

浮動小数係数での多変数ヘンゼル構成における桁落ちのメカニズム

筑波大学数学系 佐々木 建昭(Tateaki Sasaki) *

概要

1998年、SasakiとYamaguchiは、浮動小数係数の多変数多項式のHensel構成では、展開点を特異点の δ 近傍に選ぶとき、 k 次のHensel因子の係数部に $O((c\delta)^{-k})$ 、 c は定数、の巨大な桁落ち誤差が生じ得ることを理論的・実験的に示して衝撃を与えた。しかし、彼らの理論は具体的に桁落ちのメカニズムを明らかにしたものではない。本稿では、以下の3点でSasaki-Yamaguchiの理論を一般化・明解化する。a) 初期因子の個数を2から r に一般化する、b) 桁落ちのメカニズムとして四つの型を明らかにする、c) 特異点近傍のみならず遠方での桁落ちも考察する。

1 はじめに

1998年、SasakiとYamaguchi[SY98]はISSAC'98の第一講演において、浮動小数係数の多変数多項式のHensel構成では、展開点を特異点の δ 近傍に選ぶとき、 k 次のHensel因子の係数部に $O((c\delta)^{-k})$ 、 c は定数、の巨大な数値誤差が生じ得ることを理論的・実験的に示して衝撃を与えた。この誤差は主要項のキャンセレーションによる桁落ちから生じている。しかし、[SY98]の理論はべき級数の収束半径に関するCauchy-Hadamardの定理に基づいて、「級数の展開係数が漸近的にこう振舞うから、これこれの桁落ち誤差が発生するはずだ」というものであり、具体的に桁落ちのメカニズムを明らかにしたものではない。さらに、桁落ち量の解析においてはMoses-Yunの補間式が決定的役割を演じるが、[SY98]では行列式表現を用いて補間式の振舞いを調べたため、かなり複雑な解析になっている。また、展開点を特異点の近傍に限定している。

本稿では、以下の点で[SY98]の解析をより一般化・詳細化・明解化する。

- a) Hensel因子の個数は[SY98]では2だったが、本稿では個数を r に一般化する。
- b) Moses-Yunの補間式に対する解析的表現を導出し、補間式の性質を明らかにする。
- c) 拡張Hensel構成の因子との比較により、一般Hensel構成の因子の性質を明らかにする。
- d) 桁落ちのメカニズムとして四つの型を明らかにする（これで十分かどうかは不明だが）。
- e) 特異点(2章に定義する)近傍のみならず、遠方での桁落ちの様子も明らかにする。

*sasaki@math.tsukuba.ac.jp

2章では、一般 Hensel 構成の特異点を定義し、一般 Hensel 構成と拡張 Hensel 構成の算法を簡単に復習する。そして、Hensel 因子の重要な性質を明らかにする。3章では、桁落ちの解析において決定的役割を演じる Moses-Yun の補間式の性質を詳しく調べる。4章では、桁落ち誤差を引き起こす可能性のある項キャンセルーションとして、四つの型を明らかにする。5章では、展開点の特異点の近傍および遠方の場合に、これら四つの型の項キャンセルーション機構を定量的に検討する。その結果、特異点の近傍において二つの型が巨大な桁落ち誤差を引き起こしうることを明らかにする。

本稿では定理の証明等の詳細は省略する。詳細については [Sas00] を参照されたい。

2 一般 Hensel と拡張 Hensel 構成

\mathbf{C} を複素数体、 $\mathbf{C}[u_1, \dots, u_\ell]$, $\mathbf{C}(u_1, \dots, u_\ell)$, $\mathbf{C}\{u_1, \dots, u_\ell\}$ をそれぞれ \mathbf{C} 上の変数 u_1, \dots, u_ℓ の多項式環、有理式体、形式的べき級数環とする。与えられた多変数多項式を $F(x, u_1, \dots, u_\ell)$ 、展開点を (s_1, \dots, s_ℓ) とする。以下では変数の組 (u_1, \dots, u_ℓ) と数値の組 (s_1, \dots, s_ℓ) をそれぞれ (u) , (s) と略記する。 $f(u)$ を u_1, \dots, u_ℓ の多項式とすると、その全次数を $\text{tdeg}(f)$ と表し、 $f(u)$ の各項の全次数の中で最小のもの（位数と呼ぶ）を $\text{ord}(f)$ と表す。また、多項式 P のノルムを $\|P\|$ と表す。

$$\begin{cases} F(x, u) = f_n(u)x^n + f_{n-1}(u)x^{n-1} + \dots + f_0(u)x^0, \\ e_j \stackrel{\text{def}}{=} \text{ord}(f_j(u)) \leq \text{tdeg}(f_j(u)) \stackrel{\text{def}}{=} E_j \quad (j = n, \dots, 0), \\ e_n = 0, \quad e_j > 0 \quad (j = n-1, \dots, 0), \quad E \stackrel{\text{def}}{=} \max\{E_n, \dots, E_0\}. \end{cases} \quad (1)$$

$f_1(s) = f_0(s) = 0$ のとき、展開点 (s) を **Hensel 構成の特異点**、略して特異点という。この名称の由来は、 $f_{n-1}(s) = \dots = f_0(s) = 0$ のとき $F(x, s) = f_n(s)x^n$ となって、一般 Hensel 構成が破綻するからである。展開点が原点以外の場合は、 $u_i = s_i + v_i$ ($i = 1, \dots, \ell$) なる新変数 v_1, \dots, v_ℓ を導入する。

$$\begin{cases} F(x, s + v) = F^{(0)}(x) + F^{(1)}(x, v) + \dots + F^{(E)}(x, v), \\ \text{tdeg}_v(F^{(k)}(x, v)) = k \quad (k = 0, 1, \dots, E), \\ F^{(k)}(x, v) = f_n^{(k)}(v)x^n + \dots + f_1^{(k)}(v)x^1 + f_0^{(k)}(v)x^0. \end{cases} \quad (2)$$

