

移流拡散方程式の対称有限要素近似と URR

電通大 田端 正久 (Masahisa Tabata)

電通大(院) 尾関 義一 (Yoshiichi Ozeki)

電通大 浜田 穂積 (Hozumi Hamada)

0. はじめに

移流拡散問題の近似スキムは、一般に非対称である。しかし、領域が単連結で流れが非回転的な場合には、ある重み関数を用いて最小化問題の形で解くことが可能である^[1,2]。この手法により導かれる有限要素スキムは対称となり、非対称スキムに比べ計算量、必要なメモリー量が少なくて済む。しかしながら次の問題点があったため、このスキムは実用化できなかった。指数関数で重みを付けたスキムを使うので、得られる行列の成分は $\exp(Pe)$ (Pe はペクレ数) 程の数となる。従来の浮動小数点表示 (IBM表現, IEEE表現等) による計算機では、ペクレ数 10^3 ですぐオーバーフローとなり、計算不可能となった。今回、我々は事実上オーバーフロー、アンダーフローが起こらない新しい数値表現法である URR^[3,4]を用いて、十分大きいペクレ数に対してこの方法が有用であることをいくつかの数値計算結果により示す。今まであきらめられていた方法に、新しい可能性を示すものである。

1. 移流拡散問題

定常移流拡散問題は以下のように記述される。

$$-\nu \Delta u + b \cdot \text{grad } u = f, \quad x \in \Omega, \quad (1a)$$

$$u = g, \quad x \in \partial \Omega. \quad (1b)$$

ここに、 ν は拡散係数、 b は流速、 f, g は与えられた関数、 Ω は $\mathbf{R}^n$ ($n=2, 3$)の有界領域、 $\partial \Omega$ は Ω の境界である。

一般に、移流拡散問題は最小化問題の形で解くことはできない。領域 Ω が単連結で流速 b が非回転的であると、

$$b = \text{grad } \Phi \quad (2)$$

なる速度ポテンシャル Φ が存在する。このとき、問題(1)は次の最小化問題^[1]：

$V(g)$ の中で、汎関数 J を最小にする u を求めよ, (3)

におきかえることができる。ここに,

$$V(g) = \{v \in H^1(\Omega); v(x) = g(x), x \in \partial\Omega\},$$

$$J(v) = a(v, v; w)/2 - (f, v; w),$$

$$a(u, v; w) = \nu \int_{\Omega} \text{grad } u \cdot \text{grad } v \, w dx,$$

$$(f, v; w) = \int_{\Omega} f v \, w dx$$

であり, w は

$$w(x) = \exp(-\Phi(x)/\nu) \quad (4)$$

で定義される重み関数である。

2. 有限要素近似

最小化問題(3)と同値な弱形式は,

$$a(u, v; w) = (f, v; w), \quad \forall v \in V(0) \quad (5)$$

を満たす $u \in V(g)$ を求めよ, である。(5)の有限要素近似,

$$a(u_h, v_h; w) = (f, v_h; w), \quad \forall v_h \in V_h(0) \quad (6)$$

を満たす $u_h \in V_h(g)$ を求めよ, を考える。ここに,

$$V_h(g) = \{v_h \in X_h; v_h(P_i) = g(P_i), P_i \in \partial\Omega\},$$

$$X_h = \{v_h; v_h \in C(\bar{\Omega}), v_h: \text{各要素上一次式}\}$$

である。 u_h を

$$u_h(x) = \sum_{j=1}^N u_j \phi_{hj}(x) \quad (7)$$

と表現する。ここに,

$N: \Omega$ を要素分割したときの総自由度数,

$$\phi_{hi}(x) \in X_h: \text{基底関数 } (\phi_{hi}(P_j) = \delta_{ij}, \quad \forall j)$$

