

Crochet de Lie et équations différentielles linéaires

Comment la commutation explique (et remplace) la variation des constantes

Michel Hua

Octobre 2026

Résumé

On résout d'habitude l'équation $y'' + a(x)y' + b(x)y = f(x)$ par la méthode de variation des constantes, avec une condition auxiliaire, $C_1'\varphi_1 + C_2'\varphi_2 = 0$, qui paraît tomber du ciel. Cet article propose deux lectures de cette méthode fondées sur une seule notion, le *crochet de Lie* $[A, B] = AB - BA$. D'abord, en factorisant l'opérateur différentiel en deux opérateurs du premier ordre : c'est possible sans effort quand les coefficients commutent avec la dérivation, et le commutateur mesure exactement l'obstruction sinon. Ensuite, en regardant les translations par des solutions de l'équation homogène comme des transformations qui commutent : la variation des constantes est alors le changement de coordonnées qui les « redresse ». On passe ensuite à la forme matricielle, où la commutation décide si l'exponentielle donne la solution, puis on reprend les exemples d'une leçon filmée de Canal-U qui procède par « dérivée exacte » et par factorisation. On termine en réduisant les équations modulo un nombre premier p . Le crochet y fabrique un invariant, la p -courbure, et une question posée par Grothendieck, qu'on appelle aujourd'hui la conjecture de Grothendieck–Katz, demande si cet invariant détecte les équations dont les solutions sont algébriques. Elle est ouverte en général. Nous la démontrons dans les deux cadres de cet article : les coefficients constants, et l'ordre 1.

Leçons concernées. 221 (Équations différentielles linéaires, systèmes d'équations différentielles linéaires), 220 (Équations différentielles ordinaires, exemples), 156 (Exponentielle de matrices).

Développements possibles. Factorisation de L et équation de Riccati (proposition 4.1 et proposition 4.2) ; variation des constantes comme redressement de symétries qui commutent (section 5) ; condition de commutation pour l'exponentielle (proposition 6.1) et formule de Duhamel. La conjecture de Grothendieck–Katz à coefficients constants et à l'ordre 1 (théorèmes 8.11 et 8.15), avec la formule de Jacobson (proposition 8.12). Le fichier `CrochetLieED0.lean` vérifie la plupart des énoncés des sections 2 à 7 avec Mathlib ; la formalisation de la section 8 est en cours.

Prérequis. Équations différentielles linéaires d'ordre 1 et 2, wronskien, calcul matriciel (2×2), exponentielle de matrice. Les fonctions sont supposées de classe C^∞ sur un intervalle $I \subset \mathbb{R}$, ce qui évite de discuter la régularité.

Table des matières

1	Rappel : la variation des constantes	2
2	Opérateurs différentiels et commutateur	3
3	Coefficients constants : la factorisation est gratuite	4

4	Coefficients variables : le commutateur comme obstruction	4
5	Le point de vue de Lie : des symétries qui commutent	5
5.1	L'espace des phases	5
5.2	Les symétries de translation	5
5.3	Redresser les symétries, c'est faire varier les constantes	6
5.4	Quand les symétries ne commutent pas	7
6	Forme matricielle : quand l'exponentielle marche-t-elle ?	7
7	Exemples d'une leçon filmée (Canal-U)	8
7.1	Ordre 1 : la « dérivée exacte »	8
7.2	Ordre 2 : factoriser par conjugaison	9
7.3	Les exercices	9
8	Réduire modulo p : la p-courbure et une question de Grothendieck	10
8.1	Dériver modulo p	11
8.2	La p -courbure	11
8.3	Coefficients constants : le crochet est nul	12
8.4	L'ordre 1 : une formule de Jacobson	13
8.5	Conclusion : l'énoncé général	16
9	Synthèse	17
10	Exercices	17
11	Corrigés des exercices	18

1 Rappel : la variation des constantes

On considère sur I l'équation

$$y'' + a(x)y' + b(x)y = f(x), \quad (\text{E})$$

et l'équation homogène associée (H), obtenue avec $f = 0$. Si (φ_1, φ_2) est un système fondamental de solutions de (H), son wronskien

$$W = \varphi_1\varphi_2' - \varphi_2\varphi_1'$$

ne s'annule pas sur I .

La méthode classique cherche $y = C_1(x)\varphi_1 + C_2(x)\varphi_2$ en *imposant*

$$C_1'\varphi_1 + C_2'\varphi_2 = 0. \quad (1)$$

Alors $y' = C_1\varphi_1' + C_2\varphi_2'$, et en reportant dans (E) on obtient $C_1'\varphi_1' + C_2'\varphi_2' = f$. Le système de Cramer donne

$$C_1' = -\frac{\varphi_2 f}{W}, \quad C_2' = \frac{\varphi_1 f}{W}, \quad \text{d'où} \quad y = -\varphi_1 \int \frac{\varphi_2 f}{W} + \varphi_2 \int \frac{\varphi_1 f}{W}. \quad (2)$$

La méthode fonctionne, mais deux questions restent ouvertes : *pourquoi* la condition (1) est-elle légitime, et *pourquoi* tout se ramène-t-il à des primitives ? Les sections suivantes répondent à ces deux questions avec la même notion.

2 Opérateurs différentiels et commutateur

Définition 2.1. On note D l'opérateur de dérivation, $Dy = y'$. Pour une fonction q , on note M_q l'opérateur de multiplication, $M_q y = q y$. Quand il n'y a pas d'ambiguïté, on écrit simplement q au lieu de M_q .

Avec ces notations, (E) s'écrit $Ly = f$ où

$$L = D^2 + a D + b.$$

On compose les opérateurs comme des applications. *Attention*, l'ordre compte : $D \circ M_q$ n'est pas $M_q \circ D$.

Définition 2.2 (Crochet de Lie, ou commutateur). Pour deux opérateurs linéaires A et B , on pose

$$[A, B] = AB - BA.$$

On dit que A et B *commutent* si $[A, B] = 0$.

Le crochet est bilinéaire et antisymétrique ($[B, A] = -[A, B]$). Il vérifie aussi l'*identité de Jacobi*

$$[A, [B, C]] + [B, [C, A]] + [C, [A, B]] = 0,$$

qu'on peut vérifier en développant les douze termes.

Proposition 2.3 (Relation fondamentale). *Pour toute fonction q ,*

$$[D, M_q] = M_{q'}.$$

En particulier $[D, M_x] = \text{Id}$, et D commute avec M_q si et seulement si q est constante.

Démonstration. Pour toute fonction y , la règle de Leibniz donne

$$[D, M_q] y = (qy)' - qy' = q'y.$$

□

Remarque 2.4. La relation $[D, x] = 1$ est la même que la relation de Heisenberg $[\hat{p}, \hat{x}] = -i\hbar$ de la mécanique quantique. Toute la suite repose sur cette égalité d'apparence modeste.

Lemme 2.5 (Conjugaison). *Soit q une fonction et Q une primitive de q . Alors*

$$D - q = e^Q \circ D \circ e^{-Q},$$

où $e^{\pm Q}$ désigne la multiplication par $e^{\pm Q}$. Par conséquent, l'équation $(D - q)y = g$ a pour solutions

$$y = e^Q \int e^{-Q} g.$$

Démonstration. On calcule : $e^Q(e^{-Q}y)' = e^Q(e^{-Q}y' - q e^{-Q}y) = y' - qy$. Résoudre $e^Q D(e^{-Q}y) = g$ revient à résoudre $(e^{-Q}y)' = e^{-Q}g$, et il suffit de prendre une primitive. □

Le lemme 2.5 n'est rien d'autre que la méthode du facteur intégrant, écrite comme une conjugaison d'opérateurs. Il dit qu'un opérateur du premier ordre s'inverse toujours par deux multiplications et une primitive. La stratégie générale en découle : écrire L comme un produit d'opérateurs du premier ordre.

3 Coefficients constants : la factorisation est gratuite

Supposons a et b constants. Comme D commute avec les constantes, on peut manipuler L comme un polynôme en D . Si r_1, r_2 sont les racines (éventuellement complexes ou confondues) de $r^2 + ar + b$, on a

$$L = D^2 + aD + b = (D - r_1)(D - r_2).$$

Remarque 3.1. Cette factorisation n'est *pas* évidente a priori. Le développement de $(D - r_1)(D - r_2)$ fait intervenir $D r_2$, qu'on réécrit $r_2 D$, et cette réécriture est justement $[D, r_2] = 0$.

On résout donc $Ly = f$ en deux étapes du premier ordre, en posant $z = (D - r_2)y$:

$$(D - r_1)z = f, \quad (D - r_2)y = z.$$

Le lemme 2.5 donne

$$z = e^{r_1 x} \int e^{-r_1 s} f(s) ds, \quad y = e^{r_2 x} \int e^{-r_2 t} z(t) dt.$$

Exemple 3.2. On résout $y'' - 3y' + 2y = x$. On a $L = (D - 1)(D - 2)$.