2.1 一般 Hensel 構成

一般 Hensel 構成の初期因子を $G_1^{(0)}(x), \dots, G_r^{(0)}(x)$ と表し、Moses-Yun の補間式を $W_1^{(l)}(x), \dots, W_r^{(l)}(x)$ ($l = 0, 1, \dots, n-1$) と表す。初期因子と補間式は次式を満たす。

$$\begin{cases} F^{(0)}(x) = G_1^{(0)}(x) \cdots G_r^{(0)}(x), & r \geq 2, \\ \deg(G_i^{(0)}) = n_i, & \gcd(G_i^{(0)}, G_j^{(0)}) = 1 \quad (\forall i \neq j), \end{cases} \quad (3)$$

$$\begin{cases} W_1^{(l)}(x) \frac{F^{(0)}(x)}{G_1^{(0)}(x)} + \cdots + W_r^{(l)}(x) \frac{F^{(0)}(x)}{G_r^{(0)}(x)} = x^l & (l = 0, 1, \dots, n-1), \\ \deg(W_r^{(l)}(x)) < \deg(G_i^{(0)}(x)) = n_i & (i = 1, \dots, r). \end{cases} \quad (4)$$

そして、次の算式で $k = 1 \Rightarrow 2 \Rightarrow \dots$ と逐次的に $G_i^{(k)}(x, v)$ ($i = 1, \dots, r$) を計算する。

$$\begin{aligned} D^{(k)} &\equiv F(x, s+v) - G_1^{(k-1)}(x, v) \cdots G_r^{(k-1)}(x, v) \pmod{(v)^{k+1}} \\ &= d_{n-1}^{(k)}(v) x^{n-1} + \cdots + d_0^{(k)}(v) x^0, \quad k \geq 1, \end{aligned} \quad (5)$$

$$G_i^{(k)} = G_i^{(k-1)} + \delta G_i^{(k)}, \quad \delta G_i^{(k)} = W_i \otimes D^{(k)} \stackrel{\text{def}}{=} \sum_{l=0}^{n-1} W_i^{(l)}(x) \cdot d_l^{(k)}(v). \quad (6)$$

2.2 拡張 Hensel 構成

拡張 Hensel 構成とは、一般 Hensel 構成が破綻する場合に強引に Hensel 構成を実行しようというものである。2 変数に対しては 1989 年に Kuo により [Kuo89]、多変数に対しては 1993 年に Sasaki と Kako により Kuo の仕事とは無関係に考案された [SK99]。本稿では、一般性を失うことなく原点が特異点であるとし、簡単のため $f_{n-1}(0) = \cdots = f_0(0) = 0$ であるとする。多変数多項式に対する拡張 Hensel 構成のエッセンスは、 $F(x, u) \mapsto F(x, tu)$ なる変換で従変数 u_1, \dots, u_ℓ に対して全次数変数 t を導入し、初期因子として次に定める Newton 多項式の因子を使うことである。

定義 1 (Newton line L_{New} and Newton polynomial $\hat{F}_{\text{New}}(x, u)$ for $F(x, u)$)

1. For each monomial $T = c x^i t^j u_1^{j_1} \cdots u_\ell^{j_\ell}$ of $F(x, tu)$, with $c \in \mathbf{C}$ and $j = j_1 + \cdots + j_\ell$, plot a dot at the point (i, j) in the (e_x, e_t) -plane;
2. Let L_{New} be a straight line in (e_x, e_t) -plane, such that it passes the point $(n, 0)$ and another dot plotted and that any dot plotted is not below L_{New} ;
3. Construct $\hat{F}_{\text{New}}(x, tu)$ by summing all the monomials which are plotted on L_{New} .

Newton 線の傾きを $-\lambda$ とすれば、 $\hat{F}_{\text{New}}(x, tu)$ は x と t^λ の同次多項式である。さらに、イデアル I_k を次のように定める。Newton 線 L_{New} を $e_x/n + e_t/e = 1$ とし ($(0, e)$ は L_{New} と e_t -軸の交点の座標である)、正整数 \hat{n}, \hat{e} を $\hat{e}/\hat{n} = e/n$, $\gcd(\hat{n}, \hat{e}) = 1$ であるように定めるとき、

$$I_k = \langle x^n t^{(k+0)/\hat{n}}, x^{n-1} t^{(k+\hat{e})/\hat{n}}, \dots, x^0 t^{(k+n\hat{e})/\hat{n}} \rangle. \quad (7)$$

(ここで、 $\langle p_1, \dots, p_l \rangle$ は p_1, \dots, p_l を生成元とするイデアルを表す)。Newton 多項式 $\hat{F}_{\text{New}}(x, u)$ は二つ以上の項を持つから、次のように因数分解されるとする ($\hat{F}_{\text{New}}(x, u) = (x - \theta)^n$ となる場合については [SK99] を参照)。

$$\begin{cases} \hat{F}_{\text{New}}(x, tu) = \hat{G}_1^{(0)}(x, tu) \cdots \hat{G}_r^{(0)}(x, tu), & r \geq 2, \\ \deg(\hat{G}_i^{(0)}) = n_i, & \gcd(\hat{G}_i^{(0)}, \hat{G}_j^{(0)}) = 1 \quad (\forall i \neq j). \end{cases} \quad (8)$$

以後の手順は一般 Hensel 構成と同様である。すなわち、次式から Moses-Yun の補間式 $\hat{W}_i^{(l)}(x, tu)$ ($i = 1, \dots, r; l = 0, 1, \dots, n$) を計算する。

$$\begin{cases} \hat{W}_1^{(l)}(x, tu) \frac{\hat{F}_{\text{New}}(x, tu)}{\hat{G}_1^{(0)}(x, tu)} + \cdots + \hat{W}_r^{(l)}(x, tu) \frac{\hat{F}_{\text{New}}(x, tu)}{\hat{G}_r^{(0)}(x, tu)} = x^l, \\ \deg(\hat{W}_i^{(l)}(x, tu)) < \deg(\hat{G}_i^{(0)}(x, tu)) \quad (i = 1, \dots, r). \end{cases} \quad (9)$$