である。(6)に代入すると u_j に関する N 元連立一次方程式

$$\sum_{j=1}^N a(\phi_{hj}, \phi_{hi}; w) u_j = (f, \phi_{hi}; w), \quad \forall P_i \in \Omega, \quad (8a)$$

$$u_j = g(P_j), \quad \forall P_j \in \partial\Omega \quad (8b)$$

が得られる。(8a)を対称性を保って計算するために[1]では, G_e を要素 e の重心として

$$a(\phi_{hj}, \phi_{hi}; w) \sim \nu \sum_e w(G_e) \int_e \text{grad } \phi_{hj} \cdot \text{grad } \phi_{hi} dx, \quad (9)$$

を用いた。ここでは, より精密な次の近似を行う。 $a(\phi_{hj}, \phi_{hi}; w)$ の計算には

$$\Phi(x) \sim \Phi(P_{ij}) + (x - P_{ij}) b_{ij}$$

$$= \Phi(P_{ij}) + (x_{eij} - P_{ij})b_{ij} + (x - x_{eij})b_{ij} \quad (10)$$

を用いる。ここに、

$$b_{ij} = b(P_{ij}) = \text{grad}\Phi(P_{ij}),$$

$$P_{ij} = (P_i + P_j)/2: \text{節点 } P_i, P_j \text{ の中点,}$$

$$x_{eij}: \text{各要素 } e \text{ で } x \cdot b_{ij} \text{ が最小となる点}$$

である。(f, ϕ_{hi} ; w) の計算には

$$\begin{aligned} \Phi(x) &\sim \Phi(P_i) + (x - P_i)b_i \\ &= \Phi(P_i) + (x_{e_i} - P_i)b_i + (x - x_{e_i})b_i \end{aligned} \quad (11)$$

を用いる。ここに、

$$b_i = b(P_i) = \text{grad}\Phi(P_i),$$

$$x_{e_i}: \text{各要素 } e \text{ で } x \cdot b_i \text{ が最小となる点}$$

である。(10), (11) を (8a) に代入すると、

$$Au = F \quad (12)$$

が得られる。ただし、

$$(A)_{ij} = \nu \sum_e \text{grad}\phi_{hj} \cdot \text{grad}\phi_{hi} w_{ij} \bar{\int}_e \exp(-(x - x_{eij})b_{ij}/\nu) dx,$$

$$(F)_i = \sum_e w_i \bar{\int}_e f \phi_{hi} \exp(-(x - x_{e_i})b_i/\nu) dx,$$

$$w_{ij} \bar{\int} = \exp(-(\Phi(P_{ij}) + (x_{eij} - P_{ij})b_{ij})/\nu),$$

$$w_i \bar{\int} = \exp(-(\Phi(P_i) + (x_{e_i} - P_i)b_i)/\nu)$$

である。A は実対称正定値行列である。これらの式に現れる数値積分は

$$\int_T \exp(-\alpha \cdot \xi) d\xi, \quad \int_T \xi_i \exp(-\alpha \cdot \xi) d\xi \quad (13)$$

を精度良く計算するサブルーチンをまず構成してから計算する。ここに、

$$\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n), \quad \alpha_i \geq 0,$$

$$T: \text{頂点 } (0, 0, \dots, 0), (1, 0, \dots, 0), \dots, (0, 0, \dots, 1) \text{ を持つ } n \text{ 次元単体}$$

である。

移流拡散現象を特徴づけるのはペクレ数 (Peclet 数)

$$Pe = |b| L / \nu \quad (14)$$

である。ここに、L は領域代表長である。数値計算で重要なのは、L を要素代表長 h で置き換えたセル・ペクレ数 (cell Peclet 数)

$$Pc = |b| h / \nu \quad (15)$$

である。(12) の左辺の係数行列、右辺のベクトルの成分はそれぞれ $w_{ij} \bar{\int}$, $w_i \bar{\int}$ のため $\exp(Pe)$ 程の数となる。ペクレ数が大きくなると従来の計算機では内部形式の

