J. Math. Kyoto Univ. (JMKYAZ)
12-2 (1972) 369-383

On the nilpotence of the hypergeometric equation

By

William MESSING

(Communicated by Professor Nagata, December 23, 1971)

Introduction

Let T be an arbitrary scheme, S a smooth T-scheme and $\mathcal{M}$ a quasi-coherent $\mathcal{O}_S$-module. A T-connection on $\mathcal{M}$ is by definition a homomorphism of $\mathcal{O}_S$-modules:

$$\nabla: \mathcal{D}er_{\mathcal{O}_T}(\mathcal{O}_S, \mathcal{O}_S) \longrightarrow \mathcal{E}nd_{\mathcal{O}_T}(\mathcal{M})$$

which satisfies the "product formula":

$$\nabla(D)(sm)=s\nabla(D)(m)+D(s)m$$

for sections D of $\mathcal{D}er_{\mathcal{O}_T}(\mathcal{O}_S, \mathcal{O}_S)$, s of $\mathcal{O}_S$ and m of $\mathcal{M}$ over an open subset $U\subseteq S$. A section m of $\mathcal{M}$ over U is called horizontal if $\nabla(D)(m)=0$ for all D's, derivations on open subsets of U. Both $\mathcal{D}er_{\mathcal{O}_T}(\mathcal{O}_S, \mathcal{O}_S)$ and $\mathcal{E}nd_{\mathcal{O}_T}(\mathcal{M})$ are $\mathcal{O}_T$-Lie-algebras via the commutator bracket. The connection is called integrable if it is a Lie-algebra homomorphism. The obstruction to a connection being integrable is the curvature homomorphism $K: \bigwedge^2 \mathcal{D}er_{\mathcal{O}_T}(\mathcal{O}_S, \mathcal{O}_S) \rightarrow \mathcal{E}nd_{\mathcal{O}_S}(\mathcal{M})$ defined by $K(D\wedge D') = [\nabla(D), \nabla(D)]-\nabla([D, D'])$. Henceforth we will deal only with integrable connections.

A horizontal morphism ϕ between modules with connection

$\phi:(\mathcal{M}, \nabla)\to(\mathcal{M}', \nabla')$ is by definition an $\mathcal{O}_S$-linear mapping satisfying $\phi\circ\nabla(D)=\nabla'(D)\circ\phi$. Taking as objects quasi-coherent $\mathcal{O}_S$-modules with T-connections $(\mathcal{M}, \nabla)$ and as morphisms the horizontal morphisms we obtain an abelian category. This category has a partially defined internal Hom obtained by defining Hom $((\mathcal{M}, \nabla), (\mathcal{M}', \nabla'))$ as being $(\mathcal{H}om_{\mathcal{O}_S}(\mathcal{M}, \mathcal{M}'), \bar{\nabla})$ where $\bar{\nabla}(D)(\phi)=\nabla'(D)\circ\phi-\phi\circ\nabla(D)$. In particular $\check{\mathcal{M}}=\mathcal{H}om_{\mathcal{O}_S}(\mathcal{M}, \mathcal{O}_S)$ is the underlying module of Hom$((\mathcal{M}, \nabla), (\mathcal{O}_S$, standard$))$ and hence has a "dual" connection $\check{\nabla}$ which satisfies the "product formula"

$$<\check{\nabla}(D)(\phi), m> + <\phi, \nabla(D)(m)> = D<\phi, m>$$

where ϕ is a local section of $\check{\mathcal{M}}$, m of $\mathcal{M}$ and D of $\mathcal{D}er_{\mathcal{O}_T}(\mathcal{O}_S, \mathcal{O}_S)$. The category also has an internal tensor product $(\mathcal{M}, \nabla)\otimes(\mathcal{M}', \nabla')$ which by definition is $(\mathcal{M}\underset{\mathcal{O}_S}{\otimes}\mathcal{M}', \bar{\nabla})$ where $\bar{\nabla}$ is defined by $\bar{\nabla}(D)(m\otimes m')=\nabla(D)(m)\otimes m'+m\otimes\nabla(D)(m')$. As a result, we can define "induced" connections on the exterior powers of a module with connection and hence can speak of the determinant det $((\mathcal{M}, \nabla))$ provided $\mathcal{M}$ is locally free of constant (finite) rank.

If T is a scheme of characteristic p then both $\mathcal{D}er_{\mathcal{O}_T}(\mathcal{O}_S, \mathcal{O}_S)$ and $\mathcal{E}nd_{\mathcal{O}_T}(\mathcal{M})$ are p-$\mathcal{O}_T$-Lie-algebras (by $D\mapsto D^p$, $\phi\mapsto\phi^p$). We can then ask if ∇ is a homomorphism of p-Lie-algebras, i.e., if $\nabla(D^p)=(\nabla(D))^p$. The "*p-curvature*" (introduced by Deligne) is the mapping Ψ: $\mathcal{D}er_{\mathcal{O}_T}(\mathcal{O}_S, \mathcal{O}_S)\to\mathcal{E}nd_{\mathcal{O}_T}(\mathcal{M})$ defined by $\Psi(D)=(\nabla(D))^p-\nabla^p(D^p)$. It is known, [3], that the p-curvature Ψ has the following properties:

1) Ψ is additive
2) Ψ is p-linear i.e. $\Psi(sD)=s^p\Psi(D)$
3) for each D, a section of $\mathcal{D}er_{\mathcal{O}_T}(\mathcal{O}_S, \mathcal{O}_S)$ over U, $\Psi(D)$ is a horizontal endomorphism of $(\mathcal{M}, \nabla)|U$ (in particular $\Psi(D)$ is $\mathcal{O}_U$-linear).

