Approximation of Uniformly Convergent Schemes

Paul A. Farrell
Department of Mathematical Sciences,
Kent State University,
Kent, OH 44242, U.S.A.

Abstract: It is well known that only uniformly convergent schemes model the boundary layer behaviour of singularly perturbed problems well on uniform meshes. One criticism of such methods is that the evaluation of the exponential fitting factor involved is time consuming. We shall consider theoretically and computationally how accurately the boundary layer behaviour is modelled, if we approximate the exponential fitting factors by rational expressions.

We shall consider finite difference schemes for the following singularly perturbed ordinary differential equation:

$$\varepsilon u''(x) + a(x)u'(x) - b(x)u(x) = f(x), \qquad 0 < x < 1, \tag{1a}$$

$$u(0) = A, \qquad u(1) = B. \tag{1b}$$

where $a(x), b(x)$ and $f(x)$ are sufficiently smooth, and

$$0 < \underline{a} \leq a(x) \leq \bar{a} \tag{1c}$$

This problem has an exponential boundary layer at $x = 0$. It is well known ([1][7]) that standard finite difference methods such as the centered difference and upwinded schemes do not model the solution well in the region of the boundary layer. Consequently alternate methods, called exponentially fitted methods, were proposed which model the boundary layer behaviour better [6][7][8] and which satisfy a stronger convergence result known as uniform convergence. That is

$$\|u(x_i) - u^h\|_{h,\infty} \leq ch$$

where u is the solution of (1) and u^h is the solution of the difference scheme and here, as elsewhere throughout this paper, c is a generic constant **independent** of h **and** ε. The first of these methods was the Il'in-Allen-Southwell scheme, which can be written in the form

$$L^e u_i^e \equiv \varepsilon \sigma_i^e D_+ D_- u_i^e + a(x_i) D_+ u_i^e - b(x_i) u_i^e = f(x_i), \qquad 1 \leq i \leq N-1 \tag{2a}$$

$$u_0^h = A, \qquad u_N^h = B, \tag{2b}$$

where $h = 1/N$,

$$D_+ D_- h_i^h = \frac{u_{i-1}^h - 2u_i^h + u_{i+1}^h}{h} \quad , \qquad D_+ u_i^h = \frac{u_{i+1}^h - u_i^h}{h},$$

and

$$\sigma_i^e = \sigma(\rho_i) \quad \text{where} \quad \sigma(x) = \frac{x}{e^x - 1} \quad \text{and} \quad \rho_i = \frac{a(x_i)h}{\varepsilon}. \tag{3}$$

One criticism of such schemes is that they all involve the evaluation of at least one exponential at each point of the finite difference mesh. Our aim in this paper is to consider theoretically and computationally how accurately the boundary layer behaviour will be modelled, if we approximate

This paper is in final form and no version of it will be submitted for publication elsewhere

the exponential fitting factor σ_i^e by a rational expression in ρ_i. It should be noted that, since a necessary condition for uniform convergence is that

$$\lim_{h\to 0,i,\rho \text{ fixed}} \sigma_i = \sigma(\rho_i),$$

such schemes cannot be uniformly convergent [1][2][3].

The type of rational expression we wish to consider is $\sigma_i^{m,n} = p_{m,n}(\rho_i)$, where $p_{m,n}(y)$ is the Padé approximation to $\sigma(y)$, whose numerator is a polynomial of degree m and denominator is a polynomial of degree n. Table 1, below gives $p_{m,n}(y)$ for $0 \le m,n \le 3$.
Thus we consider schemes of the form

$$L^{m,n}u^{m,n} \equiv \varepsilon\sigma_i^{m,n} D_+D_-u_i^{m,n} + a(x_i)D_+u_i^{m,n} - b(x_i)u_i^{m,n} = f(x_i), \qquad 1 \le i \le N-1, \tag{4a}$$

