International Journal of Differential Equations and Applications
Volume 10 No. 3 2005, 295-305

A NOTE ON THE PERTURBED LAX-MILGRAM PROBLEM

Adam Grabarski[1 §], Alicja Smoktunowicz[2]
[1,2]Faculty of Mathematics and Information Science
Warsaw University of Technology
Pl. Politechniki 1, 00-661 Warsaw, POLAND
[1]e-mail: grab@mini.pw.edu.pl
[2]e-mail: smok@mini.pw.edu.pl

Abstract: We deal with an error estimate of the perturbed Lax-Milgram problem with respect to the perturbation in the right-hand side. Two examples of perturbed problems in Sobolev spaces and their detailed analysis are given. The problem considered here is related to the effect of quadrature errors on the finite element solution, analyzed in Ciarlet [1], Strang et al [2], Janik [3] and Kołodziejczyk [4]. However, in contrast to these papers, where (very often) tedious analysis of the effect of quadrature errors of a product of two functions, is given, a different approach is used. We consider the error by *the exact integration* of the product of a projection of function f and an element from subspace where we seek the approximate solutions. This simplifies analysis and, as indicated examples show, also gives not complicated formulae. The main characteristic of it is that numerical quadrature of a product of two functions can be interpreted as a quadrature with respect to only one function.

AMS Subject Classification: —???—
Key Words: —???—

1. Formulation and Analysis of a Problem

We begin this section with some notation which will simplify the exposition. Let U and V be two real Hilbert spaces and let (u, v) denote the inner

Received: June 15, 2006

§Correspondence author

product of $u, v \in U$. We assume that $V \subset U$ and

$$\|v\|_U \leq C\|v\|_V \text{ for any } v \in V.$$

Let $\mathbf{a} : V \times V \rightarrow \mathbb{R}$ be a bilinear form. We assume that $\mathbf{a}$ is continuous, that is

$$| \mathbf{a}(u, v) | \leq M\|u\|_V\|v\|_V \text{ for all } u, v \in V.$$

In addition, assume that $\mathbf{a}$ is V-elliptic, that is, there exists a positive constant α such that

$$\mathbf{a}(u, u) \geq \alpha\|u\|_V^2 \text{ for any } u \in V.$$

The general variational problem we consider is to find $u \in V$ such that

$$\mathbf{a}(u, v) = (f, v) \text{ for all } v \in V. \tag{1}$$

Applying the Lax-Milgram theorem ([1], Theorem 1.3), we infer that such problem has one and only unique solution u.

The Galerkin method for approximating the solution u consists in defining similar problems in finite-dimensional subspace $V_h \subset V$ ($dimV_h = N = N(h)$): find $u_h \in V_h$ such that

$$\mathbf{a}(u_h, v_h) = (f, v_h) \text{ for all } v_h \in V_h. \tag{2}$$

Clearly, the existence and uniqueness of the solution of (2) is guaranteed by the Lax-Milgram theorem. Moreover, the Céa lemma ([1], Theorem 13.1) yields

$$\|u - u_h\|_V \leq \frac{M}{\alpha} \inf_{v_h \in V_h} \|u - v_h\|_V. \tag{3}$$

From (3) it follows that in order to find an upper bound on $\|u - u_h\|_V$ it is enough to find a particular element $\bar{v}_h \in V_h$ such that the upper bound on $\|u - \bar{v}_h\|_V$ is available.

Assume now that $u \in V_0 \subset V$ is the solution of (1) and

$$\Pi_h : V_0 \rightarrow V_h \tag{4}$$

is an operator such that we can bound

$$e_{\Pi_h}(v) = \|v - \Pi_h v\|_V \text{ for } v \in V_0. \tag{5}$$

From (3) we get

$$\|u - u_h\|_V \leq \frac{M}{\alpha} e_{\Pi_h}(u). \tag{6}$$

In addition, assume that $f \in U_0 \subset U$ and

$$P_h : U_0 \to U_h \tag{7}$$

is a mapping into $U_h \subset U$ such that we can estimate the error

$$e_{P_h}(f) = \|f - P_h f\|_U. \tag{8}$$

Many formulations that one will encounter can be put into the general form of perturbed problem to (2) with appropriate choices for the Hilbert space and the bilinear form **a**. We consider the following perturbed problem (2') to (2).

$$\tilde{u}_h \in V_h \qquad \mathbf{a}(\tilde{u}_h, v_h) = (P_h f, v_h) \quad \text{for all } v_h \in V_h. \tag{2'}$$

It is easily seen that there exists a unique solution to this problem.

