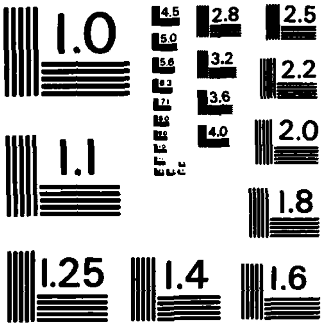
AD-A121 570 A GALERKIN METHOD ON NONLINEAR SUBSETS AND ITS APPLICATION TO A SINGULAR..(U) YALE UNIV NEW HAVEN CT DEPT OF COMPUTER SCIENCE J SUN 23 JUN 82 TR-217/82 1/1

UNCLASSIFIED N00014-82-K-0184 F/G 12/1 NL

END

FILMED

DTIC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

LUX ET VERITAS

YALE UNIVERSITY
DEPARTMENT OF COMPUTER SCIENCE

ACKNOWLEDGEMENT

The author would like to express his gratitude to Professor Martin H. Schultz for his interest in this work and valuable suggestions.

A Galerkin Method on Nonlinear Subsets and Its Application to a Singular Perturbation Problem

Jiachang Sun[1]

Technical Report #217/82

June 23, 1982

[1] Computing Center, Academia Sinica, Beijing, China. This research was supported in part by ONR Grant N00014-82-K-0184, and Yale University.

Table of Contents

A Galerkin Method on Nonlinear Subsets and Its Application to a Singular Perturbation Problem

ABSTRACT

In the Ritz-Galerkin method the linear subspace of the trial solutions is extended to a closed subset. As an example, a class of so-called sublinear approximation and interpolation is developed. Some results, such as orthogonalization and minimum property of the error function, are obtained. A second order scheme has been developed for solving the linear singular perturbation elliptic problem

$$-\epsilon u'' + p(x) u' + q(x) u = f(x), \qquad u(0) = u(1) = 0.$$

Error estimates are given for a uniform mesh size h:

$$\|u_s^h - u\|_i \leq C_i h^{1.5}, \quad \|u^{(i)}{}_s^h - u^{(i)}\|_\infty \leq C_{\infty,i} h^{1-i}, (i=0,1)$$

if $h < \dfrac{2}{\|p\|_\infty} \epsilon$, where the constants C_i and $C_{\infty,i}$ $(i=0,1)$ all are uniformly bounded for small ϵ.

For the same accuracy, the present nonlinear scheme is one order of magnitude more than the usual method used in the piecewise linear subspace. Numerical results for linear and semi-linear singular perturbation problems are included.

1. Introduction

The development of finite element methods has been successful in various fields. From a mathematical point of view, the method is one of extensions of Rayleigh-Ritz-Galerkin technique, ([1], [15], [16], [17]). Usual finite element schemes, choosing piecewise polynomials as trial functions, are very efficient when there are no steep gradients in the true solution. Otherwise, poor results might occur. In order to get accurate numerical data, one may use adaptive mesh technique(e.g. [8]) or a higher precision scheme such as h-version and p-version respectively [3]. Beyond usual polynomials, rational elements(e.g. [24]) and exponential elements [9] have been introduced to enrich the trial subspace to reduce number of parameters for a given precision. One thing in common among these techniques is that they are all reduced to a discrete linear system if the original differential equation is linear.

Nevertheless, our approach is quite different. To find a better discrete approximation of weak solutions with steep gradients, we try to relax the limitation of replacing the continuous variational problem only by a sequence of finite-dimensional subspaces. Hence, in this paper, we present an extension of the finite element method from subspace to more general subsets and adopt the method to solve singular perturbation problems (including linear and semi-linear) in one dimension. For linear problems, our aim is to solve a small semi-linear system instead of a large linear system which arises by using the usual trial subspace of piecewise polynomials for a given precision.

From a practical point of view, there are, at least, two questions which need to be answered now. First, how to find a good *non-linear* approximation of a non-linear functional space which can be devised especially for singularity problems. Secondly, how to solve the resulting discrete non-linear system efficiently. This non-linear approximation should include conventional piecewise polynomial and it is expected to be *not too far*, in some sense, from linear approximation in order to meet the theoretical demand (such as convergence and to keep some behaviors of the true solution) and to satisfy the practical aim.

The approximation used in this paper is called *piecewise mapping-polynomial* or *spline mapping-polynomial*. It means that the approximation is of piecewise, and in each subinterval a local one to one mapping is applied first, then a polynomial approximation is used in the mapping plane. The final approximation is obtained by using the inverse mapping, and the whole approximation function has some orders of smoothness according to various requirements. In particular, it reduces the usual polynomial or polynomial-spline approximation if the mapping is always equal to the identity mapping.

A large amount of attention has recently been focused on the difficult singular perturbation boundary value problems. These problems arise from some different fields, for examples, boundary layer or convective-diffusion type flows in fluid dynamics. Conventional methods applied to such problem result in unrealistic oscillation and poor approximation unless the mesh length h is excessively small. Some effects have been done by various authors using local higher order polynomial approximation with some parameter, called 'Upwinding' methods, to match the true solution better at the nodes. The method has been discussed by Christie and Mitchell [6], Barrett,Morton [4], Heinrich and el. [12], Babyska [2], etc. An "Exponentially fitted method" developed by de Groen and Hemker [9]. is to add a piecewise exponential term to enrich the subspace of piecewise polynomial.

In section 2 and 3, we generalize respectively the usual Ritz and the Galerkin method from linear trial subspaces to subsets, and derive some results such as orthogonalization and error estimations. A brief discussion about 'sub-linear' operator and its approximation is given in section

4. In section 5, the semi-linear finite element technique is studied by solving singular perturbation problems in one-dimension: $-\epsilon u''+pu'+qu=f$, $u(0)=u(1)=0$. The results show an improvement over one more order precision than the corresponding scheme of using piecewise linear subspace and that the constraint of mesh size h is relaxed from $O(\epsilon^2)$ to $O(\epsilon)$. A linear and semi-linear test singular perturbation problems are given in section 6. Computational result agree with the above theoretical analysis.

Some research results on the same topic in two-dimensions will be reported separately [22].

2. A Ritz method on subsets for self-adjoint equations

Consider a self-adjoint elliptic linear differential equation

$$Lu = f \tag{1}$$

$a(u,v) = (Lu,v)$ is a positive quadratic form in a real Hilbert space H with an inner product $(*,*)$ and a norm $||*||$:

$$C_2 ||u||^2 \leq a(u,u) \leq C_1 ||u||^2, \quad \text{for all } u \epsilon H \tag{2}$$

where C_1 and C_2 are positive constants. It implies that

$$|a(u,v)| \leq C_1 ||u|| \, ||v||, \text{ for all } u,v \in H. \tag{3}$$

u is defined as a weak solution of (1) if it satisfies

$$a(u,v) = (f,v) \quad \text{for all } v \in H. \tag{4}$$

It is well known that u is a weak solution of (1) if and only if u is the unique minimum solution of a quadratic functional I, i.e.

$$I(u) = \inf_{v \epsilon H} I(v) = \inf_{v \epsilon H} \{a(v,v) - 2(f,v)\} \tag{5}$$

As a well-known discrezation, H in the variational problem (5) is replaced by a sequence of finite-dimensional subspaces V^h contained in H:

$$I(u^h) = \inf_{v \epsilon V^h} I(v)$$

which is equivalent to the following weak solution

$$a(u^h, v^h) = (f, v^h), \text{ for all } v^h \in V^h. \tag{6}$$

Now we replace H in (5) by a sequence of closed subsets S^h with the same finite-dimensional parameters. Let T be an one-to-one continuous mapping from an open convex set V_1^h of V^h onto S^h: $TV_1^h = S^h$.

Definition 1: [25] The mapping T: $V_1^h \rightarrow S^h$ is differentiable in the open convex set V_1^h, if for each $v \epsilon V_1^h$ there is a Jacobian matrix $T'(v)$ such that

$$\lim_{\alpha\to 0} ||\frac{T(v+\alpha\eta) - T(v)}{\alpha} - T'(v)\eta|| = 0, \text{ for each } \eta \in V_1^h. \tag{7}$$

In particular, T' = T if T is a linear mapping.

Consider a restricted variational problem on the closed subset S^h:

$$I(u_s) = \inf_{u\in S^h} I(v) \tag{8}$$

Since S^h is closed, so there exists a solution of (8) in S^h. If u_s minimizes I over S^h, u_s=Tw, then for any $\alpha\geq 0$ and $\eta \in V_1^h$

$$I(u_s) \leq I(T(w+\alpha\eta)). \tag{9}$$

Let

$$T(w+\alpha\eta) = Tw + \alpha T\eta + \kappa(\alpha)$$

where T is positive-homogeneous and

$$\kappa(\alpha) = T(w+\alpha\eta) - Tw - \alpha T\eta$$

The right side of (9) is

$$I(u_s) + 2\alpha[a(u_s,T\eta) - (f,T\eta)] + 2[a(u_s,\kappa(\alpha)) - (f,\kappa(\alpha))]$$
$$+ \alpha^2 a(T\eta,T\eta) + 2\alpha a(T\eta,\kappa(\alpha)) + a(\kappa(\alpha),\kappa(\alpha)) = I(\alpha)$$

As a function of the parameter α, the fact that u_s minimizes I over S requires $\lim_{\alpha\to 0} I'(\alpha) = 0$. Observing that

$$\kappa(0) = 0,\ \kappa'(0) = (T'(T^{-1}u_s) - T)\eta,$$

and

$$0 = I'(\alpha)|_{\alpha\to 0} = 2\ \{\ a(u_s,T\eta) - (f,T\eta) + a(u_s,\ \kappa'(0)\ ) - (f,\ \kappa'(0)\ )\}$$

hence, it yields

$$a(u_s,T'(T^{-1}u_s)\eta) = (f,T'(T^{-1}u_s)\eta) \text{ for all } \eta \in V_1^h. \tag{10}$$

Therefore

Theorem 2: *If (i) V^h is a subspace of H. (ii) S^h is a close subset of H. (iii) T is an one-*

to-one positive homogeneous and differentiable mapping from an open convex set V_1 of V^h onto S^h: $TV_1 = S^h$. Then (iv) There exists a solution u_e of (8) and (10) holds.

The above Theorem shows that the nonlinear system (10) has at least a solution which minimizes the variational problem (8). Usually, it does not mean they are equivalent each other. Because there are no guarantee of unique solution in general case. However, we have the following conclusion:

Theorem 3: *If V_1 contains u^h defined in (6), then for the mapping T which is sufficient close to a linear mapping, i.e., $\|T - T'\|$ is sufficient small, the nonlinear system (10) has unique solution which minimizes the variational problem (8).*

Proof: In fact, (10) can be rewritten as

$$a(u, v^h) = (f, v^h) + Q(u,v^h)$$

where

$$Q(u,v^h) = a(u, [T-T'(T^{-1}u)]v^h) + (f, [T-T'(T^{-1}u)]v^h))$$

Since there is unique solution in (6), hence, the above equations system also has unique solution if $\|T - T'\|$ is sufficient small.

Now we suppose that the generalized coordinates (real parameters) of the subset S^h are $q_1,...,q_n$, then the first variational equations of $I(w)$ in S^h must be vanished

$$\frac{1}{2}\frac{\partial I}{\partial q_i} = a(w,\frac{\partial w}{\partial q_i}) - (f,\frac{\partial w}{\partial q_i}) = 0, \text{ for } i = 1,...,n. \tag{11}$$

and the determinant of the second variational matrix at the point of the solution is positive

$$\det(\frac{\partial^2 I}{\partial q_i \partial q_j}) > 0. \tag{12}$$

Let $\{B_j\}$ be a basis, then for each $w \epsilon S^h$

$$w = T^{-1}w + w^*, \frac{\partial w}{\partial q_i} = B_i + \frac{\partial w^*}{\partial q_i}, \text{ where } T^{-1}w = \Sigma\, q_j B_j,\ w^* = w - T^{-1}w.$$

Substituting the above formulas into (11) yields

$$\Sigma\, a(B_i,B_j)q_j = (f,B_i) + G_i(q) \tag{13}$$

where $G_i = (f,\frac{\partial w^*}{\partial q_i}) - a(w^*,\frac{\partial w}{\partial q_i}) - \Sigma\, q_j a(B_j,\frac{\partial w^*}{\partial q_i})$.

Hence, the equations of the weak solution in subsets are different from ones in subspaces only by the last extra term which tends zero when the subset S^h tends a subspace. Also, the system (11) can be written as

$$a(w,B_i) = (f,B_i) + G^*_i(q), \quad \text{where } G^*_i = (f,\frac{\partial w^*}{\partial q_i}) - a(w,\frac{\partial w^*}{\partial q_i}). \tag{14}$$

Hence, for each $v \epsilon V^h$, ignoring the extra terms, we get an approximate equations

$$a(u_s,v) = (f,v) \quad \text{for all } v \epsilon V^h \tag{15}$$

Because w in (14) corresponds to the unique solution of the variational problem (8) for the positive quadratic form a(u,u) restricted in the subset S^h, being the continuity of solutions with the system, there also exists a solution u_s of the system (15) in S^h, if the distance between V^h and S^h is sufficient small. Geometrically, it is obvious. In fact, from (11) and (12), it means that, as a hypersurface in the n dimension of $(q_1,...q_n)$, $z = \partial I/\partial q_i$ is separated by a hyperplane $z = 0$ and they have only one intersection point. Moving this hypersurface, there still exists a unique intersection point if the moving distance is sufficient small.

For practical aim there is another approximation versions of (14): Find $u_s \epsilon S^h$ such that

$$a(u_s,\ u_s - v^h) = (f,\ u_s - v^h), \text{ for all } v^h \ \epsilon\ V^h. \tag{16}$$

Suppose u_s is the unique solution of (16). From (4), for any v^h in V^h, $a(u,u_s - v^h) = (f,u_s - v^h)$, subtracting (16) leads to $a(u - u_s,u_s - v^h) = 0$. Hence

$$a(u - v^h,u - v^h) = a(u - u_s,u - u_s) + a(v^h - u_s,v^h - u_s).$$

Using (2), furthermore, for any v^h in V^h,

$$C_2 \, ||u - u_s||^2 \leq a(u - u_s, u - u_s) \leq a(u - v^h, u-v^h) \leq C_1 \, ||u - v^h||^2.$$

There are similar formulas for the case of (15). Thus, we have proved the following fundamental theorem of the Ritz method on subsets which is an extension of the Theorem 1.1 in [2] for subspaces.

