer Series in **C**omputational **M**athematics

Galerkin Finite Element Methods for Parabolic problems

抛物型问题的伽略金有限元方法

Vider Thomée

Springer-Verlag
世界图书出版公司

Vidar Thomée

Galerkin Finite Element Methods for Parabolic Problems

Springer

书　　名：Galerkin Finite Element Methods for Parabolic Problems

作　　者：V. Thomée

中 译 名：抛物型问题的伽略金有限元方法

出 版 者：世界图书出版公司北京公司

印 刷 者：北京世图印刷厂

发　　行：世界图书出版公司北京公司（北京朝内大街 137 号　100010）

联系电话：010-64015659, 64038347

电子信箱：kjsk@vip.sina.com

开　　本：24　　　印　张：13

出版年代：2003 年 6 月

书　　号：7-5062-5957-5/ O・376

版权登记：图字: 01-2003-3599

定　　价：48.00 元

Vidar Thomée
Department of Mathematics
Chalmers University of Technology
S-41296 Göteborg
Sweden
e-mail: thomee@math.chalmers.se

Cataloging-in-Publication Data applied for

Die Deutsche Bibliothek - CIP-Einheitsaufnahme
Thomée, Vidar:
Galerkin finite element methods for parabolic problems / Vidar Thomée. - Berlin ; Heidelberg ; New York ; Barcelona ; Budapest ; Hong Kong ; London ; Milan ; Paris ; Santa Clara ; Singapore ; Tokyo : Springer, 1997
(Springer series in computational mathematics ; 25)
ISBN 3-540-63236-0

Mathematics Subject Classification (1991): 65M60, 65M12, 65M15

ISSN 0179-3632
ISBN 3-540-63236-0 Springer-Verlag Berlin Heidelberg New York

Preface

My purpose in this monograph is to present an essentially self-contained account of the mathematical theory of Galerkin finite element methods as applied to parabolic partial differential equations. The emphases and selection of topics reflects my own involvement in the field over the past 25 years, and my ambition has been to stress ideas and methods of analysis rather than to describe the most general and farreaching results possible. Since the formulation and analysis of Galerkin finite element methods for parabolic problems are generally based on ideas and results from the corresponding theory for stationary elliptic problems, such material is often included in the presentation.

The basis of this work is my earlier text entitled *Galerkin Finite Element Methods for Parabolic Problems,* Springer Lecture Notes in Mathematics, No. 1054, from 1984. This has been out of print for several years, and I have felt a need and been encouraged by colleagues and friends to publish an updated version. In doing so I have included most of the contents of the 14 chapters of the earlier work in an updated and revised form, and added four new chapters, on semigroup methods, on multistep schemes, on incomplete iterative solution of the linear algebraic systems at the time levels, and on semilinear equations. The old chapters on fully discrete methods have been reworked by first treating the time discretization of an abstract differential equation in a Hilbert space setting, and the chapter on the discontinuous Galerkin method has been completely rewritten.

The following is an outline of the contents of the book:

In the introductory Chapter 1 we begin with a review of standard material on the finite element method for Dirichlet's problem for Poisson's equation in a bounded domain, and consider then the simplest Galerkin finite element methods for the corresponding initial-boundary value problem for the linear heat equation. The discrete methods are based on associated weak, or variational, formulations of the problems and employ first piecewise linear and then more general approximating functions which vanish on the boundary of the domain. For these model problems we demonstrate the basic error estimates in energy and mean square norms, in the parabolic case first for the semidiscrete problem resulting from discretization in the spatial variables only, and then also for the most commonly used fully discrete schemes obtained by discretization in both space and time, such as the backward Euler and Crank-Nicolson methods.

In the following five chapters we study several extensions and generalizations of the results obtained in the introduction in the case of the spatially semidiscrete approximation, and show error estimates in a variety of norms. First, in Chapter 2, we formulate the semidiscrete problem in terms of a more general approximate solution operator for the elliptic problem in a manner which does not require the approximating functions to satisfy the homogeneous boundary conditions. As an example of such a method we discuss a method of Nitsche based on a nonstandard weak formulation. In Chapter 3 more precise results are shown in the case of the homogeneous heat equation. These results are expressed in terms of certain function spaces $\dot{H}^s(\Omega)$ which are characterized by both smoothness and boundary behavior of its elements, and which will be used repeatedly in the rest of the book. We also demonstrate that the smoothing property for positive time of the solution operator of the initial value problem has an analogue in the semidiscrete situation, and use this to show that the finite element solution converges to full order even when the initial data are nonsmooth. The results of Chapters 2 and 3 are extended to more general linear parabolic equations in Chapter 4. Chapter 5 is devoted to the derivation of stability and error bounds with respect to the maximum-norm for our plane model problem, and in Chapter 6 negative norm error estimates of higher order are derived, together with related results concerning superconvergence.

