MATHEMATICS OF COMPUTATION
Volume 74, Number 250, Pages 1033–1052
S 0025-5718(04)01757-0
Article electronically published on November 22, 2004

REVIEWS AND DESCRIPTIONS OF TABLES AND BOOKS

The numbers in brackets are assigned according to the American Mathematical Society classification scheme. The 2000 Mathematics Subject Classification can be found in print starting with the 1999 annual index of *Mathematical Reviews*. The classifications are also accessible from `www.ams.org/msc/`.

3[05C35, 90C27, 52B60]—*Report on global methods for combinatorial isoperimetric problems*, by L. H. Harper, Cambridge Studies in Advanced Mathematics, Vol. 90, Cambridge University Press, 2004, xiv+231 pp., hardcover, \$60.00, ISBN 0-521-83268-3

It is a very nice and useful book, written by a real expert in the field. A typical problem can be described as follows. Let G be a graph with the edge set E and the vertex set V, and let k be a given positive integer. The task is to choose a k-element set $S \subseteq V$ that minimizes the number of edges which are defined by one vertex from S and by one vertex outside of S.

The book has been based on many years of teaching this material to graduate students and this is certainly reflected in the style in which it is written. In a series of lemmas and theorems, the author leads the reader through rather simple cases to more complicated concepts. On the other hand, it offers a rich and varied selection of problems from this beautiful branch of combinatorial optimization to which a certain unifying "global" approach is developed. Informal comments at the end of each chapter provide a nice supplement to the main text and also help to gain some historical perspective on the subject. I believe that both specialists in the area and mathematicians with other backgrounds will find lots of new interesting material in this book.

Igor Shparlinski
E-mail address: `igor@comp.mq.edu.au`

4[65L60, 65L70, 65M60, 65Nxx, 74S05, 76M10]—*Adaptive finite element methods for differential equations*, by Wolfgang Bangerth and Rolf Rannacher, Lectures in Mathematics ETH Zürich, Birkhäuser Verlag, Basel, 2003, viii+207 pp., softcover, EUR 22.00/SF 35.00, ISBN 3-7643-7009-2

Finite element methods are, undoubtedly, one of the most general and powerful techniques for the numerical solution of partial differential equations. Their historical roots can be traced back to the 1943 paper of Richard Courant [6] on variational methods for the approximation of problems of equilibrium and vibration. Given that V is an infinite dimensional Hilbert space, $a(\cdot,\cdot)$ is a continuous and coercive bilinear functional on $V \times V$ and $\ell(\cdot)$ is a continuous linear functional on V, the archetypal linear variational problem consists of finding u in V such that

$$\text{(P):} \qquad a(u,v) = \ell(v) \quad \text{for all } v \in V.$$

Boundary value problems for scalar linear elliptic partial differential equations or elliptic systems, such as the Stokes problem modelling the flow of a viscous incompressible fluid in a bounded open set $\Omega \subset \mathbb{R}^n$, naturally fit into this abstract variational framework.

In engineering and scientific applications it is frequently the case that, instead of the field u itself, the quantity of interest is a certain *output functional* $u \mapsto J(u)$; typical examples include the weighted integral-mean-value of u, a point-value of u, the normal flux of u through (part of) the boundary $\partial\Omega$ of Ω, or, in problems that arise from fluid mechanics, the lift and the drag exerted on a body that is immersed into a viscous or inviscid fluid.

The finite element approximation of the variational problem (P) consists of selecting a finite dimensional space V_h (of dimension $N = N(h,p)$) of the space V consisting of a piecewise polynomial function of a certain degree p on a triangulation $\mathcal{T}_h$ of granularity h of the computational domain Ω, and seeking $u_h \in V_h$ such that

$$(\mathrm{P}_h): \qquad a(u_h, v_h) = \ell(v_h) \quad \text{for all } v_h \in V_h.$$

Adaptive finite element methods, driven by *a posteriori* error bounds, aim to automatically adapt the local mesh-size h or the local polynomial degree p, or both h and p, so as to accurately capture the analytical solution u, or a certain functional $u \mapsto J(u)$ of the solution.

It is this topic that forms the subject of the book by Bangerth and Rannacher under review. The book grew out of a lecture series given by the second author during the summer of 2002 at the Department of Mathematics of the ETH in Zürich. It comprises a brief Preface, followed by twelve chapters, a 24-page Appendix, a Bibliography with 138 entries, and a 5-page Index of terms; each chapter is about 15 pages long and is supplemented by computational examples as well as exercises whose model solutions are supplied in the Appendix.