Theorem 4: *Suppose u_s is the unique solution of (16) or (15) in a closed subset S^h, then it satisfies the following properties:*

(a) Minimum property

$$a(u - u_s, u - u_s) = \inf_{v^h \in V^h} a(u - v^h, u - v^h), \tag{17}$$

or

$$a(u - u_s, u - u_s) = \inf_{v^h \in V^h} a(u - u_s - v^h, u - u_s - v^h), \tag{18}$$

and

$$||u - u_s|| \leq C \inf_{v^h \in V^h} ||u - v^h|| \tag{19}$$

or

$$||u - u_s|| \leq C \inf_{v^h \in V^h} ||u - u_s - v^h|| \tag{20}$$

where C is a constant.

(b) Orthogonalization

$$a(u - u_s, u_s - v^h) = 0, \text{ for all } v^h \text{ in } V^h. \tag{21}$$

or

$$a(u - u_s, v^h) = 0, \text{ for all } v^h \text{ in } V^h. \tag{22}$$

In practical view, as a system for the weak solution, (15) is more attractive than (16). And the difference between them could be small if the subset is 'not far' from a subspace in some sense.

3. A Galerkin Method on a closed nonlinear subset

The analysis in section 2 can be extended from the Ritz to the Galerkin method. Suppose that the operator L in (1) is not self-adjoint in which derivatives of odd order spoil the self-adjointness of an elliptic equation and the associated quadratic functional I(v) defined in (5) is not positive definite. The problem now is to find a stational point rather than a minimum of I(v). There are some results on the existence of the weak solution (4), e.g. Babuska and Aziz [1], Strang and Fix [17]. Let us quote a few results of Galerkin Method first.

Theorem 5: *Let H_1 and H_2 be two real Hilbert spaces with inner products $(*,*)_{H_1}$ and $(*,*)_{H_2}$, respectively, (f,v) be a continuous linear functional on H_2 and a(u,v) a bilinear form with three inequalities*

(i) $|a(u,v)| \leq C_1 ||u||_{H_1} \, ||v||_{H_2}$ for all $u \in H_1$ and $v \in H_2$, where $C_1 < \infty$.

(ii) $\inf_{u \in H_1} \sup_{v \in H_2} \dfrac{|a(u,v)|}{||u||_{H_1} ||v||_{H_2}} \geq C_2 > 0$.

(iii) $\sup_{u \in H_1} |a(u,v)| > 0, \; v \neq 0$

Then there exists one and only one weak solution u_0 of the functional equation Lu = f such that

$$a(u_0,v) = (f,v) \qquad \text{for all } v \in H_2 \tag{23}$$

and

$$||u_0||_{H_1} \leq C_2^{-1} \, ||f||_{H'_2}$$

A proof of this result can be found in [1], theorem 5.2.1. Galerkin's method is the natural discretization of the weak form. In general it involves two families of functions __ a subspace S^h of the solution space (or trial space) H_1 and a subspace V^h of the test space H_2. Then the Galerkin solution u^h is the element of S^h which satisfies

$$a(u^h, v^h) = (f,v^h) \qquad \text{for all } v^h \in V^h \tag{24}$$

Since both S^h and V^h are linear subspaces, if $\{s_j\}$ is a basis for S^h and $\{v_j\}$ is a basis for V^h, the solution $u^h = \Sigma q_j s_j$ satisfies a linear system

$$A\,q = d \tag{25}$$

where

$$A = (\ a(s_i, v_j)\ ), \qquad d = (f, v_j) \tag{26}$$

If A^{-1} exists, there is a unique solution u^h of (24). Also, there are some error estimations of the Galerkin method, say, see Strang, Fix [17] and Aziz [1]. However, if there is an odd-derivative term of the bilinear form with significant size, the Galerkin method is usually unsatisfactory. The essential reason is that the approximation in linear functional space is not good enough in this singular case. Probably, that is one way to overcome the difficulty is to extend the trial solution space to a nonlinear subset.

Now, suppose that S^h which is a closed subset of H_1 has the same number of freedoms with V^h and that there exists a element $u^h \epsilon S^h$ such that (24) still holds true. Being (23), subtraction yields the following Lemma.

Lemma 6: *For any subset of H_1, if there exists an element $u^h \epsilon S^h$ which satisfies the relation (24), then with respect to the energy inner product, u^h is the projection of u onto S^h, or, the error $u - u^h$ is orthogonal to V^h*

$$a(u - u^h, v^h\ ) = 0 \qquad \text{for all } v^h \epsilon V^h \tag{27}$$

Let the notation u_I denote an interpolation of any $u \epsilon H_1$ in the subspace V^h. Since for any $u_J \epsilon S^h$,

$$a(u - u^h, u - u^h) = a(u - u^h, u - u_J) + a(u - u^h, u_J\text{-}u^h)$$

being (27),

$$a(u - u^h, u_J\text{-}u^h) = a(u - u^h, (u_J\text{-}u^h)\text{-}(u_J\text{-}u^h)_I)$$

or

$$a(u - u^h, u_J\text{-}u^h) = a(u - u^h,(u - u^h)\text{-}(u - u^h)_I) - a(u - u^h, (u - u_J)\text{-}(u - u_J)_I)$$

So, *from the inequalities of* Theorem 5

$$C_2||u - u^h||^2_{H_1} \leq C_1||u - u^h||_{H_1}\{||u - u_J||_{H_2} + ||(u_J-u^h)-(u_J-u^h)_I||_{H_2}\}$$

Therefore, we proved following error estimations.

Theorem 7: *Suppose the conditions in Theorem 5 hold as well as (24), then on the closed nonlinear set S^h the approximation solution u^h of (27), if it exists, has following estimates*

$$||u - u^h||_{H_1} \leq \frac{C_1}{C_2}\inf_{w\in V^h} ||u - u^h - w||_{H_2} \tag{28}$$

$$||u - u^h||_{H_1} \leq \frac{C_1}{C_2}\inf_{u_J\in S^h} \{||u - u_J||_{H_2} + ||(u_J-u^h)-(u_J-u^h)_I||_{H_2}\}$$

or

$$||u - u^h||_{H_1} \leq \frac{C_1}{C_2}\{\ ||(u-u^h) - (u-u^h)_I||_{H_2} + \inf_{u_J\in S^h} \{||u-u_J||_{H_2} + ||(u-u_J) - (u-u_J)_I||_{H_2}\} \tag{29}$$

Corollary 1. If the subset S^h coincides with the subspace V^h, then

$$||u - u^h||_{H_1} \leq \frac{C_1}{C_2}\inf_{u_J\in S^h} ||u - u_J||_{H_2} \tag{30}$$

(30) is just the result of the usual Galerkin method. Hence, (24) above is just a generalization of the Galerkin method.

Corolla y 2. Let u_{Jh} be an interpolation of u on the subset S^h, then

$$||u - u^h||_{H_1} \leq \frac{C_1}{C_2}\{||u - u_{Jh}||_{H_2} + ||(u - u^h)-(u - u^h)_I||_{H_2} + + ||(u - u_{Jh})-(u - u_{Jh})_I||_{H_2}\} \tag{31}$$

The bounds (29) - (31) will play a central role in error analysis . It is clear that the subset S^h should be so chosen as it can tends a denumerable dense set, as h tends zero, in the true solution space H_1, as well as V^h in H_2. In this case the limiting behaviors of the error in energy norm as h-->0 depends mainly on the approximation ability of the subset S^h.

Now we turn to discuss existence and uniqueness of solution of (24) briefly. It was considered in section 2 for self-adjoint a(u,v). When a(u,v) is not self-adjoint, in terms of variational principles, the weak solution (24) is equivalent to find the stationary point for the bilinear a(u,v) on S^h x V^h

where $S \epsilon H_1$ and $V \epsilon H_2$ are a closed approximation subset and a subspace with same finite parameters respectively. Because the existence of the stationary point in the whole space H_1 is assured by theorem 5, hence, from geometric intuition, there is a stationary point in the sense of (24) for sufficient small h, at least. Besides, if there is unique stationary point of (24) when the subspace S^h coincides with the subspace V^h, then, the stationary point still exists if the subset S^h is 'very close' to the subspace V^h. In general, we have

Theorem 8: *Suppose there exists a subspace SL^h with a basis $\{s_j\}$ in which the linear system (24) has unique solution and T is a map from the subset S^h to the subspace SL^h such that for a basis $\{v_j\}$ of the test subspace V^h*

$$\rho\,(A^{-1}J(G)\,) < 1 \tag{32}$$

where the notation ρ denotes the spectral radius of a matrix, A defined in (26), and J(G) is the Jacobi Matrix of the vector G defined

$$G = (\ a(u^h - Tu^h),\ v_j\ ),$$

then the nonlinear system (27) exists unique solution.

Proof: Let $Tu^h = \Sigma\ q_j s_j$, since $a(u^h,v_j) = a(Tu^h,v_j) - a(u^h - Tu^h,v_j)$, from (24), (27) becomes $\Sigma\ a(s_i,v_j) = (f,v_j) + G_j$, In matrix form it can be written as

$$A\ q = d + G(q). \tag{33}$$

Using the following 'simple' iterative procedure

$$A\ q^{(0)} = d,$$

$$A\ q^{(k)} = d + G(q^{(k-1)}) \tag{34}$$

which is a contraction mapping if the condition (32) is satisfied. Q.E.D.

Remark: (33) is very useful not only for proving existence of the solution, but also for computing.

For practical view, hence, the first problem for using the generalized Galerkin method is to construct an adequate nonlinear approximation subset as S^h above.

4. 'Sub-linear' approximation and interpolation

Let T(u) be a real operator of u, where u(x) be a real function defined a given vector space X and belong to a space S, T(u) belong to S, too.

Definition 9: An operator T(u) is called positive on the set X, if

$$T(u) > 0 \quad \text{for all } u(x) > 0 \text{ and } x \in X \tag{35}$$

Definition 10: An operator T(u) is called sublinear on the set X if it satisfies two following conditions

(i) Positive-homogeneous

$$T(au) = aT(u) \quad \text{for all } a > 0 \text{ in } R \text{ and } u \in U,\ x \in X \tag{36}$$

(ii) Subadditive

$$T(u+v) \geq T(u) + T(v) \quad \text{for all } u,v \in U \text{ and } x \in X$$

or

$$T(u+v) \leq T(u) + T(v) \quad \text{for all } u,v \in U \text{ and } x \in X. \tag{37}$$

Consider interpolation and approximation using sublinear piecewise positive operator. For simplicity, let the set X = [0,1], and a partition Δ be given

$$\Delta\!: 0 = x_0 < x_1 < \ldots < x_N = 1,\ h_j = x_j - x_{j-1} \tag{38}$$

Particularly, the linear positive operator, defined by Korovkin[13] is sublinear positive.

When

$$B_j(x) \geq 0, \qquad \Sigma\, B_j(x) = 1, \text{ for all x in } [0,1] \tag{39}$$

then

$$T(u) = \Sigma\, u(x_j)\, B_j(x) \tag{40}$$

is positive. As an example, B_j can be chosen as B-spline. Similar, if

$$B(t,x) \geq 0, \text{ for all t,x in } [0,1], \text{ and } \int_0^1 B(t,x)\, dt = 1, \text{ for all x in } [0,1]. \tag{41}$$

then

$$T(u) = \int_0^1 u(t)\, B(t,x)\, dt \tag{42}$$

is positive too.

Lemma 11: *If u_j, $v_j > 0$, $p > 1$, $B_j(x)$ is defined by (39), then*

$$\{\Sigma(u_j+v_j)^p B_j(x)\}^{1/p} \leq \{\Sigma(u_j)^p B_j(x)\}^{1/p} + \{\Sigma(v_j)^p B_j(x)\}^{1/p}. \tag{43}$$

The inequality direction will be opposite if $p < 1$ ($p \neq 0$). In the limit case of $p = 0$, (43) becomes

$$\Pi(u_j+v_j)^{B_j(x)} \geq \Pi(u_j)^{B_j(x)} + \Pi(v_j)^{B_j(x)}. \tag{44}$$

In each case the equality holds true if and only if the two sequences (u) and (v) are proportional.

In fact, the above inequality is just the triangular inequality for the l_p space with weight. It can be easily proved using a classical inequality, e.g., [5]. Hence, the operator

$$T(u;\, p) = \{\ \Sigma\, (u_j)^p\, B_j(x)\ \}^{1/p} \tag{45}$$

is sub-linear positive.

For instance, if we take the basic functions $\{B_j(x)\}$ as B-spline and u_j as an average of u(x) on some nodes near x_j, then (45) becomes a sublinear positive approximation operator of u(x) on [0,1], it wll be a generalization of the well-known Schoenberg approximation.

Consider a kind of piecewise interpolations using the above semi-linear positive operator. For $x_{j-1} \leq x \leq x_j$, let $t = (x - x_{j-1})/ h_j$, $u_0 = u(t)|_{t=0}$, $u_1 = u(t)|_{t=1}$, $p = p_j$, and

$$T(u;\, p) = \{\ u_0^p\, (1-t) + u_1^p\, t\ \}^{1/p}\ (p \neq 0)\quad 0 < t < 1 \qquad — \tag{46}$$

$$T(u;\, 0) = u_0^{(1-t)}\, u_1^t \qquad (p = 0)\quad 0 < t < 1 \tag{47}$$

Obviously T(u; p) is piesewise sublinear and positive.

Theorem 12: *The interpolatory operator T(u; p) is piecewise sublinear and positive, and if $u \in C^3[0,1]$, for $x_{j-1} < x < x_j$, then for $u(x) > 0$ there is a remainder expression*

$$u(x)-T(u;p) = \frac{1}{2}(x-x_{j-1})(x_j-x)(u''-(1-p_j)u'^2/u)|_{x=\xi}+O(h^3).$$

$$(x_{j-1} < \xi < x_j) \tag{48}$$

Furthermore

$$\mathrm{Max}|u(x)-T(u;\, p)| \leq \frac{h^2}{8}\mathrm{Max}|u''-(1-p_j)u'^2/u| + O(h^3). \tag{49}$$

and

$$\mathrm{Max}|u'(x)-T'(u;\, p)| \leq \frac{h}{2}\mathrm{Max}|u''-(1-p_j)u'^2/u| + O(h^2). \tag{50}$$

Proof: Since $T(u)=u$ for $t=0$ and $t=1$, using a well-known technique of error estimates in Lagrange interpolation leads to (48) directly, so

$$u(x) - T(u;\, p) = \frac{1}{2}(x-x_{j-1})(x_j-x)(\, u''-T'')|_{x=\xi}\, ,\ (x_{j-1}<\xi<x_j).$$

from the Taylor expansion $T'' = (1-p_j)u'^2/u + O(h^2)$. Hence, (49) and (50) follow.