In the next six chapters we consider fully discrete methods obtained by discretization in time of the spatially semidiscrete problem. First, in Chapter 7, we study the homogeneous heat equation and give analogues of our previous results both for smooth and for nonsmooth data. The methods used for time discretization are of one-step type and rely on rational approximations of the exponential, allowing the standard Euler and Crank-Nicolson procedures as special cases. Our approach here is to first discretize a parabolic equation in an abstract Hilbert space framework with respect to time, and then to apply the results obtained to the spatially semidiscrete problem. The analysis uses eigenfunction expansions related to the elliptic operator occurring in the parabolic equation, which we assume positive definite. In Chapter 8 we generalize the above abstract considerations to a Banach space setting and allow a more general parabolic equation, which we now analyze using the Dunford-Taylor spectral representation. The time discretization is interpreted as a rational approximation of the semigroup generated by the elliptic operator, i.e., the solution operator of the initial-value problem for the homogeneous equation. Application to maximum-norm estimates is discussed. In Chapter 9 we study fully discrete one-step methods for the inhomogeneous heat equation in which the forcing term is evaluated at a fixed finite number of points per time stepping interval. In Chapter 10 we apply Galerkin's method also for the time discretization and seek discrete solutions as piecewise polynomials in the time variable which may be discontinuous at the now not necessarily equidistant nodes. In this *discontinuous Galerkin* procedure the forcing term enters in integrated form rather than at a finite number of points. In Chapter 11 we consider multistep backward difference methods. We first study such methods with constant time steps of order at most 6, and show stability as well as smooth and nonsmooth data error estimates, and then discuss the second order backward difference method with variable time steps. In Chapter 12 we study the incomplete iterative solution of the finite dimensional linear systems of algebraic equations which need to be solved at each level of the time stepping procedure, and exemplify by the use of a V-cycle multigrid algorithm.

The next two chapters are devoted to nonlinear problems. In Chapter 13 we discuss the application of the standard Galerkin method to a model nonlinear parabolic equation. We show error estimates for the spatially semidiscrete problem as well as the fully discrete backward Euler and Crank-Nicolson methods, using piecewise linear finite elements, and then pay special attention to the formulation and analysis of time stepping procedures based on these, which are linear in the unknown functions. In Chapter 14 we derive various results in the case of semilinear equations, in particular concerning the extension of the analysis for nonsmooth initial data from the case of linear homogenous equations.

In the last four chapters we consider various modifications of the standard Galerkin finite element method. In Chapter 15 we analyze the so called lumped mass method for which in certain cases a maximum-principle is valid. In Chapter 16 we discuss the H^1 and H^{-1} methods. In the first of these, the Galerkin method is based on a weak formulation with respect to an inner product in H^1 and for the second, the method uses trial and test functions from different finite dimensional spaces. In Chapter 17, the approximation scheme is based on a mixed formulation of the initial boundary value problem in which the solution and its gradient are sought independently in different spaces. In the final Chapter 18 we consider a singular problem obtained by introducing polar coordinates in a spherically symmetric problem in a ball in $\mathbf{R}^3$ and discuss Galerkin methods based on two different weak formulations defined by two different inner products.

References to the literature where the reader may find more complete treatments of the different topics, and some historical comments, are given at the end of each chapter.

A desirable mathematical background for reading the text includes standard basic partial differential equations and functional analysis, including Sobolev spaces; for the convenience of the reader we often give references to the literature concerning such matters.

The work presented, first in the Lecture Notes and now in this monograph, has grown from courses, lecture series, summer-schools, and written material that I have been involved in over a long period of time. I wish to thank my students and colleagues in these various contexts for the inspiration and support they have provided, and for the help they have given me as discussion partners and critics. As regards this new version of my work I particularly address my thanks to Georgios Akrivis, Stig Larsson, and Per-Gunnar Martinsson, who have read the manuscript in various degrees of detail and are responsible for many improvements. I also want to express my special gratitude to Yumi Karlsson who typed a first version of the text from the old lecture notes, and to Gunnar Ekolin who generously furnished me with expert help with the intricacies of TeX.

