Annales de l'institut Fourier

Michel Langevin

Quelques remarques sur les familles canoniques de polynômes générateurs pour l'exponentielle

Annales de l'institut Fourier, tome 47, nº 1 (1997), p. 1-48

<http://www.numdam.org/item?id=AIF_1997__47_1_1_0>

Ann. Inst. Fourier, Grenoble
47, 1 (1997), 1–48

QUELQUES REMARQUES SUR LES FAMILLES CANONIQUES DE POLYNÔMES GÉNÉRATEURS POUR L'EXPONENTIELLE

par Michel LANGEVIN

---

1. Introduction.

Soit K un corps commutatif. Le premier but de ce travail est de décrire l'ensemble NW $=$ NW(K) (*polynômes de Newton*) des suites de polynômes $(H_n(X))$ de $K[X]$ vérifiant, pour $n \geq 0$

$$H_n(X+Y) = \sum_{0 \leq i \leq n} H_i(X)H_{n-i}(Y)$$

ou, par récurrence sur le nombre k d'indéterminées :

$$(\mu) \quad H_n(X_1 + \cdots + X_k) = \sum_{\alpha_1+\cdots+\alpha_k=n} H_{\alpha_1}(X_1)\cdots H_{\alpha_k}(X_k) \qquad (n,k \geq 0).$$

On conserve la notation NW pour les séries génératrices

$$H(X,T) = \sum_n H_n(X)T^n \in K[X][[T]]$$

vérifiant la relation équivalente (*relation génératrice canonique de l'exponentielle*) :

$$(\mu') \qquad \begin{cases} H(X+Y,T) = H(X,T)H(Y,T) \text{ ou} \\ H(X_1+\cdots+X_k,T) = \prod_{1\leq i\leq k} H(X_i,T)\big). \end{cases}$$

Mots-clés : Exponentielle – Séries formelles – Séries génératrices – Formule de Taylor – Réduction de Jordan – Partitions d'entiers – Coefficients binomiaux – Algorithmique combinatoire – Hauteurs des polynômes – Géométrie des polynômes.
Classification math. : 11C08 – 11J99 – 13F20 – 13F25.

Les suites (resp. séries génératrices)

$$\Big(\frac{1}{n!}X^n\Big) \quad \big(\text{resp. } \exp(XT)\big),$$
$$\Big(\frac{1}{n!}X(X-1)\cdots\big(X-(n-1)\big)\Big) \quad \big(\text{resp. } (1+T)^X\big)$$

appartiennent à NW et sont les seules d'usage courant mais on en verra beaucoup d'autres étrangères à celles-ci. À noter des procédés de construction interne dans NW comme :

- pour tout entier $k > 0$ et tout $(H_n(X)) \in \mathrm{NW}$, la suite $(K_n(X))$, où $K_n = H_{n/k}$ (resp. 0) si n est (resp. n'est pas) multiple de k, est aussi élément de NW (en termes de séries, remplacer T par T^k) ;

- si les suites (F_n) et (G_n) (ou $F(X,T) = \sum\limits_n F_n(X)T^n$ et $G(X,T) = \sum\limits_n G_n(X)T^n$) sont éléments de NW et $\lambda, \mu \in K$, alors (vérification directe)

$$\Big(L_n(X) = \sum_{0\leq i\leq n} F_i(\lambda X)G_{n-i}(\mu X)\Big)$$

(ou encore $L(X,T) = F(\lambda X,T)G(\mu X,T)$) appartient à NW.

En particulier, la loi

$$\big(F(X,T),G(X,T)\big) \longmapsto L(X,T) = F(X,T)G(X,T)$$

fait de NW, privé de l'élément trivial 0, un groupe où l'élément neutre est 1 et l'inverse de $F(X,T)$ est $F(-X,T)$.

Ce but est aussi un prétexte; prétexte à d'autres développements puisque le problème initial se ramène à une équation fonctionnelle simple $\Phi(X+Y) = \Phi(X)\Phi(Y)$ dans l'anneau des séries formelles à une indéterminée à coefficients dans un anneau commutatif unitaire intègre A (observer pour cela que $H(X,T) = \sum\limits_n H_n(X)T^n$ est un élément de $K[[X,T]] = K[[T]][[X]]$). Cette généralisation fournit un procédé de résolution (*cf.* §10) de (μ) (ou (μ')) rappelant qu'une hypothèse supplémentaire factice peut alourdir un problème... mais c'est ce dernier qui est intéressant comme le montreront ces développements par rapport au « premier but » en fait facilement atteint.

On esquisse ci-dessous le plan puis ces développements (traités directement sans appel aux considérations du §10) liés à la détermination de NW en montrant de plus quelques résultats faciles.

Des travaux préalables de réduction (§§2 à 4) — intéressants par la partie combinatoire utilisée en caractéristique finie (§ 3) — montreront qu'il suffit d'étudier le sous-ensemble NW′ de NW formé des éléments vérifiant $H_0 = 1$, $H_1 = X$. Puis, on donnera deux visions indépendantes de l'ensemble NW′, l'une « exponentielle » (§ 4) et l'autre « différentielle » (§ 5) reliées ensuite (§ 7 et 8). Ces liens permettront de munir NW′ de multiples structures affines ou de groupes, allant du canonique au commode, imbriquées de façon amusante (§ 9). Les techniques de la vision « différentielle » ouvrent des perspectives de généralisation développées pour leur intérêt propre : lemmes de commutation en algèbre linéaire et formules de Taylor généralisées à une ou deux indéterminées (§ 6). Après la résolution dans un anneau de séries formelles au § 10 d'équations comme $\Phi(X+Y) = \Phi(X)\Phi(Y)$, les généralisations, applications potentielles et l'insertion parmi les travaux du domaine seront évoquées aux § 11 et 12 ; en particulier, les relations avec les développements récents de la « théorie des hauteurs » seront détaillées dans un sous-paragraphe particulier à l'intérieur du § 11.

On décrit ces visions après avoir défini l'objet jouant un rôle central dans ce travail : le groupe GSF (GSF(K)) des séries formelles $B(T) = \sum_{n>0} b_n T^n$ inversibles pour la composition (*i.e.* vérifiant $b_1 \neq 0$) et normalisées par la condition $b_1 = 1$ (*cf.* [H] pour les calculs explicites dans ce groupe).

Ce groupe est canoniquement un espace affine de direction isomorphe à $K[[T]]$; on notera que

$$(A) \quad \begin{cases} \text{dans GSF, lequel est un groupe (pour la composition) et} \\ \text{un espace affine, la translation à gauche (resp. à droite) par} \\ B : C \mapsto B \circ C \text{ (resp. } C \mapsto C \circ B) \text{ est non affine (resp. affine).} \end{cases}$$

Les résultats de réduction dans l'étude des éléments (H_n) de NW sont de deux sortes :

1) Pour éviter le cas trivial $H_n \equiv 0$ $(n \geq 0)$, on suppose $H_0 = 1$ et K de caractéristique 0 ;

2) Tout élément de NW se déduit d'un élément de NW′, sous-ensemble de NW formé des éléments vérifiant $H_0 = 1$ et $H_1 = X$; on verra alors que $\deg(H_n) = n$ et que (H_n) est donc une base de $K[X]$.

La « vision exponentielle » de NW (utilisée en fait dans la preuve du 2 ci-dessus) est simple. Si K est de caractéristique nulle, tout élément

$B(T) = \sum_{n>0} b_n T^n$ de $K[[T]]$ définit un élément de NW par :

$$\exp(B(T)X) = \sum_{n\geq 0} (B(T))^n \frac{X^n}{n!} = \sum_{n\geq 0} H_n(X)T^n. \tag{1}$$

Réciproquement, tout élément (H_n) de NW fournit un élément de $K[[T]]$ s'écrivant, H'_n ou $\mathrm{d}H_n$ désignant la dérivée de H_n,

$$B(T) = \sum_{n\geq 0} H'_n(0)T^n.$$

Par conséquent, le groupe additif $K[[T]]$ et le groupe multiplicatif $\mathrm{NW} - \{0\}$ sont isomorphes par les correspondances ci-dessus.

La «vision différentielle» de NW est plus subtile et on en limite l'exposition au cas de NW$'$. Est associé à tout élément (H_n) de NW$'$ (lequel est une base de $K[X]$) un endomorphisme de décalage h défini par :

$$h(H_0) = 0, \quad h(H_n) = H_{n-1} \quad (\text{pour } n > 0)$$

dont on montrera qu'il est un opérateur différentiel lié à la dérivation classique d par

$$\mathrm{d} = \sum_{n\geq 0} H'_n(0)h^n.$$

On obtient une réciproque en montrant que tout endomorphisme h de $K[X]$, de noyau K, vérifiant $h(X) = 1$ et $\deg(h(P)) = \deg(P) - 1$ pour tout polynôme P non constant, est, quand il commute avec d, de forme $B(d)$ avec $B \in \mathrm{GSF}$; et on montrera qu'un unique élément (H_n) de NW$'$, base de Jordan pour h, lui est associé par :

$$H_0 = 1, \quad h(H_n) = H_{n-1} \quad \text{et} \quad H_n(0) = 0 \ \text{ pour } n > 0. \tag{J'}$$

Ce qu'on va donc décrire dans ce travail sera une double action, libre et transitive, du groupe GSF dans NW, mais l'une à gauche et l'autre à droite. Il en résulte, au plan affine, une structure canonique de NW et une famille de structures paramétrée par GSF (dont certaines très simples).

Les développements extérieurs indépendants sont de plusieurs ordres. Le § 3 rappelle des lemmes combinatoires sur les polynômes symétriques et permet (en particulier) de traiter le cas de la caractéristique $p > 0$. Le § 6 est consacré à une approche purement linéaire de la vision différentielle, laquelle permettra de dégager des mécanismes formels simples. Ces mécanismes concernent des généralisations de la formule de Taylor, laquelle est en fait liée intimement avec NW$'$. Précisément, en notant γ l'application $P(X) \mapsto P(X+Y)$, l'identité $\gamma = \exp(Y\mathrm{d})$ représente la formule de Taylor classique sur $K[X]$. On la généralisera en $\gamma = \sum_{n \geq 0} H_n(Y)h^n$ où (H_n) est un élément de NW$'$, base de Jordan (*cf.* (J$'$)) pour l'opérateur différentiel h (représentable par une série formelle en d ou γ; une seule des conditions précédentes suffisant pour assurer l'autre).

Aucune application directe de ce travail n'est abordée ici, à l'exception des développements possibles évoqués aux § 11 et 12. On s'en tient à une étude systématique *ab ovo* de NW laquelle fournit à la fois des résultats non triviaux avec des notions simples dont elle éclaire les liens comme séries génératrices, réduction de Jordan, formules de Taylor, exponentielle... et un cadre formel général enrichissant pour les applications. L'intérêt à ce titre de dégager un tel cadre est particulièrement bien illustré dans le domaine voisin — comportant de nombreuses similitudes avec ce travail — des polynômes exponentiels (*cf.* le § 2 de [BG]). Ce point de vue n'est pas celui adopté en général dans la littérature; en effet, les éléments de NW apparaissent le plus souvent comme des outils dans des ouvrages d'une autre portée; quelques exemples : en géométrie, les morphismes de classes de Chern (*cf.* [FL])), en théorie des nombres, la théorie des partitions (*cf.* [R])... sans évoquer les applications dans le domaine des « équations aux différences » en analyse et en analyse numérique. Mais, dans ces cas, les structures de NW sont limitées aux applications *ad hoc*.

2. Théorème de réduction (K quelconque).

THÉORÈME 1. — *Soient K un corps commutatif et $(H_n(X))$ une suite d'éléments de $K[X]$ vérifiant (μ).*

(i) *$H_0(X)$ est une constante égale à 0 ou 1; si $H_0 = 0$, alors $H_n = 0$ pour tout $n \geq 0$;*

(ii) *$H_n(0) = 0$ pour tout $n > 0$;*

(iii) *pour tout $\lambda \in K$, la suite $(\lambda^n H_n(X))$ vérifie également (μ);*

(iv) *si K est de caractéristique 0 et si $H_0 = 1$, $H_1 \neq 0$, alors H_n est de degré n et de coefficient dominant $(c_1)^n/n!$, où c_1 est le coefficient dominant de H_1;*

(v) *si K est de caractéristique $p > 0$, alors on a $H_n(X) = 0$ pour tout $n > 0$ (voir le paragraphe 3).*

Démonstration.

Partie (i). — Le polynôme H_0 vérifiant $H_0 = (H_0)^2$ est une constante égale à 0 ou 1. Dans le premier cas l'identité

$$H_n(X+Y) = \sum_{0\le i\le n} H_i(X)H_{n-i}(Y)$$

montre que $H_n = 0$ pour tout n par récurrence.

Partie (ii). — Il suffit de spécialiser en $Y = 0$ pour déduire de

$$H_n(X+Y) = \sum_{0\le i\le n} H_i(X)H_{n-i}(Y)$$

l'égalité $H_n(0) = 0$ par récurrence pour $n \ge 1$. Ce résultat montre que $\mathrm{NW} - \{0\}$ est bien un groupe.

Partie (iii). — Évident.

Partie (iv). — L'hypothèse montre que

$$H_n(X+x) - H_n(X) = \sum_{0\le i<n} H_i(X)H_{n-i}(x) \qquad (x \in K),$$

prouvant, K étant de caractéristique 0 et x pouvant être choisi tel que $H_1(x) \ne 0$, que $\deg(H_1) = 1$ puis, de même et par récurrence, que $H_n \ne 0$ et $\deg(H_n) = n$. Soit c_n le coefficient dominant de H_n ; comme $c_0 = 1$, on voit, en choisissant $Y = X$, que

$$2^n c_n = \sum_{0\le i\le n} c_i c_{n-i} \qquad (n \ge 0)$$

et on en déduit par récurrence que $c_n = c_1^n/n!$. □

Remarque. — On verra au § 10, remarque 1, qu'est substantiellement équivalente à celle de (μ), *i.e.* du problème initial, la résolution de l'équation de récurrence :

$$2^n c_n = \sum_{0\le i\le n} c_i c_{n-i} \qquad (n \ge 0).$$

Partie (v). — On montrera au § 3 suivant la congruence modulo p :

$$H_{np}(pX) = \big(H_n(X)\big)^p \qquad (n \ge 0).$$

3. Lemmes combinatoires multinomiaux (et preuve du théorème 1 (v)).

DÉFINITION. — *Soient $i_1, \ldots, i_k$ (avec $k > 0$) des entiers ≥ 0. On désigne par « monôme symétrique » ou « polynôme symétrique minimal » $\sum\limits_{\text{sym}} X_1^{i_1} X_2^{i_2} \cdots X_k^{i_k}$ le polynôme symétrique en les variables $X_1, \ldots, X_k$ contenant $X_1^{i_1} X_2^{i_2} \cdots X_k^{i_k}$ et de plus petite longueur (et de hauteur 1).*

On rappelle que la longueur (resp. hauteur) d'un polynôme est égale à la somme (resp. la borne supérieure) des valeurs absolues de ses coefficients.

La donnée de la suite $i_1 \geq \cdots \geq i_k$ est équivalente à celle de $\sum\limits_{\text{sym}} X_1^{i_1} X_2^{i_2} \cdots X_k^{i_k}$, ou à celle de l'ensemble $\{j_1, \ldots, j_s\} = \mathrm{Im}(i)$ (rangé dans l'ordre décroissant) des valeurs prises par les entiers $i_1, \ldots, i_k$ auquel on adjoint les multiplicités associées $\mu_1 = \#(i^{-1}(j_1)), \ldots, \mu_s$ (d'où $\sum\limits_a \mu_a = k$).

