ILLINOIS JOURNAL OF MATHEMATICS
Volume 26, Number 1, Spring 1982

ON THE p-ADIC ANALYTICITY OF SOLUTIONS OF LINEAR DIFFERENTIAL EQUATIONS

BY

E. BOMBIERI AND S. SPERBER

I. Introduction

Over the complex numbers a linear differential equation with analytic coefficients has a full set of solutions at an ordinary point, which converge up to the nearest singularity. The equation $y' - y = 0$ with solution

$$e^z = \sum_{m=0}^{\infty} \frac{z^m}{m!}$$

shows that this principle fails p-adically; indeed, the p-adic radius of convergence of the exponential series is $p^{-1/(p-1)}$. In this paper we investigate this phenomenon and relate it to the singularity structure of the differential equation; for example, solutions to equations with irregular singularities behave in this respect like e^z, for almost all p.

Our notation is as follows:

K is a number field;

$\bar{K}$ is an algebraic closure of K;

$\mathscr{D} = K(z)[D]$, where $D = d/dz$ (ring of linear differential operators with coefficients in $K(z)$);

$L = \frac{D^n}{n!} - \sum_0^{n-1} G_j(z)\frac{D^j}{j!}$, an element of $\mathscr{D}$;

Sing (L) is the set of singularities of $L \in \mathscr{D}$;

v is the non-archimedean valuation of K, with residue field of characteristic p;

Ω_v is an algebraically closed, complete extension of K with valuation extending v, containing a unit t_v whose image in the residue class field is transcendental over the residue class field of k;

$| \ |_v$ is the absolute value in Ω_v;

$D_v(t, r-) = \{x \in \Omega_v \colon |x - t|_v < r\}$, the disk of center t and radius r.

The purpose of introducing a generic unit t_v is to exploit the following property. If $f(z) = \sum_{m=0}^{N} a_m z^m \in \bar{K}[z]$, then

$$(1) \qquad |f(t_v)|_v = \sup_m |a_m|_v = \sup\{|f(z)|_v \mid |z|_v \le 1\}.$$

Received January 22, 1980.

We also define $r_v(t)$ to be the v-adic radius of convergence of ker (L) at $t \in \Omega_v$, $t \notin$ sing (L), and we let $r_v = r_v(t_v)$, the generic radius of convergence. We may note that, by (1), $r_v(t_v)$ is independent of the generic center t_v, thus justifying our notation r_v. Moreover, we shall write $r_v(t; L)$, $r_v(L)$ for $r_v(t)$, r_v if we want to emphasize their dependence on the operator L.

It follows from [4] and [7] that

(2) $$r_v \geq |p|_v^{1/(p-1)}$$

for all v for which the coefficients of L (supposed monic) are bounded by 1 in the disk $D_v(t_v, 1-)$. This is the case for all v except for an effectively computable finite set, say S_0. It was discovered by N. Katz [3] that the global nilpotence of the p-curvature for the connection defined by L (for definitions, see [3]) imposes restrictions on the analytic behavior of L. In particular, in the course of his proof of the main theorem in [3], he proves that in this case the differential operator L has only regular singular points and rational exponents. This condition (of global nilpotence of the p-curvature) can be rephrased in our terminology as

$$r_v > |p|_v^{1/(p-1)}$$

for almost all v. We have been informed by Dwork that another proof of this result appeared in some seminar notes [2] by Honda in 1974.

The object of this paper is to give a third proof of this result, based on the idea of blowing up the differential operator in a neighborhood of a singular point. Our techniques will be p-adic, rather than char p techniques as in previous proofs. More precisely, we will give new proofs of the following results:

Theorem 1. *There is an effectively computable finite set S (depending on L) such that if L has at least one irregular singular point then*

$$r_v = |p|_v^{1/(p-1)} \quad \textit{for } v \notin S.$$

Theorem 2. *If L has only regular singularities and at least one irrational exponent, then $r_v = |p|_v^{1/(p-1)}$ for infinitely many v. More precisely, the set of prime numbers p for which there is such a v has positive density.*

We thank B. Dwork and N. Katz for helpful discussions and Y. Sibuya for his help in extending the results of Theorems 1 and 2 to the case of systems treated in Section V.

