

THE FINITE DIFFERENCE VERSUS THE FINITE ELEMENT METHOD FOR THE SOLUTION OF BOUNDARY VALUE PROBLEMS

VIDAR THOMÉE

In this lecture we describe, discuss and compare the two classes of methods most commonly used for the numerical solution of boundary value problems for partial differential equations, namely, the finite difference method and the finite element method. For both of these methods an extensive development of mathematical error analysis has taken place but individual numerical analysts often express strong prejudices in favor of one of them. Our purpose is to try to convey our conviction that this attitude is both historically unjustified and inhibiting, and that familiarity with both methods provides a wider range of techniques for constructing and analyzing discretization schemes.

We shall begin by describing and discussing the two methods for a simple model problem, the Dirichlet problem for Poisson's equation in the plane and then make some historical remarks. After this we shall present some examples of recent work on discretization of boundary value problems in which a unified finite difference finite element point of view is helpful. In the first such example we consider two different finite

Received 9 November 1983. This paper is based on an invited lecture given at the 1983 Australian Mathematical Society Annual Meeting held at the University of Queensland in May 1983.

Copyright Clearance Centre, Inc. Serial-fee code: 0004-9727/84 \$A2.00 + 0.00.

difference schemes for the solution of a singular two-point boundary value problem arising from a spherically symmetric Dirichlet problem in three dimensions. We then turn to a parabolic model problem and demonstrate how the lumped mass finite element method may be used to construct an explicit finite difference type scheme adjusted to the geometry of the domain. In connection with the parabolic problem, we also comment on the concept of superconvergence, and close by a remark concerning the discretization in time.

Consider thus the Dirichlet problem

$$(1) \quad \begin{aligned} -\Delta u &= f \quad \text{in } \Omega, \\ u &= 0 \quad \text{on } \partial\Omega, \end{aligned}$$

where Ω is a bounded domain in \mathbb{R}^2 . In the finite difference method one covers Ω with a mesh $\{jh, j \in \mathbb{Z}^2\}$, where h is a small positive number, and determines the values of an approximate solution U_j at the finite number of mesh-points jh in Ω from a linear system of equations. This system is obtained by replacing derivatives in (1) by finite difference quotients,

$$(2) \quad -\Delta_h U_j = -\frac{U_{j+e_1} - 2U_j + U_{j-e_1}}{h^2} - \frac{U_{j+e_2} - 2U_j + U_{j-e_2}}{h^2} = f_j,$$

at interior mesh-points, that is, such mesh-points for which the neighbors $(j \pm e_k)h$ belong to $\bar{\Omega}$. If all mesh-points of Ω are interior in this sense, the equations (2), together with the vanishing of U_j for boundary mesh-points will produce a linear system of the form $AU = F$, with A symmetric positive definite, and using a discrete maximum principle one may show that

$$(3) \quad \sup_{jh \in \Omega} |U_j - u(jh)| \leq Ch^2 \|u\|_{W_\infty^2(\Omega)}.$$

The second order accuracy is a consequence of the fact that the difference operators in (2) approximate the corresponding derivatives to this order, or, for smooth functions,

$$\left| \frac{v(x+he_j) - 2v(x) - v(x-he_j)}{h^2} - \frac{\partial^2 v}{\partial x_j^2}(x) \right| \leq \frac{1}{12} h^2 \sup_{\Omega} \left| \frac{\partial^4 v}{\partial x_j^4} \right|.$$

For general smooth domains some mesh-points near the boundary are not interior mesh-points, and other equations have to be set up for these points, for instance by means of interpolation. This may, in fact, be accomplished in such a way that the error estimate (3) remains valid, but the proof will be somewhat more complicated and the symmetry of the matrix A is lost. Higher order of accuracy may be attained by more complicated choices of the finite difference approximations of the derivatives at the expense of more complicated boundary approximations, matrices A , and convergence proofs.

We now consider the simplest form of the finite element method for the same problem (1). Let us assume, for simplicity, that Ω is a convex domain with smooth boundary. We then partition 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 Ω_h whose boundary vertices lie on $\partial\Omega$. The maximal diameter h of the triangulation will be a measure of the fineness of the triangulation, which we shall assume to be quasiuniform in the sense that each triangle contains a disk with radius ch with c bounded below independently of h . Let S_h denote the continuous functions on $\bar{\Omega}$ which are linear in each triangle and vanish outside Ω_h . Such a function χ is uniquely determined by its values at the interior vertices $\{P_j\}_1^{N_h}$ of the triangulation and may be expressed as

$$\chi(x) = \sum_{j=1}^{N_h} \chi(P_j) \phi_j(x),$$

where the "pyramid" functions $\{\phi_j\}_1^{N_h}$, defined by $\phi_j(P_k) = \delta_{jk}$ form a basis for S_h . Letting $I_h v$ denote the obvious interpolant of the smooth function v vanishing on $\partial\Omega$, we have, for instance,

$$(4) \quad \|I_h v - v\|_{L_2(\Omega)} + h \|\nabla I_h v - \nabla v\|_{L_2(\Omega)} \leq Ch^2 \|v\|_{H^2(\Omega)},$$

and

$$\|I_h v - v\|_{L_\infty(\Omega)} \leq Ch^2 \|v\|_{W_\infty^2(\Omega)} .$$

In the finite element method one first writes the Dirichlet problem in weak form as

$$(\nabla u, \nabla \phi) = (f, \phi) \quad \text{for all } \phi \in H_0^1(\Omega) ,$$

where (\cdot, \cdot) denote the standard inner products in $L_2(\Omega)$, and $H_0^1(\Omega)$ the functions with first derivatives in $L_2(\Omega)$ and vanishing on $\partial\Omega$. One may then pose the discrete problem to find $u_h \in S_h$ such that

$$(5) \quad (\nabla u_h, \nabla \chi) = (f, \chi) \quad \text{for all } \chi \in S_h .$$