Göteborg, July 1997

Vidar Thomée

Table of Contents

1. The Standard Galerkin Method

In this introductory chapter we shall study the standard Galerkin finite element method for the approximate solution of the model initial-boundary value problem for the heat equation,

$$(1.1)\qquad \begin{aligned} u_t - \Delta u &= f \quad \text{in } \Omega, \quad \text{for } t > 0,\\ u &= 0 \quad \text{on } \partial\Omega, \quad \text{for } t > 0, \quad \text{with } u(\cdot, 0) = v \quad \text{in } \Omega, \end{aligned}$$

where Ω is a domain in $\mathbf{R}^d$ with smooth boundary $\partial\Omega$, and where $u = u(x,t)$, u_t denotes $\partial u/\partial t$, and $\Delta = \sum_{j=1}^{d} \partial^2/\partial x_j^2$ the Laplacian.

Before we start to discuss this problem we shall briefly review the finite element method for the corresponding stationary problem, the Dirichlet problem for Poisson's equation,

$$(1.2)\qquad -\Delta u = f \quad \text{in } \Omega, \quad \text{with } u = 0 \quad \text{on } \partial\Omega.$$

Using a variational formulation of this problem, we shall define an approximation of the solution u of (1.2) as a function u_h which belongs to a finite-dimensional linear space S_h of functions of x with certain approximation properties. This function, in the simplest case a piecewise linear function on some partition of Ω, will be a solution of a finite system of linear algebraic equations. We show basic error estimates for this approximate solution in energy and least square norms.

Using a variational form of (1.1) we proceed to discretize the parabolic problem first in the spatial variable x, which results in an approximate solution $u_h(\cdot, t)$ in the finite element space S_h as a solution of a finite-dimensional system of ordinary differential equations. We then define a fully discrete scheme by discretizing this system in time by various finite difference approximations. This yields an approximate solution U of (1.1) which belongs to S_h at discrete time levels. Error estimates are derived for both the spatially and fully discrete solutions.

As a preparation for the definition of the finite element solution of (1.2), we consider briefly the approximation of smooth functions in Ω which vanish on $\partial\Omega$. For concreteness, we shall exemplify by piecewise linear functions in a convex plane domain.

Let thus Ω be a convex domain in the plane with smooth boundary $\partial\Omega$, and let $\mathcal{T}_h$ denote a partition of Ω into disjoint triangles τ such that no vertex of any triangle lies on the interior of a side of another triangle and such that the union of the triangles determine a polygonal domain $\Omega_h \subset \Omega$ with boundary vertices on $\partial\Omega$.

Let h denote the maximal length of the sides of the triangulation $\mathcal{T}_h$. Thus h is a parameter which decreases as the triangulation is made finer. We shall assume that the angles of the triangulations are bounded below by a positive constant, independently of h, and sometimes also that the triangulations are *quasiuniform* in the sense that the triangles of $\mathcal{T}_h$ are of essentially the same size, which we express by demanding that the area of $\tau \in \mathcal{T}_h$ is bounded below by ch^2, with $c > 0$, independent of h.

Let now S_h denote the continuous functions on the closure $\bar{\Omega}$ of Ω which are linear in each triangle of $\mathcal{T}_h$ and which vanish outside Ω_h. Let $\{P_j\}_{j=1}^{N_h}$ be the interior vertices of $\mathcal{T}_h$. A function in S_h is then uniquely determined by its values at the points P_j and thus depends on N_h parameters. Let Φ_j be the *pyramid function* in S_h which takes the value 1 at P_j but vanishes at the other vertices. Then $\{\Phi_j\}_{j=1}^{N_h}$ forms a *basis* for S_h, and every χ in S_h admits the representation

$$\chi(x) = \sum_{j=1}^{N_h} \alpha_j \Phi_j(x), \quad \text{with } \alpha_j = \chi(P_j).$$

