V. SMIRNOV

COURS DE MATHÉMATIQUES SUPÉRIEURES

TOMO III

première partie

ÉDITIONS
MIR

В. И. СМИРНОВ

КУРС ВЫСШЕЙ МАТЕМАТИКИ

ТОМ III

часть первая

ИЗДАТЕЛЬСТВО «НАУКА»
МОСКВА

V. SMIRNOV

COURS de MATHÉMATIQUES SUPÉRIEURES

tome III

première partie

ÉDITIONS MIR • MOSCOU • 1970

CDU 510(022)=40

Traduit du russe

На французском языке

Table des matières

CHAPITRE III. ÉLÉMENTS DE LA THÉORIE DES GROUPES ET REPRÉSENTATIONS LINÉAIRES DES GROUPES

AVANT-PROPOS

La division du troisième tome du Cours de mathématiques supérieures de l'académicien Vladimir Smirnov en deux parties a été dictée par l'apport de compléments à l'une des dernières éditions de ce cours en cinq tomes.

La première partie du troisième tome contient les éléments de l'algèbre linéaire, de la théorie des formes quadratiques et de la théorie des groupes. Dans cette partie, les compléments les plus essentiels se rapportent à la théorie des groupes et, notamment, à la théorie des représentations linéaires des groupes. De plus, on y donne un exposé rigoureux de certains renseignements sur l'espace des suites l_2 et sur son lien avec l'espace L_2, ce qui est conditionné par l'exposé de la théorie des intégrales de Lebesgue et des propriétés de l'espace des fonctions L_2 que l'on trouve dans le deuxième tome.

Cet ouvrage que l'on propose à l'attention des lecteurs français se distingue par sa simplicité, par son accessibilité à un large cercle de personnes qui étudient les mathématiques supérieures. Il donne des éléments de la théorie des groupes, chapitre important des mathématiques supérieures, qui a de nombreuses applications en physique moderne.

Chapitre premier

DÉTERMINANTS ET RÉSOLUTION DES SYSTÈMES D'ÉQUATIONS LINÉAIRES

I-1. Déterminants et leurs propriétés

I-1-1. Notion de déterminant. Dans ce paragraphe, nous nous proposons de traiter un problème simple d'algèbre: celui de la résolution des systèmes d'équations du premier degré. La considération de ce problème nous conduira à la notion importante de déterminant.

Soit au préalable quelques cas particuliers très simples. Considérons d'abord un système de deux équations à deux inconnues:

$$a_{11}x_1 + a_{12}x_2 = b_1,$$

$$a_{21}x_1 + a_{22}x_2 = b_2.$$

Adoptons pour les coefficients a_{ik} des inconnues une notation à deux indices, le premier indiquant le numéro de l'équation qui contient ce coefficient et le second l'indice de l'inconnue.

Comme on le sait, la solution de ce système est de la forme:

$$x_1 = \frac{b_1a_{22} - a_{12}b_2}{a_{11}a_{22} - a_{12}a_{21}}; \qquad x_2 = \frac{a_{11}b_2 - b_1a_{21}}{a_{11}a_{22} - a_{12}a_{21}}.$$

Soit à présent un système de trois équations à trois inconnues:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 &= b_1, \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 &= b_2, \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 &= b_3. \end{aligned}$$

Notons que l'on conserve pour les coefficients la convention précédente sur les indices. Mettons les deux premières équations sous la forme:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 &= b_1 - a_{13}x_3, \\ a_{21}x_1 + a_{22}x_2 &= b_2 - a_{23}x_3. \end{aligned}$$

En les résolvant par rapport aux inconnues x_1 et x_2, nous obtenons d'après les formules précédentes:

$$x_1 = \frac{(b_1 - a_{13}x_3)\,a_{22} - a_{12}\,(b_2 - a_{23}x_3)}{a_{11}a_{22} - a_{12}a_{21}};$$

$$x_2 = \frac{a_{11}\,(b_2 - a_{23}x_3) - (b_1 - a_{13}x_3)\,a_{21}}{a_{11}a_{22} - a_{12}a_{21}}.$$

En substituant ces expressions dans la dernière équation du système, nous avons une équation qui nous permet de déterminer l'inconnue x_3:

$$x_3 = \frac{a_{11}a_{22}b_3 + a_{12}b_2a_{31} + b_1a_{21}a_{32} - a_{11}b_2a_{32} - a_{12}a_{21}b_3 - b_1a_{22}a_{31}}{a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31}}. \quad (1)$$

Analysons la composition de cette expression. Remarquons d'abord que son numérateur s'obtient à partir du dénominateur en substituant, purement et simplement, aux coefficients a_{i3} de l'inconnue les b_i des seconds membres du système. Il ne reste donc plus qu'à trouver la loi de formation du dénominateur qui ne contient aucun des seconds membres du système et est uniquement constitué des coefficients des inconnues. Plaçons ces coefficients dans un tableau carré, tout en conservant l'ordre qu'ils possèdent dans le système:

$$\begin{Vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{Vmatrix}. \quad (2)$$

Le tableau ci-dessus est constitué de trois lignes et de trois colonnes. Les coefficients a_{ik} sont appelés éléments du tableau. Le premier indice indique le numéro de la ligne à laquelle appartient cet élément, le second indice le numéro de la colonne. Ecrivons à présent le dénominateur de l'expression (1):

$$a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31}. \quad (3)$$

Il est la somme algébrique de six termes dont chacun est le produit de trois éléments du tableau (2). Chaque produit contient un élément de chaque ligne et de chaque colonne. En effet, ces produits sont de la forme:

$$a_{1p}a_{2q}a_{3r}, \quad (4)$$

où p, q, r sont les nombres entiers 1, 2, 3 disposés dans un ordre déterminé. Ainsi, tant l'ensemble des premiers indices que celui des seconds sont formés de l'ensemble des nombres entiers 1, 2, 3. Pour obtenir tous les termes de l'expression (3), il faut dans le produit (4) prendre les seconds indices p, q, r dans tous les ordres différents possibles. De telles permutations possibles sont évidemment au nombre de six:

$$1, 2, 3; \quad 2, 3, 1; \quad 3, 1, 2; \quad 1, 3, 2; \quad 2, 1, 3; \quad 3, 2, 1, \tag{5}$$

et nous obtenons ainsi les six termes de l'expression (3). Remarquons maintenant que dans l'expression (3) certains des produits (4) sont précédés du signe plus, d'autres du signe moins. Il nous faut donc trouver la loi suivant laquelle on leur affecte ce signe. Comme nous le voyons, les produits (4) précédés du signe plus ont pour seconds indices les permutations suivantes:

$$1, 2, 3; \quad 2, 3, 1; \quad 3, 1, 2, \tag{5_1}$$

et ceux précédés du signe moins ont pour seconds indices les permutations:

$$1, 3, 2; \quad 2, 1, 3; \quad 3, 2, 1. \tag{5_2}$$

Déterminons à présent les lois de formation des permutations (5_1) et des permutations (5_2). Appelons *inversion* dans une permutation le fait qu'un nombre de valeur plus élevée en précède un autre et calculons le nombre d'inversions dans les permutations (5_1). Il n'y a pas d'inversions dans la première permutation. Nous voyons qu'il y a dans la seconde permutation deux inversions provenant de ce que le nombre 1 se trouve précédé du nombre 2 et du nombre 3. De la même façon, la troisième permutation (5_1) présente deux inversions. On peut donc dire que toutes les permutations (5_1) présentent un nombre pair d'inversions. En étudiant de même les permutations (5_2), nous voyons qu'elles présentent chacune un nombre impair d'inversions. Nous pouvons maintenant énoncer la règle des signes dans l'expression (3): les produits (4), pour lesquels le nombre d'inversions dans les permutations des seconds indices est un nombre pair, sont précédés du signe plus dans l'expression (3). Les produits (4), pour lesquels les permutations des seconds indices ont un nombre impair d'inversions, sont précédés du signe moins dans l'expression (3). L'expression (3) est appelée *déterminant du troisième ordre* associé au tableau des nombres (2). Il est aisé de généraliser ce qui vient d'être exposé au cas des déterminants d'ordre quelconque.

Soient donnés n^2 nombres disposés sous la forme d'un tableau carré à n lignes et n colonnes:

$$\left\| \begin{matrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{matrix} \right\| . \tag{6}$$

Les éléments a_{ik} de ce tableau sont des nombres complexes donnés, de plus les indices i et k indiquent respectivement les numéros de la ligne et de la colonne à l'intersection desquelles se trouve le nombre a_{ik}. Formons tous les produits possibles à l'aide des éléments du tableau (6) de telle sorte que ces produits contiennent un élément de chaque ligne et de chaque colonne. Ces produits sont de la forme :

$$a_{1p_1}a_{2p_2}\ldots a_{np_n}, \tag{7}$$

où $p_1, p_2, \ldots, p_n$ sont les nombres $1, 2, \ldots, n$ disposés dans un certain ordre. Pour obtenir tous les produits possibles de la forme (7), il nous faut effectuer toutes les permutations possibles des seconds indices. Comme nous l'indique l'algèbre élémentaire, le nombre de ces permutations distinctes est égal à la factorielle n :

$$1\cdot 2\cdot 3\ldots n = n!$$

Chacune de ces permutations présente un certain nombre d'inversions par rapport à la permutation fondamentale :

$$1, 2, 3, \ldots, n.$$

Les produits de la forme (7), dont les seconds indices engendrent des permutations telles que le nombre d'inversions soit pair, sont précédés du signe plus, tandis que les produits de la forme (7), dont les seconds indices engendrent des permutations telles que le nombre d'inversions soit impair, sont précédés du signe moins. La somme algébrique de tous les produits ainsi obtenus est appelée *déterminant d'ordre n* associé au tableau (6). Cette somme contient évidemment $n!$ termes. Il est aisé d'exprimer cette définition sous la forme d'une formule. Introduisons pour cela la notation suivante. Soit p_1, $p_2, \ldots, p_n$ une des permutations des nombres $1, 2, \ldots, n$. Désignons le nombre d'inversions que présente cette permutation par le symbole :

$$[p_1, p_2, \ldots, p_n].$$

Alors, la définition du déterminant associé au tableau (6) donnée ci-dessus peut être transcrite sous la forme de la formule suivante :

$$\begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix} =$$

$$= \sum_{(p_1, p_2, \ldots, p_n)} (-1)^{[p_1, p_2, \ldots, p_n]} a_{1p_1}a_{2p_2}\ldots a_{np_n}. \tag{8}$$

Pour noter le déterminant, nous écrivons le tableau (6) entre deux traits verticaux.

La sommation porte ici sur toutes les permutations possibles des seconds indices $(p_1, p_2, \ldots, p_n)$. En parlant du tableau comme tel, et non du déterminant qui lui est associé, nous l'écrivons entre deux doubles traits verticaux.

Remarquons que dans l'expression (3) nous avons disposé les facteurs de chaque produit dans un ordre tel que les premiers indices donnent la permutation fondamentale 1, 2, 3, de sorte que tous nos raisonnements se rapportent aux permutations des seconds indices. Inversement, on peut disposer les facteurs de chaque produit de telle sorte que les seconds indices se suivent dans l'ordre naturel ou ordre des grandeurs croissantes; l'expression (3) s'écrit alors sous la forme:

$$a_{11}a_{22}a_{33}+a_{31}a_{12}a_{23}+a_{21}a_{32}a_{13}-a_{11}a_{32}a_{23}-a_{21}a_{12}a_{33}-a_{31}a_{22}a_{13}. \tag{9}$$

Ici, les premiers indices donnent toutes les permutations possibles des p, q, r, de plus, il est facile de vérifier que pour chaque terme de l'expression (9) la règle des signes s'énonce exactement de la même façon que précédemment, mais porte sur les premiers indices. Cela nous conduit à considérer, outre la somme (8), une somme analogue de la forme:

$$\sum_{(p_1, p_2, \ldots, p_n)} (-1)^{[p_1, p_2, \ldots, p_n]} a_{p_1 1} a_{p_2 2} \ldots a_{p_n n}. \tag{10}$$

Il est évident que cette somme est composée des mêmes termes que la somme (8). Dans la suite, nous verrons que les signes des termes analogues dans ces deux sommes sont les mêmes, c'est-à-dire que la somme (10) coïncide avec la somme (8).

Considérons enfin le cas $n = 2$. Le tableau est alors de la forme:

$$\begin{Vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{Vmatrix}$$

et la formule (8) donne l'expression suivante pour le déterminant du second ordre associé à ce tableau:

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22}-a_{12}a_{21}. \tag{11}$$

Il découle de ce qui précède que pour aborder l'étude des propriétés des déterminants, il nous faut étudier d'abord en détail les propriétés des permutations.

I-1-2. Permutations. Soient n éléments quelconques donnés dans un ordre déterminé. Nous dirons, par définition, que c'est une *permutation* engendrée par ces éléments. Démontrons d'abord que les permutations distinctes sont au nombre de $n!$. Pour $n =$ $= 2$ cela est évident, car deux éléments peuvent engendrer deux permutations distinctes. Pour $n = 3$ cela découle directement du calcul des permutations (5), où les nombres 1, 2, 3 tiennent le rôle d'éléments. Il est facile de voir que (5) donne toutes les permutations possibles de ces trois éléments. Démontrons par récurrence notre proposition pour un entier quelconque n. Supposant que notre proposition soit vraie pour un n quelconque, démontrons qu'elle l'est aussi pour $(n+1)$ éléments. Posons que n éléments donnent $n!$ permutations et considérons $(n+1)$ éléments que nous notons

$$C_1, C_2, \ldots, C_{n+1}.$$

Considérons d'abord les permutations dont le premier élément est C_1. Pour obtenir toutes ces permutations, il faut mettre C_1 en première place, puis à la suite écrire toutes les permutations possibles des n autres éléments restants. Par hypothèse, le nombre de ces permutations est égal à $n!$. Ainsi, les permutations des éléments C_k, dont le premier élément est C_1, sont au nombre de $n!$. De la même façon, les permutations des éléments C_k, dont le premier élément est C_2, sont aussi au nombre de $n!$. D'une façon générale, les permutations distinctes des éléments C_k sont au nombre de:

$$n!\cdot(n+1) = 1\cdot 2\cdot \ldots \cdot n\cdot(n+1) = (n+1)!,$$

ce qu'il fallait démontrer.

Nous pouvons supposer que les éléments sont les nombres entiers successifs, à partir de l'unité, ce que nous convenons de conserver par la suite. Appelons *transposition l'opération qui consiste à changer les places de deux éléments dans une permutation.* On voit immédiatement que l'on peut obtenir de toute permutation toute autre permutation au moyen de plusieurs transpositions. Considérons, par exemple, les deux permutations des quatre éléments

$$1, 3, 4, 2, \quad 2, 4, 1, 3.$$

On peut passer de la première de ces permutations à la seconde au moyen de transpositions de la façon suivante:

$$1, 3, 4, 2 \to 2, 3, 4, 1 \to 2, 4, 3, 1 \to 2, 4, 1, 3.$$

Ici, nous avons besoin de trois transpositions pour passer de la première permutation à la seconde. Si nous avions appliqué d'autres transpositions, nous aurions pu aussi, par une autre voie, passer de la première permutation à la seconde, ce qui signifie que le nombre de transpositions nécessaires pour passer d'une permutation à une autre n'est pas un nombre strictement déterminé. Il est essentiel de démontrer que ces différents nombres, pour deux permutations données, sont soit tous pairs, soit tous impairs. Autrement dit, ces nombres ont toujours même parité. Pour le voir, introduisons la notion d'inversion que nous avons déjà utilisée dans [I-1-1]. Soient données des permutations des n éléments $1, 2, \ldots, n$. Appelons *permutation fondamentale* la permutation:

$$1, 2, \ldots, n, \tag{12}$$

dans laquelle les nombres se suivent dans l'ordre naturel. Nous appelons inversion dans une permutation le fait que deux éléments de cette permutation ne se suivent pas dans le même ordre que dans la permutation fondamentale (12), c'est-à-dire que le plus grand nombre précède le plus petit. Nous appelons *permutation paire* toute permutation dans laquelle le nombre d'inversions est pair et *permutation impaire* celle pour laquelle ce nombre est impair. Enonçons le théorème suivant fondamental pour la suite:

Une transposition change d'un nombre impair le nombre d'inversions.