As is highlighted by the authors in Chapter 1 of the book, the goal of adaptivity is the "optimal" use of computing resources according to either one of the following principles:

- Minimal work N subject to a prescribed positive tolerance `TOL`: $N \to \min$, `TOL` given; or,
- Maximal accuracy subject to prescribed work: $\mathtt{TOL} \to \min$, N given.

These goals are, traditionally, approached by mesh adaptivity driven by "local refinement indicators" based on the computed solution u_h. The process of adaptivity has three main ingredients:

- a rigorous *a posteriori* bound on the error in the quantity of interest in terms of the data and the computed solution;
- a local refinement indicator extracted from the *a posteriori* error bound;
- automatic mesh adaptation (in the form of local h-refinement, or local p-refinement, or their combination referred to as hp-refinement) according to certain refinement strategies based on the local refinement indicators.

The idea of *a posteriori* error estimation stems from the early work of Babuška and Rheinboldt [2, 3]; see also the monographs of Ainsworth and Oden [1], Babuška and Strouboulis [4], and Verfürth [21] for further detail on the subject of *a posteriori* error analysis of the finite element method. The focus of this book by Bangerth and Rannacher is a general technique for goal-oriented *a posteriori* error estimation for finite element approximations of differential equations, called the dual-weighted-residual (DWR) method, and the implementation of this technique into adaptive finite element algorithms.

To give a brief sketch of the DWR method, consider the variational problem (P) and its finite element approximation (P_h) and suppose that the goal of the computation is to find an accurate approximation to the real number $J(u)$ where $J : V \to \mathbb{R}$ is (for the sake of simplicity of presentation) a linear functional and u is the solution to problem (P).

The derivation of an *a posteriori* bound by means of the DWR method on the error $J(u) - J(u_h)$ between the unknown value $J(u)$ and its known finite element approximation $J(u_h)$ rests on considering the associated *dual* problem: find $z \in V$ such that

$$\text{(D):} \qquad a(w,z) = J(w) \quad \text{for all } w \in V.$$

Clearly, setting $w = u - u_h$ in (D), we deduce that

$$\begin{aligned} J(u) - J(u_h) = J(u - u_h) &= a(u - u_h, z) \\ &= a(u - u_h, z - v_h) \end{aligned}$$

for all $v_h \in V_h$, where, in the transition to the last line, we made use of the *Galerkin orthogonality* property: $a(u - u_h, v_h) = 0$ for all $v_h \in V_h$, which is a straightforward consequence of subtracting (P_h) from (P) with $v = v_h \in V_h \subset V$. Proceeding then, using (P), we obtain

$$J(u) - J(u_h) = \ell(z - v_h) - a(u_h, z - v_h) \quad \forall v_h \in V_h.$$

Thus we have eliminated the analytical solution u, at the expense of involving the dual solution z. The last identity can be written in a more compact form on introducing the linear functional $R(u_h) : V \to \mathbb{R}$, defined by

$$R(u_h)(v) = \ell(v) - a(u_h, v) \quad \forall v \in V,$$

referred to as the *finite element residual*, or, simply, *residual*; it measures the extent to which the numerical solution u_h fails to satisfy the equation (P). Hence,

$$\begin{aligned} J(u) - J(u_h) &= R(u_h)(z - v_h) \\ &= \langle R(u_h), z - v_h \rangle \quad \forall v_h \in V_h, \end{aligned}$$

where $\langle \cdot, \cdot \rangle$ denotes the duality pairing between the dual space V' of V and V. This *error representation formula* is at the heart of the DWR method, highlighting the fact that the error in the approximation of the value $J(u)$ depends on the interplay between the finite element residual $R(u_h)$ and the error $z - v_h$, with $v_h \in V_h$, in the approximation of the dual solution z, which acts as a weight function for the residual. Hence the terminology *dual-weighted-residual* method. In particular, the last identity implies that

$$|J(u) - J(u_h)| = \inf_{v_h \in V_h} |\langle R(u_h), z - v_h \rangle|. \tag{1}$$

