Advanced Studies in Pure Mathematics 17, 1989
Algebraic Number Theory — in honor of K. Iwasawa
pp. 89–96

On the Uniqueness of Frobenius Operator on Differential Equations

B. Dwork

Dedicated to K. Iwasawa

Let k be an algebraically closed field of characteristic zero complete under an ultrametric norm with residue class field of characteristic p. Let $\mathscr{H}$, $\mathscr{G}$ be $d\times d$ matrices with coefficients in $k(x)$, the field of rational functions in one variable. We consider differential equations

$$\frac{dy}{dx}=y\mathscr{H} \tag{1}$$

$$\frac{dy}{dx}=y\mathscr{G} \tag{2}$$

with a "Frobenius operation" mapping solutions of (1) into solutions of (2), i.e., we assume the existence of a $d\times d$ matrix A with non-trivial determinant and coefficients analytic (in the sense of Krasner) on the complement of a finite number of disks each lying properly in a residue class such that for each solution matrix, Y, of (1) analytic in a region meeting the support of A, the product

$$Y(x^p)A=Y_2(x) \tag{3}$$

is a solution matrix of (2). This is equivalent to the differential relation

$$\frac{dA}{dX}=A\mathscr{G}-px^{p-1}\mathscr{H}(x^p)A. \tag{4}$$

Our object is to state conditions which imply the uniqueness of A up to a constant (scalar) factor.

Lemma. *Suppose*

(α) *Equation* (1) *has a solution matrix analytic on the generic open*

Received December 15, 1987.

unit disk, $D(t, 1^-)$.

(β_1) *The singularities of* (2) *are all regular.*

(β_2) *The exponents of* (2) *lie in* $\boldsymbol{Q} \cap \boldsymbol{Z}_p$.

(β_3) *Each residue class contains at most one singularity of* (2).

(γ) *Equation* (2) *is irreducible over* $k(x)$.

Then we conclude that A is unique up to a constant factor.

Remark. Condition (α) is trivial if equation (2) has a "strong" Frobenius structure, e.g. if equations (1) and (2) are equivalent or more generally if equation (4) with $\mathscr{H} = \mathscr{G}$ and p replaced by an integral power has a solution satisfying the indicated conditions of analyticity.

Proof. Condition (α) implies that equation (2) also has a solution matrix analytic on the generic unit disk and hence by our transfer theorem [DR] for each residue class containing no singularity of (2) there is a solution matrix analytic on the residue class. Let A_1, A_2 be two matrices mapping solutions of (1) into solutions of (2). If Y_1 is a solution matrix of (1) converging on a residue class $D(a^p, 1^-)$ and if Y is an arbitrary solution matrix of (2) at a then there exist $C_1, C_2 \in Gl(d, k)$ such that

$$C_1 Y = Y_1(x^p) A_1$$

$$C_2 Y = Y_1(x^p) A_2$$

Letting $C = C_2^{-1} C_1$, $B = A_2^{-1} A_1$ we have

$$B = Y^{-1} C Y. \tag{5}$$

Here B is invertible over the ring of functions analytic on the complement of a finite set of small disks.

We assert that hypotheses (β_1)—(β_3) imply $B \in GL(d, k(x))$. Indeed (5) shows that B is analytic on each residue class containing no singularity of (2). It is well known that at each singularity a there is a solution matrix of the form

$$Y = (x-a)^H U \tag{6}$$

where H is a constant matrix with eigenvalues equal to the exponents of (2) at a and U is analytic at $x = a$.

By Christol's Transfer Theorem [Ch] we know that U is analytic on $D(a, 1^-)$ and hence has a common domain of definition with B. We deduce from (5) that

$$U B U^{-1} = (x-a)^{-H} C (x-a)^H \tag{7}$$

for some $C \in Gl(d, k)$. The left side of (7) is analytic on an annulus

$$1 > |x-a| > 1-\varepsilon$$

and hence by Lemma 5 [Dw 0] the right side must be free of logarithms and indeed must be a matrix whose coefficients are of the form $c(x-a)^s$ where $c \in k$, $s \in \boldsymbol{Z}$. Thus B has at worst a pole at $x=a$ and no other singularity in the singular disk. We conclude that B is indeed rational.