A given smooth function v on Ω which vanishes on $\partial\Omega$ may now be approximated by, for instance, its interpolant $I_h v$ in S_h, which we define as the element of S_h which agrees with v at the interior vertices, i.e.,

$$I_h v(x) = \sum_{j=1}^{N_h} v(P_j)\Phi_j(x). \tag{1.3}$$

For a general $\Omega \subset \mathbf{R}^d$ we denote below by $\|\cdot\|$ the norm in $L_2 = L_2(\Omega)$ and by $\|\cdot\|_r$ that in the Sobolev space $H^r = H^r(\Omega) = W_2^r(\Omega)$, so that for real-valued functions v,

$$\|v\| = \|v\|_{L_2} = \Big(\int_\Omega v^2\,dx\Big)^{1/2},$$

and, for r a positive integer,

$$\|v\|_r = \|v\|_{H^r} = \Big(\sum_{|\alpha|\le r} \|D^\alpha v\|^2\Big)^{1/2}, \tag{1.4}$$

where, with $\alpha = (\alpha_1, \dots, \alpha_d)$, $D^\alpha = (\partial/\partial x_1)^{\alpha_1} \cdots (\partial/\partial x_d)^{\alpha_d}$ denotes an arbitrary derivative with respect to x of order $|\alpha| = \sum_{j=1}^d \alpha_j$, so that the sum in (1.4) contains all such derivatives of order at most r. We recall that for functions in $H_0^1 = H_0^1(\Omega)$, i.e., the functions v with $\nabla v = \operatorname{grad} v$ in L_2

and which vanish on $\partial\Omega$, $\|\nabla v\|$ and $\|v\|_1$ are equivalent norms (Friedrichs' lemma, see, e.g., [33] or [41]),

$$(1.5) \qquad c\|v\|_1 \le \|\nabla v\| \le \|v\|_1, \quad \forall v \in H_0^1, \quad \text{with } c > 0.$$

Throughout this book c and C will denote positive constants, not necessarily the same at different occurrances, which are independent of the parameters and functions involved.

Using this notation in our plane domain Ω, the following error estimates for the interpolant defined in (1.3) are well known (see, e.g., [33] or [41]), namely, for $v \in H^2 \cap H_0^1$,

$$(1.6) \qquad \|I_h v - v\| \le Ch^2\|v\|_2 \quad \text{and} \quad \|\nabla(I_h v - v)\| \le Ch\|v\|_2.$$

They may be derived by showing the corresponding estimate for each $\tau \in \mathcal{T}_h$ and then taking squares and adding. For an individual $\tau \in \mathcal{T}_h$ the proof is achieved by means of the Bramble- Hilbert lemma (cf. [33] or [41]), noting that $I_h v - v$ vanishes on τ for v linear.

We shall now return to the general case of a domain Ω in $\mathbf{R}^d$ and assume that we are given a family $\{S_h\}$ of finite-dimensional subspaces of H_0^1 such that, for some integer $r \ge 2$ and small h,

$$(1.7) \qquad \inf_{\chi \in S_h} \{\|v - \chi\| + h\|\nabla(v - \chi)\|\} \le Ch^s\|v\|_s, \quad \text{for } 1 \le s \le r,$$

when $v \in H^s \cap H_0^1$. The number r is referred to as the order of accuracy of the family $\{S_h\}$. The above example of piecewise linear functions in a plane domain corresponds to $d = r = 2$. In the case $r > 2, S_h$ often consists of piecewise polynomials of degree at most $r-1$ on a triangulation $\mathcal{T}_h$ as above. For instance, $r = 4$ in the case of piecewise cubic polynomial subspaces. Also, in the general situation estimates such as (1.7) may often be obtained by exhibiting an *interpolation operator* $I_h : H^r \cap H_0^1 \to S_h$ such that

$$(1.8) \qquad \|I_h v - v\| + h\|\nabla(I_h v - v)\| \le Ch^s\|v\|_s, \quad \text{for } 1 \le s \le r.$$

When $\partial\Omega$ is curved and $r > 2$ there are difficulties in the construction and analysis of such operators near the boundary, but the above situation may be accomplished, in principle, by mapping a curved triangle onto a straight-edged one (isoparametric elements). We shall not dwell on this.