If for every section D of $\mathcal{D}er_{\mathcal{O}_T}(\mathcal{O}_S, \mathcal{O}_S)$ (over an open set U), $\Psi(D)$ is a nilpotent endomorphism, then we say the connection is nilpotent (a notion introducted by Berthelot [2], in the context of crystalline cohomology).

We observe that there is defined a notion of "inverse image" for modules with connection. Namely, if T, S, $(\mathcal{M}, \nabla)$ are given as above and if we are given a base change $T' \to T$, then there is associated with ∇ a T'-connection, ∇', on the $S' = S \times_T T'$ module $\mathcal{M}' = \mathcal{M} \otimes_{\mathcal{O}_S} \mathcal{O}_{S'}$. Locally we can give an explicit description of ∇':
If we choose affine open sets Spec(A), Spec(A′), Spec(B) of T(resp. T', resp. S) so as to obtain a commutative diagram

$$\begin{array}{ccc} \mathrm{B} & \to & \mathrm{B}' = \mathrm{B} \otimes_{\mathrm{A}} \mathrm{A}' \\ \uparrow & & \uparrow \\ \mathrm{A} & \to & \mathrm{A}' \end{array}$$

and if M is a B-module with connection ∇: $\mathrm{Der}_{\mathrm{A}}(\mathrm{B}, \mathrm{B}) \to \mathrm{End}_{\mathrm{A}}(\mathrm{M})$ then the connection ∇' on the module $\mathrm{M}' = \mathrm{M} \otimes_{\mathrm{A}} \mathrm{A}'$ is defined as the canonical mapping $\nabla \otimes 1$: $\mathrm{Der}_{\mathrm{A}'}(\mathrm{B}', \mathrm{B}') = \mathrm{Der}_{\mathrm{A}}(\mathrm{B}, \mathrm{B}) \otimes_{\mathrm{A}} \mathrm{A}' \to \mathrm{End}_{\mathrm{A}}(\mathrm{M}) \otimes_{\mathrm{A}} \mathrm{A}' \to \mathrm{End}_{\mathrm{A}'}(\mathrm{M}')$.

Now let $T = \mathrm{Spec}$ (A), where A is a ring of finite type over $\boldsymbol{Z}$ and $S = \mathrm{Spec}$ (B) when B is a smooth A-algebra. If M is an S-module with connection, we say M is *globally nilpotent* if for each closed point $\mathfrak{p}$ of T the induced connection on the module $\mathrm{M} \otimes k(\mathfrak{p})$ is nilpotent.

Let us recall that if X is a smooth S-scheme $\pi: X \to S$, then the De-Rham cohomology $\mathcal{H}_{D.R.}(X/S) \overset{\text{def.}}{=\!=} \boldsymbol{R}\pi_* (\Omega^{\cdot}_{X/S})$ has a "canonical" integrable connection: the Gauss-Manin connection [3, 4]. If T is of characteristic p, Katz and Berthelot [2, 3] proved that the Gauss-Manin connection is nilpotent. Using this result Katz [3], gave a beautiful arithmetic proof of the local monodromy theorem.

Let $a, b, c \in \boldsymbol{Q}$, n be a common denominator, $T = \mathrm{Spec}\left(\boldsymbol{Z}\left[\frac{1}{n}\right]\right)$, $S = \mathrm{Spec}\ \boldsymbol{Z}\left[\lambda, \frac{1}{n\lambda(1-\lambda)}\right]$ where λ is an indeterminate. Associated to the hypergeometric differential equation

$$\lambda(1-\lambda)\frac{d^2u}{d\lambda^2} + [c-(a+b+1)\lambda]\frac{du}{d\lambda} - abu = 0$$

is an S-module, $M_{a,b,c}$, with integrable T-connection: It is the free rank 2 module with base $\{e_1, e_2\}$ where

$$\begin{pmatrix} \nabla\left(\frac{d}{d\lambda}\right)(e_1) \\ \nabla\left(\frac{d}{d\lambda}\right)(e_2) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ \frac{ab}{\lambda(1-\lambda)} & \frac{(a+b+1)\lambda-c}{\lambda(1-\lambda)} \end{pmatrix} \begin{pmatrix} e_1 \\ e_2 \end{pmatrix}$$

We refer to $M_{a,b,c}$ as the hypergeometric module.

Katz has conjectured that the hypergeometric module, $M_{a,b,c}$, is globally nilpotent. In the first section we prove that for a "large class" of $\{a,b,c\}$ $M_{a,b,c}$ occurs as a direct factor (as module with connection) in the De Rham cohomology of a suitable family of curves. As a corollary, each of these hypergeometric modules reduces (for almost all primes p) modulo p to a nilpotent module. In the second section we prove the conjecture. The proof is based on the observation that in characteristic p, any hypergeometric equation has a nontrivial polynomial solution.

I wish to thank N. Katz for his help and encouragement during both the research and preparation of the manuscript. Also, several discussions with Professor B. Dwork proved invaluable.

Relation to De Rham Cohomology

Let n be a positive integer, ζ_n a primitive n^{th} root of 1 and λ an indeterminate. Assume a,b,c are positive integers such that $(n, a) =(n, b)=(n, c)=(n, a+b+c)=1$ and $n>a+b+c$. Let X be the curve defined over $\boldsymbol{Q}(\zeta_n, \lambda)$ which is the normalization of the projective closure of the affine curve $y^n=x^a(x-1)^b(x-\lambda)^c$. The group $\boldsymbol{\mu}_n$ of n^{th} roots of 1 operates on X. Explicitly $\boldsymbol{\mu}_n$ operates on the function field $\boldsymbol{Q}(\zeta_n, \lambda)(x, y)$ via $\sigma\cdot(x, y)=(x, \sigma y)$ where $\sigma\in\boldsymbol{\mu}_n$ because $(\sigma y)^n=\sigma^n y^n= y^n=x^a(x-1)^b(x-\lambda)^c$. Thus $\boldsymbol{\mu}_n$ operates by functoriality on $H_{D.R.}(X)$, the De Rham cohomology of X. Since we are in characteristic zero we may calculate $H^1_{D.R.}(X)$ as the factor space of differentials of the second kind modulo exact differentials. If we extend the action of $\boldsymbol{\mu}_n$ to $\Omega^{\cdot\, rat}_X$ by defining $\sigma\cdot(udx)=(\sigma\cdot u)dx$, then this mapping preserves both differentials of the second kind and exact differentials, and hence by passage to the quotient gives the action of $\boldsymbol{\mu}_n$ on $\mathrm{H}^1_{D.R.}(X)$.