Soit donc $\mathrm{BM}(n,k)$ l'ensemble des couples $\{j, \mu\}$ (plutôt que (j, μ) pour des raisons typographiques) formés d'une application strictement décroissante j d'un intervalle entier $[1, s]$ (avec $1 \leq s \leq k$) à valeurs dans les entiers ≥ 0 et d'une application μ définie sur le même ensemble $[1, s]$ et à valeurs dans les entiers > 0 tels que $\sum\limits_a \mu_a j_a = n$ et $\sum\limits_a \mu_a = k$. On écrira ainsi

$$\mathrm{SYM}_{\{j,\mu\}} \quad (\text{où } \{j, \mu\} \text{ appartient à } \mathrm{BM}(n,k))$$

le monôme symétrique $\sum\limits_{\text{sym}} X_1^{i_1} X_2^{i_2} \cdots X_k^{i_k}$ homogène de degré n.

LEMME 1. — *La longueur de* $\mathrm{SYM}_{\{j,\mu\}} = \sum\limits_{\text{sym}} X_1^{i_1} X_2^{i_2} \cdots X_k^{i_k}$ *est*

$$\frac{k!}{\mu_1! \ldots \mu_s!}.$$

Démonstration. — Cette longueur est égale à l'ordre du quotient du groupe symétrique $\mathfrak{S}_k$ par le sous-groupe laissant invariant les blocs de cardinaux $\mu_1, \ldots, \mu_s$ associés à $\{j_1, \ldots, j_s\}$. Par exemple, la longueur du polynôme symétrique élémentaire de degré global q en les indéterminées $X_1, \ldots, X_k$ (lequel est minimal et correspond au cas $j_1 = i_1 = \cdots = i_q = 1$, $j_2 = i_{q+1} = \cdots = i_k = 0$) est $\binom{k}{q}$.

Démonstration du théorème 1, partie (v).

Le lemme 1 permet de démontrer l'identité $H_{np}(pX) = (H_n(X))^p$ annoncée dans la démonstration de ce (v) au § 2. En effet, dans le contexte ci-dessus, l'égalité générale :

$$H_n(kX) \;=\; \sum_{\{j,\mu\}\in \mathrm{BM}(n,k)} \text{longueur}\big(\mathrm{SYM}_{\{j,\mu\}}\big)\big(H_{j_1}(X)\big)^{\mu_1}\cdots\big(H_{j_s}(X)\big)^{\mu_s}$$

s'écrit alors d'après le lemme 1, $n > 0$ étant un entier et $p > 0$ la caractéristique de K :

$$H_{np}(pX) = \sum_{\{j,\mu\}\in \mathrm{BM}(np,p)} \frac{p\,!}{\mu_1!\ldots\mu_s!}\big(\mathrm{SYM}_{\{j,\mu\}}\big)\big(H_{j_1}(X)\big)^{\mu_1}\cdots\big(H_{j_s}(X)\big)^{\mu_s}.$$

Comme

$$\sum_{1\le a\le s} \mu_a = p \text{ et } \sum_{1\le a\le s} \mu_a j_a = np \quad \text{avec } \mu_a > 0 \ (1 \le a \le s)$$

le coefficient multinomial $p\,!/(\mu_1!\ldots\mu_s!)$ est nul en caractéristique p sauf lorsque $s = 1$, d'où $j_1 = n$ et $\mu_1 = p$. □

On conclut ce paragraphe avec des relations trouvant une interprétation naturelle dans la suite.

La dimension de l'espace $\mathrm{SYM}(n,k)$ des polynômes symétriques homogènes à k indéterminées de degré n est égale au nombre de monômes symétriques à k indéterminées de degré n, donc au cardinal $\#\mathrm{BM}(n,k)$, autrement dit au nombre $p_k(n)$ de partitions de n en au plus k parts (quand $k \ge n$, le nombre $\#\mathrm{BM}(n,k) = p(n)$, nombre de partitions de n, est indépendant de k), ou, par conjugaison, au nombre de partitions de n en parts de cardinal $\le k$ (*cf.* [R] pour les théorèmes d'Euler sur l'écriture en produit (infini quand $k \ge n$) des séries génératrices associées à ces fonctions).

Les autres applications s'énoncent ainsi (on renvoie à [L1] ou [L2] pour d'autres correspondances entre polynômes et polynômes symétriques) :

COROLLAIRE. — *Soient $k > 0$ et $n \ge 0$ des entiers.*

(i) $$\sum_{\{j,\mu\}\in \mathrm{BM}(n,k)} \frac{k\,!}{\mu_1!\cdots\mu_s!} = \binom{n+k-1}{n}.$$

(ii) $$\sum_{\{j,\mu\}\in \mathrm{BM}(n,k)} \frac{n\,!\,k\,!}{(j_1!)^{\mu_1}\cdots(j_s!)^{\mu_s}\mu_1!\cdots\mu_s!} = k^n.$$

(iii) *Soit $M(s,k)$ l'ensemble (abrégé en $M(s)$) des applications μ de $[1,s]$ dans les entiers > 0 vérifiant $\sum_{1\le a\le s} \mu_a = k$.*

1) *$M(s)$ a pour cardinal $\#M(s) = \binom{k-1}{s-1}$;*

2) *les dimensions de l'espace des polynômes symétriques en k indéterminées de degré $< n$ en chacune et de l'espace des polynômes en k indéterminées de degré global $< n$ sont égales à*

$$\sum_{1 \leq s \leq k} \binom{n}{s} (\#M(s)) = \binom{n+k-1}{k}.$$

(iv) *Les entiers $a_{s,k} = \sum_{\mu \in M(s)} \frac{k!}{\mu_1! \cdots \mu_s!}$ possèdent les propriétés :*

1) $n^k = \sum_{1 \leq s \leq k} a_{s,k} = \sum_{s \geq 0} \binom{n}{s}$;

2) $a_{s,k} = \sum_{0 \leq i \leq s} (-1)^i \binom{s}{i} (s-i)^k.$

Démonstration.

Partie (i). — L'ensemble des monômes composant les $\mathrm{SYM}_{\{j,\mu\}}$ quand $\{j,\mu\}$ décrit $\mathrm{BM}(n,k)$ est l'ensemble des monômes de degré n en k indéterminées, d'où le résultat puisque $\binom{n+k-1}{n}$ est la dimension de l'espace des polynômes homogènes de degré n en k indéterminées (*i.e.* le nombre des applications i des entiers de $[1,k]$ dans les entiers ≥ 0 vérifiant $\sum_{1 \leq a \leq k} i_a = n$ dont on vérifie classiquement qu'il satisfait à la relation constitutive du triangle de Pascal).

On peut aussi énoncer ce résultat avec les notations du (iii) ; en notant $J(\mu)$ l'ensemble des applications strictement décroissantes j de $[1,s]$ dans $[0,n]$ vérifiant $\sum_a \mu_a j_a = n$, il vient :

$$\sum_{1 \leq s \leq k} \Big(\sum_{\mu \in M(s)} (\#J(\mu)) \frac{k!}{\mu_1! \ldots \mu_s!} \Big) = \binom{n+k-1}{n}.$$

Partie (ii). — Il suffit d'écrire

$$(X_1 + \cdots + X_k)^n = \sum_{\{j,\mu\} \in \mathrm{BM}(n,k)} \frac{n!}{(j_1!)^{\mu_1} \cdots (j_s!)^{\mu_s}} \mathrm{SYM}_{\{j,\mu\}} .$$

Partie (iii).

1) Ce résultat se déduit de celui rappelé dans (i) puisque qu'on passe d'une application i des entiers de $[1,k]$ dans les entiers ≥ 0 vérifiant

$\sum_{1\le a\le k} i_a = n$ à une application i des entiers de $[1,k]$ dans les entiers > 0 vérifiant $\sum_{1\le a\le k} i_a = n+k$ en ajoutant l'application constante 1.

2) L'espace des polynômes symétriques en k indéterminées de degré $< n$ par rapport à chacune est engendré par les $\mathrm{SYM}_{\{j,\mu\}}$ quand, pour tout intervalle entier $[1,s]$, j et μ varient indépendamment et respectivement dans l'ensemble de cardinal $\binom{n}{s}$ des applications strictement décroissantes de $[1,s]$ dans les entiers de l'intervalle $[0,n[$ et dans l'ensemble $M(s)$. Cet espace est donc de dimension

$$\sum_{1\le s\le k} \binom{n}{s}\binom{k-1}{s-1} = \binom{n+k-1}{k}$$

(appliquer l'identité $(1+X)^n(1+X)^{k-1} = (1+X)^{n+k-1}$ ou noter l'isomorphisme avec l'espace des polynômes en k indéterminées de degré global $< n$ (de dimension $\binom{n+k-1}{k}$ par le calcul rappelé dans (i)) obtenu en associant au monôme courant $X_1^{i_1}X_2^{i_2}\cdots X_k^{i_k}$ de degré $i_1+\cdots+i_k < n$ le monôme symétrique en $X_1,\ldots,X_k$ construit à partir de $Y_1^{j_1}Y_2^{j_2}\cdots Y_k^{j_k}$ par la substitution $Y_1 = X_1$, $Y_2 = X_1X_2$, $\ldots$, $Y_k = X_1X_2\cdots X_k$ (en termes de suites (d'exposants), on passe d'une suite d'entiers ≥ 0 de somme $< n$ à la suite monotone de ses sommes partielles)).

Partie (iv). — L'espace des polynômes en k indéterminées de degré $< n$ en chacune est engendré par les monômes entrant dans les $\mathrm{SYM}_{\{j,\mu\}}$ lorsque, pour tout intervalle entier $[1,s]$, j et μ varient indépendamment et respectivement dans l'ensemble des applications strictement décroissantes de $[1,s]$ dans les entiers de l'intervalle $[0,n[$ et dans l'ensemble $M(s)$. Par conséquent, la dimension n^k de cet espace est égale à

$$\sum_{1\le s\le k} \binom{n}{s} \sum_{\mu\in M(s)} \frac{k\,!}{\mu_1!\ldots\mu_s!},$$

somme des longueurs des $\mathrm{SYM}_{\{j,\mu\}}$. Si $k > 0$, la condition $s \le k$ est inutile (et demeure la limitation $\binom{n}{s} = 0$ si $s > n$) puisque $a_{s,k} = 0$ si $s > k$ et on peut admettre pour s la valeur 0 en écrivant $a_{0,k} = 0$. Le résultat s'étend enfin au cas $k = 0$ en posant $a_{0,0} = 1$.

La démonstration de la partie 1) est donc claire et analogue aux précédentes. En voici une seconde basée sur l'égalité

$$s^k = \sum_{\mu} \frac{k\,!}{\mu_1!\ldots\mu_s!},$$

μ parcourant les applications de $[1, s]$ dans les entiers ≥ 0 vérifiant la condition $\sum_{1 \leq a \leq s} \mu_a = k$. En groupant les termes du second membre en fonction de $\#\mu^{-1}(0)$, il vient l'expression coïncidant, à des modifications simples près, avec la partie 1) de (iv) lue de droite à gauche...

$$s^k = a_{s,k} + \binom{s}{1} a_{s-1,k} + \binom{s}{2} a_{s-2,k} + \cdots + s a_{1,k}.$$

Reste à calculer

$$\sum_{\mu \in M(s)} \frac{k!}{\mu_1! \ldots \mu_s!} = a_{s,k}.$$

On montre pour cela que les parties 1) et 2) de (iv) sont «inverses» l'une de l'autre (et donc équivalentes). Avec les notations ci-dessus, on passe de la rangée (où $k > 0$)

$$(a_{0,k} = 0, a_{1,k}, \ldots, a_{s,k}) \quad \text{à} \quad (0^k = 0, 1^k, \ldots, s^k)$$

en faisant le produit par la matrice carrée d'ordre $s + 1$ associée au triangle de Pascal (*i.e.* représentative de l'application $P(X) \mapsto P(X + 1)$ dans l'espace des polynômes) et on conclut en multipliant par la matrice inverse.

□

Le corollaire précédent n'intervient pas directement dans la suite de ce travail mais les «formules de Taylor» des paragraphes suivants sont un guide parmi ces identités combinatoires. Par exemple, en lisant (iv-1) comme une identité polynomiale spécialisée en $X = n$, cette formule fournit la «matrice triangulaire infinie» (explicite) de passage entre les éléments $(X^n/n!)$ et $(X(X-1)\cdots(X-(n-1))/n!)$ de NW′ cités au § 1. On verra aux §§ 6 et 12 comment écrire globalement ces formules de passage en termes de séries génératrices, mais on va résumer auparavant dans la proposition indépendante suivante ces propriétés des coefficients $a_{s,k}$, par ailleurs bien connus (voir la remarque 3 ci-après).

PROPOSITION. — *Soit*

$$a_{s,k} = \sum_{\mu \in M(s)} \frac{k!}{\mu_1! \ldots \mu_s!}$$

(où $s > 0$ et $k > 0$), $M(s) = M(s, k)$ étant l'ensemble des applications μ de $[1, s]$ dans les entiers > 0 vérifiant

$$\sum_{1 \leq a \leq s} \mu_a = k$$

(par exemple, $a_{s,k} = 0$ si $s > k$) et on étend la définition à $s \geq 0$, $k \geq 0$ par $a_{0,0} = 1$, $a_{0,k} = 0$ si $k > 0$.

Alors :

(i) *Pour tout* $s \geq 0$, *on a* $\sum_{k\geq 0} a_{s,k} \dfrac{T^k}{k\,!} = (\mathrm{e}^T - 1)^s$.

(ii) *Pour tout* $s > 0$, *on a* $s(a_{s,k} + a_{s-1,k}) = a_{s,k+1}$.

(iii) *Pour tout couple (s,k), on a* $a_{s,k} = \sum_{0\leq i\leq s} (-1)^{s-i} \binom{s}{i} i^k$.

(iv) *Soient* $c_0, \dots, c_s$ *les entiers définis par*

$$(X-1)\cdots(X-s) = \sum_{0\leq i\leq s} c_i X^{s-i}.$$

Alors, pour tout entier $k > 0$, *on a*

$$c_0 a_{s,k+s} + c_1 a_{s,k+s-1} + \cdots + c_s a_{s,k} = 0.$$

(v) *Soient* $(x_{s,k})$ *les suites en* k *définie par les relations (où les* c_i *sont ceux définis ci-dessus)*

$$c_0 x_{s,k+s} + c_1 x_{s,k+s-1} + \cdots + c_s x_{s,k} = 0,$$

avec $x_{s,0} = (-1)^{s+1}$ *et* $x_{s,1} = \cdots = x_{s,s-1} = 0$. *Alors, on a* $a_{s,k} = x_{s,k}$ *pour* $s \geq 0$ *et* $k \geq 1$.

(vi) *Pour tout* $k > 0$, *on a*

$$X^k = \sum_{s\geq 0} a_{s,k} \frac{X(X-1)\cdots(X-(s-1))}{s\,!}.$$

Démonstration. — La formule (i) définissant les $a_{s,k}$ par le biais d'une série génératrice est une conséquence claire de la définition. Un calcul analogue est évidemment possible pour $\sum_{\mu} k\,!/(\mu_1! \dots \mu_s!)$ où μ ne prend que des valeurs au moins égales à un entier donné préalablement.