制約のためオーバーフローとなる。

ペクレ数が大きくなると係数行列の成分の値は不揃いとなるので、コレスキー分解の過程で桁落ちが起こり易くなる。それを防ぐために、行列Aの左右から正の成分を持つ対角行列をかけるスケーリング^[5]を行う、

$$D^{-1}AD^{-1} \cdot Du = D^{-1}F. \quad (16)$$

ここに

$$\begin{aligned} (D^{-1}AD^{-1})_{ij} &= \nu \sum_{\theta} \text{grad} \phi_{hj} \cdot \text{grad} \phi_{hi} w_{ij}'' \int_{\theta} \exp(-(x-x_{\theta ij})b_{ij}/\nu) dx, \\ (D^{-1}F)_i &= \sum_{\theta} w_i'' \int_{\theta} f \phi_{hi} \exp(-(x-x_{\theta i})b_i/\nu) dx, \\ w_{ij}'' &= \exp(-(2\Phi(P_{ij}) - \Phi(P_i) - \Phi(P_j) + 2(x_{\theta ij} - P_{ij})b_{ij})/(2\nu)), \\ w_i'' &= \exp(-(\Phi(P_i) + 2(x_{\theta i} - P_i)b_i)/(2\nu)), \\ (D)_i &= \exp(-\Phi(P_i)/(2\nu)) \end{aligned}$$

である。係数行列Aは対称であるからスケーリングを行っても対称性は保たれる。得られた行列 $D^{-1}AD^{-1}$ の成分は、 w_{ij}'' のため $\exp(P_c/2)$ 程の数となる。スケーリングによって係数行列の成分を $\exp(P_e)$ から $\exp(P_c/2)$ に減少することが可能となった。(16)式は(12)式に比べて計算し易い。速度ポテンシャル Φ は、(2)より

$$\Phi(x) = \int^P b ds + C \quad (17)$$

と書ける。Cは積分定数である。Cは Ω の閉包で $|\min \Phi(x)|$ と $|\max \Phi(x)|$ が等しくなるようにとる。コレスキー分解によって得られた解Duからuを求めるときの D^{-1} の成分は $\exp(P_e/2)$ 程の数となり、この部分の計算を行い易くするためである。

注意1.

従属変数の変換により、移流拡散方程式(1a)を対称化する。未知関数uを

$$v = uw, \quad w = \exp(-\Phi/(2\nu)) \quad (18)$$

でvに変換すると、問題(1)は

$$-\nu \Delta v + \{ |b|^2/(4\nu) - \text{div } b/2 \} v = fw, \quad x \in \Omega, \quad (19a)$$

$$v = gw, \quad x \in \partial\Omega, \quad (19b)$$

となる。

問題(19)の弱形式は

$$\tilde{a}(v, v^*) = (f, v^*; w), \quad \forall v^* \in V(0) \quad (20)$$

を満たす $v \in V(gw)$ を求めよ、である。ここに、

$$\begin{aligned}\tilde{a}(u, v) &= \nu \int_{\Omega} \{\text{grad } u \cdot \text{grad } v + cuv\} dx, \\ c &= |b|^2 / (4\nu) - \text{div } b / 2\end{aligned}\quad (21)$$

である。(20)より得られるスキームは対称となる。(20)の有限要素近似は対称となり一見有望な方法に見えるが、その計算結果は芳しくない(4節注意2参照)。

3. URR

URR^[3,4]は、Universal Representation of Real numbersの略の新しい数値表現法である。従来の計算機は、指数部、仮数部のビット数が固定であるのに対し、URRは数値の大小によって可変となる(図1)。URRは基本的には2進法の体系を用い、数値の区間の分割法(大まかな分割、二重指数分割、等比分割、等差分割)により内部形式を決定する。例えば、1.0、 10^{100} の単精度の内部形式は図1のようになる。

数 値	内 部 形 式
1.0	0:10:000000000000000000000000000000
10^{100}	0:11111111111001001100:001001001010
2^{29}	0:11111111111111111111111111111111

図1. URR単精度(32ビット)の内部形式の例

注) 内部形式は、左から順に符号ビット、URR指数部、仮数部を表す。
 仮数部は、正規表現($-2 \leq f < -1$, $1 \leq f < 2$ とするとき小数点の左の1桁を取り除いた表示形式)を採用している。
 2^{29}
 2^{29} の仮数部はない。