In particular, if $p = 1$ everywhere, (46) becomes piecewise linear interpolation. Hence, in general case, now the piecewise linear operator operates on u^p instead of on u itself. In another words, (46) is a generalized means replacing the arithmetic means with weights for the linear case (Jiachang Sun [19]). Furthermore, if we choose the piecewise constant p_j such that

$$p_j = \frac{1}{h_j}(\frac{u_j}{u'_j} - \frac{u_{j-1}}{u'_{j-1}}\ ) \tag{51}$$

the resulting interpolation (46) will have one more order of precision.

Theorem 13: *If $u \in C^4[0,1]$, then for the piesewise interpolation (46) with (51)*

$$\mathrm{Max}|u^{(i)} - T^{(i)}(u)| \leq C_i h^{3-i}\mathrm{Max}|W(u)| + O(h^{4-i}).\ \ i=0,1,2.$$

$$||u - T(u)||_{L^2} \leq Ch^3\mathrm{Max}|W(u)| + O(h^4) \tag{52}$$

where

$$C_0 = \frac{3^{1/2}}{216},\ C_1 = \frac{1}{12},\ C_2 = \frac{1}{2},\ C = \frac{1}{6(210)^{1/2}},\ W(u) = \frac{u'^2}{u}(\frac{u}{u'})''. \tag{53}$$

Proof: Set $j=1$, $p=p_1$. Applying the Taylor expansion yields

$$p-1 = -\frac{uu''}{u'^2} + (\frac{x_1+x_0}{2} - x)\,(\frac{u}{u'})'' + O(h^2),$$

$$(u^{p-1}u')' = u^{p-2}u'^2(\frac{x_1+x_0}{2} - x)\,(\frac{u}{u'})'' + O(h^2),$$

$$(u^{p-1}u')'' = -u^{p-2}u'^2(\frac{u}{u'})'' + O(h).$$

Using these results, a straightforward computation leads to

$$u_1^p \frac{x-x_0}{h} + u_0^p \frac{x_1-x}{h} = u^p + \frac{p}{2}(x_1-x)(x-x_0)(u^{p-1}u')'$$
$$+ \frac{p}{6}(x_1-x)(x-x_0)(x_1-2x+x_0)(u^{p-1}u')'' + O(h^4)$$
$$= u^p + \frac{p}{6}(x_1-x)(x-x_0)(\frac{x_1+x_0}{2} - x)u^{p-2}u'^2(\frac{u}{u'})'' + O(h^4).$$

and

$$\frac{u_1^p-u_0^p}{ph} = u^{p-1}u' + (\frac{x_1+x_0}{2} - x)(u^{p-1}u')' +$$
$$\frac{1}{6}(u^{p-1}u')''\{(x_1-x)^2 - (x_1-x)(x-x_0) + (x_0-x)^2\} + O(h^3)$$
$$= u^{p-1}u'\{1+\frac{u'}{12u}(\frac{u}{u'})''[(x_1-x)^2+(x_0-x)^2-4(x_1-x)(x-x_0)]\} + O(h^3)$$

Hence

$$T(u;p) = \{ u_1^p \frac{x-x_0}{h} + u_0^p \frac{x_1-x}{h} \}^{1/p}$$
$$= u \{ 1 + \frac{1}{6}(x_1-x)(x-x_0)(\frac{x_1+x_0}{2} - x)u^{-2}u'^2(\frac{u}{u'})''\} + O(h^4).$$

Therefore

$$T(u;p) - u(x) = \frac{1}{6}(x_1-x)(x-x_0)(\frac{x_1+x_0}{2} - x)\frac{u'^2}{u}(\frac{u}{u'})'' + O(h^4). \tag{54}$$

Since the function

$$h^{-3}|\frac{1}{6}(x_1-x)(x-x_0)(\frac{x_1+x_0}{2} - x)| = \frac{1}{12}t(1-t)|1-2t|$$

has maximum value $\frac{3^{1/2}}{216}$ at $t=\frac{1}{2} + \frac{3^{1/2}}{6}$ (or $\frac{1}{2} - \frac{3^{1/2}}{6}$), hence we get (52) for $i=0$. Similarly

$$T'(u(x);p) = \frac{u_1^p-u_0^p}{ph}\{u_1^p \frac{x-x_0}{h} + u_0^p \frac{x_1-x}{h}\}^{(1-p)/p}$$
$$= u'\{ 1 + \frac{u'}{12u}(\frac{u}{u'})''[(x_1-x)^2+(x_0-x)^2-4(x_1-x)(x-x_0)] \} + O(h^3)$$

and

$$T''(u(x);p) = (1-p)[\frac{u_1^p-u_0^p}{ph}]^2\{u_1^p \frac{x-x_0}{h} + u_0^p \frac{x_1-x}{h}\}^{(1-2p)/p}$$
$$= u''\{ 1 - (\frac{x_1+x_0}{2} - x)\frac{u'^2}{u}(\frac{u}{u'})'' + O(h^2)\}\{1 +$$
$$\frac{(x_1-x)(x-x_0)}{2hu^p}(u^{p-1}u')'[(x_1-x)u_1^p+(x-x_0)u_0^p] + O(h^3)\}.$$

Hence

$$T'(u(x);p) - u'(x) = \frac{u'^2}{12u}(\frac{u}{u'})''[h^2 - 6(x_1-x)(x-x_0)] + O(h^3), \tag{55}$$

$$T''(u(x);p) - u''(x) = -\left(\frac{x_1+x_0}{2} - x\right)\frac{u'^2}{u}\left(\frac{u}{u'}\right)'' + O(h^2). \tag{56}$$

(55) and (56) lead to the bounds (52) for $i=1,2$, respectively. In order to prove the last estimate of (52), using (54) we find

$$||T - u||^2_{L^2} = \int_0^1[T(u(x)\ ) - u(x)]^2dx \leq \frac{h^6}{36}A^2\int_0^1 t^2(1-t)^2(1-2t)^2dt = \frac{h^6}{36}A^2\frac{1}{210}.$$

Remark: It is interesting that the coefficients C_0,C_1,C_2 in (53) are just as same as C_1,C_2,C_3 respectively in an error estimation which is for the cubic Hermite interpolation in the case of $u \epsilon C^4$. e.g. [23] (theorem 2.21). Hence, it is reasonable to call this kind of piecewise interpolation a quansi-cubic-hermitian.

Corollary. If $u \epsilon C^3$, then the main orders in (52) still keep, however, the constants before the orders are need to change now.

We may extend the interpolation form (46) further. In general, suppose $\{F_j\}$ is a sequence of piesewise one-to-one mappings in the subinterval $[x_{j-1},x_j]$ respectively. For a fixed j, say $j = 1$, let $F=F_j$, we define

$$T(u;p) = F\{tF^{-1}u_1 + (1-t)F^{-1}u_0\} \tag{57}$$

where the notation

$$F^{-1}u_i = F^{-1}u(x)|_{x=x_i}\ (i = 0, 1),\ t = \frac{x-x_0}{h}$$

It means that now the piecewise linear interpolation operator does operate on the map of $F^{-1}u$ instead of on u itself directly. In this sense, (46) is merely an example of (57) where $Fu = u^{1/p}$.

Since $T(u(t);F) = u(t)$ at two ends $t=0$ and $t=1$, it leads to

Lemma 14: *Let u, $F \epsilon C^{k+1}(x_0\ x_1)$, where $k=3$ or 4, $F^{-1}u$ is any one-to-one mapping, say $\frac{dF}{du} > 0$, then*

$$||F^{-1}T(u) - F^{-1}u||_\infty \leq \frac{h^2}{8}\ ||\frac{d^2F^{-1}u(x)}{dx^2}||_\infty + O(h^k),$$

$$||\frac{d}{dx}\{F^{-1}T(u)-F^{-1}u\}||_{\infty} \leq \frac{h}{2}||\frac{d^2F^{-1}u(x)}{dx^2}||_{\infty} + O(h^{k-1}).$$

Theorem 15: *Suppose the hypotheses of Lemma 14 hold, then there are remainder formulas and error estimates*

$$u(x) - T(u;p) = \frac{1}{2}(x-x_{j-1})(x_j-x)(u'' - \frac{d^2F(TF^{-1}u(x))}{dx^2})|_{\xi},$$

$$u(x) - T(u;p) = \frac{1}{2}(x-x_{j-1})(x_j-x)\{u''(\xi) + u'^2\frac{d^2F^{-1}u(x)}{dx^2}(\frac{dF^{-1}u(x)}{dx})^{-3}|_{\xi}(\frac{dF^{-1}u(x)}{dx})^2|_{\eta}\},$$

$$(x_{j-1} < \xi,\eta < x_j) \tag{58}$$

$$||T(u) - u||_{\infty} \leq \frac{h^2}{8}||\frac{d^2F^{-1}u(x)}{dx^2}||_{\infty}\frac{dF(g)}{dg}|_{g=F^{-1}u} + O(h^k),$$

$$||\frac{d}{dx}\{T(u)-u\}||_{\infty} \leq \frac{h}{2}||\frac{d^2F^{-1}u(x)}{dx^2}||_{\infty}\frac{dF(g)}{dg}|_{g=F^{-1}u} + O(h^{k-1}). \tag{59}$$

Proof: The first remainder is obvious. The second one needs differential formulas in implicit form

$$\frac{d^2F(TF^{-1}u(x))}{dx^2} = \frac{d}{dz}(\frac{dF}{dz})|_{z=F^{-1}}(\frac{dTF^{-1}u(x)}{dx})^2 = -\frac{d^2F^{-1}u(x)}{dx^2}(\frac{dF^{-1}u(x)}{dx})^{-3}(\frac{dTF^{-1}u(x)}{dx})^2$$

and the Mean value theorem

$$(\frac{dTF^{-1}u(x)}{dx})^2 = [\frac{F^{-1}u_1-F^{-1}u_0}{h}]^2 = u'(\frac{dF^{-1}u(x)}{dx})^2|_{\eta}.$$

Hence, (58) is proved. From Lemma 14,

$$F^{-1}T(u) \leq F^{-1}u + \frac{h^2}{8}||\frac{d^2F^{-1}u(x)}{dx^2}||_{\infty} + O(h^k)$$

Being monotony increase of F^{-1}

$$T(u) \leq F\{F^{-1}u + \frac{h^2}{8}||\frac{d^2F^{-1}u(x)}{dx^2}||_{\infty} + O(h^k)\}$$

using the Taylor expansion completes the proof of the theorem.

Furthermore

Theorem 16: *Let u, $F \epsilon C^4[x_0, x_1]$, if there exists a $\xi \epsilon (x_0, x_1)$ such that*

$$\frac{d^2F^{-1}u(x)}{dx^2}|_{x=\xi} = 0,$$

then

$$||T(u) - u||_{\infty} \leq \frac{h^3}{8} ||\frac{d^3F^{-1}u(x)}{dx^3}||_{\infty} \frac{dF(g)}{dg}|_{g=F^{-1}u} + O(h^4). \qquad (60)$$

Besides, if $\xi = (x_1+x_0)/2$, the coefficient '8' in the dominator of (60) can be improved by '16'.

Corollory. T(u) = u for all x if and only if

$$\frac{d^2F^{-1}u(x)}{dx^2} = 0,$$

i.e., $u(x) = F(C_0+C_1x)$, where C_0,C_1 -- constant.

Observing that the above piecewise interpolatory functions (46) and (57) all only belong to C^0, nevertheless, we have also designed a piecewise sublinear positive interpolatory function which belongs to C^1 [22].

5. An application to a singular perturbation boundary value problem

Consider the following boundary value problem

$$Lu = -\epsilon u'' + p(x) u' + q(x) u = f(x),$$

$$u(0) = u(1) = 0 \qquad (61)$$

where ϵ is a small positive parameter and $p(x)$,$q(x)$ and $f(x)$ are so sufficient smooth that their derivatives until second order are uniformly bounded for all x in $[0,1]$ and for all $\epsilon > 0$, besides, $p(x) \geq p^* > 0$, $q(x) \geq \max(0, p'(x))$ on $[0,1]$.

Let H_m be Sobolev space of m-order with the norm

$$||u||_m = \{ \int_0^1 \Sigma_{i \leq m} (D^i u)^2 \, dx \}^{1/2}$$

and $a(u,v)$ be the unsymmetric bilinear form

$$a(u,v) = \int_0^1 \{ \epsilon u'v' + pu'v + quv \} \, dx \qquad (62)$$

With these notations the weak solution of (61) can be written as : Find $u \in H^o{}_1[0,1]$ so that

$$a(u,v) = (f,v) \text{ for all } v \in H^o{}_1[0,1] \qquad (63)$$

where $H^o{}_1[0,1] = \{ v | \, v \in H_1[0,1] \text{ and } v(0) = v(1) = 0 \}$

Existence and uniqueness of solutions to (63) follow from Theorem 5 using the following Lemma:

Lemma 17: *[11] There exists a constant $C>0$ which is independent of ϵ such that*

$$|a(u,v)| \leq C \, ||u||_{1,\epsilon} \, ||v||_1 \quad \text{for all } u,v \in H^o{}_1$$

$$|a(u,v)| \leq C \, ||u||_{1,\epsilon} \, ||v||_{1,\epsilon,1/\epsilon} \quad \text{for all } u,v \in H^o{}_1 \qquad (64)$$

and

$$|a(u,u)| \geq C^{-1} \, ||u||_{1,\epsilon}{}^2, \text{ for all } u \in H^o{}_1 \qquad (65)$$

where

$$||u||_{1,\epsilon} = \{\int_0^1 (\epsilon u'^2 + u^2) dx\}^{1/2} \qquad (66)$$

$$||u||_{1,\epsilon,1/\epsilon} = \{\int_0^1 (\epsilon u'^2 + \tfrac{1}{\epsilon} u^2) dx\}^{1/2} \qquad (67)$$

Now we apply the generalized Galerkin method described in Section 3. Because the singularity of the solution u(x) of (61) is only near x=1, the width of the boundary layer in which u(x) has large derivatives is less than k times ϵ, where k is a constant no matter how ϵ is small, and on $[0,1-k\epsilon]$ u(x) and its some first derivatives are uniformly bounded.

