JEAN DIEUDONNÉ

La théorie des invariants au XIXe siècle

Séminaire N. Bourbaki, 1971, exp. n° 395, p. 257-274

<http://www.numdam.org/item?id=SB_1970-1971__13__257_0>

Séminaire BOURBAKI
23e année, 1970/71, n° 395 Juin 1971

LA THÉORIE DES INVARIANTS AU XIX^e SIÈCLE

par Jean DIEUDONNÉ

Introduction

Vers le milieu du XVIII^e siècle, on savait que si, dans une forme quadratique $ax^2 + bxy + cy^2$, on fait un changement de variables $x = \alpha x' + \beta y'$, $y = \gamma x' + \delta y'$ avec $\alpha\delta - \beta\gamma \neq 0$, on obtient une forme quadratique $a'x'^2 + b'x'y' + c'y'^2$ et les discriminants des deux formes sont liés par la formule $b'^2 - 4a'c' = (b^2 - 4ac)(\alpha\delta - \beta\gamma)^2$; Gauss connaissait la formule analogue pour les formes quadratiques à trois variables, et un peu plus tard Gauchy démontra la formule générale reliant les discriminants de deux formes quadratiques quelconques déduites l'une de l'autre par changement de variables linéaires. La proportionnalité des discriminants correspond d'ailleurs au fait que les formes de discriminant nul sont exactement les formes dégénérées.

C'est en partant de ce cas particulier que Cayley, vers 1845, pose le problème général de la recherche des "invariants relatifs" d'un polynôme homogène de degré r à n variables $P(x) = \sum_{\alpha} a_{\alpha} x^{\alpha}$, où les coefficients a_{α} sont considérés comme des indéterminées ("forme n-aire de degré r", " n-ary quantic" dans la terminologie de Cayley). Il s'agit de trouver les polynômes homogènes $I((a_{\alpha}))$ en les coefficients a_{α} ayant la propriété suivante : pour toute transformation $s \in GL(n)$, soit $P(s.x) = \sum_{\alpha} a'_{\alpha} x^{\alpha}$; on doit alors avoir $I((a'_{\alpha})) = \chi(s) I((a_{\alpha}))$ avec un facteur ne dépendant que de s. Il est remarquable qu'en une vingtaine d'années les efforts de Cayley lui-même et Sylvester en Angleterre, de Hermite en France, de Clebsch, Gordan et Aronhold en Allemagne aient abouti à donner une solution théoriquement complète du problème général ainsi posé et même de problèmes beaucoup plus généraux, au moyen de ce qu'on appelait à l'époque la "méthode symbolique". On va exposer les grandes lignes de cette méthode, en suivant essentiellement les idées introduites vers la fin du siècle par Frobenius, A. Young et I. Schur, qui font apparaître la "méthode symbolique" comme un chapitre de la théorie des représentations linéaires des groupes classiques.

§ 1. Généralités

Lorsqu'un groupe Γ opère sur un ensemble E, il opère de façon évidente sur l'ensemble de parties $\underline{P}(E)$; lorsque Γ opère sur une famille d'ensembles (E_λ), il opère de façon évidente sur le produit de cette famille. De ces deux observations, on déduit que si Γ opère sur E et F, il opère canoniquement sur l'ensemble $\mathcal{F}(E,F)$ des applications de E dans F, en identifiant ces applications à leurs graphes (parties de $E \times F$) ; pour toute application $u : E \to F$, et tout $\sigma \in \Gamma$, $\sigma.u$ est l'application telle que $(\sigma.u)(x) = \sigma.(u(\sigma^{-1}.x))$.

Un <u>invariant</u> pour l'action d'un groupe Γ dans un ensemble E est un élément $x \in E$ tel que $\sigma.x = x$ pour tout $\sigma \in \Gamma$. Dans le cas vu plus haut de l'opération de Γ dans $\mathcal{F}(E,F)$, les invariants dans $\mathcal{F}(E,F)$ prennent le nom de <u>concomitants</u>: ce sont donc les applications de E dans F telles que $u(\sigma.x) = \sigma.u(x)$ pour $x \in E$ et $\sigma \in \Gamma$. Lorsque E est un produit $\prod_{k=1}^{r} E_k$ et que Γ agit sur chacun des E_k, cette équation s'écrit

$$u(\sigma.x_1, \dots, \sigma.x_r) = \sigma.u(x_1, \dots, x_r) ,$$

et on dit parfois que u est un concomitant simultané de r variables x_k .