- $(D - 1)z = x$: $z = e^x \int s e^{-s} ds = -x - 1 + A e^x$.
- $(D - 2)y = z$: une intégration par parties donne

$$y = e^{2x} \int e^{-2t} (-t - 1 + A e^t) dt = \frac{x}{2} + \frac{3}{4} - A e^x + B e^{2x}.$$

Vérification : pour $y_p = \frac{x}{2} + \frac{3}{4}$, on a $y_p'' - 3y_p' + 2y_p = 0 - \frac{3}{2} + x + \frac{3}{2} = x$.

On n'a eu besoin ni de deviner la forme de la solution particulière, ni d'imposer la condition (1). On obtient néanmoins le même résultat que par (2), avec $\varphi_1 = e^x$ et $\varphi_2 = e^{2x}$, ce que le lecteur peut vérifier.

4 Coefficients variables : le commutateur comme obstruction

Si a et b dépendent de x , on cherche encore une factorisation $L = (D - p)(D - q)$ avec p et q des *fonctions*. Développons en utilisant la proposition 2.3, qui donne $Dq = qD + q'$:

$$\begin{aligned} (D - p)(D - q) &= D^2 - Dq - pD + pq \\ &= D^2 - (qD + \underbrace{q'}_{=[D, q]}) - pD + pq \\ &= D^2 - (p + q)D + (pq - q'). \end{aligned}$$

Par rapport au cas constant, le seul terme nouveau est $-q'$, qui provient *exactement* du commutateur.

En identifiant avec $L = D^2 + aD + b$, on obtient $p = -a - q$ et $pq - q' = b$, c'est-à-dire :

Proposition 4.1. *L'opérateur $L = D^2 + aD + b$ se factorise en $(D - p)(D - q)$ si et seulement si q vérifie l'équation de Riccati*

$$q' + q^2 + aq + b = 0, \tag{3}$$

et $p = -a - q$.

L'équation (3) n'est pas linéaire, et en général on ne sait pas la résoudre par quadratures. Elle a cependant une interprétation simple.

Proposition 4.2. Si y_1 est une solution de (H) qui ne s'annule pas sur I , alors $q = y_1'/y_1$ vérifie (3). Autrement dit, connaître une solution de l'équation homogène, c'est connaître une factorisation de L .

Démonstration. On a $q' = \frac{y_1''}{y_1} - \frac{y_1'^2}{y_1^2}$, donc $q' + q^2 + aq + b = \frac{y_1'' + ay_1' + by_1}{y_1} = 0$. $\square$

Une fois la factorisation connue, on résout $Ly = f$ par deux applications du lemme 2.5, comme au paragraphe précédent. C'est la *méthode de réduction de l'ordre* (d'Alembert), présentée ici comme une factorisation.

Exemple 4.3. On résout $x^2y'' - 2xy' + 2y = x^3$ sur $]0, +\infty[$. Après division par x^2 , on a $a = -2/x$, $b = 2/x^2$ et $f = x$. La fonction $y_1 = x$ est solution évidente de (H), d'où $q = 1/x$ et $p = -a - q = 1/x$. Ainsi

$$L = \left(D - \frac{1}{x}\right)\left(D - \frac{1}{x}\right), \quad \text{avec} \quad D - \frac{1}{x} = x \circ D \circ \frac{1}{x}.$$

- $(D - \frac{1}{x})z = x : z = x \int \frac{1}{s} s \, ds = x^2 + Ax$.
- $(D - \frac{1}{x})y = z : y = x \int \frac{t^2 + At}{t} \, dt = \frac{x^3}{2} + Ax^2 + Bx$.

La deuxième solution homogène x^2 apparaît d'elle-même, sans qu'on l'ait cherchée.

Remarque 4.4. Les deux facteurs sont égaux ici, mais ils ne commutent pas en général : $(D - p)(D - q) \neq (D - q)(D - p)$ dès que $p' \neq q'$. On vérifie en effet que

$$[D - p, D - q] = -[D, q] - [p, D] = p' - q'.$$

L'ordre des facteurs, indifférent à coefficients constants, devient donc significatif.

5 Le point de vue de Lie : des symétries qui commutent

Cette section explique d'où vient la condition (1). L'idée remonte à Sophus Lie (vers 1870–1880) : une équation différentielle qui possède assez de *symétries* s'intègre par quadratures, et c'est la structure de ces symétries pour le crochet qui dit comment.

5.1 L'espace des phases

On transforme (E) en système du premier ordre en posant $Y = \begin{pmatrix} y \\ y' \end{pmatrix}$:

$$Y' = A(x)Y + F(x), \quad A = \begin{pmatrix} 0 & 1 \\ -b & -a \end{pmatrix}, \quad F = \begin{pmatrix} 0 \\ f \end{pmatrix}. \quad (4)$$

À chaque instant x , l'état du système est un point du plan (y, y') , appelé *espace des phases*.

5.2 Les symétries de translation

Soit φ une solution de (H) et $\Phi = \begin{pmatrix} \varphi \\ \varphi' \end{pmatrix}$. Si Y résout (4), alors $Y + t\Phi$ aussi, pour tout réel t , par linéarité. La transformation

$$T_t^\varphi : Y \longmapsto Y + t\Phi(x)$$

envoie donc solutions sur solutions : c'est une *symétrie* de l'équation. Quand t varie, elle forme un groupe à un paramètre, puisque $T_t^\varphi \circ T_s^\varphi = T_{t+s}^\varphi$. Le *champ de vecteurs* associé est le vecteur vitesse $\frac{d}{dt}\big|_{t=0} T_t^\varphi(Y) = \Phi$. Dans le langage de la géométrie différentielle, on l'écrit

$$X_\varphi = \varphi(x) \frac{\partial}{\partial y} + \varphi'(x) \frac{\partial}{\partial y'}.$$

Définition 5.1 (Crochet de deux champs de vecteurs). Un champ de vecteurs X agit sur les fonctions $g(y, y')$ comme un opérateur de dérivation : Xg est la dérivée de g dans la direction de X . Le crochet $[X, X']$ est le commutateur de ces opérateurs :

$$[X, X'] g = X(X'g) - X'(Xg).$$

On admettra le résultat suivant, classique en géométrie différentielle.

Théorème 5.2 (admis). *Deux champs de vecteurs ont un crochet nul si et seulement si leurs flots commutent. Si de plus ils sont linéairement indépendants en chaque point, il existe des coordonnées locales (C_1, C_2) dans lesquelles ils s'écrivent $\partial/\partial C_1$ et $\partial/\partial C_2$ (redressement simultané).*

Ici, on n'a même pas besoin du théorème : les translations $T_t^{\varphi_1}$ et $T_s^{\varphi_2}$ commutent visiblement, car $Y + t\Phi_1 + s\Phi_2 = Y + s\Phi_2 + t\Phi_1$. Les coefficients de X_{φ_1} et X_{φ_2} ne dépendent pas de (y, y') , donc $[X_{\varphi_1}, X_{\varphi_2}] = 0$. La figure 1 montre ce que cela veut dire : par les deux chemins, on arrive au même point.

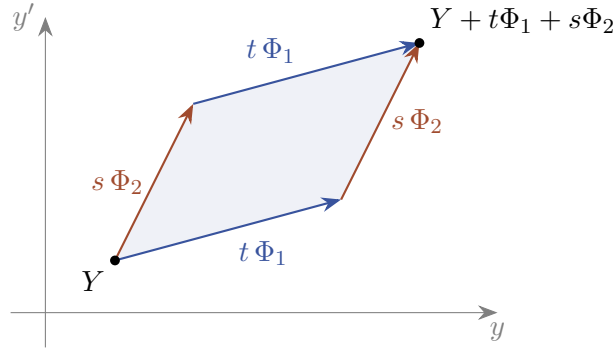


FIG. 1 : Dans le plan des phases, à x fixé : traduire d'abord par $t\Phi_1$ puis par $s\Phi_2$, ou dans l'ordre inverse, mène au même point. C'est la commutation $[X_{\varphi_1}, X_{\varphi_2}] = 0$. Les coordonnées (C_1, C_2) de la variation des constantes sont celles de ce quadrillage par Φ_1 et Φ_2 .

5.3 Redresser les symétries, c'est faire varier les constantes

Les vecteurs Φ_1 et Φ_2 sont indépendants exactement quand leur déterminant, le wronskien W , est non nul. Les coordonnées qui redressent X_{φ_1} et X_{φ_2} sont alors les coordonnées (C_1, C_2) dans la base (Φ_1, Φ_2) :

$$Y = C_1 \Phi_1 + C_2 \Phi_2, \quad \text{c'est-à-dire} \quad \begin{cases} y = C_1 \varphi_1 + C_2 \varphi_2, \\ y' = C_1 \varphi'_1 + C_2 \varphi'_2. \end{cases} \quad (5)$$

En effet, traduire Y de $t\Phi_1$ revient à traduire C_1 de t : dans ces coordonnées, $X_{\varphi_1} = \partial/\partial C_1$.

Proposition 5.3. *Dans les coordonnées (5), le système (4) devient*

$$C'_1 \Phi_1 + C'_2 \Phi_2 = F,$$

qui ne dépend plus de C_1 ni de C_2 . Il se résout donc par deux primitives, et l'on retrouve les formules (2).