We remark for later reference that if the family $\{S_h\}$ is based on a family of *quasiuniform* triangulations $\mathcal{T}_h$ and S_h consists of piecewise polynomials of degree at most $r-1$, then one has the *inverse inequality*

$$(1.9) \qquad \|\nabla\chi\| \le Ch^{-1}\|\chi\|, \quad \forall \chi \in S_h.$$

This inequality follows by taking squares and adding from the corresponding inequality for each triangle $\tau \in \mathcal{T}_h$, which in turn is obtained by a transformation to a fixed reference triangle, and using the fact that all norms on a finite dimensional space are equivalent, see, e.g., [41].

The optimal orders to which functions and their gradients may be approximated under our assumption (1.7), are $O(h^r)$ and $O(h^{r-1})$, respectively, and we shall now attempt to construct approximations of these orders for the solution of the Dirichlet problem (1.2). For this purpose we first write this problem in weak, or variational, form: We multiply the elliptic equation by a smooth function φ which vanishes on $\partial\Omega$ (it suffices to require $\varphi \in H_0^1$), integrate over Ω, and apply Green's formula on the left-hand side, to obtain

$$(1.10) \qquad (\nabla u, \nabla \varphi) = (f, \varphi), \quad \forall \varphi \in H_0^1,$$

where we have used the L_2 inner products,

$$(1.11) \qquad (v, w) = \int_\Omega vw\,dx, \quad (\nabla v, \nabla w) = \int_\Omega \sum_{j=1}^{d} \frac{\partial v}{\partial x_j}\frac{\partial w}{\partial x_j}\,dx.$$

In the finite element method we now pose the approximate problem of finding a function $u_h \in S_h$ such that

$$(1.12) \qquad (\nabla u_h, \nabla \chi) = (f, \chi), \quad \forall \chi \in S_h.$$

This way of defining an approximate solution in terms of the variational formulation of the problem is referred to as Galerkin's method, after the Russian applied mathematician Boris Grigorievich Galerkin (1871-1945).

Note that, as a result of (1.10) and (1.12),

$$(1.13) \qquad (\nabla(u_h - u), \nabla \chi) = 0, \quad \forall \chi \in S_h,$$

that is, the error in the discrete solution is orthogonal to S_h with respect to the Dirichlet inner product $(\nabla v, \nabla w)$.

In terms of a basis $\{\Phi_j\}_1^{N_h}$ for the finite element space S_h, our discrete problem may be stated: Find the coefficients α_j in $u_h(x) = \sum_{j=1}^{N_h} \alpha_j \Phi_j(x)$ such that

$$\sum_{j=1}^{N_h} \alpha_j (\nabla\Phi_j, \nabla\Phi_k) = (f, \Phi_k), \quad \text{for } k = 1, ..., N_h.$$

In matrix notation this may be expressed as

$$B\alpha = \tilde{f},$$

where $B = (b_{jk})$ is the stiffness matrix with elements $b_{jk} = (\nabla\Phi_j, \nabla\Phi_k)$, $\tilde{f} = (f_k)$ the vector with entries $f_k = (f, \Phi_k)$, and α the vector of unknowns α_j. The dimension of these arrays equals N_h, the dimension of S_h (which equals the number of interior vertices in our plane example above). The stiffness matrix B is a Gram matrix and thus in particular positive definite and invertible so that our discrete problem has a unique solution. To see that $B = (b_{jk})$ is positive definite, we note that

$$\sum_{j,k=1}^{d} b_{jk}\xi_j\xi_k = \|\nabla\Big(\sum_{j=1}^{d}\xi_j\Phi_j\Big)\|^2 \geq 0.$$

Here equality holds only if $\nabla(\sum_{j=1}^{d}\xi_j\Phi_j) \equiv 0$, so that $\sum_{j=1}^{d}\xi_j\Phi_j = 0$ by (1.5), and hence $\xi_j = 0,\ j = 1,\dots,N_h$.

When S_h consists of piecewise polynomial functions, the elements of the matrix B may be calculated exactly. However, unless f has a particularly simple form, the elements (f,Φ_j) of $\tilde{f}$ have to be computed by some quadrature formula.

We shall prove the following estimate for the error between the solutions of the discrete and continuous problems. Note that these estimates are of optimal order as defined by our assumption (1.7). Here, as will always be the case in the sequel, the statements of the inequalities assume that u is sufficently regular for the norms on the right to be finite.