Considérons une permutation quelconque

$$a, b, \ldots, k, \ldots, p, \ldots, s \tag{13}$$

et supposons que nous appliquions à cette permutation une transposition des éléments k et p, c'est-à-dire que nous intervertissions les places de ces deux éléments. Après une telle transposition, la disposition des éléments k et p par rapport à ceux qui se trouvent à gauche de k ou à droite de p reste la même. Seule la disposition de k et p, par rapport aux éléments qui se trouvent entre eux, change, ainsi que la disposition des éléments k et p l'un par rapport à l'autre. Calculons le nombre total d'inversions. Supposons que dans la permutation (13) m éléments se trouvent entre les éléments k et p et soient parmi eux, par rapport à k, α éléments dans l'ordre naturel et β inversions et, par rapport à p, α_1 éléments dans l'ordre naturel et β_1 inversions. Il est évident que nous avons:

$$\alpha+\beta=\alpha_1+\beta_1=m. \tag{14}$$

Le résultat de cette transposition est tel que l'ordre naturel se transforme en inversion et inversement, c'est-à-dire, d'une façon plus précise, si l'élément k et un élément pris entre k et p sont avant la transposition k, p dans l'ordre naturel, alors, après la transposition k, p, ils donnent une inversion et inversement; il en est de même pour l'élément p. Ainsi, le nombre total d'inversions des *éléments k et p, par rapport à ceux qui se trouvent entre eux*, est $\beta+\beta_1$ avant la transposition k, p, et $\alpha+\alpha_1$ après, c'est-à-dire que la variation du nombre d'inversions est:

$$\gamma=(\alpha+\alpha_1)-(\beta+\beta_1).$$

Tenant compte de (14), on peut écrire ce nombre sous la forme:

$$\gamma=(\alpha+\alpha_1)-(m-\alpha+m-\alpha_1)=2(\alpha+\alpha_1-m).$$

d'où il découle immédiatement que le nombre γ est pair. Il ne reste plus qu'à considérer la disposition réciproque des éléments k et p. Si avant la transposition ils sont dans l'ordre naturel, ils présentent après une inversion et inversement, c'est-à-dire que la variation du nombre d'inversions est égale à l'unité et par suite le nombre total d'inversions provenant d'une transposition est impair.

Donnons plusieurs corollaires découlant du théorème qui vient d'être démontré.

C o r o l l a i r e I. Si l'on écrit les unes à la suite des autres les $n!$ permutations et que l'on réalise la transposition de deux éléments déterminés, par exemple des éléments 1 et 3, toutes les permutations paires deviennent alors des permutations impaires et inversement; d'une façon générale on retrouve l'ensemble des $n!$ permutations. Il en découle immédiatement que *les nombres des permutations paires et impaires sont égaux.*

C o r o l l a i r e II. Toute permutation s'obtient de la permutation fondamentale au moyen de transpositions. Il découle immédiatement du théorème démontré que *les permutations paires sont composées de permutations qui s'obtiennent de la permutation fondamentale au moyen d'un nombre pair de transpositions et les permutations impaires de permutations qui s'obtiennent de la fondamentale au moyen d'un nombre impair de transpositions.*

C o r o l l a i r e III. Le choix de la permutation fondamentale est arbitraire. Nous aurions pu prendre pour permutation fondamentale (12) une permutation quelconque; en outre, pour définir les inversions, il faut évidemment comparer cette dernière à la permutation fondamentale, c'est-à-dire partir de l'ordre propre aux éléments de la permutation fondamentale. Il est facile de voir que si nous prenons à la place de la permutation (12) une permutation paire quelconque pour fondamentale, alors les permutations paires restent, comme auparavant, des permutations paires et celles impaires des permutations impaires. Inversement, si nous prenons une permutation impaire quelconque pour fondamentale, les permutations impaires par rapport à la fondamentale deviennent alors des permutations paires et celle paires des permutations impaires.

Par exemple, si pour les six permutations des éléments 1, 2, 3 nous prenons pour fondamentale la permutation 2, 1, 3, les permutations paires sont alors:

$$2, 1, 3; \quad 1, 3, 2; \quad 3, 2, 1.$$

La deuxième de ces permutations présente deux inversions: le 1 précède le 2 et le 3 précède le 2, tandis que dans la fondamentale le

2 est placé devant le 1 et le 2 avant le 3. Les permutations impaires sont :

$$1,\ 2,\ 3;\quad 2,\ 3,\ 1;\quad 3,\ 1,\ 2.$$

La première de ces permutations présente une inversion par rapport à la permutation fondamentale 2, 1, 3 : le 1 précède le 2.

Tenant compte de tout ce qui vient d'être dit, nous pouvons énoncer la règle des signes dans l'expression (8) de la façon suivante : *nous faisons précéder un produit quelconque du signe plus, si la permutation des seconds indices donne une permutation paire, et du signe moins, si la permutation des seconds indices donne une permutation impaire, en outre, la permutation* 1, 2, . . ., n *est considérée comme fondamentale.*

Etudions à présent une propriété fondamentale du déterminant. Intervertissons la première et la deuxième colonne dans le tableau qui engendre le déterminant. Les nombres que nous désignions auparavant par a_{ik} seront notés par les mêmes lettres et les mêmes indices. Le tableau (6) est remplacé par le tableau :

$$\left\|\begin{matrix} a_{12} & a_{11} & a_{13} & \ldots & a_{1n} \\ a_{22} & a_{21} & a_{23} & \ldots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{n2} & a_{n1} & a_{n3} & \ldots & a_{nn} \end{matrix}\right\|. \tag{15}$$

A l'aide de la formule (8) nous pouvons former le déterminant associé au tableau (15). Dans ce tableau, les colonnes sont numérotées dans l'ordre suivant : 2, 1, 3, . . ., n, et nous devons considérer cette permutation comme fondamentale. Elle s'obtient de la permutation (12), précédemment fondamentale, à l'aide d'une transposition, elle est donc impaire. Ainsi, les permutations qui auparavant étaient impaires deviennent, avec le nouveau choix de la permutation fondamentale, des permutations paires et inversement. Donc, le déterminant associé au tableau (15) est la somme algébrique de termes ayant même expression absolue que ceux qui entrent dans la formule (8) mais qui, en vertu du changement de parité des permutations des seconds indices qui vient d'être mentionné, ont des signes contraires, c'est-à-dire que, *lorsqu'on intervertit deux colonnes, le déterminant change de signe*. Nous avons démontré cette propriété dans le cas où l'interversion portait sur la première et la deuxième colonne. Une démonstration analogue peut être faite dans le cas où nous intervertissons deux colonnes quelconques. Par exemple.

$$\begin{vmatrix} 1 & 0 & 3 \\ 2 & 7 & 6 \\ 5 & 3 & 0 \end{vmatrix} = -\begin{vmatrix} 1 & 3 & 0 \\ 2 & 6 & 7 \\ 5 & 0 & 3 \end{vmatrix}.$$

Le second déterminant s'obtient à partir du premier en permutant la deuxième et la troisième colonne.

Considérons une autre propriété des déterminants. Prenons un terme quelconque de la somme (8):

$$(-1)^{[p_1, p_2, \ldots, p_n]} a_{1p_1} a_{2p_2} \ldots a_{np_n}. \tag{16$_1$}$$

En changeant l'ordre des facteurs, nous pouvons ramener les seconds indices à avoir l'ordre naturel, mais les premiers indices formeront alors une certaine permutation $q_1, q_2, \ldots, q_n$ et l'expression précédente s'écrit sous la forme:

$$(-1)^{[p_1, p_2, \ldots, p_n]} a_{q_1 1} a_{q_2 2} \ldots a_{q_n n}. \tag{16$_2$}$$

Le passage de (16$_1$) à (16$_2$) se réalise à l'aide de plusieurs transpositions de facteurs. Une telle transposition fait intervenir en même temps des transpositions dans les permutations des premiers indices et des seconds. Si le nombre de transpositions nécessaires au passage de (16$_1$) à (16$_2$) est pair, il en découle que la permutation $p_1, p_2, \ldots, p_n$ est paire, puisqu'elle se transforme en permutation fondamentale $1, 2, \ldots, n$ à l'aide d'un nombre pair de transpositions et, par conséquent, il est évident qu'elle peut être aussi obtenue de la fondamentale par un nombre pair de transpositions. En outre, la permutation $q_1, q_2, \ldots, q_n$ est paire, étant donné qu'elle s'obtient aussi de la fondamentale à l'aide du même nombre pair de transpositions. De la même manière, si $p_1, p_2, \ldots, p_n$ est une permutation impaire, $q_1, q_2, \ldots, q_n$ est aussi une permutation impaire. Il en découle que $(-1)^{[p_1, p_2, \ldots, p_n]} = (-1)^{[q_1, q_2, \ldots, q_n]}$, et nous pouvons donc écrire:

$$(-1)^{[p_1, p_2, \ldots, p_n]} a_{1p_1} a_{2p_2} \ldots a_{np_n} =$$
$$= (-1)^{[q_1, q_2, \ldots, q_n]} a_{q_1 1} a_{q_2 2} \ldots a_{q_n n}.$$

Ainsi, si nous comparons les termes correspondants dans les sommes (8) et (10), nous voyons que ces sommes coïncident. Dans la somme (10), les lignes jouent le même rôle que les colonnes dans la somme (8), et il découle immédiatement de nos raisonnements que *si l'on remplace dans un tableau les lignes par les colonnes, et donc les colonnes par les lignes, tout en conservant leur ordre, la valeur du déterminant ne change pas.*

Par exemple, nous avons l'égalité des deux déterminants du troisième ordre suivants:

$$\begin{vmatrix} 2 & 3 & 5 \\ 7 & 0 & 1 \\ 2 & 1 & 6 \end{vmatrix} = \begin{vmatrix} 2 & 7 & 2 \\ 3 & 0 & 1 \\ 5 & 1 & 6 \end{vmatrix}.$$

I-1-3. Propriétés fondamentales des déterminants. I. Enonçons d'abord la propriété que nous venons de démontrer: *la valeur d'un déterminant ne change pas quand on intervertit les lignes et les colonnes de même indice.* Dans la suite, tout ce qui sera démontré pour les colonnes sera valable aussi pour les lignes et inversement.

II. Nous avons vu dans [I-1-2] que lorsqu'on échange entre elles deux colonnes, le déterminant ne change que de signe, c'est-à-dire que *lorsqu'on échange entre elles deux lignes (resp. deux colonnes), le déterminant change de signe.*

III. Si un déterminant a deux lignes identiques, en les échangeant entre elles d'une part nous n'avons rien changé et d'autre part, en vertu de ce qui vient d'être démontré, nous avons changé le signe du déterminant, c'est-à-dire en désignant par Δ ce déterminant, $\Delta = -\Delta$ ou $\Delta = 0$. Ainsi, *un déterminant qui a deux lignes (resp. deux colonnes) identiques est nul.*

IV. On appelle fonction linéaire homogène des variables $x_1, x_2, \ldots, x_n$ tout polynôme du premier degré par rapport à ces variables, sans terme constant, c'est-à-dire une expression de la forme:

$$\varphi(x_1, x_2, \ldots, x_n) = a_1x_1 + a_2x_2 + \ldots + a_nx_n,$$

où les coefficients a_i ne dépendent pas des x_s. Cette fonction vérifie les deux propriétés évidentes suivantes:

$$\varphi(kx_1, kx_2, \ldots, kx_n) = k\varphi(x_1, x_2, \ldots, x_n),$$

$$\varphi(x_1+y_1, x_2+y_2, \ldots, x_n+y_n) =$$
$$= \varphi(x_1, x_2, \ldots, x_n) + \varphi(y_1, y_2, \ldots, y_n).$$

La dernière propriété est vraie pour un nombre fini quelconque de termes. Considérons à nouveau la formule (8), nous voyons que chaque terme de la somme envisagée contient comme facteur un et seulement un élément de chaque ligne. Il en découle qu'*un déterminant est une fonction linéaire homogène des éléments d'une ligne (ou resp. d'une colonne) quelconque.*

Donc, *si tous les éléments d'une ligne (resp. d'une colonne) contiennent un facteur commun, on peut le sortir du signe du déterminant.*

La valeur du déterminant associé au tableau (6) est souvent notée, comme nous l'avons déjà indiqué, sous la forme:

$$\begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix}$$

ou encore:

$$|a_{ik}| \qquad (i, k = 1, 2, \ldots, n).$$

La propriété, que nous venons de démontrer, peut s'écrire dans un cas particulier sous la forme :

$$\begin{vmatrix} ka_{11} & ka_{12} & ka_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = k \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}.$$

La seconde des propriétés énoncées des fonctions linéaires homogènes s'applique aux déterminants de la façon suivante : *si les éléments d'une ligne (resp. d'une colonne) sont les sommes d'un même nombre de termes, alors le déterminant est égal à la somme de déterminants dans lesquels les éléments de la ligne (resp. de la colonne) envisagée sont remplacés respectivement par les termes de la somme.* Par exemple,

$$\begin{vmatrix} a & b & c+c' \\ d & e & f+f' \\ g & h & i+i' \end{vmatrix} = \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} + \begin{vmatrix} a & b & c' \\ d & e & f' \\ g & h & i' \end{vmatrix}.$$

Notons encore une conséquence de la linéarité et de l'homogénéité. *Si tous les éléments d'une ligne (resp. d'une colonne) sont nuls, le déterminant est nul.*

V. Si l'on supprime dans le tableau (6) la i-ième ligne et la k-ième colonne, à l'intersection desquelles se trouve l'élément a_{ik}, il reste un tableau à $(n-1)$ lignes et colonnes. Le déterminant d'ordre $(n-1)$ ainsi associé est appelé *mineur* du déterminant fondamental d'ordre n correspondant à l'élément a_{ik}. Désignons-le par Δ_{ik} et formons le produit :

$$A_{ik} = (-1)^{i+k} \Delta_{ik} \tag{17}$$

que nous appellerons *complémentaire algébrique* ou *cofacteur* de l'élément a_{ik}. Montrons à présent que ces complémentaires algébriques sont les coefficients de la fonction linéaire homogène utilisée précédemment, c'est-à-dire que pour toute i-ième ligne, on a :

$$\Delta = A_{i1}a_{i1} + A_{i2}a_{i2} + \ldots + A_{in}a_{in} \quad (i = 1, 2, \ldots, n) \tag{18}$$

et pour toute k-ième colonne la formule :

$$\Delta = A_{1k}a_{1k} + A_{2k}a_{2k} + \ldots + A_{nk}a_{nk} \quad (k = 1, 2, \ldots, n), \tag{19}$$

où Δ est le déterminant. Autrement dit, nous devons démontrer que, si dans la somme (8) nous mettons en facteur tous les termes contenant un élément déterminé a_{ik}, son coefficient est précisément le complémentaire algébrique A_{ik} donné par la formule (17). Désignons ce coefficient par B_{ik} et remarquons tout d'abord qu'il est la somme de produits de $(n-1)$ éléments, en outre, ces produits ne contiennent déjà plus les éléments de la i-ième ligne et la k-ième colonne.

Considérons d'abord le cas $i = k = 1$ et écrivons les termes de la somme (8) qui contiennent en facteur l'élément a_{11}:

$$a_{11} \sum_{(p_2, \ldots, p_n)} (-1)^{[1, p_2, \ldots, p_n]} a_{2p_2} \ldots a_{np_n}.$$

Ici, la sommation est étendue à toutes les permutations possibles $p_2, p_3, \ldots, p_n$ engendrées par les nombres $2, 3, \ldots, n$. Dans la permutation complète $1, p_2, \ldots, p_n$ le premier élément, l'unité, se trouve dans l'ordre naturel par rapport à tous ceux qui le suivent, et pour cette raison nous avons pour le nombre d'inversions:

$$[1, p_2, \ldots, p_n] = [p_2, \ldots, p_n],$$

de plus, dans les deux cas, on prend pour permutation fondamentale celle dont les nombres se suivent dans l'ordre naturel. Ainsi, nous avons l'expression suivante pour le coefficient de a_{11}:

$$B_{11} = \sum_{(p_2, p_3, \ldots, p_n)} (-1)^{[p_2, \ldots, p_n]} a_{2p_2} \ldots a_{np_n}.$$

Cette somme ressemble à la définition du déterminant mineur de a_{11}, puisque par rapport au déterminant initial la première ligne et la première colonne sont absentes. On voit que $B_{11} = \Delta_{11} = (-1)^{1+1}\Delta_{11} = A_{11}$, c'est-à-dire pour $i = k = 1$ notre proposition est démontrée. Passons au cas où i et k sont arbitraires. Nous allons échanger la i-ième ligne avec la première. Il faut faire alors $(i - 1)$ permutations successives de la ligne envisagée. De la même façon échangeons la k-ième colonne avec la première. Après ces permutations, l'élément a_{ik} occupe la place de l'élément a_{11} au coin supérieur gauche. La ligne de numéro i et la colonne de numéro k occupent respectivement la ligne et la colonne de numéro 1 et l'ordre des autres lignes et colonnes reste inchangé. Le résultat obtenu ci-dessus montre qu'après avoir réalisé les permutations indiquées, le coefficient de a_{ik} est égal à Δ_{ik}. Mais il nous a fallu faire $(i - 1) + (k - 1)$ permutations de lignes et de colonnes, et chaque permutation pareille fait précéder le déterminant du facteur (-1), c'est-à-dire d'une manière générale du facteur $(-1)^{(i-1)+(k-1)} = (-1)^{i+k}$, donc l'expression finale du coefficient B_{ik} est:

$$B_{ik} = \frac{\Delta_{ik}}{(-1)^{i+k}} = (-1)^{i+k} \Delta_{ik} = A_{ik},$$

ce qu'il fallait démontrer. Ainsi, nous avons démontré les formules (18) et (19). Si dans le déterminant Δ nous remplaçons successivement les éléments de la i-ième ligne par des nombres $c_1, c_2, \ldots, c_n$ tout en conservant les autres lignes, alors dans la formule (18) les facteurs

A_{is} sont conservés, et la valeur du nouveau déterminant est

$$\Delta' = A_{i1}c_1 + A_{i2}c_2 + \ldots + A_{in}c_n. \tag{20}$$

En particulier, si les nombres $c_1, c_2, \ldots, c_n$ sont égaux respectivement aux éléments $a_{j1}, a_{j2}, \ldots, a_{jn}$ de la ligne de numéro j distinct de i, le déterminant Δ' a alors deux lignes identiques, la i-ième et la j-ième, et il est nul : $\Delta' = 0$, c'est-à dire :

$$A_{i1}a_{j1} + A_{i2}a_{j2} + \ldots + A_{in}a_{jn} = 0 \quad (i \neq j). \tag{21_1}$$

De même, cette propriété est vraie pour les colonnes :

$$A_{1k}a_{1l} + A_{2k}a_{2l} + \ldots + A_{nk}a_{nl} = 0 \quad (k \neq l). \tag{21_2}$$

Les formules (19) et ($21_{1,2}$) nous conduisent à la propriété suivante du déterminant importante pour la suite.