ついで、次の算式で $k = 1 \Rightarrow 2 \Rightarrow \dots$ と逐次的に $\hat{G}_i^{(k)}(x, v)$ ($i = 1, \dots, r$) を計算する。

$$\hat{D}^{(k)} \equiv F(x, tu) - \hat{G}_1^{(k-1)}(x, tu) \cdots \hat{G}_r^{(k-1)}(x, tu) \pmod{I_{k+1}} \quad (10)$$

$$= \hat{d}_{n-1}^{(k)}(tu) x^{n-1} + \cdots + \hat{d}_0^{(k)}(tu) x^0,$$

$$\hat{G}_i^{(k)} = \hat{G}_i^{(k-1)} + \delta \hat{G}_i^{(k)}, \quad \delta \hat{G}_i^{(k)} = \hat{W}_i \otimes \hat{D}^{(k)} \stackrel{\text{def}}{=} \sum_{l=0}^{n-1} \hat{W}_i^{(l)}(x, tu) \cdot \hat{d}_l^{(k)}(tu). \quad (11)$$

2.3 拡張 Hensel 構成の因子の性質

桁落ち誤差の定量的解析には拡張 Hensel 因子の性質が不可欠であるので、本章の残りが必要な性質のみを与えておく。下記の定理 3 と定理 4 は [SK99] による。

定理 2 一般 Hensel 構成と拡張 Hensel 構成の初期因子の間に $G_i^{(0)}(x) \Leftrightarrow \hat{G}_i^{(0)}(x, tu)$ なる対応があるとき、次式が成立する。

$$G_i^{(k)}(x, v) \equiv [\text{Taylor expansion of } \hat{G}_i^{(\infty)}(x, s+v)] \pmod{(v)^{k+1}}. \quad (12)$$

定理 3 (Sasaki-Kako, 1993) $\hat{G}_i^{(k)}(x, tu) = \hat{G}_i^{(k-1)}(x, tu) + \delta \hat{G}_i^{(k)}(x, tu)$ と表すとき、 $\delta \hat{G}_i^{(k)}(x, tu)$ は一般に u_1, \dots, u_ℓ の代数関数と有理関数を係数とする x と t^λ の同次多項式である。

定理 4 (Sasaki-Kako, 1993) 初期因子 $\hat{G}_i^{(0)}(x, tu)$ が x と t に関し多項式ならば、任意の k に対して $\hat{G}_i^{(k)}(x, tu)$ も x と t の多項式である。(ただし、変数 u_1, \dots, u_ℓ に関してはそうとは限らぬ)。

定理 3 によれば、拡張 Hensel 因子は全次数変数 t で見ると非常に明解になる。2 変数多項式 $F(x, u_1)$ の場合、拡張 Hensel 因子は u_1 の分数べき級数 = Puiseux 級数となるが、多変数の場合も t に関しては全く同じである。しかし、以下の例が示すように、従変数 u_1, \dots, u_ℓ に関しては複雑である。

例 1 拡張 Hensel 因子が係数に多項式と代数関数を同時に含む例

$$F(x, u) = x^2 - 2(u_1 + u_1^2)x + u_1^2 - u_2^2 - u_3^2 + 2u_1^3 + u_2^3 + u_3^3 + u_1^4. \quad (13)$$

$F(x, u)$ は原点で特異である。 $F(x, u)$ の Newton 多項式とその因数分解は以下となる。

$$\hat{F}_{\text{New}}(x, tu) = x^2 - 2tu_1x + t^2(u_1^2 - u_2^2 - u_3^2) = [x - t\theta_1(u)] \cdot [x - t\theta_2(u)].$$

ここで、 $\theta_i(u) = u_1 - (-1)^i \sqrt{u_2^2 + u_3^2}$ ($i = 1, 2$) である。上記因子を初期因子として拡張 Hensel 構成を 2 次まで行なうと次式を得る。

$$\begin{cases} \hat{G}_1^{(2)}(x, tu) = x - (tu_1 + t^2u_1^2) - \sqrt{u_2^2 + u_3^2} \cdot \left[t - t^2 \frac{u_2^3 + u_3^3}{2(u_2^2 + u_3^2)} + t^3 \frac{(u_2^3 + u_3^3)^2}{8(u_2^2 + u_3^2)^2} \right], \\ \hat{G}_2^{(2)}(x, tu) = x - (tu_1 + t^2u_1^2) + \sqrt{u_2^2 + u_3^2} \cdot \left[t - t^2 \frac{u_2^3 + u_3^3}{2(u_2^2 + u_3^2)} + t^3 \frac{(u_2^3 + u_3^3)^2}{8(u_2^2 + u_3^2)^2} \right]. \end{cases}$$

■

上記 Hensel 因子は全次数変数 t に関しては整数べき級数だが、係数が代数関数と有理関数となる例である。また、変数 u_1 に関しては多項式だが、変数 u_2 と u_3 に関して代数関数となる例でもある。そこで、因子の性質を分類するために integral と non-integral の概念を導入する。

定義 5 (integral, non-integral w.r.t. u_j) Let $\hat{G}_i^{(\infty)}(x, tu)$ be an extended Hensel factor of $F(x, u)$, and let $T_k(x, u)t^k$ be a term of $\hat{G}_i^{(\infty)}(x, tu)$, of degree k w.r.t. the total-degree variable t . If, for any k , $T_k(x, u)$ is a polynomial in variable u_j , then we say $\hat{G}_i^{(\infty)}(x, u)$ is integral w.r.t. u_j , otherwise we say $\hat{G}_i^{(\infty)}(x, u)$ is non-integral w.r.t. u_j .

(注釈) [SY98] は、拡張 Hensel 因子を Newton 線の傾きが整数か否かで crossing と branching に分類し、「crossing の場合は大きな桁落ちが生じるが、branching の場合は生じない」と結論した。しかし、以下に見るようにこれは間違いで、正しくは「integral か non-integral かで桁落ちが決まる」のである。なお、branching なら non-integral だが、crossing でも integral とは限らず non-integral の場合もある。

3 Moses-Yun の補間式の性質

Moses-Yun の補間式は、Hensel 構成で本質的な役割を演じるのみならず、桁落ち解析においても決定的に重要である。そこで、本章ではまず Moses-Yun の補間式の性質を明らかにする。そして、特異点近傍での一般 Hensel 構成の補間式 $W_i^{(l)}(x)$ と拡張 Hensel 構成の補間式 $\hat{W}_i^{(l)}(x, tu)$ との関係を与える。