Theorem 1.1. *Assume that (1.7) holds, and let u_h and u be the solutions of (1.12) and (1.2), respectively. Then, for $1 \leq s \leq r$,*

$$\|u_h - u\| \leq Ch^s\|u\|_s \quad \textit{and} \quad \|\nabla u_h - \nabla u\| \leq Ch^{s-1}\|u\|_s.$$

Proof. We start with the gradient. Using (1.13) we have for any $\chi \in S_h$

$$\begin{aligned}\|\nabla(u_h-u)\|^2 &= (\nabla(u_h-u),\nabla(u_h-u))\\ &= (\nabla(u_h-u),\nabla(\chi-u)) \leq \|\nabla(u_h-u)\|\,\|\nabla(\chi-u)\|,\end{aligned} \tag{1.14}$$

and hence by (1.7)

$$\|\nabla(u_h-u)\| \leq \inf_{\chi\in S_h}\|\nabla(\chi-u)\| \leq Ch^{s-1}\|u\|_s. \tag{1.15}$$

For the L_2−norm we proceed by a duality argument. Let φ be arbitrary in L_2, take $\psi \in H^2 \cap H_0^1$ as the solution of

$$-\Delta\psi = \varphi \quad \text{in } \Omega, \quad \text{with } \psi = 0 \quad \text{on } \partial\Omega, \tag{1.16}$$

and recall the fact that the solution ψ of (1.16) is smoother by two derivatives in L_2 than the right hand side φ, which may be expressed in terms of the *elliptic regularity* inequality

$$\|\psi\|_2 \leq C\|\Delta\psi\| = C\|\varphi\|. \tag{1.17}$$

For any $\psi_h \in S_h$ we have

$$\begin{aligned}(u_h-u,\varphi) &= -(u_h-u,\Delta\psi) = (\nabla(u_h-u),\nabla\psi)\\ &= (\nabla(u_h-u),\nabla(\psi-\psi_h)) \leq \|\nabla(u_h-u)\|\,\|\nabla(\psi-\psi_h)\|,\end{aligned} \tag{1.18}$$

and hence, using (1.15) together with (1.7) with $s = 2$ and (1.17),

$$(u_h-u,\varphi) \leq Ch^{s-1}\|u\|_s h\|\psi\|_2 \leq Ch^s\|u\|_s\|\varphi\|.$$

Choosing $\varphi = u_h - u$ completes the proof. □

Because of the variational formulation of the Galerkin method, the natural error estimates are expressed in L_2-based norms. Error analyses in other norms have also been pursued in the literature, and for later reference we quote here, without its more difficult proof, the following *maximum-norm error estimate* (see, e.g., [33]) for piecewise linear approximating functions in a plane domain Ω, where $L_\infty = L_\infty(\Omega)$ and $W^r_\infty = W^r_\infty(\Omega)$, and

$$\|v\|_{L_\infty} = \sup_{x\in\Omega} |u(x)|, \quad \|v\|_{W^r_\infty} = \max_{|\alpha|\le r} \|D^\alpha v\|_{L_\infty}.$$

We note first that the error in the interpolant introduced above is second order also in maximum-norm, so that (cf. (1.6)

$$\|I_h v - v\|_{L_\infty} \le Ch^2\|u\|_{W^2_\infty}, \quad \text{for } v \in W^2_\infty \cap H^1_0. \tag{1.19}$$

Compared to this estimate the following one for the elliptic problem contains a logarithmic factor; such a factor does not in an essentiall way influence the order of the error.

Theorem 1.2. *Let $\Omega \subset \mathbf{R}^2$ and assume that S_h consists of piecewise linear finite element functions, and that the family $\mathcal{T}_h$ is quasiuniform. Let u_h and u be the solutions of (1.12) and (1.2), respectively. Then*

$$\|u_h - u\|_{L_\infty} \le Ch^2\ell_h\|u\|_{W^2_\infty}, \quad \textit{where } \ell_h = \log\frac{1}{h}. \tag{1.20}$$

After these preparations we now turn to the initial-boundary value problem (1.1) for the heat equation. As indicated above it is convenient to proceed in two steps with the definition and analysis of the approximate solution. In the first step we shall approximate $u(x,t)$ by means of a function $u_h(x,t)$ which, for each fixed t, belongs to a finite-dimensional linear space S_h of functions of x of the type considered above. This function will be a solution of an h-dependent finite system of ordinary differential equations in time and is referred to as a *spatially discrete* , or *semidiscrete,* solution. As in the elliptic case just considered the spatially discrete problem is based on a variational formulation of (1.1). We shall then proceed to discretize this system in the time variable to produce a *fully discrete time stepping* scheme for the approximate solution of (1.1). To begin with this discretization in time will be accomplished by a finite difference approximation of the time derivative.