La somme des éléments d'une ligne (resp. d'une colonne) affectés respectivement d'un facteur égal à leur complémentaire algébrique donne la valeur du déterminant. La somme des éléments d'une ligne (resp. d'une colonne) affectés respectivement d'un facteur égal au complémentaire algébrique des éléments d'une autre ligne (resp. d'une colonne) est nulle.

VI. Ajoutons aux éléments de la première ligne d'un déterminant Δ les éléments de la deuxième ligne multipliés par un nombre p. Les éléments de la première ligne sont $a_{1s} + pa_{2s}$ ($s = 1, 2, \ldots, n$) et, en vertu de la propriété IV, le nouveau déterminant est la somme de deux déterminants : du déterminant initial et d'un second déterminant dont les éléments de la première ligne sont pa_{2s} ($s = 1, 2, \ldots, n$) et dont les autres lignes coïncident avec celles de Δ. Mettant dans la première ligne p en facteur, les première et deuxième lignes sont identiques, donc ce déterminant est nul, c'est-à-dire *la valeur d'un déterminant ne change pas si l'on ajoute aux éléments d'une ligne (resp. d'une colonne) les éléments correspondants d'une autre ligne (resp. d'une colonne) après les avoir multipliés par un même nombre.*

Indiquons plusieurs notations qui nous seront utiles dans la suite. Soit donné un tableau carré (6) de nombres et soit l un nombre entier positif non supérieur à n. Introduisons la notation suivante pour un déterminant d'ordre l constitué des lignes de numéros $p_1, p_2, \ldots, p_l$ et des colonnes de numéros $q_1, q_2, \ldots, q_l$ du tableau (6) :

$$A\begin{pmatrix} p_1 & p_2 & \ldots & p_l \\ q_1 & q_2 & \ldots & q_l \end{pmatrix} = \begin{vmatrix} a_{p_1q_1} & a_{p_1q_2} & \ldots & a_{p_1q_l} \\ a_{p_2q_1} & a_{p_2q_2} & \ldots & a_{p_2q_l} \\ \ldots & \ldots & \ldots & \ldots \\ a_{p_lq_1} & a_{p_lq_2} & \ldots & a_{p_lq_l} \end{vmatrix}. \tag{22}$$

En outre, par déterminant du premier ordre, correspondant à un nombre quelconque a, on comprend d'ordinaire ce nombre, c'est-à-dire $A\binom{p}{q} = a_{pq}$. Les suites de nombres entiers positifs p_1, $p_2, \ldots, p_l$ et $q_1, q_2, \ldots, q_l$ peuvent être ordonnées suivant un tout autre ordre que celui naturel des nombres p_s et q_s. Si dans ces deux suites, les nombres sont ordonnés suivant l'ordre naturel, alors le déterminant (22) est appelé *mineur* d'ordre l du déterminant (8). Ce déterminant (22) s'obtient de (8) en supprimant $(n-l)$ lignes et $(n-l)$ colonnes. Soient $r_1, r_2, \ldots, r_{n-l}$ et $s_1, s_2, \ldots, s_{n-l}$ les numéros dans l'ordre naturel des lignes et des colonnes supprimées. Le mineur

$$A\begin{pmatrix} r_1 & r_2 & \ldots & r_{n-l} \\ s_1 & s_2 & \ldots & s_{n-l} \end{pmatrix}$$

est appelé mineur complémentaire du mineur (22), et l'expression

$$(-1)^{p_1+p_2+\ldots+p_l+q_1+q_2+\ldots+q_l} A\begin{pmatrix} r_1 & r_2 & \ldots & r_{n-l} \\ s_1 & s_2 & \ldots & s_{n-l} \end{pmatrix} \qquad (22_1)$$

est appelée *complémentaire algébrique* du mineur (22). Pour un élément isolé a_{ik} cette définition du complémentaire algébrique coïncide avec la définition (17) donnée précédemment.

Nous désignons le complémentaire algébrique (22_1) par

$$A'\begin{pmatrix} p_1 & p_2 & \ldots & p_l \\ q_1 & q_2 & \ldots & q_l \end{pmatrix}.$$

Il est entièrement défini par la donnée du déterminant (22), c'est-à-dire par la donnée des suites de numéros des lignes $p_1, p_2, \ldots, p_l$ et des colonnes $q_1, q_2, \ldots, q_l$.

Fixons les numéros des lignes. La valeur du déterminant Δ est évidemment un polynôme homogène de degré l des éléments de ces lignes et, comme on peut le démontrer, elle s'exprime par la formule (théorème de Laplace):

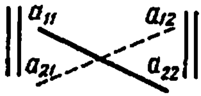
$$\Delta = \sum_{q_1<q_2<\ldots<q_l} A\begin{pmatrix} p_1 & p_2 & \ldots & p_l \\ q_1 & q_2 & \ldots & q_l \end{pmatrix} A'\begin{pmatrix} p_1 & p_2 & \ldots & p_l \\ q_1 & q_2 & \ldots & q_l \end{pmatrix}, \qquad (23)$$

où la sommation est étendue à toutes les suites croissantes de nombres $q_1, q_2, \ldots, q_l$ extraits de la suite $1, 2, \ldots, n$. Le nombre des termes dans la somme (23) est égal à celui des combinaisons de n éléments l à l:

$$C_n^l = \frac{n(n-1)\ldots(n-l+1)}{l!},$$

puisque l'ordre des nombres q_s ne joue aucun rôle, q_s n'étant pris que dans l'ordre naturel pour former la somme (23). Pour $l = 1$, nous avons $A\binom{p_1}{q_1} = a_{p_1q_1}$, et la formule (23) coïncide avec la formule (18) pour $i = p_1$. Il est facile de donner une formule analogue à (23) qui correspondrait au développement de Δ suivant les éléments de certaines colonnes. Dans la suite, nous n'aurons pas besoin de la formule (23) et, pour cette raison, nous ne la démontrons pas.

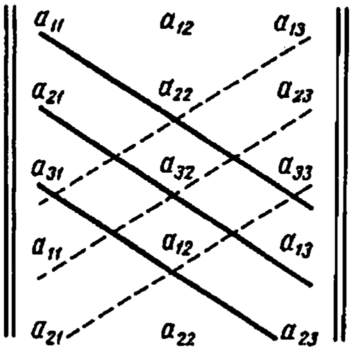
I-1-4. Calcul des déterminants. Le calcul des déterminants du deuxième ordre est très simple. D'après la formule (11), il faut écrire le tableau :



puis former le produit des éléments, qui se trouvent sur la diagonale soulignée d'un trait plein, précédé de son signe propre, et celui des éléments, qui se trouvent sur la diagonale soulignée d'un trait pointillé, affecté du signe contraire à son signe propre. La somme algébrique de ces deux produits donne la valeur du déterminant.

Passons au déterminant du troisième ordre. Sous forme développée il est donné par la formule (3). Il est facile de vérifier qu'on peut le calculer de la façon suivante : écrivons le tableau associé au déterminant et complétons-le inférieurement avec la première et la deuxième ligne. Nous avons un tableau contenant six diagonales, chacune de trois éléments.

Formons les produits des éléments, qui se trouvent sur les diagonales soulignées d'un trait plein, précédés de leur signe propre, et ceux des éléments des diagonales soulignées en pointillé, affectés du signe contraire à leur signe propre. La somme algébrique de ces six produits donne précisément la valeur du déterminant (*règle de Sarrus*).



Cette règle ne se généralise pas aux déterminants d'ordre supérieur à trois et pour ceux-ci on doit opérer différemment pour simplifier les calculs. Il est utile, par exemple, d'avoir recours à la propriété VI des déterminants indiquée dans [I-1-3]. Prenons un exemple. Considérons le déterminant du quatrième ordre:

$$\Delta=\begin{vmatrix} 3 & 5 & 1 & 0 \\ 2 & 1 & 4 & 5 \\ 1 & 7 & 4 & 2 \\ -3 & 5 & 1 & 1 \end{vmatrix}.$$

Multiplions la troisième ligne par (—2) et additionnons-la à la deuxième; puis multiplions la troisième ligne par 3, additionnons-la à la quatrième et retranchons-la de la première. Ainsi, en vertu de la propriété mentionnée, nous avons un déterminant égal au déterminant initial, qui a, en outre, trois zéros dans la première colonne:

$$\Delta=\begin{vmatrix} 0 & -16 & -11 & -6 \\ 0 & -13 & -4 & 1 \\ 1 & 7 & 4 & 2 \\ 0 & 26 & 13 & 7 \end{vmatrix}.$$

En le développant par rapport aux éléments de la première colonne, nous avons d'après la formule (19):

$$\Delta=\begin{vmatrix} -16 & -11 & -6 \\ -13 & -4 & 1 \\ 26 & 13 & 7 \end{vmatrix}.$$

Multiplions la troisième colonne par 4 et additionnons-la à la deuxième, puis multiplions-la par 13 et additionnons-la à la première. Nous obtenons ainsi:

$$\Delta=\begin{vmatrix} -94 & -35 & -6 \\ 0 & 0 & 1 \\ 117 & 41 & 7 \end{vmatrix}=-\begin{vmatrix} -94 & -35 \\ 117 & 41 \end{vmatrix}=94\cdot41-35\cdot117=-241.$$

I-1-5. Exemples. 1. Soit à calculer le volume d'un parallélépipède dont les trois arêtes issues d'un même sommet sont les vecteurs **A**, **B** et **C**. Comme on le sait (tome II, IV-1-6), le volume cherché est donné par le produit scalaire du vecteur **A** et du produit vectoriel (**B** × × **C**):

$$V=\mathbf{A}\,(\mathbf{B}\times\mathbf{C}). \tag{24}$$

Ce volume est précédé du signe plus, si les vecteurs **A**, **B**, **C** ont même orientation que les axes de coordonnées, et précédé du signe moins, si les orientations envisagées en sont distinctes. Les composantes du produit vectoriel sont

$$B_yC_z-B_zC_y,\quad B_zC_x-B_xC_z,\quad B_xC_y-B_yC_x,$$

et, par suite, le produit scalaire (24) est:

$$A_x\,(B_yC_z-B_zC_y)+A_y\,(B_zC_x-B_xC_z)+A_z\,(B_xC_y-B_yC_x).$$

Il est facile de voir que cette dernière somme est le déterminant du troisième ordre:

$$V = \begin{vmatrix} A_x & B_x & C_x \\ A_y & B_y & C_y \\ A_z & B_z & C_z \end{vmatrix}. \tag{25}$$

Si ce déterminant s'annule, cela signifie que le volume est nul, autrement dit, que les trois vecteurs **A**, **B** et **C** sont coplanaires, c'est-à-dire se trouvent dans un même plan. Si nous permutons deux lignes (resp. deux colonnes) dans le déterminant, la première et la deuxième, par exemple, par là même l'ordre des vecteurs **A**, **B**, **C** est remplacé par l'ordre des vecteurs **B**, **A**, **C**, et si les vecteurs de la première suite avaient la même orientation que les axes de coordonnées, la permutation engendre alors une autre orientation et inversement. Donc, la valeur du déterminant change de signe.

D'une façon analogue, si nous considérons dans le plan XY deux vecteurs de composantes (A_x, A_y) et (B_x, B_y), l'aire du parallélogramme construit sur ces deux vecteurs est, par définition, égale au déterminant du deuxième ordre:

$$P = \begin{vmatrix} A_x & B_x \\ A_y & B_y \end{vmatrix}.$$

Considérons maintenant un triangle dont les coordonnées des sommets sont:

$$M_1(x_1, y_1), \quad M_2(x_2, y_2), \quad M_3(x_3, y_3).$$

Soient les vecteurs $\mathbf{A} = \overline{M_1M_2}$ et $\mathbf{B} = \overline{M_1M_3}$ de composantes:

$$\overline{M_1M_2}(x_2 - x_1, y_2 - y_1), \qquad \overline{M_1M_3}(x_3 - x_1, y_3 - y_1),$$

et l'aire du triangle est exprimée sous la forme:

$$P = \frac{1}{2} \begin{vmatrix} x_2 - x_1 & x_3 - x_1 \\ y_2 - y_1 & y_3 - y_1 \end{vmatrix}.$$

Il est facile de démontrer que le déterminant du deuxième ordre écrit ci-dessus peut être remplacé par un déterminant du troisième ordre et l'on peut écrire la formule de l'aire du triangle sous la forme suivante:

$$P = \frac{1}{2} \begin{vmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ 1 & 1 & 1 \end{vmatrix}.$$

Si ce déterminant s'annule, cela indique que les points M_1, M_2, M_3 se trouvent sur une même droite. Autrement dit, l'équation d'une droite qui passe par deux points donnés (x_1, y_1) et (x_2, y_2) peut s'écrire sous la forme:

$$\begin{vmatrix} x & x_1 & x_2 \\ y & y_1 & y_2 \\ 1 & 1 & 1 \end{vmatrix} = 0.$$

2. Il est aisé de former à l'aide des déterminants les équations de certains lieux géométriques. Par exemple, soit à trouver l'équation d'un cercle qui passe par trois points donnés: (x_1, y_1), (x_2, y_2) et (x_3, y_3). Il est évident que cette équation s'écrit à l'aide d'un déterminant du quatrième ordre sous la forme suivante:

$$\begin{vmatrix} x^2+y^2 & x_1^2+y_1^2 & x_2^2+y_2^2 & x_3^2+y_3^2 \\ x & x_1 & x_2 & x_3 \\ y & y_1 & y_2 & y_3 \\ 1 & 1 & 1 & 1 \end{vmatrix} = 0. \qquad (26)$$

En effet, en développant par rapport aux éléments de la première colonne, on peut s'assurer que l'équation écrite est une équation du deuxième degré dans laquelle les coefficients de x^2 et y^2 sont identiques et le terme en xy est absent, c'est-à-dire l'équation (26) est bien celle d'un cercle. Enfin, si nous faisons $x = x_k$ et $y = y_k$ $(k = 1, 2, 3)$ dans cette équation, la première colonne du déterminant est identique alors à l'une des autres colonnes et l'équation du cercle est satisfaite, c'est-à-dire que le cercle passe réellement par les trois points donnés. Remarquons que si les trois points donnés se trouvent sur une droite, dans l'équation (26) le coefficient de $(x^2 + y^2)$ est nul.

De même, dans un espace rapporté aux trois axes OX, OY, OZ, l'équation d'un plan passant par trois points donnés (x_1, y_1, z_1), (x_2, y_2, z_2) et (x_3, y_3, z_3) peut s'écrire sous la forme d'un déterminant du quatrième ordre:

$$\begin{vmatrix} x & x_1 & x_2 & x_3 \\ y & y_1 & y_2 & y_3 \\ z & z_1 & z_2 & z_3 \\ 1 & 1 & 1 & 1 \end{vmatrix} = 0. \qquad (27)$$

Si les trois points donnés se trouvent sur une droite, l'équation (27) se transforme en l'identité $0 = 0$.