定理 6 初期因子 $G_i^{(0)}(x)$ の倍数である任意の多項式 $A(x) = a_{n-1}x^{n-1} + a_{n-2}x^{n-2} + \dots + a_0$ に対し、次式が成立する。

$$W_i \otimes A = \sum_{l=0}^{n-1} W_i^{(l)}(x) \cdot a_l = 0 \quad (14)$$

定理 7 初期因子 $G_i^{(0)}(x)$ は無平方で、その根を $\zeta_{i,1}, \dots, \zeta_{i,n_i}$ とする：

$$G_i^{(0)}(x) = (x - \zeta_{i,1}) \cdots (x - \zeta_{i,n_i}) \quad (\zeta_{i,j} \neq \zeta_{i,j'} \text{ for } \forall j \neq j'). \quad (15)$$

このとき、Moses-Yun の補間式 $W_i^{(l)}(x)$ は根 $\zeta_{i,j}$ ($j = 1, \dots, n_i$) により次式で表される。

$$W_i^{(l)}(x) = \sum_{j=1}^{n_i} \frac{\zeta_{i,j}^l}{dF^{(0)}(x)/dx|_{x=\zeta_{i,j}}} \cdot \frac{G_i^{(0)}(x)}{(x - \zeta_{i,j})}. \quad (16)$$

(この定理は $G_i^{(0)}(x)$ が無平方でない場合に拡張できるが、非常に複雑になる。)

本章の残りでは、一般 Hensel 構成の展開点は特異点の δ 近傍にあるとする： $(s_1, \dots, s_\ell) = (\delta_1, \dots, \delta_\ell)$, $\delta \stackrel{\text{def}}{=} \sqrt{\delta_1^2 + \dots + \delta_\ell^2} \ll 1$ 。さらに、各数値 δ_i/δ をパラメータとみなし、 $\delta_i = \delta \cdot (\delta_i/\delta)$ の関係から δ をパラメータ $\delta_1/\delta, \dots, \delta_\ell/\delta$ の全次数とみなす。

拡張 Hensel 構成は特異点での展開であるから、特異点近傍での一般 Hensel 構成と密接な関係を持つはずである。そのことを Moses-Yun の補間式で見ると以下となる。

定義 8 (leading order terms) Let $H(x; \delta)$ be a polynomial in x with coefficients in Taylor or Laurent series in $\delta^{e/n}$. By $\text{lot}(H(x, \delta))$, or leading order terms of H , we denote the sum of all the terms of $H(x, \delta)$, of the smallest total-degree w.r.t. x and $\delta^{e/n}$.

定理 9 数値 δ は $0 < \delta \ll 1$ を満たし、 $F(x, \delta) = \prod_{i=1}^r G_i^{(0)}(x; \delta)$, $\hat{F}_{\text{New}}(x, tu) = \prod_{i=1}^r \hat{G}_i^{(0)}(x, tu)$ とする。さらに、 $G_i^{(0)}(x; \delta) \Leftrightarrow \hat{G}_i^{(0)}(x, \delta)$ ($i = 1, \dots, r$) なる対応の下で、次の関係式が成立する。

$$\text{lot}(W_i^{(l)}(x; \delta)) = \hat{W}_i^{(l)}(x, \delta) \quad (i = 1, \dots, r; l = 0, 1, \dots, n-1), \quad (17)$$

$$\|x^j\text{-term}[W_i^{(l)}(x; \delta)]\| = O(\delta^{-(n-l-n_i+j)e/n}) \quad (j = 0, 1, \dots, n_i-1). \quad (18)$$

4 項キャンセルレーション：四つの型

4.1 第一の型の項キャンセルレーション

この型は $r \geq 3$ の場合にのみ現れる。まず、算式 (5) と (6) より次の補題が得られる。

補題 10 $D^{(0)}(x) = F^{(0)}(x)$, $D^{(1)}(x, v) = F^{(1)}(x, v)$ とし、 $W_i \otimes F^{(0)} = G_i^{(0)}$ とするならば、 $D^{(k)}(x, v)$ ($k = 2, 3, \dots$) は次式により再帰的に計算できる (次式で、 $\sum_{k_1 + \dots + k_r = k}^{k-1}$ は制約 $k_1 + \dots + k_r = k$ をの下で各 k_i につき 0 から $k-1$ までの和をとることを表す)。

$$D^{(k)}(x, v) = F^{(k)}(x, v) - \sum_{k_1 + \dots + k_r = k}^{k-1} [W_1 \otimes D^{(k_1)} \dots W_r \otimes D^{(k_r)}]. \quad (19)$$

上式において、右辺 $[\dots]$ の中には、 $r \geq 3$ の場合、 $k_i = 0$ となるものが存在する (k_1, \dots, k_r に対する制約のため、 $r = 2$ の場合には存在しない)。たとえば $k_1 = 0$ ならば

$$[W_1 \otimes D^{(0)} \cdot W_2 \otimes D^{(k_2)} \dots W_r \otimes D^{(k_r)}] = G_1^{(0)} \cdot [W_2 \otimes D^{(k_2)} \dots W_r \otimes D^{(k_r)}]$$

となる。右辺は $G_1^{(0)}$ の倍数であるから、定理 6 より $W_1 \otimes [\text{上式}] = 0$ となる。すなわち、 $\delta G_1^{(k)}(x, v)$ の計算において上式にあたる部分はキャンセルする。

4.2 第二の型の項キャンセルレーション

この型は $G_i^{(0)}(x)$ と $G_j^{(0)}(x)$ ($i \neq j$) が近接根を持つときのみ現れる。算式 (6) の $W_i^{(l)}$ に表現 (16) を代入すれば次式が得られる。(16) は $G_i^{(0)}$ が無平方な場合のものであるが、無平方でない場合、重根は“非常に近接した根”であると見做すことにする)。

$$G_i^{(k)}(x, v) = G_i^{(k-1)}(x, v) + \sum_{j=0}^{n_i} \frac{D^{(k)}(\zeta_{i,j}, v)}{dF^{(0)}(x)/dx|_{x=\zeta_{i,j}}} \cdot \frac{G_i^{(0)}(x)}{(x - \zeta_{i,j})}. \quad (20)$$

これを算式 (19) 中の積項 $[W_1 \otimes D^{(k_1)} \dots W_r \otimes D^{(k_r)}]$ に代入すれば次の補題が得られる。

