数理解析研究所講究録
第 422 巻 1981 年 104-118

平野の変形Newton法の大域的収束性

東大 工学部　室田 一雄

1. はじめに

代数方程式の解を求めるために，古来，多くの算法が提案されてきた．最近，研究と改良が行われている連立型の方法（いわゆるDKA法など）は，他の方法に比べて種々の点で優れた算法であるようである([1],[2],[3])．そして，殆んどすべての初期値から出発して解に収束するという大域的収束性も示されている．一方，Newton法の改良版として平野によって提案され，変形Newton法とも呼ばれる算法([4],[5],[6];[7]も参照）は，Newton法の局所的な二次収束性を保ちつつ，Newton法に見られる逐次近似解の不安定な振舞いを抑えるように工夫したもので，任意の初期値から出発して，必ずどれか1つの解に収束する列を生成することが経験的に知られているが，大域的収束性についての厳密な数学的証明は与えられていない。

本論文では，平野の方法が基本的には複素関数論の最大値の原理に立脚していることに注意しつつ，この方法が厳密な

意味での大域的収束性を持つこと，すなわち，任意の初期値から出発して必ずどれか1つの解に収束する近似解の系列を作り出すことを示す．さらに，近似解の系列に対応する関数値の系列は，0に少なくとも一次収束して，その収束率および1つの近似解から次の近似解を求めるのに必要な計算量が，初期値のとり方や，個々の方程式の係数とは無関係に，方程式の次数だけで定まる上界をもつことを証明する．

2．変形Newton法の算法

複素数体上での一変数n次多項式

$$p(z) = z^n + a_{n-1} z^{n-1} + \cdots + a_1 z + a_0$$

の零点の1つを求めるために，任意の初期値$z^{(0)}$から始めて，近似解の列$\{z^{(\nu)}\}_{\nu=0}^{\infty}$を生成する．第$\nu$近似解$z^{(\nu)}$から，第$(\nu+1)$近似解$z^{(\nu+1)}$を求める手順（これを第$\nu$段と呼ぶ）は，以下に示す通りである．

S1: $p(z^{(\nu)} + \Delta z) = \sum_{k=0}^{n} a_k^{(\nu)} (\Delta z)^k$

の係数$a_n^{(\nu)} (=1)$, $a_{n-1}^{(\nu)}$, $\cdots$, $a_0^{(\nu)}$を計算する．

S2: $\mu := 1$

S3: $\zeta_k(\mu) := (-\mu\, a_0^{(\nu)} / a_k^{(\nu)})^{1/k}$ $(k=1, \cdots, n)$

（ここで$(\ )^{1/k}$は任意の分枝を選んでよい）

S4: $m := \arg\min_{1 \le k \le n} |\zeta_k(\mu)|$ （最小値を与えるkをmとする）

S5：$\left|\sum_{k=0}^{n} a_k^{(\nu)} \zeta_m(\mu)^k\right| \leqq \left(1-(1-\beta)\mu\right)\left|a_0^{(\nu)}\right|$

ならば，$z^{(\nu+1)} := z^{(\nu)} + \zeta_m(\mu)$として第$\nu$段を終了する．そうでなければ，$\mu := \mu/(1+\delta)$としてS3に戻る．ただし，$\beta$ $(0<\beta<1)$およびδ (>0)は予め与えられた定数である．

この算法において，S5の判定条件が満たされて，第ν段が終了したときのμの値を$\mu^{(\nu)}$と書くと，明らかに$0<\mu^{(\nu)}\leqq 1$で

$$|p(z^{(\nu+1)})| \leqq \left(1-(1-\beta)\mu^{(\nu)}\right)|p(z^{(\nu)})| \qquad (2.1)$$

が成り立つ．この不等式は，各段が終了したとすれば，関数値が確実に減少することを示している．文献によっては，S5の判定基準を

$$|p(z^{(\nu+1)})| < |p(z^{(\nu)})|$$

としているものもあり，従来，この不等式だけを根拠として，平野の変形Newton法の大域的収束性が主張されてきた．しかし，関数値の列が減少列を成すからといって，それが0に収束する保証はない．各段ごとに(2.1)が満たされたとしても，$\mu^{(\nu)}$がνの増加とともに十分速く0に収束してしまうときには，$p(z^{(\nu)})$が0に収束しないこともありうる．しかし，不等式(2.1)によれば，変形Newton法の大域的収束性を示すには，各段内のS3～S5の反復回数が一様に有界であること，すなわち，νについて一様な下界θ (>0)が存在して，$\mu^{(\nu)} \geqq \theta$とな