Let us explicitly construct the Gauss-Manin connection on $\mathrm{H}^1_{D.R.}(X)$. Let D denote the unique derivation of the function field of X which extends the action of $\frac{d}{d\lambda}$ on $\boldsymbol{Q}(\zeta_n, \lambda)$ and kills x. Extend D to a derivation of $\Omega^{\cdot\, rat}_X$ by defining $D(fdg)=D(f)\cdot dg+fd(Dg)$. Under this derivation the differentials of the second kind and exact differentials are stable. The induced action of D on $\mathrm{H}^1_{D.R.}(X)$ $=$d.s.k./exact is $\nabla\left(\frac{d}{d\lambda}\right)$.

We observe that for $\sigma\in\boldsymbol{\mu}_n$ $D\circ\sigma-\sigma\circ D$ is a derivation of the function field of X. Since it kills λ and x, it is zero. This means that $\boldsymbol{\mu}_n$ actually operates via horizontal automorphisms on $\mathrm{H}^1_{D.R.}(X)$. Let us denote by χ the inverse of the principal character of $\boldsymbol{\mu}_n$ and by $\mathrm{H}^1_{D.R.}(X)^{\chi}$ the sub-module consisting of elements which transform according to χ.

Proposition *The module* $\mathrm{M}_{\frac{c}{n},\, \frac{a+b+c}{n}-1,\, \frac{a+c}{n}}$ *is isomorphic (as module with connection) to* $\mathrm{H}^1_{D.R.}(X)^{\chi}$, *and hence is a direct factor of* $\mathrm{H}^1_{D.R.}(X)$.

Proof: Consider X as lying over $\boldsymbol{P}^1$ via the morphism induced by the inclusion of function fields $\boldsymbol{Q}(\zeta_n, \lambda)(x)\to\boldsymbol{Q}(\zeta_n, \lambda)(x, y)$. The assumptions made on the four integers n, a, b, c imply that lying over each of the four points $0, 1, \lambda, \infty$ of $\boldsymbol{P}^1$ there is exactly one point of X(denoted respectively $\mathfrak{p}_0, \mathfrak{p}_1, \mathfrak{p}_\lambda, \mathfrak{p}_\infty$). We have $\mathrm{ord}_{\mathfrak{p}_0}(x)=n$, $\mathrm{ord}_{\mathfrak{p}_0}(y)=a$; $\mathrm{ord}_{\mathfrak{p}_1}(x)=n$, $\mathrm{ord}_{\mathfrak{p}_1}(y)=b$; $\mathrm{ord}_{\mathfrak{p}_\lambda}(x)=n$, $\mathrm{ord}_{\mathfrak{p}_\lambda}(y)=c$; $\mathrm{ord}_{\mathfrak{p}_\infty}(x)=-n$, $\mathrm{ord}_{\mathfrak{p}_\infty}(y)=-(a+b+c)$. This implies that both $\frac{dx}{y}$ and $\frac{xdx}{y}$ have poles only at $\mathfrak{p}_\infty$ and hence are differentials of the second kind (because the sum of the residues of any differential is zero). Let ω_1

and ω_2 denote the classes of $\frac{dx}{y}$ and $\frac{xdx}{y}$ in $\mathrm{H}^1_{D.R.}(X)$.

The proof breaks up into three parts;

1) We show ω_1 and ω_2 span $\mathrm{H}^1_{D.R.}(X)^\chi$

2) We define a surjective horizontal homomorphism $\mathrm{M}_{\frac{c}{n}, \frac{a+b+c}{n}-1, \frac{a+c}{n}} \to \mathrm{H}^1_{D.R.}(X)^\chi$

3) We prove this horizontal morphism is injective.

1) Represent $\mathrm{H}^1_{D.R.}(X)$ as a factor space of the space of differentials having poles only at $\mathfrak{p}_\infty$ and of some bounded order $\leq N$ (by Riemann-Roch Theorem this is possible). Then $\boldsymbol{\mu}_n$ operates on this space in a manner compatible with its action on $\mathrm{H}^1_{D.R.}(X)$. Both this space of differentials and $\mathrm{H}^1_{D.R.}(X)$ decompose into direct sums where the summands are the spaces of differentials (resp. cohomology classes) which transform according to a given character of $\boldsymbol{\mu}_n$. Thus any cohomology class which transforms according to χ is represented by a differential, regular except at $\mathfrak{p}_\infty$, which transforms according to χ.

Since $\mathrm{Spec}\,\boldsymbol{Q}(\zeta_n, \lambda)\left[x, y, \frac{1}{y}\right]$ (where $y^n = x^a(x-1)^b(x-\lambda)^c$) is non-singular any differential regular except at $\mathfrak{p}_\infty$ can be written $\frac{R(x,y)}{y^{\text{some power}}}dx$, where $R(x, y) \in (\zeta_n, \lambda)\,[x, y]$. By the division algorithm we can write it as $\left(R_0(x) + \frac{R_1(x)}{y} + \ldots + \frac{R_{n-1}(x)}{y^{n-1}}\right)dx$ where the $R_i \in \boldsymbol{Q}(\zeta_n, \lambda)(x)$. It can transform according to χ if and only if it is $\frac{R_1(x)}{y}dx$. Because this differential is regular except at $\mathfrak{p}_\infty$, $R_1(x)$ must be a polynomial. To conclude the first part, it remains to prove the following lemma.