Theorem 1. *Under the above assumptions the following inequality holds*

$$\|u - \tilde{u}_h\|_V \leq \frac{1}{\alpha}(M\ e_{\Pi_h}(u) + C\ e_{P_h}(f)). \tag{9}$$

Proof. The triangle inequality and (6) yield

$$\|u - \tilde{u}_h\|_V \leq \|u - u_h\|_V + \|u_h - \tilde{u}_h\|_V \leq \frac{M}{\alpha} e_{\Pi_h}(u) + \|u_h - \tilde{u}_h\|_V. \tag{10}$$

From (2), (2'), using the mentioned properties of the form $\mathbf{a}(u, v)$ and the following estimation of the inner-product

$$|\ (f, v)\ | \leq \|f\|_U \|v\|_U \leq C\|f\|_U \|v\|_V$$

we obtain

$$\alpha\|u_h - \tilde{u}_h\|_V^{\ 2} \leq a(u_h - \tilde{u}_h, u_h - \tilde{u}_h) = a(u_h, u_h - \tilde{u}_h) - a(\tilde{u}_h, u_h - \tilde{u}_h) =$$

$$(f, u_h - \tilde{u}_h) - (P_h f, u_h - \tilde{u}_h) = (f - P_h f, u_h - \tilde{u}_h) \leq C\|f - P_h f\|_U \|u_h - \tilde{u}_h\|_V.$$

From this it follows that for $u_h \neq \tilde{u}_h$ we have

$$\|u_h - \tilde{u}_h\|_V \leq \frac{C}{\alpha}\|f - P_h f\|_U = \frac{C}{\alpha} e_{P_h}(f).$$

Therefore, using (10) we get (9). ■

Now we will formulate four remarks. They will be very useful in analysis of examples presented in Section 2.

Remark 1 If $V = U, V_0 = U_0$ and $\mathbf{a}(u, v) = (u, v)$, then $u = f$ is the solution of (1). Here $\alpha = M = C = 1$, therefore

$$\|u - \tilde{u}_h\|_V = \|f - \tilde{u}_h\|_V \leq e_{\Pi_h}(f) + e_{P_h}(f). \tag{11}$$

Remark 2 Under the assumptions of **Remark 1**, and, in addition, if $U_h = V_h$, then $\tilde{u}_h = P_h f$ is the solution of (2'). This implies

$$\|u - \tilde{u}_h\|_V = \|f - P_h f\|_U = e_{P_h}(f). \tag{12}$$

Remark 3 From (10) it follows that in order to achieve the same order of approximation of elements of a space V_0 by the elements of a space V_h, we should apply approximate U_0 at least on the same level. In particular, these requirements will be fulfilled if $V_0 = U_0$ and $\Pi_h = P_h$. Clearly, usually smaller error $e_{P_h}(f)$ gives better approximation of u.

Remark 4 If $v_1, \ldots, v_N$ is a basis in a space V_h, then our problem (2') is equivalent to solving of a linear system of equations

$$A\alpha = F, \tag{13}$$

where $A = \{\mathbf{a}(v_i, v_j)\}_{i,j=1}^N$, $F = [(P_h f, v_1), \ldots, (P_h f, v_N)]^T$, and the vector $\alpha = [\alpha_1, \ldots, \alpha_N]^T$ is a vector of the coefficients of $\tilde{u}_h$ in the basis $v_1, \ldots, v_N$, i.e.