ることを示せば十分であることがわかる．以下の証明においては，これよりさらに強く，この下界θは，多項式の次数nおよびパラメータβ，δだけによって定まり，個々の多項式の係数や初期値の与え方とは無関係であることが示される．

3．予備的考察

平野の変形Newton法の基本方針は，十分良い近似解が得られた後はNewton法と同じ挙動をするようにしながら，しかも近似解が真の解からいかに遠い所にあっても，反復1段ごとに関数の絶対値が徐々に減少するようにすることによって，局所的収束率を損わずに，大域的に確実に収束させようとすることである．この基本方針は，複素関数論における"最大値の原理"（に同値な命題）「正則関数の絶対値は零以外の極小値を有しない」によって正当化される．すなわち，ある初期値から始めて関数の絶対値を減らしていくならば，0でない極小値に落ち込んで動きがとれなくなるようなことは決して起こらない．この事実は平野法にとって最も基本的であることはいうまでもないが，従来の文献では明示的には述べられていなかったようである．

次に，算法の各段内で反復されるS3～S5のループについては，μが十分小さくなると，S4で$m=k_0$（k_0は$a_k^{(\lambda)}\neq 0$なる

最小の k）が選ばれ，S5の判定基準を満足して脱出できる．すなわち，各段は有限回の反復の後に終了する．次節において，2次方程式を例にしてこの部分の挙動を具体的に解析して，第ν段内のループの反復回数は，νについて一様な上界をもつこと，言いかえれば，νに依らない$\theta\,(>0)$が存在して，$\mu^{(\nu)} \geqq \theta$となることを示し，一般の場合への布石とする．

この算法は，元来Newton法の改良版として考案されたものであるが，実際，近似解が真の解に十分近くなるとNewton法に一致する．$p(z)$が，$z=0$ をg重根にもつ場合：

$$p(z) = \sum_{k=g}^{n} a_k z^k \qquad (\ a_n = 1,\quad a_g \neq 0,\quad g \geqq 1\ )$$

を考える．S1に従って，これを$z^{(\nu)}$のまわりで展開すると，

$$p(z^{(\nu)}+\Delta z) = \sum_{k=0}^{n} a_k^{(\nu)} (\Delta z)^k ,$$

$$a_k^{(\nu)} = \sum_{j=\max(k,g)}^{n} a_j \binom{j}{k} \left(z^{(\nu)}\right)^{j-k}$$

となる．ここで，$z^{(\nu)}$が十分0に近いならば，

$$a_k^{(\nu)} \sim \begin{cases} a_g \binom{g}{k} \left(z^{(\nu)}\right)^{g-k} & (k \leqq g) \\ a_{j_k} \binom{j_k}{k} \left(z^{(\nu)}\right)^{j_k-k} & (k > g), \end{cases}$$

ただし，j_kは $j \geqq k$ で $a_j \neq 0$ となる最小のjである．特に，$k=0$として，

$$a_0^{(\nu)} \sim a_g \left(z^{(\nu)}\right)^g$$

に注意すると，

$$|\zeta_k(\mu)| \sim \begin{cases} \mu^{\frac{1}{k}} \binom{g}{k}^{-\frac{1}{k}} \, |z^{(\nu)}| & (k \leqq g) \\[2ex] \left|\dfrac{\mu\, a_g}{a_{j_k}}\right|^{\frac{1}{k}} \binom{j_k}{k}^{-\frac{1}{k}} |z^{(\nu)}|^{1-\frac{j_k - g}{k}} & (k > g) \end{cases}$$