II. Formal theory

The object of this section is the study of the differential operator

$$L = \frac{D^n}{n!} - \sum_{j=0}^{n-1} G_j \frac{D^j}{j!}$$

with $G_j = G_j(z) \in K(z)$, in the neighborhood of the origin. Let us write

$$G_j = \frac{\lambda_j}{z^{\delta_j}} + \text{higher order terms}$$

with $\lambda_j \neq 0$, and let us define with Poincaré [5] the rank ρ of L at 0 to be

$$\rho = \max_{0 \leq j < n} \left(\frac{\delta_j}{n-j}, 1 \right).$$

The well-known criterion of Fuchs says that 0 is an irregular singular point if and only if $\rho > 1$; we may think of ρ as giving a measure of the complication of the singularity of L at 0. We also define J to be the set

$$J = \{0 \leq j < n \colon \delta_j = \rho(n-j)\}.$$

J is non-empty if $\rho > 1$.

Let L_m, $m \geq 0$, be the unique differential operator defined by the two conditions

(a) $L_m \in \mathscr{D}L$,

(b) $\deg \left(L_m - \dfrac{D^m}{m!} \right) < n$;

we have $L_n = L$ and $L_m = 0$ if $m < n$. We write

$$L_m = \frac{D^m}{m!} - \sum_{j=0}^{n-1} G_{m,j} \frac{D^j}{j!},$$

so that $G_{n,j} = G_j$ for all j and $G_{m,m} = 1$, $G_{m,j} = 0$ if $j \neq m$, in the case $m < n$. It is immediate that the $G_{m,j}$ satisfy the recurrence

$$G_{m+1,j} = \frac{1}{m+1} \{DG_{m,j} + jG_{m,j-1} + nG_{m,n-1} G_j\}.$$

The recursion implies by induction on m that $\delta_{m,j} \leq \rho(m-j)$, hence we may write

$$G_{m,j} = \frac{\gamma_{m,j}}{z^{\rho(m-j)}} + \text{higher order terms},$$

where of course $\gamma_{m,j} = 0$ if $\rho(m-j)$ is not an integer and $\gamma_{n,j} = \lambda_j$ if $j \in J$; moreover for $m < n$ we have $\gamma_{m,m} = 1$ and $\gamma_{m,j} = 0$ if $j \neq m$.

We define the blowing up of L at $z = 0$ to be the constant coefficient differential operator

$$\tilde{L} = \frac{D^n}{n!} - \sum_{j \in J} \lambda_j \frac{D^j}{j!}$$

in the case $\rho > 1$, and the differential operator of Euler type

$$\tilde{L} = \frac{D^n}{n!} - \sum_{j \in J} \frac{\lambda_j}{(1+z)^{n-j}} \frac{D^j}{j!}$$

in the case $\rho = 1$.

LEMMA 1. *The formal power series*

$$v_j(z) = z^j + \gamma_{n,j} z^n + \gamma_{n+1,j} z^{n+1} + \cdots$$

satisfy $\tilde{L} v_j = 0$, *for* $j = 0, 1, \ldots, n-1$.

Proof. A simple-minded argument consists in noting that the statement of Lemma 1 is an assertion that certain recurrences for the $\gamma_{m,j}$ hold and deducing them from the recurrences for the $G_{m,j}$.

We present instead another proof, which will justify our terminology for $\tilde{L}$. Let t be an indeterminate and let us consider a set of formal solutions u_j, $j = 0, 1, \ldots, n-1$ of $Lu = 0$, with the initial conditions at $z = t$

$$\left(\frac{D^h}{h!} u_j\right)(t) = \begin{cases} 1 & \text{if } h = j, \\ 0 & \text{if } h \neq j, \end{cases}$$

for $h = 0, 1, \ldots, n-1$. Now $L_m u_j = 0$, hence

$$\left(\frac{D^m}{m!} u_j\right)(t) = \sum_{h=0}^{n-1} G_{m,h}(t) \left(\frac{D^h u_j}{h!}\right)(t) = G_{m,j}(t),$$

and Taylor's formula yields the formal series expansion

$$\begin{aligned} u_j(t + \zeta) &= \sum_{m=0}^{\infty} G_{m,j}(t) \zeta^m \\ &= \sum_{m=0}^{\infty} \left(\frac{\gamma_{m,j}}{t^{\rho(m-j)}} + \text{higher order terms}\right) \zeta^m. \end{aligned}$$