$$\tilde{u}_h = \sum_{j=1}^{N} \alpha_j v_j. \tag{14}$$

2. Some Examples

In order to illustrate our problem we give a few examples. **Example 1** deals with the commonly used least squares problem. In this case the form $\mathbf{a}(u, v)$ is the inner product. Here we consider two cases. In **Case A** ($U_h = V_h$ and $P_h = \Pi_h$) the perturbed solution $\tilde{u}_h$ can be found easily, however in Case B ($U_h \neq V_h$) we have to solve a linear system of equations. Finally, we present two tables which contain comparisons of the errors in **Case A** and **B** (see **Table 1**) and empiric verification of the order of convergence (see **Table 2**). **Example 2** deals with a boundary value problem for a linear differential equation of the second order. In **Case A** and **B** we use similar subspaces as in **Example 1**. In these two cases the problem of finding the perturbed solution $\tilde{u}_h$ is equivalent to solving a linear system of equations. We finish example with a table for these two cases (**A** and **B**) for two different functions $f(x)$.

2.1. Example 1

Let $V = U = L^2(a,b)$, $(u,v) = \int_a^b u(x)v(x)dx$, $\|u\|_{L^2} = (\int_a^b u^2(x)dx)^{1/2}$ and $\mathbf{a}(u,v) = (u,v)$. Let us define a space

$$V_h = \{v \in C[a,b] : v_{|_{I_i}} \in P_1(I_i),\ \ i = 1,\ldots,n\},$$

where $I_i = [x_{i-1}, x_i]$, $x_i = a + ih$, $h = (b-a)/n$, $P_1(I_i)$ being the space of all polynomials of degree at most one in the interval I_i.

Let $\Pi_h : C[a,b] \to V_h$ be a mapping defined by the interpolation conditions

$$(\Pi_h v)(x_i) = v(x_i),\ \ i = 0,\ldots,n.$$

It is known ([1], Theorem 16.2) that

$$e_{\Pi_h}(v) \le C_k h^k |v|_{H^k} \text{ for } \ v \in H^k(a,b),\ k = 1,2, \tag{15}$$

where C_k are constants independent of h, v and

$$|v|_{H^k} = (\int_a^b [v^{(k)}(x)]^2 dx)^{1/2} \tag{16}$$

denotes the semi-norm on Sobolev space $H^k(a,b)$.

Now we consider two important cases.

2.1.1. Case A

Let $V_0 = U_0 = H^k(a,b)$ for $k=1$ or $k=2$. Let $U_h = V_h$ and $P_h = \Pi_h$. Then , according to **Remark 2**,

$$\|u - \tilde{u}_h\|_{L^2} \le C_k h^k |f|_{H^k} \text{ for } k = 1 \text{ or for } k = 2 \tag{17}$$

and $\tilde{u}_h = \Pi_h f$. Thus, we can write $\tilde{u}_h(x)$ in the following form

$$\tilde{u}_h(x) = \sum_{j=0}^{n} f(x_j) v_j(x),$$

where $v_0(x), \ldots, v_n(x)$ form the basis in a space V_h and are piecewise linear. Let

$$v_0(x) = \begin{cases} (x_1 - x)/h & \text{for } x \in [x_0, x_1], \\ 0 & \text{otherwise,} \end{cases}$$

$$v_i(x) = \begin{cases} (x - x_{i-1})/h & \text{for } x \in [x_{i-1}, x_i], \\ (x_{i+1} - x)/h & \text{for } x \in [x_i, x_{i+1}], \\ 0 & \text{otherwise,} \end{cases}$$

for $i = 1, \ldots, n-1$, and

$$v_n(x) = \begin{cases} (x - x_{n-1})/h & \text{for } x \in [x_{n-1}, x_n], \\ 0 & \text{otherwise.} \end{cases}$$

2.1.2. Case B

Let $U_h = \{v \in C[a,b] : v_{|_{I_i}} \in P_2(I_i)\}$, where $P_2(I_i)$ is the space of all polynomials of degree at most 2 in the interval I_i.

Let us define a mapping $P_h : C[a,b] \to U_h$ satisfying the following interpolation conditions

$$\begin{cases} (P_h v)(x_i) = v(x_i), & i = 0, \ldots, n, \\ (P_h v)(x_{i+1/2}) = v(x_{i+1/2}), & i = 0, \ldots, n-1, \end{cases}$$

where $x_{i+1/2} = x_i + h/2$.