となるので，$\min_{1 \leqq k \leqq n} |\zeta_k(\mu)| = |\zeta_1(\mu)|$ となる．すなわち，平野の変形Newton法は，解の十分近くでは解の重複度とは無関係に，S4において $m=1$ を選び，Newton法に一致する．したがって，近似解の真解への局所的収束性は，単根ならば二次収束，重根ならば一次収束である．

4. 2次方程式の場合

平野の変形Newton法（$\beta=3/4$，$\delta=37/27$）を2次方程式に適用したときの，第ν段の計算の流れを図1に示す．上添字$^{(\nu)}$は省略した．S1で計算された係数 a_0，a_1 に従い，3つの場合が起こりうる：

(i) $|a_0| \leqq \frac{3}{4}|a_1|^2$ のとき　　$\Delta z = \dfrac{-a_0}{a_1}$，$\mu = 1$，

(ii) $|a_0| \geqq \frac{16}{9}|a_1|^2$ のとき　　$\Delta z = (-a_0)^{1/2}$，$\mu = 1$，

(iii) $\frac{3}{4}|a_1|^2 < |a_0| < \frac{16}{9}|a_1|^2$ のとき　　$\Delta z = \dfrac{-27 a_0}{64 a_1}$，$\mu = \dfrac{27}{64}$，

ただし，Δz は近似解の修正量である．いずれの場合にも，S3～S5のループは高々1回余分に実行されるだけで，$\mu^{(\nu)}$ に対する一様な下界 $\theta = 27/64$ が存在することがわかる．そし