このため、従来の計算機で表現できなかった指数部の大きな数値が表現可能となり、事実上オーバーフロー、アンダーフローが生じない。さらにURRでは、 $1/256 \leq |A| \leq 256$ なるAはIEEE表現より精度が良く、IEEE表現では実現不可能な $|A| \geq 10^{308}$ 、 $|A| \leq 10^{-308}$ なるAもURRでは表現できる。

今回の対称スキムは、URRを用いて初めて実行可能となる。

4. 数値計算例と計算結果

4.1 数値計算例

領域 $\Omega = (0, 1) \times (0, 1)$, $\nu = 1.0, 10^{-1}, \dots, 10^{-9}$ として、次の2種類のデータにより問題(1)を解く。

$$\begin{aligned} 1) & b = (1, 0), \quad g = 0, \quad f = 1, \\ 2) & b = (\cos(\pi/8), \sin(\pi/8)), \quad g = 0, \\ & f = \begin{cases} 1, & x_2 \geq x_1 \tan(\pi/8) + 0.2, \\ 0, & \text{その他,} \end{cases} \end{aligned}$$

Friedricks-Keller型分割により、 Ω を3200要素（各辺40等分）に分割する（図2）。

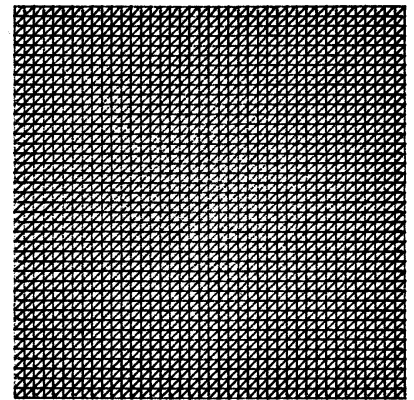


図2. Friedrichs-Keller型40分割

4.2 誤差と計算限界

図3, 図4は、1), 2)それぞれの問題に対して、単精度解と倍精度解の誤差のペクレ数依存性を示したものである。非対称スキム（従来の計算機で計算するために、対称スキムを修正したスキム^[2]）による倍精度解を厳密解代りに用い、絶対値最大ノルムで誤差を計算した。縦軸、横軸とも10を底とする対数値である。URRはURR実数表現による計算を、APOLLO DN4000はアポロ社製ワークステーション（IEEE数値表現を採用）による計算を、HITAC M680Hは日立社製汎用機による計算を示している。