補題 11 算式 (19) 中の $[W_1 \otimes D^{(k_1)} \dots]$ において、 $k_1 \geq 1, \dots, k_q \geq 1$ ($q < r$)、かつ $k_{q+1} = \dots = k_r = 0$ であるならば、次式が成立する。

$$[W_1 \otimes D^{(k_1)} \dots W_r \otimes D^{(k_r)}] = \sum_{j_1=1}^{n_1} \dots \sum_{j_q=1}^{n_q} \left[\prod_{i=1}^q \frac{D^{(k_i)}(\zeta_{i,j_i}, v)}{dF^{(0)}(x)/dx|_{x=\zeta_{i,j_i}}} \right] \frac{G_1^{(0)}(x) \dots G_r^{(0)}(x)}{\prod_{i=1}^q (x - \zeta_{i,j_i})}. \quad (21)$$

$q \geq 2$ とし、簡単のため $G_1^{(0)}(x)$ と $G_2^{(0)}(x)$ が互いに近接根を持つとして、それらの近接根を $\zeta_{i,j_i}, \zeta_{i,j'_i}, \dots$ ($i = 1, 2$) とする。このとき、恒等式 $x - \zeta_{2,j'_2} = x - \zeta_{1,j_1} + \zeta_{1,j_1} - \zeta_{2,j'_2}$ を利用すれば、(21) の最後の因子は次のように分割できる。

$$\begin{aligned} \frac{G_1^{(0)}(x) \cdots G_r^{(0)}(x)}{\prod_{i=1}^q (x - \zeta_{i,j_i})} &= \frac{x - \zeta_{1,j_1} + \zeta_{1,j_1} - \zeta_{2,j'_2}}{x - \zeta_{2,j'_2}} \cdot \frac{G_1^{(0)}(x)}{x - \zeta_{1,j_1}} \cdot \frac{G_2^{(0)}(x) \cdots G_r^{(0)}(x)}{\prod_{i=2}^q (x - \zeta_{i,j_i})} \\ &= G_1^{(0)}(x) \cdot \frac{G_2^{(0)}(x) \cdots G_r^{(0)}(x)}{(x - \zeta_{2,j'_2}) \prod_{i=2}^q (x - \zeta_{i,j_i})} \\ &\quad + \frac{\zeta_{1,j_1} - \zeta_{2,j'_2}}{(x - \zeta_{2,j'_2})} \cdot \frac{G_1^{(0)}(x) \cdots G_r^{(0)}(x)}{\prod_{i=1}^q (x - \zeta_{i,j_i})}. \end{aligned} \quad (22)$$

定理 6 より、上式右辺の第一項は $G_1^{(k)}(x, v)$ の計算においてキャンセルする。一方、右辺の第二項は $\zeta_{1,j_1} \simeq \zeta_{2,j'_2}$ のため第一項に比べれば微小項である。すなわち、(21) 式右辺のいくつかの項は $G_i^{(k)}(x, v)$ の計算において“近似的にキャンセルする”と言える。

4.3 第三の型の項キャンセルレーション

多変数多項式の根として定まる代数関数は互いに共役なクラスに分割できる。同様に、拡張 Hensel 構成の因子全体も互いに共役なクラスに分けることができる。まず、共役因子の概念を定義する。

定義 12 (conjugate Hensel factors) For $\hat{n} \in \mathbf{Z}$, $\hat{n} \geq 2$, we define a mapping $\mathcal{R}_{\hat{n}}$ as follows, where \mathbf{A} is a field of algebraic functions in u_1, \dots, u_ℓ .

$$\begin{aligned} \mathcal{R}_{\hat{n}} : \mathbf{A}\{t^{1/\hat{n}}\}[x] &\longrightarrow \mathbf{A}\{t^{1/\hat{n}}\}[x], \\ G(x, t^{1/\hat{n}}, u) &\mapsto G(x, e^{2\pi i/\hat{n}} t^{1/\hat{n}}, u). \end{aligned} \quad (23)$$

Since $\mathcal{R}_{\hat{n}} F(x, tu) = F(x, tu) \in \mathbf{C}[x, tu]$, we may have $\hat{G}_{i_1}^{(k)}, \dots, \hat{G}_{i_m}^{(k)}$ satisfying

$$[\mathcal{R}_{\hat{n}} \hat{G}_{i_1}^{(k)}(x, tu)] \cdots [\mathcal{R}_{\hat{n}} \hat{G}_{i_m}^{(k)}(x, tu)] = \hat{G}_{i_1}^{(k)}(x, tu) \cdots \hat{G}_{i_m}^{(k)}(x, tu) \in \mathbf{A}[x, t]. \quad (24)$$

The Hensel factors $\hat{G}_{i_1}^{(k)}, \dots, \hat{G}_{i_m}^{(k)}$ which satisfy (24) and each of which contains truly fractional power terms w.r.t. t are called conjugate.

例 2 例 1 における $\hat{G}_1^{(2)}(x, tu)$ と $\hat{G}_2^{(2)}(x, tu)$ は互いに共役である。 ■

さて、展開点が特異点近傍の場合、一般 Hensel 因子は拡張 Hensel 因子と似た振舞いをするはずである。したがって、拡張 Hensel 構成で共役部分が完全にキャンセルすることに対応して、一般 Hensel 構成の対応する部分が近似的にキャンセルすることになる。

4.4 第四の型の項キャンセレーション

第三の型と同様、この型も展開点が特異点近傍の場合にのみ現れるが、どのように項がキャンセルするのか非常に分りにくい。3章におけると同様、本節においても展開点を $(s_1, \dots, s_\ell) = (\delta_1, \dots, \delta_\ell)$, $\delta \stackrel{\text{def}}{=} \sqrt{\delta_1^2 + \dots + \delta_\ell^2} \ll 1$ とし、 $\delta_i = \delta \cdot (\delta_i/\delta)$ の関係から δ をパラメータ見做して、 $\delta \rightarrow 0$ の極限を考える。

補題 13 $\delta \rightarrow 0$ のとき、ある $i \in \{1, \dots, r\}$ とあらゆる $l < \min\{n - n_1, \dots, n - n_r\}$ に対し、 $W_i^{(l)}(x)$ の係数のうち少なくとも一つは発散する。

補題 14 (5)において $D^{(k)}(x, v) \neq 0$ ならば、どれかの $i \in \{1, \dots, r\}$ に対し $W_i \otimes D^{(k)} \neq 0$ である。