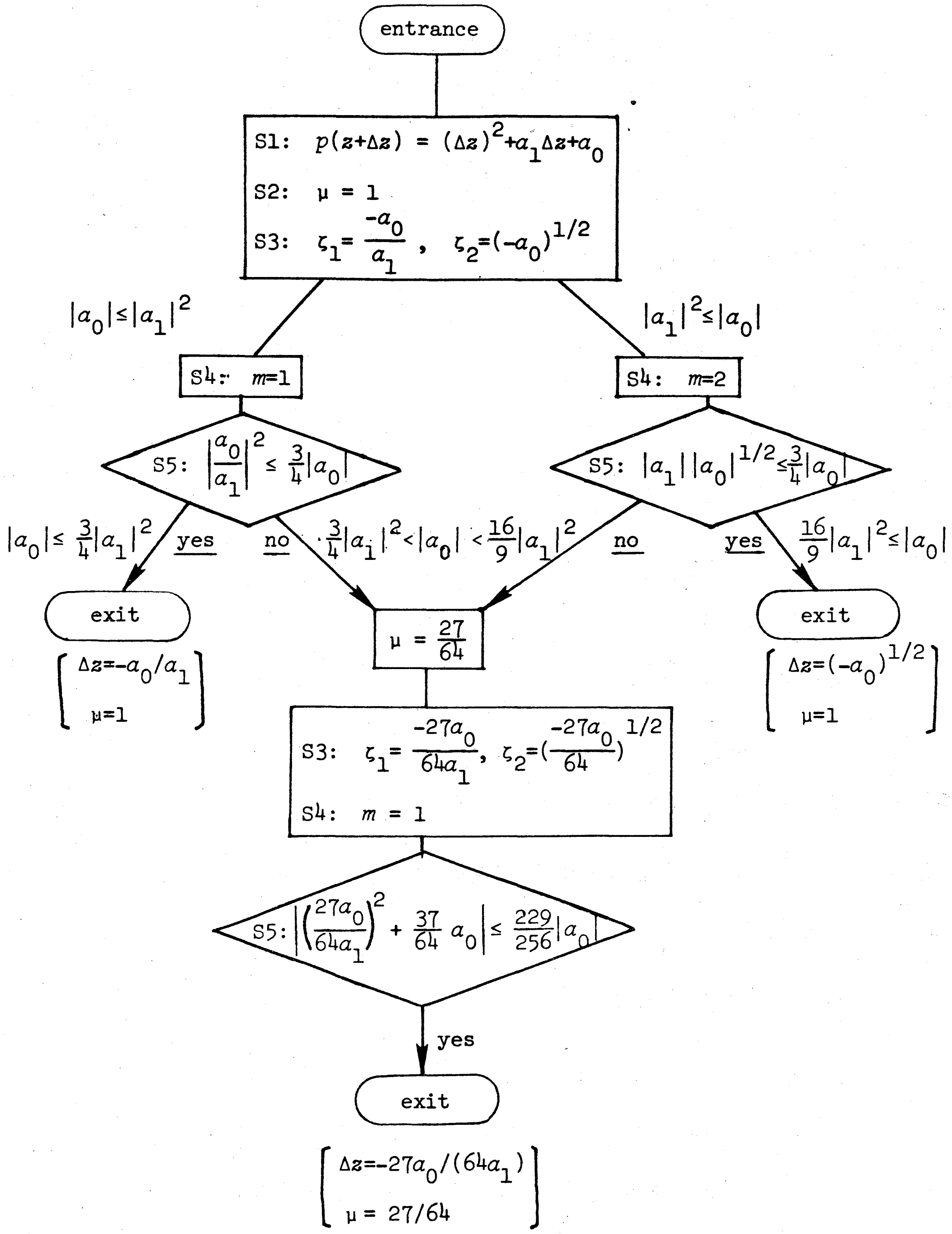

図1. 2次方程式の場合の平野法の計算の流れ $(\beta=\frac{3}{4},\delta=\frac{37}{27})$

て，関数の絶対値の減少率 $|p(z^{(\nu+1)})/p(z^{(\nu)})| \leqq 1-(1-\beta)\theta = 229/256$ が保証されるので，$\nu\to\infty$ のとき $p(z^{(\nu)})\to 0$ となる．

5. 大域的収束性の証明

平野の変形Newton法の第 ν 段が終了したときの μ の値を $\mu^{(\nu)}$ とするとき，ν に依らない $\theta(>0)$ が存在して，$\mu^{(\nu)}\geqq\theta$ となることを示すことによって，平野の変形Newton法の大域的収束性を証明する．以下，第 ν 段だけを考えるので，上添字 $^{(\nu)}$ は省略し，$a_0\neq 0$ と仮定する．

まず，次の記号を用意する：

$$M_m \equiv \{\mu>0 \mid |\zeta_m(\mu)| \leqq |\zeta_k(\mu)|,\ k=1,\cdots,n\},$$

$$D_m \equiv \{\mu>0 \mid |\sum_{k=0}^{n} a_k\zeta_m(\mu)^k| \leqq (1-(1-\beta)\mu)|a_0|\}.$$

ここで，M_m はS4において，$m=m$ が選ばれるような μ の範囲を表わし，D_m はS5の判定条件が満たされるような μ の範囲を表わす．S3~S5のループを脱出するのは，等比的に減少する μ の数列 $\{(1+\delta)^{-k}\}_{k=0}^{\infty}$ が，はじめて $\bigcup_{m=1}^{n}(M_m\cap D_m)$ にはいったときである．

$\zeta_k(1)\equiv\zeta_k$ と書くと，$\zeta_k(\mu)=\mu^{1/k}\zeta_k$ となり，上で定義した集合 M_m，D_m は，

$$M_m=\left\{\mu>0 \,\middle|\, \mu^{\frac{1}{m}-\frac{1}{k}} \leqslant \left|\frac{\zeta_k}{\zeta_m}\right| ,\ k=1,\cdots,n\right\},$$

$$D_m=\left\{\mu>0 \,\middle|\, \left|1-\sum_{k=1}^{n}\mu^{\frac{k}{m}}\left(\frac{\zeta_m}{\zeta_k}\right)^k\right| \leqslant 1-(1-\beta)\mu\right\}$$

と表わされる．μ の区間 B_m を

$$B_m \equiv \left\{\mu>0 \,\middle|\, \mu^{\frac{1}{m}-\frac{1}{k}} \leqslant c\left|\frac{\zeta_k}{\zeta_m}\right| ,\ k=1,\cdots,n;\ k\neq m\right\}$$

$$c\equiv \beta/(1+\beta), \tag{5.1}$$

で定義すると，$B_m\subset M_m\cap D_m$ となる．実際，$B_m\subset M_m$ は，$0<c<1/2$ であることにより明らかであり，$B_m\subset D_m$ は，