In earlier incarnations of duality-based error estimation—particularly in the pioneering research pursued by the Gothenburg school (see, for example, the articles by Johnson [16], Eriksson and Johnson [13, 14], and the illuminating survey paper by Eriksson, Estep, Hansbo, and Johnson [12])—the objective was to eliminate the explicit appearance of the dual solution z from the right-hand side of (1) through a succession of upper bounds. The first of these upper bounds involved making a particular choice of v_h such as the finite element interpolant or quasi-interpolant $P_h z$

of z; this step was followed by localizing the expression $|\langle R(u_h), z - P_h z\rangle|$ through decomposing it, as a sum of analogous terms defined locally, over the elements T in the triangulation; the next step was to apply the Cauchy–Schwarz inequality to each of these local terms in tandem with an interpolation-error bound such as $\|z - P_h z\|_{L_2(T)} \leq C_{\text{int}} h_T^s \|z\|_{H^s(T)}$, where $h_T = \text{diam}(T)$, with $T \in \mathcal{T}_h$, and C_{int} is an interpolation constant; and, finally, to exploit the strong stability of the dual problem to bound the Sobolev norm $\|z\|_{H^s(\Omega)}$ in terms of the data of the dual problem and the stability constant C_{stab} of the dual problem, resulting in an *a posteriori* error bound of the form

$$|J(u) - J(u_h)| \leq C_{\text{int}} C_{\text{stab}} \left(\sum_{T \in \mathcal{T}_h} h_T^{2s} \|R(u_h)\|^2_{L^2(T)} \right)^{1/2}$$

with no explicit dependence on the dual solution. While such an *a posteriori* error bound is *reliable* in the sense that the right-hand side of the inequality is a guaranteed upper bound on the left-hand side, numerical experiments will quickly reveal that, typically, the right-hand side will overestimate the left-hand side—sometimes by orders of magnitude—even if the sharpest available values of the constants C_{int} and C_{stab} are used. A further observation in connection with the last bound is that the original feature of (1), namely that it is the interplay between $R(u_h)$ and $z - v_h$, with $v_h \in V_h$, that governs the error $J(u) - J(u_h)$, rather than the size of $R(u_h)$ alone, is completely lost through successive applications of the Cauchy–Schwarz inequality aimed at eliminating the presence of the dual solution z. The importance of preserving the dual solution z as a locally varying weight to the residual is particularly important in instances when the dual solution exhibits complex behavior over the computational domain Ω. Whether or not this is so, of course, depends entirely on the nature of the problem (P) and the choice of the output functional J. For example, when (P) is the weak formulation of an elliptic convection-dominated diffusion equation and $J(u) = u(\mathbf{x}_0)$, $\mathbf{x}_0 \in \Omega$, the dual solution z will contain a thin internal layer which will be aligned with the subcharacteristic curve passing through $\mathbf{x}_0$. It would be unreasonable to expect that the presence of such a localized and anisotropic structure in the dual solution could be represented by, or encoded into, a single constant, C_{stab}, the stability constant of the dual problem featuring in the last *a posteriori* error bound.

These recognitions motivated, in the mid-1990s, the work of Becker and Rannacher [7] where the dual-weighted-residual method was first introduced (see also [8] and the survey articles [8] and [15]). At about the same time, other researchers have also embarked on closely related investigations (see, for example, [17], [19] and [20]).

In particular, in order to derive a sharp *a posteriori* error bound from the error representation formula (1) while retaining the presence of the dual solution in the bound as a local weight to the finite element residual, it was recognized in [7] that the number of applications of the Cauchy-Schwarz inequality in the derivation of the bound has to be kept to the minimum. An *a posteriori* error bound based on the DWR method which meets these objectives can be inferred from (1); it has the form

$$|J(u) - J(u_h)| \leq \sum_{T \in \mathcal{T}_h} |\langle R(u_h)|_T, z - P_h z\rangle_T|, \tag{2}$$

where $\langle \cdot, \cdot \rangle_T$ is a localized counterpart of the duality pairing $\langle \cdot, \cdot \rangle$, $R(u_h)|_T$ is the restriction of the (global) finite element residual $R(u_h)$ to element $T \in \mathcal{T}_h$, and $P_h z \in V_h$ is the finite element interpolant or quasi-interpolant of z.