3. Considérons le déterminant D_n d'ordre n, dont chaque ligne est composée de toutes les puissances entières d'un nombre, de la

$(n-1)$-ième puissance jusqu'à la puissance zéro incluse:

$$D_n = \begin{vmatrix} x_1^{n-1} & x_1^{n-2} & \ldots & x_1 & 1 \\ x_2^{n-1} & x_2^{n-2} & \ldots & x_2 & 1 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ x_n^{n-1} & x_n^{n-2} & \ldots & x_n & 1 \end{vmatrix}. \qquad (28)$$

Pour $n=1$ et 2 nous avons:

$$D_1 = 1; \qquad D_2 = x_1 - x_2.$$

Pour développer le déterminant D_3 nous remplaçons dans la première ligne de ce déterminant le nombre x_1 par la lettre x. Nous obtenons un déterminant de la forme:

$$D_3(x) = \begin{vmatrix} x^2 & x & 1 \\ x_2^2 & x_2 & 1 \\ x_3^2 & x_3 & 1 \end{vmatrix}$$

En le développant par rapport aux éléments de la première ligne, nous voyons que $D_3(x)$ est un polynôme du deuxième degré en x. Si nous faisons $x = x_2$ ou $x = x_3$ dans le déterminant, la première ligne est identique à la deuxième ou à la troisième ligne et le déterminant est nul, c'est-à-dire que le trinôme $D_3(x)$ a pour racines x_2 et x_3 et peut donc être représenté sous la forme:

$$D_3(x) = A_3(x-x_2)(x-x_3),$$

où A_3 est le coefficient de x^2, c'est-à-dire le complémentaire algébrique correspondant à l'élément x^2 du déterminant $D_3(x)$ qui se trouve dans l'angle supérieur gauche. Il en résulte:

$$A_3 = \begin{vmatrix} x_2 & 1 \\ x_3 & 1 \end{vmatrix},$$

c'est-à-dire que A_3 est le déterminant D_2 formé à partir des nombres x_2 et x_3. Soit:

$$D_3(x) = (x_2 - x_3)(x-x_2)(x-x_3).$$

Faisant $x = x_1$, nous obtenons pour D_3 une expression sous la forme d'un produit de trois facteurs:

$$D_3 = \begin{matrix} (x_1 - x_2)(x_1 - x_3) \\ (x_2 - x_3). \end{matrix}$$

D'une façon analogue, connaissant l'expression de D_3, on peut obtenir l'expression de D_4. Elle a la forme d'un produit de six facteurs :

$$D_4 = \begin{array}{l} (x_1-x_2)(x_1-x_3)(x_1-x_4) \\ \qquad\qquad (x_2-x_3)(x_2-x_4). \\ \qquad\qquad\qquad\qquad (x_3-x_4) \end{array}$$

Et, pour n quelconque, nous trouvons l'expression suivante du déterminant D_n, appelé déterminant de Vandermonde :

$$D_n = \begin{array}{l} (x_1-x_2)(x_1-x_3)\dots(x_1-x_n) \\ \qquad\qquad (x_2-x_3)\dots(x_2-x_n). \\ \dots\dots\dots\dots\dots\dots \\ \qquad\qquad\qquad\qquad (x_{n-1}-x_n) \end{array} \tag{29}$$

L'expression écrite ci-dessus possède un lien intéressant avec la définition fondamentale des déterminants. Tout déterminant d'ordre n peut s'écrire sous la forme :

$$\begin{vmatrix} x_{1n} & x_{1\,n-1} & \dots & x_{11} \\ x_{2n} & x_{2\,n-1} & \dots & x_{21} \\ \dots & \dots & \dots & \dots \\ x_{nn} & x_{n\,n-1} & \dots & x_{n1} \end{vmatrix}. \tag{30}$$

Remplaçons-y formellement chaque élément x_{ik} par x_i^{k-1}. Après cette substitution, le déterminant (30) coïncide, comme il est facile de le voir, avec le déterminant de Vandermonde (28). Il en découle immédiatement la règle suivante de formation de la somme qui donne la valeur du déterminant (30) : ouvrons les parenthèses dans l'expression (29) et remplaçons chaque terme de la forme x_k^{s-1} par x_{ks}, de plus, si un tel terme ne contient pas les puissances d'un nombre quelconque x_k, il faut dans le produit correspondant ajouter le facteur x_k^0 qui, après la substitution mentionnée, devient x_{k1}. Remarquons que cette dernière règle peut être adoptée pour définition des déterminants.

4. Considérons une expression dont nous aurons besoin par la suite :

$$\Delta(x) = \begin{vmatrix} a_{11}+x & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22}+x & a_{23} & \dots & a_{2n} \\ a_{31} & a_{32} & a_{33}+x & \dots & a_{3n} \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nn}+x \end{vmatrix} \tag{31}$$

et développons-la suivant les puissances de x. A cet effet écrivons-la d'abord sous la forme suivante:

$$\Delta(x) = \begin{vmatrix} a_{11}+x & a_{12}+0 & a_{13}+0 & \dots & a_{1n}+0 \\ a_{21}+0 & a_{22}+x & a_{23}+0 & \dots & a_{2n}+0 \\ a_{31}+0 & a_{32}+0 & a_{33}+x & \dots & a_{3n}+0 \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1}+0 & a_{n2}+0 & a_{n3}+0 & \dots & a_{nn}+x \end{vmatrix}. \qquad (32)$$

Chacune des colonnes de ce déterminant est la somme de deux termes et, en appliquant plusieurs fois la propriété IV des déterminants, nous le représentons sous la forme d'une somme de 2^n déterminants dont les colonnes ne sont plus des sommes. Si dans toutes les colonnes de l'expression (32) nous supprimons les seconds termes, nous obtenons alors un terme indépendant de la lettre x, c'est-à-dire le terme constant du développement de $\Delta(x)$:

$$\Delta = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}. \qquad (33)$$

Inversement, si nous supprimons dans toutes les colonnes les premiers termes, nous obtenons alors le terme de plus haut degré du polynôme $\Delta(x)$ qui est égal à:

$$\begin{vmatrix} x & 0 & 0 & \dots & 0 \\ 0 & x & 0 & \dots & 0 \\ 0 & 0 & x & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & x \end{vmatrix} = x^n.$$

Considérons à présent les autres termes du polynôme. Supposons que nous conservions les seconds termes dans les colonnes de numéros $p_1, p_2, \dots, p_s$ et les premiers termes dans les autres colonnes. Chaque colonne de numéro p_k ($k = 1, 2, \dots, s$) n'a donc que des zéros pour éléments sauf pour un, égal à x, qui se trouve sur la diagonale principale, c'est-à-dire à l'intersection de la ligne et de la colonne de même numéro. En développant le déterminant ainsi obtenu successivement suivant les éléments des colonnes de numéros $p_1, p_2, \dots, p_s$, nous obtenons le facteur x^s et nous devons supprimer les lignes de numéros $p_1, p_2, \dots, p_s$, ainsi que les colonnes de mêmes numéros. Après une telle suppression, les complémentaires algébriques correspondant à ces éléments sont égaux aux mineurs, puisque les numéros de la ligne et de la colonne qu'on supprime coïncident. Ainsi, pour tout choix des numéros des colonnes p_k ($k = 1, 2, \ldots, s$) le déterminant contient x^s avec un coefficient égal au déterminant d'ordre $(n - s)$ obtenu à partir du déterminant fondamental (33) en supprimant les lignes et les colonnes à l'intersection desquelles se trouvent les éléments de la diagonale principale $a_{p_1p_1}, a_{p_2p_2}, \ldots, a_{p_sp_s}$. Notons par $\Delta_{p_1p_2\ldots p_s}^{p_1p_2\ldots p_s}$ ce déterminant d'ordre $(n - s)$. On convient de l'appeler *mineur principal d'ordre* $(n - s)$ *du déterminant* Δ. En choisissant les $p_1, p_2, \ldots, p_s$ de différente manière, nous obtenons le coefficient définitif de x^s dans l'expression de $\Delta(x)$ sous la forme d'une somme de tous les mineurs principaux possibles d'ordre $(n - s)$, c'est-à-dire:

$$\Delta(x) = x^n + S_1x^{n-1} + S_2x^{n-2} + \ldots + S_{n-1}x + S_n,$$

où S_k est la somme de tous les mineurs principaux d'ordre k du déterminant Δ et, en particulier, S_n est égal à Δ. Ecrivons l'expression explicite d'un coefficient quelconque:

$$S_k = \sum_{p_1<p_2<\ldots<p_{n-k}}^{(1,2,\ldots,n)} \Delta_{p_1p_2\ldots p_{n-k}}^{p_1p_2\ldots p_{n-k}} =$$

$$= \sum_{q_1<q_2<\ldots<q_k} \begin{vmatrix} a_{q_1q_1} & a_{q_1q_2} & \ldots & a_{q_1q_k} \\ a_{q_2q_1} & a_{q_2q_2} & \ldots & a_{q_2q_k} \\ \ldots & \ldots & \ldots & \ldots \\ a_{q_kq_1} & a_{q_kq_2} & \ldots & a_{q_kq_k} \end{vmatrix}. \tag{34}$$

Ici, la sommation s'étend à toutes les combinaisons possibles de k nombres $q_1, q_2, \ldots, q_k$, pris dans l'ordre naturel et extraits de l'ensemble des nombres $1, 2, \ldots, n$. Si dans l'expression (34) nous sommions séparément suivant chaque indice q_j, de 1 à n, alors dans la permutation $q_1, q_2, \ldots, q_k$ les nombres entiers seraient non seulement disposés dans l'ordre naturel, mais le seraient aussi dans tous les autres ordres possibles. D'une façon plus précise, chaque suite croissante donnerait en tout $k!$ permutations lors de la sommation suivant tous les q_j de 1 à n. Remarquons maintenant que la valeur du déterminant qui intervient dans la formule (34) ne change pas lorsqu'on permute deux nombres quelconques q_i et q_j. En effet, lorsque nous permutons, par exemple q_1 et q_2, nous permutons par là même la première et la deuxième ligne ainsi que la première et la deuxième colonne dans le déterminant envisagé, ce qui n'influe pas sur la valeur du déterminant. Ainsi, il découle de ce qui précède que si dans l'expression (34) nous sommons suivant chaque q_j variant de 1 à n, chaque terme de la somme (34) est répété $k!$ fois et nous pouvons donc écrire l'expression du coefficient S_k sous la forme:

$$S_k=\frac{1}{k!}\sum_{q_1=1}^{n}\sum_{q_2=1}^{n}\cdots\sum_{q_k=1}^{n}A\begin{pmatrix}q_1 & q_2 & \dots & q_k\\ q_1 & q_2 & \dots & q_k\end{pmatrix}. \tag{35}$$

I-1-6. Théorème sur la multiplication des déterminants. Nous allons à présent établir la formule qui donne le produit de deux déterminants de même ordre.

Soient deux déterminants d'ordre n:

$$\Delta=|a_{ik}|_1^n \tag{36$_1$}$$

et

$$\Delta_1=|b_{ik}|_1^n. \tag{36$_2$}$$

Formons un nouveau déterminant dont les éléments sont donnés par les formules:

$$c_{ik}=\sum_{s=1}^{n}a_{is}b_{sk}\qquad (i,\ k=1,\ 2,\ \dots,\ n), \tag{37}$$

et démontrons que ce déterminant est égal au produit des déterminants (36_1) et (36_2). Commençons par le cas $n = 2$. Tenant compte des formules (37) et, en vertu de la propriété IV de [I-1-3], en décomposant ce déterminant en déterminants plus simples, nous obtenons:

$$\begin{vmatrix}c_{11} & c_{12}\\ c_{21} & c_{22}\end{vmatrix}=\begin{vmatrix}a_{11}b_{11}+a_{12}b_{21} & a_{11}b_{12}+a_{12}b_{22}\\ a_{21}b_{11}+a_{22}b_{21} & a_{21}b_{12}+a_{22}b_{22}\end{vmatrix}=$$

$$=\begin{vmatrix}a_{11}b_{11} & a_{11}b_{12}\\ a_{21}b_{11} & a_{21}b_{12}\end{vmatrix}+\begin{vmatrix}a_{11}b_{11} & a_{12}b_{22}\\ a_{21}b_{11} & a_{22}b_{22}\end{vmatrix}+$$

$$+\begin{vmatrix}a_{12}b_{21} & a_{11}b_{12}\\ a_{22}b_{21} & a_{21}b_{12}\end{vmatrix}+\begin{vmatrix}a_{12}b_{21} & a_{12}b_{22}\\ a_{22}b_{21} & a_{22}b_{22}\end{vmatrix}.$$

En mettant en facteur les termes communs aux éléments d'une même colonne, nous obtenons, dans les premier et quatrième déterminants du troisième membre, des déterminants qui ont des colonnes identiques, donc nuls. En échangeant entre elles les colonnes dans l'un des déterminants restants, nous en déduisons:

$$\begin{vmatrix}c_{11} & c_{12}\\ c_{21} & c_{22}\end{vmatrix}=b_{11}b_{22}\begin{vmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{vmatrix}+b_{12}b_{21}\begin{vmatrix}a_{12} & a_{11}\\ a_{22} & a_{21}\end{vmatrix}=$$

$$=b_{11}b_{22}\begin{vmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{vmatrix}-b_{12}b_{21}\begin{vmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{vmatrix}=$$

$$=\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}(b_{11}b_{22}-b_{12}b_{21})=\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}\cdot\begin{vmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{vmatrix},$$

ce qu'il fallait démontrer.

Dans le cas général d'ordre n, après application de la propriété IV de [I-1-3] nous avons:

$$|c_{ik}|_1^n=\sum_{(s_1, s_2, \ldots, s_n)}\begin{vmatrix} a_{1s_1}b_{s_11} & a_{1s_2}b_{s_22} & \ldots & a_{1s_n}b_{s_nn} \\ a_{2s_1}b_{s_11} & a_{2s_2}b_{s_22} & \ldots & a_{2s_n}b_{s_nn} \\ \cdot & \cdot & \cdot & \cdot \\ a_{ns_1}b_{s_11} & a_{ns_2}b_{s_22} & \ldots & a_{ns_n}b_{s_nn} \end{vmatrix}, \tag{38}$$

où les indices de sommation $s_1, s_2, \ldots, s_n$ prennent les valeurs entières $1, 2, \ldots, n$. Les déterminants de cette somme peuvent être mis sous la forme:

$$b_{s_11}b_{s_22}\ldots b_{s_nn}\begin{vmatrix} a_{1s_1} & a_{1s_2} & \ldots & a_{1s_n} \\ a_{2s_1} & a_{2s_2} & \ldots & a_{2s_n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{ns_1} & a_{ns_2} & \ldots & a_{ns_n} \end{vmatrix}. \tag{39}$$

Si parmi les indices $s_1, s_2, \ldots, s_n$ il y en a d'identiques, le déterminant (39) a alors des colonnes identiques et est donc nul. Ainsi, nous pouvons nous limiter à la considération des déterminants pour lesquels les indices s_k sont tous distincts et, donc, la suite des indices $s_1, s_2, \ldots, s_n$ n'est autre qu'une parmi les permutations des nombres $1, 2, \ldots, n$. Multiplions deux fois l'expression (39) par $(-1)^{[s_1, s_2, \ldots, s_n]}$, ce qui évidemment ne la change pas, et cette expression peut être écrite sous la forme d'un produit de deux facteurs:

$$(-1)^{[s_1, s_2, \ldots, s_n]}\begin{vmatrix} a_{1s_1} & a_{1s_2} & \ldots & a_{1s_n} \\ a_{2s_1} & a_{2s_2} & \ldots & a_{2s_n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{ns_1} & a_{ns_2} & \ldots & a_{ns_n} \end{vmatrix}\times$$
$$\times(-1)^{[s_1, s_2, \ldots, s_n]}\, b_{s_11}b_{s_22}\ldots b_{s_nn}.$$

Dans le premier facteur, à l'aide de plusieurs transpositions, la permutation $s_1, s_2, \ldots, s_n$ se ramène à la forme $1, 2, \ldots, n$. Pour chaque transposition la quantité $(-1)^{[s_1, s_2, \ldots, s_n]}$ et le déterminant considéré ne changent que de signe et le premier facteur reste inchangé.

Ainsi, l'expression (39) s'écrit sous la forme:

$$\begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix} \cdot (-1)^{[s_1, s_2, \ldots, s_n]} b_{s_1 1} b_{s_2 2} \ldots b_{s_n n},$$

et, par conséquent, revenant à la somme (38), nous obtenons:

$$|c_{ik}|_1^n = \Delta \sum_{(s_1, s_2, \ldots, s_n)} (-1)^{[s_1, s_2, \ldots, s_n]} b_{s_1 1} b_{s_2 2} \ldots b_{s_n n},$$

où la sommation est étendue à toutes les permutations $s_1, s_2, \ldots, s_n$ des nombres 1, 2, . . ., n. La somme écrite est un déterminant Δ_1, c'est-à-dire $|c_{ik}|_1^n = \Delta\Delta_1$, ce qu'il fallait démontrer. La formule (37) des c_{ik} se ramène à ce qui suit: les éléments de la i-ième ligne du déterminant Δ sont multipliés par les éléments correspondants de la k-ième colonne du second déterminant, puis on somme ces produits Nous savons que l'on peut échanger entre elles lignes et colonnes dans un déterminant sans changer sa valeur. Donc, la règle précédente de multiplication d'une ligne par une colonne peut être remplacée par trois autres règles: les règles d'une ligne par une ligne, d'une colonne par une colonne et d'une colonne par une ligne.

Enonçons alors le théorème suivant: *soient deux déterminants d'ordre n:*

$$|a_{ik}| \quad \text{et} \quad |b_{ik}|.$$

Formons un nouveau déterminant

$$|c_{ik}|,$$

dont les éléments sont donnés par l'une des formules suivantes:

$$c_{ik} = \sum_{s=1}^{n} a_{is} b_{sk}, \tag{40_1}$$

$$c_{ik} = \sum_{s=1}^{n} a_{is} b_{ks}, \tag{40_2}$$

$$c_{ik} = \sum_{s=1}^{n} a_{si} b_{sk}, \tag{40_3}$$

$$c_{ik} = \sum_{s=1}^{n} a_{si} b_{ks} \quad (i, k = 1, 2, \ldots, n). \tag{40_4}$$

La valeur du déterminant $|c_{ik}|$ est égale alors au produit des déterminants $|a_{ik}|$ et $|b_{ik}|$.