In this case, see ([1], Theorem 16.2), we get

$$e_{P_h}(v) \leq \tilde{C}_k h^k |v|_{H^k} \text{ for } v \in H^k(a,b),\ k = 1,2,3, \tag{18}$$

where $\tilde{C}_k$ are constants independent of h and v.

Here we set $V_0 = U_0 = H^k(a,b)$ for $k = 1, 2$ or 3.

Due to **Remark 1**, for $k = 1, 2$ we obtain

$$\|u - \tilde{u}_h\|_{L^2} \leq (C_k + \tilde{C}_k) h^k |f|_{H^k} \tag{19}$$

and

$$\|u - \tilde{u}_h\|_{L^2} \leq C_3 h^2 |f|_{H^2} + \tilde{C}_3 h^3 |f|_{H^3}. \tag{20}$$

Notice that a form of $\tilde{u}_h$ seems to be more complicated now. **Remark 4** yields to the solution of a linear system of equations. If we take v_j,

$j = 0, \ldots, n$, defined by **Case A**, as a basis of our space, then we can write $\tilde{u}_h$ as follows

$$\tilde{u}_h(x) = \sum_{j=0}^{n} \alpha_j v_j(x), \tag{21}$$

where the coefficients $\alpha_0, \ldots, \alpha_n$ can be determined from the linear system of equations $A\alpha = F$, where

$$A = \frac{h}{6} \begin{pmatrix} 2 & 1 & & & \\ 1 & 4 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & 4 & 1 \\ & & & 1 & 2 \end{pmatrix} \in \mathbb{R}^{(n+1)\times(n+1)} \tag{22}$$

and

$$F = \frac{h}{6}[f_0 + 2f_{1/2}, 2(f_{1/2} + f_1 + f_{3/2}), \ldots,$$
$$2(f_{n-3/2} + f_{n-1} + f_{n-1/2}), 2f_{n-1/2} + f_n]^T, \tag{23}$$

where $f_i = f(x_i)$, $i = 0, \ldots, n$; $f_{i+1/2} = f(x_i + h/2)$, $i = 0, \ldots, n-1$.

As was pointed in **Remark 3**, in **Case B** we can expect better approximation. Numerical tests, performed in MATLAB for a function $f(x) = \sin x$ in the interval $[0, \pi]$, confirm that. **Table 1** contains the computed least squares errors e_i for $i = 1, 2$. More precisely, e_1 is used in **Case A** and e_2 in **Case B**. In addition, we computed also the quantities e_1/e_2. **Table 2** gives information on upper error bounds for values $h = \pi/n$ and for $h/2$. It is pointed out that in these cases the chosen methods behave like methods of the second order and the quantities e_1/e_2 are very close to $4 = h^2/(h/2)^2$ in floating point arithmetic.

n	e_1	e_2	e_1/e_2
5	$0.2525e{-1}$	$0.1077e{-1}$	2.34
10	$0.6354e{-2}$	$0.2628e{-2}$	2.41
20	$0.1591e{-2}$	$0.6530e{-3}$	2.44
40	$0.3980e{-3}$	$0.1639e{-3}$	2.43
80	$0.9949e{-4}$	$0.4095e{-4}$	2.43
100	$0.2487e{-4}$	$0.1024e{-4}$	2.43

Table 1

n	${e_1}^h/{e_1}^{h/2}$	${e_2}^h/{e_2}^{h/2}$
5	–	–
10	3.9739	4.0981
20	3.9937	4.0245
40	3.9975	3.9841
80	4.0004	4.0024
100	4.0004	3.9990

Table 2

Notice that a mapping (connected to Gauss-Legendre quadrature) $P_h : C[a,b] \to U_h = \{v \in L^2(a,b) : v_{|_{I_i}} \in P_1(I_i)\}$ by formulae

$$\begin{cases} (P_h v)(y_i) = v(y_i), & i = 1, \ldots, n, \\ (P_h v)(z_i) = v(z_i), & i = 1, \ldots, n, \end{cases}$$

where $y_i = x_{i-1} + ht_1$, $z_i = x_{i-1} + ht_2$, $t_{1,2} = (3 \pm \sqrt{3})/6$ (here t_1, t_2 are the roots of the Legendre polynomial of the second degree in the interval $[0,1]$), is almost identical as in **Case B**. Also the computational cost is the same as in **Case B**.