Let Δ_h denote a partition of the interval [0,1] into N subintervals $[x_{j-1},x_j]$, j=1,2,...N with $x_0=0,x_N=1$. For convenience, we will consider only the case of uniform mesh : $x_j-x_{j-1} = h$, j = 1,2,...,N. Associated with Δ_h we have two subsets with same freedoms of $H^o{}_1[0,1]$, one is the usual piecewise linear space P^h , another is called $SP_1{}^h$ which is defined by that if $u_s{}^h(x) \epsilon SP_1{}^h$, then for $x_{j-1} \leq x \leq x_j$, $t = (x - x_{j-1})/h$,

$$u_s{}^h(x) = \{ \begin{array}{ll} u_{j-1}(1-t) + u_jt & \text{if } |u_j - u_{j-1}|/h < d1 \\ (u_{j-1}+c)\{(u_j+c)/(u_{j-1}+c)\}^t - c & \text{Otherwise} \end{array} \qquad (68)$$

where c is a parameter to be such chosen that it makes the formula to be well defined and to get better approximation for the special problem, d1 is a controllable constant.

For a fixed u(x), the interval [0,1] now divides into two subintervals : $[0,1] = I_r + I_s$, where I_r will be be called regular on which the first derivative of u(x) is bounded by a control number, I_s - singular subinterval in which u'(x) could be very large.

Being Theorem 12, for fixed c and d1, $SP_1{}^h$ consisted by all admissible elements of (68) is a sublinear set of $H^o{}_1$, it is differs from the corresponding linear space V^h only where the element has large first derivative.

Let $\{v_j\}$ be the 'roof' basis of the test function space

$$v_j^h(x) = \begin{cases} (x - x_{j-1})/h & x_{j-1} \leq x \leq x_j \\ & (\, j = 1,2,\ldots,N-1 \,) \\ (x_{j+1} - x)/h & x_j \leq x \leq x_{j+1} \end{cases} \tag{69}$$

For the sake of simplity we first suppose that the coefficients p and q in (61) are constant. In order to get the integration (62) we need the following Lemma which can be convinced by part integration

Lemma 18: *For ab > 0,*

$$I_0 = \int_0^1 a^{1-t} b^t dt = \frac{b - a}{\mathrm{Log}(b/a)},$$
$$I_k = \int_0^1 a^{1-t} b^t t^k dt = I_0 \frac{b - kI_{k-1}}{b - a}, \quad k = 1,2,\ldots \tag{70}$$

In particular

$$I_1 = \frac{1}{\mathrm{Log}(b/a)}\{b - \frac{b - a}{\mathrm{Log}(b/a)}\}.$$

There are some inequalities in [21] about I_0 and I_1 which will be used later:

Lemma 19: *Suppose a,b > 0, then*

$$(ab)^{1/2} \leq I_0 \leq \frac{a+b}{2},$$
$$\frac{1}{2}(ab)^{1/2} \min(1, a^{-1/4} b^{1/4}) \leq I_1 \leq \frac{a+b}{4} \max(1, \frac{b+(ab)^{1/2}}{b + a}). \tag{71}$$

with '=' iff a = b.

The corresponding integral of linear interpolation to I_k is

$$LI_k = \int_0^1 [a(1-t)+bt] t^k dt = \frac{a + (k+1)b}{(k+1)(k+2)}.$$

Therefore we have estimates

$$0 \leq LI_0 - I_0 \leq \frac{1}{2}(b^{1/2}-a^{1/2})^2,$$
$$\frac{1}{12}(b^{1/2}-a^{1/2})(b^{1/2}-2a^{1/2}) \leq LI_1 - I_1 \leq \frac{1}{6}(b^{1/2}-a^{1/2})(2b^{1/2}-a^{1/2}),\ (b \geq a > 0),$$
$$-\frac{1}{12}(a-b) \leq LI_1 - I_1 \leq \frac{1}{6}(a^{1/4}-b^{1/4})(a^{3/4}+a^{1/2}b^{1/4}+a^{1/4}b^{1/2}- 2b^{3/4}),\ (a \geq b > 0),$$
$$\sup_{0<a,b\leq 1} |LI_0 - I_0| = \frac{1}{2},\ \sup_{0<a,b\leq 1} |LI_1 - I_1| = \frac{1}{3}. \tag{72}$$

Integrating (62) from x_{j-1} to x_j yields

$$a(u_s{}^h, v_{j-}{}^h) =$$

$$\int_0^1 (\frac{\epsilon}{h}+pt)(c+u_{j-1})^{1-t}(c+u_j)^t \mathrm{Log}\frac{c+u_j}{c+u_{j-1}}dt + hq \int_0^1 \{(c+u_{j-1})^{1-t}(c+u_j)^t - c\}tdt$$

$$= \frac{\epsilon}{h}(u_j-u_{j-1}) + p\{c + u_j - \frac{u_j-u_{j-1}}{\mathrm{Log}((c+u_j)/(c+u_{j-1}))}\} +$$

$$hq\{-\frac{c}{2} + \frac{u_j + c}{\mathrm{Log}((c+u_j)/(c+u_{j-1}))} - \frac{u_j-u_{j-1}}{(\mathrm{Log}((c+u_{j-1})/(c+u_j)))^2}\}.$$

Similarly, integrating (62) from x_j to x_{j+1}

$$a(u_s{}^h, v_{j+}{}^h) =$$

$$\int_0^1 (-\frac{\epsilon}{h}+p(1-t))(c+u_j)^{1-t}(c+u_{j+1})^t \mathrm{Log}\frac{c+u_{j+1}}{c+u_j}dt + hq \int_0^1 \{(c+u_j)^{1-t}(c+u_{j+1})^t - c\}tdt$$

$$= \frac{\epsilon}{h}(u_j-u_{j+1}) - p\{(c + u_j) - \frac{u_{j+1}-u_j}{\mathrm{Log}((c+u_{j+1})/(c+u_j))}\} +$$

$$hq_j\{-\frac{c}{2} + \frac{u_j + c}{\mathrm{Log}((c+u_j)/(c+u_{j+1}))} - \frac{u_j-u_{j+1}}{(\mathrm{Log}((c+u_{j+1})/(c+u_j)))^2}\}.$$

For $[x_{j-1}, x_j] \in I_r$, a straightforward computation yields

$$a(u_s{}^h, v_j{}^h) = \frac{\epsilon}{h}[2u_j-u_{j-1}-u_{j+1}] + \frac{1}{2}p(u_{j+1}-u_{j-1}) + \frac{h}{6}q(u_{j+1}+4u_j+u_{j-1}) \tag{73}$$

and for $[x_{j-1}, x_j] \in I_s$

$$a(u_s{}^h, v_j{}^h) = a(u_s{}^h, v_{j-}{}^h) + a(u_s{}^h, v_{j+}{}^h)$$

$$= \frac{\epsilon}{h}(2u_j-u_{j-1}-u_{j+1}) + p\{\frac{u_{j+1} - u_j}{\mathrm{Log}((c+u_{j+1})/(c+u_j))} - \frac{u_j - u_{j-1}}{\mathrm{Log}((c+u_j)/(c+u_{j-1}))}\} +$$

$$qh\{(c+u_j)[\frac{1}{\mathrm{Log}((c+u_j)/(c+u_{j+1}))} + \frac{1}{\mathrm{Log}((c+u_j)/(c+u_{j-1}))}]$$

$$- c - \frac{u_j - u_{j+1}}{(\mathrm{Log}((c+u_{j+1})/(c+u_j)))^2} - \frac{u_j - u_{j-1}}{(\mathrm{Log}((c+u_{j-1})/(c+u_j)))^2}\} \tag{74}$$

or

$$a(u_s^{\,h},v_j^{\,h}) = \frac{\epsilon}{h}(2u_j-u_{j-1}-u_{j+1}) + \frac{1}{2}p(u_{j+1}-u_{j-1}) + \frac{h}{6}q(u_{j+1}+4u_j+u_{j-1}) + g_j \tag{75}$$

where g_j is the difference of the right parts between (73) and (74).

Denote $\frac{\epsilon}{h} = \alpha$, substituting (73) and (75) into the generalized Galerkin method, i.e.

$$a(u_s^{\,h}, v_j^{\,h}) = (f,v_j^{\,h}) \qquad \text{for } j = 1,2,\ldots N\text{-}1 \tag{76}$$

leads to

$$L_hU^h = \begin{cases} (f,v_j^{\,h}) & \text{if } j \in I_r \\ (f,v_j^{\,h}) - g(U^h_{\ j-1}, U^h_{\ j}, U^h_{\ j+1}) & \text{if } j \in I_s \end{cases} \tag{77}$$

where the left side

$$L_hU^h = -(\alpha + \frac{p}{2} - \frac{h}{6}q)U^h_{\ j-1} + (2\alpha + \frac{2h}{3}q)U^h_{\ j} - (\alpha - \frac{p}{2} - \frac{h}{6}q)U^h_{\ j-1}$$

which is exactly the same to the scheme from usual piecewise linear subspace.

With matrix form it can be written as the special form as (33).

$$A\,U = d + Q(U) \tag{78}$$

where A is a tridiagonal matrix $A = (\ \alpha_{i,j}\ )$

$$\alpha_{i,j} = \begin{cases} -(\alpha + \frac{p}{2} - \frac{h}{6}q) & i>j \\ 2\alpha + \frac{2h}{3}q & i=j \\ -(\alpha - \frac{p}{2} - \frac{h}{6}q) & i<j \end{cases} \tag{79}$$

Denote the determinants of the first j and the last N-i principal determinants of A by D_j and $D_{i,N-1}$, respectively, set

$$\beta_n = \frac{D_{n-1}}{D_n},\ \beta_{j,N-1} = \frac{D_{j+1,N-1}}{D_{j,N-1}}.$$

Due to the recursion formula

$$\beta_n = \{2\alpha + \frac{2h}{3}q - (\alpha - \frac{p}{2} - \frac{h}{6}q)(\alpha + \frac{p}{2} - \frac{h}{6}q)\beta_n\}^{-1},$$

therefore

Lemma 20: *If* $\alpha = \frac{\epsilon}{h} \geq \frac{p}{2} + \frac{h}{6}q$, *then*

$$\beta_n \leq \{\alpha + \frac{p}{2} - \frac{h}{6}q\}^{-1}, \quad \text{for all } n < N\text{-}1.$$

$$\beta_{n,N-1} \leq \{\alpha + \frac{p}{2} - \frac{h}{6}q\}^{-1}, \quad \text{for all } n < N\text{-}1. \quad — \tag{80}$$

Meanwhile, we have

Theorem 21: *When*

$$\alpha = \frac{\epsilon}{h} \geq \frac{p}{2} + \frac{h}{6}q, \tag{81}$$

$A^{-1} = (\alpha^{-1}_{i,j})$ is a good discrete Green function in the following sense: A^{-1} is non-negative and

$$\alpha^{-1}_{i,j} \geq \alpha^{-1}_{i,j-1} \text{ if } i \geq j \quad \text{or} \quad \leq \alpha^{-1}_{i,j-1} \text{ if } i < j \tag{82}$$

Proof: In fact, in this case $A^{-1} = (\alpha^{-1}_{i,j})$

$$\alpha^{-1}_{i,j} = \begin{cases} (\alpha + \frac{p}{2} - \frac{h}{6}q)^{i-j} D_{j-1} D_{N-1-i}/D_{N-1} & \text{if } i \geq j \\ (\alpha - \frac{p}{2} - \frac{h}{6}q)^{i-j} D_{i-1} D_{N-1-j}/D_{N-1} & \text{if } i \leq j \end{cases} \tag{83}$$

Since A^{-1} exists, (78) can be written as

$$U = A^{-1}(d + Q(U)) \tag{84}$$

Now we look for an estimate of $||A^{-1}J(Q(U))||$, where $J(Q)$ is the Jacobi matrix of Q. The main idea of the derivation is the same to our another paper [21] in which the scheme based on a second order semi-linear numerical differentiation formulas has the same form (84) with slight different A and Q. Thus we only need to explain the outline of proofs which are different here. First, we prove that the following important 'semi-linearity' of Q defined in [21]:

Lemma 22: *For $Q(u)$ defined by the difference between the linear scheme (78) and the semilinear scheme (74), there exits the following identities:*

$$\{J(Q(u))(u+c)\}_j = \{Q(u)\}_j, \quad \text{if } j < N\text{-}1. \tag{85}$$

Proof: Denote the nonlinear term of third term in the right of (74) by

$$F(u_{-1},u_0,u_1) = c + u_0 - (c+u_0)\{\frac{1}{\mathrm{Log}((c+u_0)/(c+u_{-1}))} + \frac{1}{\mathrm{Log}((c+u_0)/(c+u_1))}\} +$$

$$\frac{u_1 - u_0}{(\mathrm{Log}((c+u_1)/(c+u_0)))^2} + \frac{u_0 - u_{-1}}{(\mathrm{Log}((c+u_0)/(c+u_{-1})))^2}.$$

but

$$\frac{\partial F}{\partial u_{-1}} = -(c+u_0)\frac{1/(c + u_{-1})}{(\mathrm{Log}((c+u_0)/(c+u_{-1})))^2} - \frac{1}{(\mathrm{Log}((c+u_0)/(c+u_{-1})))^2}$$

$$- \frac{2(u_0-u_{-1})/(c+u_{-1})}{(\mathrm{Log}((c+u_0)/(c+u_{-1})))^3}.$$

$$\frac{\partial F}{\partial u_0} = 1 - [\frac{1}{\mathrm{Log}((c+u_0)/(c+u_1))} + \frac{1}{\mathrm{Log}((c+u_0)/(c+u_{-1}))}] +$$

$$(c+u_0)[\frac{1/(c + u_0)}{(\mathrm{Log}((c+u_0)/(c+u_{-1})))^2} + \frac{1/(c + u_0)}{(\mathrm{Log}((c+u_0)/(c+u_1)))^2}]$$

$$+ \frac{1}{(\mathrm{Log}((c+u_0)/(c+u_{-1})))^2} - \frac{1}{(\mathrm{Log}((c+u_0)/(c+u_1)))^2}$$

$$- \frac{2}{c+u_0}[\frac{u_1-u_0}{(\mathrm{Log}((c+u_1)/(c+u_0)))^2} + \frac{u_0-u_{-1}}{(\mathrm{Log}((c+u_0)/(c+u_{-1})))^2}].$$

$$\frac{\partial F}{\partial u_1} = -(c+u_0)\frac{1/(c + u_1)}{(\mathrm{Log}((c+u_0)/(c+u_1)))^2} + \frac{1}{(\mathrm{Log}((c+u_0)/(c+u_1)))^2}$$

$$- \frac{2(u_1-u_0)/(c+u_1)}{(\mathrm{Log}((c+u_0)/(c+u_1)))^3}.$$

Hence

$$\frac{\partial F}{\partial u_{-1}}(c+u_{-1}) + \frac{\partial F}{\partial u_0}(c+u_0) + \frac{\partial F}{\partial u_1}(c+u_1) = F(u_{-1},u_0,u_1).$$

It has been proved in [21] that the second term in the right of (74) satisfies (85), due to the linearity of the 'semi-linear' relation we get (85).