On va ne considérer que le cas où E (ou les E_k) sont des espaces vectoriels sur un corps K , et où Γ opère linéairement dans E (ou les E_k) ; si μ_σ est l'application linéaire $x \mapsto \sigma.x$ de E dans lui-même, $\sigma \mapsto \mu_\sigma$ est donc une représentation linéaire de Γ dans E (homomorphisme de Γ dans $GL(E)$). Lorsque $E = K$, μ_σ est une homothétie de rapport $\chi(\sigma) \in K^*$, et χ est un homomorphisme de Γ dans K^*, qu'on appelle encore caractère abélien de Γ à valeurs dans K . Un invariant relatif de poids χ de Γ , défini dans E , est une application $u : E \to K$ telle que $u(\sigma.x) = \chi(\sigma)u(x)$ pour $x \in E$ et $\sigma \in \Gamma$, autrement dit un concomitant dans $\mathcal{F}(E,K)$ pour l'action donnée de Γ sur E et l'action de Γ sur K définie par le caractère χ . Lorsque $\chi = 1$, on parle d'invariant absolu.

§ 2. La linéarisation classique

On se borne à partir de maintenant au cas où K est de caractéristique 0 (il n'y a pas de théorie en caractéristique $p > 0$) et les espaces vectoriels sur K sont de dimension finie. En outre, on va seulement considérer des concomitants rationnels, c'est-à-dire des concomitants $u : E \to F$ tels que, pour deux bases (e_i) , (f_j) de E et F , on ait

$$u\left(\sum_{i=1}^{m} \xi_i e_i\right) = \sum_{j=1}^{n} u_j(\xi_1, \dots, \xi_m) f_j$$

où les u_j sont des fonctions rationnelles des ξ_i . Cette définition est indépendante des bases ; une telle fonction u n'est pas partout définie, et la relation $u(\sigma.x) = \sigma.u(x)$ doit donc être entendue comme ayant lieu seulement là où les deux membres sont définis. Mais en vertu du principe de prolongement des identités algébriques (le corps K étant infini), les relations algébriques entre les ξ_i que l'on obtient ainsi sont encore valables lorsqu'on remplace les ξ_i par des indéterminées.

On dira qu'un concomitant rationnel u est <u>polynomial</u> si les u_j sont des polynomes homogènes de degré m ; cela est encore indépendant des bases.

Cela étant, on a les réductions successives suivantes (à peu près triviales) :

A) Tout concomitant rationnel peut s'écrire $u = (p/q)v$, où $v : E \to F$ est un concomitant polynomial, et p,q des invariants relatifs polynomiaux.

B) Les parties homogènes d'un concomitant polynomial sont des concomitants.

C) Pour toute application polynomiale homogène $u : E \to F$ de degré r , on peut écrire, pour $\lambda_1,\dots,\lambda_r$ dans K , $x_1,\dots,x_r$ dans E

$$u(\lambda_1 x_1+\dots+\lambda_r x_r) = \sum_{|\alpha|=r} c_\alpha \lambda^\alpha u_\alpha(x_1,\dots,x_r)$$

où les c_α sont des entiers ; en particulier, si $v = u_{(1,1,\dots,1)}$, v est une application <u>multilinéaire</u> de E^r dans F , et on a

$$u(x) = (r!)^{-1} v(x,x,\dots,x)$$

("polarisation" de u). Cela étant, si u est un concomitant, v est un concomitant simultané de $x_1,\dots,x_r$, multilinéaire.

D) Si $\varphi : E^r \to E^{\otimes r}$ est l'application canonique, on peut écrire $v = w \circ \varphi$ où cette fois $w : E^{\otimes r} \to F$ est un concomitant <u>linéaire</u>.

E) Finalement, comme $\mathrm{Hom}(E^{\otimes r},F)$ s'identifie canoniquement à $E^{*\otimes r} \otimes F$, on est ramené à trouver les <u>invariants</u> pour l'action (linéaire) de Γ dans cet espace vectoriel déduite canoniquement des actions de Γ dans E et dans F .

Ce problème à son tour est résolu si, pour une action linéaire de Γ sur un espace vectoriel M, on connaît les sous-espaces de dimension 1 de M invariants par Γ. Autrement dit, on est ramené à l'étude des représentations linéaires du groupe Γ.

§ 3. Représentations linéaires rationnelles du groupe linéaire général

On va maintenant supposer que K est en outre algébriquement clos, et que $\Gamma = GL(n,K)$. Pour appliquer ce qui précède, on va déterminer explicitement, par la méthode de Frobenius-Young-I. Schur, les représentations linéaires rationnelles de $GL(n,K)$: on entend par là un homomorphisme $\underline{X} \mapsto \underline{F}(\underline{X})$ de $GL(n,K)$ dans un $GL(N,K)$, où les éléments de la matrice $\underline{F}(\underline{X})$ d'ordre N sont fonctions rationnelles des éléments de la matrice $\underline{X}$ d'ordre n (cela semble encore soulever la même difficulté que plus haut, la relation $\underline{F}(\underline{XY}) = \underline{F}(\underline{X})\underline{F}(\underline{Y})$ n'ayant lieu que pour les matrices $\underline{X},\underline{Y}$ telles que $\underline{F}$ soit définie pour $\underline{X},\underline{Y}$ et $\underline{XY}$; mais ici encore, cette relation équivaut à celle qu'on obtient en remplaçant les éléments de $\underline{X}$ et de $\underline{Y}$ par des indéterminées). On dira encore que $\underline{F}$ est polynomiale (resp. polynomiale homogène) si les éléments de $\underline{F}(\underline{X})$ sont des polynômes (resp. des polynômes homogènes de même degré) en les éléments de $\underline{X}$). On a encore les réductions suivantes (un peu moins triviales que ci-dessus) :

A) Si $\underline{X} \mapsto \underline{F}(\underline{X})$ est une représentation rationnelle de $GL(n,K)$, on peut écrire $\underline{F} = (p/q)\underline{G}$, où p et q sont des caractères abéliens polynomiaux et $\underline{G}$ une représentation polynomiale.