Démonstration. On dérive $Y = C_1 \Phi_1 + C_2 \Phi_2$: $Y' = C'_1 \Phi_1 + C'_2 \Phi_2 + C_1 \Phi'_1 + C_2 \Phi'_2$. Comme $\Phi'_i = A\Phi_i$, les deux derniers termes valent AY , et (4) se réduit à $C'_1 \Phi_1 + C'_2 \Phi_2 = F$. Ce système de Cramer, de déterminant W , s'écrit composante par composante

$$C'_1 \varphi_1 + C'_2 \varphi_2 = 0, \quad C'_1 \varphi'_1 + C'_2 \varphi'_2 = f.$$

□

Remarque 5.4 (La réponse à la question de départ). La première ligne est *exactement* la condition (1). Elle n'a rien d'arbitraire : elle dit qu'on prend aussi des coordonnées pour y' , et pas seulement pour y , de façon à redresser les symétries dans tout l'espace des phases. Le fait qu'il ne reste que des primitives est la traduction de l'invariance par translation : l'équation ne peut pas dépendre de C_i puisque translater C_i ne change rien. Et on peut traiter C_1 et C_2 *indépendamment* parce que les deux symétries *commutent*.

5.4 Quand les symétries ne commutent pas

L'équation homogène (H) possède une symétrie de plus : la dilatation $Y \mapsto e^t Y$, de champ $S = y \partial_y + y' \partial_{y'}$. Un calcul direct donne

$$[X_\varphi, S] = X_\varphi.$$

Ces deux symétries ne commutent pas, mais le crochet de X_φ et S redonne X_φ : on dit que X_φ engendre un *idéal* et que l'algèbre $\langle X_\varphi, S \rangle$ est *résoluble*.

Le théorème d'intégration de Lie affirme qu'une équation d'ordre n possédant une algèbre résoluble de n symétries s'intègre par quadratures *successives* : d'abord en utilisant l'idéal, puis le quotient. Avec une seule solution φ connue, c'est précisément la situation de la section précédente : la factorisation $L = (D - p)(D - q)$ et les deux intégrations emboîtées traduisent la suite $\langle X_\varphi \rangle \subset \langle X_\varphi, S \rangle$.

En résumé :

- symétries qui *commutent* (deux solutions connues) : primitives indépendantes, c'est la variation des constantes ;
- symétries formant une algèbre *résoluble* (une solution connue) : primitives emboîtées, c'est la réduction de l'ordre.

6 Forme matricielle : quand l'exponentielle marche-t-elle ?

Pour un système $Y' = A(x)Y$, on est tenté d'écrire, comme en dimension 1,

$$Y(x) = \exp\left(\int_{x_0}^x A(s) ds\right) Y(x_0).$$

Proposition 6.1. *Posons $B(x) = \int_{x_0}^x A(s) ds$. Si $A(x)$ et $B(x)$ commutent pour tout x , ce qui est le cas en particulier si $[A(x), A(s)] = 0$ pour tous x, s , alors $\frac{d}{dx} e^{B(x)} = A(x) e^{B(x)}$, et la formule ci-dessus est correcte.*

Idée de preuve. On dérive terme à terme $e^B = \sum_k B^k/k!$. La dérivée de B^k est $\sum_j B^j A B^{k-1-j}$, et elle vaut $k A B^{k-1}$ uniquement si l'on peut faire passer A à gauche de chaque B^j , c'est-à-dire si $[A, B] = 0$. $\square$

- **Coefficients constants** : A est constante, donc commute avec elle-même, et la formule de Duhamel

$$Y(x) = e^{(x-x_0)A} \left(Y(x_0) + \int_{x_0}^x e^{-(s-x_0)A} F(s) ds \right)$$

donne directement la solution. C'est encore la variation des constantes, avec la matrice fondamentale $e^{(x-x_0)A}$.

- **Coefficients variables** : en général $[A(x), A(s)] \neq 0$ et l'exponentielle est fausse. Le *développement de Magnus* (hors programme) corrige l'exposant par des termes qui sont précisément des crochets :

$$\Omega(x) = \int_{x_0}^x A + \frac{1}{2} \int_{x_0}^x \int_{x_0}^{s_1} [A(s_1), A(s_2)] ds_2 ds_1 + \dots$$

7 Exemples d'une leçon filmée (Canal-U)

Une leçon vidéo de Canal-U, « Équations différentielles » (série *Mathématiques DEUG 1* des Amphis de France 5), consacrée aux équations différentielles linéaires d'ordre 1 et 2, ne parle ni de variation des constantes ni de commutateur. Elle utilise pourtant exactement les deux idées du lemme 2.5 et de la section 3 : la conjugaison pour l'ordre 1, la factorisation pour l'ordre 2. On reprend ici ses exemples dans le langage de cet article.

7.1 Ordre 1 : la « dérivée exacte »

Pour $a(x)y' + b(x)y = f(x)$, avec a qui ne s'annule pas, la leçon multiplie l'équation par une fonction $u(x)$ choisie pour que le membre de gauche devienne la dérivée d'un produit $(yB)' = y'B + yB'$. Il faut donc

$$B = u a, \quad B' = u b, \quad \text{d'où} \quad (ua)' = ub \quad \text{et} \quad u(x) = \frac{1}{a(x)} \exp \int_{x_0}^x \frac{b(t)}{a(t)} dt.$$

L'équation devient $(yB)' = uf$, et une intégration donne

$$y(x) = \underbrace{\frac{C}{B(x)}}_{\text{homogène}} + \underbrace{\frac{1}{B(x)} \int_{x_0}^x u(t) f(t) dt}_{\text{particulière}}.$$

Remarque 7.1 (Lien avec le lemme 2.5). Multiplier par u puis reconnaître $(yB)'$, c'est écrire

$$a y' + b y = \frac{1}{u} D(B y), \quad \text{soit} \quad a \left(D + \frac{b}{a} \right) = a B^{-1} \circ D \circ B,$$

c'est-à-dire la conjugaison du lemme 2.5 avec $q = -b/a$ et $e^{-Q} = B$, à constante multiplicative près (B/a est le multiplicateur u). La variation de la constante sur $y = C/B$ mène à $C' = uf$, qui est la même équation : les deux méthodes ne diffèrent que par la présentation.

Remarque 7.2 (Le facteur intégrant et le crochet de Lie). Le lien avec le crochet est double.

Côté opérateurs. Pour deux opérateurs A et B , la formule de Hadamard donne

$$e^A B e^{-A} = B + [A, B] + \frac{1}{2!} [A, [A, B]] + \dots$$

Avec $A = M_Q$ et $B = D$, la proposition 2.3 donne $[Q, D] = -M_{Q'}$, puis $[Q, [Q, D]] = -[Q, Q'] = 0$, car deux multiplications commutent. La série s'arrête après un seul crochet :

$$e^Q D e^{-Q} = D - Q' = D - q.$$

Le facteur intégrant fonctionne donc parce que le crochet de D avec une multiplication est encore une multiplication, et que les multiplications commutent entre elles.

Côté symétries. Lie a montré en 1874 que si l'équation $M dx + N dy = 0$ admet la symétrie $X = \xi \partial_x + \eta \partial_y$, alors $\mu = 1/(\xi M + \eta N)$ est un facteur intégrant. Pour $y' + p y = f$, on a $M = p y - f$ et $N = 1$, et la symétrie est la translation $X = \varphi \partial_y$ par une solution homogène $\varphi = e^{-\int p}$. On trouve

$$\mu = \frac{1}{\varphi} = e^{\int p},$$

qui est exactement le facteur intégrant de la leçon (à la division par a près). Le facteur intégrant est donc la version en dimension 1 du redressement de la section 5.

Exemple 7.3. On résout $y' \sin x + y \cos x = \cos x$ sur un intervalle où $\sin x \neq 0$. Le membre de gauche est déjà une dérivée exacte ($u = 1$, $B = \sin x$) :

$$(y \sin x)' = \cos x \implies y \sin x = \sin x + K \implies y = 1 + \frac{K}{\sin x}.$$

La formule générale donne bien $u = \frac{1}{\sin x} \exp \int \frac{\cos}{\sin} = 1$.

Exemple 7.4. On résout $y' \cos x + y \sin x = \sin x$ sur un intervalle où $\cos x \neq 0$. On cherche u tel que $(u \cos x)' = u \sin x$, soit $u' \cos x = 2u \sin x$, d'où $u = \frac{1}{\cos^2 x}$ et $B = \frac{1}{\cos x}$:

$$\left(\frac{y}{\cos x} \right)' = \frac{\sin x}{\cos^2 x} \implies \frac{y}{\cos x} = \frac{1}{\cos x} + C \implies y = 1 + C \cos x.$$

7.2 Ordre 2 : factoriser par conjugaison

Pour $ay'' + by' + cy = f(x)$ avec a, b, c constants et $a \neq 0$, la leçon propose d'abord le « petit exercice » suivant : pour toute fonction u ,

$$e^{px} \frac{d}{dx} \left(e^{-qx} \left[e^{qx} \frac{d}{dx} (e^{-qx} u) \right] \right) = u'' - (p + q) u' + pq u.$$

Avec les notations de l'article, chaque bloc $e^{px} \frac{d}{dx} e^{-px}$ est l'opérateur $D - p$ (lemme 2.5), et l'identité dit simplement

$$(D - p)(D - q) = D^2 - (p + q) D + pq.$$

Elle n'est vraie que parce que $[D, q] = 0$, q étant constant.