Since the singularity is only near x=1 and the width of the boundary layer is less than $k\epsilon$, using the inequalities (72, (82), (81) and (83), similarly to [21] a straightforward computation yields

$$J(Q(u)\ )(u+c) = \{0,\ldots,0,Q_n,\ldots,Q_{N-2},Q^*_{N-1}\},$$

$$\{A^{-1}J(Q(u)\,)u\}_i = \sigma_i + hq\tau_i$$

where

$$-\frac{3p}{4} < (\alpha + \frac{p}{2} - \frac{h}{6}q)\sigma_i < \frac{p}{2}, - \frac{2}{3}qk\epsilon < (\alpha + \frac{p}{2} - \frac{h}{6}q)hq\tau_i < \frac{2}{3}qk\epsilon.$$

Therefore

Theorem 23: *When the mesh size condition (81), i.e.,*

$$h \leq \frac{2\epsilon}{p}\{\frac{1}{2} + [\frac{1}{4} + \frac{2}{3}(\frac{\epsilon q}{p})^2]^{1/2}\}^{-1} - \frac{2\epsilon}{p}\{1 - \frac{2}{3}(\frac{\epsilon q}{p})^2\} \tag{86}$$

holds as well as

$$k\,\epsilon \leq \frac{3p}{8q} \tag{87}$$

then the mapping $A^{-1}Q(u)$ is contractive, in the meantime the semi-linear system (78) can be solved by the following convergent 'simple' iteration

$$A\ U^{(0)} = d$$

$$A\ U^{(k)} = d + Q(U^{(k-1)})\ \ (k = 1,2,...) \tag{88}$$

Remark: When ϵ is small,in practice, the mesh condition (86) can be simplified by $h < \frac{2\epsilon}{p}$.

Now we consider error estimations. Let u be the true solution of (61), there exist a decomposition [11]

$$u(x) = \gamma\{W(x) + Z(x)\}$$

$$W(x) = e^{-p(1)(1-x)/\epsilon} - x - (1-x)e^{-p(1)/\epsilon} \tag{89}$$

where $\gamma = \lim_{\epsilon \to 0} \lim_{x \to 1} \epsilon u'(x)/p(1)$ is a constant bounded uniformly for all $0<\epsilon<1$, and

$$|Z(x)| \leq C,\ \ |Z'(x)| \leq C,\ \ |Z''(x)| \leq C\{1+\frac{1}{\epsilon}e^{-\beta(1-x)/\epsilon}\}$$

C is a constant independent of ϵ, and $0<\beta\leq p^*$.

Set c in (89) equal to γ which can be found in computing test. We proved in [21] that

Lemma 24: *Let u be the true solution of (61), if h and ϵ are of the same order, then*

$$||u^{(j)}||_\infty = O(h^{-j}),\ \ (j = 1,2,...)$$

$$||u'' - \frac{u'^2}{c+u}||_\infty = O(h^{-1}), \ \ ||\{u'' - \frac{u'^2}{c+u}\}'||_\infty = O(h^{-2}). \tag{90}$$

where $c = \lim_{\epsilon \to 0} \lim_{x \to 1} \epsilon u'(x)/p(1)$.

Being (49) and (50), hence, for the interpolation function $u^h{}_J$ of $u(x)$ in $SP_1{}^h$ we have error bounds

$$||u^h{}_J - u||_\infty = O(h), \ \ ||u'^h{}_J - u'||_\infty = O(1). \tag{91}$$

Moreover, since the width of the boundary layer is the same order of ϵ, keeping h as the same order of ϵ too, it leads to

$$\begin{aligned} &||u^h{}_J - u||_0 = O(h^{3/2}), \quad ||u^h{}_J - u||_1 = O(h^{1/2}), \\ &||u^h{}_J - u||_{1,\epsilon} = O(h), \quad ||u^h{}_J - u||_{1,\epsilon,1/\epsilon} = O(h). \end{aligned} \tag{92}$$

Let H_1 and H_2 be the Hilbert space with the norm (66) and (67), respectively. Using Lemma 17 and (31) yields

$$||u - u^h||_{1,\epsilon} \leq \frac{C_1}{C_2}\{||u - u_{Jh}||_{1,\epsilon,1/\epsilon} + ||(u - u^h)\text{-}(u - u^h)_I||_{1,\epsilon,1/\epsilon} + ||(u - u_{Jh})\text{-}(u - u_{Jh})_I||_{1,\epsilon,1/\epsilon}\}$$

where the subscript I denotes the interpolation in the test space V^h – piecewise linear function subspace. On the right side of the above inequality, the first term is the major one, others are of higher power of h. Hence, being (92), we get the main error estimation for the scheme (76)

Theorem 25: *If the mesh size condition (86) holds, then*

$$||u_s{}^h - u||_{1,\epsilon} = O(h). \tag{93}$$

where coefficients before powers of h are uniformly bounded for all small ϵ satisfying (87).

Applying the Taylor expansion and using the equation (61) itself and (72), substituting the true solution u into the scheme (77) yields

$$L_h u_j = (f,v_j{}^h) + O(h^2) + + \frac{h^3}{12}\{pu^{(3)} + qu''\} + \ldots$$

if $j \in I_r$, and

$$L_h u_j = (f, {v_j}^h) + O(h^2) - g_j(u_{j-1}, u_j, u_{j+1}) + Tr_j(u),$$
$$Tr_j = \frac{h^3}{12}\{pu^{(3)} - p(\frac{u'^2}{c+u})' + 2qu''\} + \ldots$$

Using (90) and noting the fact that the width of the boundary layer in only $k\epsilon$, similar to [21] we get

$$||A^{-1}Tr(x)||_\infty = O(h) \qquad — \qquad (94)$$

furthermore

$$||{u_s}^h - u||_\infty = O(h), \; ||{u'_s}^h - u'||_\infty = O(1).$$

Similarly

$$||{u_s}^h - u||_0 = O(h^{1.5}), \; ||{u_s}^h - u||_1 = O(h^{0.5}).$$

Summarizing the above results, finally we obtain the following theorem of error estimations

Theorem 26: *For small ϵ satisfying (87), when the mesh size condition (86) holds, then the generalized Galerkin method on the subset (76) has one more order of precision than its corresponding scheme of piecewise linear subspace, i.e., in this case we have there exist constants C_0, C_1, C_∞ and C'_∞ which are uniformly bounded for all small ϵ such that*

$$||{u_s}^h - u||_0 \leq C_0 h^{1.5}, \; ||{u_s}^h - u||_1 \leq C_1 h^{0.5},$$
$$||{u_s}^h - u||_\infty \leq C_\infty h, \; ||{u'_s}^h - u'||_\infty \leq C'_\infty. \qquad (95)$$

In the case of the general variable coefficients p and q, it can be proved that the above conclusion still holds true for small ϵ if two extra requires are satisfied:

$$h < \frac{2}{||p||_\infty}\epsilon, \text{ and } \frac{1}{6}(q_{j-1}+4q_j+q_{j+1}) \geq \frac{1}{2h}(p_{j+1}-p_{j-1}), \; j = 1,2,\ldots \qquad (96)$$

the last one is a discrete form for the elliptic condition of $q(x) \geq p'(x)$.

As a matter of fact, we only need to point that, being smoothness of p and q, substituting their piecewise linear interpolations into the integral form (62), (73) becomes

$$a({u_s}^h, {v_j}^h) = \frac{\epsilon}{h}[2u_j - u_{j-1} - u_{j+1}] + \frac{1}{6}[u_{j+1}(2p_j+p_{j+1}) + u_j(p_{j-1}-p_{j+1}) - u_{j-1}(2p_j+p_{j-1})] +$$
$$\frac{h}{12}[u_{j+1}(q_j+q_{j+1}) + u_j(q_{j-1}+6q_j+q_{j+1}) + u_{j-1}(q_j+q_{j-1})] + O(h^2) \qquad (97)$$

The related tridiagonal matrix A = $(\alpha_{i,j})$ in (79) now is

$$\alpha_{i,j} = \left\{ \begin{array}{ll} -[\alpha + \frac{1}{6}(2p_j+p_{j-1}) - \frac{h}{12}(q_j+q_{j-1})], & i=j+1 \\ 2\alpha - \frac{1}{6}(p_{j+1}-p_{j-1}) + \frac{h}{12}(q_{j+1}+6q_j+q_{j-1}), & i=j \\ -[\alpha - \frac{1}{6}(2p_j+p_{j+1}) - \frac{h}{12}(q_j+q_{j+1})], & i=j-1 \end{array} \right. \qquad (98)$$

The rest derivation is similar to [21], we omit it in detail.

For higher order schemes based on the interpolation described in section 4, the similar analysis can be also done.

6. Numerical Results

In this section, we give two examples to show how well the numerical results match the conclusions in Theorem 25 and 26. In the following Tables, the notation $N = 1/h$, SL and L represent the subset scheme (76) and its corresponding linear scheme, respectively, Er(Max) is the maximum error with sign of the discrete solution and it has occupied on the node xM, Er(H 1,eps), Er(H0) and Er(H1) are the approximation values of errors in $H_{1,\epsilon}$,H_0 and H_1, respectively. CPU - the CPU time in terms of seconds. The Fortran program was run in double-precision, on a DEC-System 2060 computer. The iterative error for (88) is equal to 10^{-5}.

EXAMPLE 1. A linear singular perturbation problem with constant coefficients

$$Lu = - \epsilon u'' + u' + (1+\epsilon)u = f(x), \quad \text{in } (0,1)$$

$$u(0) = u(1) = 0$$

where $f(x) = (1+\epsilon)(a-b)x - \epsilon a - b$, $a = 1 + e^{-(1+\epsilon)/\epsilon}$, $b = 1 + e^{-1}$, with true solution (see. Figure 1)

$$u(x) = e^{-(1+\epsilon)(1-x)/\epsilon} + e^{-x} - a + (a-b)x$$

In our case, set the constant $c = 1$ in the scheme (68), see [21].

The results listed in Table 1-4 (or Figure 2-4) show that :

1. The iteration of (88) monotony converges if the ratio $h/\epsilon < 2$, the results match with the theoretical analysis above, the SL- scheme is much better than L-scheme with little more CPU time cost (about 20% for small ϵ) for the same mesh size h.
2. When $2 \leq h/\epsilon \leq 2.25$, the iteration seems still convergent, but oscillation is occupied now, and the error is getting more than the above estimates, CPU time is more, too.
3. If the ratio increases again, the iteration (88) does not converge.
4. For a given level of accuracy, the CPU time costs much less using the SL-scheme than using the L-scheme, and more small ϵ there is, more advantage the SL-scheme has. For instance, given an admissible maximum error at knots ≤ 0.005, their CPU time ratio are about 0.3 : 1.1 and 3 : 15, for $\epsilon = 0.01$ and 0.001, respectively.

EXAMPLE 2. A semi-linear singular perturbation problem

$$Lu = - \epsilon u'' + p(x)u' + q(x)u = f(x,u), \quad \text{in } (0,1)$$

$$u(0) = u(1) = 0$$

where

$$f(x,u) = a - b - (1+\epsilon)\{e^{-x} - u + \frac{c}{u+a-(a-b)x-e^{-x}}\},$$
$$a=1+e^{-(1+\epsilon)/\epsilon},\ b=1+e^{-1}, c = e^{2(1+\epsilon)(1-x)/\epsilon},$$

$$p(x) = 1,\ q(x) = 1+\epsilon$$

with the same solution as example 1.

In the semi-linear case, the advantage of SL scheme over L-scheme is more obvious than in linear case, the results of SL-scheme still match the Theorem 25 and they are much better than L-scheme with same conditions to obtain higher accuracy and save computer time both (see Table 5-7, or Figure. 5-7).