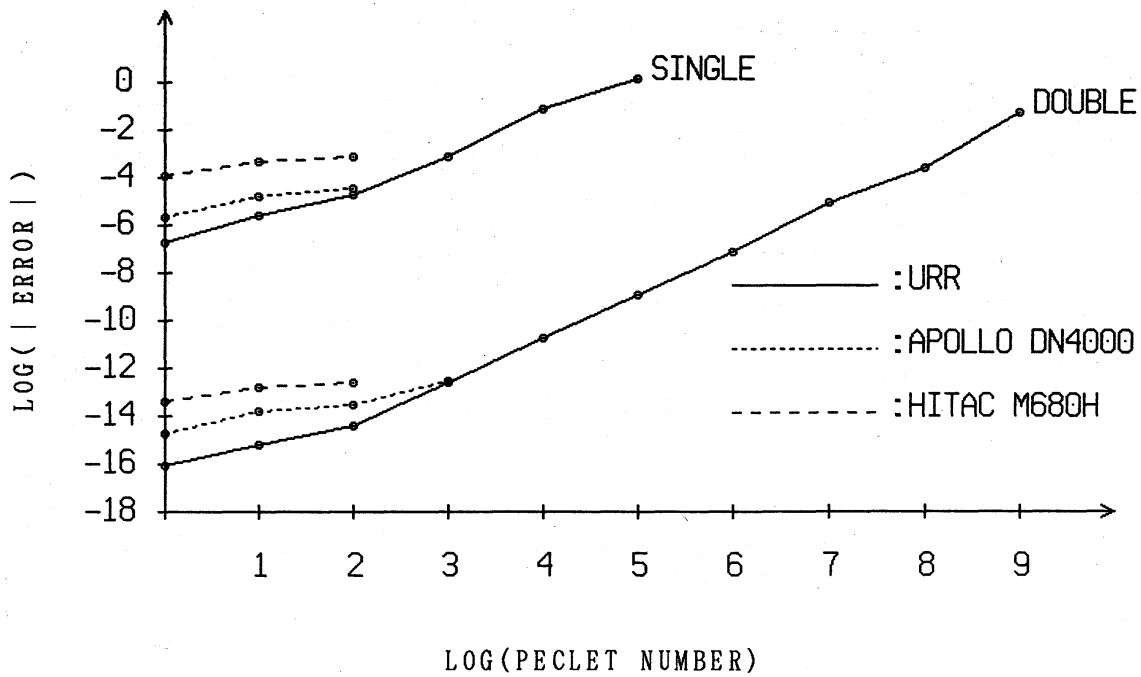


図 3. 1)の相対誤差

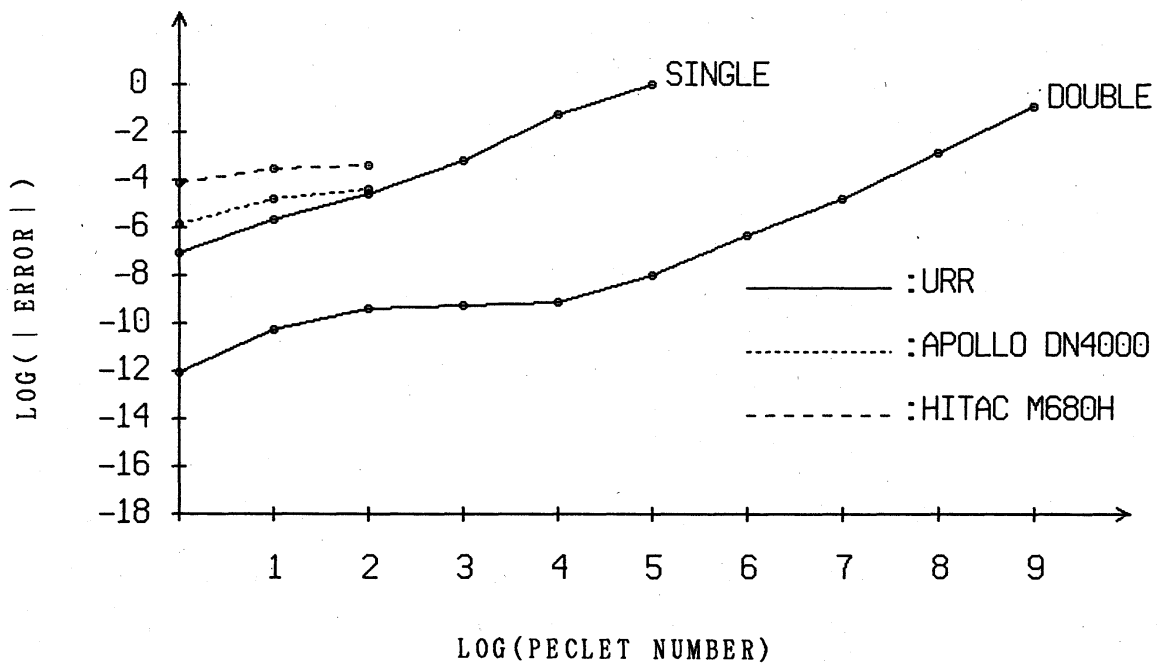


図 4. 2)の相対誤差

図3で、URR単精度解、倍精度解ともペクレ数が大きくなるにつれ誤差は単調に増加している。従来の計算機では、ペクレ数が 10^3 でオーバーフローにより計算が中断するのに対し、URRでは事実上オーバーフローが生じないのでいくらかでも計算できるが、解の精度は序々に落ちていくことが観察できる。アポロ計算機の倍精度解でペクレ数が 10^3 まで解けるのは、計算機の内部形式の違いによる。ペクレ数が小さい範囲でURR計算の精度が良いのは、この範囲では、URR計算の方が仮数部が長くとれていることに起因している。図4でも、同様な結果が得られている。図では見づらいが、倍精度計算でアポロDN4000による結果は 10^3 まで、日立M680Hによる結果は 10^2 までURRのグラフと重なっている。