On obtient (ii) en dérivant par rapport à T les deux membres de l'identité (i) :

$$\sum_{k\geq 0} a_{s,k+1} \frac{T^k}{k\,!} = s(\mathrm{e}^T - 1)^s + s(\mathrm{e}^T - 1)^{s-1}.$$

Le résultat (iii) (*i.e.* le (iv-2) du corollaire précédent) se déduit de (ii) par récurrence sur l'entier $(k+s)$ puisque $s\left(\binom{s}{i}-\binom{s-1}{i}\right)=i\binom{s}{i}$, relation restituant (ii) à partir de (iii).

Quand $k>0$, on a

$$a_{s,k}=\sum_{0\leq i\leq s}(-1)^{s-i}\binom{s}{i}i^k=\sum_{0<i\leq s}(-1)^{s-i}\binom{s}{i}i^k$$

(égalité fausse si $k=0$ puisque le membre de gauche est nul et celui de droite égal à $(-1)^{s+1}$). Mais, la structure du terme $\sum_{0<i\leq s}(-1)^{s-i}\binom{s}{i}i^k$ montre que cette suite en k vérifie une relation de récurrence linéaire d'ordre s dont le polynôme caractéristique est $(X-1)\cdots(X-s)$; les énoncés (iv) et (v) sont alors clairs.

On obtient une variante plus générale de cette preuve en introduisant les polynômes de degré s et multiples de X :

$$A_{s,k}(X)=\sum_{0<i\leq s}(-1)^{s-i}\binom{s}{i}i^kX^i\,;$$

en revenant à la définition des coefficients c_i, on voit que

$$\sum_{0\leq i\leq s}c_iA_{s,k+s-i}=0$$

(les coefficients de $X,X^2,\ldots,X^s$ sont nuls; à noter aussi qu'une relation de dépendance était attendue entre ces $(s+1)$ polynômes); comme $A_{s,k}(1)=a_{s,k}$ pour $k>0$, il vient $\sum_{0\leq i\leq s}c_ia_{s,k+s-i}=0$.

Au prix d'un changement de notations, on voit que les formules de récurrence de (iv) et (v) peuvent être écrites matriciellement sous la forme

$$CA=\mathrm{Id}$$

(définir A (resp. C) comme étant la matrice triangulaire dont le coefficient de $(j+1)$-ième ligne et de la $(i+1)$-ième colonne est $a_{i,j}$ (resp. $c_{i,j}$ où $c_{i,j}$ est défini par $X(X-1)\cdots(X-(j-1))/j\,!=\sum_i c_{i,j}X^i$)) et (vi) se déduit de la relation $A=C^{-1}$.

Remarques.

1) On avait déjà donné trois preuves de la relation (vi) (*cf.* les deux preuves du (iv-1) du corollaire par les arguments de dimension ou les

arguments multinomiaux et l'équivalence entre ce (iv-1) et le (iv-2) dont on vient de donner ci-dessus (*cf.* (iii)) une preuve directe). On en obtiendra une quatrième avec l'identité démontrée dans la suite (et aussi établie par la proposition précédente) :

$$\sum_n \frac{(XT)^n}{n!} = \sum_n \frac{X\cdots(X-(n-1))}{n!}(\mathrm{e}^T-1)^n$$
$$\Big(=\sum_{s,k} a_{s,k}\frac{X\cdots(X-(s-1))}{s!}\frac{T^k}{k!}\Big).$$

2) Les $|j!\,c_{i,j}|$ sont les fonctions symétriques élémentaires de $1,2,\ldots,(j-1)$ dont les valeurs (exemples : $\sum_{1\le i<j} i = j(j-1)/2$, $\sum_{1\le i<i'<j} ii' = (j-2)(j-1)j(3j-1)/4!,\ldots$) s'obtiennent par récurrence plus efficacement qu'en appliquant $\sum_{i,j} c_{i,j}X^iT^j = (1+T)^X$; de même, les $a_{i,j}$:

$$\begin{aligned}
a_{1,j} &= 1 && (j>0),\\
a_{2,j} &= 2^j-2 && (j>1),\\
a_{3,j} &= 3^j-3\cdot 2^j+3 && (j>2),\\
a_{j,j} &= j!,\\
a_{j,j+1} &= (j+1)!(j/2),\\
a_{j,j+2} &= j(3j+1)((j+2)!/4!)
\end{aligned}$$

par exemple se déduisent de (ii) ou (v) plus aisément que de

$$\sum_{s,k} a_{s,k}\frac{X^sT^k}{s!\,k!} = \mathrm{e}^{X(\mathrm{e}^T-1)}.$$

3) Soient μ un élément de $M(s,k)$ et $\{\sigma(1),\ldots,\sigma(k)\}$ un élément du groupe symétrique $\mathfrak{S}_k$. En partageant cet ensemble ordonné de k éléments en les s parties formées en considérant les μ_1 premiers puis les μ_2 suivants et ainsi de suite jusqu'aux μ_s derniers, on définit sur $\mathfrak{S}_k$ une application montrant qu'à μ est attaché le nombre $k!/(\mu_1!\ldots\mu_s!\,s!)$, égal au nombre de partitions d'un ensemble de cardinal k en s sous-ensembles de cardinaux respectifs $\mu_1,\ldots,\mu_s$. Il en résulte que $a_{s,k}$ est un multiple de $s!$, ce qu'on peut aisément vérifier par récurrence en utilisant les parties (ii) ou (iv) de la proposition précédente. Les nombres $(a_{s,k}/s!)$ qu'on définit le plus souvent grâce à (vi) sont donc des entiers et sont appelés *nombres de Stirling de deuxième espèce.* Ils jouissent de nombreuses autres propriétés pour lesquelles on renvoie au classique *Analyse combinatoire* de L. Comtet et à sa bibliographie.

4. Vision exponentielle (K corps commutatif de caractéristique 0).

THÉORÈME 2. — *On suppose K de caractéristique* 0; *la dérivée de $P \in K[X]$ est notée* $\mathrm{d}P$ *ou P'.*

(i) *Soit $(H_n(X))$ une suite d'éléments de $K[X]$ vérifiant (μ) et $H_0 \neq 0$. La série formelle*

$$B(T) = \sum_{n\geq 0} H_n'(0)T^n = \sum_{n>0} H_n'(0)T^n \in K[[T]]$$

caractérise $(H_n(X))$ par l'identité

$$\exp\big(B(T)X\big) = \sum_{n\geq 0} \big(B(T)\big)^n \frac{X^n}{n!} = \sum_{n\geq 0} H_n(X)T^n. \tag{1}$$

Réciproquement, toute série $B(T) = \sum_{n>0} b_n T^n$ définit par (1) *un élément $(H_n(X)) \neq 0$ de* NW.

(ii) *Pour toute solution $(H_n(X))$, il existe un élément $x \neq 0$ de K, un entier $k \geq 0$ et une solution $(K_n(X))$ vérifiant $K_1 = X$ et*

$$\sum_{n\geq 0} H_n(X)T^n = \sum_{n\geq 0} K_n(xXT^k)T^n.$$

Démonstration.

Partie (i). — Soit $B(T) = \sum_{i\geq 0} b_i X^i$ un élément de $K[[T]]$ avec $b_0 = 0$. Il est clair à partir des propriétés fonctionnelles de l'exponentielle que la formule (1) définit un élément $(H_n(X))$ de NW.

Réciproquement, soient un élément $(H_n(X)) \neq 0$ de NW et $H(X,T) = H_n(X)T^n$ la série génératrice associée. En dérivant l'identité (μ'), il vient

$$\frac{\partial H(X+Y,T)}{\partial Y} = H(X,T)\frac{\partial H(Y,T)}{\partial Y}$$

et, en spécialisant en $Y = 0$, on voit que $H(X,T)$ est proportionnel à $\exp\big(X\big(\sum_{n\geq 0} H_n'(0)T^n\big)\big)$, le facteur multiplicatif associé à cette relation appartenant à $K[[T]]$, et en fait à K comme le montre (μ); on conclut en notant que les deux termes de la relation coïncident pour $X = 0$ et sont donc égaux.

Une variante de cette réciproque est la suivante. On forme

$$\frac{\partial^k H_n(X_1 + \cdots + X_k)}{\partial X_1 \cdots \partial X_k},$$

puis on spécialise en 0 les indéterminées $X_1, \ldots, X_k$; il vient grâce à (μ) et au théorème 1 :

$$\sum_{\alpha_1 + \cdots + \alpha_k = n} \mathrm{d}H_{\alpha_1}(0) \cdots \mathrm{d}H_{\alpha_k}(0) = H_n^{(k)}(0)\,;$$

on reconstitue ensuite $H_n(X)$ par la formule de Taylor (classique) et, en posant $B(T) = \sum\limits_n \mathrm{d}H_n(0)T^n$, on obtient bien l'égalité (1) annoncée.

Partie (ii). — À la suite $(H_n(X)) \in \mathrm{NW}$, on associe par (1) une série $B(T)$ — qu'on peut écrire $(xT^kC(T))$ (avec $k \geq 0$, $x \in K$, $x \neq 0$, $C(0) = 0$ et $C'(0) = 1$) — ce qui permet d'introduire $(K_n(X)) \in \mathrm{NW}'$ par $K(X,T) = \sum\limits_n K_n(X)T^n = \exp(XC(T))$ et d'établir l'identité à démontrer. □

Remarques.

- On peut retrouver le (iv) du théorème 1 à partir du théorème 2.
- Les ensembles NW et $(K^* \times N \times \mathrm{NW}')$ sont équipotents (en caractéristique 0). En tant que groupe multiplicatif, $\mathrm{NW} - \{0\}$ est isomorphe au groupe additif $K[[T]]$ comme annoncé.
- L'auteur remercie J.-P. Bézivin de sa contribution pour les nombreuses preuves possibles du théorème 2.

Le théorème 2 (ii) et le théorème 1 (iii) et (v) permettent de ramener l'étude de NW (quand est écarté le cas trivial) au sous-ensemble NW' des suites $(H_n(X))$ vérifiant $H_0 = 1$ et $H_1 = X$. C'est cet ensemble NW' qui est intéressant et qu'on étudie dans la suite de ce travail. Le théorème 2 (i) établit une bijection, notée w, entre GSF et NW', laquelle appartient en fait à un ensemble de bijections paramétré par GSF. Ainsi :

COROLLAIRE.

(i) *Soient* $B(T) \in \mathrm{GSF}$ *et* (H_n) *la suite polynomiale définie par* (1); *alors, la correspondance* w *ainsi introduite de* GSF *sur* NW' *est bijective.*

(ii) *Soient* $F(X,T)$, $G(X,T)$ *deux éléments de* NW' ; *il existe un élément unique* $B \in \mathrm{GSF}$ *vérifiant* $F(X, B(T)) = G(X,T)$.

5. Vision différentielle générale (K corps commutatif quelconque).

Notations. — On note T_1 l'ensemble des endomorphismes f de $K[X]$ de noyau K vérifiant

$$f(X) = 1 \quad \text{et} \quad \deg\big(f(P)\big) = \deg(P) - 1$$

pour tout polynôme P de degré positif («deg» désigne la fonction degré, la condition $f(X) = 1$ est une convention de normalisation, le point important est $f(X) \neq 0$). Par exemple, en caractéristique 0, la dérivation d (et donc $B(\mathrm{d})$, où $B \in \mathrm{GSF}$) est élément de T_1.

On note J l'application associant à $f \in T_1$ la base $(F_n(X))$ de $K[X]$ définie par récurrence par :

$$(\mathrm{J}'') \qquad F_0 = 1, \quad f(F_n) = F_{n-1}, \quad F_n(0) = 0 \quad \text{pour} \quad n > 0.$$

Par J, l'ensemble T_1 correspond bijectivement aux bases $(F_n)_{n \geq 0}$ de $K[X]$ vérifiant :

$$(\mathrm{J}') \qquad F_0 = 1, \quad F_1 = X, \quad \deg(F_n) = n, \quad F_n(0) = 0 \quad \text{pour} \quad n > 0.$$

Soient $f \in T_1$, $(F_n) = J(f)$ et $F(X,T) = \sum_n F_n(X)T^n$; on peut prolonger f à $K[X][[T]]$ pour obtenir la relation :

$$f\big(F(X,T)\big) = TF(X,T)$$

(par exemple, $f = \mathrm{d} = \partial/\partial X$, $F_n(X) = X^n/n!$ et $F(X,T) = \exp(XT)$).

Quand $(F_n) = J(f)$ appartient à NW, (F_n) appartient à NW$'$ et la caractéristique de K est 0. Soit Y une nouvelle indéterminée; pour tout élément $f \in T_1$, soient f^* le prolongement $K[Y]$-linéaire de f à $K[Y][X]$ et γ l'application de $K[X]$ dans $K[X,Y]$ définie par

$$\gamma\big(P(X)\big) = P(X+Y).$$

On note T_2 la partie de T_1 formée des éléments f vérifiant :

$$(\Gamma) \qquad \gamma \circ f = f^* \circ \gamma.$$

Exemples. — En caractéristique 0, la dérivation d et δ_x (définie par $\delta_x(P) = P(X+x) - P(X)$ où $x \in K$) appartiennent à T_2 mais Δ, définie par $\Delta(P) = (P(X) - P(0))/X$, est un élément de $(T_1 - T_2)$.

Remarque. — Quand f vérifie (Γ), on peut spécialiser X en 0 dans $f^* \circ \gamma(P(X))$ et donner le sens $f(P)(Y)$ à cette expression.

Soit $B(T)$ un élément de $K[[T]]$; la spécialisation de T en un point quelconque f de T_1 permet de définir un endomorphisme$B(f)$. En se restreignant à GSF, on obtient une action de ce groupe dans T_1. Soit $f^{\#}$ l'application de GSF dans T_1 associant l'endomorphisme

$$B(f) = f^{\#}(B)$$

de $K[X]$ à $B \in$ GSF. La partie T_2 est clairement stable dans cette action de groupe.

THÉORÈME 3 (*cf.* §7 pour la démonstration).

(i) *Tout endomorphisme g du K-espace $K[X]$ commutant avec un élément f de T_1 s'écrit*

$$g = \sum_n g(F_n)(0) f^n \, ; \tag{S}$$

(ii) *Dans T_1, la relation «f et g commutent» est une relation d'équivalence (si f est dans T_1, l'application $f^{\#} : B \mapsto B(f)$ est une bijection de* GSF *sur la classe de f).*

(iii) *Soit f un élément de T_1 ; les conditions suivantes sont équivalentes :*

$$\gamma \circ f = f^* \circ \gamma \quad (i.e.\ f \in T_2), \tag{Γ}$$

$$\gamma = F(Y, f), \tag{Γ'}$$

$$F(X+Y, T) = F(X, T) F(Y, T). \tag{μ'}$$

En particulier, si K est de caractéristique 0, *J établit une bijection entre T_2 et* NW′ *(et T_2 est vide en caractéristique* > 0 *d'après le* (v) *du théorème* 1).