2.2. Example 2

Consider the simple boundary value problem on $[a,b]$ ($\lambda \geq 0$).

$$\begin{cases} -u" + \lambda u = f(x), & x \in (a,b), \\ \quad u(a) = 0, & u(b) = 0. \end{cases}$$

Take $U = L^2(a,b)$ with the same inner-product and norm chosen in **Example 1**. Write $V = H^1(a,b)$ with the norm

$$\|v\|_{H^1} = (|v|^2_{H^1} + \|v\|^2{}_{L^2})^{1/2}.$$

Let $\mathbf{a}(u,v) = \int_a^b u'(x)v'(x)dx + \lambda \int_a^b u(x)v(x)dx$. The form $\mathbf{a}$ is bilinear, continuous and H^1-elliptic, where $M = \max(1, \lambda)$. Here $\alpha = \frac{\pi^2}{\pi^2 + (b-a)^2}$, by the Friedrichs inequality ([5], p. 264): $\|v\|_{L^2} \leq \frac{b-a}{\pi} | \, v \, |_{H^1}$ for a function v vanishes at the ends of $[a,b]$.

Let, using the same notation as in **Example 1**,

$$V_h = \{v \in C[a,b] : v_{|_{I_i}} \in P_1(I_i), \;\; i = 1, \ldots, n, \;\; v(a) = v(b) = 0\}$$

and let $\Pi_h : C[a,b] \to V_h$ be a mapping given by the interpolation conditions

$$(\Pi_h v)(x_i) = v(x_i), \;\; i = 1, \ldots, n-1.$$

Then, see ([1], Theorem 16.2),

$$e_{\Pi_h}(v) \leq C_0 h |v|_{H^2} \text{ for } \; v \in H^2(a,b). \tag{24}$$

Using the same approach as in **Example 1**, we consider two cases.

2.2.1. Case A

Let $V_0 = H^2(a,b) \cap H_0^1(a,b)$ and $U_0 = H^k(a,b) \cap H_0^1(a,b)$ for $k = 1$ or $k = 2$, where $H_0^1(a,b) = \{v \in H^1(a,b) : \ v(a) = v(b) = 0\}$. Moreover, let $U_h = V_h$ i $P_h = \Pi_h$. Then, (9), (15) i (24) yield

$$\|u - \tilde{u}_h\|_{H^1} \leq \frac{h}{\alpha}(MC_0|u|_{H^2} + C_k h^{k-1}|f|_{H^k}) \text{ for } k = 1 \text{ or } k = 2. \tag{25}$$

We take the same basis functions $v_1, \ldots, v_{n-1}$ as in **Example 1**. Then a perturbed problem is equivalent to a linear system of equations $A\alpha = F$, where $A \in \mathbb{R}^{(n-1)\times(n-1)}$ has the following form

$$A = \frac{1}{6h} \times \begin{pmatrix} 12+4\lambda h^2 & -6+\lambda h^2 & & & \\ -6+\lambda h^2 & 12+4\lambda h^2 & -6+\lambda h^2 & & \\ & \ddots & \ddots & \ddots & \\ & & -6+\lambda h^2 & 12+4\lambda h^2 & -6+\lambda h^2 \\ & & & -6+\lambda h^2 & 12+\lambda h^2 \end{pmatrix}. \tag{26}$$

Here

$$F = \frac{h}{6}[f_0 + 4f_1 + f_2, \ldots, f_{n-2} + 4f_{n-1} + f_n]^T, \tag{27}$$

where $f_i = f(x_i)$, $i = 0, \ldots, n$.

The solution of a perturbed problem is given by

$$\tilde{u}(x) = \sum_{j=1}^{n-1} \alpha_j v_j(x). \tag{28}$$

Notice that now the right-hand side vector F is computed by Simpson formula (paraboloids).