$$\mu\in B_m \Rightarrow \mu^{\frac{k}{m}}\left|\frac{\zeta_m}{\zeta_k}\right|^k \leqslant \mu\, c^k\ (1\leqslant k\leqslant n,\ k\neq m)$$

$$\Rightarrow \left|1-\sum_{k=1}^{n}\mu^{\frac{k}{m}}\left(\frac{\zeta_m}{\zeta_k}\right)^k\right| = \left|1-\mu-\sum_{\substack{1\leqslant k\leqslant n\\ k\neq m}}\mu^{\frac{k}{m}}\left(\frac{\zeta_m}{\zeta_k}\right)^k\right|$$

$$\leqslant 1-\mu+\sum_{\substack{1\leqslant k\leqslant n\\ k\neq m}}\mu^{\frac{k}{m}}\left|\frac{\zeta_m}{\zeta_k}\right|^k$$

$$\leqslant 1-\mu+\mu\sum_{\substack{1\leqslant k\leqslant n\\ k\neq m}}c^k$$

$$\leqslant 1-\mu+\mu\cdot\frac{c}{1-c} = 1-(1-\beta)\mu$$

による．ゆえに，$\bigcup_{m=1}^{n} B_m \subset \bigcup_{m=1}^{n}(M_m\cap D_m)$ となる。

次の定理は，任意の $(\zeta_1,\cdots,\zeta_n)$ に対して，$\{B_m\cap[\theta,1]\}_{m=1}^{n}$

の中に，対数的長さ(=上限と下限の比)が$(1+\delta)$以上の区間が存在することを主張する．

<u>定理</u>　任意のβ $(0<\beta<1)$，任意のδ (>0)，任意の$(\zeta_1,\cdots,\zeta_n)$に対して，あるm $(1\leq m\leq n)$が存在して，

$$B_m\cap[\theta,1]\neq\phi$$

であって，

$$\frac{\sup(B_m\cap[\theta,1])}{\inf(B_m\cap[\theta,1])}\geqq 1+\delta, \tag{5.2}$$

ただし，

$$\theta=\left(\frac{\beta}{1+\beta}\right)^{2n^3}(1+\delta)^{-n}. \tag{5.3}$$

任意の$(\zeta_1,\cdots,\zeta_n)$を考えることは，任意の係数$(a_0,\cdots,a_n)$を考えることと等価であるから，この定理は，変形Newton法のパラメータβ $(0<\beta<1)$，δ (>0)を任意に指定するとき，各段において定まる$\mu^{(\nu)}$が，初期値のとり方や，解くべき個々の方程式の係数とは無関係に，方程式の次数だけによって定まる下界θをもつこと，すなわち，すべてのνに対して，$\mu^{(\nu)}\geqq\theta$を意味し，したがって，変形Newton法の大域的収束性を保証する．同時に，各段におけるS3～S5のループの反復回数が

$$N\equiv\frac{-\log\theta}{\log(1+\delta)}=n+\frac{2n^3\log(1+1/\beta)}{\log(1+\delta)} \tag{5.4}$$

で上から抑えられていることも示す.

<u>定理の証明</u>　　B_m の定義(5.1)により,

$$\mu \in B_m \iff$$
$$\max_{1\leq j\leq m-1}\left(\frac{1}{c}\left|\frac{\zeta_m}{\zeta_j}\right|\right)^{\frac{mj}{m-j}} \leq \mu \leq \min_{m+1\leq k\leq n}\left(c\left|\frac{\zeta_k}{\zeta_m}\right|\right)^{\frac{km}{k-m}}. \tag{5.5}$$

定理が成り立たないと仮定すると，ある $\beta\,(0<\beta<1)$, $\delta\,(>0)$, $(\zeta_1,\cdots,\zeta_n)$ が存在して，すべての $m\,(1\leq m\leq n)$ に対して，

(i)　$B_m \neq \phi$ で $B_m \subset [0, (1+\delta)\theta)$,

(ii)　$B_m \neq \phi$ で $B_m \subset ((1+\delta)^{-1}, \infty)$,

(iii)　B_m の対数的長さが $(1+\delta)$ 未満 ($B_m=\phi$ も含むものと解釈する)

の少なくとも1つが成り立つ．B_mの定義(5.5)を用いると，これらは次のように書ける．

(i)　$B_m \neq \phi$ で, $\exists k\,(>m)$: $\left(c\left|\frac{\zeta_k}{\zeta_m}\right|\right)^{\frac{km}{k-m}} < (1+\delta)\theta$.