(iv) *Si f, g sont deux éléments de T_1, deux des hypothèses suivantes impliquent la troisième :*

$$f^* \circ \gamma = \gamma \circ f, \quad g^* \circ \gamma = \gamma \circ g, \quad f \circ g = g \circ f.$$

(v) *Si K est de caractéristique nulle, T_2 est la classe de* d *dans T_1 pour la relation de commutation et l'orbite de* d *dans l'action de* GSF *sur T_1.*

(vi) *Si K est de caractéristique* 0, *la classe de* d *dans* $\mathrm{End}(K[X])$ *pour la relation de commutation est formée des opérateurs différentiels polynomiaux $B(\mathrm{d})$ (l'ordre de cet opérateur coïncidant avec la valuation de la série formelle B) ; par suite, si $f \in \mathrm{End}(K[X])$ vérifie $f(1) = 0$, $f(X) = 1$ et commute avec* d *ou γ, alors f appartient à T_2 (et réciproquement).*

Exemples (cf. notations précédentes).

$$J(\mathrm{d}) = \frac{X^n}{n\,!}, \quad J(\delta_x) = \frac{X(X-x)\cdots(X-(n-1)x)}{x^n\, n\,!},$$
$$x\,\mathrm{d} = \sum_{n>0} \frac{(-1)^{n-1}}{n}(\delta_x)^n, \quad \delta_x = \sum_{n>0} x^n \frac{\mathrm{d}^n}{n\,!}.$$

Nota Bene. — C'est (sans ce vocabulaire) par

$$\delta_Y = \sum_{n>0} Y^n \frac{\mathrm{d}^n}{n\,!}$$

qu'on définit d (Bourbaki, *Algèbre*, chap. 4). L'exemple ci-dessus repris dans le *Nota Bene* n'est rien d'autre que la classique formule de Taylor. Plus généralement, le théorème suivant (où la partie (iii) se déduit du théorème 3 (iv)) montre qu'à tout élément de T_1 (resp. T_2) est associée une formule de Taylor généralisée à une (resp. deux) variables. On met ainsi en évidence la différence de nature entre ces formules.

THÉORÈME 4 (*cf.* §7 pour la démonstration).

(i) *Soient $f \in T_1$ et $(F_n) = J(f)$. Alors :*

$$(\Theta_1) \qquad P(X) = \sum_n f^n(P)(0)F_n(X)$$

est vérifiée pour tout $P \in K[X]$;

(ii) *Soient $f \in T_2$ et $(F_n) = J(f)$. Alors, pour tout $P \in K[X]$,*

$$(\Theta_2) \qquad P(X+Y) = \sum_{i\geq 0} F_i(Y)f^i(P)(X) = \sum_{i\geq 0} F_i(X)f^i(P)(Y)$$

(iii) *T_2 est l'ensemble des éléments de T_1 vérifiant (Θ_2) pour tout $P \in K[X]$;*

(vi) *les relations* : (Γ), (Γ'), (Θ_2) *et* (μ') *sont équivalentes pour tout élément de T_1.*

Remarques.

1) Soient $f \in T_1$ et $(F_n) = J(f)$; la formule (Θ_1) permet de résoudre par décalage l'équation $P = f(Q)$ où le polynôme P est donné et le polynôme Q inconnu (la difficulté pratique se trouve en général dans le calcul

matriciel de (F_n), le cas favorable pour cela est celui où $f \in T_2$ et où la série B vérifiant $\mathrm{d} = B(f)$ est connue puisqu'alors $\sum_n F_n(X)T^n = \exp(XB(T)))$. Une application classique de cette méthode est le calcul de $\sum_{0 \leq i \leq n} P(X+ix)$ grâce à l'opérateur δ_x ; par exemple, si $P(X) = X^k$, le polynôme

$$Q(X) = \sum_s a_{s,k} \frac{X(X-1)\cdots(X-s)}{(s+1)!}$$

(*cf.* §3) vérifie $Q(X+1) - Q(X) = X^k$, et on retrouve les habituelles sommes des puissances k-ièmes d'un ensemble d'entiers consécutifs.

2) Si $f \in T_1$, alors $\sum_{0 \leq s \leq n} \binom{n}{s} (f - \mathrm{Id})^s = f^n$ annule tout polynôme de degré au plus $(n-1)$.

3) Le résultat principal est clairement la réduction à l'hypothèse (Γ) pour obtenir (Θ_2), laquelle s'écrit encore $\gamma = F(Y, f)$ et généralise la formule de Taylor pour les polynômes : $\gamma = \exp(Y\mathrm{d})$. En d'autres termes, un endomorphisme f du K-espace $K[X]$ vérifiant

$$f^*\big(P(X+Y)\big) = \big(f(P)\big)(X+Y)$$

est de la forme $B(\mathrm{d})$ avec

$$B(T) = \sum_n \Big[f\Big(\frac{X^n}{n!}\Big)\Big]_{X=0} T^n.$$

Cette dernière identité est issue de (S) (*cf.* théorème 3) mais on peut la déduire de la formule de Taylor classique

$$P(X+Y) = \sum_n \frac{X^n}{n!} \mathrm{d}^n P(Y).$$

En effet, l'application de f conduit, en utilisant (Γ) et en spécialisant en $X = 0$, à :

$$f(P(Y)) = \sum_n \Big[f\Big(\frac{X^n}{n!}\Big)\Big]_{X=0} \mathrm{d}^n P(Y).$$

Enfin, de l'identité

$$\gamma = \exp(Y\mathrm{d}) = \sum_n F_n(Y) f^n$$

(pour tout élément f de T_2), on déduit aussitôt

$$\mathrm{d} = \sum_n F_n'(0) f^n$$

(*cf.* théorème 2), identité d'ailleurs incluse dans

$$Y\mathrm{d} = \sum_{i>0} \frac{(-1)^{i-1}}{i} \Big(\sum_{n>0} F_n(Y) f^n\Big)^i.$$

Pour établir ces résultats, on va dégager quelques lemmes d'algèbre linéaire dans le § 6 autonome suivant. L'utilisation directe des arguments du § 6 dans le contexte est laissée aux soins du lecteur.

6. Bases de Jordan et commutation (K corps commutatif quelconque).

Tous les résultats qui suivent ont des traductions pratiques matricielles ou algorithmiques ainsi que les applications mettant en jeu le groupe GSF. Comme déjà signalé, on renvoie pour cela à [H] (voir notamment dans cette référence les théorèmes de Schur-Jabotinski et de Lagrange-Bürmann).

Soient K un corps et E un K-espace vectoriel (à gauche) de dimension infinie dénombrable. On abrège dimension en « dim ». Si un endomorphisme f de E vérifie

$$E = \bigcup_n \operatorname{Ker} f^n \quad \text{et} \quad \dim(\operatorname{Ker} f) = 1,$$

alors $\operatorname{Ker} f^n$ est de dimension n et f est surjective (pour tout couple d'endomorphismes (g_1, g_2) d'un module, $\operatorname{Ker}(g_2 \circ g_1)/\operatorname{Ker} g_1$ s'injecte dans $\operatorname{Ker} g_2$, d'où l'inégalité $\dim(\operatorname{Ker} f^n) \leq n \dim(\operatorname{Ker} f)$ devenant une égalité lorsque la suite $(\operatorname{Ker} f^n)$ est non stationnaire).

On peut donc associer dans ce cas à tout élément x de E un degré noté $\deg_f(x)$, égal au plus petit entier m vérifiant $f^{m+1}(x) = 0$.

De plus, si φ est une forme linéaire de noyau égal à un supplémentaire de $\operatorname{Ker} f$, on associe à f une base (e_n) « de Jordan » construite par récurrence par l'ensemble de relations $(j_\varphi) = (j)$:

$$(j) \qquad \begin{cases} \varphi(e_0) = 1, & f(e_0) = 0, \\ f(e_n) = e_{n-1}, & \varphi(e_n) = 0 \quad (n > 0). \end{cases}$$

Avec ces notations — pour lesquelles φ est la forme duale e_0^* — on va établir par récurrence des identités en montrant que la différence entre les deux membres est nulle car annulée à la fois par f (ceci grâce à l'hypothèse de récurrence) et par φ (ceci par un calcul immédiat).

Ainsi, dans cette base (e_n), une récurrence sur $\deg_f(x)$ montre que le vecteur courant x de E s'écrit

$$(\Theta_1') \qquad x = \sum_n \varphi\big(f^n(x)\big)e_n\,;$$

la formule (Θ_1') représente une « formule de Taylor » (à une indéterminée) pour le couple (f, φ).

Réciproquement, si un espace vectoriel admet une base dénombrable (e_n), la formule (j), où φ est la forme duale e_0^*, définit un endomorphisme f satisfaisant aux hypothèses ci-dessus et restituant la base (e_n) donnée par la construction précédente. On peut donc ainsi, dans un espace de dimension dénombrable, attacher une « formule de Taylor » à un couple (f, φ) ou à une base (e_n).

L'hypothèse $E = \bigcup_n \operatorname{Ker} f^n$ montre que, pour toute série formelle

$$B(T) = \Big(\sum_{i\geq 0} b_i T^i\Big) \in K[[T]],$$

un endomorphisme $\sum_n b_n f^n$ est univoquement défini, commutant avec f, d'où une structure de $K[[T]]$-module sur E ; on va montrer que les endomorphismes de ce module sont les homothéties.

LEMME 2. — *Soit E un K-espace vectoriel de dimension infinie dénombrable. Pour tout endomorphisme f de E vérifiant* $\dim(\operatorname{Ker} f) = 1$ *et $E = \bigcup_n \operatorname{Ker} f^n$, le centralisateur de f dans $\operatorname{End}_K E$ est formé par les éléments de la forme*

$$g = B(f) = \sum_n b_n f^n.$$

L'application $B \mapsto B(f)$ de $K[[T]]$ sur le centralisateur de f est bijective et l'hypothèse $f \circ g = g \circ f$ (où $g \in \operatorname{End}_K(E)$) implique, en associant à f une base (e_n) par (j_φ)

$$g = \sum_n \varphi\big(g(e_n)\big)f^n$$

(soit $k = \inf\{n\,;\ \varphi(g(e_n)) \neq 0\}$, g est injective si $k = 0$ et $\operatorname{Ker} g$ a pour base $\{e_0, \dots, e_{k-1}\}$ si $k > 0$).

Si $B(T) = \sum_{n>0} b_n f^n$ *appartient à* GSF, *on a* $\operatorname{Ker} f^n = \operatorname{Ker}(B(f))^n$ *pour* $n \geq 0$ *et la base* (e'_n) *associée à* $B(f)$ *par* (j_φ) *est liée à* (e_n) *par la relation de passage formelle :*

$$\sum_n e_n T^n = \sum_n e'_n \big(B(T)\big)^n.$$

Démonstration. — Si $g \in \operatorname{End}_K(E)$ vérifie $f \circ g = g \circ f$, alors, suivant l'idée générale exposée dans l'introduction de ce paragraphe,

$$\Big(g - \sum_n \varphi\big(g(e_n)\big) f^n\Big)(e_k)$$

est, pour tout entier k, annulé par φ (vérification facile) et f (par récurrence sur k). La suite de l'énoncé est alors de vérification facile. □

Remarque. — L'exemple du couple d'endomorphismes $(\partial/\partial X, \partial/\partial Y)$ de $E = K[X, Y]$ commutant entre eux montre l'importance de l'hypothèse sur la dimension du noyau.

COROLLAIRE. — *Soit* TE_1 *l'ensemble des endomorphismes* f *de l'espace* E *vérifiant* $\dim(\operatorname{Ker} f) = 1$ *et* $E = \bigcup_n \operatorname{Ker} f^n$. *Dans* TE_1, *la relation de commutation est une relation d'équivalence et* GSF *opère librement et transitivement dans chacune des classes associées.*

Remarque. — Le lemme 2 et son corollaire sont des exemples d'extensions en dimension infinie du résultat classique (dont la réciproque est exacte) suivant :

Soit f *un endomorphisme d'un espace vectoriel* E *de dimension finie sur un corps* K *commutatif. Si le polynôme minimal et le polynôme caractéristique de* f *coïncident (autrement dit, si* f *est représentée par une matrice de Jordan, celle-ci ne comporte que des* 1 *le long de la diagonale), seuls les endomorphismes de* E *de la forme* $P(f)$ *(avec* $P \in K[X]$*) commutent avec* f.

On suppose maintenant K commutatif et soient K' une K-algèbre, f^*, φ^* les prolongements de f, φ par extension des scalaires. Dans ce contexte, s'établit comme le lemme 2 le résultat suivant :

LEMME 3 (notations ci-dessus). — *Soient ξ un morphisme de E dans son extension par K' et $f \in \mathrm{End}_K E$; alors, une condition nécessaire et suffisante pour que f vérifie $\xi \circ f = f^* \circ \xi$ est :*

$$(\Theta'_2) \qquad \xi = \sum_n \varphi^*\big(\xi(e_n)\big) f^n.$$

En particulier, si $\varphi^*(\xi(e_1)) \neq 0$ et si K' est un surcorps de K, tout endomorphisme g de E vérifiant $\xi \circ g = g^* \circ \xi$ commute avec f.

7. Preuves des résultats du paragraphe 5.

On applique les résultats du §6 avec $E = K[X]$; dans ce cas, les éléments de T_1 (*cf.* §5) satisfont aux hypothèses du lemme 3 et le choix de $\varphi : P(X) \mapsto P(0)$ montre que l'application J est celle par laquelle on a construit la base «de Jordan» (e_n) du §6. L'identité (Θ'_1) appliquée au cas présent établit la relation (Θ_1) : $P(X) = \sum_n f^n(P(0))F_n(X)$ du théorème 4 (i). Le lemme 2 montre que, si un endomorphisme g commute avec un élément f de T_1, il existe $B(T) = \sum_{i \geq 0} b_i T^i \in K[[T]]$ unique vérifiant $g = B(f)$ (avec B dans GSF si g appartient à T_1); en explicitant B, on obtient (S) (*cf.* théorème 3 (i), §5); de plus, la relation de commutation dans T_1 est transitive (*cf.* théorème 3 (ii) et corollaire, §6).

Soit $f \in T_2$; l'identité (Θ'_2), appliquée au cas présent ($E = K[X]$, $\varphi : P(X) \mapsto P(0)$, $\gamma = \xi$) s'écrit, pour $P \in K[X]$:

$$P(X+Y) = \sum_{i \geq 0} F_i(Y) f^i(P)(X)$$

($= \sum_{i \geq 0} F_i(X) f^i(P)(Y)$ en intervertissant X et Y), *i.e.* (Θ_2) ou (Γ'), ou encore (μ') en choisissant $P = F_n$.

Par ailleurs, comme (F_n) est une base, il suffit que la relation

$$P(X+Y) = \sum_{i \geq 0} F_i(Y) f^i(P)(X)$$

soit vérifiée lorsque $P = F_n$ (avec $n \geq 0$) pour qu'elle le soit pour tout P et on voit en fait que (Γ') équivaut à (μ').

Cette dernière équivalence est aussi claire en termes de séries génératrices ; on observe en effet que l'on a

$$f\big(F(X,T)\big) = TF(X,T)$$

(d'où $F(Y,f)(F(X,T)) = F(Y,T)F(X,T)$) et

$$\gamma\big(F(X,T)\big) = F(X+Y,T).$$

Quant à l'implication $(\Gamma') \Rightarrow (\Gamma)$, elle est évidente.

Les preuves du théorème 3 (iii) et du théorème 4 sont ainsi achevées et il suffit de montrer que deux éléments f et g de T_2 commutent pour terminer celle de (iv) (les énoncés (v), (vi) deviennent alors clairs puisque d appartient à T_2). Pour cela, on applique la fin du lemme 3. Précisément, comme $\varphi^*(\xi(e_1)) = Y$, en passant au corps $K(Y)$, on peut inverser l'écriture de $\gamma - \mathrm{Id}$ (correspondant à $\xi - \mathrm{Id}$) en fonction de f en une écriture de f en fonction de $\gamma - \mathrm{Id}$, ce qui prouve que la commutation de g avec γ implique sa commutation avec f ; tout cela s'écrit explicitement puisque $\gamma - \mathrm{Id} = \delta_Y$.