Lemma: *The differentials* $x^l\frac{dx}{y}$ $(l \geq 2)$ *are linearly dependent on* $\frac{dx}{y}$ *and* $\frac{xdx}{y}$ *modulo exact differentials.*

Proof: (By induction on l). We have

$$d\left(\frac{x^{l-1}(x-1)(x-\lambda)}{y}\right) = (l+1)x^l\frac{dx}{y} - l(1+\lambda)x^{l-1}\frac{dx}{y} + (l-1)\lambda x^{l-2}\frac{dx}{y}$$

$$+x^{l-1}(x-1)(x-\lambda)\left(\frac{-c}{n(x-\lambda)}+\frac{-b}{n(x-1)}+\frac{-a}{nx}\right)\frac{dx}{y}$$

$$=\left(l+1-\frac{a+b+c}{n}\right)x^{l}\frac{dx}{y}+P(x)\frac{dx}{y}$$

where $P(x)$ is a polynomial of degree $\leq l-1$. As $l+1-\frac{a+b+c}{n}\neq 0$ we are done.

2) The existence and the surjectivity of a horizontal morphism $M_{\frac{c}{n},\frac{a+b+c}{n}-1,\frac{a+c}{n}}\rightarrow H^{1}_{D.R.}(X)^{\chi}$ will follow immediately from the following three lemmas. Explicitly the mapping will be defined by $e_1\rightarrow\omega_1$, $e_2\rightarrow\omega_1'$ where "$'$" stands for the action of $\nabla\left(\frac{d}{d\lambda}\right)$.

Let us write "$\equiv$" to denote congruence modulo exact.

Lemma: $$D\left(\frac{xdx}{y}-\frac{dx}{y}\right)\equiv\left[\left(1-\frac{a+b}{n}\right)+\frac{1}{\lambda}\left(\frac{c+\lambda b+a(1+\lambda)-n(1+\lambda)}{n}\right)\right]\frac{dx}{y}+\frac{1}{\lambda}\left(\frac{2n-(a+b+c)}{n}\right)\frac{xdx}{y}$$

Proof: We compute:

$$D(y^{n})=-cx^{a}(x-1)^{b}(x-\lambda)^{c-1}$$

$$ny^{n-1}D(y)=-cx^{a}(x-1)^{b}(x-\lambda)^{c-1}$$

$$D(y)=\frac{-c}{n}\frac{x^{a}(x-1)^{b}(x-\lambda)^{c-1}}{y^{n-1}}$$

$$D\left(\frac{1}{y}\right)=\frac{c}{n}\frac{x^{a}(x-1)^{b}(x-\lambda)^{c-1}}{y^{n+1}}=\frac{cx^{a}(x-1)^{b}(x-\lambda)^{c}}{ny^{n+1}(x-\lambda)}=\frac{c}{n}\cdot\frac{1}{(x-\lambda)y}$$

Therefore $D\left(\frac{dx}{y}\right)=\frac{c}{n}\left(\frac{1}{x-\lambda}\right)\frac{dx}{y}$, $D\left(\frac{xdx}{y}\right)=\frac{c}{n}\left(\frac{x}{x-\lambda}\right)\frac{dx}{y}$ and hence $D\left(\frac{xdx}{y}-\frac{dx}{y}\right)=\frac{c}{n}\left(\frac{x-1}{x-\lambda}\right)\frac{dx}{y}$

Now writing $f(x)=x^a(x-1)^b(x-\lambda)^c$ we have:

$$d(y^n)=f'(x)dx=[cx^a(x-1)^b(x-\lambda)^{c-1}+bx^a(x-1)^{b-1}(x-\lambda)^c$$
$$+ax^{a-1}(x-1)^b(x-\lambda)^c]dx$$

$$d\left(\frac{1}{y}\right)=-\frac{d(y)}{y^2}=-\frac{f'(x)dx}{ny^{n+1}}=-\frac{f'(x)}{ny^n}\cdot\frac{dx}{y}$$
$$=\left(\frac{-c}{n(x-\lambda)}+\frac{-b}{n(x-1)}+\frac{-a}{nx}\right)\frac{dx}{y}$$

$$d\left(\frac{x-1}{y}\right)=\frac{dx}{y}+(x-1)\left(\frac{-c}{n(x-\lambda)}-\frac{b}{n(x-1)}-\frac{a}{nx}\right)\frac{dx}{y}$$
$$=\frac{dx}{y}-\frac{c}{n}\left(\frac{x-1}{x-\lambda}\right)\frac{dx}{y}-\frac{b}{n}\frac{dx}{y}-\frac{a}{n}\left(\frac{x-1}{x}\right)\frac{dx}{y}$$
$$=\left(1-\frac{a+b}{n}\right)\frac{dx}{y}-\frac{c}{n}\left(\frac{x-1}{x-\lambda}\right)\frac{dx}{y}+\frac{a}{n}\frac{dx}{xy}$$

In order to eliminate (modulo exact) $\frac{dx}{xy}$, we calculate

$$d\left(\frac{(x-1)(x-\lambda)}{y}\right)=[2x-(1+\lambda)]\frac{dx}{y}-(x-1)(x-\lambda)$$
$$\times\left(\frac{c}{n(x-\lambda)}+\frac{b}{n(x-1)}+\frac{a}{nx}\right)\frac{dx}{y}$$
$$=[2x-(1+\lambda)]\frac{dx}{y}-\frac{c}{n}(x-1)\frac{dx}{y}-\frac{b}{n}(x-\lambda)\frac{dx}{y}$$
$$-\frac{(x-1)(x-\lambda)a}{nx}\frac{dx}{y}$$
$$=\left[2x-(1+\lambda)-\frac{(x-1)c}{n}-\frac{(x-\lambda)b}{n}\right]\frac{dx}{y}$$
$$-\frac{a}{n}\left(\frac{x^2-(1+\lambda)x+\lambda}{x}\right)\frac{dx}{y}$$
$$=\left[\frac{2n-(a+b+c)}{n}\right]\frac{xdx}{y}$$
$$+\left[\frac{c+\lambda b+a(1+\lambda)-n(1+\lambda)}{n}\right]\frac{dx}{y}-\frac{a}{n}\lambda\frac{dx}{xy}$$