両辺の対数をとると,

$$-\frac{\log|\zeta_m|-\log|\zeta_k|}{\frac{1}{m}-\frac{1}{k}} < \log\theta + \log(1+\delta) + \frac{\log(1/c)}{\frac{1}{m}-\frac{1}{k}}.$$

したがって

$$-\Delta(m,k) < \log\theta + \log(1+\delta) + n^2\log(1/c), \tag{5.6}$$

ただし,

$$\Delta(m,k) \equiv \frac{\log|\zeta_m|-\log|\zeta_k|}{\frac{1}{m}-\frac{1}{k}}$$

(ii) $B_m \neq \phi$ で，$^{\exists} j(<m)$: $(1+\delta)^{-1} < \left(\frac{1}{c}\left|\frac{\zeta_m}{\zeta_j}\right|\right)^{\frac{mj}{m-j}}$，

これより

$$\Delta(j, m) < \log(1+\delta) + n^2 \log(1/c). \tag{5.7}$$

(iii) $^{\exists} j(<m)$, $^{\exists} k(>m)$:

$$(1+\delta)\left(\frac{1}{c}\left|\frac{\zeta_m}{\zeta_j}\right|\right)^{\frac{mj}{m-j}} > \left(c\left|\frac{\zeta_k}{\zeta_m}\right|\right)^{\frac{km}{k-m}}.$$

上と同様にして，

$$\Delta(j, m) - \Delta(m, k) < \log(1+\delta) + 2n^2 \log(1/c). \tag{5.8}$$

さて，図2のように，xy 平面上の n 個の点 $(1/m, \log|\zeta_m|)$ $(m=1, \cdots, n)$，および2つの半直線 $x = 1/n$ $(y \geqq \log|\zeta_n|)$，$x = 1/k_0$ $(y \geqq \log|\zeta_{k_0}|)$ (k_0 は，$a_k \neq 0$ なる最小の k) の凸閉包の頂点の x 座標を $1/m_0, 1/m_1, \cdots, 1/m_L$ $(m_0 < m_1 < \cdots < m_L)$ とする．明らかに $m_0 = k_0$, $m_L = n$ である．m_l $(0 \leqq l \leqq L)$ の中で，(i) を満たす最大のものを m_s，(ii) を満たす最小のものを m_f とする．(i) は $m = m_0$ で，(ii) は $m = m_L$ で成り立つから，m_s, m_f ともに定義され，さらに，$m < m'$ $(B_m \neq \phi, B_{m'} \neq \phi)$ ならば，(5.5) により，