If we set $\zeta = t^\rho z$ we obtain

$$t^{-\rho j} u_j(t + t^\rho z) = v_j(z) + t^{1/b} \psi_j(t^{1/b}, z)$$

where the natural integer b is such that $b\rho$ is an integer and where $\psi_j(\tau, z)$ is a formal power series in τ and z. Now we see that $v_j(z)$ is obtained from $t^{-\rho j} u_j(t + t^\rho z)$ by specializing t to 0. Since $Lu_j = 0$, the change of variables $z \to t + t^\rho z$ yields

$$\tilde{L}_t(t^{-\rho j} u_j(t + t^\rho z)) = 0$$

where

$$\tilde{L}_t = \frac{D^n}{n!} - \sum_{h=0}^{n-1} G_h(t + t^\rho z) t^{(n-h)\rho} \frac{D^h}{h!}.$$

The lemma follows by noting that $\tilde{L}_t$ specializes to $\tilde{L}$ when t specializes to 0.

We conclude this section with the remark that we may interpret the above proof as performing a blowing up transformation

$$(z, u) \to \left(\frac{z-t}{t^{\rho}},\ t^{-\rho j}u\right),$$

for $t \to 0$, on the graph of the mapping $z \to u_j(z)$ and obtaining the differential operator associated to the blown up graph.

III. Local theory

We work here over the field Ω_v. Our aim is to obtain information on the generic radius r_v of L, in terms of the simpler operator $\tilde{L}$ defined in the previous section.

If $\xi \in \operatorname{Sing}(L)$, $\xi \neq \infty$, we define

$$V_1(\xi) = \{v\colon |\xi|_v \leq 1, \text{ and } |\eta - \xi|_v \geq 1 \text{ for } \eta \in \operatorname{Sing}(L), \eta \neq \xi, \infty\};$$

while if $\xi = \infty \in \operatorname{Sing}(L)$, we define

$$V_1(\xi) = \{v\colon |\eta|_v \leq 1 \text{ for all } \eta \in \operatorname{Sing}(L), \eta \neq \infty\}.$$

In either case we define the finite ser $S_1(\xi)$ to be the complement of $V_1(\xi)$. Our purpose in making these definitions is the following. Let $f \in \bar{K}(z)$ be a rational function with poles contained in Sing (L) and let the Laurent expansion of f at ξ be

$$f(z) = \begin{cases} \sum_{-N}^{\infty} a_m(z-\xi)^m & \text{if } \xi \neq \infty, \\ \sum_{-N}^{\infty} a_m z^{-m} & \text{if } \xi = \infty. \end{cases}$$

Let $v \notin S_1(\xi)$. Then we have

$$|f(t_v)|_v \geq \sup_m |a_m|_v, \tag{3}$$

which we explain as follows. If we factor out the pole ξ from f, then

$$f(z) = \frac{1}{(z-\xi)^N} h(z)$$

with $h(z)$ analytic in $D_v(\xi, 1^-)$. Since $|t_v - \xi|_v = 1$, we have for all r, $0 < r < 1$,

$$|f(t_v)|_v = |h(t_v)|_v \geq \sup_{|z-\xi|_v = r} |h(z)|_v = \sup_{m \geq -N} \{|a_m|_v r^{m+N}\}. \tag{3'}$$

The inequality in (3′) follows from (1) and the maximum principle applied to the numerator of $h(z)$; the absolute values of the denominator of $h(z)$ are the same on both sides of the inequality by virtue of the assumption $v \notin S_1(\xi)$. The inequality (3) then follows from the continuity of the right side of (3′) as $r \to 1^-$.

If we apply the results of the previous section we obtain information about the behavior of L at $z = 0$, and we want to do so for every $z = \xi$, $\xi \in \text{Sing}(L)$. We denote by L_ξ the differential operator obtained from L after the change of variable $z' = z - \xi$ if $\xi \neq \infty$, $z' = 1/z$ if $\xi = \infty$. We recall that by our notation $r_v(L_\xi)$ is the generic radius of convergence of L_ξ and $r_v(0; \tilde{L}_\xi)$ is the radius of convergence of Ker $(\tilde{L}_\xi)$ at the origin.