2.2.2. Case B

Let $U_h = \{v \in C[a,b] : v_{|_{I_i}} \in P_2(I_i), \ \ v(a) = v(b) = 0\}$ and $P_h : C[a,b] \to U_h$

be a mapping given by

$$\begin{cases} (P_h v)(x_i) = v(x_i), & i = 1, \dots, n-1, \\ (P_h v)(x_{i+1/2}) = v(x_{i+1/2}), & i = 0, \dots, n-1, \end{cases}$$

where $x_{i+1/2} = x_i + h/2$.

Then, from (9), (18) i (24) we get

$$\|u - \tilde{u}_h\|_{H^1} \leq \frac{h}{\alpha}(MC_0|u|_{H^2} + \tilde{C}_k h^{k-1}|f|_{H^k}), \text{ for } k = 1 \text{ or } k = 2 \text{ or } k = 3. \tag{29}$$

The space V_0 is the same as chosen in **Case A**, and $U_0 = H^k(a,b) \cap H_0^1(a,b)$ for $k = 1$ or $k = 2$, or $k = 3$.

For the basis defined in **Case A** the matrix A of a linear system of equations $A\alpha = F$ is given by (26) and right-hand vector is equal to

$$F = \frac{h}{3}[f_{1/2} + f_1 + f_{3/2}, \dots, f_{n-3/2} + f_{n-1} + f_{n-1/2}]^T, \tag{30}$$

where $f_i = f(x_i)$, $i = 1, \dots, n-1$; $f_{i+1/2} = f(x_i + h/2)$, $i = 0, \dots, n-1$.

In this case, similarly as in **Example 1**, we can expect better approximation than in **Case A**. Numerical tests performed in MATLAB confirm that. We took $\lambda = 1$ and $f(x) = 2\sin x$ in the interval $[0, \pi]$ (then $u(x) = \sin x$ is the solution of our problem)-see **Table 3** and $\lambda = 1$ and $f(x) = (1-x)arctgx + \frac{2(x+1)}{(x^2+1)^2}$ in the interval $[0, 1]$ (then $u(x) = (1-x)arctgx$ is the solution of our problem)-see **Table 4**.

The notation we use below is the same as in **Example 1**.

Table 3 **Table 4**

n	e_1	e_2	e_1/e_2
5	$0.3569e-1$	$0.1598e-1$	2.23
10	$0.9054e-2$	$0.3924e-2$	2.31
20	$0.2272e-2$	$0.9768e-3$	2.33
40	$0.5683e-3$	$0.2442e-3$	2.33
80	$0.1421e-3$	$0.6104e-4$	2.33
100	$0.3553-4$	$0.1526e-4$	2.33

n	e_1	e_2	e_1/e_2
5	$0.7499e-2$	$0.6273e-2$	1.20
10	$0.1883e-2$	$0.1564e-2$	1.20
20	$0.4712e-3$	$0.3907e-3$	1.21
40	$0.1178e-3$	$0.9766e-4$	1.21
80	$0.2946e-4$	$0.2441e-4$	1.21
100	$0.7365e-5$	$0.6103e-5$	1.21

References

[1] P.G. Ciarlet, *Basic Error Estimates for Elliptic Problems*, Handbook of Numerical Analysis, Volume II: Finite Element Methods, Part 1, North-Holland, Amsterdam (1991).

[2] G. Strang, G.J. Fix, *An Analysis of the Finite Element Method*, Englewood Cliffs (1973).

[3] T.J. Janik, The effect of quadrature errors on the finite element solution of some nonlinear elliptic problems, *Demonstratio Mathematica*, **XIX**, No. 4 (1986).

[4] J.A. Kołodziejczyk, The effect of quadrature errors on the collocation-Galerkin solution of the two-point boundary value problem, *Demonstratio Mathematica*, **XXII**, No. 3 (1989).

[5] G. Hämmerlin, K.-H. Hoffmann, *Numerical Mathematics*, Springer-Verlag, N. York (1991).