B) Toute représentation polynomiale $\underline{F}$ est semblable à une somme directe de représentations polynomiales homogènes $\underline{F}_j$ (on utilise la forme de Jordan d'une matrice).

C) Supposons donc que $\underline{F}$ soit une représentation polynomiale homogène de degré f de $GL(n,K)$ dans K^N. Soient x_{ij} $(1 \leq i,j \leq n)$ n^2 indéterminées sur K ; pour tout couple de multiindices $\alpha = (\alpha_1, \ldots, \alpha_f)$, $\beta = (\beta_1, \ldots, \beta_f)$, posons $x_{\alpha\beta} = x_{\alpha_1\beta_1} \ldots x_{\alpha_f\beta_f}$; on voit aisément que l'on a $x_{\alpha\beta} = x_{\alpha'\beta'}$ si et seulement s'il existe une permutation $\pi \in \mathfrak{S}_f$ telle que $\pi.\alpha = \alpha'$ et $\pi.\beta = \beta'$ ($\pi.\alpha$ est le multiindice $(\alpha_{\pi^{-1}(1)}, \ldots, \alpha_{\pi^{-1}(f)})$). Soit $\underline{X} = (x_{ij})$ et soient $F_{hk}(\underline{X})$ $(1 \leq h,k \leq N)$ les éléments de la matrice $\underline{F}(\underline{X})$; par hypothèse on peut écrire d'une seule manière $F_{hk}(\underline{X}) = \sum_{\alpha,\beta} a_{hk\alpha\beta} x_{\alpha\beta}$, si l'on impose la condition $a_{hk,\pi.\alpha,\pi.\beta} = a_{hk\alpha\beta}$ pour toute permutation $\pi \in \mathfrak{S}_f$. Or, les $x_{\alpha\beta}$ sont les éléments de la matrice $\underline{X}^{\otimes f}$.

On est donc conduit à introduire, dans l'anneau $\mathrm{End}((K^n)^{\otimes f})$, le sous-anneau A_f engendré par les matrices $\underline{U}^{\otimes f}$, où $\underline{U}$ parcourt $\mathrm{End}(K^n) = M_n(K)$ (ou seulement $GL(n,K)$) ; on montre aisément que c'est l'anneau des matrices $\underline{T} = (t_{\alpha\beta})$ telles que $t_{\pi.\alpha,\pi.\beta} = t_{\alpha\beta}$ pour $\pi \in \mathfrak{S}_f$. Revenant à la représentation $\underline{F}$, on pose $G_{hk}(\underline{T}) = \sum_{\alpha,\beta} a_{hk\alpha\beta} t_{\alpha\beta}$, de sorte que la matrice $\underline{G}(\underline{T}) = (G_{hk}(\underline{T}))$ est telle que $\underline{G}(\underline{X}^{\otimes f}) = \underline{F}(\underline{X})$, et on prouve aisément que $\underline{G}$ est un <u>homomorphisme de l'algèbre</u> A_f <u>dans l'algèbre</u> $M_N(K)$.

D) On est maintenant en état de faire donner la grosse artillerie des algèbres semi-simples. Soit $(e_j)_{1\leq j\leq n}$ la base canonique de K^n ; la base canonique de $(K^n)^{\otimes f}$ est formée des $e_\alpha = e_{\alpha_1} \otimes e_{\alpha_2} \otimes \ldots \otimes e_{\alpha_f}$, et le groupe $\mathfrak{S}_f$ opère linéairement sur $(K^n)^{\otimes f}$ par $\pi.e_\alpha = e_{\pi.\alpha}$; on en déduit une action de l'algèbre de groupe $K[\mathfrak{S}_f]$ sur $(K^n)^{\otimes f}$. Cela étant, on voit aussitôt que l'algèbre A_f est la <u>commutante</u> dans $\mathrm{End}((K^n)^{\otimes f})$ de l'image de $K[\mathfrak{S}_f]$ pour cette action. Or, par le th. de Maschke, $K[\mathfrak{S}_f]$ est <u>semi-simple</u>, donc aussi tous ses quotients. En second lieu, par le th. de commutation de I. Schur, la commutante dans $\mathrm{End}((K^n)^{\otimes f})$ d'une sous-algèbre semi-simple contenant le centre (i.e. les matrices scalaires) est une sous-algèbre semi-simple. d'où tout d'abord le

THÉORÈME 1. Toute représentation linéaire rationnelle de $GL(n,K)$ est complètement réductible.