Exemple. Soit le déterminant fondamental:

$$\Delta = |a_{ik}|,$$

considérons en outre le déterminant constitué des complémentaires algébriques de ses éléments:

$$|A_{ik}|.$$

D'après le théorème énoncé plus haut, exprimons le produit $|a_{ik}|\cdot|A_{ik}|$ sous forme de déterminant en multipliant les lignes par les lignes. Nous avons pour ce déterminant les éléments suivants:

$$c_{ik} = \sum_{s=1}^{n} a_{is}A_{ks}.$$

En vertu de la propriété V des déterminants, nous avons:

$$c_{ik} = 0 \text{ pour } i \neq k; \quad c_{ii} = \Delta,$$

c'est-à-dire:

$$|a_{ik}|\cdot|A_{ik}| = \begin{vmatrix} \Delta & 0 & 0 & \dots & 0 \\ 0 & \Delta & 0 & \dots & 0 \\ 0 & 0 & \Delta & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \dots & \Delta \end{vmatrix}$$

ou

$$|a_{ik}|_1^n\,|A_{ik}|_1^n = \Delta^n, \text{ c'est-à-dire } \Delta\,|A_{ik}|_1^n = \Delta^n.$$

Supposons Δ non nul et en simplifiant par Δ nous avons:

$$|A_{ik}|_1^n = \Delta^{n-1}. \tag{41}$$

Si les éléments $a_{ik} = a_{ik}^{(0)}$ sont tels que le déterminant Δ est nul, on peut alors trouver des quantités a_{ik}, aussi voisines que l'on veut de $a_{ik}^{(0)}$, pour lesquelles le déterminant Δ est non nul. Pour ces valeurs de a_{ik} la formule (41) est valable et, par suite, à la limite pour $a_{ik} \to$ $\to a_{ik}^{(0)}$ cette formule reste vraie aussi pour $a_{ik} = a_{ik}^{(0)}$, c'est-à-dire que cette formule est vraie aussi dans le cas où $\Delta = 0$. Si l'on exprime Δ et A_{ik} en fonction des éléments a_{ik}, la formule (41) est alors une identité par rapport aux a_{ik}.

I-1-7. Tableaux rectangulaires. Nous rencontrerons dans la suite des tableaux d'éléments tels que les nombres de lignes et de colonnes puissent être différents. Considérons donc le tableau le plus général:

$$\begin{Vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{Vmatrix}. \tag{42}$$

Il contient m lignes et n colonnes, en outre les nombres m et n peuvent être distincts ou égaux. En supprimant dans ce tableau plusieurs lignes et colonnes nous pouvons former un déterminant tel que le nombre de lignes et celui de colonnes restantes soient identiques. Ces déterminants sont appelés déterminants extraits du tableau (42). Il est évident que l'ordre le plus élevé qu'ils puissent avoir est égal au plus petit des nombres m et n et que l'ordre le plus petit de ces déterminants est l'unité, les éléments n'étant alors autres que ceux du tableau (42). Supposons que tous les déterminants d'un ordre l extraits du tableau soient nuls. Il est alors aisé de voir que tous les déterminants d'ordre $(l+1)$ extraits du tableau sont également nuls. En effet, tout déterminant d'ordre $(l+1)$ peut être représenté sous la forme d'une somme de produits d'éléments d'une ligne par les complémentaires algébriques de ces éléments. Mais ces derniers coïncident, au signe près, avec certains déterminants d'ordre l extraits du tableau et sont donc tous nuls. Puisque tous les déterminants d'ordre $(l+1)$ sont nuls, en reprenant le raisonnement ci-dessus, tous les déterminants d'ordre $(l+2)$ sont également nuls, etc. Ainsi, si tous les déterminants d'un ordre déterminé, extraits du tableau (42), sont nuls, alors tous les déterminants d'ordre supérieur extraits de ce tableau sont aussi nuls.

Introduisons la notion importante de rang d'un tableau ou, comme on dit encore, d'une matrice (42). On appelle *rang d'une matrice* (42) l'ordre le plus élevé du déterminant non nul extrait de ce tableau, c'est-à-dire que si le rang de la matrice est k, alors parmi les déterminants d'ordre k extraits de ce tableau, il en existe au moins un non nul, mais tous les déterminants d'ordre $(k+1)$ extraits du tableau sont nuls.

Outre le tableau (42) soit le tableau :

$$\begin{Vmatrix} b_{11} & b_{12} & \dots & b_{1m} \\ b_{21} & b_{22} & \dots & b_{2m} \\ \cdot & \cdot & \cdot & \cdot \\ b_{n1} & b_{n2} & \dots & b_{nm} \end{Vmatrix}, \tag{43}$$

à n lignes et m colonnes. Formons les m^2 nombres :

$$c_{ik} = \sum_{s=1}^{n} a_{is} b_{sk} \quad (i, k = 1, 2, \dots, m). \tag{44}$$

On convient d'appeler le tableau carré composé des nombres c_{ik} produit des tableaux rectangulaires (42) et (43).

Démontrons un théorème qui généralise le théorème de la multiplication des déterminants.

Théorème. *Si $m \leqslant n$, alors*

$$|c_{ik}|_1^m = \sum_{r_1 < r_2 < \ldots < r_m} A\begin{pmatrix} 1 & 2 & \ldots & m \\ r_1 & r_2 & \ldots & r_m \end{pmatrix} B\begin{pmatrix} r_1 & r_2 & \ldots & r_m \\ 1 & 2 & \ldots & m \end{pmatrix}, \quad (45)$$

où la sommation est étendue à toutes les valeurs de r_k prises dans la suite des nombres 1, 2, . . ., n vérifiant l'inégalité indiquée. Si $m > n$, alors le déterminant $|c_{ik}|_1^m$ est nul.

La signification des symboles

$$A\begin{pmatrix} 1 & 2 & \ldots & m \\ r_1 & r_2 & \ldots & r_m \end{pmatrix} \quad \text{et} \quad B\begin{pmatrix} r_1 & r_2 & \ldots & r_m \\ 1 & 2 & \ldots & m \end{pmatrix}$$

a été définie dans [I-1-3]. La seconde de ces notations désigne un déterminant composé d'éléments du tableau (43) appartenant aux lignes d'indices $r_1, r_2, \ldots, r_m$ et aux colonnes 1, 2, . . ., m. Pour $m = n$, la somme de la formule (45) ne contient qu'un terme, celui qui correspond à $r_1 = 1$, $r_2 = 2$, . . ., $r_m = m$, et la formule (45) exprime le théorème de la multiplication des déterminants.

Considérons le cas $m < n$. La démonstration de la formule (45) est analogue à celle du théorème de la multiplication des déterminants. De même que dans cette démonstration, nous avons:

$$|c_{ik}|_1^m = \sum_{(s_1, \ldots, s_m)} A\begin{pmatrix} 1 & 2 & \ldots & m \\ s_1 & s_2 & \ldots & s_m \end{pmatrix} b_{s_1 1} b_{s_2 2} \ldots b_{s_m m}, \quad (46)$$

de plus chacun des nombres s_k peut prendre les valeurs 1, 2, . . ., n et on peut exclure les termes parmi lesquels des nombres s_h sont égaux, puisque ces termes sont nuls. Considérons une suite déterminée des nombres $r_1 < r_2 < \ldots < r_m$ extraite de la suite des nombres 1, 2, . . ., n et mettons en évidence les termes de la somme (45) dont l'ensemble des nombres $s_1, s_2, \ldots, s_m$ coïncide avec celui des nombres $r_1, r_2, \ldots, r_m$. Nous obtenons ainsi une partie de la somme (46):

$$\sum_{(t_1, t_2, \ldots, t_m)} A\begin{pmatrix} 1 & 2 & \ldots & m \\ t_1 & t_2 & \ldots & t_m \end{pmatrix} b_{t_1 1} b_{t_2 2} \ldots b_{t_m m}, \quad (47)$$

où la sommation est étendue à toutes les permutations possibles $(t_1, t_2, \ldots, t_m)$ des nombres $r_1, r_2, \ldots, r_m$. Multipliant deux fois par $(-1)^{[t_1, t_2, \ldots, t_m]}$ chaque terme de la somme (47), nous démontrons, de la même façon que dans [I-1-6], que cette somme est égale à:

$$A\begin{pmatrix} 1 & 2 & \ldots & m \\ r_1 & r_2 & \ldots & r_m \end{pmatrix} B\begin{pmatrix} r_1 & r_2 & \ldots & r_m \\ 1 & 2 & \ldots & m \end{pmatrix}.$$

Pour obtenir la somme complète de la formule (46) il suffit de sommer ce produit pour tous les $r_1 < r_2 < \ldots < r_m$, ce qui donne précisément la formule (45). Supposons enfin que $m > n$. Nous pouvons ajouter au tableau (42) $(m - n)$ colonnes constituées de zéros et au tableau (43) $(m - n)$ lignes également constituées de zéros. Si après cette addition nous calculons les c_{ik} non pas d'après les formules (44), mais d'après les formules:

$$c_{ik} = \sum_{s=1}^{m} a_{is} b_{sk} \qquad (i, k = 1, 2, \ldots, m), \tag{48}$$

nous obtenons alors les valeurs précédentes de c_{ik}, puisque les termes qu'on ajoute dans le second membre de (48) sont nuls. D'autre part, après l'addition indiquée, les tableaux (42) et (43) deviennent des tableaux carrés auxquels correspondent des déterminants nuls, et le déterminant $|c_{ik}|_1^m$ est nul en vertu du théorème de la multiplication des déterminants. Le théorème est donc démontré.

Remarque. Soient deux tableaux rectangulaires ayant chacun m lignes et n colonnes. Multipliant ligne par ligne:

$$c_{ik} = \sum_{s=1}^{n} a_{is} b_{ks} \qquad (i, k = 1, 2, \ldots, m),$$

nous obtenons un déterminant $|c_{ik}|_1^m$ dont la valeur est nulle pour $m > n$ et pour $m < n$ s'exprime par la formule:

$$|c_{ik}|_1^m = \sum_{r_1 < r_2 < \ldots < r_m} A \begin{pmatrix} 1 & 2 & \ldots & m \\ r_1 & r_2 & \ldots & r_m \end{pmatrix} B \begin{pmatrix} 1 & 2 & \ldots & m \\ r_1 & r_2 & \ldots & r_m \end{pmatrix}.$$

Corollaire. Soient deux tableaux carrés d'ordre n composés des éléments a_{ik} et b_{ik}, et les nombres c_{ik} sont donnés par les formules (44). Exprimons un mineur quelconque $C \begin{pmatrix} p_1 & p_2 & \ldots & p_l \\ q_1 & q_2 & \ldots & q_l \end{pmatrix}$ du déterminant $|c_{ik}|_1^n$ à l'aide des mineurs des déterminants $|a_{ik}|_1^n$ et $|b_{ik}|_1^n$. Il est facile de voir que le tableau carré qui engendre le mineur $C \begin{pmatrix} p_1 & p_2 & \ldots & p_l \\ q_1 & q_2 & \ldots & q_l \end{pmatrix}$ est le produit des tableaux rectangulaires:

$$\begin{Vmatrix} a_{p_1 1} & a_{p_1 2} & \ldots & a_{p_1 n} \\ a_{p_2 1} & a_{p_2 2} & \ldots & a_{p_2 n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{p_l 1} & a_{p_l 2} & \ldots & a_{p_l n} \end{Vmatrix} \quad \text{et} \quad \begin{Vmatrix} b_{1q_1} & b_{1q_2} & \ldots & b_{1q_l} \\ b_{2q_1} & b_{2q_2} & \ldots & b_{2q_l} \\ \cdot & \cdot & \cdot & \cdot \\ b_{nq_1} & b_{nq_2} & \ldots & b_{nq_l} \end{Vmatrix}.$$

Appliquant le théorème que nous venons de démontrer, nous obtenons l'expression cherchée:

$$C \begin{pmatrix} p_1 & p_2 & \ldots & p_l \\ q_1 & q_2 & \ldots & q_l \end{pmatrix} = \sum_{r_1 < r_2 < \ldots < r_l} A \begin{pmatrix} p_1 & p_2 & \ldots & p_l \\ r_1 & r_2 & \ldots & r_l \end{pmatrix} B \begin{pmatrix} r_1 & r_2 & \ldots & r_l \\ q_1 & q_2 & \ldots & q_l \end{pmatrix}, \tag{49}$$

où r_k prend ses valeurs dans la suite 1, 2, . . ., n. Soient R_A, R_B, R_C les rangs des tableaux $\| a_{ik} \|_1^n$, $\| b_{ik} \|_1^n$ et $\| c_{ik} \|_1^n$. Si par exemple $R_A < n$, et si dans la formule (49) nous prenons tout $l > R_A$, alors tous les $A \begin{pmatrix} p_1 p_2 \dots p_l \\ r_1 r_2 \dots r_l \end{pmatrix}$ sont, en vertu de la définition de R_A, nuls, donc tous les $C \begin{pmatrix} p_1 p_2 \dots p_l \\ q_1 q_2 \dots q_l \end{pmatrix}$ sont nuls aussi. Il en découle que $R_C < l$, c'est-à-dire $R_C \leqslant R_A$. Si le rang du tableau $\| a_{ik} \|_1^n$ est n, il est alors évident que $R_C \leqslant R_A$, car $R_C \leqslant n$. D'une façon analogue, $R_C \leqslant R_B$. Dans la suite nous montrerons que si le déterminant $| b_{ik} |_1^n \neq 0$, alors $R_C = R_A$, et si $| a_{ik} |_1^n \neq 0$, alors $R_C = R_B$.

I-2. Résolution des systèmes d'équations

I-2-1. Théorème de Cramer. La notion de déterminant étant établie et les propriétés fondamentales données, appliquons cette notion à la résolution des systèmes d'équations du premier degré. On étudiera d'abord le cas fondamental où le nombre d'équations coïncide avec le nombre d'inconnues. On peut écrire un système de n équations à n inconnues sous la forme:

$$\left.\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2, \\ \dots\dots\dots\dots\dots\dots\dots \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n = b_n, \end{array}\right\} \qquad (1)$$

où la notation des coefficients se fait comme en [I-1-1], dans le cas de trois équations à trois inconnues. Supposons, par hypothèse, que le déterminant du système, c'est-à-dire le déterminant correspondant au tableau des coefficients a_{ik} du système, ne soit pas nul:

$$\Delta = | a_{ik} | \neq 0. \qquad (2)$$

On multiplie les deux membres des équations du système (1) par les complémentaires algébriques des éléments de la colonne d'indice k de ce déterminant, autrement dit, on multiplie les deux membres de la première équation du système par A_{1k}, ceux de la seconde par A_{2k}, etc. En additionnant les équations ainsi obtenues, on arrive à une équation dont le second membre est la somme:

$$A_{1k}b_1 + A_{2k}b_2 + \dots + A_{nk}b_n,$$

et le coefficient de l'inconnue x_l dans le premier membre est la somme:

$$A_{1k}a_{1l} + A_{2k}a_{2l} + \dots + A_{nk}a_{nl} \qquad (l = 1, 2, \dots, n).$$

Cette somme est nulle pour $l \neq k$ et égale au déterminant Δ pour $l = k$, c'est-à-dire :

$$\Delta \cdot x_k = A_{1k}b_1 + A_{2k}b_2 + \ldots + A_{nk}b_n.$$

Si l'on procède de la même façon pour chaque indice k, on obtient à partir des équations (1) un nouveau système d'équations :

$$\Delta \cdot x_k = A_{1k}b_1 + A_{2k}b_2 + \ldots + A_{nk}b_n \qquad (k = 1, 2, \ldots, n). \qquad (3)$$

Réciproquement, il est facile de démontrer que l'on obtient le système (1) à partir du système (3). En effet, si l'on multiplie les deux membres de l'équation (3) par a_{lk} et si l'on somme par rapport à tous les k de 1 à n, on obtient, en utilisant la propriété V des déterminants, l'équation :

$$\Delta \cdot (a_{l1}x_1 + a_{l2}x_2 + \ldots + a_{ln}x_n) = \Delta \cdot b_l, \qquad (4)$$

ce qui donnera, lorsqu'on aura simplifié par le déterminant Δ non nul, la l-ième équation du système (1), l étant arbitraire. Ainsi, les systèmes (1) et (3) sont équivalents, et l'on peut résoudre le système (3) au lieu du système (1). Le système (3) se résout directement et ne donne qu'une seule solution qu'on calcule suivant les formules :

$$x_k = \frac{A_{1k}b_1 + A_{2k}b_2 + \ldots + A_{nk}b_n}{\Delta} \qquad (k = 1, 2, \ldots, n). \qquad (5)$$

D'après [I-1-3], le numérateur de l'expression ci-dessus est le déterminant obtenu à partir du déterminant Δ lorsqu'on remplace les éléments de la colonne d'indice k, c'est-à-dire les coefficients a_{ik} de x_k, par les termes constants b_i. On a donc le théorème suivant :

Théorème de Cramer. *Si le déterminant Δ du système (1) est non nul, ce système a une solution déterminée, donnée par les formules (5). D'après ces formules, chacune des inconnues est le quotient de deux déterminants ; le dénominateur est le déterminant du système, et le numérateur le déterminant obtenu à partir de celui-ci en remplaçant les coefficients de l'inconnue à déterminer par les seconds membres.*

Lorsqu'on a affaire à un grand nombre d'équations, le théorème de Cramer n'est pas d'application commode, et il existe des méthodes approchées pour résoudre des systèmes à grand nombre d'équations et d'inconnues que nous n'exposerons pas.