Equation (5) means that $y \mapsto yB$ is an endomorphism of solutions of (2). Let a be an eigenvalue and u a corresponding eigenvector. Then

$$u(B-aI)=0 \tag{8}$$

If $B \neq aI$ this gives a non-trivial homogeneous relation over $k(x)$ among the components of u. This implies reducibility contrary to hypothesis (γ).

Remark. A. Without hypothesis (γ) we may use the differential equation

$$\frac{dB}{dx} = B\mathcal{G} - \mathcal{G}B \tag{9}$$

to conclude that B is completely specified by its specialization at one non-singular point.

Furthermore the mapping $y \mapsto yB$ respects the decomposition of the corresponding differential module. A knowledge of this decomposition can be used to reduce the number of parameters needed to specify B.

Thus for $d=2$ there are four cases of reducibility. If M_1 is a one dimensional submodule of the two dimensional differential module M then the number of parameters for B is

4 if M_1 is a direct summand *and* $M/M_1 \simeq M$,
1 if neither of the above conditions hold,
2 if precisely one of the above conditions hold.

B. In the case of irregular singularities, local solutions no longer have the form (6) but we do have the Hukuhara-Turrittin form [L]

$$z^H \exp \Delta(1/z) U(z) \tag{10}$$

where Δ is a diagonal polynomial matrix, H is again a constant matrix and U is a formal power series matrix and z is a suitable root of the local parameter $x-a$.

Christol's Transfer Theorem is not available in this case. We know only that U has a non-trivial radius of convergence subject to some hypothesis on the eigenvalues of H (always true if the differential equation is defined over an algebraic number field). However we shall give below an example in which the method of proof may be applied to such a differential equation.

Example.

1. Some time ago [Dw 1] we examined the middle dimensional cohomology of the family of K-3 surfaces

$$x_1^4+x_2^4+x_3^4+x_4^4-4\lambda x_1x_2x_3x_4=0$$

and found that this cohomology split into a number of subspaces of dimensions 1 and 2 and one interesting subspace, V, of dimension 3 spanned by the derivative (with respect to λ) of the holomorphic 2-form. The corresponding differential equation was identified with a ${}_3F_2$ generalized hypergeometric function which in fact by a formula of Clausen,

$${}_2F_1\left(a, b, a+b+\frac{1}{2}; z\right)^2 = {}_3F_2\left(\begin{matrix} 2a, a+b, 2b; z \\ a+b+\frac{1}{2}, 2a+2b\end{matrix}\right),$$

may be identified with the symmetric square of the second order differential equation satisfied by ${}_2F_1(1/8, 3/8, 1, t)$, $t=\lambda^4$. (The symmetric square is no accident, by Fano [F] such a relation must exist for third order differential equation whose solutions satisfy a homogeneous quadratic relation and this is just the consequence of duality in the present example).

By the lemma the Frobenius matrix for V may be computed (up to a constant factor) from the symmetric square of the Frobenius matrix of the ${}_2F_1$, which in turn has a known relation with a subspace of the cohomology of a curve. The constant may be computed by a specific calculation at $\lambda=0$, the matrix for the surface at $\lambda=0$ being computable in terms of gauss sums while for ${}_2F_1$ we may for example use the asymptotic calculations of [Dw 2, Chapter 26]. In this way the Riemann hypothesis for this family of K-3 surfaces could be deduced from the corresponding hypothesis for curves.

2. We have studied [Dw 3] the Frobenius matrix $\gamma(a, a', \lambda)$ mapping (for $a=(a_1, a_2, a_3)$, $a'=(a'_1, a'_2, a'_3)$, $pa'-a=\mu\in \boldsymbol{Z}^3$) branches of ${}_2F_1(a', \lambda)$ into branches of ${}_2F_1(a, \lambda)$. Kummer's formulae for 24 solutions of each of these equations gives rise to transformation properties of the Frobenius matrix. We have determined [Dw 4] these transformations by using the fact that the Frobenius matrix gives the action of Frobenius in mapping the obvious two dimensional cohomology space containing ω_{a',λ^p} into the corresponding space for $\omega_{a,\lambda}$ where

$$\omega_a = x^{a_2}(1-x)^{a_3-a_2-1}(1-\lambda x)^{-a_1}dx/x$$

and then using the well known fact that the 24 solutions of Kummer correspond precisely to the transformation of ω_a under the 24 fractional linear mappings which map $\{0, 1, \infty\}$ into a subset of $\{0, 1, \infty, 1/\lambda\}$. The present article gives a different approach to this question.