$$u_0^{m,n} = A, \qquad u_N^{m,n} = B. \tag{4b}$$

m	n = 0	n = 1	n = 2	n = 3
0	1	$1/(1+\frac{y}{2})$	$1/(1+\frac{y}{2}+\frac{y^2}{6})$	$1/(1+\frac{y}{2}+\frac{y^2}{6}+\frac{y^3}{24})$
1	$1-\frac{y}{2}$	$(1-\frac{y}{3})/(1+\frac{y}{6})$	$(1-\frac{y}{4})/(1+\frac{y}{4}+\frac{y^2}{24})$	$(1-\frac{y}{5})/(1+\frac{3y}{10}+\frac{y^2}{15}+\frac{y^3}{120})$
2	$1-\frac{y}{2}+\frac{y^2}{12}$	$1-\frac{y}{2}+\frac{y^2}{12}$	$(1-\frac{2y}{5}+\frac{y^2}{20})/(1+\frac{y}{10}+\frac{y^2}{60})$	$(1-\frac{y}{3}+\frac{y^2}{30})/(1+\frac{y}{6}+\frac{y^2}{30}+\frac{y^3}{360})$
3	$1-\frac{y}{2}+\frac{y^2}{12}$	$1-\frac{y}{2}+\frac{y^2}{12}$	$(1-\frac{y}{2}+\frac{y^2}{10}-\frac{y^3}{120})/(1+\frac{y^2}{60})$	$(1-\frac{3y}{7}+\frac{y^2}{14}-\frac{y^3}{210})/(1+\frac{y}{14}+\frac{y^2}{42}+\frac{y^3}{840})$

Table 1: Padé approximants $p_{m,n}(y)$ to $\sigma(y)$

The Padé approximation is guaranteed to be a good approximation to $\sigma(y)$ in the region near $y = 0$. However, they do not necessarily exhibit the correct behaviour for $y \gg 0$. It is clear that to preserve the correct behaviour of the difference scheme we require $\varepsilon\sigma_i^{m,n} \to 0$, as $\varepsilon \to 0$, and hence $\rho_i \to \infty$. If $m < n$ then on the contrary $\varepsilon\sigma_i^{m,n} \to \infty$, as $\rho_i \to \infty$. Consequently only the cases where $m \ge n$ are of interest. We should note that the cases (0,1), (0,2) and (1,0) are all well known schemes. In the case (0,1) we have $\varepsilon\sigma_i^{0,1} = \varepsilon$ and hence the scheme is just the standard upwinded scheme. The case (0,2) has $\varepsilon\sigma_i^{0,2} = \varepsilon/(1+\frac{\rho_i}{2})$ which is just Samarskii's scheme [7][2][3]. Finally one can easily show that (1,0) is, in fact, the standard centered difference scheme. It is well known that the centered difference scheme is unstable, for singularly perturbed problems, and can give rise to significant oscillations. This warns us that higher order approximations may not always be more desirable. On the other hand upwinding does not model the boundary layer well, whereas Samarskii's scheme approximates it considerable better. The maximum error arises in general at $x = h$ when $\rho_i \approx 1$. This case is illustrated in Figure 1 for the problem:

$$\varepsilon u''(x) + (1+x^2)u'(x) - (x-.5)^2u(x) = -(e^x + x^2) \tag{5a}$$

$$u(0) = -1, \qquad u(1) = 0. \tag{5b}$$

In [7], it is shown that the error is given by

$$\begin{aligned} \|u(x_i) - u_i^{0,1}\| &\le ch[1+\varepsilon^{-1}\exp(-\bar{a}\varepsilon^{-1}(1-x_i))], & h &\le \varepsilon \\ \|u(x_i) - u_i^{0,1}\| &\le c[h+\exp(-\underline{a}(1-x_i)/(\underline{a}h+\varepsilon))], & h &\ge \varepsilon \end{aligned}$$

and

$$\begin{aligned} \|u(x_i) - u_i^{0,2}\| &\le \tfrac{ch^2}{h+\varepsilon}[1+\varepsilon^{-1}\exp(-\bar{a}\varepsilon^{-1}(1-x_i))], & h &\le \varepsilon \\ \|u(x_i) - u_i^{0,2}\| &\le c[\tfrac{h^2}{h+\varepsilon}+\exp(-\underline{a}(1-x_i)/(\underline{a}h+\varepsilon))], & h &\ge \varepsilon \end{aligned}$$

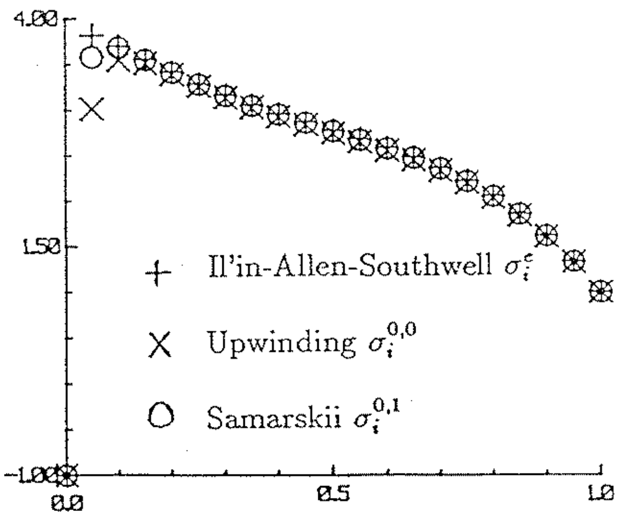

Figure 1: Difference solution of (5)