I-2-2. Cas général des systèmes d'équations. Soit le cas général de m équations à n inconnues:

$$\left.\begin{aligned}
X_1 &= a_{11}x_1 + a_{12}x_2 + \ldots + a_{1k}x_k + a_{1,k+1}x_{k+1} + \ldots + a_{1n}x_n = b_1,\\
X_2 &= a_{21}x_1 + a_{22}x_2 + \ldots + a_{2k}x_k + a_{2,k+1}x_{k+1} + \ldots +\\
&\qquad + a_{2n}x_n = b_2,\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
X_k &= a_{k1}x_1 + a_{k2}x_2 + \ldots + a_{kk}x_k + a_{k,k+1}x_{k+1} + \ldots +\\
&\qquad + a_{kn}x_n = b_k,\\
X_{k+1} &= a_{k+1,1}x_1 + a_{k+1,2}x_2 + \ldots + a_{k+1,k}x_k +\\
&\qquad + a_{k+1,k+1}x_{k+1} + \ldots + a_{k+1,n}x_n = b_{k+1},\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
X_m &= a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mk}x_k + a_{m,k+1}x_{k+1} + \ldots +\\
&\qquad + a_{mn}x_n = b_m.
\end{aligned}\right\} \tag{6}$$

Dans les calculs qui vont suivre, on désigne, pour simplifier, par X_s tout le premier membre de l'équation d'indice s. Les coefficients a_{ik} de ce système forment un tableau rectangulaire de rang k. En inversant les lignes et les colonnes, c'est-à-dire en substituant la numérotation des équations à celle des inconnues et inversement, on peut arriver à ce qu'un déterminant d'ordre k, appartenant au tableau et non nul, se trouve dans le coin supérieur gauche. On l'appellera *déterminant principal du système*. Il aura la forme:

$$\Delta = \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1k}\\ a_{21} & a_{22} & \ldots & a_{2k}\\ \cdot & \cdot & \cdot & \cdot\\ a_{k1} & a_{k2} & \ldots & a_{kk} \end{vmatrix}. \tag{7}$$

Formons les $(m - k)$ déterminants d'ordre $(k + 1)$ appelés *déterminants caractéristiques du système*; on obtient ces déterminants à partir du déterminant principal en lui ajoutant une ligne composée des coefficients des équations, dont l'indice est supérieur à k, et une colonne composée des termes constants. Ces déterminants caractéristiques sont définis par les formules:

$$\Delta_{k+s} = \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1k} & b_1\\ a_{21} & a_{22} & \ldots & a_{2k} & b_2\\ \cdot & \cdot & \cdot & \cdot & \cdot\\ a_{k1} & a_{k2} & \ldots & a_{kk} & b_k\\ a_{k+s,1} & a_{k+s,2} & \ldots & a_{k+s,k} & b_{k+s} \end{vmatrix} \tag{8}$$

$$(k+s = k+1,\ k+2,\ \ldots,\ m).$$

Si $k = m$, c'est-à-dire si le rang est égal au nombre d'équations, on n'aura pas de déterminants caractéristiques. Considérons, parallèlement aux déterminants caractéristiques, d'autres déterminants que l'on obtient à partir des premiers en remplaçant dans la dernière colonne les termes constants par le premier membre des équations:

$$\begin{vmatrix} a_{11} & a_{12} & \dots a_{1k} & X_1 \\ a_{21} & a_{22} & \dots a_{2k} & X_2 \\ \dots & \dots & \dots & \dots \\ a_{k1} & a_{k2} & \dots a_{kk} & X_k \\ a_{k+s,1} & a_{k+s,2} & \dots a_{k+s,k} & X_{k+s} \end{vmatrix}. \tag{9}$$

Ces déterminants contiennent x_j, en plus des coefficients donnés a_{ik}. Il est facile de démontrer que les déterminants (9) sont identiquement nuls. En effet, les éléments de la dernière colonne de ces déterminants sont, d'après:

$$X_i = a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n,$$

composés de n termes, et, par conséquent, chaque déterminant (9) peut être représenté, d'après la propriété IV de [I-1-3], par une somme de termes de la forme:

$$\begin{vmatrix} a_{11} & a_{12} & \dots a_{1k} & a_{1j} \\ a_{21} & a_{22} & \dots a_{2k} & a_{2j} \\ \dots & \dots & \dots & \dots \\ a_{k1} & a_{k2} & \dots a_{kk} & a_{kj} \\ a_{k+s,1} & a_{k+s,2} & \dots a_{k+s,k} & a_{k+s,j} \end{vmatrix} \cdot x_j.$$

Il est facile de montrer que le déterminant, facteur de x_j, est nul. En effet, si $j \leqslant k$, la dernière colonne de ce déterminant coïncide avec l'une des précédentes. Si $j > k$, ce déterminant est d'ordre $(k+1)$, il appartient au tableau (5) et, par conséquent, il est nul, puisque par hypothèse le rang de ce tableau est k. En retranchant les déterminants (9), identiquement nuls, des déterminants caractéristiques et en utilisant la propriété IV des déterminants, on peut mettre les déterminants caractéristiques sous la forme:

$$\Delta_{k+s} = \begin{vmatrix} a_{11} & a_{12} & \dots a_{1k} & b_1 - X_1 \\ a_{21} & a_{22} & \dots a_{2k} & b_2 - X_2 \\ \dots & \dots & \dots & \dots \\ a_{k1} & a_{k2} & \dots a_{kk} & b_k - X_k \\ a_{k+s,1} & a_{k+s,2} & \dots a_{k+s,k} & b_{k+s} - X_{k+s} \end{vmatrix} \tag{10}$$

$$(k+s = k+1,\ k+2,\ \dots,\ m),$$

et sous cette forme ils ne dépendent que formellement des x_j. Supposons maintenant que le système (6) ait la solution:

$$x_1 = x_1^{(0)},\ x_2 = x_2^{(0)},\ \dots,\ x_n = x_n^{(0)}.$$

En faisant $x_j = x_j^{(0)}$ dans la dernière colonne des déterminants (10), on a des zéros dans cette dernière colonne, c'est-à-dire que tous les déterminants caractéristiques doivent être nuls.

Théorème I. *Pour que le système* (6) *ait au moins une solution, il faut que tous les déterminants caractéristiques* (8) *soient nuls.*

Démontrons que cette condition est suffisante et donnons un procédé pour trouver toutes les solutions du système. Supposons que tous les déterminants caractéristiques soient nuls. Prenons-les sous la forme (10) et développons-les suivant les éléments de la dernière colonne. Il est facile de voir que le complémentaire algébrique de l'élément $(b_{k+s} - X_{k+s})$ est égal au déterminant principal Δ, non nul, et nous pouvons écrire la condition précédente (que tous les déterminants caractéristiques sont nuls) sous la forme:

$$\begin{aligned} &a_1^{(k+s)}(b_1 - X_1) + a_2^{(k+s)}(b_2 - X_2) + \ldots + a_k^{(k+s)}(b_k - X_k) + \\ &+ \Delta(b_{k+s} - X_{k+s}) = 0 \qquad (k+s = k+1, k+2, \ldots, m), \end{aligned} \tag{11}$$

où $a_p^{(q)}$ sont des coefficients numériques ne présentant pour nous aucun intérêt.

Soit maintenant une quelconque solution des k premières équations du système; substituons cette solution à x_j dans les identités (11). Toutes les différences

$$b_1 - X_1, \quad b_2 - X_2, \quad \ldots, \quad b_k - X_k$$

deviennent alors nulles, et l'on obtient après cette substitution:

$$\Delta \cdot (b_{k+s} - X_{k+s}) = 0$$

ou, d'après $\Delta \neq 0$:

$$b_{k+s} - X_{k+s} = 0 \qquad (k+s = k+1, k+2, \ldots, m),$$

c'est-à-dire que si tous les déterminants caractéristiques sont nuls, toute solution des k premières équations du système vérifie toutes les équations suivantes, et, dans ce cas, il ne reste donc qu'à résoudre les k premières équations.

Portons toutes les inconnues d'indice supérieur à k dans le second membre, les équations sont alors de la forme:

$$\left.\begin{aligned} &a_{11}x_1 + a_{12}x_2 + \ldots + a_{1k}x_k = b_1 - a_{1,\,k+1}x_{k+1} - \ldots - a_{1n}x_n, \\ &a_{21}x_1 + a_{22}x_2 + \ldots + a_{2k}x_k = b_2 - a_{2,\,k+1}x_{k+1} - \ldots - a_{2n}x_n, \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ &a_{k1}x_1 + a_{k2}x_2 + \ldots + a_{kk}x_k = b_k - a_{k,\,k+1}x_{k+1} - \ldots - a_{kn}x_n. \end{aligned}\right\} \tag{12}$$

Considérons ces équations comme un système pour déterminer $x_1, x_2, \ldots, x_k$. Le déterminant Δ de ce système étant non nul, nous pouvons donc obtenir une solution définie au moyen de la formule de Cramer. Il faut remarquer seulement que les termes constants du dernier système contiennent $x_{k+1}, \ldots, x_n$ auxquels on peut donner des valeurs arbitraires. Il résulte directement des formules de Cramer que la solution du système (12) est de la forme:

$$x_j = \alpha_j + \beta_{k+1}^{(j)} x_{k+1} + \ldots + \beta_n^{(j)} x_n \qquad (j = 1, 2, \ldots, k), \tag{13}$$

où α_s et $\beta_p^{(q)}$ sont des coefficients numériques et $x_{k+1}, \ldots, x_n$ sont arbitraires. Il découle de ce qui précède que ces formules donnent la solution la plus générale du système (6) lorsqu'on suppose que tous les déterminants caractéristiques sont nuls.

Théorème II. *Si tous les déterminants caractéristiques du système sont nuls, il suffit de résoudre les équations du système qui contiennent le déterminant principal par rapport aux inconnues dont les coefficients composent ce déterminant principal. Cette solution peut être donnée par les formules de Cramer et on a l'expression des k inconnues (où k est le rang du tableau de coefficients) sous forme de fonctions linéaires* (13) *des* (n — k) *autres inconnues dont les valeurs sont arbitraires.* On obtient ainsi toutes les solutions du système (6).

Si on compare les théorèmes I et II, il vient:

Théorème III. *La condition nécessaire et suffisante, pour que le système* (6) *ait des solutions, est que tous les déterminants caractéristiques de ce système soient nuls.*

Remarque: si $k = n$, c'est-à-dire si le rang est égal au nombre d'inconnues, le second membre des formules (13) ne contient pas de x_j et toutes les inconnues de x_1 à x_n peuvent être déterminées.

Théorème IV. *Pour qu'un système ait une solution définie, il faut et il suffit que tous les déterminants caractéristiques soient nuls et que le rang du tableau de ses coefficients soit égal au nombre d'inconnues.*

Les raisonnements précédents sont évidemment valables dans le cas où le nombre d'équations est égal au nombre d'inconnues, c'est-à-dire lorsque $m = n$.

Exemple. Soit un système de quatre équations à trois inconnues:

$$\begin{aligned} x - 3y - 2z &= -1, \\ 2x + y - 4z &= 3, \\ x + 4y - 2z &= 4, \\ 5x + 6y - 10z &= 10. \end{aligned}$$

Ecrivons le tableau des coefficients:

$$\left\| \begin{matrix} 1 & -3 & -2 \\ 2 & 1 & -4 \\ 1 & 4 & -2 \\ 5 & 6 & -10 \end{matrix} \right\|.$$

Il est facile de montrer que tous les déterminants du troisième ordre qui font partie de ce tableau sont nuls et que le déterminant du deuxième ordre qui se trouve dans le coin supérieur gauche n'est pas nul. On peut donc le prendre comme déterminant principal, et le rang du système est égal à deux. Formons les déterminants caractéristiques, au nombre de deux dans ce cas:

$$\Delta_3=\begin{vmatrix} 1 & -3 & -1 \\ 2 & 1 & 3 \\ 1 & 4 & 4 \end{vmatrix}=0\,;\quad \Delta_4=\begin{vmatrix} 1 & -3 & -1 \\ 2 & 1 & 3 \\ 5 & 6 & 10 \end{vmatrix}=0.$$

Ils sont tous deux nuls et, par conséquent, le système donné est compatible. C'est pourquoi il suffit de résoudre les deux premières équations par rapport à x et y, en portant z dans le second membre:

$$x-3y=2z-1,$$
$$2x+y=4z+3.$$

On obtient la solution sous la forme:

$$x=\frac{\begin{vmatrix} 2z-1 & -3 \\ 4z+3 & 1 \end{vmatrix}}{\begin{vmatrix} 1 & -3 \\ 2 & 1 \end{vmatrix}}=2z+\frac{8}{7},\qquad y=\frac{\begin{vmatrix} 1 & 2z-1 \\ 2 & 4z+3 \end{vmatrix}}{\begin{vmatrix} 1 & -3 \\ 2 & 1 \end{vmatrix}}=\frac{5}{7},$$

où z est arbitraire.

1-2-3. Systèmes homogènes. On dit qu'un système est homogène si tous ses termes constants b_i sont nuls. Si ce système homogène a des déterminants caractéristiques, leur dernière colonne est composée de zéros, et les déterminants sont tous nuls. Il est évident que tout système homogène possède la solution:

$$x_1=x_2=\ldots=x_n=0,$$

qu'on appellera par la suite solution triviale. La question fondamentale, pour un système homogène, est de savoir s'il possède des solutions non triviales et, s'il en a, quel sera l'ensemble de ces solutions. Soit d'abord le cas où le nombre d'équations est égal au nombre d'inconnues. Le système est de la forme:

$$\left.\begin{array}{l} a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n=0, \\ a_{21}x_1+a_{22}x_2+\ldots+a_{2n}x_n=0, \\ \cdots\cdots\cdots\cdots\cdots \\ a_{n1}x_1+a_{n2}x_2+\ldots+a_{nn}x_n=0. \end{array}\right\} \qquad (14)$$

Si le déterminant de ce système est différent de zéro, le système possède, d'après le théorème de Cramer, une solution définie, précisément, dans ce cas, la solution triviale. Si ce déterminant est nul, le rang k du tableau des coefficients est inférieur au nombre d'inconnues n et, par conséquent, les valeurs des $(n-k)$ inconnues sont arbitraires, et l'on aura un nombre infini de solutions non nulles. On arrive ainsi au théo-rème fondamental suivant.

Théorème I. *Pour que le système* (14) *ait une solution non nulle, il faut et il suffit que son déterminant soit nul.*

Etablissons un parallèle entre les résultats obtenus pour le système non homogène (1) et pour le système homogène (14). Si le déterminant du système n'est pas nul, le système non homogène (1) a une solution définie et le système homogène n'a que la solution triviale. Si le déterminant du système est nul, le système homogène (14) a des solutions non nulles, mais avec cette condition le système (1) ne possède en général pas de solutions, puisque pour qu'il en ait une, il faut que les seconds membres soient tels qu'ils annulent tous les déterminants caractéristiques. Ce parallélisme jouera un très grand rôle par la suite. On rencontrera en physique des systèmes homogènes dans l'étude des oscillations propres et des systèmes non homogènes dans l'étude des oscillations forcées ; le cas qu'on vient d'indiquer, à savoir celui de déterminant nul, caractérisera pour les systèmes homogènes les oscillations propres et pour les systèmes non homogènes le phénomène de résonance.

Passons maintenant à l'étude plus détaillée des solutions du système (14) dans le cas où son déterminant principal est nul. Soit k le rang du tableau des coefficients, il est évident que $k < n$. D'après un théorème démontré au paragraphe précédent, on doit prendre les k équations, qui contiennent le déterminant principal, et les résoudre par rapport aux k inconnues. Sans restreindre la généralité, on peut supposer que ces inconnues sont $x_1, \ldots, x_k$. Les solutions sont obtenues sous la forme:

$$x_j = \beta_{k+1}^{(j)} x_{k+1} + \ldots + \beta_n^{(j)} x_n \qquad (j = 1, 2, \ldots, k), \tag{15}$$

où $\beta_p^{(q)}$ sont des coefficients numériques définis et $x_{k+1}, \ldots, x_n$ peuvent prendre des valeurs arbitraires.