On choisit alors $p + q = -b/a$ et $pq = c/a$: p et q sont les racines de l'équation caractéristique $ar^2 + br + c = 0$. L'équation s'écrit

$$e^{qx} \frac{d}{dx} \left(e^{-qx} \left[e^{px} \frac{d}{dx} (e^{-px} y) \right] \right) = \frac{f(x)}{a},$$

et l'on « épiluche » les deux couches par deux intégrations successives. On obtient

$$e^{-px} y(x) = m + m' \int_{x_0}^x e^{(q-p)t} dt + \int_{x_0}^x g(t) dt, \quad g(t) = \frac{e^{(q-p)t}}{a} \int_{x_0}^t f(s) e^{-qs} ds.$$

- Si $p \neq q$, l'intégrale vaut $\frac{e^{(q-p)x} - e^{(q-p)x_0}}{q-p}$ et la partie homogène est $A e^{px} + B e^{qx}$.
- Si $p = q$, l'intégrale vaut $\int_{x_0}^x e^{0 \cdot t} dt = x - x_0$ et la partie homogène est $A e^{px} + B x e^{px}$. Le facteur x de la racine double n'a pas besoin d'être deviné : il sort du calcul.

La leçon s'arrête aux coefficients constants. À coefficients variables, la même tentative bute sur le terme $[D, q] = q'$ et conduit à l'équation de Riccati (section 4).

7.3 Les exercices

Exemple 7.5. On résout $y'' + 4y' + 3y = e^{-x} + \sin x$. Les racines de $r^2 + 4r + 3$ sont -3 et -1 . Dans la leçon, l'étudiant découpe le second membre par superposition, écrit $\sin x = \frac{e^{ix}}{2i} - \frac{e^{-ix}}{2i}$, et devine une solution particulière de chaque morceau : Axe^{-x} pour e^{-x} (il trouve $A = \frac{1}{2}$), $Be^{\pm ix}$ pour les deux autres. Il obtient

$$y = a e^{-3x} + b e^{-x} + \frac{x}{2} e^{-x} - \frac{1}{5} \cos x + \frac{1}{10} \sin x, \quad (a, b) \in \mathbb{R}^2.$$

Par factorisation, le morceau e^{-x} ne demande aucune intuition. On a $L = (D+3)(D+1)$; on pose $z = (D+1)y$:

$$(D+3)z = e^{-x} \implies z = e^{-3x} \int e^{2s} ds = \frac{1}{2} e^{-x}, \quad (D+1)y = \frac{1}{2} e^{-x} \implies y = e^{-x} \int \frac{1}{2} dt = \frac{x}{2} e^{-x}.$$

La résonance, c'est-à-dire le facteur x , apparaît d'elle-même, alors que l'étudiant devait savoir à l'avance qu'il fallait essayer Axe^{-x} .

Exemple 7.6. On résout $y'' + y' + y = \cos x$. Les racines de $r^2 + r + 1$ sont $j = -\frac{1}{2} + i\frac{\sqrt{3}}{2}$ et $j^2 = \bar{j}$. Les parties réelle et imaginaire de e^{jx} donnent la solution homogène. Dans la leçon, l'étudiante teste $\sin x$: $-\sin x + \cos x + \sin x = \cos x$, donc $\sin x$ est solution particulière, et

$$y = A e^{-x/2} \cos\left(\frac{\sqrt{3}}{2}x\right) + B e^{-x/2} \sin\left(\frac{\sqrt{3}}{2}x\right) + \sin x.$$

La factorisation fonctionne aussi sur $\mathbb{C}$, avec $L = (D-j)(D-j^2)$. Comme les coefficients sont réels, on résout avec e^{ix} et l'on prend la partie réelle :

$$(D-j)z = e^{ix} \Rightarrow z = \frac{e^{ix}}{i-j}, \quad (D-j^2)y = z \Rightarrow y = \frac{e^{ix}}{(i-j)(i-j^2)}.$$

Or $(i-j)(i-j^2) = i^2 - i(j+j^2) + j^3 = -1 + i + 1 = i$, donc $y = -i e^{ix}$, dont la partie réelle est $\sin x$.

8 Réduire modulo p : la p -courbure et une question de Grothendieck

Les équations rencontrées jusqu'ici ont des solutions de deux natures très différentes. L'équation $y' = \frac{y}{2x}$ a pour solution $\sqrt{x}$, qui vérifie l'équation polynomiale $y^2 = x$: on dit que c'est une fonction *algébrique*. L'équation $y' = y$ a pour solution e^x , qui ne vérifie aucune équation de ce genre. Voici la question de cette section : peut-on voir la différence sur l'équation elle-même, sans la résoudre ?

Une idée de Grothendieck répond, au moins en partie. Elle consiste à lire l'équation *modulo un nombre premier* p . Pour s'en faire une intuition, regardons les séries. Celle de $e^x = \sum x^n/n!$ a pour dénominateurs les factorielles, et tout nombre premier p divise $n!$ dès que $n \geq p$: cette série n'a aucun sens modulo p , quel que soit p . Celle de $\sqrt{1+u} = \sum \binom{1/2}{n} u^n$ n'a au dénominateur que des puissances de 2 : elle a un sens modulo tout p impair. Les solutions algébriques semblent « survivre » à la réduction modulo p , pour presque tout p , et les autres non.

Le crochet de Lie va transformer cette intuition en un calcul. Modulo p , la puissance p -ième de l'opérateur $D - A$ cesse d'être un opérateur différentiel : c'est une simple matrice, la *p -courbure*, et on peut la calculer sans résoudre l'équation. La question est alors de savoir si cette matrice nulle pour presque tout p suffit à garantir des solutions algébriques. Grothendieck l'a posée dans ses notes sur les connexions de Gauss–Manin ; on l'appelle aujourd'hui la conjecture de Grothendieck–Katz, et elle est ouverte. Nous allons la démontrer dans les deux cadres de cet article : les coefficients constants (section 6), et l'ordre 1 (lemme 2.5).

Notations. On note $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$ et $\mathbb{F}_p(x)$ le corps des fractions rationnelles à coefficients dans $\mathbb{F}_p$. La dérivation D y est définie par les formules habituelles, $D(x^n) = nx^{n-1}$ et la règle du quotient. Une fonction y définie sur un intervalle I est *algébrique* s'il existe un polynôme non nul $P \in \mathbb{C}[x, Y]$ tel que $P(x, y(x)) = 0$ pour tout $x \in I$.

8.1 Dériver modulo p

Modulo p , la dérivée se comporte bizarrement : $D(x^p) = p x^{p-1} = 0$, alors que x^p n'est pas constante. Ce défaut est en fait une force, et voici sa forme précise.

Lemme 8.1. *Soit $\mathcal{R}$ un anneau commutatif muni d'une dérivation D .*

1. *Pour tous $f, g \in \mathcal{R}$ et $n \geq 0$, $D^n(fg) = \sum_{k=0}^n \binom{n}{k} D^k f \cdot D^{n-k} g$ (formule de Leibniz).*
2. *Si $p \cdot 1 = 0$ dans $\mathcal{R}$, alors D^p est encore une dérivation : $D^p(fg) = f D^p g + g D^p f$.*
3. *Sur $\mathbb{F}_p(x)$, on a $D^p = 0$.*

Démonstration. 1. Par récurrence sur n , avec la relation de Pascal $\binom{n}{k} + \binom{n}{k-1} = \binom{n+1}{k}$.

2. Pour $0 < k < p$, le nombre premier p divise $\binom{p}{k} = \frac{p!}{k!(p-k)!}$, car il divise le numérateur et pas le dénominateur. Il ne reste que les termes $k = 0$ et $k = p$.

3. D'après le point 2, D^p est une dérivation de $\mathbb{F}_p(x)$, et $D^p(x) = 0$ puisque $D^2 x = 0$. Une dérivation nulle sur x est nulle sur tout polynôme, par la règle de Leibniz, puis sur tout quotient : en dérivant $g \cdot \frac{1}{g} = 1$, on trouve $D^p\left(\frac{1}{g}\right) = -\frac{D^p g}{g^2}$. $\square$

Le second outil est une propriété du crochet lui-même. On note ad_X l'opérateur $Y \mapsto [X, Y]$.