It is easy to see that this problem admits a unique solution for each f in $L_2(\Omega)$. This solution may also be defined as the function in S_h which minimizes the functional

$$(6) \quad J(\chi) = \|\nabla \chi\|^2 - 2(f, \chi) = \|\nabla(\chi - u)\|^2 - \|\nabla u\|^2 .$$

We note that this method may be written in matrix form as $A\alpha = F$ where A has elements $(\nabla\phi_j, \nabla\phi_k)$ and thus is symmetric positive definite, and where α is the vector of nodal values of u_h .

The standard L_2 error estimates for this method are

$$(7) \quad \|u_h - u\|_{L_2(\Omega)} + h \|\nabla u_h - \nabla u\|_{L_2(\Omega)} \leq Ch^2 \|u\|_{H^2(\Omega)} ,$$

and it has been shown that

$$(8) \quad \|u_h - u\|_{L_\infty(\Omega)} \leq Ch^2 \log \frac{1}{h} \|u\|_{W_\infty^2(\Omega)} .$$

The L_2 estimates are of the same order in h as the best approximation possible (cf. (4)) and so is the maximum norm estimate, modulo the factor $\log 1/h$. In order to increase the accuracy one may employ functions which reduce to polynomials of higher degree than first in the triangles. Thus, for instance, with piecewise quadratics and cubics one would expect $O(h^3)$

and $O(h^4)$ error estimates, respectively. This will meet with some difficulties in the case of a curved boundary, because of the discrepancy between Ω_h and Ω , but different ways have been devised to overcome these.

One way of performing a triangulation of Ω is to start with the three families of straight lines $x_1 = jh$, $x_2 = jh$, $x_1 + x_2 = jh$, $j \in Z$. The triangles thus formed may be used in the interior of Ω and then supplemented by other triangles near $\partial\Omega$ to form the desired triangulation. If this is done, the equation corresponding to an interior vertex takes the form, with $U_j = u_h(jh)$ and ϕ_j the corresponding basis function

$$-\frac{U_{j+e_1} - 2U_j + U_{j-e_1}}{h^2} - \frac{U_{j+e_2} - 2U_j + U_{j-e_2}}{h^2} = h^{-2}(f, \phi_j).$$

We recognize this as essentially the same equation as our previous finite difference equation, but with the right hand side $f_j = f(jh)$ replaced by an average of f over a neighborhood of the point $x = jh$. In practice this average may have to be evaluated by a numerical quadrature rule, and taking the value of f at $x = jh$ is a possible such rule.

Thus the finite difference method consists in replacing derivatives by finite differences with some *ad hoc* modification near the boundary, whereas the finite element method uses a variational formulation in a way that automatically accommodates the boundary conditions. The error analysis for the finite difference method uses a local accuracy condition, usually based on Taylor expansion, and some stability property, in our example in the form of a discrete maximum principle. The finite element error analysis is in essence based on the variational formulation itself and is global in nature. We note that the finite element error estimates above require the solution to have two derivatives, whereas four derivatives are needed in the finite difference method. This advantage of finite elements stems from the use of averages and disappeared when a quadrature rule is used.

The difficulties in the construction of finite difference equations near the boundary are even greater for Neumann type boundary conditions, whereas in the variational or finite element approach these are natural

boundary conditions which do not have to be imposed on the approximating functions and are thus very simple to treat.

We pause now for some historical remarks.

The idea of using a variational formulation of a boundary value problem for its numerical solution goes back to Lord Rayleigh [31, 32] and Ritz [33]. In Ritz's approach the approximate solution was sought as a finite linear combination of functions in a sequence $\{\phi_j\}_1^\infty$ which in some sense spans the space of admissible functions. Such functions could be, for instance, polynomials or trigonometrical polynomials. The idea to use an orthogonality condition such as (5) rather than the minimization of $J(\chi)$ in (6) is attributed to Galerkin [19]. This latter approach is in some sense more general than the Rayleigh-Ritz procedure; its use for time-dependent problems is sometimes referred to as the Faedo-Galerkin method [15].

The defining feature of the finite element method may be considered to be the use of approximating functions which are piecewise polynomials, on a partitioning of the domain under consideration into small "elements", such as triangles or quadrilaterals, and with the property that the polynomial in an element is determined by local parameters. The approximating space then admits a basis of functions with small support and, as a consequence, the matrices occurring in the corresponding algebraic problems will be sparse, which facilitates the computational work. The use in this context of piecewise linear approximating functions based on triangulations adapted to the geometry of the domain was suggested by Courant in an address delivered to the American Mathematical Society in 1941 [8], but the further development and analysis of the method would occur much later.

Already in the fundamental 1928 paper by Courant, Friedrichs and Lewy [9], finite difference equations of the form considered above were analyzed and, in fact, deduced from a variational principle, although with the Dirichlet integral replaced by a finite sum over the points of a uniform mesh of the sum of squares of difference quotients. Convergence was shown by compactness arguments and no error estimates or convergence rates were given. Error estimates of the type described above in (3) were shown by Gerschgorin in 1930 [20]. For the following several decades the finite difference method continued to develop as a basic technique in numerical

work and is still by some considered to be the most effective, particularly for hyperbolic problems. The analysis has grown increasingly sophisticated mathematically, but has generally moved away from the variational approach.