LEMMA 2. *If $\xi \in \text{Sing}(L)$ and $v \notin S_1(\xi)$ then*

$$r_v(0; \tilde{L}_\xi) \geq r_v(L_\xi).$$

Proof. For notational convenience, we may and shall assume that $\xi = 0$ and hence $L_\xi = L$. By the hypothesis $v \notin S_1(\xi)$, inequality (3) yields

$$|G_{m,j}(t_v)|_v \geq |\gamma_{m,j}|_v$$

for every m, j. Now Lemma 2 follows from Lemma 1 and Hadamard's formula.

LEMMA 3. *Let us assume $v \notin S_1(\xi)$. If $\xi \in \text{Sing}(L)$, $\xi \neq \infty$, then*

$$r_v(L_\xi) = r_v.$$

If $\xi = \infty \in \text{Sing}(L)$, then

$$\min(1, r_v(L_\infty)) = \min(1, r_v).$$

Proof. The first assertion is a consequence of the fact that ξ is algebraic and $|\xi|_v \leq 1$, hence $\xi + t_v$ remains a generic unit. In order to prove the second assertion we note that t_v is a generic unit if and only if t_v^{-1} is a generic unit. It follows that $|z - t_v|_v = r < 1$ if and only if

$$\left|\frac{1}{z} - \frac{1}{t_v}\right|_v = r < 1.$$

This shows that $r_v(L_\infty) \geq \min(1, r_v)$. Since the change of variables $z' = 1/z$ is involutory, this inequality proves Lemma 3.

COROLLARY. *Let v, ξ, L, L_ξ, $\tilde{L}_\xi$ be as before. Then*

$$\min(1, r_v) \leq \min(1, r_v(0; \tilde{L}_\xi))$$

for all $\xi \in \text{Sing}(L)$.

IV. Proof of theorems

Let us assume $\xi \in \text{Sing}(L)$ is an irregular singular point. If so, $\tilde{L}_\xi$ is a constant coefficient differential operator and a basis of Ker $(\tilde{L}_\xi)$ is $\{e^{\alpha z} z^k\}$ where α runs over the roots of the characteristic polynomial

$$P(x) = \frac{x^n}{n!} - \sum_{j \in J} \lambda_j \frac{x^j}{j!}$$

and $k = 0, 1, \ldots, k_\alpha - 1$ with k_α the multiplicity of α. We remarked earlier that, since $\rho > 1$, the set J is non-empty, thus $P(x)$ has at least one non-zero root α. Now $e^{\alpha z}$ has radius of convergence exactly $|\alpha|_v^{-1} |p|_v^{1/(p-1)}$, while α is a unit for all v outside an effectively computable finite set $S_2(\xi)$. This shows that

$$r_v(0; \tilde{L}_\xi) \leq |p|_v^{1/(p-1)} \quad \text{if } v \notin S_2(\xi).$$

If we combine this last inequality with the corollary to Lemma 3 and inequality (2), we obtain Theorem 1.

The proof of Theorem 2 is along similar lines. We assume that ξ is a regular singular point, hence $\rho = 1$. Now the differential operator $\tilde{L}_\xi$ is of Euler type and a basis of Ker $(\tilde{L}_\xi)$ is

$$\{(1 + z)^\alpha (\log (1 + z))^k\},$$

where α runs over the roots of the indicial polynomial of $\tilde{L}_\xi$ at $z = -1$ (which is identical with the indicial polynomial of L at ξ) and $k = 0, 1, \ldots, k_\alpha - 1$, with k_α the multiplicity of α. In order to compute the radius of convergence of

$$(1 + z)^\alpha (\log (1 + z))^k,$$

we note that $\log (1 + z)$ has radius of convergence equal to 1 for all v, which reduces the question to the study of the binomial series for $(1 + z)^\alpha$. If

$$(4) \qquad |\alpha - l|_v = 1 \quad \text{for every } l \in Z$$

then $(1 + z)^\alpha$ has radius of convergence exactly $|p|_v^{1/(p-1)}$ and, exactly as before, the proof of Theorem 2 will be complete, provided the set of such v has positive density whenever α is irrational. In that case $Q(\alpha) \neq Q$ and v will satisfy (4) whenever the reduction $\bar{\alpha}$ of α mod v will not be in the prime field F_p. If p does not split completely in $Q(\alpha)$, at least one v over p will verify (4) and the set of such p has density

$$1 - \frac{1}{[E: Q]} \geq 1 - \frac{1}{[Q(\alpha): Q]},$$

where E is the smallest Galois extension of Q containing $Q(\alpha)$.