Lemme 8.2. *Soit $\mathcal{A}$ un anneau (non nécessairement commutatif) dans lequel $p \cdot 1 = 0$. Pour tous $X, Y \in \mathcal{A}$,*

$$\text{ad}_X^p(Y) = [X^p, Y].$$

Démonstration. Écrivons $\text{ad}_X = \lambda - \rho$, où $\lambda(Y) = XY$ et $\rho(Y) = YX$. Les deux opérations λ et ρ commutent, par associativité : $X(YX) = (XY)X$. On peut donc développer $(\lambda - \rho)^p$ par la formule du binôme, et les coefficients intermédiaires sont divisibles par p comme dans le lemme précédent : $(\lambda - \rho)^p = \lambda^p + (-1)^p \rho^p = \lambda^p - \rho^p$ (pour $p = 2$, $-1 = 1$). Or $\lambda^p(Y) - \rho^p(Y) = X^p Y - Y X^p$. $\square$

8.2 La p -courbure

Considérons un système $Y' = A(x)Y$, où A est une matrice $n \times n$ à coefficients dans $\mathbb{Q}(x)$. Une équation d'ordre 2 s'y ramène par la section 5.1. Pour presque tout p , aucun dénominateur des coefficients de A ne s'annule modulo p ; on dit alors que p est *bon*, et on note A_p la matrice réduite, à coefficients dans $\mathbb{F}_p(x)$.

Définition 8.3. Pour p bon, on note $\nabla_p = D - A_p$ l'opérateur qui agit sur les vecteurs $Y \in \mathbb{F}_p(x)^n$, D dérivant chaque coordonnée. La p -courbure du système est l'opérateur $\psi_p = (\nabla_p)^p$.

En caractéristique nulle, la puissance n -ième d'un opérateur du premier ordre est un opérateur différentiel d'ordre n . Modulo p , la puissance p -ième perd toutes ses dérivées, et c'est le crochet qui le montre.

Proposition 8.4. *La p -courbure est $\mathbb{F}_p(x)$ -linéaire : $\psi_p(fY) = f \psi_p(Y)$ pour toute $f \in \mathbb{F}_p(x)$. C'est donc la multiplication par une matrice $n \times n$ à coefficients dans $\mathbb{F}_p(x)$, encore notée ψ_p .*

Démonstration. La relation fondamentale (proposition 2.3) donne $[D, M_f] = M_{f'}$, et A_p commute avec la multiplication par la fonction f (une matrice commute avec les scalaires). Donc $[\nabla_p, M_f] = M_{f'}$, puis, en itérant, $\text{ad}_{\nabla_p}^k(M_f) = M_{D^k f}$. Avec $k = p$ et le lemme 8.1, $\text{ad}_{\nabla_p}^p(M_f) = M_{D^p f} = 0$. Le lemme 8.2 dit que c'est $[\nabla_p^p, M_f] : \psi_p$ commute avec toutes les multiplications. $\square$

Pourquoi cette matrice parle-t-elle de solutions ? Parce qu'elle tue toutes celles qui existent modulo p .

Proposition 8.5. *S'il existe une matrice Φ inversible, à coefficients dans $\mathbb{F}_p(x)$, dont les colonnes sont solutions ($\nabla_p \Phi = 0$), alors $\psi_p = 0$.*

Démonstration. On a $\psi_p \Phi = \nabla_p^p \Phi = 0$, et ψ_p est une matrice d'après la proposition 8.4. Comme Φ est inversible, $\psi_p = 0$. $\square$

La réciproque est vraie aussi : c'est le théorème de descente de Cartier, que nous n'utiliserons pas. Ainsi $\psi_p = 0$ signifie exactement que le système a, modulo p , autant de solutions indépendantes que d'inconnues. Une solution algébrique se réduisant, pour presque tout p , en une solution modulo p , on s'attend à ce qu'un système à solutions algébriques ait une p -courbure nulle pour presque tout p , et c'est bien un théorème (N. Katz). La question de Grothendieck porte sur la réciproque.

Conjecture 8.6 (Grothendieck–Katz). *Soit $A \in M_n(\mathbb{Q}(x))$. Le système $Y' = AY$ admet n solutions indépendantes dont toutes les coordonnées sont algébriques si et seulement si $\psi_p = 0$ pour presque tout nombre premier p .*

Remarque 8.7. N. Katz a démontré la conjecture pour les connexions de Gauss–Manin, c'est-à-dire pour les équations de Picard–Fuchs que satisfont les périodes d'une famille de variétés algébriques (*Invent. Math.* 18, 1972). Le cas d'une seule équation d'ordre 1 est connu en général (D. et G. Chudnovsky, 1985 ; J.-B. Bost, 2001 ; Y. André, 2004). Le cas général est ouvert. Les deux théorèmes qui suivent en sont des cas particuliers, dans les cadres de cet article, avec des démonstrations complètes et élémentaires.

8.3 Coefficients constants : le crochet est nul

Quand A est constante, tout se simplifie, pour la même raison qu'à la section 3 : D commute avec A .

Proposition 8.8. *Si $A \in M_n(\mathbb{Q})$ est constante et p bon, alors $\psi_p = -A_p^p$.*

Démonstration. La matrice A_p est constante, donc $[D, A_p] = 0$ (proposition 2.3, coefficient par coefficient). Pour deux opérateurs qui commutent, la formule du binôme s'applique, et les coefficients intermédiaires sont divisibles par p : $(D - A_p)^p = D^p + (-A_p)^p = D^p - A_p^p$. Enfin $D^p = 0$ par le lemme 8.1. $\square$

Exemple 8.9. Pour $y' = y$, on a $\psi_p = -1$ pour tout p : la p -courbure ne s'annule jamais, comme les factorielles de e^x l'annonçaient. Pour $y'' - 3y' + 2y = 0$ (exemple 3.2), la matrice est $A = \begin{pmatrix} 0 & 1 \\ -2 & 3 \end{pmatrix}$, de valeurs propres 1 et 2. Pour $p \geq 3$, ces valeurs propres sont distinctes dans $\mathbb{F}_p$, donc A_p est diagonalisable, et A_p^p a pour valeurs propres $1^p = 1$ et $2^p = 2$ (petit théorème de Fermat) : $A_p^p = A_p$ et $\psi_p = -A_p \neq 0$.

Pour énoncer le théorème, il nous faut un dernier fait, qui dit que l'exponentielle n'est pas algébrique.

Lemme 8.10. *Pour $\lambda \in \mathbb{C}$ non nul, la fonction $x \mapsto e^{\lambda x}$ n'est algébrique sur aucun intervalle.*

Démonstration. Supposons $P(x, e^{\lambda x}) = 0$ sur un intervalle I , avec $P = \sum_{k=0}^d c_k(x) Y^k$ non nul, et $c_d \neq 0$. La fonction $z \mapsto \sum_k c_k(z) e^{k\lambda z}$ est définie sur $\mathbb{C}$ tout entier et développable en série entière ; nulle sur I , elle est nulle partout, par le principe des zéros isolés. Prenons $z = s\bar{\lambda}$ avec s réel, de sorte que $e^{k\lambda z} = e^{k|\lambda|^2 s}$, et divisons par $e^{d|\lambda|^2 s}$:

$$c_d(s\bar{\lambda}) = - \sum_{k < d} c_k(s\bar{\lambda}) e^{-(d-k)|\lambda|^2 s} \xrightarrow{s \rightarrow +\infty} 0,$$

car l'exponentielle l'emporte sur tout polynôme. Mais un polynôme non nul ne tend pas vers 0 à l'infini : contradiction. $\square$

Théorème 8.11 (Grothendieck–Katz à coefficients constants). *Soit $A \in M_n(\mathbb{Q})$. Les conditions suivantes sont équivalentes :*

1. *toutes les solutions de $Y' = AY$ sont algébriques ;*
2. *A est nilpotente ;*
3. *$\psi_p = 0$ pour tout p bon et $\geq n$, donc pour presque tout p ;*
4. *$\psi_p = 0$ pour une infinité de nombres premiers p .*

L'idée est simple : d'après la proposition 8.8, la p -courbure est $-A_p^p$, et il s'agit de voir que « A^p est nul modulo p pour une infinité de p » force A à être nilpotente.

Démonstration. $2 \Rightarrow 1$. Si $A^n = 0$, la série de l'exponentielle s'arrête : $e^{xA} = \sum_{k < n} \frac{x^k}{k!} A^k$ est une matrice de polynômes, et les solutions $e^{xA} Y_0$ (section 6) aussi.

$2 \Rightarrow 3$. Si $A^n = 0$, alors $A_p^n = 0$, donc $A_p^p = 0$ dès que $p \geq n$. Et $3 \Rightarrow 4$ est évident.

$4 \Rightarrow 2$. Si $\psi_p = 0$, alors $A_p^p = 0$: la matrice A_p est nilpotente, donc toutes ses valeurs propres (dans une clôture algébrique de $\mathbb{F}_p$) sont nulles, et son polynôme caractéristique est t^n . Or ce polynôme caractéristique est la réduction modulo p de celui de A , $\chi_A(t) = t^n + c_{n-1}t^{n-1} + \dots + c_0$. Chaque coefficient c_j est donc, pour une infinité de p , un rationnel dont le numérateur est divisible par p . Un entier divisible par une infinité de nombres premiers est nul : $\chi_A = t^n$, et $A^n = 0$ par le théorème de Cayley–Hamilton.