In the same way the present article may be used to deduce the transformation laws of $\gamma(a, a', \lambda)$ corresponding to the quadratic transformations of ${}_2F_1$. Recent work of I. Pastro has given a cohomological explanation of these quadratic transformations. This would give a cohomological method for solving the same problem.

3. We now give an example which uses uniqueness to deduce duality for the L-function associated with hyperkloosterman sums and at the same time illustrate our method in the case of an irregular singularity. We shall use the notation of Sperber [Sp 1, 2] and slightly simplify and extend one of his results.

Let $\pi^{p-1} = -p$, $\delta = x(d/dx)$ and we can consider

$$G_\pi = \begin{pmatrix} 0 & 0 \cdots 0 & \pi^{n+1}x \\ 1 & 0 \cdots 0 & 0 \\ 0 & 1 \cdots 0 & 0 \\ \cdots & \cdots & \cdots \\ 0 & 0 \cdots 1 & 0 \end{pmatrix}$$

so that the equation

$$\delta y = y G_\pi \tag{11}$$

is the system associated in the usual way with the scaler equation

$$(\delta^{n+1} - \pi^{n+1}x)u = 0. \tag{12}$$

Let A_π be the Frobenius matrix associated with (11) so by [Sp 1, Theorem 1.3.9], A_π is analytic for

$$\text{ord } x > -(n+1)(p-1)/p^2. \tag{13}$$

Let

$$\Theta=\begin{pmatrix} 0 & 0\cdots\cdot 0 & 1 \\ 0 & 0\cdot\cdot -1 & 0 \\ \cdots & \cdots & \cdots \\ (-1)^n & 0\cdots\cdot 0 & 0 \end{pmatrix}$$

Then

(14) $$G_\pi\Theta+\Theta G^t_{-\pi}=0$$

and hence if Y_π (resp: $Y_{-\pi}$) is solution matrix of (11) (resp: the transformation of (11) under $\pi\mapsto-\pi$) at some point a say $|a|=1$, then $d/dx(Y_\pi\Theta Y^t_{-\nu})=0$ and hence

(15) $$Y_\pi\Theta Y^t_{-\nu}=\theta,$$

a constant invertible matrix which shows if $A_{-\pi}$ is a Frobenius matrix for the differential equation involving $G_{-\pi}$ then $\Theta A^*_{-\pi}\Theta^{-1}$ is a Frobenius matrix of (11) and hence if the lemma were applicable we could conclude that

(16) $$A_\pi=\gamma\Theta A^*_{-\pi}\Theta^{-1}$$

for some scaler γ, where $A^*_{-\pi}=(A^t_{-\pi})^{-1}$. We write this in the form

(17) $$A_\pi\Theta A^t_{-\pi}=\gamma\Theta.$$

We specialize the equation at $x=0$ and compute the coefficient of the first row, $(n+1)^{st}$ column on both sides of (17). By [Sp 1, Theorem 4.2.11] $A_\pi(0)$ is lower triangular with $(1, p, p^2, \cdots, p^{n+1})$ as diagonal elements. We readily compute $\gamma=p^n$.

To complete the proof we must show the validity of the uniqueness lemma in this case. Let then

(18) $$B=A_\pi^{-1}\Theta A^*_{-\pi}\Theta^{-1}.$$

Since (11) has a regular singularity at $x=0$ and no other singularity except the irregular one at infinity we know by the proof of the lemma that B is analytic everywhere (except for a pole at $x=0$) in the region (13). We show for p prime to $(n+1)$ that B extends to the sphere punctured at infinity.

To do this we put $x=z^{n+1}$ so that (12) becomes

(19) $$(\delta_z^{a+1}-(az)^{n+1})u=0$$

with $a=(n+1)\pi$.