4.3 C P U - T i m e

表1は、ペクレ数100で各辺を40分割したとき、計算機の種類によるCPU-Timeの違いを示している。

表1. 計算機の種類によるCPU-Time (Pe=100)

	対称単精度	対称倍精度	非対称倍精度
U R R	71.01	75.74	100.52
アポロ	108.06	113.76	160.57
M 6 8 0 H	2.50	2.61	3.85

注) 単位: 秒

対称スキムの計算時間は、非対称スキムの約75%である。現在URR計算は、M680H上のエミュレータで行われているのでかなり時間がかかっている。将来その計算速度は、日立Mシリーズの120%程度まで向上されると予想されるので、対称スキムは非対称スキムの約90%で計算可能となる。

図5は、各辺の分割数による対称スキムと非対称スキムの計算時間の比を示している。

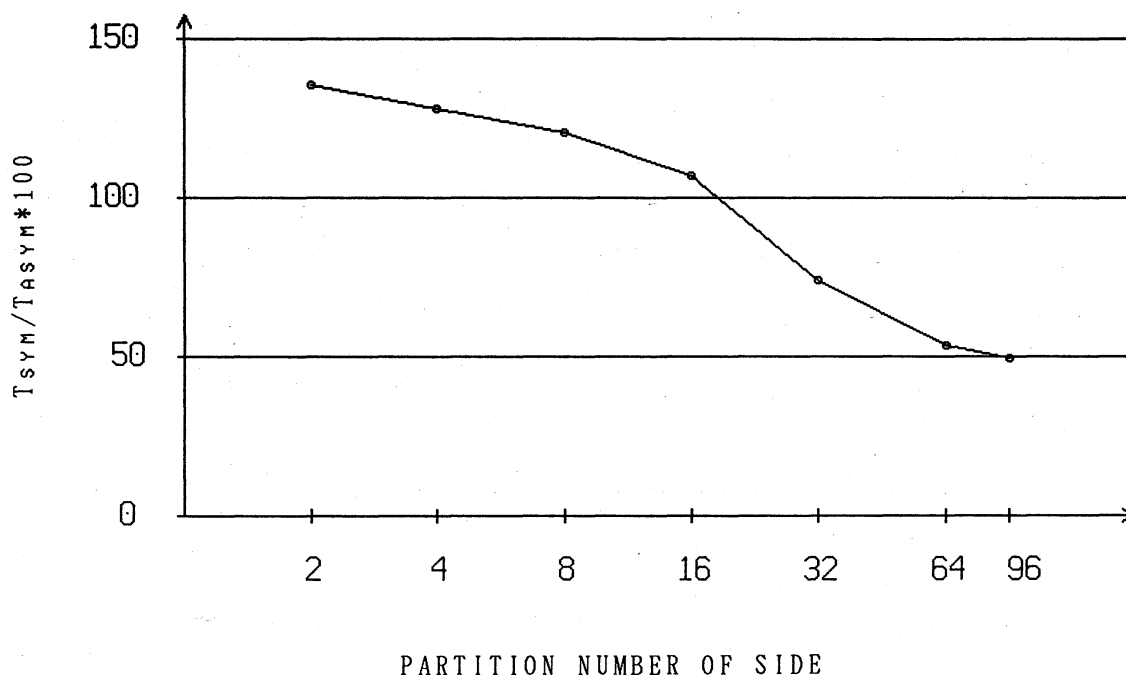


図5. 対称スキムと非対称スキムの計算時間の比 ($Pe=100$)

分割数が少ないとき、剛性行列作成時間が支配的である。分割数が多くなると、連立一次方程式を解く時間が支配的となり、その解法時間の理論比率50%が得られている^[6]。分割数が多くなると、対称スキムは非対称スキムより計算時間がかからず有効である。

注意2.