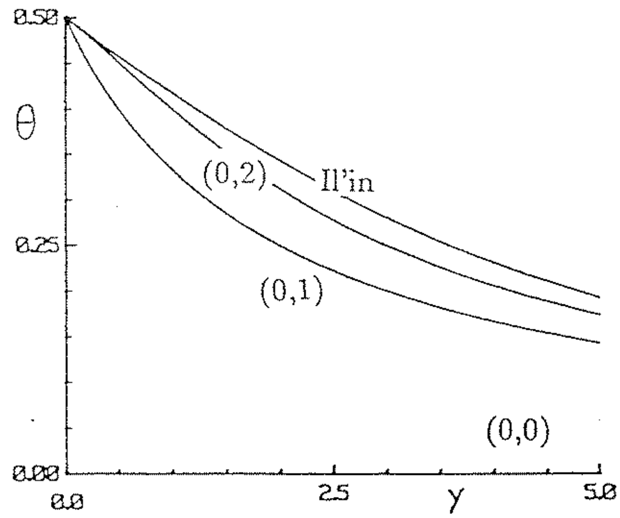

Figure 2 : $\theta^e(y)$ and $\theta^{0,m}(y)$

respectively. It is clear that these estimates reflect the fact that Samarskii's scheme is classically order h^2 as opposed to upwinding, which is classically order h. Similar estimates can be derived for the higher order Padé approximants and indicate that these are also classically order h^2. Numerical evidence confirms this result. Unfortunately it is not, in general, easy to derive any estimate of the quality of approximation for $h \approx \varepsilon$.

We wish to take an alternative approach in which we compare the solution of the Padé scheme to that of the uniformly convergent Il'in-Allen-Southwell scheme.

Let us first note that schemes of the form (2) and (4) can be rewritten in the form

$$L^\theta u_i \equiv \varepsilon D_+ D_- u_i + a(x_i)\left[(1-\theta_i)D_+ u_i + \theta_i D_- u_i\right] + b(x_i)u_i = f(x_i), 1 \le i \le N-1, \tag{5b}$$

where

$$\theta_i = \frac{1}{\rho_i}(1-\sigma_i). \tag{6}$$

By rewriting L^e in the form L^θ it can be shown, using the criteria in [5], that, the Il'in-Allen-Southwell scheme is uniformly $(\infty, 1)$ stable, that is

$$\|u_i\|_{h,\infty} \le c\left\{\|L^e u_i\|_{h,1} + |u_0| + |u_N|\right\}. \tag{7}$$

The same approach is not easily applicable to the approximate schemes, since it involves uniform consistency of the scheme. The following argument gives a sufficient, but certainly not necessary condition, for uniform $(\infty, 1)$ stability of the approximate scheme.

In a similar manner to (7), one can show

$$\max\{\|u_i\|_{h,\infty}, \|D_+ u_i\|_{h,1}\} \le c\left\{\|L^e u_i\|_{h,1} + |u_0| + |u_N|\right\}.$$

Hence

$$\begin{aligned}
\|u_i\|_{h,\infty} &\le c\left\{\|L^e u_i - L^{m,n} u_i\|_{h,1} + \|L^{m,n} u_i\|_{h,1} + |u_0| + |u_N|\right\} \\
&= c\left\{\|a(x_i)\left[(1-\theta_i^e) - (1-\theta_i^{m,n})\right] D_+ u_i^e + a(x_i)[\theta_i^e - \theta_i^{m,n}] D_- u_i^e\|_{h,1}\right. \\
&= \qquad \left. + \|L^{m,n} u_i\|_{h,1} + |u_0| + |u_N|\right\} \\
&= c\left\{\|a(x_i)(\theta_i^{m,n} - \theta_i^e)[D_+ u_i^e - D_- u_i^e]\|_{h,1} + \|L^{m,n} u_i\|_{h,1} + |u_0| + |u_N|\right\} \\
&= c\left\{2\|a(x_i)(\theta_i^{m,n} - \theta_i^e)\|_{h,\infty}\|[D_+ u_i^e\|_{h,1} + \|L^{m,n} u_i\|_{h,1} + |u_0| + |u_N|\right\}.
\end{aligned}$$

Thus, if there exists a δ independent of i, h and ϵ, such that

$$2c\|a(x_i)(\theta_i^{m,n} - \theta_i^e)\|_{h,\infty} \le \delta < 1, \tag{8}$$

we have,

$$\|u_i\|_{h,\infty} \le (c/(1-\delta))\,\{\|L^{m,n}u_i\|_{h,1} + |u_0| + |u_N|\},$$

which is a uniform $(\infty,1)$ stability result for $L^{m,n}$.