V. First-order systems

Let $A(z)$ be an $n \times n$ matrix with entries in $K(z)$ and consider the differential system

$$(5) \qquad DY = A(z)Y$$

where Y denotes an n-vector. The existence of a "cyclic vector" [1, Chapter II, Lemma 1.3] implies that the system may be reduced to a scalar equation by a change of variables $W = B(z)Y$ where W is an n-vector and $B(z) \in GL(n, K(z))$. The drawback of this procedure is that determination of $B(z)$ is not presently effective.

One may also proceed by means of elementary divisors (cf. [6, Chapter III, Section 11]) determining effectively $n \times n$ matrices G and H each having entries in the ring $\mathscr{D}$ and determinant in $K(z)$ not identically zero, such that

$$C = G \cdot (D - A(z)) \cdot H$$

is a diagonal matrix with entries in $\mathscr{D}$. In fact, since the kernel of $D - A(z)$ is a vector space of dimension n over K, we may conclude that the sum of the degrees (in D) of the diagonal entries of C is n. The next two theorems follow immediately.

THEOREM 1A. *There is an effectively computable finite set of primes S such that if the system* (5) *has an irregular singular point, then $r_v = |p|_v^{1/(p-1)}$ for $v \notin S$.*

THEOREM 2A. *If the system* (5) *has only regular singularities and at least one irrational exponent, then $r_v = |p|_v^{1/(p-1)}$ for infinitely many v. More precisely, the set of prime numbers p for which there is such a v has positive density.*

One may also proceed effectively by the results of Hukuhara and Turittin. By their work, there exists a positive integer b and an $n \times n$ matrix $P(\tau)$ with entries in $F[\tau]$ (where F is a finite extension of K) satisfying:

(i) det $P(z^{1/b})$ is not identically zero.
(ii) The transformation $Y = P(z^{1/b})X$ takes the given system (5) into

$$z^{\rho}DX = B(z^{1/b})X \tag{6}$$

where $\rho \geq 1$, $b\rho$ is an integer, and $B(\tau)$ is an $n \times n$ matrix with entries in $F(\tau)$ and analytic at $\tau = 0$.
(iii) If $\rho > 1$, then $B(0)$ is a non-nilpotent matrix.

The singularity at $z = 0$ is irregular if and only if $\rho > 1$.

We may perform our blowing-up transformation on the system (6); in particular sending $z \to t + t^{\rho}z$ and letting $t \to 0$, we obtain

$$(1 + z)DX = B(0)X \quad \text{if } \rho = 1;$$

if instead $\rho > 1$ then $DX = B(0)X$ and $B(0)$ is non-nilpotent.

The main purpose in working with the transformed system (6) rather than the given system (5) is to ensure that in the case $\rho > 1$ the constant matrix $B(0)$ has a non-zero eigenvalue. Now the proof of Theorems 1A and 2A for system (6) can be carried out exactly as in the scalar case. The proof of these results for system (5) will follow from the fact that the integer b, the finite extension F and the matrix $P(\tau)$ are all effectively computable.

REFERENCES

1. P. DELIGNE, *Equations Differentielles à Points Singuliers Réguliers*, Lecture Notes in Mathematics 163, Springer-Verlag, Berlin, 1970.
2. T. HONDA, *Algebraic differential equations*. Proceedings of the Severi Centennial Conference, Rome (1979), to appear.

3. N. KATZ, *Nilpotent connections and the monodromy theorem: applications of a result of Turrittin*, Publ. Math. I.H.E.S., vol. 39 (1970), pp. 175–232.
4. E. LUTZ, *Sur l'equation* $y^2 = x^3 - Ax - B$ *dans les corps* p*-adiques*, J. Reine Angew. Math., vol. 177 (1937), pp. 238–247.
5. H. POINCARÉ, *Sur les Integrales Irregulières des Équations Linéaires*, Acta Math, Vol. 8 (1886), pp. 295–344.
6. E. G. C. POOLE, *Introduction to the theory of linear differential equations*, Oxford University Press, London, 1936.
7. A. WEIL, *Sur les fonctions elliptiques* p*-adiques*, C. R. Acad. Sci. Paris, vol. 203 (1936), pp. 22–24.

INSTITUTE FOR ADVANCED STUDY
PRINCETON, NEW JERSEY

UNIVERSITY OF MINNESOTA
MINNEAPOLIS, MINNESOTA