The finite element method as it is known today originated in the engineering literature, where in the mid 1950s structural engineers had connected the well established framework analysis with variational methods in continuum mechanics into a discretization method in which a structure is thought of as divided into elements with locally defined strains or stresses. Some of the pioneering work was done by Turner, Clough, Martin and Topp [40] and Argyris [1] and the name of the finite element method appeared first in a 1960 paper by Clough [7]. The method was later applied to other classes of problems in continuum mechanics, see, for example, Zienkiewicz and Cheung [43]; a standard reference from the engineering literature is Zienkiewicz [42].

In the mid 1960s, a number of papers appeared independently in the mathematical literature, which were concerned with the construction and analysis of finite difference schemes by the variational principle. Thus Friedrichs and Keller [17] applied the Rayleigh-Ritz procedure with piecewise linear approximating functions to the Neumann problem and gave the natural error estimates in the energy norm, and Feng [16] pursued a similar program for a somewhat more general class of problems, including the equations of elasticity. Other related work was published by Céa [4], Demjanovič [11] and Oganessian [28]. The methods discussed were, in fact, special cases of the finite element method but were thought of by the respective authors as finite difference methods; they were referred to in the Russian literature as variational difference schemes.

In the period immediately following this, the theory established itself through contributions such as Birkhoff, Schultz and Varga [3], in which the theory of splines was brought to bear on the development, and Zlámal [44], with the first stringent error analysis of more complicated finite elements. The L_2 error estimate in (7) was proved by a now well-known duality argument independently by Aubin [2], Nitsche [26] and Oganessian and Ruchovets [29] and later maximum-norm error estimates such as (8) were shown by Scott [36], Natterer [25] and Nitsche [27]. The theory attracted the attention of many mathematicians including several previously working with finite differences and is now a very active field. We shall

refrain from describing the more recent history as the main purpose of our above remarks was to point out how early work on finite differences was related to variational principles and that the pioneers in the mathematical analysis of the finite element method thought in terms of finite differences. We refer to the books by Ciarlet [6] and Strang and Fix [37] for more references.

We shall now return to the mathematics and cite some more recent work.

Consider first the Dirichlet problem

$$\begin{aligned} -\Delta u + q(x)u &= f \quad \text{in } B, \\ u &= 0 \quad \text{on } \partial B, \end{aligned}$$

where $B = B_1(0)$ is the unit ball in R^3 and where q is non-negative and q and f are smooth and depend only on $|x|$. Then the solution also depends only on $|x|$ and introducing polar coordinates, one finds that $u = u(|x|)$ satisfies the singular two-point boundary value problem

$$\begin{aligned} (9) \quad -u'' + \frac{2}{x} u' + q(x)u &= f(x) \quad \text{for all } x \in (0, 1), \\ u'(0) &= u(1) = 0. \end{aligned}$$

The differential equation may be written in the form

$$-(x^2 u')' + x^2 q u = x^2 f,$$

and one finds hence that u solves the variational problem

$$A(u, \phi) = (x^2 u', \phi') + (x^2 q u, \phi) = (x^2 f, \phi) \quad \text{for all } \phi \in \dot{H}^1,$$

where \dot{H}^1 denotes the space of all $v \in C([0, 1])$ which vanish at $x = 1$ and for which $xv' \in L_2$. One may therefore seek an approximate solution in the finite dimensional space S_h of continuous functions which vanish at $x = 1$ and which are linear in each interval $I_j = ((j-1)h, jh)$, $j = 1, \dots, N$, where $h = 1/N$, and which satisfies

$$(10) \quad A(u_h, \chi) = (x^2 f, \chi) \quad \text{for all } \chi \in S_h.$$

Letting $\{\phi_j\}_0^{N-1}$ be the basis of S_h defined by $\phi_i(x_j) = \delta_{ij}$ and

$U_i = u_h(x_i)$ ($x_i = ih$) this problem reduces to the difference scheme

$$\begin{aligned} & \frac{h}{3} (U_0 - U_1) + \left(x^2 q \phi_0, \phi_0 \right) U_0 + \left(x^2 q \phi_1, \phi_0 \right) U_1 = \left(x^2 f, \phi_0 \right), \\ & -\frac{1}{h} \left(\frac{x^2}{x_{i+\frac{1}{2}}} \frac{U_{i+1} - U_i}{h} - \frac{x^2}{x_{i-\frac{1}{2}}} \frac{U_i - U_{i-1}}{h} \right) + \frac{1}{h} \sum_{k=i-1}^{i+1} \left(x^2 q \phi_k, \phi_i \right) U_k = \frac{1}{h} \left(x^2 f, \phi_i \right) \\ & \qquad \qquad \qquad \text{for } i = 1, \dots, N-1, \\ & U_N = 0, \end{aligned}$$

where $\frac{x^2}{x_{i+\frac{1}{2}}} = \frac{1}{12} \left(5x_i^2 + 2x_i x_{i+1} + 5x_{i+1}^2 \right)$.