さて、 $\hat{G}_i^{(\infty)}(x, tu)$ がたとえば変数 u_1 に関して integral である場合を考えよう。すると、 $D^{(k)}(x, v)$ は $x^l t^k u_1^k$ ($0 \leq l < n$) なる非零項を持つであろう。拡張 Hensel 構成においては $\delta \hat{G}_i^{(k)} = \hat{W}_i \otimes \hat{D}^{(k)}$ であり、この計算においては特異な部分はどこにも現れない。もしも $D^{(k)}$ の計算で主要項のキャンセレーションが起きなければ、 $\delta \rightarrow 0$ の極限を考えると、補題 13 と補題 14 によれば、一般 Hensel 構成の $\delta G_i^{(k)}$ の計算には発散する項が現れる。これは矛盾である。なぜなら、仮定より $\hat{G}_i^{(\infty)}(x, tu)$ は変数 u_1 に関して解析的だから、

$$\lim_{\delta \rightarrow 0} [G_i^{(\infty)}(x, v) \text{ の変数 } v_1 \text{ の部分}] = \lim_{\delta \rightarrow 0} [\hat{G}_i^{(\infty)}(x, v + \delta) \text{ の変数 } v_1 \text{ の部分}]$$

が成立するからである。このことから次の定理を得る。

定理 15 拡張 Hensel 構成の因子 $\hat{G}_i^{(\infty)}(x, tu)$ が u_1 に関して integral であり、算式 (10) の $\hat{D}^{(k)}(x, u)$ が非零項 $x^l t^k u_1^k$ ($0 \leq l < n$) を持つならば、算式 (5) の $D^{(k)}(x, v)$ の計算において δ の低次項のどれかがキャンセルする。

5 桁落ち誤差のオーダー解析

まず、前章に与えた四つの型の項キャンセレーション機構が大きな桁落ち誤差を引き起こすか否かを検討する。

第一の型は、例外的な場合を除き、大きな桁落ち誤差を引き起こさない。なぜなら、(19) の積項 $[W_1 \otimes D^{(k_1)} \dots W_r \otimes D^{(k_r)}]$ において、たとえば $k_1 = 0$ となる項があれば必ず $k_1 \neq 0$ となる項もあり、 $G_1^{(k)}(x, v)$ の計算で前者はキャンセルされるが後者はキャンセルされない。そして、例外的な場合を除き、両者は同じ程度の大きさの項を与えるからである。

第二の型は、実は大きな桁落ち誤差を引き起こさない (一見するとそうは思えないが)。理由は以下である。(22) において、右辺第二項は左辺に比べて次数が一つ低下しており、

$G_i^{(0)}(x)$ ($i = 1, \dots, r$) が近接根を持つとの条件下で x^j -項の係数を比較してみれば、低次項 (j が小さい項) は同じ程度の大きさになる。一方、算式 (6) によれば、 $\delta G_i^{(k)}$ は $D^{(k)}$ の各係数 $d_i^{(k)}$ に $W_i^{(l)}$ を掛けて計算される。したがって、算式 (6) の計算法では桁落ち誤差は実際には生じないのである。

第三の型は、ある条件下で大きな桁落ち誤差を引き起こす。共役因子どうしの積において、キャンセルする項がキャンセルしない項と重複しない場合には、最大の桁落ち誤差が生じる (すなわち、キャンセレーションの結果、有効数字が消えて誤差のみの係数が生じる)。 そうでなければ、桁落ち誤差が生じても、その大きさは展開次数 k には依存しない。その理由については [Sas00] を参照されたい。

第四の型については以下で詳細な解析を行なう。 ただし、展開点を特異点の近傍 (距離 δ) に選ぶ場合と、遠方 (距離 L) に選ぶ場合に限定し、展開点までの距離 $\delta \ll 1$ と $L \gg 1$ に関して誤差をオーダー評価する。さらに、一般の多項式に対する解析は非常に繁雑になるので、数学的本質を失わない範囲で $F(x, u)$ の構造を簡単なものに限定する。

5.1 $F(x, u)$ に対する制約

$F(x, 0) = f_n(0)x^n + \dots + f_m(0)x^m$ ($f_m(0) \neq 0$) の場合、 $\tilde{F}^{(0)} = f_n(0)x^{n-m} + \dots + f_m(0)$, $\hat{F}^{(0)} = x^m$ を初期因子として一般 Hensel 構成することにより、

$$F(x, u) \equiv \tilde{F}^{(k)}(x, u)\hat{F}^{(k)}(x, u) \pmod{(u)^{k+1}}$$

を満たす $\tilde{F}^{(k)}, \hat{F}^{(k)}$ を計算できる。前章の議論によれば、項キャンセレーションは特異点の近傍で生じ易い。したがって、桁落ちの観点からは $\hat{F}^{(k)}(x, u)$ のみを調べておけば数学的にはほぼ十分である。そこで、以下では $F(x, u)$ が $\hat{F}^{(k)}(x, u)$ と同じ振舞いをする と仮定する。このとき、 $|\delta| \ll 1$ に対し $F(x, \delta)$ は原点近傍に n 個の根を持つが、それらは同じオーダーの大きさであると仮定する。また、 $F(x, u)$ の数係数はほぼ 1 の大きさに規格化されているとする。以上をまとめて、 $F(x, u)$ に以下の制約を課す。

Restriction A $F(x, \delta)$ is square-free, i.e., it has no multiple root.

Restriction B The Newton line for $F(x, u)$ is $e_x/n + e_t/e_0 = 1$, that is, for each $i \in \{0, \dots, n\}$, e_i in (1) satisfies $i/n + e_i/e_0 \geq 1$. Below, we put $e = e_0$.

Restriction C The absolute value of each numerical coefficient c of $F(x, u)$ is not much different from 1, say, $|c| \leq 3$ and $|1/c| \leq 3$.