Mais en fait la méthode de démonstration donne beaucoup plus que ce théorème qualitatif, comme on va le voir.

§ 4. Les diagrammes d'Young

Le th. de commutation de Schur affirme, non seulement la semi-simplicité de la commutante B' d'une sous-algèbre semi-simple B de $M_p(K)$ contenant le centre, mais encore décrit explicitement les B'-modules simples (ou, ce qui revient au même, les idéaux minimaux à gauche de B') à l'aide des idéaux à gauche minimaux de B : on les obtient en prenant un générateur c d'un idéal à gauche minimal de B, et on forme le B'-module $c.K^p$.

Pour avoir les A_f-modules simples, qui correspondent aux représentations irréductibles de $GL(n,K)$, on doit donc commencer par déterminer les idéaux minimaux de l'algèbre de groupe $B = K[\mathfrak{S}_f]$. C'est ce qu'on peut faire explicitement à l'aide des diagrammes d'Young. On définit un tel diagramme à l'aide d'une suite décroissante $\alpha = (\alpha_1,\ldots,\alpha_r)$ d'entiers tels que

$$\alpha_1 \geqq \alpha_2 \geqq \ldots \geqq \alpha_r \geqq 1 , \quad \text{et} \quad \alpha_1 + \alpha_2 + \ldots + \alpha_r = f ;$$

un diagramme d'Young Σ_α d'indice α est une suite double d'entiers (m_{i,j_i}) où $1 \leqq i \leqq r$, $1 \leqq j_i \leqq \alpha_i$, les m_{i,j_i} étant deux à deux distincts et prenant des valeurs entre 1 et f ; on représente Σ_α comme un tableau rectangulaire où les entiers entre 1 et f sont arrangés en r lignes de longueurs $\alpha_1, \alpha_2, \dots, \alpha_r$:

$$\begin{array}{l} m_{11} \quad m_{12} \cdots\cdots\cdots m_{1\alpha_1} \\ m_{21} \quad m_{22} \cdots\cdots m_{2\alpha_2} \\ \cdots\cdots\cdots\cdots \\ m_{r1} \cdots\cdots m_{r\alpha_r} \; . \end{array}$$

Le groupe symétrique $\mathfrak{S}_f$ agit sur l'ensemble des diagrammes d'Young : $\pi.\Sigma_\alpha$ est le diagramme (m'_{ij_i}) correspondant à la <u>même</u> suite α , tel que $m'_{ij_i} = \pi(m_{ij_i})$.

Un diagramme Σ_α détermine deux sous-groupes de $\mathfrak{S}_f$; le premier, $\underline{L}(\Sigma_\alpha)$ est formé des $\pi\in\mathfrak{S}_f$ qui laissent invariante <u>chaque colonne</u>. Il est clair que $\underline{L}(\Sigma_\alpha) \cap \underline{C}(\Sigma_\alpha) = \{1\}$, et que pour toute permutation $\pi\in\mathfrak{S}_f$,

$$\underline{L}(\pi.\Sigma_\alpha) = \pi.\underline{L}(\Sigma_\alpha).\pi^{-1} \quad \text{et} \quad \underline{C}(\pi.\Sigma_\alpha) = \pi.\underline{C}(\Sigma_\alpha).\pi^{-1} \; .$$

A l'aide de ces groupes, on forme dans l'algèbre $B = K[\mathfrak{S}_f]$ les éléments suivants :

$$a_\alpha = \sum_{p\in\underline{L}(\Sigma_\alpha)} p \; , \qquad b_\alpha = \sum_{q\in\underline{C}(\Sigma_\alpha)} \varepsilon_q . q \; , \qquad c_\alpha = a_\alpha b_\alpha$$

et on démontre que $c_\alpha \neq 0$ et que $B.c_\alpha$ est un <u>idéal à gauche minimal</u> de B ; en outre, tout idéal à gauche minimal de B est isomorphe à un des idéaux $B.c_\alpha$, et les idéaux correspondant à Σ_α et Σ'_β sont isomorphes si et seulement si $\alpha = \beta$.

On peut alors revenir au problème considéré dans le § 3, la détermination des A_f-modules simples : ce sont les sous-A_f-modules $c_\alpha.(K^n)^{\otimes f}$ qui ne sont pas réduits à 0, correspondant aux diagrammes d'Young Σ_α. Deux tels sous-espaces de $(K^n)^{\otimes f}$ correspondant à deux diagrammes Σ_α et Σ'_α de même indice α sont des A_f-modules simples isomorphes ; les sous-espaces des $(K^n)^{\otimes f}$ obtenus ainsi (pour toutes les valeurs de f) sont ce qu'on appelle les <u>espaces de tenseurs irréductibles</u> sur K^n.