Using a straightforward variational technique, Schreiber and Eisenstadt [35] have shown that

$$\|x(u_h - u)\|_{L_2} \leq Ch^2 \|xu''\|_{L_2},$$

and by a somewhat more refined analysis, Jespersen [22] was able to derive the uniform error bound

$$\|u_h - u\|_{L_\infty} \leq Ch^2 \log \frac{1}{h} \|u''\|_{L_\infty}.$$

Numerical experiments using the above procedure show a marked loss of accuracy near $x = 0$, which is not unexpected as the underlying norm contains the weight function x . In order to remedy this disadvantage we shall now consider an alternative variational formulation with a weaker weight near the origin (*cf.* Eriksson and Thomée [14]). For this purpose we first observe that the differential equation (9) may also be written

$$-(xu')' - u' + xqu = xf,$$

and that u thus satisfies

$$B(u, \phi) = (xu', \phi') - (u', \phi) + (xqu, \phi) = (xf, \phi) \text{ for all } \phi \in \dot{H}^1.$$

We therefore propose as an alternative to (10) the discrete method

$$(11) \quad B(u_h, \chi) = (xf, \chi) \text{ for all } \chi \in S_h.$$

We note that $B(\cdot, \cdot)$ is non-symmetric but for v vanishing at $x = 1$, we have

$$B(v, v) = (xv', v') + \frac{1}{2}v(0)^2 + (xqv, v) > 0 \text{ for } v \neq 0,$$

so that, in particular, the discrete problem (11) admits a unique solution. The natural norm associated with this formulation is $\|x^{\frac{1}{2}}v\|$ and one may demonstrate the uniform error estimate

$$\|u_h - u\|_{L_\infty} \leq Ch^2 \|u''\|_{L_\infty} .$$

Numerical experiments show a more uniform error distribution over the interval than for the above symmetric method.

This time the corresponding difference scheme is

$$U_0 - U_1 + (xq\phi_0, \phi_0)U_0 + (xq\phi_1, \phi_0)U_1 = (xf, \phi_0) ,$$

$$-x_i \frac{U_{i+1} - 2U_i + U_{i-1}}{h^2} - 2 \frac{U_{i+1} - U_{i-1}}{2h} + \frac{1}{h} \sum_{k=i-1}^{i+1} (xq\phi_k, \phi_i)U_k = \frac{1}{h} (xf, \phi_i) ,$$

$i = 1, \dots, N-1 ,$

$$U_N = 0 ,$$

or, using Simpson's rule for the local quadratures,

$$U_0 - U_1 + \frac{1}{3} x_{\frac{1}{2}} q_{\frac{1}{2}} \frac{U_0 + U_1}{2} = \frac{h}{3} x_{\frac{1}{2}} f_{\frac{1}{2}} ,$$

$$-x_i \frac{U_{i+1} - 2U_i + U_{i-1}}{h^2} - 2 \frac{U_{i+1} - U_{i-1}}{2h}$$

$$+ \frac{1}{6} (x_{i-\frac{1}{2}} q_{i-\frac{1}{2}} U_{i-1} + (x_{i-\frac{1}{2}} q_{i-\frac{1}{2}} + 2x_i q_i + x_{i+\frac{1}{2}} q_{i+\frac{1}{2}}) U_i + x_{i+\frac{1}{2}} q_{i+\frac{1}{2}} U_{i+1})$$

$$= \frac{1}{3} (x_{i-\frac{1}{2}} f_{i-\frac{1}{2}} + x_i f_i + x_{i+\frac{1}{2}} f_{i+\frac{1}{2}}) , \quad i = 1, \dots, N-1 ,$$

$$U_N = 0 ,$$

where $x_{i+\frac{1}{2}} = \frac{1}{2}(x_i + x_{i+1})$, $q_{i+\frac{1}{2}} = q(x_{i+\frac{1}{2}})$, and $f_{i+\frac{1}{2}} = f(x_{i+\frac{1}{2}})$.

We note again that both procedures are finite difference schemes and that they have been arrived at by a variational philosophy rather than by an *ad hoc* approximation of derivatives by finite difference quotients.

We shall briefly consider the application of the finite difference and the finite element methods to the initial boundary value problem for the simple heat equation,

$$(12) \quad \begin{aligned} u_t - \Delta u &= f \quad \text{in } \Omega \times (0, \infty), \\ u &= 0 \quad \text{on } \partial\Omega \times (0, \infty), \\ u(x, 0) &= v(x). \end{aligned}$$

In the finite difference approach to this problem we may replace the differential equation by

$$(13) \quad \frac{U_j^{n+1} - U_j^n}{k} - \frac{U_{j+e_1}^n - 2U_j^n + U_{j-e_1}^n}{h^2} - \frac{U_{j+e_2}^n - 2U_j^n + U_{j-e_2}^n}{h^2} = f_j^n,$$

where U_j^n is the approximation at (jh, nk) with h and k the mesh sizes in space and time. This equation may be written in the explicit form

$$U_j^{n+1} = \lambda \left(U_{j+e_1}^n + U_{j-e_1}^n + U_{j+e_2}^n + U_{j-e_2}^n \right) + (1-4\lambda)U_j^n + kf_j^n,$$

where $\lambda = k/h^2$. Combined with discrete initial and boundary conditions this defines a method for finding an approximate solution of our problem.