REFERENCES

1. Aziz, A.Z., (ed.) "The mathematical Foundations of The Finite Element With Applications to Partial Differential Equations ". AP, 1972.
2. Babuska, I., Szymczak, W. G., An Error Analysis for the Element Method Applied to Convection Diffusion Problems, Technical Note BN-962, 1981, March. Institute for Physical Science and Technology, University of Maryland.
3. Babuska, I., Szabo,B.A., and Katz,I.N., The p-version of the finite element method. SIAM Journal on Numerical Analysis 18(1981), 515-545.
4. Barrett, J.W., Morton, K.W., Optimal Finite Element Solutions to Diffusion-Convection Problems in One Dimension, Int. J. Num. Meth. Engng. 15(1980), 1457-1474.
5. Beckenbach, E.F., Bellman R., " Inequalities ". Spinger. 1961.
6. Christie, J., Mitchell, A.R., Upwinding of High Order Gelerkin Methods in Conduction-Convection Problems, Ins. J. Num. Meths. in Engng. 12(1978), 1764-1771.
7. Davis, P.J., Interpolation and Approximation, Blaisdell, New York, 1963.
8. Davis, S.F., Flaherty, J.E., An Adaptive Finite Element Method for Initial-Boundary Value Problems for Partial Differential Equations. Technical Report 81-13, 1981, March, Universities Space Research Association.
9. DeGroen, P.P.N., Hemker,P.W., Error Bounds for Exponentially Fitted Galerkin Methods Applied to Stiff Two-point Boundary Value Problems. (Eds. P.W.Hemker and J.J.H.Miller), Academic Press (1979)
10. Griffiths, D.F., Lorentz, J., An Analysis of the Petrov-Galerkin Finite Element Method, Comp. Meth. Appl. Mech. Engng., 14(1978), 39-64.
11. Kellogg, R.B., Han Honde, The Finite Element Method for a Singular Perturbation Problem Using Enriched Subspaces. Technical Note BN-978 Semptember 1981. University of Maryland.
12. Heinrich, J.C., Huyakorn, P.S., Zienkienwicz, O.C., Mitchell, A.R., An "Upwind" Finite Element Scheme for Two-Dimensional Convective Transport Equation, Int. J. Num. Meths. in Engng. 11(1977), 131-143.
13. Korovkin,P.P., "Linear Operator and Approximation Theory ", PHYSIMATH.1959. in Russian.
14. Rice, J.R., "The Approxmation of Functions", Addison-Wesley Publishing Company, 1969.
15. Schultz, M.H., "Spline Analysis", Prentice-Hall, 1973.
16. Schultz, M.H., (ed.) "Elliptic Problem Solvers", Academic Press, 1981.
17. Strang, G., Fix, G.J., "An Analysis of the Finite Element Method", Prentice-Hall, Inc., 1973.
18. Sun Jiachang, The Spline Functions in Local Coordinates and Circular Spline Curve, Mathematicae Sinica, 20(1977), 28-40.
19. Sun Jiachang, Generalizations of the Means and their Inequalities. Department of Mathematics, University of California, Santa Barbara, May, 1981. (to appear).
20. Sun Jiachang, Semi-linear One-step Implicit Schemes for Solving Initial Value Problem

of Ordinary Differential Equations with Steep Gradients. Department of Computer Science, Yale University, Technical Report #215. 1982.

21. Sun Jiachang, Semi-linear Difference Schemes for Singular Perturbation Problems in one dimension. Department of Computer Science, Yale University, Technical Report #216. 1982.

22. Sun Jiachang, Semi-linear Difference Schemes and Semi-linear Finite Element Methods for Singular Perturbation Problems in two dimensions. (in preparing.)

23. Sun Jiachang, "Spline Functions and Computational Geometry", Science Press, 1982. Beijing.

24. Wachspress, E.L., "A Rational Finite Element Basis", Academic Press, 1975.

25. Scarpini, F., Some Nonlinear Complementarity Systems Algorithms and Application to Unilateral Boundary-value Problem. BOLOGNA NICOLA ZANCHELLI EDITORE. 1980.

Table 1-1

SL: h/ε = 1.5

N	xM	Er(Max)	Er(H1,eps)	Er(H0)	Er(H1)	CPU
25	0.920	-0.5837D-02	0.3508D-01	0.1740D-02	0.2146D+00	0.09
50	0.960	-0.6023D-02	0.2079D-01	0.1239D-02	0.1798D+00	0.16
100	0.970	-0.4562D-02	0.1555D-01	0.7205D-03	0.1902D+00	0.50
200	0.985	-0.1810D-02	0.8410D-02	0.2019D-03	0.1456D+00	1.07
400	0.993	-0.1239D-02	0.5495D-02	0.1031D-03	0.1346D+00	2.17
800	0.996	-0.5259D-03	0.3000D-02	0.3223D-04	0.1039D+00	4.47
1600	0.998	-0.3363D-03	0.1823D-02	0.1483D-04	0.8928D-01	8.56

Table 1-2 L: h/ε = 1.5

N	xM	Er(Max)	Er(H1,eps)	Er(H0)	Er(H1)	CPU
25	0.960	-0.8199D-01	0.1216D+00	0.1742D-01	0.7371D+00	0.04
50	0.980	-0.8112D-01	0.1183D+00	0.1223D-01	0.1019D+01	0.06
100	0.990	-0.8070D-01	0.1167D+00	0.8620D-02	0.1425D+01	0.33
200	0.995	-0.8048D-01	0.1159D+00	0.6084D-02	0.2004D+01	0.75
400	0.998	-0.8038D-01	0.1155D+00	0.4299D-02	0.2827D+01	1.53
800	0.999	-0.8033D-01	0.1153D+00	0.3038D-02	0.3992D+01	3.23
1600	0.999	-0.8030D-01	0.1152D+00	0.2148D-02	0.5641D+01	6.54

Table 2 SL: h/ε = 1.75

N	xM	Er(Max)	Er(H1,eps)	Er(H0)	Er(H1)	CPU
25	0.920	-0.1899D-01	0.4066D-01	0.4869D-02	0.2671D+00	0.11
50	0.960	-0.6682D-02	0.2341D-01	0.1254D-02	0.2187D+00	0.25
100	0.980	-0.4199D-02	0.1643D-01	0.6130D-03	0.2171D+00	0.51
200	0.990	-0.1913D-02	0.1043D-01	0.2120D-03	0.1950D+00	1.11
400	0.993	-0.1162D-02	0.6063D-02	0.8863D-04	0.1604D+00	2.48
800	0.996	-0.7638D-03	0.3817D-02	0.4342D-04	0.1428D+00	4.93
1600	0.998	-0.3355D-03	0.2080D-02	0.1374D-04	0.1101D+00	10.03

Table 3-1

SL: h/ε = 2.0

N	xM	Er(Max)	Er(H1,eps)	Er(H0)	Er(H1)	CPU
25	0.920	0.6469D-02	0.2969D-01	0.1423D-02	0.2097D+00	0.18
50	0.960	-0.7583D-02	0.2592D-01	0.1366D-02	0.2588D+00	0.36
100	0.970	0.3605D-02	0.1359D-01	0.3779D-03	0.1921D+00	0.66
200	0.990	-0.1235D-02	0.1032D-01	0.1033D-03	0.2064D+00	1.47
400	0.990	0.1452D-02	0.5153D-02	0.7843D-04	0.1457D+00	2.93
800	0.996	-0.6512D-03	0.4100D-02	0.3331D-04	0.1640D+00	5.81
1600	0.997	0.8136D-03	0.1324D-02	0.3186D-04	0.7485D-01	12.21

Table 3-2 L: h/ε = 2.0

N	xM	Er(Max)	Er(H1,eps)	Er(H0)	Er(H1)	CPU
25	0.960	-0.1366D+00	0.1647D+00	0.2753D-01	0.1149D+01	0.03
50	0.980	-0.1360D+00	0.1614D+00	0.1939D-01	0.1602D+01	0.21
100	0.990	-0.1356D+00	0.1597D+00	0.1369D-01	0.2250D+01	0.37
200	0.995	-0.1355D+00	0.1589D+00	0.9668D-02	0.3171D+01	0.77
400	0.998	-0.1354D+00	0.1584D+00	0.6833D-02	0.4477D+01	1.55
800	0.999	-0.1354D+00	0.1582D+00	0.4830D-02	0.6326D+01	3.35
1600	0.999	-0.1354D+00	0.1581D+00	0.3415D-02	0.8942D+01	6.61

Table 4 SL: h/ε = 2.25

N	xM	Er(Max)	Er(H1,eps)	Er(H0)	Er(H1)	CPU
25	0.920	0.1618D-01	0.2526D-01	0.3461D-02	0.1877D+00	0.17
50	0.940	0.2635D-01	0.2103D-01	0.5016D-02	0.2167D+00	0.48
100	0.980	-0.4804D-02	0.1985D-01	0.5968D-03	0.2976D+00	0.86
200	0.990	-0.1223D-02	0.1143D-01	0.9819D-04	0.2424D+00	1.46
400	0.985	0.1082D-01	0.7342D-02	0.9483D-03	0.2184D+00	3.86
800	0.988	0.2192D-02	0.4271D-02	0.9035D-04	0.1812D+00	8.98
1600	0.997	0.4134D-02	0.4955D-02	0.2310D-03	0.2970D+00	16.28

Table 5-1

SL: $h/\varepsilon = 1.5$

N	xM	Er(Max)	Er(H1,eps)	Er(H0)	Er(H1)	CPU
25	0.920	−0.1930D−01	0.3845D−01	0.4679D−02	0.2337D+00	0.34
50	0.960	−0.5220D−02	0.1843D−01	0.1040D−02	0.1593D+00	0.98
100	0.970	−0.4529D−02	0.1436D−01	0.6564D−03	0.1757D+00	2.42
200	0.985	−0.1742D−02	0.7681D−02	0.1787D−03	0.1330D+00	5.40
400	0.993	−0.1201D−02	0.5141D−02	0.9504D−04	0.1259D+00	11.58
800	0.996	−0.5050D−03	0.2820D−02	0.2971D−04	0.9767D−01	23.48
1600	0.998	−0.3337D−03	0.1738D−02	0.1397D−04	0.8515D−01	47.13

Table 5-2 L: $h/\varepsilon = 1.5$

N	xM	Er(Max)	Er(H1,eps)	Er(H0)	Er(H1)	CPU
25	0.960	−0.8244D−01	0.1217D+00	0.1773D−01	0.7377D+00	2.64
50	0.980	−0.8125D−01	0.1184D+00	0.1231D−01	0.1020D+01	9.63
100	0.990	−0.8100D−01	0.1167D+00	0.8685D−02	0.1426D+01	25.26
200	0.995	−0.8059D−01	0.1159D+00	0.6102D−02	0.2005D+01	56.46
400	0.998	−0.8043D−01	0.1155D+00	0.4305D−02	0.2827D+01	11.24
800	0.999	−0.8035D−01	0.1153D+00	0.3041D−02	0.3992D+01	23.49
1600	0.999	−0.8031D−01	0.1152D+00	0.2149D−02	0.5641D+01	48.78

Table 6-1 SL: $h/\varepsilon = 1.75$

N	xM	Er(Max)	Er(H1,eps)	Er(H0)	Er(H1)	CPU
25	0.920	−0.1968D−01	0.3901D−01	0.4724D−02	0.2562D+00	2.81
50	0.960	−0.6017D−02	0.2080D−01	0.1041D−02	0.1943D+00	1.14
100	0.980	−0.3879D−02	0.1503D−01	0.5494D−03	0.1986D+00	3.26
200	0.985	−0.3431D−02	0.1136D−01	0.3500D−03	0.2124D+00	7.61
400	0.993	−0.1143D−02	0.5662D−02	0.8080D−04	0.1498D+00	14.67
800	0.996	−0.7534D−03	0.3625D−02	0.4076D−04	0.1356D+00	29.78
1600	0.998	−0.3246D−03	0.1988D−02	0.1290D−04	0.1052D+00	65.90

Table 6-2

L: h/ε = 1.75

N	xM	Er(Max)	Er(H1,eps)	Er(H0)	Er(H1)	CPU
25	0.960	-0.1163D+00	0.1453D+00	0.2433D-01	0.9480D+00	2.73
50	0.980	-0.1077D+00	0.1411D+00	0.1569D-01	0.1312D+01	10.29
100	0.990	-0.1104D+00	0.1397D+00	0.1145D-01	0.1842D+01	26.24
200	0.995	-0.1076D+00	0.1387D+00	0.7839D-02	0.2590D+01	56.17
400	0.998	-0.1075D+00	0.1382D+00	0.5538D-02	0.3655D+01	119.92
800	0.999	-0.1073D+00	0.1380D+00	0.3908D-02	0.5163D+01	140.12
1600	0.999	-0.1072D+00	0.1379D+00	0.2760D-02	0.7297D+01	496.64

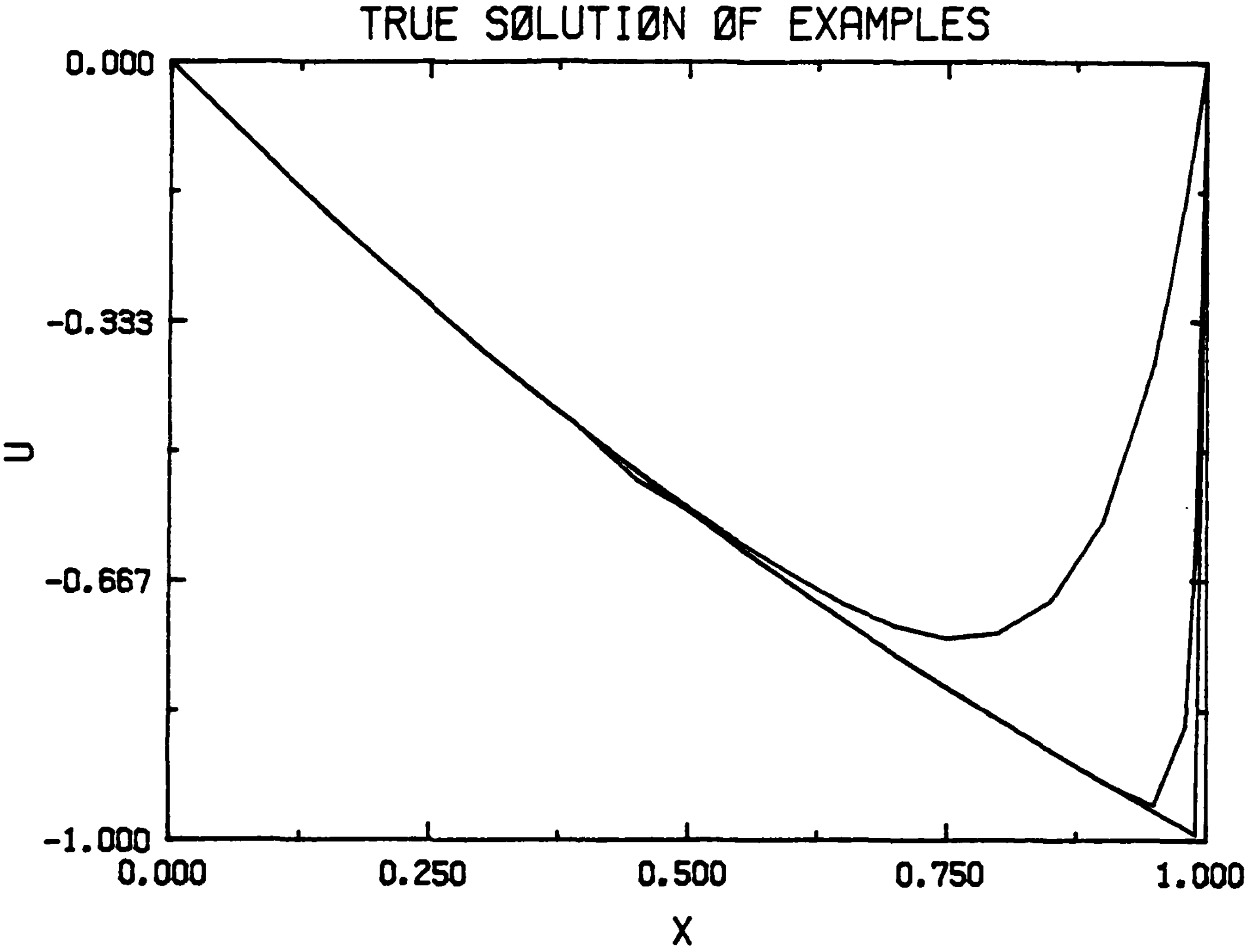
Table 7-1 SL: h/ε = 2.0

N	xM	Er(Max)	Er(H1,eps)	Er(H0)	Er(H1)	CPU
25	0.920	-0.1928D-01	0.3572D-01	0.4404D-02	0.2507D+00	3.42
50	0.960	-0.7551D-02	0.2358D-01	0.1238D-02	0.2354D+00	10.95
100	0.980	-0.3771D-02	0.1592D-01	0.4465D-03	0.2251D+00	25.85
200	0.985	-0.2508D-02	0.8006D-02	0.2353D-03	0.1601D+00	55.86
400	0.993	-0.1181D-02	0.6546D-02	0.8375D-04	0.1851D+00	116.21
800	0.996	-0.7179D-03	0.4017D-02	0.3538D-04	0.1607D+00	236.06