$1 \Rightarrow 2$. Si A n'est pas nilpotente, elle a une valeur propre $\lambda \neq 0$ et un vecteur propre $v \in \mathbb{C}^n$. Alors $Y(x) = e^{\lambda x} v$ est solution, et l'une de ses coordonnées est $v_j e^{\lambda x}$ avec $v_j \neq 0$. Elle n'est pas algébrique, d'après le lemme 8.10. $\square$

8.4 L'ordre 1 : une formule de Jacobson

Passons à l'équation $y' = a(x)y$, c'est-à-dire à l'opérateur $D - a$ du lemme 2.5. Ici D ne commute plus avec a , et la formule du binôme ne s'applique pas. Il faut calculer $(D + b)^p$ autrement ; le résultat est remarquablement simple.

Proposition 8.12 (formule de Jacobson en rang 1). *Soit $\mathcal{R}$ un anneau commutatif muni d'une dérivation D , avec $p \cdot 1 = 0$, et $b \in \mathcal{R}$. Comme opérateurs sur $\mathcal{R}$,*

$$(D + b)^p = D^p + (b^p + D^{p-1}b).$$

On procède en trois étapes. La première ramène tout à une seule fonction, $u_m = (D + b)^m(1)$. La deuxième écrit u_m comme une somme sur les façons de découper $\{1, \dots, m\}$ en paquets. La troisième fait tourner ces découpages, et la plupart disparaissent modulo p .

Démonstration. Étape 1. Pour tout n et toute $y \in \mathcal{R}$, on a

$$(D + b)^n(y) = \sum_{k=0}^n \binom{n}{k} u_{n-k} D^k y.$$

C'est vrai pour $n = 0$. Si c'est vrai au rang n , on applique $D + b$: le terme $u_{n-k} D^k y$ donne $(Du_{n-k} + b u_{n-k}) D^k y + u_{n-k} D^{k+1} y = u_{n-k+1} D^k y + u_{n-k} D^{k+1} y$, et la relation de Pascal regroupe les termes. Pour $n = p$, les coefficients $\binom{p}{k}$ avec $0 < k < p$ sont nuls dans $\mathcal{R}$, et il reste $(D + b)^p(y) = u_p y + D^p y$.

Étape 2. Notons Π_n l'ensemble des partitions de $\{1, \dots, n\}$ en parties non vides, qu'on appelle des blocs. Alors

$$u_n = \sum_{\pi \in \Pi_n} \prod_{B \in \pi} D^{|B|-1} b.$$

Par récurrence, à partir de $u_{n+1} = Du_n + b u_n$. Une partition de $\{1, \dots, n+1\}$ s'obtient à partir d'une partition π de $\{1, \dots, n\}$ de deux façons. Ou bien $n+1$ forme un bloc à lui seul,

ce qui multiplie le produit par b : c'est le terme $b u_n$. Ou bien on ajoute $n + 1$ à l'un des blocs B de π , ce qui remplace le facteur $D^{|B|-1}b$ par $D^{|B|}b$; en sommant sur les blocs B , on obtient exactement la dérivée du produit, par la règle de Leibniz : c'est le terme Du_n .

Étape 3. Prenons $n = p$ et identifions $\{1, \dots, p\}$ à $\mathbb{Z}/p\mathbb{Z}$. La rotation $\sigma : i \mapsto i + 1$ agit sur les partitions, et ne change pas le produit associé, qui ne dépend que des tailles des blocs. Les orbites ont 1 ou p éléments, puisque $\sigma^p = \text{id}$ et que p est premier. Une orbite à p éléments contribue p fois le même produit, soit 0 dans $\mathcal{R}$. Restent les partitions fixées par σ . Si π est fixée et si B est le bloc qui contient 0, ou bien $\sigma(B) = B$, et alors B contient $0, 1, 2, \dots$, donc $B = \mathbb{Z}/p\mathbb{Z}$ et π n'a qu'un bloc; ou bien les translatés $\sigma^k(B)$, $0 \leq k < p$, sont p blocs distincts (si deux coïncidaient, B serait stable par une puissance de σ qui engendre le groupe, donc par σ), ce qui, dans un ensemble à p éléments, impose des singletons. Il ne reste que deux termes : un seul bloc, qui donne $D^{p-1}b$, et p singletons, qui donnent b^p . $\square$

Exemple 8.13. Pour $p = 3$: les cinq partitions de $\{1, 2, 3\}$ donnent $u_3 = b^3 + 3bb' + b''$. Les trois partitions en un bloc de deux et un singleton forment une orbite de la rotation, et leur contribution $3bb'$ disparaît modulo 3.

Appliquons la proposition à l'équation $y' = ay$, avec $\mathcal{R} = \mathbb{F}_p(x)$, $b = -a_p$ et $D^p = 0$. Notons $\Psi(b) = b^p + D^{p-1}b$. L'application Ψ est additive, parce que $(b+c)^p = b^p + c^p$ modulo p (binôme encore) et que D^{p-1} est linéaire. La p -courbure est donc

$$\psi_p = \Psi(-a_p) = -\Psi(a_p).$$

Tout se ramène à calculer Ψ sur les éléments simples d'une fraction rationnelle.

Lemme 8.14. Soit p premier, $\alpha, r, c \in \mathbb{F}_p$, et n, k des entiers avec $0 \leq n \leq p-2$ et $2 \leq k \leq p-1$. Dans $\mathbb{F}_p(x)$:

$$\Psi(\alpha x^n) = \alpha x^{np}, \quad \Psi\left(\frac{r}{x-c}\right) = 0, \quad \Psi\left(\frac{r}{(x-c)^k}\right) = \frac{r}{(x^p-c)^k}.$$

Démonstration. Les puissances p -ièmes se calculent avec le petit théorème de Fermat ($\alpha^p = \alpha$ dans $\mathbb{F}_p$) et la formule du binôme modulo p , qui donne $(x-c)^p = x^p - c^p = x^p - c$. Restent les dérivées $(p-1)$ -ièmes. D'abord $D^{p-1}(x^n) = n(n-1) \cdots (n-p+2) x^{n-p+1}$, et ce produit contient le facteur 0 puisque $n-p+2 \leq 0 \leq n$. Ensuite

$$D^{p-1}((x-c)^{-k}) = (-1)^{p-1} k(k+1) \cdots (k+p-2) (x-c)^{-k-p+1}.$$

Pour $k \geq 2$, les $p-1$ entiers consécutifs $k, \dots, k+p-2$ contiennent p , et le produit est nul. Pour $k = 1$, il vaut $(p-1)!$, qui est congru à -1 modulo p (théorème de Wilson); on trouve $D^{p-1}\left(\frac{r}{x-c}\right) = -\frac{r}{(x-c)^p}$, qui compense exactement $\left(\frac{r}{x-c}\right)^p = \frac{r}{(x-c)^p}$. (Pour $p = 2$, le signe $(-1)^{p-1}$ vaut $-1 = 1$ et le calcul est le même.) $\square$

Le lemme dit, en une phrase, que la p -courbure ignore les pôles simples et voit tout le reste. C'est exactement ce qui sépare x^r , dont la dérivée logarithmique r/x n'a qu'un pôle simple, de e^x ou $e^{1/x}$.

Théorème 8.15 (Grothendieck–Katz à l'ordre 1). Soit $a \in \mathbb{Q}(x)$ dont tous les pôles sont rationnels. Les conditions suivantes sont équivalentes :

1. l'équation $y' = ay$ a une solution non nulle algébrique, sur un intervalle où a n'a pas de pôle ;
2. $a = \sum_i \frac{r_i}{x-c_i}$, avec des $r_i \in \mathbb{Q}$ et des $c_i \in \mathbb{Q}$ distincts ;
3. $\psi_p = 0$ pour presque tout p ;
4. $\psi_p = 0$ pour une infinité de nombres premiers p .

Démonstration. $2 \Rightarrow 1$. Soit N un dénominateur commun des r_i . La fonction $y = \prod_i |x - c_i|^{r_i}$ vérifie $y'/y = \sum_i r_i/(x - c_i) = a$, et $y^N = \pm \prod_i (x - c_i)^{Nr_i}$ est une fraction rationnelle : y est algébrique.

$1 \Rightarrow 2$. Soit y une solution algébrique non nulle, et $P(Y) = Y^d + c_{d-1}Y^{d-1} + \dots + c_0$, à coefficients $c_k \in \mathbb{C}(x)$, un polynôme unitaire de degré minimal tel que $P(y) = 0$; il est irréductible. Dérivons l'identité $P(y(x)) = 0$, en utilisant $y' = ay$:

$$\sum_{k=0}^d (c'_k + k a c_k) y^k = 0 \quad (c_d = 1).$$

Le polynôme $Q(Y) = \sum_k (c'_k + k a c_k) Y^k$ a pour coefficient dominant $d a$, et $Q - d a P$, de degré $< d$, s'annule en y : il est nul par minimalité. Donc $c'_k = (d - k) a c_k$ pour tout k , et en particulier $c'_0 = d a c_0$. De plus $c_0 \neq 0$, sinon Y diviserait le polynôme irréductible P , qui serait Y , et y serait nulle. Écrivons $c_0 = \mu \prod_j (x - \gamma_j)^{m_j}$, avec $m_j \in \mathbb{Z}$ non nuls : alors $d a = c'_0/c_0 = \sum_j \frac{m_j}{x - \gamma_j}$. Ainsi $a = \sum_j \frac{m_j/d}{x - \gamma_j}$: ses pôles γ_j sont rationnels par hypothèse, et ses résidus m_j/d sont rationnels.