If $k \leq \frac{1}{4}h^2$ all the above coefficients are nonnegative and, assuming the boundary conditions are well behaved, maximum norm stability holds, so that in the case of the homogeneous equation ($f = 0$),

$$\max_j \left| U_j^{n+1} \right| \leq \max_j \left| U_j^n \right| \leq \dots \leq \max_j \left| U_j^0 \right|.$$

In practice, the requirement $k \leq \frac{1}{4}h^2$ often presents too severe a restriction for the time step and one is lead to using instead of a forward difference quotient in (13) a backward difference quotient, which leads to the equation

$$(1+4\lambda)U_j^{n+1} - \lambda \left(U_{j+e_1}^{n+1} + U_{j-e_1}^{n+1} + U_{j+e_2}^{n+1} + U_{j-e_2}^{n+1} \right) = U_j^n + kf_j^{n+1}.$$

In contrast to our previous explicit scheme, this method is implicit in that a linear system of equations has to be solved at each time level, but no restriction on the mesh-ratio λ is required for stability. In both cases, in the same way as for the elliptic problem, complications may arise in the construction of the difference equations near the boundary and the symmetry of the matrices may be lost.

Taylor expansion around the point (jh, nk) shows that a smooth solution of (12) satisfies the equation (13) up to an error of order $O(k+h^2)$ and similarly for the backward Euler equation. Together with the stability of the method this will show a global error estimate of the same order. Since these schemes are only first order in k , one often prefers the Crank-Nicolson scheme

$$\frac{U_j^{n+1} - U_j^n}{k} - \frac{1}{2} \sum_{l=1}^2 \left(\frac{U_{j+e_l}^{n+1} - 2U_j^{n+1} + U_{j-e_l}^{n+1}}{h^2} + \frac{U_{j+e_l}^n - 2U_j^n + U_{j-e_l}^n}{h^2} \right) = f_j^{n+\frac{1}{2}},$$

which is symmetric around $(jh, (n+\frac{1}{2})k)$ and has local and global errors of order $O(k^2+h^2)$, without any mesh-ratio restrictions for stability. We refer to Richtmyer's book [33] for more details about the finite difference method.

In the Galerkin finite element method for the same problem one first writes the problem in weak form

$$(u_t, \phi) + (\nabla u, \nabla \phi) = (f, \phi) \quad \text{for } \phi \in H_0^1(\Omega), \quad t \geq 0,$$

and then reduces the problem to a system of ordinary differential equations by setting $u_h(t) \in S_h$, with S_h as for the elliptic problem above, and

$$(u_{h,t}, \chi) + (\nabla u_h, \nabla \chi) = (f, \phi) \quad \text{for } \chi \in S_h, \quad t \geq 0,$$

$$u_h(0) = v_h \approx v.$$

This may then be followed by a discretization in time, for instance by forward or backward differencing,

$$(14) \quad \left(\frac{U^{n+1} - U^n}{k}, \chi \right) + (\nabla U^n, \nabla \chi) = (f^n, \chi) \quad \text{for } \chi \in S_h,$$

$$\left(\frac{U^{n+1} - U^n}{k}, \chi \right) + (\nabla U^{n+1}, \nabla \chi) = (f^{n+1}, \chi) \quad \text{for } \chi \in S_h,$$

or by the Crank-Nicolson method

$$\left(\frac{U^{n+1} - U^n}{k}, \chi \right) + \left(\nabla \frac{U^{n+1} + U^n}{2}, \nabla \chi \right) = (f^{n+\frac{1}{2}}, \chi) \quad \text{for all } \chi \in S_h.$$

In matrix notation the semidiscrete equation may be written

$$(15) \quad B\alpha'(t) + A\alpha(t) = F(t) \quad \text{for } t \geq 0, \quad \text{with } \alpha(0) = \gamma \text{ given,}$$

where B is the mass matrix with elements $b_{jk} = (\phi_j, \phi_k)$, $A = (a_{jk})$ the stiffness matrix with $a_{jk} = (\nabla\phi_j, \nabla\phi_k)$, $F = (f_k)$ with $f_k = (f, \phi_k)$ and $\alpha(t)$ vector of nodal values for the approximate solution. The completely discrete schemes may be correspondingly represented, as for instance for the forward Euler method,

$$(16) \quad B \frac{\alpha^{n+1} - \alpha^n}{k} + A\alpha^n = F^n.$$

We note that since B is not diagonal, this is not a completely explicit method as was the case for the forward Euler finite difference method.

The most natural analysis of these finite element methods take place in L_2 based spaces and in the L_2 norm one may show, for instance, for the homogeneous semidiscrete problem with suitable choice of discrete initial data,

$$\|u_h(t) - u(t)\|_{L_2(\Omega)} \leq Ch^2 \|\Delta v\|_{L_2(\Omega)},$$

and for the backward Euler method with an additional term $ck \|\Delta v\|_{L_2(\Omega)}$.

Maximum-norm error estimates are also available, but as before their proofs now require more refined tools.

We note that in the same way as for the elliptic problem, in the case of a triangulation which is uniform in the interior of Ω , these equations may be interpreted as finite difference equations, but even for non-uniform triangulations, the error stemming from the discretization in space is $O(h^2)$. For a survey of estimates such as the ones just quoted, see Thomée [39].

As remarked above, the most natural finite element discretization of the heat equation leads to the semidiscrete problem (15) where B is positive definite but not diagonal. We shall now briefly consider a modification of this method, the lumped mass method. This method may be most easily defined by simply replacing the matrix B in (15) by the diagonal matrix \bar{B} having for its diagonal elements the row sums of B . The homogeneous semidiscrete problem then takes the form

$$(17) \quad \bar{B}\alpha'(t) + A\alpha(t) = 0 ,$$

in which \bar{B} may trivially be inverted.