Notons une propriété générale de la solution du système (14) qui résulte directement de la linéarité et de l'homogénéité de ce système et qui est le principe de superposition des solutions : si on a plusieurs solutions du système:

$$x_s = x_s^{(1)}\,;\ x_s = x_s^{(2)}\,;\ x_s = x_s^{(3)}\,;\ \ldots\,;\ x_s = x_s^{(l)} \qquad (s = 1, 2, \ldots, n), \tag{16}$$

en les multipliant par des constantes arbitraires et en additionnant, on obtient encore une solution du système:

$$x_s = C_1 x_s^{(1)} + C_2 x_s^{(2)} + C_3 x_s^{(3)} + \ldots + C_l x_s^{(l)} \qquad (s = 1, 2, \ldots, n).$$

En procédant comme on l'a fait pour les équations différentielles linéaires (tome II, [II-1-3]), on dit que les solutions (16) sont linéairement indépendantes s'il n'existe pas de constantes C_i, non toutes nulles, telles que, pour tout s, on ait les égalités:

$$\sum_{i=1}^{l} C_i x_s^{(i)} = 0.$$

Il est facile de construire $(n - k)$ solutions linéairement indépendantes d'un système, telles que, si on les multiplie par des constantes arbitraires et qu'on les additionne, on obtienne toutes les solutions du système. En effet, en partant des formules (15) qui donnent la solution générale du système, construisons des solutions de la façon suivante : dans le premier cas, posons $x_{k+1} = 1$, et tous les autres x_{k+s} sont nuls ; dans le deuxième posons $x_{k+2} = 1$, et tous les autres x_{k+s} sont nuls, etc., et enfin, dans la dernière $(n - k)$-ième solution, posons $x_n = 1$, et tous les autres x_{k+s} sont nuls. Il est facile de voir que les solutions obtenues sont linéairement indépendantes puisque chacune d'elles contient une des inconnues égale à l'unité qui, dans les autres solutions, est nulle. On désigne les solutions obtenues par :

$$x_s = x_s^{(k+1)}; \quad x_s = x_s^{(k+2)}; \quad \ldots, \quad x_s = x_s^{(m)} \quad (s = 1, 2, \ldots, n).$$

Soit maintenant une solution quelconque du système (14). Elle est obtenue d'après les formules (15) pour certaines valeurs particulières :

$$x_{k+1} = \gamma_{k+1}; \quad x_{k+2} = \gamma_{k+2}; \quad \ldots; \quad x_n = \gamma_n.$$

Il est évident que cette solution est une combinaison linéaire des solutions que l'on a écrites :

$$x_s = \gamma_{k+1} x_s^{(k+1)} + \gamma_{k+2} x_s^{(k+2)} + \ldots + \gamma_n x_s^{(n)} \quad (s = 1, 2, \ldots, n).$$

Revenons à l'étude des solutions du système homogène (14) et montrons que, quelles que soient les solutions linéairement indépendantes que l'on a choisies, leur nombre total est égal à $(n - k)$.

Soit à nouveau le cas général de m équations homogènes à n inconnues. Si $m < n$, le rang k, qui ne peut pas être plus grand que m, est également inférieur à n, et $(n - k)$ inconnues sont arbitraires, c'est-à-dire que *si le nombre d'équations homogènes est inférieur au nombre d'inconnues, le système possède des solutions non nulles.*

En général, $k \leqslant n$, et pour $k = n$ le système n'a que la solution triviale.

I-2-4. Formes linéaires. La résolution des systèmes d'équations du premier degré est étroitement liée à celle des systèmes de formes linéaires. On entend par forme linéaire des variables $x_1, x_2, \ldots, x_n$ une fonction linéaire homogène de ces variables. Soient m formes linéaires :

$$y_s = a_{s1}x_1 + a_{s2}x_2 + \ldots + a_{sn}x_n \quad (s = 1, 2, \ldots, m). \qquad (17)$$

On dit que ces formes sont *linéairement dépendantes*, s'il existe des constantes $\alpha_1, \alpha_2, \ldots, \alpha_m$, non toutes nulles, telles que l'on ait, identiquement par rapport aux variables $x_1, x_2, \ldots, x_n$, la relation :

$$\alpha_1 y_1 + \alpha_2 y_2 + \ldots + \alpha_m y_m = 0.$$

S'il n'y a pas de telles constantes, les formes (17) sont dites *linéairement indépendantes*. Dans cette identité, on doit annuler les coefficients de toutes les variables x_l. Ainsi, l'identité précédente est équivalente au système suivant de n égalités :

$$\left.\begin{array}{l} \alpha_1 a_{11} + \alpha_2 a_{21} + \ldots + \alpha_m a_{m1} = 0, \\ \alpha_1 a_{12} + \alpha_2 a_{22} + \ldots + \alpha_m a_{m2} = 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ \alpha_1 a_{1n} + \alpha_2 a_{2n} + \ldots + \alpha_m a_{mn} = 0. \end{array}\right\}$$

Les formes y_s sont linéairement indépendantes si et seulement si ce système d'équations homogènes par rapport aux $\alpha_1, \alpha_2, \ldots, \alpha_m$ n'a que la solution triviale. Les résultats obtenus entraînent une série de conséquences concernant la dépendance linéaire des formes. Si $m > n$, le système homogène écrit a certainement des solutions non triviales, et les formes sont linéairement dépendantes. Pour que ces formes soient indépendantes, il faut et il suffit que le rang k du tableau des coefficients a_{pq} soit égal au nombre de formes m. *Si $m = n$, c'est-à-dire si le nombre de formes est égal au nombre de variables, il faut et il suffit, pour qu'elles soient linéairement indépendantes, que le déterminant de tout le tableau carré a_{pq} $(m = n)$ ne soit pas nul.* On dit dans ce cas qu'on a un *système complet de formes linéairement indépendantes*. Si $m \leqslant n$ et si les formes (17) sont linéairement indépendantes (c'est-à-dire si $k = m$), le système d'équations (17) peut, pour des valeurs quelconques de y_s, être résolu par rapport aux variables x_l dont les coefficients forment un déterminant d'ordre k, non nul, c'est-à-dire que *les formes linéairement indépendantes peuvent prendre un ensemble quelconque de valeurs de y_s*. Si $k = m = n$, toutes les variables x_l sont définies pour les y_s donnés.

Supposons à présent que $k < m$. Si en numérotant on fait correspondre les formes y_s et les variables x_l, on peut considérer qu'on a dans le coin supérieur gauche du tableau a_{pq} un déterminant d'ordre k, non nul. Les k premières formes : $y_1, y_2, \ldots, y_k$, sont alors linéairement indépendantes, et chacune des autres formes y_{k+l} est exprimée linéairement en fonction des k premières formes. En effet, pour les k premières formes, le rang du tableau des coefficients, égal à k, coïncide avec le nombre de formes, et c'est pourquoi les formes sont linéairement indépendantes. Si on prend $(k + 1)$ formes $y_1, y_2, \ldots$ $\ldots, y_k, y_{k+l}$, le rang du tableau de leurs coefficients, qui est encore égal à k, est inférieur au nombre de formes, de sorte que les formes sont linéairement dépendantes, c'est-à-dire qu'il existe des constantes β_s telles que $\beta_1 y_1 + \ldots + \beta_k y_k + \beta_{k+l} y_{k+l} = 0$. Dans cette relation, le coefficient β_{k+1} ne doit pas être nul, car dans le cas contraire les formes $y_1, y_2, \ldots, y_k$ seraient linéairement dépendantes, et on en déduirait que y_{k+1} s'exprime linéairement en fonction des k premières formes:

$$y_{k+1} = -\frac{\beta_1}{\beta_{k+1}} y_1 - \frac{\beta_2}{\beta_{k+1}} y_2 - \ldots - \frac{\beta_k}{\beta_{k+1}} y_k.$$

Le nombre k est le *rang du système de formes* (17). Ce nombre, égal au rang du tableau des coefficients, est, d'autre part, égal au plus grand nombre de formes linéairement indépendantes du système de formes (17).

Supposons qu'on ait k formes linéairement indépendantes y_1, $y_2, \ldots, y_k$, avec $k < n$. On peut considérer que le déterminant d'ordre k, se trouvant dans le coin supérieur gauche du tableau a_{pq}, n'est pas nul. Il est facile de voir qu'on peut compléter le système de ces k formes jusqu'à ce qu'on ait un système complet de n formes linéairement indépendantes. En effet, il suffit pour cela de supposer, par exemple, que:

$$y_{k+1} = x_{k+1}; \ \ldots; \ y_n = x_n.$$

Le déterminant des k formes obtenues sera:

$$\begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1k} & a_{1\,k+1} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2k} & a_{2\,k+1} & \ldots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{k1} & a_{k2} & \ldots & a_{kk} & a_{k\,k+1} & \ldots & a_{kn} \\ 0 & 0 & \ldots & 0 & 1 & \ldots & 0 \\ 0 & 0 & \ldots & 0 & 0 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \ldots & 0 & 0 & \ldots & 1 \end{vmatrix}.$$

Si l'on développe ce déterminant, en commençant par les éléments de la dernière ligne, puis suivant les éléments de l'avant-dernière ligne, etc., on verra qu'il est égal au déterminant d'ordre k qui se trouve dans le coin supérieur gauche, c'est-à-dire qu'il n'est pas nul. Par conséquent, les formes $y_1, y_2, \ldots, y_n$ sont effectivement linéairement indépendantes. *On peut compléter ainsi tout système de formes linéairement indépendantes en un système complet de formes linéairement indépendantes.*

I-2-5. Espace vectoriel à n dimensions. Donnons aux résultats obtenus ci-dessus un énoncé géométrique, dont on se servira par la suite. Pour cela, il faut introduire la notion de vecteur dans un espace à n dimensions: un vecteur est un ensemble ordonné de n nombres (complexes). Ainsi, *tout vecteur* $\mathbf{x}$ *est caractérisé par une suite de* n

nombres complexes que l'on appelle composantes de ce vecteur : $\mathbf{x}(x_1, x_2, \ldots, x_n)$. *L'ensemble de tous ces vecteurs forme un espace vectoriel* R_n *à n dimensions.*

On dit que deux vecteurs sont égaux si et seulement si toutes leurs composantes sont égales, c'est-à-dire que si on a deux vecteurs $\mathbf{u}(u_1, u_2, \ldots, u_n)$ *et* $\mathbf{v}(v_1, v_2, \ldots, v_n)$, *l'égalité vectorielle* $\mathbf{u} = \mathbf{v}$ *est équivalente aux n égalités scalaires suivantes :* $u_1 = v_1$; $u_2 = v_2$; . . . ; $u_n = v_n$. Définissons l'opération de multiplication d'un vecteur par un nombre et l'addition des vecteurs. La multiplication d'un vecteur par un nombre est la multiplication de toutes les composantes du vecteur par ce nombre, c'est-à-dire que si un vecteur $\mathbf{x}$ a les composantes $(x_1, x_2, \ldots, x_n)$, le vecteur $k\mathbf{x}$ a les composantes $(kx_1, kx_2, \ldots, kx_n)$. L'addition des vecteurs revient à additionner les composantes, c'est-à-dire que si l'on a les vecteurs $\mathbf{x}(x_1, x_2, \ldots, x_n)$ et $\mathbf{y}(y_1, y_2, \ldots, y_n)$, la somme $\mathbf{x}+\mathbf{y}$ a, par définition, les composantes $(x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n)$. On appelle vecteur nul un vecteur $(0, 0, \ldots, 0)$ dont toutes les composantes sont nulles. On désignera ce vecteur par le symbole θ. On a évidemment $\theta = 0\mathbf{x}$, où $\mathbf{x}$ est un vecteur quelconque, et $\mathbf{x} + \theta = \mathbf{x}$. Etudions maintenant la soustraction des vecteurs : le vecteur $\mathbf{x} - \mathbf{y}$ a les composantes $(x_1 - y_1, x_2 - y_2, \ldots, x_n - y_n)$. On a évidemment $\mathbf{x} - \mathbf{x} = \theta$ et $\mathbf{x} - \mathbf{y} = \mathbf{x} + (-1)\,\mathbf{y}$, c'est-à-dire que soustraire le vecteur $\mathbf{y}$ équivaut à l'additionner après l'avoir multiplié par le nombre (-1). On aura par la suite souvent l'occasion d'écrire des égalités vectorielles. Toute égalité de ce type est équivalente à n égalités scalaires, qui signifient que les composantes correspondantes des deux parties de l'égalité sont identiques. Par la suite, on n'utilisera pas le symbole θ pour le vecteur nul, mais il ne faut pas oublier que si dans une égalité vectorielle, on a zéro d'un côté, il faut considérer ce zéro comme le vecteur nul. Les propriétés habituelles de l'addition et de la multiplication résultent des propriétés données ci-dessus :

$$\mathbf{x}+\mathbf{y}=\mathbf{y}+\mathbf{x}\,; \qquad \mathbf{x}+(\mathbf{y}+\mathbf{z})=(\mathbf{x}+\mathbf{y})+\mathbf{z}\,;$$
$$(k_1+k_2)\,\mathbf{x}=k_1\mathbf{x}+k_2\mathbf{x}\,; \quad k\,(\mathbf{x}+\mathbf{y})=k\mathbf{x}+k\mathbf{y}\,; \quad k_1\,(k_2\mathbf{x})=(k_1k_2)\,\mathbf{x}.$$

On peut ainsi, dans une somme de vecteurs contenant un nombre quelconque de termes, les intervertir et les regrouper. Il résulte de l'égalité $\mathbf{x} + \mathbf{y} = \mathbf{z}$ que $\mathbf{x} = \mathbf{z} - \mathbf{y}$ et $\mathbf{y} = \mathbf{z} - \mathbf{x}$ et, réciproquement, il résulte de $\mathbf{x} - \mathbf{y} = \mathbf{z}$ que $\mathbf{x} = \mathbf{y} + \mathbf{z}$.

Introduisons à présent les notions de dépendance et d'indépendance linéaires de vecteurs. On dira que les vecteurs :

$$\mathbf{x}^{(1)}, \quad \mathbf{x}^{(2)}, \ \ldots, \ \mathbf{x}^{(l)} \tag{18}$$

sont des *vecteurs linéairement dépendants*, s'il existe des constantes $C_1, \ldots, C_l$, non toutes nulles, telles que :

$$C_1\mathbf{x}^{(1)}+C_2\mathbf{x}^{(2)}+\ldots+C_l\mathbf{x}^{(l)}=0. \tag{19}$$

Si ces constantes n'existent pas, on dira que les vecteurs (18) sont des *vecteurs linéairement indépendants*. On désigne les composantes des vecteurs $\mathbf{x}^{(j)}$ par $(x_1^{(j)}, x_2^{(j)}, \ldots, x_n^{(j)})$. La condition (19) est évidemment équivalente au système de n équations à l inconnues C_1, $C_2, \ldots, C_l$:

$$\left.\begin{aligned} x_1^{(1)}C_1+x_1^{(2)}C_2+\ldots+x_1^{(l)}C_l=0,\\ x_2^{(1)}C_1+x_2^{(2)}C_2+\ldots+x_2^{(l)}C_l=0,\\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\ x_n^{(1)}C_1+x_n^{(2)}C_2+\ldots+x_n^{(l)}C_l=0. \end{aligned}\right\} \tag{20}$$

En utilisant les résultats obtenus ci-dessus pour un système homogène, il est facile de tirer une série de conséquences et de donner à ces résultats un énoncé géométrique. Supposons d'abord que $l > n$, c'est-à-dire que le nombre de vecteurs soit plus grand que le nombre de dimensions de l'espace. Le nombre d'équations est alors inférieur au nombre d'inconnues dans le système homogène (20) et ce système, comme on le sait, aura certainement pour les inconnues C_j des solutions non triviales, c'est-à-dire que les vecteurs sont certainement linéairement dépendants. Autrement dit, *le plus grand nombre de vecteurs linéairement indépendants est égal au nombre de dimensions de l'espace*. Soit maintenant le cas où $l = n$. Le système (20) contient alors le même nombre d'équations et d'inconnues et a des solutions non triviales si et seulement si son déterminant est nul, c'est-à-dire que si l'on a dans un espace à n dimensions n vecteurs, et si l'on forme un déterminant à partir des n^2 composantes de ces vecteurs, en plaçant, par exemple, les composantes de chaque vecteur dans une colonne déterminée et en considérant que l'indice de la ligne correspond à l'indice de la composante, alors pour que les vecteurs soient linéairement indépendants, il faut et il suffit que ce déterminant ne soit pas nul. La valeur de ce déterminant est analogue au volume d'un parallélépipède dans un espace réel à trois dimensions.

Nous pouvons considérer, dans un déterminant quelconque $|b_{ik}|$ d'ordre n, les éléments de chaque colonne $(b_{1k}, b_{2k}, \ldots, b_{nk})$ comme les composantes d'un vecteur $\mathbf{b}^{(k)}$, et la valeur du déterminant sera alors fonction des n vecteurs $\mathbf{b}^{(1)}, \ldots, \mathbf{b}^{(n)}$. Si ce déterminant est nul, ces vecteurs sont linéairement dépendants.

La valeur du déterminant peut être considérée comme une fonction des vecteurs $\mathbf{b}^{(k)}$:

$$|b_{ik}|=\Delta(\mathbf{b}^{(1)}, \mathbf{b}^{(2)}, \ldots, \mathbf{b}^{(n)}).$$

Puisque la valeur du déterminant varie lorsqu'on intervertit deux colonnes, on peut affirmer que la fonction Δ ne changera que de signe si l'on intervertit deux de ses variables. On dit d'habitude que cette fonction est *antisymétrique*. Il est facile de voir, par exemple, que le déterminant de Vandermonde D_n que l'on a étudié en [I-1-5] est également une fonction antisymétrique de ses variables $x_1, \ldots, x_n$.