Si on regarde les définitions de a_α et b_α, on voit que cela donne sur (K^n) des opérations de <u>symétrisation</u> et d'<u>antisymétrisation partielles</u>. On voit alors aisément que :

1° pour que $c_\alpha.(K^n)^{\otimes f} = \{0\}$, il faut et il suffit que la première colonne du tableau α ait une longueur $> n$;

2° pour que $c_\alpha.(K^n)^{\otimes f}$ soit de dimension 1, il faut et il suffit que <u>toutes</u> les colonnes du tableau α aient pour longueur n ; cela entraîne évidemment que $f = gn$, et le caractère abélien de $GL(n,K)$ qui lui correspond est $\underline{X} \mapsto (\det \underline{X})^g$.

§ 5. <u>La méthode symbolique pour les invariants de tenseurs contravariants</u>

Ce dernier résultat fournit déjà les <u>invariants relatifs simultanés</u> multilinéaires d'un certain nombre f de vecteurs de K^n (pour le groupe $GL(n,K)$).

THÉORÈME 2 (<u>premier théorème fondamental de la théorie des invariants</u>). <u>Il n'existe d'invariants relatifs</u> multilinéaires <u>de</u> f <u>vecteurs</u> x_j $(1 \leq j \leq f)$ <u>de</u> K^n <u>que si</u> f <u>est un multiple</u> gn <u>de</u> n. <u>Ces invariants sont les combinaisons linéaires des invariants de la forme</u>

(1) $$[x_{i_1}\dots x_{i_n}][x_{i_{n+1}}\dots x_{i_{2n}}]\dots[x_{i_{f-n+1}}\dots x_{i_f}]$$

<u>où</u> $(i_1,\dots,i_f)$ <u>est une permutation arbitraire de</u> $\mathfrak{S}_f$ <u>et</u> $[z_1 z_2 \dots z_n]$ <u>est</u>

<u>le déterminant de la matrice dont les colonnes sont les vecteurs</u> z_i . <u>Ces invariants sont tous de poids</u> g .

Il n'y a donc pas d'invariant absolu multilinéaire ; mais on obtient aisément à partir de là des invariants absolus rationnels, par exemple pour $n = 2$ et $f = 4$, le <u>birapport</u> $[x_1x_2][x_3x_4]/[x_1x_3][x_2x_4]$.

A partir de là, en reprenant le processus de linéarisation du § 2, on obtient la règle donnant tous les <u>invariants relatifs polynomiaux homogènes</u> d'un certain nombre de tenseurs. Pour fixer les idées, supposons qu'on cherche les invariants relatifs de trois tenseurs contravariants

$$x \in (K^n)^{\otimes p} \quad , \quad x' \in (K^n)^{\otimes q} \quad , \quad x'' \in (K^n)^{\otimes r}$$

dont on désignera les composantes par

$$\zeta^{k_1k_2\ldots k_p} \quad , \quad \zeta'^{k_1k_2\ldots k_q} \quad , \quad \zeta''^{k_1k_2\ldots k_r} \qquad (1 \leqq k_i \leqq n) \; .$$

On considère les invariants polynomiaux en x,x',x'' qui sont homogènes de degré h par rapport aux composantes de x , de degré h' par rapport à celles de x', de degré h'' par rapport à celles de x'' .

On commence par rechercher les invariants relatifs <u>multilinéaires</u> par rapport à h tenseurs $\overset{j}{x}$ $(1 \leqq j \leqq h)$ de $(K^n)^{\otimes p}$, h' tenseurs $\overset{j'}{x'}$ $(1 \leqq j' \leqq h')$ de $(K^n)^{\otimes q}$, h'' tenseurs $\overset{j''}{x''}$ $(1 \leqq j'' \leqq h'')$ de $(K^n)^{\otimes r}$; les composantes se noteront comme celles de x,x',x'' en insérant un indice inférieur j,j' ou j'' sous ζ,ζ',ζ'' .

Puis, par définition du produit tensoriel, il revient au même de chercher les invariants relatifs multilinéaires de $hp + h'q + h''r$ <u>vecteurs de</u> K^n, qu'on désigne par

$$\overset{j\ell}{x} \quad , \qquad \overset{j'\ell'}{x'} \quad , \qquad \overset{j''\ell''}{x''}$$

avec $1 \leq j \leq h$, $1 \leq \ell \leq p$, $1 \leq j' \leq h'$, $1 \leq \ell' \leq q$, $1 \leq j'' \leq h''$, $1 \leq \ell'' \leq r$; on pose

$$\overset{j\ell}{x} = \sum_{k=1}^{n} \xi_{j}^{\ell k} e_k \ , \qquad \overset{j'\ell'}{x'} = \sum_{k=1}^{n} \xi'^{\,\ell' k}_{j'} e_k \ , \qquad \overset{j''\ell''}{x''} = \sum_{k=1}^{n} \xi''^{\,\ell'' k}_{j''} e_k \ .$$