Another way of defining this method is as follows (cf. Raviart [30]): for each triangle τ in the triangulation (with vertices $\{P_{\tau,j}\}_1^3$) let us introduce the quadrature formula

$$Q_{\tau}(f) = \frac{1}{3} \text{area } \tau \sum_{j=1}^3 f(P_{\tau,j}) \approx \int_{\tau} f dx ,$$

and set

$$(v, w)_h = \sum_{\tau} Q_{\tau}(vw) ,$$

where the summation extends over the triangulation. Then the equation (17) is equivalent to

$$(18) \quad (u_{h,t}, \chi)_h + (\nabla u_h, \nabla \chi) = 0 \text{ for all } \chi \in S_h .$$

This modification may be applied as well to the totally discrete schemes discussed above. For instance, the forward Euler method (14) or (16) now corresponds to

$$(19) \quad \left(\frac{U^{n+1} - U^n}{k}, \chi \right)_h + (\nabla U^n, \nabla \chi) = 0 \text{ for all } \chi \in S_h ,$$

or in matrix form

$$\bar{B}\alpha^{n+1} = (\bar{B} + kA)\alpha^n .$$

In this latter case, as \bar{B} is diagonal, we may think of (19) as a purely explicit method, a generalization to non-uniform meshes of the forward Euler finite difference scheme. As earlier the stability of this method requires the mesh-ratio k/h^2 to be sufficiently small.

It may be shown that the error in the lumped mass method is again of second order. More precisely, if the initial data are chosen appropriately, we have for the semidiscrete problem

$$\|u_h(t) - u(t)\|_{L_2(\Omega)} \leq C h^2 \|(-\Delta)^{3/2} v\|_{L_2(\Omega)} .$$

For the forward and backward Euler methods an additional term

$Ck \|(-\Delta)^{3/2} v\|_{L_2(\Omega)}$ appears and in the case of the Crank-Nicolson method,

$Ck^2 \|(-\Delta)^{5/2} v\|_{L_2(\Omega)}$. Here the fractional powers of $-\Delta$ are defined by spectral representation.

An interesting feature of the lumped mass method is that, in contrast to the situation for the standard Galerkin method, an exact maximum principle holds, under the specific assumption that the triangulation contains no obtuse triangles, so that (*cf.* Fujii [18])

$$\|u_h(t)\|_{L_\infty(\Omega)} \leq \|v\|_{L_\infty(\Omega)},$$

and similarly for the discretizations in time. This maximum norm stability may be applied, as in the finite difference theory, together with a local approximation property to yield maximum-norm convergence of the method (*cf.* Ushijima [4]). However, as the local order of approximation will generally only be $O(h)$, the same will hold for the global error bound. Using energy methods suggested by the Galerkin formulation (18), it is shown in Chen and Thomée [5], without the use of the maximum principle and the non-obtuseness of the triangles, that

$$\|u_h(t) - u(t)\|_{L_\infty} = O\left(h^2 \left(\log \frac{1}{h}\right)^{\frac{1}{2}}\right),$$

and correspondingly for the completely discrete variants.

We may thus think of this example as a way of constructing a finite difference scheme for a non-regular mesh, and where the finite element mode of thinking produces a better error bound than the straight-forward finite difference approach.

The theory of finite difference schemes has obviously influenced the analysis of the finite element method and many results for the latter may be seen as generalizations of earlier results for finite differences. We shall mention one particular such instance when finite difference analysis was instrumental in the development. This concerns a type of result which has later become characterized by the term superconvergence (*cf.* Douglas and Dupont [12]). Consider thus the pure initial-value problem for the heat equation

$$u_t = u_{xx} \quad \text{for } -\infty < x < \infty, \quad t \geq 0,$$

and let u_h denote the approximate solution by Galerkin's method in the set S_h of smooth splines of order r (that is, of C^{r-2} piecewise polynomials of degree $r-1$), based on a uniform mesh with subinterval length h . Then the best order of global approximation is $O(h^r)$ and such a bound may also be shown for the global error $u_h - u$. However, interpreting the discrete equation as a finite difference equation and using established knowledge for these, it was shown in Thomée [38] that at the mesh-points, the error is indeed of higher order, or

$$u_h(jh, t) - u(jh, t) = O(h^{2r-2}) \quad \text{as } h \rightarrow 0.$$

For instance, for $r = 4$, which corresponds to piecewise cubics, the global error is $O(h^4)$ and the error at the mesh-points is $O(h^6)$.

In most work on finite element methods for time dependent problems the discretization in time has been accomplished by a finite difference approximation, such as the forward or backward Euler or the Crank-Nicolson method. It is natural to consider also the use of the Galerkin method in this context and results in this direction have been presented in Jamet [21], Lesaint and Raviart [24] and Johnson, Nävert and Pitkäranta [23]. We shall quote briefly some work in preparation for parabolic problems by Eriksson, Johnson and Thomée [13].