For the preliminary step of defining a spatially semidiscrete approximate solution to our initial boundary value problem, we thus write the problem in weak form: We multiply the heat equation by a smooth function φ which vanishes on $\partial\Omega$ (or $\varphi \in H^1_0$), integrate over Ω, and apply Green's formula to the second term, to obtain, with (v,w) and $(\nabla v, \nabla w)$ as in (1.11),

$$(u_t, \varphi) + (\nabla u, \nabla\varphi) = (f, \varphi), \quad \forall \varphi \in H^1_0,\ t > 0.$$

We may then pose the approximate problem to find $u_h(t) = u_h(\cdot, t)$, belonging to S_h for each t, such that

(1.21) $(u_{h,t},\chi)+(\nabla u_h,\nabla\chi)=(f,\chi),\quad \forall\,\chi\in S_h,\ t>0,\quad u_h(0)=v_h,$

where v_h is some approximation of v in S_h.

In terms of the basis $\{\Phi_j\}_1^{N_h}$ for S_h, our semidiscrete problem may be stated: Find the coefficients $\alpha_j(t)$ in $u_h(x,t)=\sum_{j=1}^{N_h}\alpha_j(t)\Phi_j(x)$ such that

$$\sum_{j=1}^{N_h}\alpha_j'(t)(\Phi_j,\Phi_k)+\sum_{j=1}^{N_h}\alpha_j(t)(\nabla\Phi_j,\nabla\Phi_k)=(f,\Phi_k),\quad k=1,\dots,N_h,$$

and, with γ_j the components of the given initial approximation v_h, $\alpha_j(0)=\gamma_j$ for $1,\dots,N_h$. In matrix notation this may be expressed as

$$A\alpha'(t)+B\alpha(t)=\tilde{f}(t),\quad \text{for } t>0,\quad \text{with } \alpha(0)=\gamma,$$

where $A=(a_{jk})$ is the mass matrix with elements $a_{jk}=(\Phi_j,\Phi_k)$, $B=(b_{jk})$ the stiffness matrix with $b_{jk}=(\nabla\Phi_j,\nabla\Phi_k)$, $\tilde{f}=(f_k)$ the vector with entries $f_k=(f,\Phi_k)$, $\alpha(t)$ the vector of unknowns $\alpha_j(t)$, and $\gamma=(\gamma_k)$. The dimension of all these items equals N_h, the dimension of S_h.

Since like the stiffness matrix B the mass matrix A is a Gram matrix, and thus in particular positive definite and invertible, the above system of ordinary differential equations may be written

$$\alpha'(t)+A^{-1}B\alpha(t)=A^{-1}\tilde{f}(t),\quad \text{for } t>0,\quad \text{with } \alpha(0)=\gamma,$$

and hence obviously has a unique solution for t positive.

Our first aim is to prove the following estimate in L_2 for the error between the solutions of the semidiscrete and continuous problems.

Theorem 1.3. *Let u_h and u be the solutions of (1.21) and (1.1). Then*

$$\|u_h(t)-u(t)\|\le\|v_h-v\|+Ch^r\Big(\|v\|_r+\int_0^t\|u_t\|_r\,ds\Big),\quad \textit{for } t\ge 0.$$

Here as earlier we require that the solution of the continuous problem has the regularity implicitly assumed by the presence of the norms on the right, and that v vanishes on $\partial\Omega$. Note also that if (1.8) holds and $v_h=I_hv$, then the first term on the right is dominated by the second. The same holds true if $v_h=P_hv$, where P_h denotes the orthogonal projection of v onto S_h with respect to the inner product in L_2, since this choice is the best approximation of v in S_h with respect to the L_2 norm. Another such optimal order choice for v_h is the projection to be defined next.

For the purpose of the proof of Theorem 1.3 we introduce the so called *elliptic* or *Ritz projection* R_h onto S_h as the orthogonal projection with respect to the inner product $(\nabla v,\nabla w)$, so that

(1.22) $(\nabla R_hv,\nabla\chi)=(\nabla v,\nabla\chi),\quad \forall\,\chi\in S_h,\quad \text{for } v\in H_0^1.$