$2 \Rightarrow 3$. Pour p ne divisant aucun dénominateur des r_i et des c_i , l'additivité de Ψ et le lemme 8.14 donnent $\psi_p = -\sum_i \Psi\left(\frac{r_i}{x - c_i}\right) = 0$. Et $3 \Rightarrow 4$ est évident.

$4 \Rightarrow 2$. Décomposons a en éléments simples sur $\mathbb{Q}$, ce qui est possible puisque ses pôles sont rationnels :

$$a = P(x) + \sum_i \sum_{k=1}^{m_i} \frac{r_{ik}}{(x - c_i)^k}.$$

Pour p assez grand ($p > \deg P + 1$, $p > m_i$, p ne divisant aucun dénominateur, et les c_i distincts modulo p), l'additivité de Ψ et le lemme 8.14 donnent

$$\Psi(a_p) = R(x^p), \quad \text{où} \quad R(t) = P(t) + \sum_i \sum_{k=2}^{m_i} \frac{r_{ik}}{(t - c_i)^k} \in \mathbb{F}_p(t).$$

Les pôles simples ont disparu. Si $\psi_p = 0$, alors $R(x^p) = 0$, donc $R = 0$ (un polynôme $N(t)$ tel que $N(x^p) = 0$ est nul). Par unicité de la décomposition en éléments simples sur $\mathbb{F}_p$, les coefficients de P et les r_{ik} avec $k \geq 2$ sont nuls modulo p . Si cela arrive pour une infinité de p , ces rationnels sont nuls, et $a = \sum_i \frac{r_{i1}}{x - c_i}$. $\square$

Exemple 8.16. Pour $y' = \frac{r}{x} y$, avec $r \in \mathbb{Q}$, on trouve $\psi_p = 0$ pour presque tout p , et en effet x^r est algébrique. Pour $y' = -\frac{1}{x^2} y$, le lemme donne $\psi_p = -\Psi(-x^{-2}) = x^{-2p} \neq 0$, et en effet la solution $e^{1/x}$ n'est pas algébrique. Dans l'exemple 4.3, l'opérateur se factorise en deux facteurs $D - \frac{1}{x}$, de p -courbure nulle, et les solutions x et x^2 sont des polynômes. On peut le vérifier directement : la matrice $\Phi = \begin{pmatrix} x & x^2 \\ 1 & 2x \end{pmatrix}$ des solutions et de leurs dérivées a pour déterminant x^2 , inversible dans $\mathbb{F}_p(x)$, et la proposition 8.5 donne $\psi_p = 0$ pour tout p .

Remarque 8.17 (le lien avec la cote 93). La formule $\Psi(a) = R(x^p)$ dit que la p -courbure garde la partie non logarithmique de a , lue en x^p . Grothendieck écrit la même chose page 207 du dossier 93, pour la connexion $d + \omega$ sur une variété : la p -courbure y est $\omega^{\otimes p} - F^*(\tilde{C}\omega)$, où $\tilde{C}$ est l'opération de Cartier. Pour $\omega = a dx$ sur une droite, $\tilde{C}$ garde exactement les termes $\frac{r dx}{x - c}$ (et ceux en $x^{p-1} dx$), et le terme $F^*(\tilde{C}\omega)$ correspond à notre $D^{p-1}a$, au signe près. Quand les pôles de a ne sont pas rationnels, ses résidus ρ sont des nombres algébriques. Le même calcul montre alors que $\rho^p \equiv \rho$ modulo les idéaux premiers au-dessus de presque tout p , et un théorème d'arithmétique (Kronecker, Chebotarev) en déduit que ρ est rationnel. C'est la seule étape qui manque ici pour le cas général de l'ordre 1.

Revenons à la question du début. La p -courbure se lit sur l'équation, sans la résoudre : il suffit de calculer une puissance de matrice ou une dérivée $(p-1)$ -ième. Dans les deux cadres de cet article, elle répond exactement : nulle pour presque tout p si et seulement si les solutions sont algébriques. Et c'est le crochet de Lie, à travers les lemmes 8.2 et 8.1, qui fait de ∇_p^p une matrice plutôt qu'un opérateur différentiel.

8.5 Conclusion : l'énoncé général

La conjecture 8.6 est le cas d'une seule variable. Grothendieck la pose pour plusieurs variables, et c'est encore le crochet de Lie qui en fixe le cadre. Un système d'équations aux dérivées partielles

$$\frac{\partial Y}{\partial x_i} = A_i(x_1, \dots, x_m) Y \quad (1 \leq i \leq m)$$

ne peut avoir assez de solutions que si les opérateurs $\nabla_i = \partial_i - A_i$ commutent : en dérivant $\partial_i Y = A_i Y$ par rapport à x_j et en échangeant les rôles de i et j , on trouve la condition $[\nabla_i, \nabla_j] = 0$. Ce crochet s'appelle la *courbure* du système, et le système est dit *intégrable* quand elle est nulle. Modulo p , chaque ∇_i^p est encore linéaire, par le même argument qu'à la proposition 8.4 : ce sont les p -courbures $\psi_{p,i}$. Le mot « p -courbure » vient de là.

Le cadre naturel est celui d'une variété algébrique lisse X sur un corps de nombres K (par exemple le plan privé de quelques courbes, sur $\mathbb{Q}$) et d'un fibré vectoriel E sur X muni d'une connexion intégrable ∇ : localement, c'est exactement un système comme ci-dessus, à coefficients fonctions rationnelles. Pour presque tout idéal premier $\mathfrak{p}$ de l'anneau des entiers de K , on peut réduire (E, ∇) modulo $\mathfrak{p}$, et la p -courbure de la réduction est l'application $\partial \mapsto \nabla_{\partial}^p - \nabla_{\partial[p]}$, où $\partial[p]$ est la dérivation ∂^p (nulle pour $\partial = \partial_i$, ce qui redonne les $\psi_{p,i}$).

Conjecture 8.18 (Grothendieck–Katz, énoncé général). *Soit X une variété lisse sur un corps de nombres K , et (E, ∇) un fibré vectoriel sur X muni d'une connexion intégrable. Les conditions suivantes sont équivalentes :*

1. *(E, ∇) admet localement une base de solutions algébriques ; de façon équivalente, (E, ∇) devient trivial après un revêtement fini étale de X ou, sur $\mathbb{C}$, sa monodromie est un groupe fini ;*
2. *pour presque tout idéal premier $\mathfrak{p}$ de l'anneau des entiers de K , la p -courbure de la réduction de (E, ∇) modulo $\mathfrak{p}$ est nulle.*

Le sens $1 \Rightarrow 2$ est démontré ; c'est l'autre sens que Grothendieck demande, page 194 du dossier 93. N. Katz l'a démontré pour les connexions de Gauss–Manin (1972), et le cas où la monodromie est un groupe résoluble, qui contient le rang 1, est connu (Chudnovsky, Bost, André). Le cas général est l'un des grands problèmes ouverts sur les équations différentielles. Les théorèmes 8.11 et 8.15 en donnent une idée concrète : dans chacun, la p -courbure a mesuré exactement l'écart entre une équation et une équation à solutions algébriques.

9 Synthèse

Méthode	Ce qui doit commuter	Résultat
Factorisation de L	D et les coefficients	immédiate si constants; sinon équation de Riccati
Variation des constantes	des translations $X_{\varphi_1}, X_{\varphi_2}$	deux primitives indépendantes, formule avec W
Réduction de l'ordre	$\langle X_\varphi, S \rangle$ seulement résoluble	deux primitives emboîtées
Exponentielle de matrice	$A(x)$ et $A(s)$	formule fermée si le crochet est nul
p -courbure	∇ avec les multiplications, modulo p	∇_p^p est une matrice; nulle pour presque tout p si et seulement si les solutions sont algébriques (démontré à coefficients constants et à l'ordre 1)

Dans chaque cas, le crochet de Lie mesure ce qui empêche d'intégrer « d'un seul coup ». S'il est nul, tout se découple. Si les symétries forment une algèbre résoluble, on intègre par étapes. Sinon, il faut une information supplémentaire, typiquement une solution de l'équation de Riccati. Et modulo p , le même crochet fait apparaître un invariant qui, d'après la conjecture de Grothendieck–Katz, dirait si les solutions sont algébriques.

10 Exercices

Exercice 1. Montrer que pour tout entier $n \geq 1$, $[D, M_{x^n}] = n M_{x^{n-1}}$, puis que $[D^2, M_q] = 2q'D + q''$.

Exercice 2. Résoudre $y'' - 2y' + y = e^x$ par factorisation $(D - 1)^2$, puis par variation des constantes, et comparer les calculs. *Indication : on trouve $y_p = \frac{x^2}{2}e^x$.*

Exercice 3. Vérifier que $y_1 = e^x$ est solution de $xy'' - (x+1)y' + y = 0$ sur $]0, +\infty[$. En déduire une factorisation de l'opérateur associé, puis toutes les solutions. *Indication : on obtient $q = 1$ et $p = 1/x$, puis la solution $x + 1$.*

Exercice 4. Calculer $[X_\varphi, S]$ en appliquant les deux champs à une fonction $g(y, y')$ quelconque, et retrouver $[X_\varphi, S] = X_\varphi$.