Chapters 2–4 of the book are devoted to explaining the application of the DWR method to an ODE model problem (Chapter 2) and a PDE model problem (Chapter 3), and to discussing practical aspects of the method (Chapter 4), including the evaluation of the DWR error bound (2) and other DWR error bounds akin to (2). For, strictly speaking, inequality (2), as it stands, is *not* an *a posteriori* error bound in the classical sense of the word, given that it involves the unknown analytical solution z to the dual problem (D). Clearly, z has to be computed numerically; in particular, if a finite element method is used to compute an approximation to z, then a finite element space different from V_h must be used for this purpose; once such an approximation to z is available, it has to be projected onto V_h to obtain $z_h \in V_h$ which can be used *in lieu* of $P_h z$ in (2). The additional errors incurred through the numerical approximation of the dual solution are difficult to quantify unless one embarks on reliable *a posteriori* error estimation for the dual problem; for reasons of economy, this is rarely attempted in practice. Indeed, there is very little in the current literature in the way of rigorous analytical quantification of the impact of replacing the exact dual solution z in the DWR error bound by its numerical approximation; see, however, the recent analytical work of Carstensen [5] on the estimation of higher Sobolev norm from lower order approximation, and the application of this in the context of the DWR method. A second issue is that the necessity to compute a "reasonably" accurate approximation to the dual solution results in added computational work. The authors of the book provide a convincing computational demonstration through a wide range of model problems that, except on very coarse meshes, *a posteriori* error bounds obtained by the DWR method remain reliable and very sharp even on replacement of z by its numerical approximation. In addition, when implemented into adaptive finite element algorithms, error bounds derived by the DWR method lead to economical computational meshes.

An analysis aimed at gaining further theoretical insight into the performance of the DWR method is performed in Chapter 5 of the book. The chapter also discusses the current limits of theoretical analysis of the method focusing, in particular, on convergence under mesh refinement of the finite element residual and of the weights which incorporate the numerical approximation to the dual solution z. As is noted by the authors at the end of Section 5.3, further challenges include the convergence analysis of the method on locally refined meshes, particularly in the presence of singularities in the solutions to the primal problem (P) and/or the dual problem (D). Indeed, the convergence analysis of adaptive algorithms has been the subject of active research in recent years (see, for example, the papers of Morin, Nochetto, and Siebert [18], Cohen, Dahmen, and DeVore [11], and Binev, Dahmen, and DeVore [10] in this direction in the context of energy-norm-based *a posteriori* error estimation and adaptivity for elliptic problems).

Chapter 6 is concerned with the extension of the DWR method to nonlinear variational problems. A particularly appealing feature of the DWR method from the practical point of view is that, when applied to nonlinear PDEs, the dual problem, which is simply the adjoint of the linearization of the primal problem, is still a linear problem. Hence the computational overhead of obtaining an approximate dual solution is merely a fraction of the computational complexity of solving the primal nonlinear problem itself.

Chapters 7 to 11 discuss the application of the DWR method to, respectively, eigenvalue problems, optimization problems, time-dependent problems, linear and nonlinear problems in structural mechanics, and problems in fluid dynamics including the computation of drag and lift coefficients in a viscous incompressible flow.

The book closes, in Chapter 12, with an overview of miscellaneous and open problems, including historical remarks and a survey of current developments. Some of the open problems identified by the authors include the use of the DWR method for multidimensional time-dependent problems, its application in the context of the *hp*-version finite element method, the organization of anisotropic mesh refinement, the effective control of variational crimes, the control of the error incurred in the solution of algebraic equations which result from finite element discretizations of differential equations, the application of the DWR method to nonvariational problems, and, finally, the solution of the theoretical problems raised in Chapter 5 so as to provide complete theoretical underpinning of the DWR method. Some of these are already the subject of ongoing research.

This well-written book is highly suitable as supporting text for an advanced undergraduate or a basic graduate course on adaptive finite element methods for partial differential equations. The material is clearly structured and well organized, and the numerous computational examples and exercises induce the reader to further explore the subject. The discussions of open or incompletely understood problems are particularly stimulating and raise the understanding of the reader to the forefront of current research in the field. I warmly recommend this book to anyone with interest in the analysis of finite element methods and their application to partial differential equations.

References

[1] M. Ainsworth and J.T. Oden. *A posteriori Error Estimation in Finite Element Analysis.* Wiley, New York, 2000.

[2] I. Babuška and W.C. Rheinboldt. Error estimates for adaptive finite element computations. *SIAM J. Numer. Anal.*, 15:736–754, 1978.

[3] I. Babuška and W.C. Rheinboldt. A posteriori error estimation for the finite element method. *Int. J. Numer. Meth. Eng.*, 12:1597–1615, 1978.

[4] I. Babuška and T. Strouboulis. *The Finite Element Method and Its Reliability.* Oxford University Press, Oxford, 2001.

[5] C. Carstensen. Estimation of higher Sobolev norm from lower order approximation. *SIAM J. Numer. Anal.* (Accepted for publication, 2004).

[6] R. Courant. Variational methods for the solution of problems of equilibrium and vibration. *Bull. Amer. Math. Soc.*, **49**, 1–23, 1943.