Therefore $\frac{c}{n}\left(\frac{x-1}{x-\lambda}\right)\frac{dx}{y}=\left(1-\frac{a+b}{n}\right)\frac{dx}{y}+\frac{a}{n}\frac{dx}{xy}-d\left(\frac{x-1}{y}\right)$

$$\equiv\left(1-\frac{a+b}{n}\right)\frac{dx}{y}+\frac{1}{\lambda}\Big[\frac{2n-(a+b+c)}{n}\frac{xdx}{y}$$

$$+\left(\frac{c+\lambda b+a(1+\lambda)-n(1+\lambda)}{n}\right)\frac{dx}{y}$$

$$-d\frac{(x-1)(x-\lambda)}{y}\Big]$$

$$\equiv\left(\frac{a+c-n}{n\lambda}\right)\frac{dx}{y}+\left(\frac{2n-(a+b+c)}{n\lambda}\right)\frac{xdx}{y}$$

Lemma:

$$\begin{cases} D\left(\frac{dx}{y}\right)\equiv\left(\frac{n-(a+c)+c\lambda}{n\lambda(1-\lambda)}\right)\frac{dx}{y}+\left(\frac{a+b+c-2n}{n\lambda(1-\lambda)}\right)\frac{xdx}{y} \\ D\left(\frac{xdx}{y}\right)\equiv\left(\frac{n-a}{n(1-\lambda)}\right)\frac{dx}{y}+\left(\frac{a+b+c-2n}{n(1-\lambda)}\right)\frac{xdx}{y} \end{cases}$$

Proof: $d\left(-\frac{x}{y}\right)=-\frac{dx}{y}+x\left(\frac{c}{n(x-\lambda)}+\frac{b}{n(x-1)}+\frac{a}{nx}\right)\frac{dx}{y}$

$$=\left(\frac{a}{n}-1\right)\frac{dx}{y}+\frac{c}{n}\left(1+\frac{\lambda}{x-\lambda}\right)\frac{dx}{y}+\frac{b}{n}\left(1+\frac{1}{x-1}\right)\frac{dx}{y}$$

$$=\left(\frac{a+b+c}{n}-1\right)\frac{dx}{y}+\frac{c}{n}\left(\frac{\lambda}{x-\lambda}\right)\frac{dx}{y}+\frac{b}{n}\left(\frac{1}{x-1}\right)\frac{dx}{y}$$

$$d\left(\frac{x(x-\lambda)}{y}\right)=(2x-\lambda)\frac{dx}{y}$$

$$+x(x-\lambda)\left(-\frac{c}{n(x-\lambda)}-\frac{b}{n(x-1)}-\frac{a}{nx}\right)\frac{dx}{y}$$

$$=(2x-\lambda)\frac{dx}{y}-\frac{c}{n}x\frac{dx}{y}-\frac{a}{n}(x-\lambda)\frac{dx}{y}$$

$$-\frac{b}{n}\left(x+(1-\lambda)+\frac{1-\lambda}{x-1}\right)\frac{dx}{y}$$

$$=\left(2-\frac{a+b+c}{n}\right)\frac{xdx}{y}+\left(\frac{a\lambda}{n}+\frac{b(\lambda-1)}{n}-\lambda\right)\frac{dx}{y}$$

$$-\frac{b}{n}\left(\frac{1-\lambda}{x-1}\right)\frac{dx}{y}$$

Therefore $-\frac{b}{n}\left(\frac{1}{x-1}\right)\frac{dx}{y}\equiv\frac{1}{1-\lambda}\Big[\left(\frac{a+b+c}{n}-2\right)\frac{xdx}{y}$

$$+\left(\lambda-\frac{a\lambda}{n}+\frac{b(1-\lambda)}{n}\right)\frac{dx}{y}\Big]$$

But we have

$$D\left(\frac{dx}{y}\right)=\frac{c}{n}\left(\frac{1}{x-\lambda}\right)\frac{dx}{y}$$

$$\equiv\frac{1}{\lambda}\Big[\left(1-\frac{a+b+c}{n}\right)\frac{dx}{y}-\frac{b}{n}\left(\frac{1}{x-1}\right)\frac{dx}{y}\Big]$$

$$\equiv\Big[\frac{1}{\lambda}\left(1-\frac{a+b+c}{n}\right)+\frac{1}{\lambda(1-\lambda)}\left(\lambda-\frac{a\lambda}{n}+\frac{b(1-\lambda)}{n}\right)\Big]\frac{dx}{y}$$

$$+\frac{1}{\lambda(1-\lambda)}\left(\frac{a+b+c-2n}{n}\right)\frac{xdx}{y}$$

Combining this expression for $D\left(\frac{dx}{y}\right)$ with the result of the preceeding lemma, we find the desired formulae.