On peut aussi donner une démonstration de ce dernier point utilisant l'argument de la remarque terminant le §5 : on écrit (Θ_2) pour f et on intervertit les indéterminées X et Y ; en appliquant g^*, il vient :

$$g^*\big(P(X+Y)\big) = \sum_n g\big(F_n(X)\big) f^n(P)(Y).$$

Grâce à l'hypothèse $g^* \circ \gamma = \gamma \circ g$, on voit que le premier membre est $g(P)(X+Y)$, et on peut maintenant spécialiser en $X = 0$ afin de retrouver l'égalité attendue $g = \sum_n g(F_n(0)) f^n$.

8. Lien entre les visions exponentielle et différentielle (K corps commutatif de caractéristique 0).

On rappelle dans l'énoncé ci-dessous les notations pour décrire de façon plus autonome ce lien :

THÉORÈME 5. — *Soient K un corps commutatif de caractéristique nulle,* NW′ *l'ensemble des suites $(H_n(X))$ vérifiant $H_0 = 1$, $H_1 = X$ et la condition (μ). Soit* GSF *le groupe (pour la composition) des séries formelles*

$B(T) = \sum_{i>0} b_i T^i$, *avec* $b_1 = 1$ *et à coefficients dans* K. *Soit* $B(T) \in$ GSF ; *pour tout* $n \geq 0$, *les suites* (H_n) *et* (K_n) *définies par* 1) *et* 2) *sont égales :*

1) $$\exp\big(B(T)X\big) = \sum_{n\geq 0}\big(B(T)\big)^n \frac{X^n}{n!} = \sum_{n\geq 0} H_n(X)T^n\,;$$

2) $$K_0 = 1,\quad B^{-1}(\mathrm{d})(K_n) = K_{n-1} \ \text{ et } \ K_n(0) = 0 \text{ pour } n > 0.$$

L'élément obtenu appartient à NW$'$ *et les applications de* GSF *dans* NW$'$ *ainsi définies sont bijectives.*

COROLLAIRE 1. — *Soient* K *un corps commutatif de caractéristique nulle et* f *un endomorphisme du* K*-espace* $K[X]$ *vérifiant*

$$\operatorname{Ker} f = K,\quad f(X) = 1,\quad \deg\big(f(P)\big) = \deg(P) - 1$$

pour tout P *de degré non nul. Alors*

$$\exp\Big(X\Big(\sum_n \Big[f\Big(\frac{X^n}{n!}\Big)\Big]_{X=0} T^n\Big)^{-1}\Big)$$

s'écrit sous la forme $\sum_n F_n(X)T^n$ *où* (F_n) *est une base de* $K[X]$ *vérifiant*

$$F_0 = 1,\quad f(F_n) = F_{n-1} \quad \text{et} \quad F_n(0) = 0 \text{ pour } n > 0$$

(*i.e. la formule* (J'') *du* § 5).

COROLLAIRE 2 (notations du § 5). — *Si* $J(f) = F(X,T)$ *et* $J(g) = G(X,T)$ *appartiennent à* NW$'$, *il existe* $B \in$ GSF *pour laquelle il est équivalent d'écrire* $G(X,T) = F(X, B(T))$ *ou* $f = B(g)$.

Démonstration du théorème 5 et de ses corollaires. — Elle s'écrit plus généralement en termes d'action de groupe ainsi : on identifie comme ci-dessus NW$'$ et T_2 par J^{-1}. Soient $B(T) \in$ GSF et $J(f) = F(X,T) \in$ NW$'$; à ce couple, on peut associer naturellement les éléments de NW$'$:

- $F(X, B(T))$ (opération à droite de GSF dans NW$'$) ;
- $B(f)$ (opération à gauche de GSF dans NW$'$).

De $f(F(X,T)) = TF(X,T)$, on déduit les deux relations :

$$B(f)\big(F(X,T)\big) = B(T)F(X,T),$$

$$f\big(F(X, B(T)\big) = B(T)\big(F(X, B(T)\big).$$

Soient maintenant

$$G(X,T) = F\big(X, B(T)\big) \quad \text{et} \quad g = J^{-1}\big(G(X,T)\big).$$

La seconde relation ci-dessus s'écrit $f(G(X,T)) = B(T)G(X,T)$ qui est égale à $B(g)(G(X,T))$ d'après la première relation, d'où $f = B(g)$. □

Le théorème 5 où $F(X,T) = \exp(XT) = J(\mathrm{d})$ est donc démontré avec ses corollaires puisque les autres assertions ont été établies antérieurement.

On aurait pu déduire aussi cette démonstration de la formule

$$B(T) = \sum_{n \geq 0} H'_n(0) T^n$$

(*cf.* §1) et (S) (*cf.* théorème 3 (i) avec $g = \mathrm{d}$). Mais il est plus important d'unifier les points de vue des §§4 et 5 en ne considérant que l'action (à gauche) associant l'élément $F(X, B^{-1}(T)) = J(B(f))$ de NW$'$ à l'élément $(B, F(X,T) = J(f))$ de GSF $\times$ NW$'$; de plus, on ne privilégie aucune application, et notamment pas l'exponentielle comme usuellement dans la littérature.

Ce qui précède éclaire le §9 étudiant les possibles structures affines et de groupes sur NW$'$.

9. Stuctures affines et structures de groupes sur NW$'$.

Les résultats précédents permettent d'identifier NW$'$ avec GSF et T_2 qui sont naturellement des espaces affines. Mais, la multiplicité des structures affines n'est qu'apparente puisque GSF opère librement et transitivement sur ces deux structures. En fait, « affinement parlant », l'identification « exponentielle » de NW$'$ avec GSF n'est pas « canonique » (d'où une infinité de structures affines sur NW$'$, paramétrées par NW$'$ ou GSF), contrairement à celle, canonique au sens général du terme, avec T_2 par J^{-1} (d'où une structure affine de référence sur NW$'$ non modifiable par le jeu des identifications (affines et paramétrées par T_2) de T_2 avec GSF).

Ce paradoxe, alors que les deux visions du problème semblent équivalentes puisque se ramenant au passage à l'inverse dans le groupe GSF, est issu du fait simple (A) énoncé dans l'introduction (§1). Précisément, l'identification de l'élément $H(X,T)$ de NW$'$ (écrit sous forme de série génératrice) avec l'élément B de GSF suppose le choix préalable dans NW$'$ d'un élément de référence $H_0(X,T)$ pour écrire la relation $H(X,T) = H_0(X,B(T))$ ($H_0(X,T)$ étant $\exp(XT)$ dans le cas de la bijection w introduite au §4 et dans le corollaire du théorème 2) ; cette dernière relation montre que le changement d'élément de référence revient à effectuer une translation à *gauche* dans le groupe GSF.

À l'opposé, quel que soit l'élément de référence $f_0 \in T_2$ choisi pour identifier GSF à T_2 par $f_0^{\#}$, cette dernière identification $f_0^{\#}$ est un isomorphisme affine de GSF sur T_2 et le changement d'élément de référence se traduit pour GSF par une translation à *droite.*

Dans le théorème 6 suivant,

• le (i) — rattaché au théorème 3 — décrit dans NW$'$ la construction barycentrique « canonique » ;

• le (ii) (rattaché au corollaire 1) décrit la construction barycentrique, de forme agréable mais pas vraiment canonique, liée à l'identification de NW$'$ à GSF par w ;

• le (iii) décrit les constructions barycentriques de type « exponentiel » paramétrées par NW$'$ (ou GSF...) ;

• enfin, le (iv) rappelle comment « bien » écrire l'action de GSF dans l'ensemble NW$'$ (*i.e.* pour que cette action induise une famille d'identifications affines de GSF avec NW$'$) et permet de relier (i), (ii) et (iii).

THÉORÈME 6. — *Soient* (F_n) (*ou* $F(X,T) = \sum_n F_n(X)T^n$) *et* (G_n) (*ou* $G(X,T) = \sum_n G_n(X)T^n$) *deux éléments de* NW$'$ *et* $\lambda \in K$, *alors* :

(i) $J\big(\lambda J^{-1}(F_n(X)) + (1-\lambda)J^{-1}(G_n(X))\big)$
appartient à NW$'$ (*barycentre canonique*) ;

(ii) $\big(L_n(X) = \sum_{0\le i\le n} F_i(\lambda X)G_{n-i}((1-\lambda)X)\big)$
(*ou encore* $L(X,T) = F(\lambda X,T)G((1-\lambda)X,T)$)
appartient à NW$'$ (*barycentre* « *exponentiel* » *lié à* w) ;

(iii) *pour tout* $H_0(X,T) \in$ NW$'$, $L(X,T) = H_0(X, \lambda B(T) + (1-\lambda)C(T))$ *avec* $(B,C) \in$ GSF^2 *défini par* $H_0(X,B(T)) = F(X,T)$ *et* $H_0(X,C(T)) = G(X,T)$ (*si* $H_0(X,T) = \exp(XT)$, *on retrouve* (ii)) *est un barycentre des éléments* $F(X,T)$ *et* $G(X,T)$ *affectés des masses* λ *et* $(1-\lambda)$;

(iv) *on identifie encore les ensembles*

- NW$'$,
- *celui des séries* $\big(1 + XT + \sum_{n>1} H_n(X)T^n\big)$ *vérifiant* (μ'),
- T_2 (*par* J^{-1}).

L'application associant l'élément $H(X, B^{-1}(T)) = J(B(h))$ *de* NW$'$ *à l'élément* $(B, H(X,T) = J(h))$ *de* GSF $\times$ NW$'$ *est une action libre et transitive dans* NW$'$ *du groupe* GSF.

En identifiant ainsi NW$'$ et GSF par le choix d'un élément de référence $H_0(X,T)$ de NW$'$, la structure affine sur NW$'$ déduite de celle de GSF est indépendante du référentiel $H_0(X,T)$ et coïncide avec celle donnée par (i). Par exemple, sous forme de série, le barycentre du (i) s'écrit :

$$\exp\Big(X\Big(\sum_n \Big[(\lambda f + (1-\lambda)g)\Big(\frac{X^n}{n!}\Big)\Big]_{X=0} T^n\Big)^{-1}\Big)$$

avec $J(f) = (F_n)$ et $J(g) = (G_n)$.

Aux notations près (*cf.* §8), le corollaire suivant restitue l'identité

$$F(X,T) = \sum_n H_{0,n}(X)\Big(\sum_n h_0(F_m)(0)T^m\Big)^n$$

et résume les liens entre ces structures affines et celles de groupe déduites de GSF.

COROLLAIRE. — *Soit* $h_0 \in T_2$. *L'application associant au couple* $((F_n),(G_n))$ *d'éléments de* NW$'$ *l'élément* $\sum_n F_n(X)(\sum_m h_0(G_m)(0)T^m)^n$ *munit* NW$'$ *d'une loi* $(*)$ *de groupe d'élément neutre* $J(h_0)$ *et isomorphe à* GSF. *En notant* $+_{\mathrm{exp}}$ (*resp.* $+_{\mathrm{can}}$) *les lois de composition barycentrique définies dans* (ii) (*resp.* (i) *ou* (iv)) *et* $*$ *la loi de groupe précédente, on obtient, si* $(H_n) \in$ NW$'$:

$$\begin{aligned}\big(\lambda(F_n) +_{\mathrm{exp}} (1-\lambda)(G_n)\big) * (H_n)& \\ = \lambda\big((F_n) * (H_n)\big) +_{\mathrm{exp}} (1-\lambda)\big((G_n) * (H_n)\big) &\quad \textit{si}\ \ h_0 = \mathrm{d},\end{aligned}$$

$$\begin{aligned}(H_n) * \big(\lambda(F_n) +_{\mathrm{can}} (1-\lambda)(G_n)\big)& \\ = \lambda\big((H_n) * (F_n)\big) +_{\mathrm{can}} (1-\lambda)\big((H_n) * (G_n)\big) &\quad \textit{pour tout}\ \ h_0.\end{aligned}$$

10. Résolution de l'équation fonctionnelle $\Phi(X+Y)=\Phi(X)\Phi(Y)$ dans $A[[X]]$ et d'équations apparentées.

LEMME 4. — *Soient p un nombre premier et n un entier > 0. Le coefficient multinomial $\dfrac{(pn)!}{(n\,!)^p}$ est un multiple de p.*

Démonstration. — Soit ν_p la valuation p-adique. De

$$\nu_p(n) \;=\; \sum_{i>0,\ p^i\,|\,n} 1,$$

on déduit, en notant [ ] la fonction partie entière :

$$\nu_p(n\,!) = \sum_{i>0}\left[\frac{n}{p^i}\right].$$

Il vient :

$$\nu_p\Big(\frac{(pn)!}{(n\,!)^p}\Big) = \sum_{i>0}\left[\frac{n}{p^{i-1}}\right] - p\left[\frac{n}{p^i}\right]$$

qui apparaît comme une somme d'entiers ≥ 0 dans laquelle la contribution d'indice $\nu_p(n)+1$ est > 0. □

THÉORÈME 7. — *Soient A un anneau commutatif unitaire intègre et $\Phi(X) = \sum_{n\geq 0} a_n X^n$ une solution dans $A[[X]]$ de l'équation fonctionnelle*

$$\Phi(X+Y) = \Phi(X)\Phi(Y).$$

Alors, on a ou $\Phi(X)=0$ ou $a_0 = 1$. Dans ce dernier cas, on obtient :

(i) *si A est de caractéristique $p>0$, on a $a_n = 0$ pour tout $n>0$;*

(ii) *si A est de caractéristique nulle, on a $a_n = a_1^n/n\,!$ pour tout $n>0$ avec a_1 arbitraire dans A (autrement dit, il n'y a pas d'autres solutions que les séries exponentielles $\exp(a_1X)$).*

Démonstration. — L'équation fonctionnelle équivaut, dans $A[[X]]$, à

$$a_{i+j}\binom{i+j}{i} = a_i a_j \quad \text{pour tout couple } (i,j) \text{ d'entiers}$$

(et, plus généralement, à $n\,!\,a_n/i_1!\ldots i_k! = a_{i_1}\cdots a_{i_k}$ si $n = i_1+\cdots+i_k$).

En particulier, a_0 est idempotent et est donc égal à 1 ou 0 (ainsi que tous les a_n dans ce second cas d'après la même relation).

En caractéristique $p > 0$, la relation $a_{np}(np)!/(n!)^p = a_n^p$ montre que $a_n = 0$ pour tout $n > 0$ puisque le coefficient multinomial $(np)!/(n!)^p$ est un multiple de p d'après le lemme 4. En caractéristique nulle, la relation s'écrit $\alpha_{i+j} = \alpha_i \alpha_j$ avec $\alpha_i = i!\,a_i$, d'où $a_i = a_1^i/i!$ (avec $i > 0$ et a_1 arbitraire).

Remarque. — Plus généralement, la relation $\Phi(pX) = (\Phi(X))^p$ déduite de l'équation fonctionnelle $\Phi(X+Y) = \Phi(X)\Phi(Y)$ montre que Φ est constante en caractéristique p.

Soit K un corps commutatif. On relie maintenant le théorème 7 et la détermination de NW.

Si $H(X,T)$ satisfait à (μ'), l'équation fonctionnelle est vérifiée avec $A = K[[T]]$. Soit, dans ce dernier cas, EXP_T l'ensemble des solutions

$$\sum_{i,j} \varphi_{i,j} X^i T^j$$

de l'équation fonctionnelle du théorème 7, lequel contient donc NW.

Si K est de caractéristique finie, EXP_T est réduit aux constantes 0 et 1 tandis qu'en caractéristique 0, l'ensemble EXP_T est formé par les séries $\exp(B(T)X)$ où $B(T) \in K[[T]]$. Avec ces notations et en caractéristique 0, on voit que NW est inclus strictement dans EXP_T.