[7] R. Becker and R. Rannacher. Weighted a posteriori error control in FE methods. Lecture at ENUMATH-95, Paris, Sept. 18-22, 1995, Preprint 96-01, SFB 359, University of Heidelberg, PROC. *ENUMATH'97* (H.G. Brock et al., eds.), pp.621–637, World Scientific, Singapore, 1998.

[8] R. Becker and R. Rannacher. A feed-back approach to error control in finite element methods: Basic analysis and examples. *East–West J. Numer. Math.,* 4:237–264, 1996.

[9] R. Becker and R. Rannacher. An optimal control approach to error estimation and mesh adaptation in finite element methods. *Acta Numerica*, Vol. 10, (A. Iserles, ed.) pp.1–101, Cambridge University Press, 2001.

[10] P. Binev, W. Dahmen, and R. DeVore. Adaptive finite element methods with convergence rates. *Numerische Mathematik,* 97(2):219–268, 2004.

[11] A. Cohen, W. Dahmen, and R. DeVore. Adaptive wavelet methods for elliptic operator equations—Convergence rates. *Math. Comp.,* 70:22–75, 2001.
[12] K. Eriksson, D. Estep, P. Hansbo, and C. Johnson. Introduction to adaptive methods for differential equations. *Acta Numerica*, Vol. 4 (A. Iserles, ed.), pp.105–158, Cambridge University Press, 1995.
[13] K. Eriksson and C. Johnson. An adaptive finite element method for linear elliptic problems. *Math. Comp.,* 50:361–383, 1988.
[14] K. Eriksson and C. Johnson. Adaptive finite element methods for parabolic problems, I: linear model problem. *SIAM J. Numer. Anal.,* 28:43–77, 1991.
[15] M.B. Giles and E. Süli. Adjoint methods for PDEs: a posteriori error analysis and postprocessing by duality. *Acta Numerica*, Vol. 11 (A. Iserles, ed.), pp.145–236, Cambridge University Press, 2002.
[16] C. Johnson. Adaptive finite element methods for diffusion and convection problems. *Comput. Methods Appl. Mech. Eng.*, 82:301–322, 1990.
[17] L. Machiels, A.T. Patera, and J. Peraire. Output bound approximation for partial differential equations; applications to the incompressible Navier–Stokes equations. In S. Biringen, editor, *Industrial and Environmental Applications of Direct and Large Eddy Numerical Simulation.* Springer, Berlin, Heidelberg, New York, 1998.
[18] P. Morin, R. Nochetto, and K. Siebert. Data oscillation and convergence of adaptive FEM. *SIAM J. Numer. Anal.*, 38:466–488, 2000.
[19] J.T. Oden and S. Prudhomme. On goal-oriented error estimation for elliptic problems: Application to the control of pointwise errors. *Comput. Methods Appl. Mech. Eng.,* 176:313–331, 1999.
[20] M. Paraschivoiou and A.T. Patera. Hierarchical duality approach bounds for the outputs of partial differential equations. *Comput. Methods Appl. Mech. Eng.,* 158:389–407, 1998.
[21] R. Verfürth. *A Review of A Posteriori Error Estimation and Adaptive Mesh-Refinement Techniques.* Wiley/Teubner, New York, Stuttgart, 1996.

Endre Süli
University of Oxford

5[11B85, 11Z05, 37A45, 37B10, 68Q45, 68R15, 94A45]—*Automatic sequences. Theory, applications, generalizations*, by Jean-Paul Allouche and Jeffrey Shallit, Cambridge University Press, Cambridge, 2003, xvi+571 pp., $50.00, ISBN 0-521-82332-3

> Sequences come in all flavors. Some, such as periodic sequences, are highly organized, while others are unordered and have no simple description. The subject of this book is *automatic sequences* and their generalizations. Automatic sequences form a class of sequences somewhere between simple order and chaotic disorder. This class contains such celebrated sequences as the Thue–Morse sequence and the Rudin–Shapiro sequence. ... [from the Introduction]

The subjects of this fine book include combinatorics on words, formal languages, useful parts of number theory, formal power series, ... The chapter titles give some hint of the breadth of material appropriately touched upon: Stringology, Number Theory and Algebra, Numeration Systems, Finite Automata and Other Models of Computation, Automatic Sequences, Uniform Morphisms and Automatic Sequences, Cobham's Theorem for (k, l) Numeration Systems, Morphic Sequences, Frequency of Letters, Characteristic Words, Subwords, Cobham's Theorem, Formal Power Series, Automatic Real Numbers, Multidimensional Automatic Sequences, Automaticity, k-Regular Sequences, Physics.

Automatic Sequences is both an introduction to the study of the said sequences and related mathematics and a careful survey of known results and applications