Let us denote by "$'$" the action of $\nabla\left(\frac{d}{dx}\right)$ on $H^1_{D.R.}(X)$. Then we have the following

Lemma: $\lambda(1-\lambda)\omega_1''+\Big[\frac{a+c}{n}-\left(\frac{a+b+2c}{n}\right)\lambda\Big]\omega_1'-\left(\frac{a+b+c-n}{n}\right)\frac{c}{n}\omega_1$
$=0$

Proof: Using the previous lemma we find:

$$\omega_2'-\lambda\omega_1'=\Big[\frac{n-a}{n(1-\lambda)}-\frac{n-(a+c)+c\lambda}{n(1-\lambda)}\Big]\omega_1=\frac{c}{n}\omega_1$$

$$\lambda\omega_1'+\frac{c}{n}\omega_1=\left(\frac{n-a}{n(1-\lambda)}\right)\omega_1+\left(\frac{a+b+c-2n}{n(1-\lambda)}\right)\omega_2$$

$$\frac{n\lambda\omega_1'+c\omega_1}{n}=\left(\frac{(n-a)\lambda}{n(1-\lambda)\lambda}\right)\omega_1+\left(\frac{a+b+c-2n}{n(1-\lambda)\lambda}\right)\omega_2$$

$$\lambda-\lambda^2(n\lambda\omega_1'+c\omega_1)=(n-a)\lambda\omega_1+(a+b+c-2n)\lambda\omega_2$$

$$(n\lambda^2-n\lambda^3)\omega_1'=[(n-a)\lambda-c(\lambda-\lambda^2)]\omega_1+(a+b+c-2n)\lambda\omega_2$$

$$(n\lambda-n\lambda^2)\omega_1'=[n-a-c(1-\lambda)]\omega_1+(a+b+c-2n)\omega_2$$

Therefore $(n-2n\lambda)\omega_1'+(n\lambda-\lambda^2)\omega_1''=c\omega_1+[n-a-c(1-\lambda)]\omega_1'+(a+b+c-2n)\omega_2'$. But $\omega_2'=\frac{c}{n}\omega_1+\lambda\omega_1'$. Therefore we obtain:

$$[n\lambda(1-\lambda)]\omega_1''+[n-2n\lambda-(n-a-c(1-\lambda))]\omega_1'-c\omega_1 \\ -(a+b+c-2n)\left(\lambda\omega_1'+\frac{c}{n}\omega_1\right)=0$$

and hence

$$\lambda(1-\lambda)\omega_1''+\left[\frac{a+c}{n}-\left(\frac{a+b+2c}{n}\right)\lambda\right]\omega_1'-\left(\frac{a+b+c-n}{n}\right)\frac{c}{n}\omega_1=0$$

3) We now show that our mapping is injective.

If not, there exist $\alpha, \beta\in \boldsymbol{Q}(\zeta_n, \lambda)$ such that $(\alpha x+\beta)\frac{dx}{y}$ is exact. Then at $\mathfrak{p}_0$ $\operatorname{ord}(\alpha x+\beta)\frac{dx}{y}\geq n-1-a$; at $\mathfrak{p}_1$ $\operatorname{ord}\left((\alpha x+\beta)\frac{dx}{y}\right)\geq n-1-b$; at $\mathfrak{p}_\lambda$ $\operatorname{ord}(\alpha x+\beta)\frac{dx}{y}\geq n-1-c$. But at $\mathfrak{p}_\infty$ $\operatorname{ord}(\alpha x+\beta)\frac{dx}{y}=a+b+c-n-1$ if $\alpha=0$ and $\operatorname{ord}(\alpha x+\beta)\frac{dx}{y}=a+b+c-2n-1$ if $\alpha\neq 0$. Let g be a function such that $dg=(\alpha x+\beta)\frac{dx}{y}$. Because $(\alpha x+\beta)\frac{dx}{y}$ has a pole at $\mathfrak{p}_\infty$ (as $n>a+b+c$), either $\operatorname{ord}_{\mathfrak{p}_\infty}(g) = a+b+c-n$ or $\operatorname{ord}_{\mathfrak{p}_\infty}(g)=a+b+c-2n$ depending on whether $\alpha=0$ or $\alpha\neq 0$.

Just as in part 1) above we have $g\in \boldsymbol{Q}(\zeta_n, \lambda)\left[x, y, \frac{1}{y}\right]$ because $(\alpha x+\beta)\frac{dx}{y}$ is regular except at $\mathfrak{p}_\infty$. Writing $g=P_0(x)+\frac{P_1(x)}{y}+\dots+\frac{P_{n-1}(x)}{y^{n-1}}$ with $P_i(x)\in \boldsymbol{Q}(\zeta_n, \lambda, x)$ and using the projection $\pi_\chi=\frac{1}{n}\Sigma\bar{\chi}(\sigma)\cdot\sigma$ on the relation $dg = (\alpha x+\beta)\frac{dx}{y}$ we find $d\left(\frac{P_1(x)}{y}\right)=(\alpha x+\beta)\frac{dx}{y}$. Thus we may assume $g = \frac{P_1(x)}{y}$. As $(\alpha x+\beta)\frac{dx}{y}$ is regular except at $\mathfrak{p}_\infty$, so is $\frac{P_1(x)}{y}$, hence also $P_1(x)$ and therefore $P_1(x)$ is a polynomial.

Now $\mathrm{ord}_{\mathfrak{p}_\infty}(x)=-n$ and thus $\mathrm{ord}_{\mathfrak{p}_\infty}(P_1(x))=-n\cdot\deg(P_1(x))$. Thus $P_1(x)$ has degree ≤ 2. As $\frac{P_1(x)}{y}$ is regular at $\mathfrak{p}_0$, $\mathfrak{p}_1$, $\mathfrak{p}_\lambda$ we find $x(x-1)(x-\lambda)$ divides $P_1(x)$. Thus $P_1(x)=0$ and $(\alpha x+\beta)\frac{dx}{y}=0$ which implies $\alpha=\beta=0$. This concludes the proof that $\mathrm{M}_{\frac{c}{n}, \frac{a+b+c}{n}-1, \frac{a+c}{n}} \to \mathrm{H}^1_{D.R.}(X)^\chi$ is injective.