On peut alors appliquer le th. 2 ; il n'y a d'invariant relatif du type voulu que si

$$(2) \qquad hp + h'q + h''r = gn$$

et les invariants relatifs des gn vecteurs précédents sont combinaisons linéaires de ceux du type (1). Il s'agit ensuite de revenir au problème posé : dans l'expression de la fonction multilinéaire obtenue, on remplace

$$\xi_{j}^{1k_1} \ \xi_{j}^{2k_2} \cdots \xi_{j}^{pk_p} \quad \text{par} \quad \zeta_{j}^{k_1k_2\ldots k_p}$$

$$\xi'^{\,1k_1}_{j'} \ \xi'^{\,2k_2}_{j'} \cdots \xi'^{\,qk_q}_{j'} \quad \text{par} \quad \zeta'^{\,k_1k_2\ldots k_q}_{j'}$$

$$\xi''^{\,1k_1}_{j''} \ \xi''^{\,2k_2}_{j''} \cdots \xi''^{\,rk_r}_{j''} \quad \text{par} \quad \zeta''^{\,k_1k_2\ldots k_r}_{j''}$$

puis on supprime les indices inférieurs j, j', j'', ce qui donne les invariants polynomiaux cherchés.

Pour avoir un exemple simple, cherchons les invariants de degré 2 d'un seul tenseur $x \in (K^n)^{\otimes 2}$, avec $n = 2$; on a alors $p = 2$, $h = 2$, et la relation (2) est vérifiée avec $g = 2$. On doit donc introduire 4 vecteurs de K^2,

$\overset{rs}{x} = \underset{r}{\xi}^{s1} e_1 + \underset{r}{\xi}^{s2} e_2$. Partons par exemple de l'invariant multilinéaire

$$[\overset{11}{x}\ \overset{22}{x}][\overset{12}{x}\ \overset{21}{x}] = (\underset{1}{\xi}^{11}\underset{2}{\xi}^{22} - \underset{1}{\xi}^{12}\underset{2}{\xi}^{21})(\underset{1}{\xi}^{21}\underset{2}{\xi}^{12} - \underset{1}{\xi}^{22}\underset{2}{\xi}^{11}).$$

On en tire d'abord l'invariant multilinéaire

$$\underset{1}{\zeta}^{11}\underset{2}{\zeta}^{22} - \underset{1}{\zeta}^{12}\underset{2}{\zeta}^{12} - \underset{1}{\zeta}^{21}\underset{2}{\zeta}^{21} + \underset{1}{\zeta}^{22}\underset{2}{\zeta}^{11}$$

de $\overset{1}{x}$ et $\overset{2}{x}$, puis l'invariant homogène de degré 2

$$2\zeta^{11}\zeta^{22} - (\zeta^{12})^2 - (\zeta^{21})^2$$

de x . Si on part de $[\overset{11}{x}\ \overset{12}{x}][\overset{21}{x}\ \overset{22}{x}]$, on obtient $(\zeta^{12} - \zeta^{21})^2$ et si on part de $[\overset{11}{x}\ \overset{21}{x}][\overset{12}{x}\ \overset{22}{x}]$, on obtient $2(\zeta^{11}\zeta^{22} - \zeta^{12}\zeta^{21})$; tout invariant homogène de degré 2 est donc combinaison linéaire de ces trois invariants qui ne sont pas linéairement indépendants d'ailleurs.

§ 6. Extension de la méthode aux tenseurs mixtes

En pratique, on n'a pas seulement à chercher les invariants simultanés de tenseurs contravariants, mais aussi de tenseurs mixtes. Pour cela on est ramené par les procédés du § 2, à déterminer d'abord les invariants <u>multilinéaires</u> d'un certain nombre de vecteurs $\overset{1}{x},\dots,\overset{p}{x}$ de K^n et d'un certain nombre de vecteurs $\underset{1}{y'},\dots,\underset{q}{y'}$ du dual $(K^n)^*$. On utilise le fait que la représentation linéaire naturelle de $GL(n,K)$ dans $(K^*)^n$ est équivalente à la <u>puissance extérieure</u> $(n-1)$-<u>ème</u> de la représentation naturelle de $GL(n,K)$ dans K^n , à l'aide de l'isomorphisme (canonique à facteur scalaire près) entre vecteurs covariants et $(n-1)$-vecteurs contravariants. Des manipulations élémentaires d'algèbre multilinéaire conduisent alors au

THÉORÈME 3 (<u>premier théorème fondamental de la théorie des invariants, seconde version</u>). <u>Les invariants multilinéaires des</u> $\overset{i}{x}$ $(1 \leq i \leq p)$ <u>et des</u> y'_j $(1 \leq j \leq q)$ <u>sont des combinaisons linéaires de produits d'invariants simples de trois types</u> :

(i) <u>crochets</u> $[\overset{i_1}{x}\ \overset{i_2}{x} \dots \overset{i_n}{x}]$ <u>de poids</u> 1 ;

(ii) <u>crochets</u> $[y'_{j_1}\ y'_{j_2} \dots y'_{j_n}]$ <u>de poids</u> -1 ;

(iii) <u>produits scalaires</u> $\langle \overset{i}{x}, y'_j \rangle$ <u>de poids</u> 0 .