変数変換による対称化弱形式(20)の有限要素計算結果を図6, 図7に示す。これは、図3, 図4に対応している。図から明らかなように、誤差が急激に大きくなりこの方法は有効でない。

5. おわりに

今回、領域が単連結で流れが非回転的な場合、移流拡散問題を対称スキムで解いた。これは、最小化問題から導かれたスキムであり、解の正定値性を維持する。このスキムは、ペクレ数が大きくなると非常に大きい数（ペクレ数が 10^9 のとき $\exp(Pe)$ は約 10^{10^8} 桁(約10の1億乗))を含む数値計算となり、従来の計算機では内部形式の制約上取り扱うことができなかった。URR計算では、そのような大きな

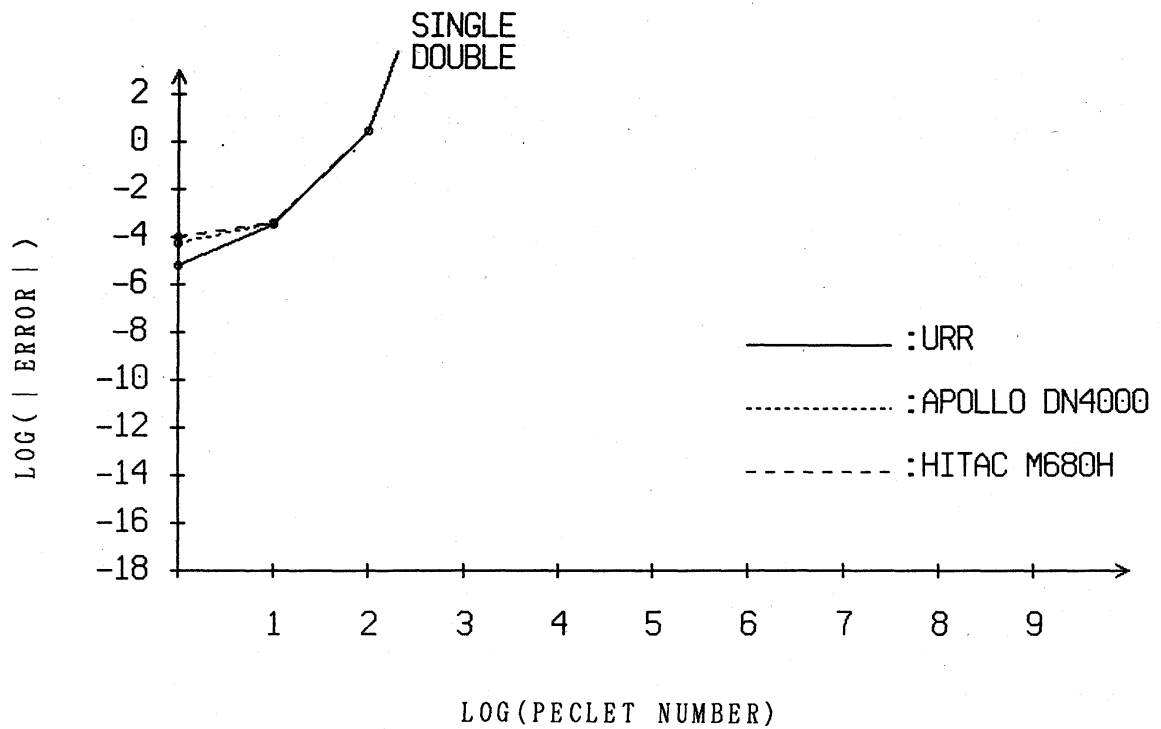


図 6. 1)の相対誤差

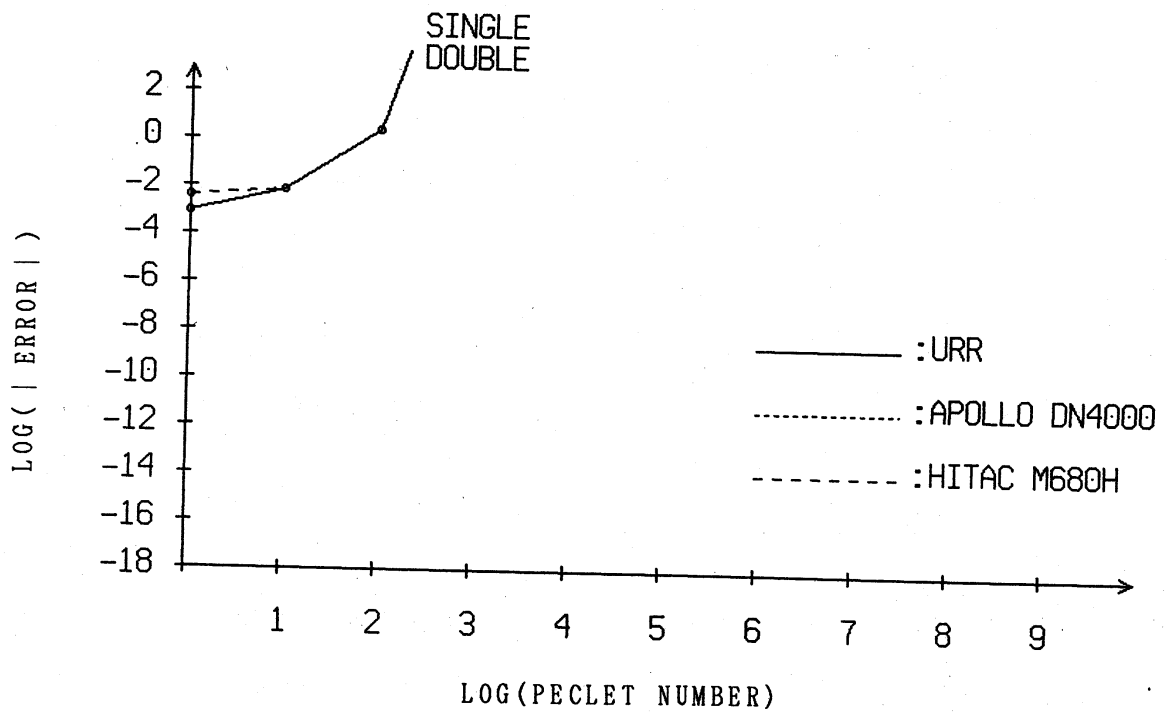


図 7. 2)の相対誤差

数も計算可能となり，高ペクレ数に対しても十分実用的な範囲内の精度で解けるようになった。計算時間に関しては，分割数を多くすると連立一次方程式の解法時間が支配的となるため，対称スキムの所要時間は非対称スキムの半分に近づくことを確認した。

この研究を行うにあたって，日立中央研究所第8部長島重夫部長をはじめ同部の皆様に大変お世話になった。ここに感謝の意を表す。URR計算には，日立中央研究所のM680H上のエミュレータを使用した。また，岡山理科大学理学部榊原道夫助手には，注意1に関する助言を貰った。合せて感謝する。

参考文献

- [1]Tabata, M.: "Symmetric Finite Element Approximation for Convection-Diffusion Problems", Theoretical and Applied Mechanics, Vol.33 (University of Tokyo Press), 445-453(1985).
- [2]Tabata, M.: "A Numerical Algorithm for an Upwind-Type Finite Element Method Using Exponential Functions", Theoretical and Applied Mechanics, Vol.34 (University of Tokyo Press), 371-376(1986).
- [3]浜田穂積: "二重指数分割に基づくデータ長独立実数値表現法", 情報処理学会論文誌, Vol.22, No.6, 521-526(1981).
- [4]Hamada, H.: "A New Real Number Representation and its operation", Proceedings of the eighth symposium on Computer Arithmetic, (IEEE Computer Society Press), 153-157(1987).
- [5]村田健郎, 小国力, 唐木幸比古: "スーパーコンピュータ～科学技術計算への応用～", 丸善, 1985.
- [6]田端正久, 尾関義一: "対称スキムによる移流拡散問題の数値計算", 第3回数値流体力学シンポジウム講演論文集, 177-180(1989).