

A NOTE ON THE p -ADIC CONVERGENCE OF SOLUTIONS OF LINEAR DIFFERENTIAL EQUATIONS

D. N. CLARK

1. Introduction. In this note we consider an ordinary linear differential equation

$$(1) \quad l(y) = \sum_{j=0}^n A_j(x) \frac{d^j y}{dx^j} = 0$$

whose coefficients are power series in one variable with nonzero radius of convergence. Our main interest lies in the p -adic behavior of formal power series solutions of (1) near a singular point. In the classical case it is known [5, Chapter IV, Theorem 11.3] that in the vicinity of a regular singular point, each formal power series solution has a radius of convergence for which a nonzero lower bound may be given. It is also known in the classical case that this need not hold if the singularity is not regular. We shall show that in the p -adic case, a formal power series solution has a nonzero radius of convergence (provided, for example, the coefficients of each A_j are algebraic numbers). Results of this type are useful in the application of Dwork's deformation theory [3] to singular hypersurfaces.

It has been noted by Igusa [6] that the classical existence theorems are valid in the p -adic case in the vicinity of an ordinary point. Professor Igusa's brief reference to the "calcul des limites" may give the erroneous impression that a comparison with a "dominating" differential equation is useful. We take this opportunity to clarify the theory of ordinary points.

Let Z denote the ordinary integers, Z_+ the positive integers, Z_- the negative integers, Q' the rational p -adic numbers, \mathfrak{O}' the ring of integers of Q' , Ω the completion of the algebraic closure of Q' and \mathfrak{O} the ring of integers of Ω . The p -adic valuation of Ω will be denoted multiplicatively ($|p| = 1/p$) in §2 and additively ($\text{ord } p = 1$) in the remainder of this note.

2. Ordinary points. Let $a > 0$, $b > 0$ and let U be the polycylinder of all $(x, y) \in \Omega \times \Omega$ such that $|x| \leq a$, $|y| \leq b$. Let F be the space of all holomorphic functions, f , on U . It is known that under the sup norm, $\|f\| = \sup_{(x, y) \in U} |f(x, y)|$, F is a p -adic Banach space [8]. Letting $D_x = \partial/\partial x$, $D_y = \partial/\partial y$, we note that the well-known Cauchy inequality means that if (x_0, y_0) is a center of U then the linear function

Received by the editors January 25, 1965.

$f \rightarrow (i!j!)^{-1}(D_x^i D_y^j f)(x_0, y_0)$ has norm not greater than $a^{-i}b^{-j}$ for each pair (i, j) of non-negative integers. The ultra-metric norm of Ω implies that each point of U is a center of U and hence the endomorphism $f \rightarrow D_x^i D_y^j f$, has operator norm, $\|D_x^i D_y^j\|$, with upper bound,

$$(2) \quad \|D_x^i D_y^j\| \leq |i!j!| a^{-i} b^{-j}.$$

Thus, in particular, if $g \in F$, $\|g\| = M$, then $\|D_x\| \leq a^{-1}$, $\|gD_y\| \leq Mb^{-1}$ and hence

$$(3) \quad \|D_x + gD_y\| \leq \text{Max}(a^{-1}, Mb^{-1}).$$

Certainly, the differential equation

$$(4) \quad \frac{dy}{dx} = g(x, y),$$

has unique formal power series solution with no constant term,

$$(5) \quad y = B_1x + B_2x^2 + \dots$$

and precisely as in the "calcul des limites" [4, p. 46, equation (2)], for each $j \geq 1$,

$$(6) \quad j!B_j = \phi[(D_x + gD_y)^{j-1}g],$$

where ϕ is the linear function: $f \rightarrow f(0, 0)$ on F . Since $|j!| \geq p^{-j/(p-1)}$, it follows from (3) and (6) that the power series (5) converges for $|x| \leq p^{-1/(p-1)} \text{Min}(a, b/M)$ and assumes values in the disk $|y| < b$. The generalization may be stated without proof.