このとき、 $\text{ord}(f_n) = 0$, $\text{ord}(f_0) = e$ ゆえ、 $F(x, \delta)$ の根は $O(\delta^{e/n})$ となる。

$$\begin{cases} \zeta_{i,j} = O(\delta^{e/n}) \neq o(\delta^{e/n}) & (j=1, \dots, n_i), \quad i=1, \dots, r, \\ \zeta_{i,j} \neq \zeta_{i',j'} & (\forall i \neq i' \text{ and/or } \forall j \neq j'). \end{cases} \quad (25)$$

5.2 展開点が特異点近傍の場合

まず、初期因子 $G_i^{(0)}(x)$ と補間式 $W_i^{(l)}(x)$ の数係数のオーダー評価を行なう。 $G_i^{(0)}(x)$ と $W_i^{(l)}(x)$ を次のように表す。

$$\begin{cases} G_i^{(0)}(x) = g_{i,n_i}^{(0)} x^{n_i} + g_{i,n_i-1}^{(0)} x^{n_i-1} + \dots + g_{i,0}^{(0)} x^0, \\ W_i^{(l)}(x) = w_{i,n_i-1}^{(l)} x^{n_i-1} + w_{i,n_i-2}^{(l)} x^{n_i-2} + \dots + w_{i,0}^{(l)} x^0. \end{cases} \quad (26)$$

表現 (16) と (25) を用いれば、数係数 $g_{i,j}^{(0)}$ と $w_{i,j}^{(l)}$ の大きさは次のようにオーダー評価できる。(下式 (28) は (18) と同じである。すなわち、我々は二通りの方法で $w_{i,j}^{(l)}$ のオーダー評価を簡明に行なったのである)。

$$\begin{cases} g_{i,j}^{(0)} = O(\delta^{(n_i-j)e/n}) & (j=0, \dots, n_i), \\ g_{i,n_i}^{(0)} \neq o(\delta^0), \quad g_{i,0}^{(0)} \neq o(\delta^{n_i e/n}), \end{cases} \quad i=1, \dots, r, \quad (27)$$

$$w_{i,j}^{(l)} = O(\delta^{le/n - (n-1)e/n}) \cdot O(\delta^{(n_i-1-j)e/n}) = O(\delta^{-(n-l-n_i+j)e/n}). \quad (28)$$

次に $S^{(k)}(x, v) \stackrel{\text{def}}{=} [W_1 \otimes D^{(k_1)} \dots W_r \otimes D^{(k_r)}]$ と表し、 $S^{(k)}$ と $W_i \otimes S^{(k)}$ の数係数のオーダー評価を行なう。まず、(2) の $f_j^{(k)}(v)$ に対しては次のオーダー評価が成立する。

$$\|f_j^{(k)}\| = O(\delta^{\max\{0, e_j - k\}}), \quad j=0, 1, \dots, n. \quad (29)$$

このことから、 $W_i \otimes F^{(k)}$ の主要項 (数係数が大きな項) をなすのは積 $W_i^{(l)}(x) f_l^{(k)}(v)$ のうち l の小さなものであることが分り、 $W_i \otimes F^{(k)}$ は次のようにオーダー評価できる。

$$\begin{cases} W_i \otimes F^{(k)} = \phi_{i,n_i-1}^{(k)}(v) x^{n_i-1} + \phi_{i,n_i-2}^{(k)}(v) x^{n_i-2} + \dots + \phi_0^{(k)}(v) x^0, \\ \text{where } \|\phi_{i,j}^{(k)}\| = O(\delta^{(n_i-j)e/n - \min\{e, k\}}). \end{cases} \quad (30)$$

以上のオーダー評価より、次の補題が帰納的に証明できる。

補題 16 Restriction A, B, C の下で、もしも $D^{(1)}, D^{(2)}, \dots, D^{(k)}$ の主要項にキャンセルが生じなければ、 $S^{(k)}(x, v) \stackrel{\text{def}}{=} [W_1 \otimes D^{(k_1)} \dots W_r \otimes D^{(k_r)}]$ と $W_i \otimes S^{(k)}$ の係数部のオーダーは次式のようになる。

$$\begin{cases} S^{(k)}(x, v) = s_{n-q}^{(k)}(v) x^{n-q} + s_{n-q-1}^{(k)}(v) x^{n-q-1} + \dots + s_0^{(k)}(v) x^0, \\ \text{where } \|s_j^{(k)}\| = O(\delta^{(n-j)e/n - k}), \end{cases} \quad (31)$$

$$\begin{cases} W_i \otimes S^{(k)} &= \sigma_{i,n_i-1}^{(k)}(v)x^{n_i-1} + \sigma_{i,n_i-2}^{(k)}(v)x^{n_i-2} + \cdots + \sigma_0^{(k)}(v)x^0, \\ \text{where} & \|\sigma_{i,j}^{(k)}\| = O(\delta^{(n_i-j)e/n-k}). \end{cases} \quad (32)$$

$\hat{G}_i^{(\infty)}(x, tu)$ が変数 u_j に関して non-integral ならば、 $\hat{G}_i^{(\infty)}(x, tu)$ の項 $\hat{T}_\kappa(x, t\delta + tv)$, ただし $\text{tdeg}_t(\hat{T}_\kappa) = \kappa$, を変数 v_1, \dots, v_ℓ に関して Taylor 展開すると、展開項 $v_1^{e_1} \cdots v_\ell^{e_\ell}$ の数係数 c は $c = O(\delta^{\kappa - e_1 - \cdots - e_\ell})$ となる。この事より、 $\|\delta G_i^{(k)}(x, v)\| = O(\delta^{-k})$ となる。このオーダーは補題 16 で求めた $W_i \otimes S^{(k)}$ のオーダーと $O(\delta^c)$, ただし c は定数、の違いしかない。一方、 $\hat{G}_i^{(\infty)}(x, tu)$ が変数 u_j に関して integral ならば、上記 $\hat{T}_\kappa(x, t\delta + tv)$ の変数 u_j に関する部分は $\delta \rightarrow 0$ とともに有限の一定値に収束する。したがって、補題 16 を考慮すると、その部分の数係数には $O(\delta^{-k})$ の桁落ちが発生する。

定理 17 Restriction A, B, C の下で、もしも $\hat{G}_i^{(\infty)}(x, tu)$ が変数 u_j に関して integral ならば、 $D^{(k)}(x, v)$ の計算において変数 x_j の部分に $O(\delta^{-k})$ の桁落ち誤差が発生する。 $\hat{G}_i^{(0)}(x, tu)$ があらゆる変数 u_j ($1 \leq j \leq \ell$) に関して non-integral ならば、桁落ち誤差は高々 $O(\delta^c)$ 、ただし c は定数、の程度である。