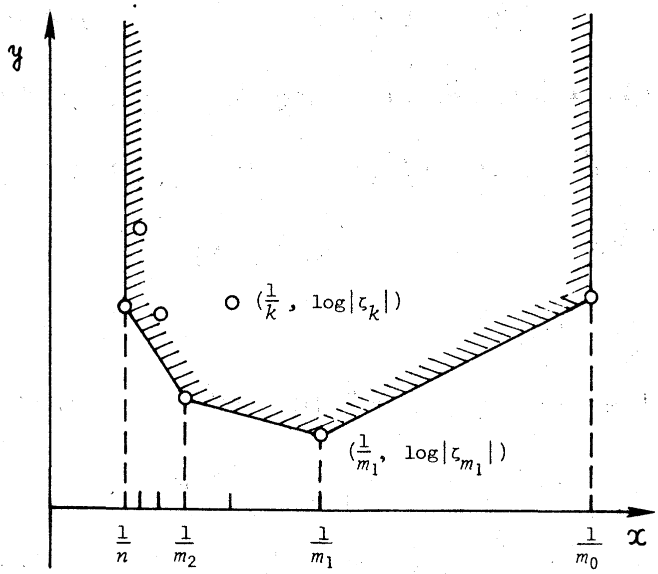

図2. 証明に用いる凸閉包

$$\sup B_m = \min_{m+1 \leq k \leq n} \left(c \left| \frac{\zeta_k}{\zeta_m} \right| \right)^{\frac{km}{k-m}}$$
$$\leq \left(c \left| \frac{\zeta_{m'}}{\zeta_m} \right| \right)^{\frac{m'm}{m'-m}}$$
$$\leq \left(\frac{1}{c} \left| \frac{\zeta_{m'}}{\zeta_m} \right| \right)^{\frac{m'm}{m'-m}}$$
$$\leq \max_{1 \leq j \leq m'} \left(\frac{1}{c} \left| \frac{\zeta_{m'}}{\zeta_j} \right| \right)^{\frac{m'j}{m'-j}} = \inf B_{m'}$$

であり，$n \geqq 2$ ならば，$(1+\delta)\theta < (1+\delta)^{-1}$ となるので，$m_s < m_f$ である．m_s と m_f の定め方によって，$m = m_l$ $(s < l < f)$ に対しては(iii)が成り立っている．

ここで，上記の不等式が，特に $m = m_s, m_{s+1}, \cdots, m_{f-1}, m_f$ に対して成立することに注目する．まず $m = m_s$ は(i)の場合であり，(5.6) の左辺は $k = m_{s+1}$ のとき最小となるので，存在が保証されている k は m_{s+1} にとることができる．すなわち，

$$-\Delta(m_s, m_{s+1}) < \log\theta + \log(1+\delta) + n^2 \log(1/c). \quad (5.9)$$

$m = m_l$ $(s < l < f)$ は(iii)の場合で，同様に，(5.8)において，$j = m_{l-1}$，$k = m_{l+1}$ とできるので，$s < l < f$ に対し，

$$\Delta(m_{l-1}, m_l) - \Delta(m_l, m_{l+1})$$
$$< \log(1+\delta) + 2n^2 \log(1/c). \quad (5.10)$$

$m = m_f$ では(ii)が成り立ち，(5.7)で $j = m_{f-1}$ にとれるので，

$$\Delta(m_{f-1}, m_f) < \log(1+\delta) + n^2 \log(1/c). \quad (5.11)$$

以上の不等式(5.9)，(5.10)，(5.11)を加え合わせると，

$$0 < \log\theta + (f-s+1)\log(1+\delta) + 2(f-s)n^2\log(1/c).$$

ゆえに，

$$0 < \log\theta + n\log(1+\delta) + 2n^3\log(1/c).$$

ところが，これはθの定義(5.3)と矛盾する．　証明終

7. むすび

本論文では，いわゆる"平野法"の大域的収束性を論じ，特に，関数値の減少率に関して，方程式には無関係に次数のみで定まる限界を与えることに成功した．この理論的限界はまだかなり甘く，平野法の実際の挙動をよく反映しているとは言い難い．さらに実用的な限界を探究することが今後の課題である．

最後に，本研究の機会を与えて下さり，また，有益な助言を多く与えて下さった東京大学工学部伊理正夫教授に感謝します．

参考文献

[1] 山本哲朗：ある代数方程式解法と解の事後評価法．数理科学，Vol. 14, No.7, pp. 52-57 (1976)．

[2] 山本哲朗，古金卯太郎，野倉久美：代数方程式を解くDurand-Kerner法とAberth法．情報処理，Vol. 18, No.6,

pp. 566-571 (1977).

[3] 伊理正夫，山下浩，寺野隆雄，小野令美：大域的収束性をもつ代数方程式の解法．数値計算のアルゴリズムとコンピューター，数理科学講究録339，pp. 43-69 (1978).

[4] 平野菅保：代数方程式の解法及び誤差，第8回プログラミングシンポジウム報告集，情報処理学会 (1967).

[5] 木村久男他：電気学会大学講座，電子演算工学概論，2版，pp. 110-123，電気学会，東京 (1976).

[6] 情報処理学会：情報処理ハンドブック，10編7章3節，オーム社，東京 (1980).

[7] Nickel, K: Numerische Berechnung der Wurzeln eines Polynoms. *Numerische Mathematik*, Vol. 9, pp. 80-98 (1966).