THEOREM 1. *Let $a > 0$, $b_1 > 0$, \dots , $b_n > 0$ and let U be the polycylinder $U = \{(x, y_1, \dots, y_n) \mid |x| \leq a, |y_i| \leq b_i, 1 \leq i \leq n\}$. Let f_1, \dots, f_n be holomorphic functions on U , $|f_i| \leq M_i$ for $i = 1, 2, \dots, n$. Let $(\xi, \eta_1, \dots, \eta_n) \in U$, and let $y = (y_1, \dots, y_n)$, $\eta = (\eta_1, \dots, \eta_n)$, then the initial value problem*

$$\frac{dy}{dx} = (f_1, f_2, \dots, f_n)(x, y),$$

$$y(\xi) = \eta,$$

has a unique solution holomorphic in the disk,

$$|x| \leq p^{-1/(p-1)} \text{Min}(a, b_1/M_1, \dots, b_n/M_n).$$

3. Asymptotic estimate. For $\alpha \in \Omega$, $s \in \mathbb{Z}_+$, let

$$(7) \quad \theta(\alpha, s) = \alpha^{-1} \prod_{j=0}^s (\alpha + j),$$

it being understood that $\theta(0, s) = s!$ and that in the product the factor corresponding to $j = -\alpha$ is to be deleted if $j < 0$, $j \in \mathbb{Z}_-$.

The object of this paragraph is to find an asymptotic formula for $\text{ord } \theta(\alpha, s)$.

DEFINITION 1. An element, $\alpha \in \Omega$, is said to be (p -adically) non-Liouville if for $s \in \mathbb{Z}$, $s \rightarrow +\infty$, we have

$$(8) \quad \text{ord}(\alpha + s) \leq O(\log s).$$

Let \mathfrak{L} be the set of all non-Liouville elements of Ω . Clearly \mathfrak{L} contains the complement of \mathfrak{D}' in Ω . Furthermore the p -adic Liouville theorem shows that if β is algebraic over Q , then there exists a constant k' such that $\text{ord}(\beta + s) > k' + k \log s$ for all $s \in \mathbb{Z}$, where $k \log p$ is the degree of β over Q . (The p -adic Roth Theorem, [7], shows that k may be replaced by any real number greater than $2/\log p$.) Hence β is certainly non-Liouville.

To each element of \mathfrak{L} we assign weights.

DEFINITION 2. For $\alpha \in \mathfrak{L}$ we define the *weight*, $w(\alpha)$, of α to be

$$w(\alpha) = 0 \quad \text{if } \alpha \notin \mathfrak{D},$$

$$w(\alpha) = (p-1)^{-1} \quad \text{if } \alpha \in \mathfrak{D}',$$

$$w(\alpha) = (1 - p^{-[r]})/(p-1) + \langle r \rangle p^{-1-[r]} \quad \text{if } \alpha \in \mathfrak{D}, \alpha \notin \mathfrak{D}',$$

where in the last case, $r = r(\alpha) = \text{Max}_{t \in \mathbb{Z}} \text{ord}(\alpha + t)$, and $r = [r] + \langle r \rangle$ is the usual decomposition of r into integral and fractional parts.

We extend this last definition to polynomials.

DEFINITION 3. The *weight*, $w(g)$, of a polynomial, g , in one variable whose roots lie in \mathfrak{L} is the sum of the weights of the roots (counting multiplicities).

Let π be an element of \mathfrak{D} such that $\text{ord } \pi = (p-1)^{-1}$. For $\alpha \in \mathbb{Z}_-$, let

$$G_1(t) = \sum_{j=0}^{\infty} (-\pi t)^j / (j!(j+\alpha)), \quad G_2(t) = \sum_{j=0}^{\infty} (\pi t)^j / \theta(\alpha, j).$$

We need the formal identity

$$(9) \quad \alpha^{-1} G_2(t) = (\exp \pi t) G_1(t),$$

which may be checked most easily by noting that $\alpha^{-1} G_2(t/\pi)$ is a formal power series solution of the differential equation $tl_1(y) = 1$, where $l_1 = d/dt - 1 + \alpha/t$ and by further noting that $l_1 = g(t) \cdot d/dt \cdot g(t)^{-1}$ where $g(t) = t^{-\alpha} \exp t$ and using this second form to check that $(\exp t) G_1(t/\pi)$ is also a formal solution of the same differential equation.