Condition (8) essentially says that $L^{m,n}$ must be sufficiently close to L^e. However the condition is difficult to verify and does not guarantee that the solution of the difference scheme will satisfy a maximum principle. Assuming that $b(x) \ge 0$, the condition, in [2][3], under which the solution satisfies a maximum principle is that $\sigma_i \ge 0$ for all i. This would not be satisfied by any of the cases with $n = 1$ or by $\sigma_i^{3,2}$ or $\sigma_i^{3,3}$. On the other hand, all the cases with $n = 0$ or $n = 2$ do satisfy $\sigma_i \ge 0$ and hence satisfy a maximum principle. Consideration of all these factors suggests that one should chose an approximation that is super-diagonal, that is has $m \ge n$, and in addition has $n = 0$ or $n = 2$. Numerical results will confirm this opinion. We shall now prove a result which describes the error of one of these approximate schemes in terms of the error of the Il'in-Allen-Southwell scheme, and a term which measures the degree to which $p_{m,n}(y)$ approximates $\sigma(y)$.

Theorem: Let $L^{m,n}$ be a uniformly $(\infty,1)$ stable operator and let $u(x_i)$ be the solution of (1) and $u_i^{m,n}$ be the solution of (4), for $1 \le i \le N-1$, then

$$\|u(x_i) - u_i^{m,n}\|_{h,\infty} \le c_1 h + c_2 \|\frac{1}{\rho_i}(\sigma_i^e - \sigma_i^{m,n})\|_{h,\infty} \tag{13}$$

where c_1, c_2 are independent of i, h and ε.

Proof: Let $e_i = u_i^e - u_i^{m,n}$. Then, in a similar manner to above we have,

$$L^{m,n}e_i = L^{m,n}u_i^e - L^{m,n}u_i^{m,n} \quad = \quad a(x_i)(\theta_i^{m,n} - \theta_i^e)[D_+u_i^e - D_-u_i^e]$$

Now

$$\|L^{m,n}e_i\|_{h,1} \le \|a(x_i)(\theta_i^{m,n} - \theta_i^e))\|_{h,\infty} \max\left(\|D_+u_i^e\|_{h,1}, \|D_-u_i^e\|_{h,1}\right).$$

Further, we can show, in a manner analogous to that for the continuous problem in [4], using estimates from [5], that

$$\|D_+u_i^e\|_{h,1} \le c \text{ and } \|D_-u_i^e\|_{h,1} \le c$$

and, hence by the $(\infty,1)$ stability of $L^{m,n}$, that

$$\|e_i\|_{h,\infty} \le c\,\{\|L^{m,n}e_i\|_{h,1} + |e_0| + |e_N|\} = c\|L^{m,n}e_i\|_{h,1}.$$

Hence

$$\|u_i^e - u_i^{m,n}\|_{h,\infty} \le c\|\theta_i^{m,n} - \theta_i^e\|_{h,\infty}$$
$$= c\|\frac{1}{\rho_i}(\sigma_i^e - \sigma_i^{m,n})\|.$$

However, as is well known, the Il'in-Allen-Southwell scheme is uniformly convergent [2][3][6][7][8], that is

$$\|u(x_i) - u_i^e\|_{h,\infty} \le ch.$$

Thus

$$\|u(x_i) - u_i^{m,n}\|_{h,\infty} \le c_1 h + c_2 \|\frac{1}{\rho_i}(\sigma_i^e - \sigma_i^{m,n})\|_{h,\infty}.$$