<u>Cela entraîne que</u> $p-q$ <u>doit être multiple</u> (<u>positif ou négatif</u>) <u>de</u> n .

A partir de ce résultat, on procède comme dans le § 5 pour obtenir tous les invariants polynomiaux homogènes simultanés d'un certain nombre de tenseurs mixtes de types donnés. Au lieu de tenseurs quelconques d'un espace $T^r_s(K^n)$, on peut aussi considérer les tenseurs d'un des sous-espaces de tenseurs <u>irréductibles</u> en lesquels se décompose cet espace : si π est le projecteur canonique de $T^r_s(K^n)$ sur ce sous-espace, il suffit de remplacer dans l'invariant obtenu, le tenseur $x \in T^r_s(K^n)$ par $\pi(x)$.

Un cas particulier de ce procédé donne la solution générale du problème de Cayley, les "formes n-aires de degré r " n'étant autres que les tenseurs <u>symétriques</u> dans $T^o_r(K^n) = ((K^n)^*)^{\otimes r}$; ces derniers forment bien un espace de tenseurs irréductibles, correspondant au tableau d'Young formé d'une seule ligne. L'application du th. 3 à ce cas donne alors la "règle d'Aronhold" que nous énoncerons, pour simplifier, dans un cas particulier. On cherche les invariants polynomiaux homogènes simultanés de deux formes n-aires f,g de degrés respectifs r,s , et d'un vecteur $x \in K^n$; si un tel invariant est de degré p par rapport aux coefficients de f , de degré q par rapport aux coefficients de g, et de degré m par rapport aux coordonnées de x, on a

$$(3) \qquad m - pr - qs = gn$$

où g (entier positif ou négatif) est le <u>poids</u> de l'invariant. On considère alors $p+q$ vecteurs covariants y'_i, z'_j ($1 \leq i \leq p$, $1 \leq j \leq q$) et on forme un produit de crochets $[w'_1\, w'_2 \,\ldots\, w'_n]$, où chaque w'_k est un y'_i <u>ou</u> un z'_j, et de produits scalaires $< x, y'_i >$ ou $< x, z'_j >$, de sorte que ψ soit de degré r par rapport à chacun des y'_i, de degré s par rapport à chacun des z'_j, et de degré m par rapport à x. On obtient un invariant φ du type voulu (qui peut éventuellement être identiquement nul !) en remplaçant dans ψ chaque monôme $(\overset{i}{\eta}_1)^{r_1} \ldots (\overset{i}{\eta}_n)^{r_n}$ de degré $r = r_1 + r_2 + \ldots + r_n$ en les coordonnées de y'_i par

$$\frac{r_1! \ldots r_n!}{r!} a_{r_1 r_2 \ldots r_n} \quad \text{si} \quad f(T) = \Sigma\, a_{r_1 r_2 \ldots r_n} T_1^{r_1} \ldots T_n^{r_n} ,$$

pour $1 \leq i \leq p$,

et de même pour les monômes en les coordonnées des z'_j, en remplaçant f par g. On dit que ψ est l'<u>expression symbolique</u> de φ.

<u>Exemples</u>. 1) $f(x)$ est un invariant (absolu) de f et de x, dont l'expression symbolique est $< x, y'_1 >^r$.

2) Le discriminant d'une forme n-aire quadratique a pour expression symbolique $n^{-n}[y'_1\, y'_2 \,\ldots\, y'_n]^2$.

3) Prenons $r = 2$, $s = 1$, $m = 0$ (invariants d'une forme quadratique et d'une forme linéaire) ; on satisfait à (3) en prenant $p = n-1$, $q = 2$, $g = -2$; on obtient ainsi un invariant dont l'expression symbolique est

$$[y'_1\, y'_2 \dots y'_{n-1}\; z'_1][y'_1\, y'_2 \dots y'_{n-1}\; z'_2] .$$

Lorsqu'on annule cet invariant, on obtient la condition pour que dans $P_{n-1}(K)$ l'hyperplan défini par la forme linéaire soit tangent à la quadrique définie par la forme quadratique.