Consider first for simplicity an initial-value problem for a system of ordinary differential equations of the form

$$(20) \quad u' + Au = f \quad \text{for } t \geq 0, \quad u(0) = v,$$

where A is a constant matrix. Letting $0 = t_0 < t_1 < \dots$ be a partition of the positive real axis we shall approximate the solution of (20) by a function U which reduces to a polynomial of degree less than or equal to $q-1$ ($U \in \Pi_{q-1}$) on each of the intervals (t_n, t_{n+1}) and which is determined from the equations

$$\int_{t_n}^{t_{n+1}} (U' + AU, \psi) dt + (U_+^n, \psi_+^n) = \int_{t_n}^{t_{n+1}} (f, v) dt + (U_-^n, \psi_+^n)$$

for all $\psi \in \Pi_{q-1}$,

$$U_-^0 = v,$$

where U_{\pm}^n denote the limits from above and below at t_n . (Note that continuity across the points t_n is not required from U .) For the case of piecewise constant approximating functions ($q = 1$) this procedure reduces to a backward Euler method and for $q > 1$ it is associated with a higher order subdiagonal Padé approximation of the exponential. With $k = \max(t_{n+1} - t_n)$ one may show the expected error estimate

$$\sup_t |U(t) - u(t)| \leq Ck^q,$$

but also a "superconvergent" order error estimate at the partition points,

$$\sup_n |U_-^n - u(t_n)| \leq Ck^{2q-1}.$$

Using piecewise linear approximating functions, for instance, the global error estimate is $O(k^2)$ and the error at the nodes is $O(k^3)$.

This procedure, which was analyzed for ordinary differential equations by Delfour, Hager and Trochu [10] may now be applied to systems such as (15) obtained from discretization of parabolic equations. One may then show, for instance, in the case of the homogeneous equation, with suitable choice of semidiscrete initial data, that the error between the completely discrete and the semidiscrete solutions is bounded as

$$\|U_-^n - u_h(t_n)\|_{L_2(\Omega)} \leq Ck^{2q-1} \|\Delta^{2q-1} v\|_{L_2(\Omega)}.$$

Such estimates may be combined with error bounds for the discretization in space to yield optimal order error estimates for the totally discrete scheme. Note that in the application to a nonhomogeneous equation this method uses averages in time of f rather than point-values as was the case in the finite difference approach.

This concludes our series of examples of work on the border line between the finite difference and finite element methods. In summary, our examples propose to demonstrate that often an efficient way of generating a (generalized) finite difference scheme for a boundary value problem is to apply the finite element Galerkin approach, combined with numerical quadrature. Such a method may be easily adapted to the geometry of the problem and its error analysis naturally offers itself by the variational formulation.

References

- [1] J.H. Argyris, *Energy theorems and structural analysis. A generalised discourse with applications on energy principles of structural analysis including the effects of temperature and non-linear stress-strain relations* (Butterworths, London, Toronto, Sydney, 1960).
- [2] Jean Pierre Aubin, "Behavior of the error of the approximate solutions of boundary value problems for linear elliptic operators by Galerkin's and finite difference methods", *Ann. Scuola Norm. Sup. Pisa Cl. Sci.* (3) 21 (1967), 599-637.
- [3] G. Birkhoff, M.H. Schultz, and R.S. Varga, "Piecewise Hermite interpolation in one and two variables with applications to partial differential equations", *Numer. Math.* 11 (1968), 232-256.
- [4] Jean Céa, "Approximation variationnelle des problèmes aux limites", *Ann. Inst. Fourier (Grenoble)* 14 (1964), 345-444.
- [5] Chuan-miao Chen and Vidar Thomée, "The lumped mass finite element method for a parabolic problem" (Report No. 1983-1. Department of Mathematics, Göteborg, Sweden).
- [6] Philippe G. Ciarlet, *The finite element method for elliptic problems* (Studies in Mathematics and its Applications, 4. North-Holland, Amsterdam, New York, Oxford, 1978).
- [7] R.W. Clough, "The finite element method in plane stress analysis", *Proc. Second ASCE Conf. on Electronic Computation*, 1960, 345-378 (Pittsberg, Pennsylvania, 1960).

- [8] R. Courant, "Variational methods for the solution of problems of equilibrium and vibrations", *Bull. Amer. Math. Soc.* **49** (1943), 1-23.
- [9] R. Courant, K. Friedrichs und H. Lewy, "Über die Partiellen Differenzgleichungen der Mathematischen Physik", *Math. Ann.* **100** (1928), 32-74.
- [10] M. Delfour, W. Hager and F. Trochu, "Discontinuous Galerkin methods for ordinary differential equations", *Math. Comp.* **36** (1981), 455-473.
- [11] Ю.К. Демьянович [Ju.K. Demjanovič], "Метод сеток для некоторых задач математической физики" [The net method for some problems in mathematical physics], *Dokl. Akad. Nauk SSSR* **159** (1964), 250-253.
- [12] Jim Douglas, Jr. and Todd Dupont, "Some superconvergence results for Galerkin methods for the approximate solution of two-point boundary value problems", *Topics in numerical analysis*, 89-92 (Proc. Roy. Irish Acad. Conf., University College, Dublin, 1972. Academic Press, New York, London, 1973).
- [13] K. Eriksson, C. Johnson and Vidar Thomée, "The discontinuous Galerkin method for time discretization of parabolic problems" (in preparation).
- [14] K. Eriksson and Vidar Thomée, "Galerkin methods for singular boundary value problems in one space dimension" (Report No. 1982-11. Department of Mathematics, Goteborg, Sweden, 1982).
- [15] Sandro Faedo, "Un nuovo metodo der l'analisi esistenziale e quantitativa dei problemi di propagazione", *Ann. Scuola Norm. Sup. Pisa Cl. Sci.* (3) **1** (1949), 1-40.
- [16] K. Feng, "Finite difference schemes based on variational principles" (Chinese), *Appl. Math. Comput. Math.* **2** (1965), 238-262.
- [17] K.O. Friedrichs and H.B. Keller, "A finite difference scheme for generalized Neumann problems", *Numerical solution of partial differential equations*, 1-19 (Academic Press, New York, London, 1966).