Substituting $u=w\exp az$ in (19) we determine a differential equation for w which has indicial polynomial at infinity

$$(n+1)(s+n/2)$$

and hence there exists a unique formal solution for w of the form $z^{n/2}\xi$ with $\xi \in k[[1/z]]$. Putting

$$\xi=\sum_{s=0}^{\infty} \varepsilon_s/(za)^s$$

the recursion relation is of the form

$$(n+1)\varepsilon_s\left(s-\frac{n}{2}\right) \in \sum_{j=1}^{n} \varepsilon_{s-j}Z\left[\frac{1}{2}\right]$$

and so

(20) $$\operatorname{ord} \varepsilon_s/a^s \geq -2s \operatorname{ord} \pi(n+1)+O(\log s).$$

Thus $z^{n/2}(\exp az)\xi$ is a formal solution of (19) and the remaining solutions are obtained from this one by the substitutions $z \mapsto z\zeta^i$, $0\leq i\leq n$ where ζ is a primitive $(n+1)^{st}$ root of unity. This shows that (19) is irreducible over $k(z^{n+1})$ and hence (11) is irreducible over $k(x)$.

Following the method of [Sp, 1, Proposition 5.1.7] and using (20) the Frobenius structure may be used to show that if $vz^{n/2}\exp az$ is a formal solution of (11) with $v \in (k[[1/z]])^{n+1}$ then v is analytic in the disk $\operatorname{ord} z<0$. (This replaces Christol's theorem in this case). Putting $\tilde{v}=vB$ analytic on the annulus

$$0>\operatorname{ord} z>-\frac{p-1}{p^2}$$

we know that $z^{n/2}\tilde{v}\exp az$ is a formal solution of (11) and hence letting z_0 be a point in the annulus we must have for some constants $c_0, \cdots, c_n$,

$$\tilde{v}\exp a(z-z_0)=\sum_{j=1}^{n} c_j v(z\,\zeta^j)\exp a\zeta^j(z-z_0),$$

a relation of linear dependence of $\{\exp a\zeta^j(z-z_0)\}_{0\leq j\leq n}$ over the field of functions meromorphic on the disk. The quotients of these exponentials do not lie in that differential field and hence by the proof of Lemma 5 [Dw 0], the relation must be trivial. Thus $\tilde{v}=cv$ for some constant. We know that equation (11) has a solution matrix

$$z^{n/2}\exp(az\Delta)V(z)$$

where Δ is the diagonal matrix with coefficients $(1, \zeta, \cdots, \zeta^n)$ and V has coefficients in $k[[1/z]]$ and in fact converges for ord $z<0$. We have shown that for some diagonal constant matrix C, $VB=CV$ and a wronskian calculation shows that $z^{n(n+1)/2}$ det V is constant. Thus B is meromorphic as function of z in the infinite residue class with pole only at $z=\infty$. But B is analytic as function of $x=z^{n+1}$ on an annulus of that disk. We conclude that B is analytic (except for pole at infinity) as a function of x on the infinite residue class. Hence $B \in Gl(n+1, k(x))$. This together with our previous discussion of irreducibility completes our treatment. Aside from a simplification of method we have improved upon Sperber's results only by replacing his condition $p>n+1$ by the condition that p be prime to $(n+1)$.

References

[Ch] Christol, G., Un théorème de transfert pour les disques singulieres reguliers, Astérisque, **119–120** (1984), 151–168.

[Dw 0] Dwork, B., On p-adic Analysis, Science Conference Proc., **2** (1965–66) Belfor Grad. School of Science, Yeshiva Univ. NY (1969), 129–154.

[Dw 1] ——, p-adic cycles, Publ. Math. I.H.E.S., **37** (1969), 27–115.

[Dw 2] ——, Lectures on p-adic differential equations, Springer, NY 1982.

[Dw 3] ——, On the Boyarsky principle, Amer. J. Math., **105** (1983), 115–156.

[Dw 4] ——, On Kummer's 24 solutions of the hypergeometric differential equation, Trans. Amer. Math. Soc., **285** (1984), 497–521.

[D-R] —— with Robba, Ph., Effective p-adic bounds for the solutions of linear homogeneous differential equations, Trans. Amer. Math. Soc., **259** (1980), 559–577.

[F] Fano, G., Ueber lineare homogene Differentialgleichungen mit algebraischen Relationen zwischen den Fundamentallösungen, Math. Ann., **53** (1900), 493–590.

[L] Levelt, A., Jordan decomposition for a class of singular differential operators, Ark. Mat., **13** (1975), 1–27.

[Sp 1] Sperber, S., p-adic hypergeometricfunctions and their cohomology, Duke Math. J., **44** (1977), 535–589.

[Sp 2] ——, Congruence properties of the hyperkloosterman sum, Compositio Math., **40** (1980), 3–33.

Department of Mathematics
Princeton University
Princeton, NJ 08544
U.S.A.