Let S be a principal open set of Spec $\boldsymbol{Z}\left[\zeta_n, \lambda, \frac{1}{n\lambda(1-\lambda)}\right]$ over which there is a proper, irreducible, smooth S-scheme $\tilde{X}$ with $\tilde{X}\times_S$ Spec $\boldsymbol{Q}(\zeta_n, \lambda)=X$. We assume that S has been chosen sufficiently small so that $\mathrm{H}^*_{D.R.}(\tilde{X}/S)$ is locally free and commutes with base change. Furthermore we assume the horizontal isomorphism $\mathrm{M}_{\frac{c}{n}, \frac{a+b+c}{n}-1, \frac{a+c}{n}} \to \mathrm{H}^1_{D.R.}(X)^\chi$ extends to S. Thus we can state:

Theorem: *There is a non-empty open set S of Spec $\boldsymbol{Z}\left[\zeta_n, \lambda, \frac{1}{n\lambda(1-\lambda)}\right]$ and a horizontal isomorphism* $\mathrm{M}_{\frac{c}{n}, \frac{a+b+c}{n}-1, \frac{a+c}{n}} | S \xrightarrow{\sim} \mathrm{H}^1_{D.R.}(\tilde{X}/S)^\chi$.

Corollary: *For all but finitely many primes p,* $\mathrm{M}_{\frac{c}{n}, \frac{a+b+c}{n}-1, \frac{a+c}{n}} \otimes_{\boldsymbol{Z}} \boldsymbol{F}_p$ *is nilpotent.*

Proof: If a prime ideal $(p)(\neq 0)$ of $\boldsymbol{Z}$ belongs to the image of S, then $\mathrm{M}_{\frac{c}{n}, \frac{a+b+c}{n}-1, \frac{a+c}{n}} | S\otimes \boldsymbol{F}_p$ is a sub-module of $\mathrm{H}^1_{D.R.}(\tilde{X}\otimes \boldsymbol{F}_p/S\otimes \boldsymbol{F}_p)$. By the theorem of Katz and Berthelot: the Gauss-Manin connection (in characteristic p) is nilpotent, we see that $\mathrm{M} | S\otimes \boldsymbol{F}_p$ is nilpotent. This implies $\mathrm{M}_{\frac{c}{n}, \frac{a+b+c}{n}-1, \frac{a+c}{n}} \otimes \boldsymbol{F}_p$ is nilpotent.

The Theorem

Let us return momentarily to the general situation of the introduction; T arbitrary, S a smooth T-scheme, $\mathcal{M}$ a quasi-coherent S-module with a T-connection ∇, ... We note the following elementary facts:

1) If $(\mathcal{M}, \nabla_{\mathcal{M}})$ and $(\mathcal{N}, \nabla_{\mathcal{N}})$ are two S-modules with connection, D, m, n are sections of $\mathcal{D}er_{\mathcal{O}_T}(\mathcal{O}_S, \mathcal{O}_S)$, $\mathcal{M}$, $\mathcal{N}$ over an open subset $U \subseteq S$ and l is a strictly positive integer, then we have the Leibniz rule: $(\nabla_{\mathcal{M}\otimes\mathcal{N}}(D))^l(m\otimes n)=\sum_{i=0}^{l}\binom{l}{i}\nabla_{\mathcal{M}}(D)^i(m)\otimes\nabla_{\mathcal{N}}(D)^{l-i}(n)$ (proved as usual by induction on l)

2) Suppose $\mathcal{M}$ free of a fixed finite rank n, with base $\{e_1,\dots,e_n\}$. If D is a section of $\mathcal{D}er_{\mathcal{O}_T}(\mathcal{O}_S, \mathcal{O}_S)$ and if $\nabla(D)\begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix}=A_D\begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix}$ where $A_D\in \boldsymbol{M}_n(\mathcal{O}_S)$ is the so called "connection matrix", then $\nabla_{det(\mathcal{M})}(D)(e_1\wedge\dots\wedge e_n)=\operatorname{tr}(A_D)\cdot e_1\wedge\dots\wedge e_n$. We suppose in the next four statements that T is of characteristic p.

3) $\psi_{\mathcal{M}\otimes\mathcal{N}}(D)=\psi_{\mathcal{M}}(D)\otimes id_{\mathcal{N}}+id_{\mathcal{M}}\otimes\psi_{\mathcal{N}}(D)$
(because $(\nabla_{\mathcal{M}\otimes\mathcal{N}}(D))^p(m\otimes n)=\nabla_{\mathcal{M}}(D)^p(m)\otimes n+m\otimes\nabla_{\mathcal{N}}(D)^p(n)$ by Leibniz)

4) If ϕ: $\mathcal{M}\to\mathcal{N}$ is a horizontal morphism, $\psi_{\mathcal{N}}(D)\circ\phi=\phi\circ\psi_{\mathcal{M}}(D)$

5) Suppose $\mathcal{M}$ free of finite rank. A necessary and sufficient condition that $(\mathcal{M}, \nabla)$ be nilpotent is that for every section D of $\mathcal{D}er_{\mathcal{O}_T}(\mathcal{O}_S, \mathcal{O}_S)$, every coefficient except the leading one of the characteristic polynomial of $\psi(D)$ is nilpotent in $\mathcal{O}_S$.

6) If M is free of finite rank, then $\psi_{det(\mathcal{M})}(D)=\operatorname{tr}(\psi_{\mathcal{M}}(D))$

Having completed these preliminaries we turn to the main result. To fix notation again let $a, b, c\in \boldsymbol{Q}$, n be a common denominator and $S=\operatorname{Spec}\, \boldsymbol{Z}\left[\lambda, \frac{1}{n\lambda(1-\lambda)}\right]$. Let $\mathrm{M}_{a,b,c}$ be the hypergeometric S-module defined in the introduction.