Table 7-2 L: h/ε = 2.0

N	xM	Er(Max)	Er(H1,eps)	Er(H0)	Er(H1)	CPU
25	0.960	-0.1421D+00	0.1653D+00	0.2877D-01	0.1151D+01	3.28
50	0.980	-0.1337D+00	0.1612D+00	0.1907D-01	0.1601D+01	11.33
100	0.990	-0.1255D+00	0.1590D+00	0.1264D-01	0.2242D+01	27.03
200	0.995	-0.1369D+00	0.1590D+00	0.9775D-02	0.3173D+01	58.68
400	0.998	-0.1337D+00	0.1583D+00	0.6748D-02	0.4474D+01	124.43
800	0.999	-0.1323D+00	0.1580D+00	0.4715D-02	0.6318D+01	248.85

Figure 1



1 EPS = 0.1
2 EPS = 0.01
3 EPS = 0.001

Figure 2 EXAMPLE 1: $\frac{h}{\varepsilon} = 1.5$

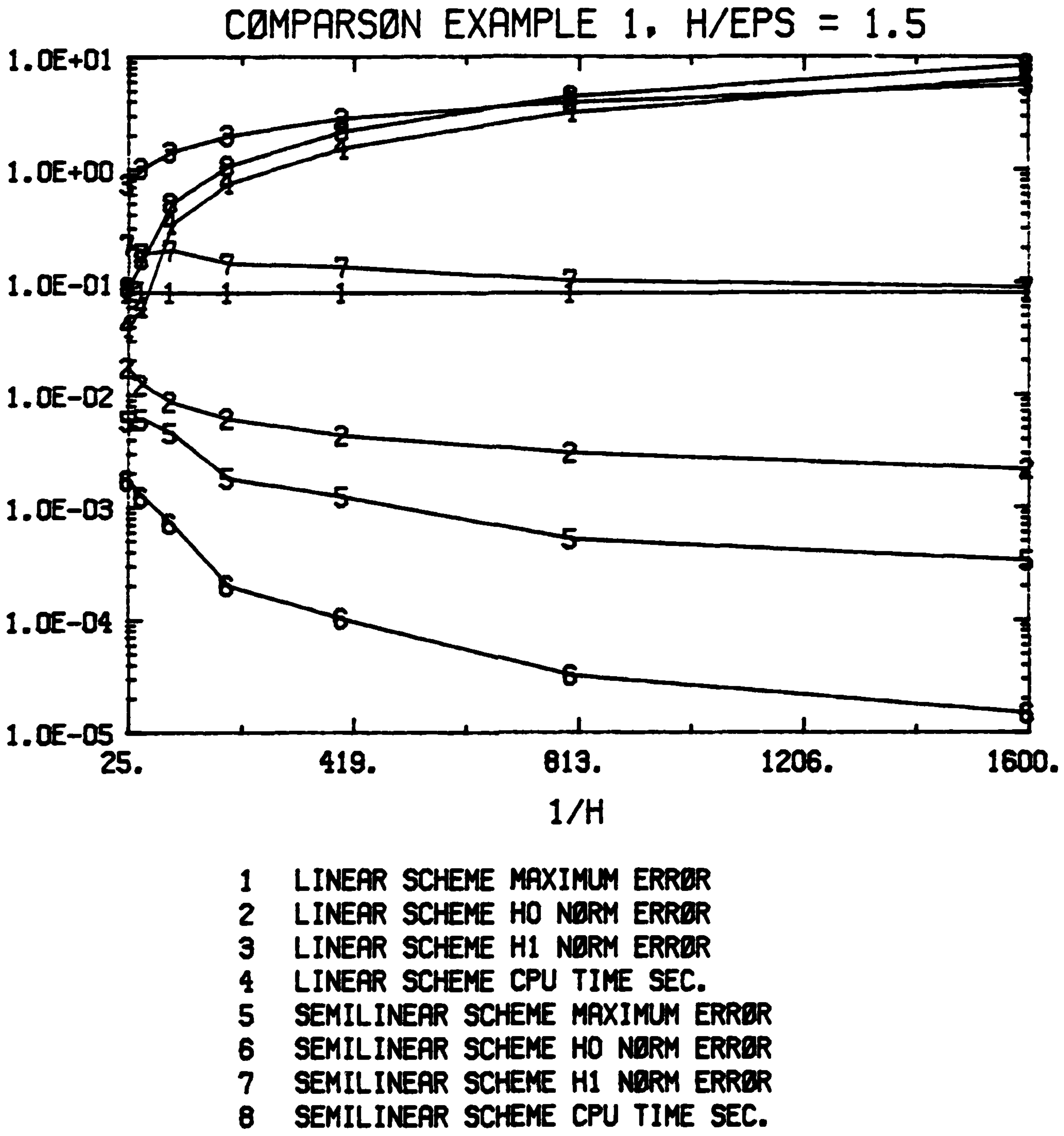


Figure 3

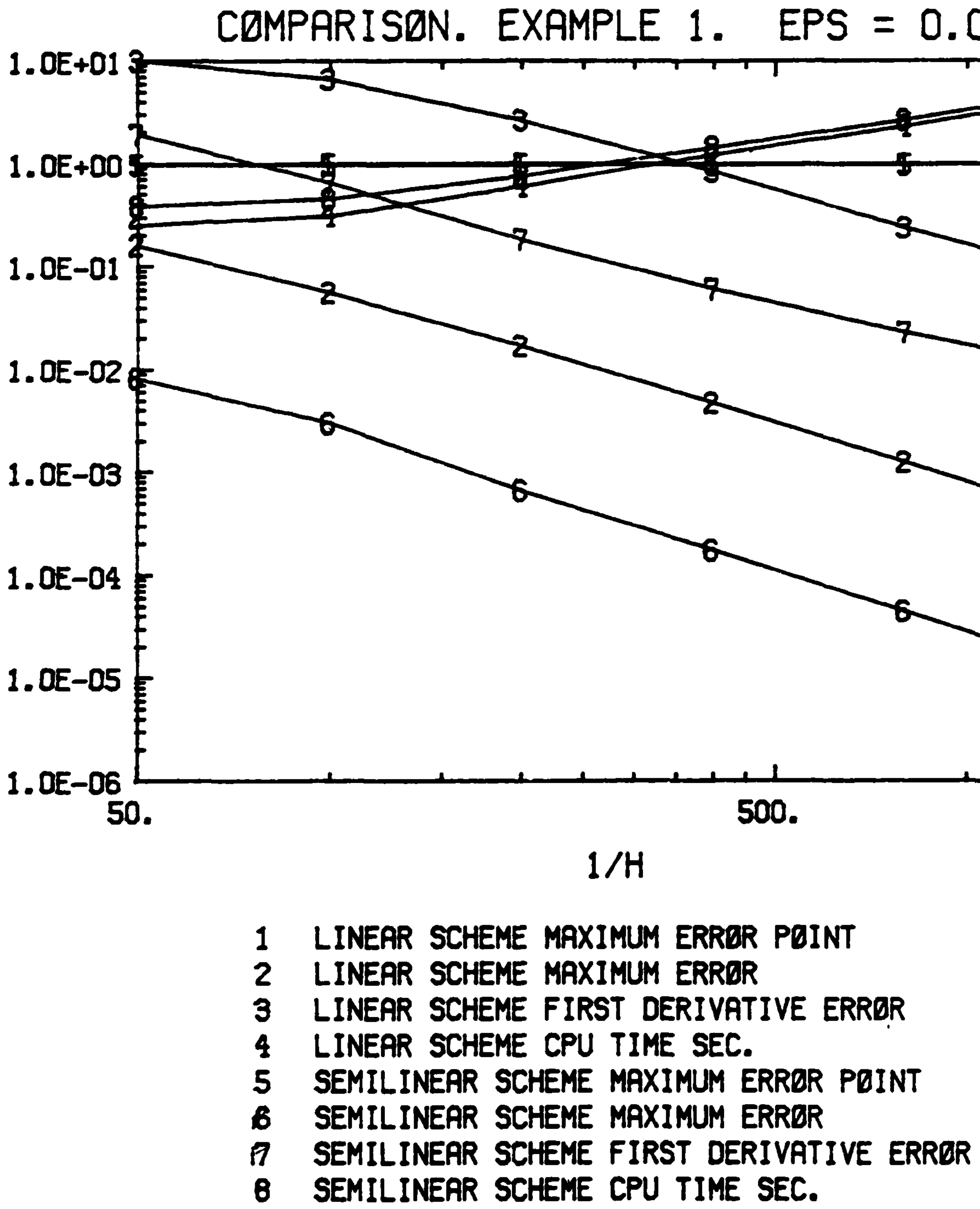

CØMPARISØN. EXAMPLE 1. EPS = 0.01
1.0E+01
1.0E+00
1.0E-01
1.0E-02
1.0E-03
1.0E-04
1.0E-05
1.0E-06
50.
500.
1/H
1 LINEAR SCHEME MAXIMUM ERRØR PØINT
2 LINEAR SCHEME MAXIMUM ERRØR
3 LINEAR SCHEME FIRST DERIVATIVE ERRØR
4 LINEAR SCHEME CPU TIME SEC.
5 SEMILINEAR SCHEME MAXIMUM ERRØR PØINT
6 SEMILINEAR SCHEME MAXIMUM ERRØR
7 SEMILINEAR SCHEME FIRST DERIVATIVE ERRØR
8 SEMILINEAR SCHEME CPU TIME SEC.

Figure 4

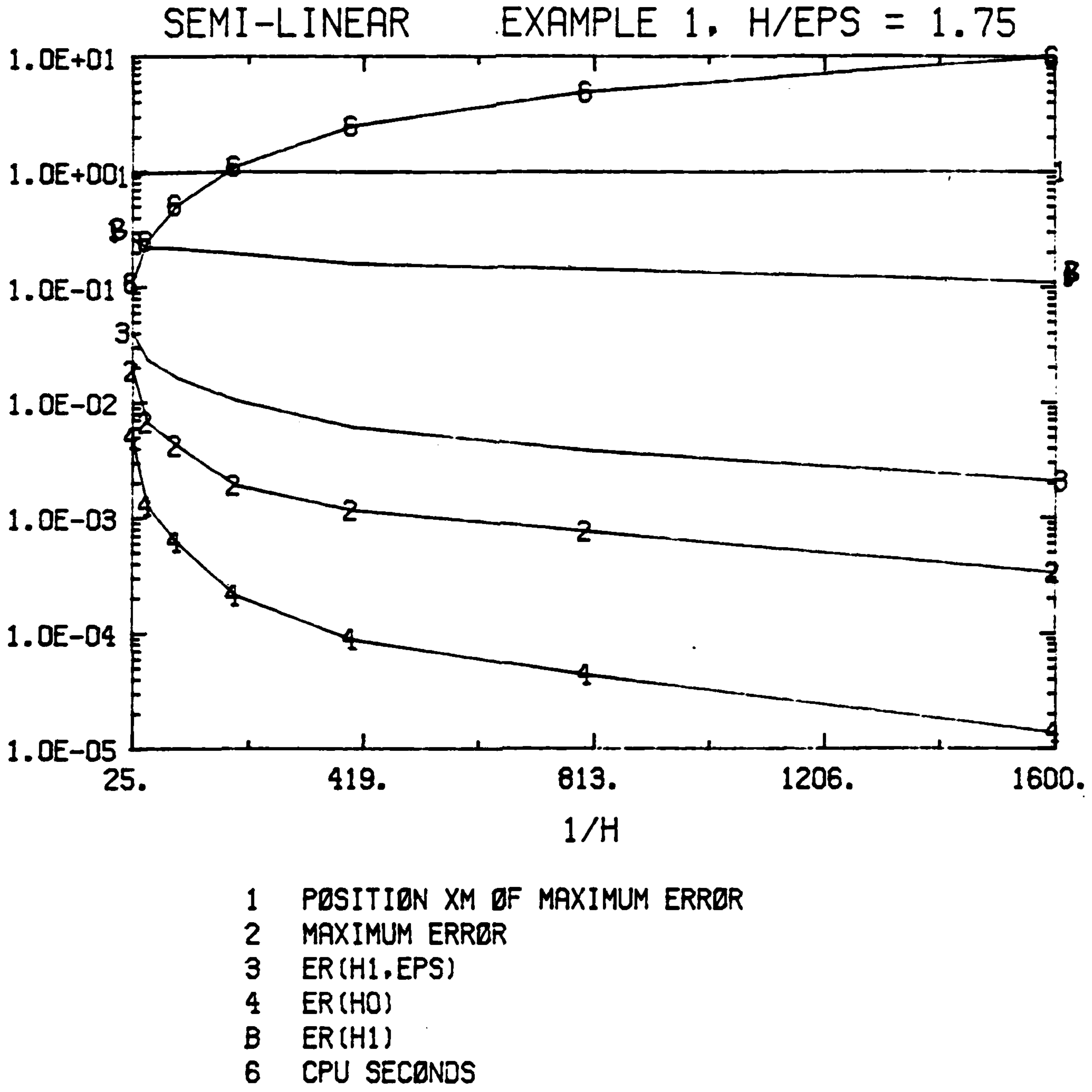

SEMI-LINEAR EXAMPLE 1. H/EPS = 1.75
1.0E+01
1.0E+00
1.0E-01
1.0E-02
1.0E-03
1.0E-04
1.0E-05
25.
419.
813.
1206.
1600.
1/H
1 PØSITIØN XM ØF MAXIMUM ERRØR
2 MAXIMUM ERRØR
3 ER(H1,EPS)
4 ER(H0)
B ER(H1)
6 CPU SECØNDS