Précisément, en écrivant les éléments de EXP_T sous la forme $\exp(XB(T))$, on obtient un élément $\sum_n H_n(X)T^n$ de $K[X][[T]]$ si et seulement si la série formelle $B \neq 0$ est de valuation $k > 0$; et ce dernier calcul montre que le degré de H_n est au plus la partie entière de n/k (avec égalité si n/k est entier).

En écrivant l'élément $H(X,T) = \sum_n H_n(X)T^n$ (supposé $\neq 0$) de EXP_T sous la forme $\sum_n A_n(T)X^n$, on déduit du théorème 7 que

- $H_0(0) = 1$,
- $H_n(0) = 0$ (spécialiser X en 0),
- $H(X,T) = \exp\big((\sum_n H_n'(0)T^n)X\big)$ (appliquer $\partial/\partial X$ et spécialiser X

en 0).

En particulier, $H_0'(0) = 0$ est une condition nécessaire et suffisante pour que $\sum_n H_n(X)T^n \in K[X][[T]]$ (et alors le degré de H_n est au plus n).

Le (i) du théorème 8 suivant qui établit l'équipotence de NW et EXP_T résume ces résultats.

THÉORÈME 8.

(i) *Toute solution* $H(X,T) = \sum_n H_n(X)T^n$ *de* (μ') *appartient à* EXP_T *(et donc aussi* $H(X,T^k)$ *si* $k > 0$*); réciproquement, si* $\sum_n H_n(X)T^n \in K[[X]][[T]]$ *appartient à* EXP_T*, alors* $\sum_n H_n(X)T^n$ *appartient à* NW *si et seulement si* $H_0'(0) \neq 0$.

(ii) *Soient* $\Phi(X,T) \in \mathrm{EXP}_\mathrm{T}$ *et* $\sum_n \Phi_n(X,T)$ *la décomposition canonique en composantes homogènes (*n *est le degré de la forme* $\Phi_n(X,T)$*) de l'élément* $\Phi(X)$*; alors,* $(\Phi_n(X,1))$ *appartient à* NW *(en termes de séries génératrices, on associe à l'élément courant* $\sum_{i,j} \varphi_{i,j} X^i T^j$ *de* EXP_T *l'élément* $\sum_{i,j} \varphi_{i,j} X^i T^{i+j} = \sum_n \big(\sum_{i+j=n} \varphi_{i,j} X^i \big) T^n$ *de* NW*).*

Démonstration. — L'application de EXP_T dans NW décrite dans (ii) est clairement injective, ce qui prouve l'équipotence de ces ensembles compte tenu du (i); elle est en fait bijective puisque tout élément $\sum_n H_n(X)T^n$ de NW vérifie $\deg(H_n) \leq n$ (l'application inverse est $\sum_n H_n(X)T^n \mapsto \sum_n H_n(X/T)T^n$). □

Remarques.

1) L'application $\big(\sum_{i,j} \varphi_{i,j} X^i T^j\big) \mapsto \big(\sum_{i,j} \varphi_{i,j} X^i T^{i+j}\big)$ du (ii) est l'identification de $K[[X,T]]$ avec la sous-algèbre de $K[X][[T]]$ formée des éléments $\sum_n H_n(X)T^n$ vérifiant $\deg(H_n(X)) \leq n$. La partie (ii) du théorème 8 justifie donc la limitation de l'étude au cas de $K[X][[T]]$ dans les paragraphes antérieurs ainsi que l'assertion correspondante du résumé comme on verra ci-dessous.

2) Un autre procédé pour passer d'un élément de $\mathrm{EXP}_T \subseteq K[[X,T]]$ à un élément de $K[X][[T]]$ consiste à l'écrire sous la forme :

$$\begin{aligned}\exp\big(B(T)X\big) &= \exp\big(B(0)X\big)\exp\big((B(T)-B(0))X\big)\\ &= \exp\big(B(0)X\big)\Big(\sum_n G_n(X)T^n\Big),\end{aligned}$$

les (G_n) étant des polynômes à coefficients dans K.

Équations apparentées.

L'équation fonctionnelle $S(X+Y,T)/S(X,T) \in K[[Y,T]]$ avec $S(X,T) \in K[[X,T]]$ se ramène, en écartant la solution triviale, à celle du théorème 7 pour

$$\Phi(Y,T) = S(X+Y,T)/S(X,T) = S(Y,T)/S(0,T).$$

On peut aussi traiter cette équation avec les considérations des §§5 et suivants. On se limite au cas de la caractéristique 0, lequel se ramène, sans perte de généralité comme on vient de le voir, à celui où les solutions appartiennent à NW′ ; on va donner des conditions suffisantes pour qu'il en soit ainsi :

THÉORÈME 9. — *Si la solution $S(X,T)$ de*

$$S(X+Y,T)/S(X,T) \in K[[Y,T]]$$

s'écrit $\sum_n S_n(X)T^n$ avec $S_n \in K[X]$, (S_n) étant une base de $K[X]$ vérifiant $S_0 = 1$, $S_1 = X$ et $S_n(0) = 0$ pour $n > 0$, alors, $S(X,T)$ appartient à NW′.

Démonstration. — Soit f l'endomorphisme défini par $f(S_0) = 0$ et $f(S_n) = S_{n-1}$, ce qui permet de prolonger f et d'écrire $f(S(X,T)) = TS(X,T)$. L'hypothèse sur $S(X,T)$ montre alors aussitôt l'égalité

$$\gamma \circ f\big(S(X,T)\big) = f^* \circ \gamma\big(S(X,T)\big),$$

d'où $\gamma \circ f = f^* \circ \gamma$ et donc, compte tenu des hypothèses, $f = \mathrm{d} + \sum_{i>1} b_i\, \mathrm{d}^i$. L'appartenance de $S(X,T)$ à NW′ est alors claire (*cf.* §§5 à 7). □

On poursuit ces raffinements du théorème 7 par restriction des hypothèses toujours en caractéristique nulle.

THÉORÈME 10. — *Soient A un anneau commutatif unitaire intègre de caractéristique nulle, $k > 1$ un entier, $\Phi(X) = \sum_{n\geq 0} a_n X^n$ une solution de l'équation fonctionnelle $\Phi(kX) = (\Phi(X))^k$ vérifiant $a_0 = 1$. Alors, on a $a_n = a_1^n/n!$ pour tout $n > 0$ avec a_1 arbitraire et $\Phi(X)$ est de la forme* $\exp(a_1 X)$.

Démonstration. — L'équation fonctionnelle est équivalente dans A aux relations, pour tout entier $n \geq 0$:

$$k^n a_n = \sum_{\alpha_1+\cdots+\alpha_k=n} a_{\alpha_1} \cdots a_{\alpha_k}.$$

On en déduit par récurrence l'égalité $a_n = a_1^n/n!$ pour $n > 0$ puisque $\mathrm{car}(A) = 0$. □

Remarques.

1) Le théorème 10 généralise le résultat du §2 relatif au coefficient dominant c_n du polynôme H_n. Mais en fait, en caractéristique 0, ce résultat facile suffit à remplir le «premier but» évoqué au §1 (résolution de la récurrence (μ')). En effet, on déduit de (μ') que, pour *tout* entier $k > 1$, la série $\Phi(X) = H(X,T)$ vérifie $\Phi(kX) = (\Phi(X))^k$ et, lorsque $\Phi(0) = 1$, il suffit d'après le théorème 10 que cette relation soit vérifiée pour *une* valeur $k > 1$ pour obtenir $H(X,T) = \exp(B(T)X)$ (s'écrivant sous la forme cherchée lorsque B est de valuation > 0). Plus généralement, d'autres équations fonctionnelles comme

$$\Phi\big((k-k')X\big) = \big(\Phi(X)\big)^k \big(\Phi(-X)\big)^{k'}$$

se déduisent de (μ') ; certaines d'entre elles apparaissent équivalentes comme dans le théorème 10 à (μ') (pour l'exemple précédent, si $|k-k'| > 1$), d'autres non (cas $|k-k'| \leq 1$ pour encore le même exemple).

2) L'ensemble de ce travail aurait pu être écrit dans un cadre homogène en considérant les séries à deux indéterminées $S(X,T) \in K[[X,T]]$ vérifiant $S(X+Y,T) = S(X,T)S(Y,T)$ et leurs décompositions homogènes canoniques. À ce titre, les seules solutions symétriques (*i.e.* vérifiant $S(X,T) = S(T,X)$) sont de la forme k^{XT} et ne correspondent qu'en apparence dans NW à des suites de polynômes réciproques (X^n correspond au polynôme homogène symétrique de degré $2n$

$$X^nT^n = \sum_{0\leq i\leq 2n} a_i X^i T^{2n-i} \quad \text{avec} \quad a_i = \begin{cases} 1 & \text{si } i = n, \\ 0 & \text{si } i \neq n\,). \end{cases}$$

On prendra garde dans le cadre homogène aux changements introduits par une telle réécriture (ce n'est rien de plus d'après la remarque 1 suivant le théorème 8). Par exemple, la décomposition homogène de

$$\sum_{n\geq 0} \frac{X(X-1)\cdots(X-(n-1))}{n!} T^n = (1+T)^X$$

donne naissance (par l'application définie dans le théorème 8 (ii)) à l'élément (C_n) de NW défini par

$$\sum_{n\geq 0} C_n(X)T^n = (1+T)^{TX}$$

tandis que c'est la décomposition homogène de $(1+T)^{X/T}$ qui fournit $(X(X-T)\cdots(X-(n-1)T)/n!)$; les «bonnes» notations (lesquelles doivent faire apparaître simplement les objets intéressants...) sont donc bien celles adoptées ici.

11. Généralisations à plusieurs variables.

La partie différentielle du présent travail s'étend sans difficulté de $K[X]$ à $K[X_1, \ldots, X_m]$. On applique pour cela, avec les notations des §§5 à 7, les formules suivantes où la seconde généralise la première :

$$\mathrm{d} = \sum_{i \geq 0} \mathrm{d}\Phi_i(0) f^i, \quad \gamma = \exp(Y\mathrm{d}) = \sum_{i \geq 0} F_i(Y) f^i.$$

On va développer maintenant ces remarques dans deux directions : formule de Taylor pour $K[X_1, \ldots, X_m]$ d'abord puis « cadre homogène » auquel on consacrera un sous-paragraphe particulier.

Formule de Taylor pour $K[X_1, \ldots, X_m]$.

Soit γ_m l'application généralisant γ définie par

$$\gamma_m\big(P(X_1, \ldots, X_m)\big) = P(X_1 + Y_1, \ldots, X_m + Y_m).$$

La formule de Taylor s'écrit :

$$\gamma_m = \exp(Y_1\,\mathrm{d}_1 + \cdots + Y_m\,\mathrm{d}_m) = \exp(Y_1\,\mathrm{d}_1) \circ \cdots \circ \exp(Y_m\,\mathrm{d}_m)$$

avec $\mathrm{d}_i = \partial/\partial X_i$ pour $1 \leq i \leq m$.

Soit f un élément de T_2 (cf. §5); on peut, à partir de f, définir $f_1, \ldots, f_m$ à la façon des dérivées partielles et obtenir

$$\gamma_m = \Big(\sum_{i \geq 0} F_i(Y_1) f_1^i\Big) \circ \cdots \circ \Big(\sum_{i \geq 0} F_i(Y_m) f_m^i\Big) \quad \text{avec} \quad (F_n) = J(f).$$

Appliquée au cas du polynôme $Q(X_1, \ldots, X_m) = P(X_1 + \cdots + X_m)$, où P est maintenant un polynôme à une seule indéterminée dans lequel les f_i ont tous la même action, le développement de ce dernier produit permet de retrouver les identités (μ) satisfaites par (F_n).

Cadre homogène et séries formelles.

Les généralisations en caractéristique 0 dans le cadre homogène ne se limitent pas aux considérations du §10; elles ouvrent une connexion entre des travaux récents de normalisation en matière de « hauteur » et des travaux très fins et plus anciens de géométrie des polynômes (de lecture assez rebutante dans la littérature classique faute d'interprétation géométrique et de normalisation convenables). C'est l'objet du sous-paragraphe suivant.

Identité d'Euler et produits scalaires sur les espaces de polynômes homogènes.

L'identité d'Euler précise pour l'espace des polynômes homogènes de degré n à k indéterminées $X_1, \dots, X_k$ le terme courant $[Y_1 \,\mathrm{d}_1 + \cdots + Y_k \,\mathrm{d}_k]^r$ de la formule de Taylor où $r = 1, 2, \dots$:

$$\Big[Y_1 \,\mathrm{d}_1 + \cdots + Y_k \,\mathrm{d}_k\Big]^r_{Y_1 = X_1, \dots, Y_k = X_k} = n(n-1)\cdots(n-r+1)\,\mathrm{Id}\,.$$

Soit $f \in T_2$; l'identité précédente se généralise en :

$$\Big[\sum_n F'_n(0)(Y_1 f_1^n + \cdots Y_k f_k^n)\Big]^r_{Y_1 = X_1, \dots, Y_k = X_k} = n(n-1)\cdots(n-r+1)\,\mathrm{Id}\,.$$

Quand $r = n$, cette identité d'Euler pour un polynôme homogène P de degré n et de k variables peut être écrite sous la forme (où i décrit l'ensemble des applications de $[1, n]$ dans $[1, k]$) :

$$P(X_1, \dots, X_k) = \frac{1}{n\,!} \sum_i \frac{\partial^n P}{\partial X_{i_1} \partial X_{i_2} \cdots \partial X_{i_n}} X_{i_1} X_{i_2} \cdots X_{i_n}$$

en développant directement le multinôme

$$\frac{1}{n\,!}\Big[Y_1 \,\mathrm{d}_1 + \cdots + Y_k \,\mathrm{d}_k\Big]^n_{Y_1 = X_1, \dots, Y_k = X_k},$$

ce qui permet d'éviter le calcul et l'écriture de factorielles (cette identité est désignée par l'appellation « formule de Taylor » chez certains auteurs).

On peut évidemment aussi l'obtenir en dérivant n fois par rapport à T l'égalité

$$P(TX_1, \dots, TX_k) = T^n P(X_1, \dots, X_k)$$

puis en spécialisant au point $T = 1$.

Cette identité a donné naissance à un produit scalaire hermitien pour deux polynômes complexes P et Q homogènes de k variables et de même degré n considéré par Bombieri, Montgomery, Enflo, Beauzamy et, indépendamment, M. Laurent :

$$\langle P, Q\rangle = \Big(\frac{1}{n\,!}\Big)^2 \sum_i \Big(\frac{\partial^n P}{\partial X_{i_1} \partial X_{i_2} \cdots \partial X_{i_n}}\Big)\Big(\frac{\partial^n \overline{Q}}{\partial X_{i_1} \partial X_{i_2} \cdots \partial X_{i_n}}\Big).$$

Ce produit scalaire coïncide avec le suivant, considéré par Kostlan, Smale, *etc.* À tout couple de polynômes homogènes

$$P(\underline{X}) = \sum_i a_{(i)}\underline{X}^{(i)}, \quad Q(\underline{X}) = \sum_i b_{(i)}\underline{X}^{(i)}$$

de degré n à k indéterminées (i décrit ici l'ensemble des applications de $[1, k]$ dans les entiers ≥ 0 vérifiant $\sum_{1\leq a\leq k} i_a = n$ et $\underline{X}^{(i)}$ représente le monôme $X_1^{i_1} X_2^{i_2} \cdots X_k^{i_k}$ avec $i_1 + \cdots + i_k = n$), on associe :

$$\sum_i \frac{a_{(i)}\bar{b}_{(i)}}{\binom{n}{(i)}} \quad \text{où} \quad \binom{n}{(i)} = \frac{n!}{i_1! \ldots i_k!}.$$

Selon S. Lang, l'utilisation de ce dernier produit scalaire hermitien (*unitaire pour le groupe unitaire*) où les factorielles ne figurent que dans les définitions (et on pourrait les en chasser, *cf.* ci-dessus...) « sans revenir plus jamais dans les énoncés » (*loc. cit.*) doit fournir de meilleures normalisations dans de nombreux travaux (Mahler, Berenstein-Yger...) et jouer un rôle naturel en tant que « hauteur » pour les polynômes.