- [18] H. Fujii, "Some remarks on finite element analysis of time-dependent field problems", *Theory and practice in finite element structural analysis*, 91-106 (University of Tokyo Press, Tokyo, 1973).
- [19] B. G. Galerkin, "Rods and plates. Series occurring in various questions concerning the elastic equilibrium of rods and plates", *Vestnik Inzh.* 19 (1915), 897-908.
- [20] S. Gerschgorin, "Fehlerabschätzung für das Differenzenverfahren zur Lösung partieller Differentialgleichungen", *Z. Angew. Math. Mech.* 10 (1930), 373-382.
- [21] Pierre Jamet, "Galerkin-type approximations which are discontinuous in time for parabolic equations in a variable domain", *SIAM J. Numer. Anal.* 15 (1978), 912-928.
- [22] Dennis Jespersen, "Ritz-Galerkin methods for singular boundary value problems", *SIAM J. Numer. Anal.* 15 (1978), 813-834.
- [23] C. Johnson, U. Nävert and J. Pitkäranta, "Finite element methods for linear hyperbolic problems", *Comput. Methods Appl. Mech. Engrg.* (to appear).
- [24] P. Lasaint and P.A. Raviart, "On a finite element method for solving the neutron transport equation", *Mathematical aspects of finite elements in partial differential equations*, 89-123 (Proc. Sympos. Math. Res. Center, University of Wisconsin, Madison, 1974, Publications No. 33. Academic Press, New York, London, 1974).
- [25] Frank Natterer, "Über die punktweise Konvergenz finiter Elemente", *Numer. Math.* 25 (1975), 67-77.
- [26] J. Nitsche, "Lineare Spline-Funktionen und die Methoden von Ritz für elliptische Randwertprobleme", *Arch. Rational Mech. Anal.* 36 (1970), 348-355.
- [27] J. Nitsche, " L_∞ -convergence of finite element approximation", *Second conference on finite elements*, (Rennes, France, 1975).

- [28] Л.А. Оганесян [L.A. Oganessian], "Сходимость разностных схем при улучшенной аппроксимации границы" [Convergence of difference schemes in case of improved approximation of the boundary], *Ж. Вычисл. Мат и Мат. Физ.* 6 (1966), 1029–1042.
- [29] Л.А. Оганесян, Л.А. Руховец [L.A. Oganessian, L.A. Rukhovets], "Исследование скорости сходимости вариационно-разностных схем для эллиптических уравнений второго порядка в двумерной области с гладкой границей" [Study of the rate of convergence of variational difference schemes for second order elliptic equations in two-dimensional regions with smooth boundaries], *Ж. Вычисл. Мат и Мат. Физ.* 9 (1969), 1102–1120.
- [30] Pierre-Arnaud Raviart, "The use of numerical integration in finite element methods for solving parabolic equations", *Topics in numerical analysis*, 233–264 (Proc. Roy. Irish Acad. Conf., University College, Dublin, 1972. Academic Press, New York, London, 1973).
- [31] Lord Rayleigh, *Theory of sound*, Volume I, Second Edition (Macmillan, London, 1894).
- [32] Lord Rayleigh, *Theory of sound*, Volume II, Second Edition (Macmillan, London, 1896).
- [33] Robert D. Richtmyer, *Difference methods for initial-value problems* (John Wiley & Sons, New York, London, 1957).
- [34] Walter Ritz, "Über eine neue Methode zur Lösung gewisser Variationsprobleme der Mathematischen Physik", *J. Reine Angew. Math.* 135 (1908), 1–61.
- [35] Robert Schreiber and Stanley C. Eisenstat, "Finite element methods for spherically symmetric elliptic equations", *SIAM J. Numer. Anal.* 18 (1981), 546–558.
- [36] Ridgway Scott, "Optimal L^∞ estimates for the finite element method on irregular meshes", *Math. Comp.* 30 (1976), 681–697.
- [37] Gilbert Strang and George J. Fix *An analysis of the finite element method* (Prentice-Hall, Englewood Cliffs, New Jersey, 1973).

- [38] Vidar Thomée, "Spline approximation and difference schemes for the heat equation", *The mathematical foundations of the finite element method with applications in partial differential equations*, 711-746 (Proc. Sympos. University of Maryland, Baltimore, Maryland, 1972. Academic Press, New York, London, 1973).
- [39] Vidar Thomée, "Galerkin-finite element methods for parabolic equations", *Proceedings of the International Congress of Mathematicians*, Helsinki, 1978, 943-592 (Acad. Sci. Fennica, Helsinki, 1980).
- [40] M.J. Turner, R.W. Clough, H.C. Martin and L. Topp, "Stiffness and deflection analysis of complex structures", *J. Aero. Sci.* **23** (1956), 805-823.
- [41] Teruo Ushijima, "Error estimates for the lumped mass approximation of the heat equation", *Mem. Numer. Math.* **6** (1979), 65-82.
- [42] O.C. Zienkiewicz, *The finite element method in engineering science*, Third Edition (McGraw-Hill, London, New York, 1977).
- [43] O.C. Zienkiewicz and Y.K. Cheung, "Finite elements in the solution of field problems", *Engineer* **220** (1965), 507-510.
- [44] Miloš Zlámal, "On the finite element method", *Numer. Math.* **12** (1968), 394-409.

Department of Mathematics,
Chalmers University of Technology,
Göteborg,
Sweden.