Figure 5

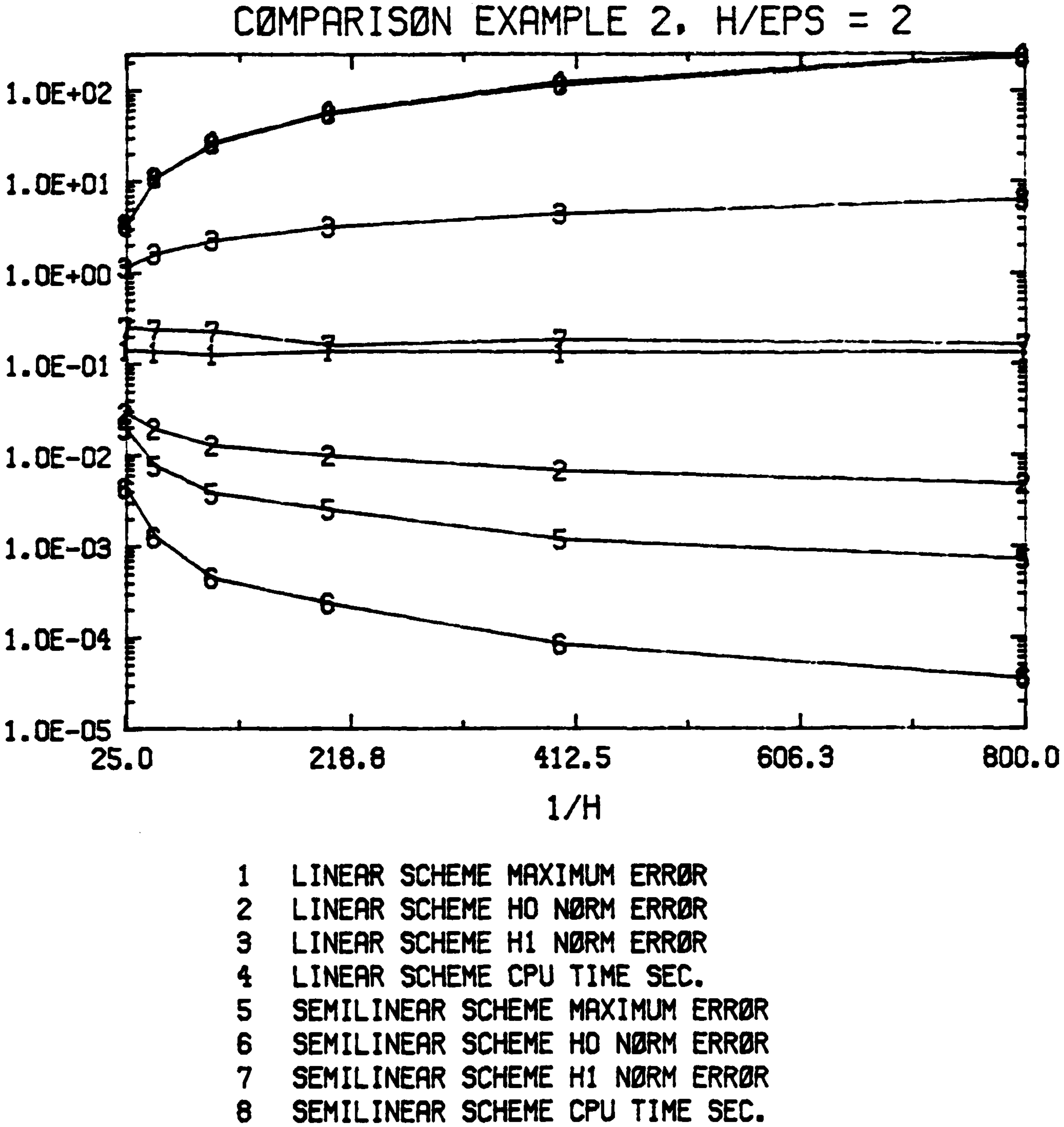

CØMPARISØN EXAMPLE 2, H/EPS = 2
1.0E+02
1.0E+01
1.0E+00
1.0E-01
1.0E-02
1.0E-03
1.0E-04
1.0E-05
25.0
218.8
412.5
606.3
800.0
1/H
1 LINEAR SCHEME MAXIMUM ERRØR
2 LINEAR SCHEME H0 NØRM ERRØR
3 LINEAR SCHEME H1 NØRM ERRØR
4 LINEAR SCHEME CPU TIME SEC.
5 SEMILINEAR SCHEME MAXIMUM ERRØR
6 SEMILINEAR SCHEME H0 NØRM ERRØR
7 SEMILINEAR SCHEME H1 NØRM ERRØR
8 SEMILINEAR SCHEME CPU TIME SEC.

Figure 6

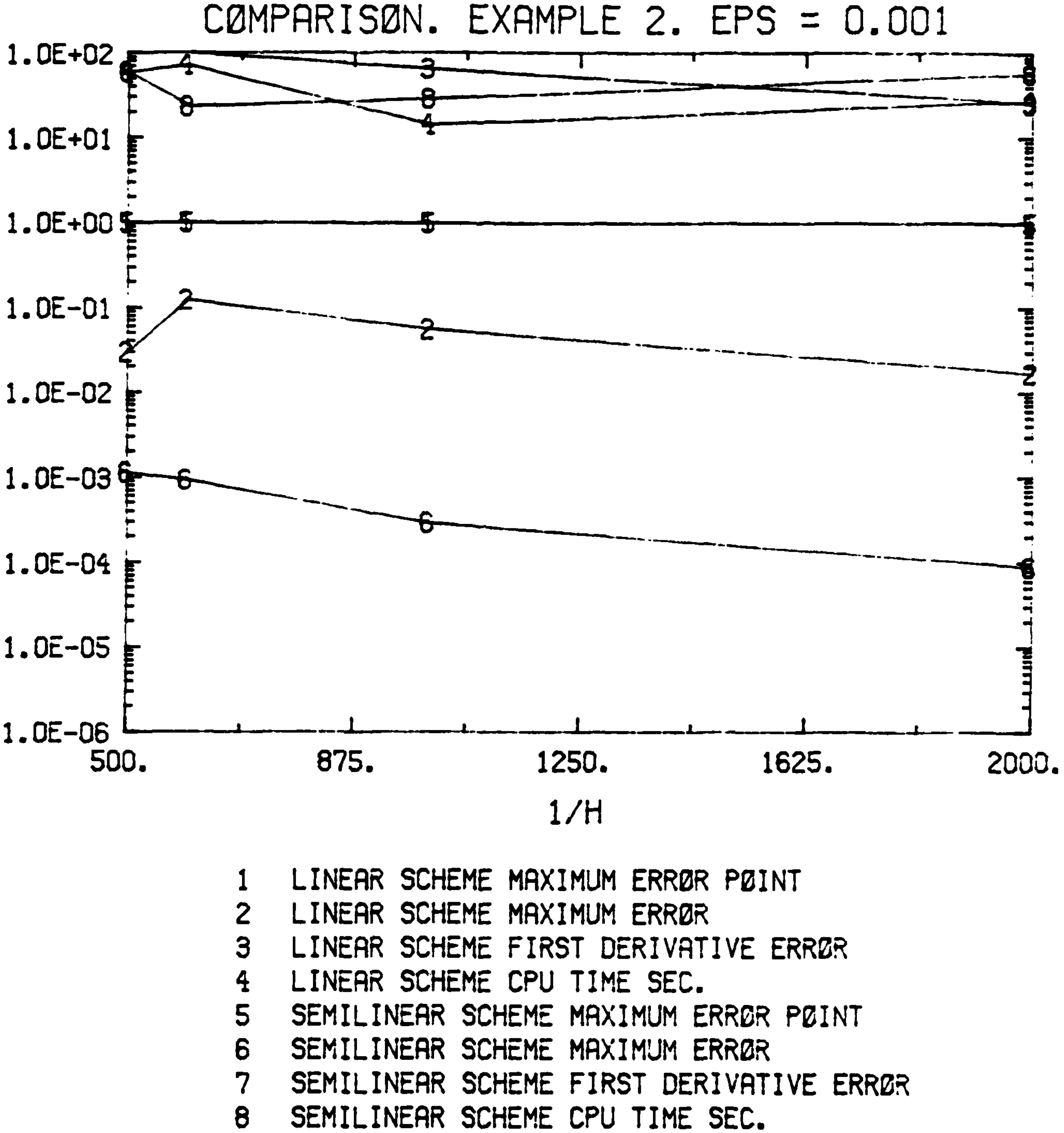
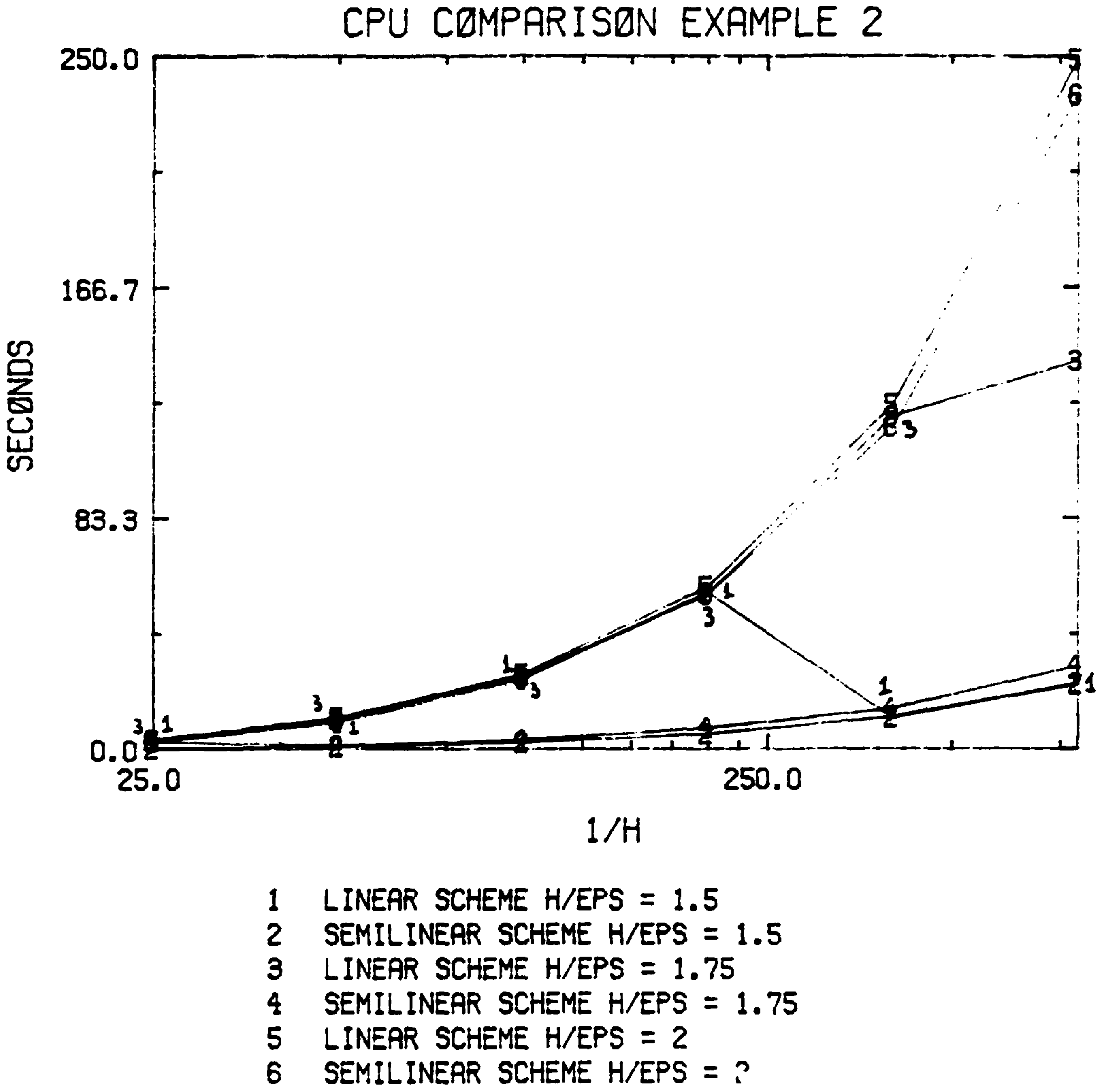

CØMPARISØN. EXAMPLE 2. EPS = 0.001
1.0E+02
1.0E+01
1.0E+00
1.0E-01
1.0E-02
1.0E-03
1.0E-04
1.0E-05
1.0E-06
500.
875.
1250.
1625.
2000.
1/H
1
2
3
4
5
6
8
1 LINEAR SCHEME MAXIMUM ERRØR PØINT
2 LINEAR SCHEME MAXIMUM ERRØR
3 LINEAR SCHEME FIRST DERIVATIVE ERRØR
4 LINEAR SCHEME CPU TIME SEC.
5 SEMILINEAR SCHEME MAXIMUM ERRØR PØINT
6 SEMILINEAR SCHEME MAXIMUM ERRØR
7 SEMILINEAR SCHEME FIRST DERIVATIVE ERRØR
8 SEMILINEAR SCHEME CPU TIME SEC.

Figure 7


CPU CØMPARISØN EXAMPLE 2
250.0
166.7
83.3
0.0
SECØNDS
25.0
250.0
1/H
1 LINEAR SCHEME H/EPS = 1.5
2 SEMILINEAR SCHEME H/EPS = 1.5
3 LINEAR SCHEME H/EPS = 1.75
4 SEMILINEAR SCHEME H/EPS = 1.75
5 LINEAR SCHEME H/EPS = 2
6 SEMILINEAR SCHEME H/EPS = ?