Theorem: $\mathrm{M}_{a,b,c}$ *is globally nilpotent.*

Proof: Fix once and for all a prime p which does not become invertible in $\boldsymbol{Z}\left[\lambda, \frac{1}{n\lambda(1-\lambda)}\right]$. Consider the $\boldsymbol{F}_p\left[\lambda, \frac{1}{\lambda(1-\lambda)}\right]$-module (with connection) $\mathrm{M}_{a,b,c}\otimes_{\boldsymbol{Z}}\boldsymbol{F}_p$. We must show that it is nilpotent. By

statement 5) above this is equivalent to showing that the determinant and trace of $\psi\left(\frac{d}{d\lambda}\right)$ are zero. It suffices to show this at the generic point of Spec $\boldsymbol{F}_p\left[\lambda, \frac{1}{\lambda(1-\lambda)}\right]$ and therefore we shall work with the module $\mathrm{M}=\mathrm{M}_{a,b,c}\underset{S}{\otimes}\boldsymbol{F}_p(\lambda)$.
First we shall deal with the determinant.

Denoting by $\check{\mathrm{M}}$ the dual module, it is immediately checked that the mapping $\phi\mapsto\langle\phi, e_1\rangle$ establishes an $\boldsymbol{F}_p(\lambda^p)$-linear isomorphism between the horizontal elements of $\check{\mathrm{M}}$ and the solutions in $\boldsymbol{F}_p(\lambda)$ of the differential equation:

$$(*)\quad \lambda(1-\lambda)u''+[c-(a+b+1)\lambda]u'-abu=0.$$

Suppose for the moment that there is a non-zero solution in $\boldsymbol{F}_p(\lambda)$ of this equation, i.e., that $\check{\mathrm{M}}$ possesses a non-zero horizontal section. Then $\psi_{\check{\mathrm{M}}}\left(\frac{d}{d\lambda}\right)$ has determinant$=0$. Applying 3) and 4) above to the canonical horizontal morphism $\check{\mathrm{M}}\otimes\mathrm{M}\to\boldsymbol{F}_p(\lambda)$ we see that $-\psi_{\check{\mathrm{M}}}\left(\frac{d}{d\lambda}\right)$ is the transpose of $\psi_{\mathrm{M}}\left(\frac{d}{d\lambda}\right)$ and hence that $\det\left(\psi_{\mathrm{M}}\left(\frac{d}{d\lambda}\right)\right)=0$.

In order to find a non-zero solution of (*) we may assume that $a, b, c\in\boldsymbol{Z}$, $-(p-1)\leq a\leq 0$; $c<a$; $b, c\neq 0$ (in $\boldsymbol{Z}$). As is "well-known" [1], the differential equation

$$\lambda(1-\lambda)u''+[c-(a+b+1)\lambda]u'-abu=0 \quad\text{over}\quad \boldsymbol{Z}\left[\lambda, \frac{1}{\lambda(1-\lambda)}\right]$$

has a non-zero solution in $\boldsymbol{Q}[\lambda]$, namely

$$F(a, b, c; \lambda)=\sum_{r=0}^{-a}\frac{(a)_r(b)_r}{(c)_r r!}\lambda^r \quad\text{where}\quad \begin{cases}(\theta)_0=1\\ (\theta)_r=\theta(\theta+1)\ldots(\theta+r-1)\\ \qquad\text{for}\quad r\neq 0.\end{cases}$$

By multiplying $F(a, b, c; \lambda)$ by the least common multiple of the denominators of its coefficients we obtain a primitive polynomial in $\boldsymbol{Z}[\lambda]$ which is still a solution of this differential equation. The reduction mod p of this polynomial is the desired polynomial solution of (*).

This completes the proof that $\det\left(\psi\left(\frac{d}{d\lambda}\right)\right)=0$.

In order to show that $\operatorname{tr}\left(\psi\left(\frac{d}{d\lambda}\right)\right)=0$ we use statement 6) above, $\operatorname{tr}\left(\psi\left(\frac{d}{d\lambda}\right)\right)=\psi_{det(M)}\left(\frac{d}{d\lambda}\right)$. We observe that $\psi_{det(M)}\left(\frac{d}{d\lambda}\right)=0$ if and only if $\det(M)$ has a non-trivial horizontal section. By 2) above $\nabla_{det(M)}\left(\frac{d}{d\lambda}\right)=\frac{d}{d\lambda}+\frac{(a+b+1)\lambda-c}{\lambda(1-\lambda)}$. Thus it suffices to find $g\in\boldsymbol{F}_p(\lambda)$, $g\neq 0$ such that $\frac{dg}{d\lambda}+\left(\frac{(a+b+1)\lambda-c}{\lambda(1-\lambda)}\right)g=0$. But $g=\lambda_c(1-\lambda)^{a+b+1-c}$ is a nonzero solution of the equation, whence $\operatorname{tr}\left(\psi\left(\frac{d}{d\lambda}\right)\right)=0$; which completes the proof of the theorem.

INSTITUT DES HAUTES ÉTUDES SCIENTIFIQUES

References

[1] Bateman Manuscript Project (Erdélyi editor), Higher Transcendental Functions, Volume 1, Chapter II, McGraw-Hill, New York, 1953.
[2] Berthelot, P., Cohomologie p-cristalline des schémas I, II, III, C. R. Acad. Sc. Paris, t. 269 Sér. A (1969) pp. 297–300, 357–360, 397–400.
[3] Katz, N., Nilpotent Connections and the Monodromy Theorem Applications of a result of Turrittin, Pub. Math. IHES No. 39.
[4] Katz, N. and Oda, T., On the differentiations of De Rham cohomology classes with respect to parameters, J. Math. Kyoto Univ., 8-2 (1968) pp. 199–213.