Exercice 5. Soit $A(x) = \begin{pmatrix} 0 & 1 \\ 0 & x \end{pmatrix}$. Calculer $[A(x), A(s)]$ et en déduire que la formule exponentielle de la proposition 6.1 n'est pas applicable directement.

Exercice 6. Calculer la p -courbure du système associé à $y'' + y = 0$ pour p impair, et montrer qu'elle n'est jamais nulle. Qu'en déduit-on sur les solutions? Retrouver le résultat avec la solution e^{ix} .

Exercice 7. Soit $a = \frac{1}{x} + \frac{1}{x^2}$. Calculer ψ_p pour l'équation $y' = ay$ et $p \geq 3$. Conclure avec le théorème 8.15, puis vérifier que $y = xe^{-1/x}$ est solution.

Exercice 8. Vérifier la formule $u_3 = b^3 + 3bb' + b''$ de l'exemple qui suit la proposition 8.12, en calculant $(D + b)^3(1)$ directement. Écrire de même u_4 à l'aide des quinze partitions de $\{1, 2, 3, 4\}$, et vérifier que $u_4 \neq b^4 + b'''$ en général : la simplification est propre aux nombres premiers.

11 Corrigés des exercices

Exercice 1. Pour toute fonction y , la règle de Leibniz donne

$$[D, M_{x^n}] y = (x^n y)' - x^n y' = n x^{n-1} y,$$

ce qui est la proposition 2.3 avec $q = x^n$. Pour la seconde formule, on développe $D^2 q - q D^2 = D(Dq - qD) + (Dq - qD)D$, soit

$$[D^2, M_q] = D[D, q] + [D, q]D = Dq' + q'D.$$

En appliquant encore la proposition 2.3 à q' , on a $Dq' = q'D + q''$, d'où $[D^2, M_q] = 2q'D + q''$.

Exercice 2. *Par factorisation.* On pose $z = (D - 1)y$. L'équation $(D - 1)z = e^x$ donne $z = e^x \int e^{-s} e^s ds = (x + A)e^x$, puis $(D - 1)y = z$ donne

$$y = e^x \int (t + A) dt = \left(\frac{x^2}{2} + Ax + B \right) e^x.$$

Par variation des constantes. Avec $\varphi_1 = e^x$ et $\varphi_2 = x e^x$, on trouve $W = e^x(1 + x)e^x - x e^x e^x = e^{2x}$. Les formules (2) donnent $C'_1 = -x$ et $C'_2 = 1$, donc $C_1 = -\frac{x^2}{2}$, $C_2 = x$ et $y_p = -\frac{x^2}{2}e^x + x \cdot x e^x = \frac{x^2}{2}e^x$.

Comparaison. La variation des constantes demande de connaître d'avance la seconde solution $x e^x$ et de calculer W . La factorisation n'utilise que la racine double 1 : le facteur x de la seconde solution et le x^2 de la solution particulière sortent des deux primitives.

Exercice 3. Avec $y_1 = e^x$, on a $x e^x - (x + 1)e^x + e^x = 0$. Après division par x , l'équation s'écrit $y'' + a y' + b y = 0$ avec $a = -\frac{x+1}{x}$ et $b = \frac{1}{x}$. La proposition 4.2 donne $q = y'_1/y_1 = 1$, puis $p = -a - q = \frac{x+1}{x} - 1 = \frac{1}{x}$. On vérifie l'équation de Riccati : $q' + q^2 + a q + b = 0 + 1 - \frac{x+1}{x} + \frac{1}{x} = 0$. Ainsi

$$L = \left(D - \frac{1}{x} \right) (D - 1).$$

On pose $z = (D - 1)y$. L'équation $(D - \frac{1}{x})z = 0$ donne $z = Ax$, puis $(D - 1)y = Ax$ donne

$$y = e^x \int A t e^{-t} dt = -A(x + 1) + B e^x.$$

Les solutions sont donc les $y = C(x + 1) + B e^x$, et l'on vérifie directement que $x + 1$ convient : $x \cdot 0 - (x + 1) \cdot 1 + (x + 1) = 0$.

Exercice 4. À x fixé, φ et φ' sont des constantes, et l'on note $g_y, g_{y'}, g_{yy}, \dots$ les dérivées partielles de g . D'une part,

$$X_\varphi(Sg) = X_\varphi(y g_y + y' g_{y'}) = \varphi g_y + \varphi' g_{y'} + T,$$

où $T = \varphi(y g_{yy} + y' g_{yy'}) + \varphi'(y g_{yy'} + y' g_{y'y'})$ regroupe les termes du second ordre. D'autre part,

$$S(X_\varphi g) = S(\varphi g_y + \varphi' g_{y'}) = y(\varphi g_{yy} + \varphi' g_{yy'}) + y'(\varphi g_{yy'} + \varphi' g_{y'y'}) = T,$$

grâce au lemme de Schwarz. La différence ne garde que les termes du premier ordre : $[X_\varphi, S]g = \varphi g_y + \varphi' g_{y'} = X_\varphi g$.

Exercice 5. On calcule

$$A(x)A(s) = \begin{pmatrix} 0 & s \\ 0 & xs \end{pmatrix}, \quad A(s)A(x) = \begin{pmatrix} 0 & x \\ 0 & xs \end{pmatrix}, \quad [A(x), A(s)] = \begin{pmatrix} 0 & s-x \\ 0 & 0 \end{pmatrix},$$

qui n'est pas nul dès que $x \neq s$. L'hypothèse de la proposition 6.1 porte en fait sur $A(x)$ et $B(x) = \int_0^x A = \begin{pmatrix} 0 & x \\ 0 & x^2/2 \end{pmatrix}$; elle échoue aussi, puisque

$$[A(x), B(x)] = \begin{pmatrix} 0 & x^2/2 \\ 0 & x^3/2 \end{pmatrix} - \begin{pmatrix} 0 & x^2 \\ 0 & x^3/2 \end{pmatrix} = \begin{pmatrix} 0 & -x^2/2 \\ 0 & 0 \end{pmatrix} \neq 0$$

pour $x \neq 0$. On ne peut donc pas conclure que $e^{B(x)}$ est une matrice fondamentale.

Exercice 6. Le système est $Y' = AY$ avec $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, constante, donc $\psi_p = -A_p^p$ (proposition 8.8). Comme $A^2 = -I$, on a $A^p = A(A^2)^{(p-1)/2} = (-1)^{(p-1)/2}A$, et $\psi_p = \pm A_p \neq 0$. Le théorème 8.11 (condition 4 fausse) dit que les solutions ne sont pas toutes algébriques. En effet $e^{ix} = \cos x + i \sin x$ est solution, et elle n'est pas algébrique d'après le lemme 8.10.

Exercice 7. D'après le lemme 8.14, $\Psi(1/x) = 0$ et $\Psi(1/x^2) = 1/x^{2p}$ (avec $c = 0$, $k = 2$), donc $\psi_p = -\Psi(a_p) = -x^{-2p} \neq 0$ pour tout $p \geq 3$. Par le théorème 8.15 (condition 4 fausse), l'équation n'a pas de solution algébrique non nulle : le pôle double de a l'interdit. Et $y = x e^{-1/x}$ vérifie bien $y'/y = \frac{1}{x} + \frac{1}{x^2}$.

Exercice 8. On a $u_1 = b$, $u_2 = (D+b)(b) = b' + b^2$, puis $u_3 = (D+b)(b' + b^2) = b'' + 2bb' + bb' + b^3 = b^3 + 3bb' + b''$. Pour $\{1, 2, 3, 4\}$: un bloc (1 partition, b'''), un bloc de trois et un singleton (4 partitions, $4bb''$), deux blocs de deux (3 partitions, $3b'^2$), un bloc de deux et deux singletons (6 partitions, $6b^2b'$), quatre singletons (b^4). Donc $u_4 = b^4 + 6b^2b' + 3b'^2 + 4bb'' + b'''$, qui est le total des 15 partitions. Modulo 2, il reste $b^4 + b'^2 + b'''$, et le terme b'^2 ne disparaît pas : la rotation de $\{1, 2, 3, 4\}$ a des orbites de taille 2, et la partition $\{1, 3\}, \{2, 4\}$ est fixée par elle.

Sources. La leçon filmée « Équations différentielles », série *Mathématiques DEUG 1* des Amphis de France 5, sur Canal-U. Pour la p -courbure, le dossier 93 du fonds Grothendieck de l'université de Montpellier, « Connexions de Gauss-Manin et équations de Picard-Fuchs. Opération de Cartier », pages 193 à 194 et 204 à 212, dont une lecture est disponible sur <https://grothendieck.commutator.io>; et N. Katz, « Nilpotent connections and the monodromy theorem », *Publ. Math. IHÉS* 39 (1970).