∎

Theorem 1 indicates that the better an approximation $\sigma_i^{m,n}$ is to σ_i^e, the closer to uniformly convergent the difference scheme will be. This explains why Samarskii's scheme is superior to upwinding since, as is illustrated in Figure 2, $p_{0,1}(y)$ is a better approximation than $p_{0,0}(y)$ to $\sigma(y)$ for all $y > 0$. In the case of Padé approximants, one is guaranteed by classical theory that $p_{m,n}(y)$ better approximates $\sigma(y)$, near $y = 0$, as $m + n$ increases. However it is clear, from Figure 2 and from Figures 3 and 4 of the error term $(\sigma^e(y) - \sigma^{m,n}(y))/y$ in (13), that, for super-diagonal entries in the Padé table, this term has a finite maximum which, for fixed n, reduces as m increases. Table 2 gives the maximum value of the error term and the point at which it is attained.

m	n							
	0		1		2		3	
	Max Error	$y_{\max}$	Max Error	$y_{\max}$	Max Error	$y_{\max}$	Max Error	$y_{\max}$
0	0.50000	0.0	0.09445	1.718	0.03210	2.457	0.01265	3.037
1	-0.5000	∞	-0.08962	7.987	-0.02343	6.439	-0.00726	6.202
2	∞	∞	∞	∞	0.0678	17.729	0.0119	12.310
3	∞	∞	∞	∞	-0.5000	∞	-0.04481	31.620

Table 2: The maximum value of $(\sigma^{m,n}(y) - \sigma^e(y))/y$ and the point $y_{\max}$ at which it is attained.

It should be noted from the table that $p_{n,n}(y) \to (-1)^n(n+1)$ as $y \to \infty$. Thus $\sigma(y) - p_{n,n}(y) \to (-1)^{n+1}(n+1)$. However the term appearing in the error estimate (13) is $(\sigma(y) - p_{n,n}(y))/y$ and this converges to 0, as $y \to \infty$, for all finite n. In general however the diagonal entries, as we might expect from the behaviour of upwinding, do not give as accurate an approximation as the superdiagonal entries with equal values of $m+n$. To state this more precisely, it is better to choose an entry $(n, n+2)$ rather than $(n+1, n+1)$. Both of these involve the same number of evaluations and hence are of comparable cost. Table 3 gives the difference between the solution of the approximate scheme and that of Il'in-Allen-Southwell's scheme for problem (5). It gives results for the case $h = \epsilon$, with a number of values of h varying between $h = 1/10$ and $h = 1/500$. It also includes the error for schemes with $h = 1/20$, and a value of ϵ chosen to make ρ_i be in the region of $y_{\max}$, thus causing the error term $(\sigma(y) - p_{n,n}(y))/y$ to be close to its maximum value.

	$h = \epsilon$				$\rho_i \approx y_{\max}$
Scheme	10	20	40	500	$h = 20$
0,0	.661(0)	.676(0)	.678(0)	.675(0)	.680(0)
0,1	.167(0)	.168(0)	.167(0)	.164(0)	.303(0)
0,2	.396(-1)	.386(-1)	.376(-1)	.365(-1)	.133(0)
0,3	.824(-2)	.776(-2)	.737(-2)	.694(-2)	.572(-1)
1,0	-.193(0)	-.188(0)	-.183(0)	-.177(0)	-.800(2)
1,1	-.255(-1)	-.242(-1)	-.231(-1)	-.218(-1)	-.502(0)
1,2	-.354(-2)	-.325(-2)	-.302(-2)	-.275(-3)	-.121(0)
1,3	-.487(-3)	-.430(-3)	-.384(-3)	-.329(-3)	-.368(-1)
2,0	.395(-2)	.355(-2)	.321(-2)	.281(-2)	.478(1)
2,2	.383(-3)	.332(-3)	.289(-3)	.235(-3)	.303(0)
2,3	.413(-4)	.330(-4)	.273(-4)	.200(-4)	.596(-1)
3,2	.466(-5)	.365(-5)	.264(-5)	.166(-5)	-.796(2)
3,3	-.419(-5)	-.327(-5)	-.231(-5)	-.131(-5)	-.239(0)

Table 3: Error in the solution for various values of h.

Note from this table that, for many values of h, the error with respect to Il'in's scheme is less than might be expected from the behaviour of $(\sigma(y) - p_{n,n}(y))/y$. Also it is slowly varying, with respect to h, as might be expected of a function which depends primarily on ρ_i. However, the final column illustrates that this error may become significant, particularly if one does not choose the approximation in accordance with the rules we suggested earlier.