Le théorème suivant donne une troisième définition (décrivant comment établir la coïncidence précédente) et ouvre une connexion avec la notion d'apolarité qu'annonce le théorème 12.

THÉORÈME 11. — *Soient*

$$P(X_1, \ldots\ldots X_n) = P(\underline{X}) = \sum_i a_{(i)}\underline{X}^{(i)},$$

$$Q(X_1, \ldots\ldots X_n) = Q(\underline{X}) = \sum_i b_{(i)}\underline{X}^{(i)}$$

deux polynômes homogènes de même degré n à coefficients dans un même corps K de caractéristique 0. On note comme précédemment d_i l'opérateur $\partial/\partial X_i$. Alors :

$$\frac{1}{n!} P(\mathrm{d}_1, \ldots, \mathrm{d}_k)(Q) = \frac{1}{n!} Q(\mathrm{d}_1, \ldots, \mathrm{d}_k)(P) = \sum_i \frac{a_{(i)}\bar{b}_{(i)}}{\binom{n}{(i)}}.$$

Démonstration. — Par symétrie et linéarité, il suffit de vérifier que

$$\frac{1}{n!} P(\mathrm{d}_1, \ldots, \mathrm{d}_k)(Q) = \sum_i \frac{a_{(i)}\bar{b}_{(i)}}{\binom{n}{(i)}}$$

est vérifié lorsque $P(\underline{X}) = \underline{X}^{(i)} = X_1^{i_1} X_2^{i_2} \cdots X_k^{i_k}$ et $Q(\underline{X}) = \underline{X}^{(j)} = X_1^{j_1} X_2^{j_2} \cdots X_k^{j_k}$. Comme $i_1 + \cdots + i_k = j_1 + \cdots + j_k = n$, on obtient 0 si $i \neq j$ et $1/\binom{n}{(i)}$ si $i = j$. □

THÉORÈME 12. — *Soient $P(X,T)$ et $Q(X,T)$ deux polynômes homogènes de même degré n. Alors :*

$$\Big(P\Big(\frac{\partial}{\partial T},-\frac{\partial}{\partial X}\Big)(Q)\Big)T^n = (-1)^n\Big(Q\Big(\frac{\partial}{\partial T},-\frac{\partial}{\partial X}\Big)(P)\Big)T^n = \sum_{0\leq i\leq n}(-1)^{n-i}\Big(\frac{\partial^i P}{\partial X^i}\Big)\Big(\frac{\partial^{n-i}Q}{\partial X^{n-i}}\Big).$$

Démonstration. — Le terme $\sum_{0\leq i\leq n}(-1)^i(\partial^i P/\partial X^i)(\partial^{n-i}Q/\partial X^{n-i})$ est bien un polynôme homogène de degré n multiple de T^n puisque

$$\frac{\partial}{\partial X}\Big(\sum_{0\leq i\leq n}(-1)^i\Big(\frac{\partial^i P}{\partial X^i}\Big)\Big(\frac{\partial^{n-i}Q}{\partial X^{n-i}}\Big)\Big) = \Big(\frac{\partial^{n+1}Q}{\partial X^{n+1}}\Big)P + (-1)^n\Big(\frac{\partial^{n+1}P}{\partial X^{n+1}}\Big)Q = 0.$$

Lorsque $P(X,T) = X^iT^{n-i}$ et $Q(X,T) = X^jT^{n-j}$,

$$\Big(P\Big(\frac{\partial}{\partial T},-\frac{\partial}{\partial X}\Big)(Q)\Big)T^n$$

s'annule si $j\neq n-i$ et prend la valeur $(-1)^{n-i}(i\,!)(n-i)!\,T^n$ si $j=n-i$ et il suffit de vérifier qu'on obtient dans ce cas le même résultat pour

$$\sum_{0\leq m\leq n}(-1)^{n-m}\Big(\frac{\partial^m P}{\partial X^m}\Big)\Big(\frac{\partial^{n-m}Q}{\partial X^{n-m}}\Big)\,;$$

la remarque faite au début de la démonstration montre qu'importe seul le coefficient de T^n ; on voit ainsi qu'on obtient 0 si $j\neq n-i$ et, si $j=n-i$, que n'intervient que la contribution attachée à $m=i$; la vérification est alors immédiate. On achève la démonstration en intervertissant les rôles de P et Q. □

COROLLAIRE. — *Soient*

$$P(X,T) = \sum_i a_iX^iT^{n-i},\quad Q(X,T) = \sum_i b_iX^iT^{n-i}$$

deux polynômes homogènes de même degré n. Soit

$$Q^*(X,T) = \overline{Q}(T,-X)$$

le polynôme «antiréciproque» associé à $Q(X,T)$. Alors on a :

$$\langle P, Q^* \rangle = \frac{(-1)^n}{n!} P\Big(\frac{\partial}{\partial T}, -\frac{\partial}{\partial X}\Big)(Q)$$
$$= \frac{1}{n!} Q\Big(\frac{\partial}{\partial T}, -\frac{\partial}{\partial X}\Big)(P) = \sum_i \frac{(-1)^i a_i b_{n-i}}{\binom{n}{i}}.$$

Il est intéressant de rapprocher ce corollaire des théorèmes 11 et 12 du résultat ci-dessous de Grace (1900) sur la localisation des racines des polynômes complexes par rapport à des lignes circulaires. Pour l'énoncé de ce théorème, quelques rappels sont nécessaires : un domaine circulaire sur la sphère de Riemann S est un domaine (pouvant être S ou l'ensemble vide) ouvert ou fermé de bord circulaire; deux polynômes P et Q de $\mathbb{C}[z]$ de même degré et qu'on écrit sous leurs formes homogènes PH, QH sont dits apolaires si $\langle PH, QH^* \rangle = 0$. Cette notion d'apolarité, découverte en 1900 par Grace, a donné lieu à de nombreux travaux (Dieudonné indique sept références en 1938 pour la preuve) dont ceux de Walsh, Szegö et Dieudonné lui-même. Dieudonné établira que la relation d'apolarité, après élimination des coefficients a_i et b_i au profit des racines (x_i) de P et (y_j) de Q — et retranscrite formellement comme polynôme en ces dernières — est «la seule relation symétrique entre les x_i et les y_j qui soit linéaire par rapport à chacune de ces variables et *invariante lorsqu'on effectue une même transformation homographique sur toutes les variables*».

THÉORÈME (Grace). — *Soient*

$$P(z) = a_n z^n + \cdots + a_0, \quad Q(z) = b_n z^n + \cdots + b_0$$

deux polynômes de même degré n; alors, s'ils vérifient la relation bilinéaire d'apolarité :

$$\sum_{0 \le i \le n} (-1)^i \frac{a_i b_{n-i}}{\binom{n}{i}} = 0,$$

tout domaine circulaire contenant les zéros de P (resp. Q) contient au moins un zéro de Q (resp. P).

Le rapprochement avec les théorèmes précédents montre bien que la relation d'apolarité s'écrit simplement $\langle P, Q^* \rangle = 0$ (après transformation homogène de P et Q). Une part des progrès conceptuels récents est cachée dans cette transformation.

En effet, la notion d'apolarité pour un couple de polynômes de même degré d'une variable dans $\mathbb{C}$ fournit une condition suffisante pour qu'aucune ligne circulaire ne sépare les zéros de ces polynômes (dans [L3], est donnée une condition nécessaire et suffisante de nature géométrique, impliquée par la relation algébrique d'apolarité) ; les nombreuses visions traditionnelles de la relation d'apolarité sont donc affines et non projectives. Plus précisément, c'est la composition de « dérivations polaires » (déduites de la dérivation par changement du pôle (*i.e.* le point à l'infini) sur la sphère de Riemann) du polynôme P, les pôles décrivant l'ensemble des zéros de Q, qui fait apparaître cette relation d'apolarité. Mais ces calculs, lourds en termes affines, s'éclairent « en projectif ». Les dérivations polaires ne sont en fait que des spécialisations, en les coordonnées homogènes des pôles, des indéterminées X et T de l'opérateur d'Euler $X\partial/\partial X + T\partial/\partial T$ et, sous cette forme, la commutativité des opérateurs de dérivation polaire est évidente. La formation (après transformation homogène de P et Q, ce qui permet de plus de traiter le cas où ces polynômes n'ont pas le même degré affine) des termes essentiels $P(\partial/\partial T, -\partial/\partial X)(Q)$ ou $Q(\partial/\partial T, -\partial/\partial X)(P)$ est alors toute naturelle (*cf.* le théorème 14 ci-dessous).

À l'exception des remarques ci-dessous, on ne poursuit pas davantage dans ce domaine pour lequel on renvoie à [L1] et [L2]. De plus, l'indispensable traitement projectif n'est que la première composante de ces progrès conceptuels. Il est néanmoins curieux et intéressant de voir que ces références, unifiant un vaste ensemble — s'étalant sur 80 ans... — de travaux sans liens apparents et d'apparences de raffinements (*cf.* [L4]), sont apparues simultanément à la fin des années 1980 avec les normalisations précitées visant à dégager de « bonnes hauteurs » en plusieurs variables.

Remarques.

1) Avec les hypothèses du théorème de Grace, le théorème 12 montre que la relation $\langle P, Q^* \rangle = 0$ est équivalente (où z est quelconque puisque éliminé) à

$$\sum_{0\le i\le n} (-1)^i P^{(i)}(z) Q^{(n-i)}(z) = 0.$$

2) L'identité générale

$$\frac{\mathrm{d}}{\mathrm{d}z}\Big[\sum_{0\le i\le n} (-1)^i\, \mathrm{d}^i P(z)\, \mathrm{d}^{n-i} Q(z)\Big] = P(z)\,\mathrm{d}^{n+1}Q(z) + (-1)^n Q(z)\,\mathrm{d}^{n+1}P(z)$$

qu'on vient d'appliquer ci-dessus est l'une des formes de la formule de Taylor à une variable...

3) La forme ci-dessus de la formule de Taylor ne s'étend pas par substitution de $f = B(\mathrm{d})$ (avec $B \in \mathrm{GSF}$) à d, mais les autres considérations de ce sous-paragraphe s'étendent puisqu'obtenues par compositions successives de l'opérateur d'Euler $[Y_1\,\mathrm{d}_1+\cdots+Y_N\,\mathrm{d}_N]^r$ auquel on peut substituer, en définissant $(F_n) = J(f)$:

$$\Big[\sum_n F_n'(0)(Y_1 f_1^n + \cdots + Y_N f_N^n)\Big]^r.$$

D'autres extensions viendront aux 8) et 9).

4) Quand le nombre k d'indéterminées est 2, tout produit scalaire $\langle P, Q\rangle$ (*cf.* le théorème 11 et les définitions le précédant) apparaît comme le membre non nul d'une relation d'apolarité et peut donc être écrite comme telle. En particulier, le théorème de coïncidence de Walsh (*i.e.* le théorème 3 de [L2]) montre que, pour tout couple de polynômes P, Q de $\mathbb{C}[z]$ et tout domaine circulaire D contenant les zéros de P, il existe $x \in D$ vérifiant $\langle P, Q\rangle = a_0 Q^*(x)$, a_0 désignant le coefficient dominant de P et $\langle P, Q\rangle$ le résultat obtenu après « passage en homogène » de P et Q. Mais il n'est pas souhaitable en général d'inclure l'une de ces deux notions de natures différentes dans l'autre (*cf.* les remarques 7, 8 et 9) : apolarité et produit scalaire.

5) La remarque précédente se déduit du théorème (*cf.* [L2]) qui est, pour $m = 1$, la forme raffinée de Szegö du théorème de Bernstein (*cf.* [L4]) et, pour $m = n$, une forme du théorème de Walsh :

THÉORÈME 13. — *Soient U un domaine circulaire, $P \in \mathbb{C}[z]$ un polynôme de degré n, $PH(X,T)$ le polynôme homogène défini par $T^nP(X/T)$. Pour tout entier naturel $m \le n$ et tout $(m+1)$-uplet $\xi = (x_1, \ldots, x_m, z)$, soit*

$$Q_\xi(X,T) = (X - zT)^{n-m} \prod_{1\le i\le m} (X - x_iT).$$

Alors, pour tout entier $m \le n$, l'application associant au $(m+1)$-uplet $\xi = (x_1, \ldots, x_m, z)$:

$$\begin{aligned}&\frac{1}{n!}Q_\xi\Big(\frac{\partial}{\partial T}, -\frac{\partial}{\partial X}\Big)(PH)\\ &\quad= \frac{1}{n!}\Big(\Big(\prod_{1\le i\le m}\Big(\frac{\partial}{\partial T} + x_i\frac{\partial}{\partial X}\Big)\Big)\Big(\frac{\partial}{\partial T} + z\frac{\partial}{\partial X}\Big)^{n-m}\Big)(T^nP(X/T))\end{aligned}$$

a pour image $P(U)$ lorsque ξ décrit U^{m+1}.

6) On revient sur les propriétés d'invariance des notions précédentes soulignées par S. Lang et J. Dieudonné respectivement (en italique ci-dessus). Ces propriétés sont distinctes puisque se rattachant à des groupes distincts (groupe unitaire — et pour toute dimension — pour le produit scalaire hermitien cité par S. Lang, groupe $\mathrm{SL}_2(\mathbb{C})$ pour la relation d'apolarité étudiée par J. Dieudonné) ; on peut toutefois les traiter simultanément par le lemme suivant de vérification automatique :

LEMME. — *Soient $\underline{X} = (X_1, \ldots, X_k)$ et $\underline{Y} = (Y_1, \ldots, Y_k)$ deux matrices unilignes liées par la relation matricielle $\underline{X} = \underline{Y}M$ et $P = P(\underline{X})$ un polynôme en les indéterminées $X_1, \ldots, X_k$, la matrice M et le polynôme P étant à coefficients dans un même anneau commutatif. On définit $P_M(\underline{Y}) = P(\underline{Y}M)$. Alors, avec des notations évidentes, on a :*

$$\frac{\partial}{\partial \underline{Y}} = \frac{\partial}{\partial \underline{X}} \widetilde{M} \quad \text{et} \quad P_M\Big(\frac{\partial}{\partial \underline{Y}}\Big) = P\Big(\frac{\partial}{\partial \underline{X}} \widetilde{M} M\Big)$$

où $\widetilde{M}$ désigne la transposée de M.

Applications.

— Pour la forme bilinéaire symétrique introduite dans le théorème 11 et appliquée à un couple de polynômes homogènes de même degré $(P(\underline{X}), Q(\underline{X}))$, le lemme montre l'invariance pour toute matrice orthogonale M du terme

$$P_M\Big(\frac{\partial}{\partial \underline{Y}}\Big)(Q_M(\underline{Y})) = P\Big(\frac{\partial}{\partial \underline{X}}\Big)(Q\underline{X})).$$

— Quand $\mathbb{C}$ est le corps des scalaires, on voit de même en ajoutant la conjugaison complexe que le produit scalaire hermitien $\langle P, Q\rangle$ est inchangé quand on effectue une transformation unitaire sur les indéterminées (et c'est ainsi qu'il faut comprendre et prouver l'assertion précitée de S. Lang).