THEOREM 2. If $\alpha \in \mathfrak{L} \cap \mathfrak{D}$, then

$$(10) \quad \text{ord } \theta(\alpha, s) = w(\alpha) + O(\log s).$$

PROOF. We consider three cases:

Case 1. $\alpha \notin \mathfrak{D}'$. There exists $m \in \mathbb{Z}_+$, $\beta \in \mathfrak{D}$ such that $\alpha = -m + p^r \beta$, where $r = r(\alpha)$ (cf. Definition 2) and $r(\beta) = 0$. Putting $\sigma = [r(\alpha)]$, we have

$$\text{ord } \prod_{j=1}^s (j + p^r \beta) = \sum_{i=0}^{\sigma} [s/p^i] + \langle r \rangle [s/p^{\sigma+1}] = O(1) + w(\alpha)s.$$

Since $\theta(\alpha, s) = \theta(\alpha, m) \prod_{j=1}^m (j + p^r \beta)$, the verification of (10) (with $O(\log s)$ replaced by $O(1)$) is immediate.

Case 2. $\alpha \in \mathbb{Z}$. In this case $\theta(\alpha, s)$ differs by a constant from $s!$ and hence (10) follows immediately from the Gauss formula which we write in the form,

$$(11) \quad s/(p-1) - \log(1+s)/\log p \leq \text{ord } s! \leq s/(p-1).$$

Case 3. $\alpha \in \mathfrak{D}'$, $\alpha \notin \mathbb{Z}$. It is well known (and easily checked by means of [1, Lemma 1]) that $(1-t)^{-\alpha-1} \in \mathfrak{D}'[t]$, while the coefficient of t^s in this series is $\theta(\alpha, s)/s!$. Thus $\text{ord } \theta(\alpha, s) \geq \text{ord } s!$ and hence using (11), we have

$$(12) \quad s/(p-1) - \log(1+s)/\log p \leq \text{ord } \theta(\alpha, s).$$

A more lengthy argument will be needed to establish an upper bound. Letting B_j be the coefficient of t^j in the series G_1 , we deduce from (8) (since $\alpha \in \mathfrak{X}$) and (11) that $j^{-1} \text{ord } B_j \rightarrow 0$ and hence the domain of convergence of G_1 is the "open" disk: $\text{ord } t > 0$. The same holds for $\exp \pi t$ so that (9) shows that G_2 certainly converges in the open disk. We may now conclude from (12) that this open disk is also the domain of convergence of G_2 . Thus the Newton polygons [2, §1] of G_1 and G_2 have no sides of strictly positive slope, the Newton polygon of $\exp \pi t$ is the x -axis and hence (9) shows that G_1 and G_2 have the same Newton polygons. Using (8), there exist real numbers k, k' such that if v is the function on the positive real line, $v(x) = -k \log(1+x) - k'$, then

$$(13) \quad j/(p-1) - \text{ord } j! - \text{ord}(j + \alpha) \geq v(j)$$

for all $j \in \mathbb{Z}_+$. The graph of v (since it is convex upward) must therefore bound the Newton polygon of G_1 from below and thus

$$(14) \quad j/(p-1) - \text{ord } \theta(\alpha, j) \geq v(j)$$

for all $j \in \mathbb{Z}_+$. Equation (10) follows from (14) and (12). This completes the proof.

The following statement is an immediate consequence of (10).

COROLLARY. *Let h be a monic polynomial whose roots lie in $\mathfrak{D} \cap \mathfrak{E}$ and let t be a fixed positive integer greater than all roots of h in Z . If $s, s' \in Z, s > s' > t$, then*

$$(15) \quad \sum_{j=s'}^s \text{ord } h(j) = w(h)(s - s') + O(\log s).$$

Since $\text{ord } h(s) \geq 0$ for all $s \in Z$, we conclude under the hypotheses of the Corollary that if s_1, s_2, \dots, s_e is any set of distinct integers in the interval $[s', s]$, then

$$(16) \quad - \sum_{j=1}^e \text{ord } h(s_j) \geq -w(h)(s - s') + O(\log s).$$

Note: The verification of (15) and (16) required both upper and lower bounds for $\text{ord } \theta(\alpha, s)$.