It is clear that one can achieve good approximation to the boundary layer behaviour using low order rational approximations. These schemes are not uniformly convergent but, for moderate values of h, the additional error due to the approximation is of the order of the discretization error. Better results could be obtained by using a judicious choice of approximation depending on the value of ρ_i, since the maximum error of the different schemes occur at different points. We have discussed here only the Padé approximants. However other rational approximations might also lead to good schemes. In particular the Best Rational Approximation which minimises the L_∞ error on the half-line $(0, \infty)$ might be a more appropriate choice. It should be remarked these schemes

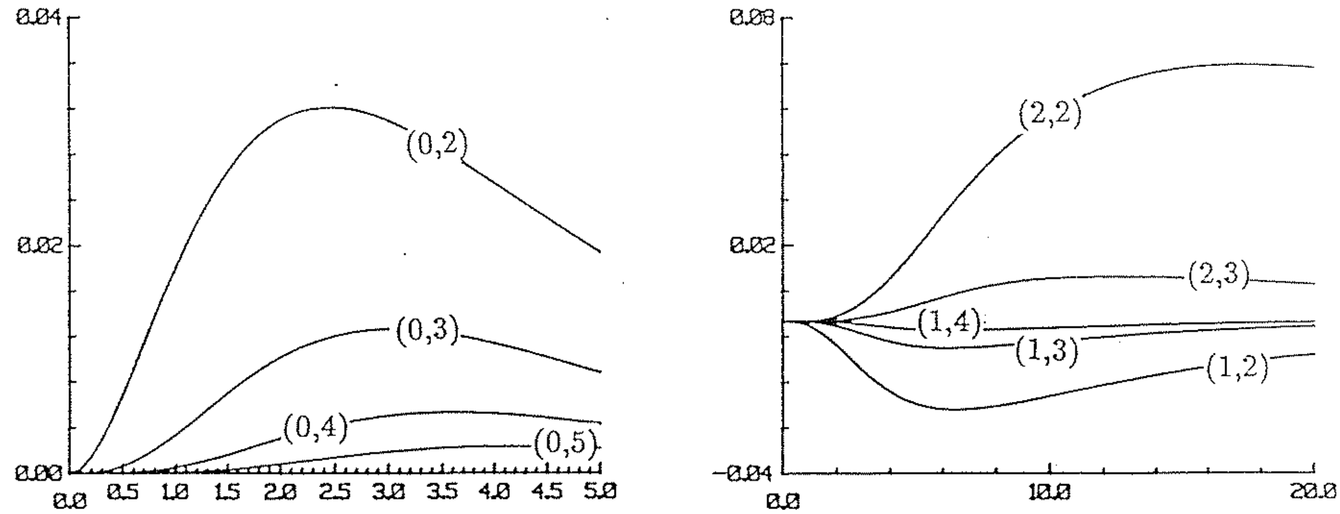

Figure 3: $(\sigma^e(y) - \sigma^{m,n}(y))/y$

behave similarly for simple turning-point problems and higher-order turning point problems, where $a(x)$ has simple and repeated zeros in the interval $(0, 1)$.

References

[1] E.P.Doolan, J.J.H.Miller, W.H.A. Schilders, *Uniform numerical methods for problems with initial and boundary layers*, Boole Press, Dublin, (1980)

[2] P.A. Farrell, *Uniformly convergent difference schemes for singularly perturbed turning and non-turning point problems*, Ph.D. thesis, Trinity College Dublin (1983)

[3] P.A. Farrell, *Sufficient conditions for uniform convergence of a class of difference schemes for a singularly perturbed problem*, IMA Jour. of Numer. Anal., **7**, pp 459-472, (1987)

[4] E.C. Gartland, Jr., *An analysis of a uniformly convergent finite difference/ finite element scheme for a model singular-perturbation problem* Math. Comp. **51**, No. 183 (1988)

[5] E.C. Gartland, Jr., *Graded-mesh difference schemes for singularly perturbed two-point boundary value problems*, Math. Comp. **51**, No. 184 (1988)

[6] A.M. Il'in, *Differencing scheme for a differential equation with a small parameter affecting the highest derivative*, Math. Notes Vol. 6, No. 2, pp. 596-602 (1969)

[7] R.B.Kellogg, A.Tsan, *Analysis of some finite difference approximations for a singular perturbation problem without turning points*, Math. Comp., **32**, pp 1025-1039 (1978)

[8] J.J.H.Miller, *Sufficient conditions for the convergence, uniformly in epsilon, of a 3-point difference scheme for a singular perturbation problem*, Numerical treatment of differential equations, Lect. Notes in Math. 679, eds. R.Ansorge, W.Tornig, pp 51-91 (1978).