— Toujours dans le cas complexe et quand le nombre k d'indéterminées est 2, on voit, pour tout polynôme P homogène de degré n et par une modification immédiate de la preuve du lemme, que la relation matricielle $(X, T) = (\xi, \tau)M$ implique

$$P_M\Big(\frac{\partial}{\partial \tau}, -\frac{\partial}{\partial \xi}\Big) = \big(\det(M)\big)^n P\Big(\frac{\partial}{\partial T}, -\frac{\partial}{\partial X}\Big).$$

Cette remarque rend claire l'assertion de J. Dieudonné, une homographie sur les variables dans la relation d'apolarité revenant après transformation homogène à effectuer l'homographie inverse sur le couple des polynômes associés ; il suffit donc de vérifier qu'on a, pour toute matrice M de déterminant $\det(M) = 1$:

$$P\Big(\frac{\partial}{\partial T}, -\frac{\partial}{\partial X}\Big)(Q(X,T)) = P_M\Big(\frac{\partial}{\partial \tau}, -\frac{\partial}{\partial \xi}\Big)(Q_M(\xi,\tau))$$

comme on vient de le voir.

7) On revient dans le cas des polynômes $P(z) = \sum_{0\le i\le n} a_i z^i$ à coefficients complexes sur la norme hermitienne

$$\sqrt{\langle P,P\rangle} = \sum_i \frac{a_i\bar{a}_i}{\binom{n}{i}}$$

notée $|[P]|$. C'est une « hauteur » aisément comparable à

$$\begin{aligned} L_2(P) &= \Big(\sum_i a_i\bar{a}_i\Big)^{1/2} \\ &= \Big(\int_0^1 \big|P(\mathrm{e}^{2i\pi t})\big|^2 \,\mathrm{d}t\Big)^{1/2} \\ &= \Big(\frac{1}{n!}\Big[\frac{\mathrm{d}^n(P(z)z^n\overline{P}(z^{-1}))}{\mathrm{d}z^n}\Big]_{z=0}\Big)^{1/2} \end{aligned}$$

mais qui se comporte assez différemment vis-à-vis des hauteurs classiques ; par exemple, alors que toutes ces dernières sont inchangées (ou croissent) lors du passage de $P(z)$ à $P(z^m)$, l'entier m allant en croissant, on voit que :

$$\lim_{m\to\infty} \big|[P(z^m)]\big| = \sqrt{|a_0|^2 + |a_d|^2}.$$

À noter toutefois, outre les résultats de Beauzamy, Bombieri, Enflo et Montgomery [BBEM] sur les minorations des normes de produits tels que $|[PQ]|$, les inégalités $|[P']| \le n\,|[P]|$ (analogue à celle de Bernstein) ou encore, pour $i \ge 0$:

$$\begin{aligned} \|\mathrm{d}^i P\| &= \sup_{|z|=1} \big|\mathrm{d}^i P(z)\big| \\ &\le n(n-1)\cdots(n-i+1)\bigg(\sum_j \Big(\frac{(i+j)!}{j!}\Big)^2\binom{n}{i+j}\bigg)^{1/2} |[P]| \end{aligned}$$

déduites, par la majoration de Cauchy-Schwarz et pour tout nombre complexe α, de l'identité :

$$\mathrm{d}^i P(\alpha) = P^{(i)}(\alpha) = \big\langle P(z), n(n-1)\cdots(n-i+1)z^i(1+\bar{\alpha}z)^{n-i}\big\rangle.$$

8) Soient

$$P(X) = a_pX^p + \cdots + a_0, \quad Q(X) = b_qX^q + \cdots + b_0$$

deux polynômes à coefficients dans un corps commutatif K de caractéristique 0. On leur associe les polynômes homogènes de même degré projectif $n \geq \sup(p,q)$

$$PH(X,T) = \sum_i a_iX^iT^{n-i}, \quad QH(X,T) = \sum_i b_iX^iT^{n-i}$$

(en posant $a_i = 0$ (resp. $b_i = 0$) pour $i > p$ (resp. $i > q$)) et le composé par apolarité de ces polynômes homogènes :

$$\{P(X), Q(X)\}_n = \sum_{0 \leq i \leq n} \frac{(-1)^i a_{n-i} b_i}{\binom{n}{i}} = \frac{1}{n!} \sum_{0 \leq i \leq n} (-1)^i P^{(i)}(z) Q^{(n-i)}(z).$$

Avec cette dernière définition, on vérifie aisément, par linéarité à partir des dérivations, le théorème suivant qui montre comment cette notion permet, pour un élément f quelconque de T_2 (*cf.* §5), de ramener le calcul de $f(P)$ (où $P \in K[X]$) en un point α à celui de $f((\alpha - X)^n)$, ce dernier polynôme étant considéré comme un polynôme en l'indéterminée α :

THÉORÈME 14.

(i) *Soient* $P(X) \in K[X]$ *un élément de degré au plus* n *et* $B \in$ GSF ; *alors,* d *désignant comme précédemment la dérivation, on peut écrire pour tout élément* α *de* K :

$$\{(B(\mathrm{d})P)(X), (\alpha - X)^n\}_n = \{P(X), (B(-\mathrm{d})(\alpha - X)^n)\}_n = \big(B(\mathrm{d})P\big)(\alpha).$$

(ii) *En particulier, pour tout* $f \in T_2$ *et en notant* $f/\partial\alpha$ *l'opérateur* f *pour l'indéterminée* α :

$$\big(f(P)\big)(\alpha) = \{(f(P))(X), (\alpha - X)^n\}_n = \{P(X), f/\partial\alpha(\alpha - X)^n)\}_n.$$

9) Les remarques précédentes rendent clair l'énoncé suivant, écrit en termes affines mais de nature projective, qui est derechef une généralisation de la formule de Taylor relative à une variante des formes «apolarité» et $\langle\ ,\ \rangle$:

THÉORÈME 15. — *Soient*

$$P(X) = a_p X^p + \cdots + a_0, \quad Q(X) = b_q X^q + \cdots + b_0$$

deux polynômes à coefficients dans un corps commutatif K de caractéristique 0. *Soit un entier $n \geq \sup(p,q)$. En posant $a_i = 0$ (resp. $b_i = 0$) pour $i > p$ (resp. $i > q$), on définit la forme bilinéaire symétrique :*

$$\begin{aligned}\{\{P(X), Q(X)\}\}_n &= \sum_{0\leq i\leq n} \frac{a_{n-i} b_i}{\binom{n}{i}} \\ &= \frac{1}{n!}\Big[\sum_i \mathrm{d}^i(P(X))\mathrm{d}^{n-i}(Q(X))\Big]_{X=0}.\end{aligned}$$

Alors, pour tout opérateur $B(\mathrm{d}) = \sum_{i\geq 0} a_i \mathrm{d}^i$ et tout couple P, Q de polynômes comme ci-dessus :

$$\{\{B(\mathrm{d})(P), Q\}\}_n = \{\{P, B(\mathrm{d})(Q)\}\}_n.$$

En particulier, pour tout $\alpha \in K$ et pour tout $f \in T_2$:

$$\{\{f(P(X)), (X+\alpha)^n\}\}_n = \{\{P(X), f((X+\alpha)^n)\}\}_n = f(P)(\alpha).$$

12. Conclusion.

La présente étude est partie de celle des « familles canoniques de polynômes générateurs pour l'exponentielle » (ces familles sont nombreuses, d'où l'inhabituel pluriel pour l'adjectif « canonique » rattaché aux identités (μ) ; à noter toutefois que, dans le cas complexe, les seules approximations polynomiales de e^z (dans un sens et des domaines convenables) satisfont à (μ)). Cette étude a de fait débouché sur une description formelle détaillée de l'application $U(T) \mapsto (U(T))^X$ définie sur le groupe des unités (*i.e.* $U(0) \neq 0$) de $K[[T]]$.

Ceci est un cas particulier de l'action (à droite) « par substitution $T \mapsto B(T)$ » du groupe GSF dans $A[[T]]$, où A désigne une K-algèbre, définie par $(B(T), S(T)) \in \mathrm{GSF} \times A[[T]] \mapsto S(B(T))$; action qu'il est préférable, comme on l'a vu au §9, de remplacer par l'action à gauche associant $S(B^{-1}(T))$ au même couple $(B(T), S(T))$.

Les formules permettant de calculer explicitement l'orbite d'un élément $\sum_{n\geq 0} a_n T^n$ sont liées à des suites emboîtées de matrices triangulaires pour lesquelles on a renvoyé à [H]. Ce sont ces formules qu'il convient d'appliquer pour passer d'un élément à un autre dans NW′, ensemble qui n'est autre dans le cas $A = K[X]$ que l'orbite des éléments $X^n/n!$ ou $X(X-1)\cdots(X-(n-1))/n!$ donnés comme exemples au § 1. De fait, le passage de cet exemple à l'autre se résume à écrire les identités suivantes (équivalentes) déjà vues de façon très détaillée au § 3 par des arguments combinatoires :

$$\begin{aligned}\sum_{n\geq 0}\frac{(XT)^n}{n!} &= \sum_{n\geq 0}\frac{X(X-1)\cdots(X-(n-1))}{n!}\Big(\sum_{k>0}\frac{T^k}{k!}\Big)^n\\ &= \exp(XT),\end{aligned}$$

$$\begin{aligned}\sum_{n\geq 0}\frac{X(X-1)\cdots(X-(n-1))}{n!}T^n &= \sum_{n\geq 0}\frac{1}{n!}\Big(X\Big(\sum_{k>0}(-1)^{k+1}\frac{T^k}{k}\Big)\Big)^n\\ &= (1+T)^X.\end{aligned}$$

L'action de groupe ci-dessus est aussi une généralisation de celle du § 10 où l'on a examiné les liens entre l'étude de NW et celle d'équations fonctionnelles convenables. L'intérêt de relier relations de récurrence et équations fonctionnelles *via* des séries génératrices convenables est bien connu. Plus précisément, il existe dans les anneaux de séries formelles tout un ensemble de situations analogues à celles traitées ici et se ramenant à l'étude d'une action du groupe GSF.

La « vision différentielle » du problème a permis d'aborder beaucoup d'autres problèmes. Elle a conduit à examiner le sens général de la formule de Taylor et à montrer que sa composante « non linéaire » (ou à deux variables) (la composante « linéaire » (ou à une variable) fait l'objet du § 6, le lemme 3 qui le termine préparant au passage « non linéaire ») se ramène d'abord à la commutation (*cf.* (Γ)) d'un endomorphisme f du K-espace $K[X]$ abaissant le degré d'une unité (et normalisé par $f(X)=1$) avec l'opérateur γ défini par $\gamma(P(X)) = P(X+Y)$ (avec $P \in K[X]$). Plus précisément, pour tout polynôme P et tout endomorphisme f abaissant le degré d'une unité, on peut écrire $P(X+Y)$ de façon unique sous la forme $\sum_{i\geq 0} F_i(Y)f^i(P)(X)$, les coefficients $F_i(Y)$ dépendants de P ; et c'est l'existence d'une suite $(F_i(Y))$ indépendante de P qui caractérise l'existence d'une formule de Taylor (ou, si l'on préfère, de Taylor-Newton). On sait qu'alors (*cf.* (J″), § 5) cette suite est liée à f par une réduction de Jordan de sorte qu'on peut dire que toute « vraie » formule de Taylor — décomposition linéaire de $P(X+Y)$ suivant des termes en X et en Y, le polynôme P décrivant $K[X]$ — est attachée à un élément de NW'.

Le théorème suivant résume brièvement (et très partiellement) ce dernier point de vue :

THÉORÈME 16. — *Soit $(F_n(X))$ une base de $K[X]$ avec $F_0(X) = 1$ et $F_1(X) = X$. Soit Y une nouvelle indéterminée. S'il existe un endomorphisme f de $K[X]$ tel que, pour tout $P \in K[X]$*

$$P(X+Y) = \sum_{i\geq 0} F_i(Y) f^i(P)(X)$$

alors on a $f = B(\mathrm{d}) = \sum\limits_{i>0} b_i \, \mathrm{d}^i$ avec $b_1 = 1$ et la suite (F_n) satisfait à $f(F_i) = F_{i-1}$ pour $i > 0$ et appartient à NW'.

Démonstration. — Avec les notations du §5, on voit (*cf.* fin du §7) que f vérifie (Γ) et donc commute avec γ et d; la relation $f = B(\mathrm{d}) = \sum\limits_{i\geq 0} b_i \, \mathrm{d}^i$ s'en déduit (*cf.* théorème 3 (i)) et il suffit de réappliquer l'hypothèse avec $P = 1$, puis $P = X$ pour obtenir $b_0 = 0$ et $b_1 = 1$. L'appartenance de f à T_2 est ainsi établie et par suite le résultat. □

On achève cette conclusion en évoquant deux autres domaines potentiels d'application de ce travail :

— Dans des applications liées aux polynômes

$$J(\delta_x) = \Big(\frac{X(X-x)\cdots(X-(n-1)x)}{x^n \, n!}\Big)_n$$

ou à l'exponentielle et, plus généralement, à l'approximation diophantienne ou aux «spécialisations multiples» (*cf.* par exemple, théorème d'irréductibilité de Hilbert), le présent cadre formel fournit un vocabulaire commode et éclairant les démonstrations.

— Dans le domaine diophantien où le «contrôle des dénominateurs» est essentiel, on peut noter le :

THÉORÈME 17 (*cf.* théorème 2 (i)). — *Si B est à coefficients entiers, il en est de même des $n!\,H_n(X)$ pour $n \geq 0$.*

L'auteur exprime au *Referee* sa vive reconnaissance pour ses commentaires et le soin apporté à la relecture du manuscrit.

BIBLIOGRAPHIE

[BBEM] B. BEAUZAMY, E. BOMBIERI, P. ENFLO, H.L. MONTGOMERY, Products of polynomials in many variables, J. Number Th., 36, n° 2 (1990), 219–245.

[BG] J.P. BÉZIVIN, F. GRAMAIN, Solutions entières d'un système d'équations aux différences, Ann. Inst. Fourier, 43, n° 3 (1993), 791–814.

[FL] W. FULTON, S. LANG, Riemann–Roch Algebra, Springer–Verlag, 1985.

[H] P. HENRICI, Applied and Computational Complex Analysis, Tome 1, Wiley & Sons, 1974.

[L1] M. LANGEVIN, Polynômes et convexité sur la sphère de Riemann, dans «Cinquante ans de polynômes», Proceedings Paris, 1988; M. Langevin et M. Waldschmidt éd., Lecture Notes n° 1415, Springer–Verlag (1990), 138–159.

[L2] M. LANGEVIN, Du théorème de Bernstein au théorème de Walsh, Sém. de théorie des nombres de Paris 91/92, Prog. in Math., Birkhäuser (1993), 73–88.

[L3] M. LANGEVIN, Convexité sur la sphère de Riemann, C. R. Acad. Sci. Paris, 310, Sér. I (1990), 37–40.

[L4] M. LANGEVIN, Géométrie autour d'un théorème de Bernstein, Sém. de théorie des nombres de Paris 82/83, Prog. in Math., Birkhäuser (1984), 143–160.

[R] H. RADEMACHER, Topics in Analytic Number Theory, Springer–Verlag, 1973.

Manuscrit reçu le 16 mai 1995,
accepté le 26 avril 1996.

Michel LANGEVIN,
Institut de Mathématiques, A2X
Université de Bordeaux (BXI)
33405 Talence (France).
lgv@math.u–bordeaux.fr