4. Formal solutions. We generalize slightly the problem stated in the introduction. We suppose the coefficients, A_j , of (1) lie in $\Omega((x))$. Multiplying the operator l by a suitable power of x , we may write for each $s \in Z$,

$$(17) \quad l(x^s) = \sum_{j=0}^{\infty} \Phi_j(s+j)x^{s+j},$$

where each Φ_j is a polynomial of degree not greater than n and Φ_0 is not identically zero. Note [5, Chapter IV, §12] that Φ_0 is the classical indicial polynomial of l (at $x=0$). We recall that the ordinal, $\text{ord } \Phi$, of a polynomial, Φ , (say in one variable) is the maximal ordinal of its coefficients and that

$$(18) \quad \text{ord } \Phi = \inf \text{ord } \Phi(x) \leq \inf_{s \in Z} \text{ord } \Phi(s),$$

the first \inf being over all $x \in \mathfrak{D}$.

Let $R(l)$ be the minimal real number such that A_0, A_1, \dots, A_n all converge in the punctured disk, $\infty > \text{ord } x > R(l)$. (In general $R(l) \neq R(A_n^{-1} \cdot l)$.) Writing $A_i = \sum_j a_{i,j} x^{i+j}$, we have $\liminf j^{-1} \text{ord } a_{i,j} \geq -R(l)$ for each i and since $\Phi_j(s+j) = \sum_{i=0}^n a_{i,j} s(s-1) \cdots (s-i+1)$, it follows that if we set $\liminf j^{-1} \text{ord } \Phi_j = -b_{\infty}(l)$, then $-b_{\infty}(l) \geq -R(l)$. Conversely, it is clear that for each $m \in Z, l(x^m)$ converges for $\infty > \text{ord } x > b_{\infty}(l)$ and hence A_0, A_1, \dots, A_n all converge on that set. Thus $R(l) \leq b_{\infty}(l)$ and hence $R(l) = b_{\infty}(l)$.

The convex closure (in Euclidean 2-space) of the points

$\{(j, \text{ord } \Phi_j) | j \in \mathbb{Z}_+\}$ will be referred to as the *Newton Polygon* of l (at $x=0$). From the above definitions, it is clear that $-b_\infty(l)$ is the asymptotic slope of this polygon. The slope, $-b_1(l)$, of the first side of the Newton polygon, is of importance in the following. The first side has minimal slope, and hence $b_1(l) \geq R(l)$. We note that

$$b_1(l) = \sup_{j \geq 1} j^{-1}(\text{ord } \Phi_0 - \text{ord } \Phi_j).$$

DEFINITION 4. The differential operator, l , is said to be *non-Liouville* (at $x=0$) if $b_1(l) < \infty$ (certainly true if the A_j all converge in a nontrivial disk) and each zero of its indicial polynomial lies in \mathfrak{R} .

With each non-Liouville, l , we associate the real number,

$$(19) \quad b(l) = b_1(l) + w(\Phi_0) \geq b_\infty(l) + w(\Phi_0).$$

THEOREM 3. Let l be non-Liouville, $b \geq b(l)$, and let $f = \sum C_j x^j$, $y = \sum B_j x^j$ be elements of $\Omega((x))$ such that

$$ly = f, \quad \text{ord } C_j + bj > O(\log j),$$

then $\text{ord } B_j + bj > O(\log j)$. In particular each power series solution of (1) converges for $\text{ord } x > b(l)$.

PROOF. We may suppose $\text{ord } \Phi_0 = 0$. We choose $t \in \mathbb{Z}$, t larger than each integer root of Φ_0 . Let I_t be the ideal $x^t \Omega[[x]]$ and let

$$T = \left\{ \sum_{s=t}^{\infty} B_s x^s \in I_t \mid \text{ord } B_s + bs > O(\log s) \right\}.$$

By (17) l is stable on both I_t and T (since $b \geq b_1(l)$) and if τ is the natural projection of $\Omega[[x]]$ onto I_t , then $\tau \cdot l$ maps polynomials into T . Thus we may reduce to the case in which $f \in T$, $y \in I_t$. The endomorphism, $l_0: x^s \rightarrow \Phi_0(s)x^s$, is invertible in I_t and by (8) is also invertible in T . Thus we may (without other alteration) change the hypothesis, $ly = f$, to read

$$(20) \quad ly = l_0 f.$$

If we write $l_1 = -(l - l_0)$, then $l_0^{-1}l = 1 - l_0^{-1}l_1$ and l_1 (and hence $l_0^{-1}l_1$) has the property