4) Le hessien d'une forme n-aire de degré r a pour expression symbolique, à un coefficient numérique près,

$$[y'_1\, y'_2 \,\dots\, y'_n]^2 < x,y'_1 >^{r-2} \dots < x,y'_n >^{r-2} .$$

5) Supposons $n = 2$ et cherchons <u>tous</u> les invariants simultanés de deux formes quadratiques

$$f = \alpha_{11}(\xi^1)^2 + 2\alpha_{12}\xi^1\xi^2 + \alpha_{22}(\xi^2)^2 \;, \qquad g = \beta_{11}(\xi^1)^2 + 2\beta_{12}\xi^1\xi^2 + \beta_{22}(\xi^2)^2$$

et d'un vecteur $x = (\xi^1,\xi^2) \in K^2$. La règle d'Aronhold prescrit de considérer les produits dont les facteurs sont de l'une des formes

$$[y'_i\, y'_h] \;,\; [z'_j\, z'_k] \;,\; [y'_i\, z'_j] \;,\; < x,y'_i > \;,\; < x,z'_j >$$

et qui sont de degré 2 par rapport à chacun des y'_i et des z'_j . Il est assez facile d'examiner les divers cas possibles, et on arrive à la conclusion suivante : les invariants cherchés sont des polynômes par rapport à 6 invariants "fondamentaux" :

f d'expression symbolique $< x,y'_1 >^2$;

g " " $< x,z'_1 >^2$;

$$h = (\alpha_{11}\beta_{12} - \alpha_{12}\beta_{11})(\xi^1)^2 + (\alpha_{11}\beta_{22} - \alpha_{22}\beta_{11})\xi^1\xi^2 + (\alpha_{12}\beta_{22} - \alpha_{22}\beta_{12})(\xi^2)^2$$

d'expression symbolique $[y'_1\, z'_1] < x,y'_1 > < x,z'_1 >$

$d_1 = \alpha_{11}\alpha_{22} - (\alpha_{12})^2$ d'expression symbolique $[y'_1\ y'_2]^2$;

$d_2 = \beta_{11}\beta_{22} - (\beta_{12})^2$ " " $[z'_1\ z'_2]^2$;

$d_{12} = \alpha_{11}\beta_{22} + \alpha_{22}\beta_{11} - 2\alpha_{12}\beta_{12}$ d'expression symbolique $[y'_1\ z'_1]^2$.

§ 7. Invariants et géométrie projective

Si l'on adopte la définition de Klein, la géométrie projective "élémentaire" est l'étude des propriétés algébriques d'objets liés à $P_{n-1}(K)$ et invariantes par le groupe projectif. Si on veut préciser cela, on doit appeler "propriété algébrique" une propriété concernant un certain nombre de tenseurs mixtes $u_1, u_2, \dots, u_r$ sur K^n et qui s'exprime en annulant un certain nombre de polynômes homogènes $P_h(u_1, \dots, u_r)$ par rapport aux composantes des u_j ; en outre, l'invariance par le groupe projectif signifie que pour tout $\sigma \in GL(n,K)$, le système des relations $P_h(u_1, \dots, u_r) = 0$ $(1 \leqq h \leqq s)$ entraîne $P_h(\sigma.u_1, \dots, \sigma.u_r) = 0$ $(1 \leqq h \leqq s)$; on dit que ce système de relations a une signification invariante.

Cela étant, la détermination de tous ces systèmes revient à celle des relations algébriques (ou "syzygies") entre invariants (rationnels) absolus de $GL(n,K)$ (théorème de Gram). Pour le prouver, soit $(e_i)_{1 \leq i \leq n}$ la base canonique de K^n , et pour toute base $y = (y_1, \dots, y_n)$ de K^n , notons σ_y la bijection linéaire telle que $\sigma_y(e_i) = y_i$. Alors, pour $u \in T^p_q(K^n)$, le tenseur $\sigma_y^{-1}.u$ a des composantes sur la base canonique qui sont des invariants rationnels absolus de $u_1, \dots, u_r$, $y_1, \dots, y_n$. En effet, pour $\tau \in GL(n,K)$, on a $\sigma_{\tau.y} = \tau\sigma_y$, donc $\sigma_{\tau.y}^{-1} = \sigma_y^{-1}\tau^{-1}$, donc $\sigma_{\tau.y}^{-1}.(\tau.u) = \sigma_y^{-1}.u$.

Cela étant, supposons que les relations $P_h(u_1,\dots,u_r) = 0$ $(1 \leqq h \leqq s)$ entraînent, pour tout $\sigma \in GL(n,K)$, $P_h(\sigma.u_1,\dots,\sigma.u_r) = 0$ $(1 \leqq h \leqq s)$; pour toute base $y = (y_1,\dots,y_n)$ de K^n, les relations $P_h(u_1,\dots,u_r) = 0$ entraînent donc $P_h(\sigma_y.u_1,\dots,\sigma_y.u_r) = 0$ $(1 \leqq h \leqq s)$, relations qui s'écrivent $Q_h(u_1,\dots,u_r, y_1,\dots,y_n) = 0$, où les Q_h sont des polynômes par rapport à des invariants absolus de $u_1,\dots,u_r$, $y_1,\dots,y_n$. Inversement, si ces relations ont lieu pour $1 \leqq h \leqq s$, et pour tout choix de $y_1,\dots,y_n$ formant une base, on peut y faire $y_j = e_j$ pour $1 \leqq j \leqq n$, et on retrouve $P_h(u_1,\dots,u_r) = 0$.