5.3 展開点が遠方の場合

前章と同様、 $F(x, u)$ は原点に特異点を持つとし、展開点 (s_1, \dots, s_ℓ) は原点から遠方にあるとする：

$$(s_1, \dots, s_\ell) = (L_1, \dots, L_\ell), \quad L \stackrel{\text{def}}{=} \sqrt{|L_1|^2 + \cdots + |L_\ell|^2} \gg 1. \quad (33)$$

$F(x, L + v)$ は (2) で定義されるが、今の場合、 $\|f_j^{(k)}\|$ のオーダー評価は次式となる。

$$\|f_j^{(k)}\| = O(L^{E_j - k}), \quad j = 0, 1, \dots, n. \quad (34)$$

特異点の近傍とは異なり、 $F^{(0)}(x), F^{(1)}(x, v), \dots$ の数係数は巨大になる。そこで、次なるスケール変換を導入し、前章までの解析と並行に解析を行なうことを試みる。

定義 18 We define a scale transformation \mathcal{T}_S , with S a positive number, which acts on $F^{(k)}(x, v)$ ($k = 0, 1, \dots, E$), as follows.

$$\mathcal{T}_S F^{(k)}(x, v) \stackrel{\text{def}}{=} S^{-\text{deg}(F)} F^{(k)}(Sx, v) = \sum_{j=0}^n S^{j-n} x^j f_j^{(k)}(v). \quad (35)$$

以後、 $F^{(k)}, G_i^{(0)}$ 等に変換 \mathcal{T}_S を施した結果を $\tilde{F}^{(k)}, \tilde{G}_i^{(0)}$ 等と表す。 \mathcal{T}_S により $G_i^{(0)}, W_i^{(l)}, W_i \otimes F^{(k)}, D^{(k)}$ はそれぞれ次のように変換される。

$$\tilde{G}_i^{(0)}(x) = S^{-n_i} G_i^{(0)}(Sx), \quad n_i = \deg(G_i^{(0)}) \quad (i = 1, \dots, r), \quad (36)$$

$$\tilde{W}_i^{(l)}(x) = S^{n-n_i-l} W_i^{(l)}(Sx) \quad (l = 0, 1, \dots, n-1; i = 1, \dots, r), \quad (37)$$

$$\tilde{W}_i \otimes \tilde{F}^{(k)} = S^{-n_i} \cdot [W_i^{(n-1)}(Sx)f_{n-1}^{(k)}(v) + \dots + W_i^{(0)}(Sx)f_0^{(k)}(v)]. \quad (38)$$

これらを用いて算式 (19) を再帰的に調べることにより、次の定理が得られる。

定理 19 2.1 節に記述した一般 Hensel 構成手順で引き起こされる桁落ち誤差は変換 \mathcal{T}_S の下で不変である。

これまで S の値は任意であったが、ここで、本章での解析が前章までの解析と並行的に行くように S の値を定める。それには、定義 1 に述べたように $F(x, tu)$ の各項を (e_x, e_t) -平面上にプロットし、全てのプロット点を囲む最小の凸包、すなわち Newton 多角形 \mathcal{N} を考える。図 1 は Newton 多角形 \mathcal{N} を図示したものである。

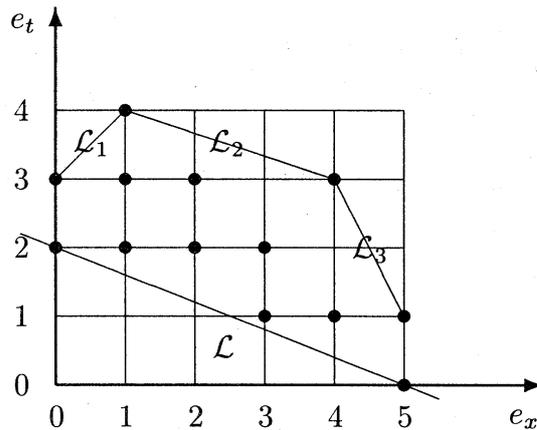


図 1. Illustration of the Newton polygon
Line \mathcal{L} is the Newton line.

Newton 多角形 \mathcal{N} の上辺の各頂点を以下とする： $0 = i_0 < i_1 < \dots < i_{\kappa-1} < i_\kappa = n$.

$$(0, E_0) = (i_0, E_0), (i_1, E_{i_1}), \dots, (i_{\kappa-1}, E_{i_{\kappa-1}}), (i_\kappa, E_{i_\kappa}) = (n, E_n), \quad (39)$$

このとき、最も右側の辺（図 1 では辺 \mathcal{L}_3 ）の傾きが 0 になるように S の値を決める：

$$S = L^{(E_{i_{\kappa-1}} - E_n)/(n - i_{\kappa-1})} \quad (40)$$

たとえば図 1 では、 $n = 5, i_{\kappa-1} = 4, E_n = 1, E_{i_{\kappa-1}} = 3$ ゆえ、 $S = L^2$ と決める。

参 考 文 献

- [Kuo89] T.-C. Kuo: *Generalized Newton-Puiseux theory and Hensel's Lemma in $C[[x, y]]$* , Can. J. Math., Vol. XLI, 1101-1116 (1989).
- [MY73] J. Moses & D. Y. Y. Yun: *The EZGCD algorithm*, Proc. 1973 ACM National Conference, 159-166 (1973).
- [Mus71] D. R. Musser: *Algorithms for polynomial factorizations*, Ph.D. Thesis, University of Wisconsin, 1971.
- [Sas00] T. Sasaki: *Mechanism of cancellation errors in multivariate Hensel construction with floating-point numbers*, Preprint of Univ. Tsukuba (in preparation).
- [SK99] T. Sasaki & F. Kako: *Solving multivariate algebraic equation by Hensel construction*, Japan J. Indus. Appl. Math., **16**, 257-285 (1999). (This paper was submitted in March, 1993, and the authors recieved referees' reports in Sep., 1996 and the letter of acceptance in June, 1998.)
- [SY98] T. Sasaki & T. Yamaguchi: *An analysis of cancellation errors in multivariate Hensel construction with floating-point number arithmetic*, Proc. of ISSAC'98, ACM Press, pp. 1-8 (1998).