$$l_1 x^m \Omega[[x]] \subset x^{m+1} \Omega[[x]]$$

for all m . Thus $l_0^{-1}l$ has inverse,

$$(21) \quad (l_0^{-1}l)^{-1} = \sum_{r=0}^{\infty} (l_0^{-1}l_1)^r,$$

in I_t and the problem is that of checking stability on T . Using (17), (20) and (21), and writing $C(s)$ for C_s , $\Phi(j, s)$ for $\Phi_j(s)$ whenever $j > 0$, we compute

$$(22) \quad B_s = \sum_{r=0}^{\infty} \sum C(s_0) E(s_0, s_1, \dots, s_r),$$

the inner sum being over all r -tuples, $(s_0, s_1, \dots, s_{r-1})$ such that $t \leq s_0 < s_1 < s_2 < \dots < s_r = s$, where

$$(23) \quad (-1)^r E(s_0, s_1, \dots, s_r) = \prod_{j=1}^r \{ \Phi(s_j - s_{j-1}, s_j) / \Phi_0(s_j) \}.$$

Note that in (22), $r \leq s$ and hence the formula for B_s is a finite sum. We factor $\Phi_0 = gh$, where h is monic with roots in \mathfrak{D} , while g has no roots in \mathfrak{D} . Since $\text{ord } \Phi_0 = 0$, we have (writing b_1 for $b_1(l)$)

$$\text{ord } g(s) = 0, \quad \text{ord } \Phi_j(s) \geq \text{ord } \Phi_j \geq -jb_1$$

for all $s \in Z$, $j \geq 1$. We now easily compute with the aid of (16),

$$\begin{aligned} \text{ord } E(s_0, s_1, \dots, s_r) &\geq -b_1(s_r - s_0) - \sum_{j=1}^r \text{ord } h(s_j) \\ &\geq -(b_1 + w(h))(s_r - s_0) + O(\log s). \end{aligned}$$

Since $s_r = s > s_0$,

$$\text{ord}(C(s_0)E(s_0, s_1, \dots, s_r)) + bs \geq (b - b_1 - w(h))(s - s_0) + O(\log s).$$

The theorem now follows from (22) using $w(h) = w(\Phi_0)$, $b \geq b(l)$. Note:

1. The classical theory suggests the conjecture that if l is normalized ($A_n = 1$), then the theorem would remain valid with $b_1(l)$ replaced by $R(l)$. This is false, even at ordinary points (e.g. $dy/dx = y$).

2. The result may be improved slightly. Everything said after (18) would remain valid if for each $j \geq 1$, the symbol $\text{ord } \Phi_j$ were replaced by $\inf_{s \in Z} \text{ord } \Phi_j(s)$.

3. The estimates $O(\log s)$ may be made precise. If the roots of Φ_0 in \mathfrak{D}' are algebraic over Q , let ρ be the number of these roots, let $k > 2/\log p$ and let $\epsilon > 0$. If the term $O(\log s)$ in the hypothesis for f is replaced by $-\epsilon k \log s + O(1)$, then in the conclusion the $O(\log s)$ term in the estimate for y may be replaced by $-(2.5\rho + \epsilon)k \log s + O(1)$ and if $ly = 0$, then the estimate becomes $-1.5\rho k \log s + O(1)$.

REFERENCES

1. B. Dwork, *Norm residue symbol in local number fields*, Abh. Math. Sem. Univ. Hamburg 22 (1958), 180–190.

2. ———, *On the zeta function of a hypersurface*, Inst. Hautes Études Sci. Publ. Math., No. 12, Paris, 5–68.
3. ———, *On the zeta function of a hypersurface*. II, Ann. of Math. (2) 80 (1964), 227–299.
4. E. Goursat, *Differential equations*, Ginn, New York, 1945.
5. P. Hartman, *Ordinary differential equations*, Wiley, New York, 1964.
6. J. Igusa, *Analytic groups over complete fields*, Proc. Nat. Acad. Sci. U.S.A. 42 (1956), 540–541.
7. D. Rideout, *The p -adic generalization of the Thue-Siegel-Roth theorem*, Mathematika 5 (1958), 40–48.
8. J. Serre, *Endomorphismes complètement continus des espaces de Banach P -adiques*, Inst. Hautes Études Sci. Publ. Math., No. 12, Paris, 69–85.

JOHNS HOPKINS UNIVERSITY