Revenons à l'étude du système (20) et à la question de l'indépendance linéaire des vecteurs $\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(l)}$ en supposant $l \leqslant n$. On désignera par k le rang du tableau formé par les composantes $x_p^{(q)}$. Si $k = l$, le système, comme on l'a vu, n'a que la solution triviale, c'est-à-dire que les vecteurs sont linéairement indépendants. Si $k < l$, le système aura certainement une solution non nulle, c'est-à-dire que *pour que des vecteurs soient linéairement indépendants, il faut et il suffit que leur nombre soit égal au rang du tableau de leurs composantes.* On suppose maintenant que $k < l$, c'est-à-dire que les vecteurs sont linéairement dépendants. On choisira k vecteurs dont les composantes contiennent un déterminant non nul d'ordre k (on peut le faire parfois de plusieurs façons différentes). D'après ce qui a été démontré ci-dessus, ces vecteurs sont linéairement indépendants. Il est facile de voir que chacun des autres vecteurs peut s'exprimer linéairement en fonction des vecteurs que l'on a choisis. En effet, soient $\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(k)}$ des vecteurs linéairement indépendants. En joignant à ces vecteurs un vecteur quelconque $\mathbf{x}^{(k+s)}$, on obtiendra $(k+1)$ vecteurs qui seront linéairement dépendants, puisque le rang k du tableau de leurs composantes est inférieur à leur nombre $l = k + 1$. On aura ainsi des constantes C_i $(i = 1, 2, \ldots, k, k+s)$, non toutes nulles, telles que:

$$C_1\mathbf{x}^{(1)} + C_2\mathbf{x}^{(2)} + \ldots + C_k\mathbf{x}^{(k)} + C_{k+s}\mathbf{x}^{(k+s)} = 0.$$

On a alors certainement $C_{k+s} \neq 0$, car, dans le cas contraire, les vecteurs $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \ldots, \mathbf{x}^{(k)}$ seraient linéairement dépendants. De l'égalité ci-dessus, il résulte que:

$$\mathbf{x}^{(k+s)} = -\frac{C_1}{C_{k+s}}\mathbf{x}^{(1)} - \frac{C_2}{C_{k+s}}\mathbf{x}^{(2)} - \ldots - \frac{C_k}{C_{k+s}}\mathbf{x}^{(k)},$$

c'est-à-dire que $\mathbf{x}^{(k+s)}$ s'exprime linéairement en fonction de $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \ldots, \mathbf{x}^{(k)}$. Soient $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \ldots, \mathbf{x}^{(n)}$ n vecteurs quelconques linéairement indépendants. On peut donner, comme exemple de vecteurs de ce type, les vecteurs:

$$(1, 0, 0, \ldots, 0); \quad (0, 1, 0, \ldots, 0); \quad \ldots; \quad (0, 0, 0, \ldots, 1). \qquad (21)$$

Si l'on prend un vecteur $\mathbf{x}$ arbitraire, les $(n+1)$ vecteurs $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \ldots, \mathbf{x}^{(n)}, \mathbf{x}$ sont, comme on l'a vu ci-dessus, certainement linéairement dépendants:

$$C_1\mathbf{x}^{(1)} + C_2\mathbf{x}^{(2)} + \ldots + C_n\mathbf{x}^{(n)} + C\mathbf{x} = 0,$$

la constante C étant alors certainement non nulle car, dans le cas contraire, les vecteurs $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \ldots, \mathbf{x}^{(n)}$ seraient linéairement dépendants. Il résulte de ce qui précède qu'*un vecteur quelconque* $\mathbf{x}$ *s'ex-*

prime en fonction de n vecteurs linéairement indépendants:

$$\mathbf{x} = \alpha_1\mathbf{x}^{(1)} + \alpha_2\mathbf{x}^{(2)} + \ldots + \alpha_n\mathbf{x}^{(n)} \left(\alpha_s = -\frac{C_s}{C}\right). \tag{22}$$

Il est facile de voir qu'*il n'existe qu'une expression de* $\mathbf{x}$ *en fonction des* $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \ldots, \mathbf{x}^{(n)}$. En effet, s'il existait en plus de l'expression (22) une autre:

$$\mathbf{x} = \beta_1\mathbf{x}^{(1)} + \beta_2\mathbf{x}^{(2)} + \ldots + \beta_n\mathbf{x}^{(n)},$$

dans laquelle on ait des coefficients β_s différents des α_s correspondants, on obtiendrait, en retranchant les deux relations membre à membre:

$$(\alpha_1 - \beta_1)\,\mathbf{x}^{(1)} + (\alpha_2 - \beta_2)\,\mathbf{x}^{(2)} + \ldots + (\alpha_n - \beta_n)\,\mathbf{x}^{(n)} = 0,$$

c'est-à-dire que les vecteurs $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \ldots, \mathbf{x}^{(n)}$ seraient linéairement dépendants. Si l'on prend pour les vecteurs $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \ldots, \mathbf{x}^{(n)}$ les vecteurs (21), les nombres α_s de la formule (22) coïncideront, évidemment, avec les composantes x_s du vecteur $\mathbf{x}\,(x_1, x_2, \ldots, x_n)$. On peut, dans le cas général, les appeler également *composantes du vecteur* $\mathbf{x}$, si $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \ldots, \mathbf{x}^{(n)}$ sont pris pour vecteurs de base. En donnant aux nombres α_s toutes les valeurs complexes possibles, on obtient tous les vecteurs de l'espace à n dimensions. Supposons que l'on ait k vecteurs linéairement indépendants:

$$\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \ldots, \mathbf{x}^{(k)}, \tag{23}$$

on a alors $k < n$. On dit que l'ensemble des vecteurs obtenus d'après la formule:

$$\mathbf{y} = C_1\mathbf{x}^{(1)} + C_2\mathbf{x}^{(2)} + \ldots + C_k\mathbf{x}^{(k)}, \tag{23$_1$}$$

où les C_s sont des constantes arbitraires, forme un *sous-espace* L_k de dimension k. On peut montrer, comme ci-dessus, que tout vecteur appartenant à L_k s'exprime de façon unique en fonction de $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \ldots, \mathbf{x}^{(k)}$. On dit encore que les *vecteurs* (23) *forment un sous-espace* L_k.

R e m a r q u e. Si un vecteur $\mathbf{z}$ appartient à L_k, c'est-à-dire s'il s'exprime par une formule du type (23$_1$), le vecteur $c\mathbf{z}$, où c est une constante quelconque, s'exprime aussi par une formule du type (23$_1$), c'est-à-dire qu'il appartient aussi à L_k. De même, si $\mathbf{z}^{(1)}$ et $\mathbf{z}^{(2)}$ appartiennent à L_k, leur somme $\mathbf{z}^{(1)} + \mathbf{z}^{(2)}$ appartient aussi à L_k. Il en résulte immédiatement une propriété plus générale: *si des vecteurs* $\mathbf{z}^{(1)}, \mathbf{z}^{(2)}, \ldots, \mathbf{z}^{(p)}$ *appartiennent à* L_k, *toute combinaison linéaire de ces vecteurs* $\gamma_1\mathbf{z}^{(1)} + \gamma_2\mathbf{z}^{(2)} + \ldots + \gamma_p\mathbf{z}^{(p)}$ *appartient aussi à* L_k.

Soient m vecteurs appartenant à L_k:

$$y^{(s)} = C_1^{(s)}\mathbf{x}^{(1)} + C_2^{(s)}\mathbf{x}^{(2)} + \ldots + C_k^{(s)}\mathbf{x}^{(k)} \quad (s = 1, 2, \ldots, m). \tag{24}$$

Puisque les vecteurs (23) sont linéairement indépendants, une relation de la forme:

$$\alpha_1 \mathbf{y}^{(1)} + \alpha_2 \mathbf{y}^{(2)} + \ldots + \alpha_m \mathbf{y}^{(m)} = 0$$

équivaut à un système de k équations homogènes:

$$\alpha_1 C_q^{(1)} + \alpha_2 C_q^{(2)} + \ldots + \alpha_m C_q^{(m)} = 0 \qquad (q = 1, 2, \ldots, k).$$

Si ce système a des solutions non nulles, les vecteurs (24) sont liné, airement dépendants. En particulier, si $m > k$, on a certainement des solutions non nulles, c'est-à-dire que tout ensemble de vecteurs dont le nombre est supérieur à k, d'un sous-espace formé par les vecteurs (23), sera un ensemble de vecteurs linéairement dépendants. Il en résulte que le sous-espace formé par les vecteurs linéairement indépendants (23) ne peut être engendré par aucun ensemble de vecteurs linéairement indépendants $\mathbf{z}^{(1)}, \ldots, \mathbf{z}^{(l)}$, dont le nombre l est inférieur à k. En effet, d'après ce qui a été démontré ci-dessus, il ne peut exister dans un tel sous-espace plus de l vecteurs linéairement indépendants, et, d'autre part, les vecteurs linéairement indépendants (23), dont le nombre k est supérieur à l, doivent appartenir à ce sous-espace. *Si l'on prend k vecteurs linéairement indépendants quelconques $\mathbf{u}^{(1)}$, $\mathbf{u}^{(2)}, \ldots, \mathbf{u}^{(k)}$ appartenant à L_k, ils forment, au sens indiqué ci-dessus, ce même sous-espace L_k.*

En effet, d'après la définition du sous-espace, toute combinaison linéaire:

$$C_1 \mathbf{u}^{(1)} + C_2 \mathbf{u}^{(2)} + \ldots + C_k \mathbf{u}^{(k)}$$

appartient à L_k. D'autre part, soit un vecteur quelconque $\mathbf{y}$ de L_k. Les vecteurs $\mathbf{u}^{(1)}, \ldots, \mathbf{u}^{(k)}, \mathbf{y}$, en nombre $(k+1)$, appartiennent à L_k et, d'après ce qui précède, doivent être linéairement dépendants:

$$\beta_1 \mathbf{u}^{(1)} + \beta_2 \mathbf{u}^{(2)} + \ldots + \beta_k \mathbf{u}^{(k)} + \gamma \mathbf{y} = 0,$$

et $\mathbf{u}^{(1)}, \ldots, \mathbf{u}^{(k)}$ étant linéairement indépendants, le coefficient γ ne doit pas être nul, c'est-à-dire que tout vecteur $\mathbf{y}$ de L_k est exprimé par $\mathbf{u}^{(1)}, \ldots, \mathbf{u}^{(k)}$; ces vecteurs forment donc L_k. Si $m = k$ dans les formules (24) et si le déterminant des coefficients $C_p^{(q)}$ n'est pas nul, les vecteurs $\mathbf{y}^{(1)}, \ldots, \mathbf{y}^{(k)}$ seront des vecteurs linéairement indépendants de L_k. Il est aisé, dans le cas général, de montrer que le nombre de vecteurs linéairement indépendants, donnés par les formules (24), est égal au rang du tableau $C_p^{(q)}$.

On a vu ci-dessus que, si le vecteur $\mathbf{z}$ appartient à un sous-espace L, le vecteur $c\mathbf{z}$ (la constante c étant quelconque) appartient aussi à L, et si $\mathbf{z}^{(1)}$ et $\mathbf{z}^{(2)}$ appartiennent à L, $(\mathbf{z}^{(1)} + \mathbf{z}^{(2)})$ appartient aussi à L. On pourrait donner une nouvelle définition du sous-espace et appeler *sous-espace* un ensemble de vecteurs tel que si $\mathbf{z}$ appartient à L, cz appartient aussi à L et si $z^{(1)}$ et $z^{(2)}$ appartiennent à L, $(z^{(1)}+z^{(2)})$ appartient aussi à L. Il en résulte immédiatement que toute combinaison linéaire de vecteurs, appartenant à L, appartient aussi à L. On vient de voir que les propriétés énoncées à l'aide de la nouvelle définition résultent de la définition précédente du sous-espace. Réciproquement, montrons que de la nouvelle définition découle la précédente, c'est-à-dire que *les deux définitions sont équivalentes.*

Soit un vecteur $x^{(1)}$ appartenant à L. D'après la définition du sous-espace L, les vecteurs $C_1x^{(1)}$, pour C_1 quelconque, appartiennent aussi à L. Si ces vecteurs épuisent le sous-espace L, on a L_1 au sens précédent. Dans le cas contraire, un vecteur $x^{(2)}$ linéairement indépendant de $x^{(1)}$ appartient à L, et les vecteurs $C_1x^{(1)}+C_2x^{(2)}$, C_1 et C_2 étant quelconques, appartiennent à L. Si ces vecteurs épuisent tout le sous-espace L, L coïncide alors avec un sous-espace L_2 au sens précédent. Dans le cas contraire, un vecteur $x^{(3)}$, tel que $x^{(1)}$, $x^{(2)}$, $x^{(3)}$ soient linéairement indépendants, appartient à L. En continuant ainsi, par adjonction d'un nombre fini de vecteurs $x^{(i)}$ linéairement indépendants, on épuise L puisqu'il existe au plus n vecteurs linéairement indépendants. Le nombre total k de ces vecteurs $x^{(i)}$ donne la dimension du sous-espace L. Si l'on a $k=n$, L coïncide avec l'espace à n dimensions tout entier.

Il faut noter un fait lié à la formation du sous-espace. Soient des vecteurs $x^{(1)}, x^{(2)}, \ldots, x^{(k)}$ linéairement dépendants. On dit dans ce cas que la formule (23_1) définit le sous-espace L. On suppose que parmi ces vecteurs, les l premiers: $x^{(1)}, x^{(2)}, \ldots, x^{(l)}$, sont linéairement indépendants et que chacun des vecteurs suivants $x^{(l+1)}, \ldots$ $\ldots, x^{(k)}$ s'exprime linéairement en fonction des l premiers. L'ensemble des vecteurs définis par la formule (23_1) coïncidera alors avec l'ensemble des vecteurs définis par la formule:

$$y=C_1x^{(1)}+C_2x^{(2)}+\ldots+C_lx^{(l)},$$

c'est-à-dire que le sous-espace L engendré par les vecteurs linéairement dépendants $x^{(1)}, x^{(2)}, \ldots, x^{(k)}$ a pour dimensions l $(l<k)$.

Soit un espace réel à trois dimensions dans lequel on repère les vecteurs à partir d'un point déterminé O (origine). Dans ce cas $n=3$. Si $k=1$, le sous-espace L_1 est une droite passant par O, et L_2 est un plan passant par O.

I-2-6. Produit scalaire. Soit α un nombre complexe, on désigne par $\bar{\alpha}$ le conjugué complexe et par $|\alpha|$ le module de α, de sorte que $\alpha\bar{\alpha}=|\alpha|^2$. Si α est réel, $\bar{\alpha}=\alpha$ et $|\alpha|^2=\alpha^2$. Introduisons à présent une notion nouvelle qui joue un grand rôle dans la suite.

D é f i n i t i o n. *On appelle produit scalaire de deux vecteurs*

$$x(x_1, x_2, \ldots, x_n) \text{ et } y(y_1, y_2, \ldots, y_n)$$

la quantité égale à la somme:

$$\sum_{s=1}^{n} x_s\bar{y}_s.$$

On désigne le produit scalaire par le symbole (x, y). On a:

$$(\mathbf{x}, \mathbf{y}) = \sum_{s=1}^{n} x_s\bar{y}_s; \qquad (\mathbf{y}, \mathbf{x}) = \sum_{s=1}^{n} y_s\bar{x}_s,$$

d'où:

$$(\mathbf{y}, \mathbf{x}) = \overline{(\mathbf{x}, \mathbf{y})}.$$

On dit que *deux vecteurs sont orthogonaux, si leur produit scalaire est nul.* Dans la mesure où le conjugué complexe de zéro est zéro. l'ordre des vecteurs dans le produit scalaire ne joue aucun rôle dans la condition d'orthogonalité. Il est facile de voir que le vecteur nul (0, 0, . . ., 0) est orthogonal à tout vecteur x.

Les propriétés suivantes résultent de la définition du produit scalaire:

$$(\alpha\mathbf{x}, \mathbf{y}) = \alpha(\mathbf{x}, \mathbf{y}); \quad (\mathbf{x}, \alpha\mathbf{y}) = \bar{\alpha}(\mathbf{x}, \mathbf{y}),$$

où α est un facteur numérique. De plus:

$$(\mathbf{x}+\mathbf{y}, \mathbf{z}) = (\mathbf{x}, \mathbf{z}) + (\mathbf{y}, \mathbf{z}); \quad (\mathbf{x}, \mathbf{y}+\mathbf{z}) = (\mathbf{x}, \mathbf{y}) + (\mathbf{x}, \mathbf{z}),$$

et la loi de distributivité est vraie quel que soit le nombre de composantes. Il en résulte, entre autres, que:

$$(\mathbf{x}+\mathbf{y}, \mathbf{u}+\mathbf{v}) = (\mathbf{x}, \mathbf{u}) + (\mathbf{x}, \mathbf{v}) + (\mathbf{y}, \mathbf{u}) + (\mathbf{y}, \mathbf{v}).$$

Formons le produit scalaire du vecteur $\mathbf{x}(x_1, x_2, \ldots, x_n)$ par lui-même:

$$(\mathbf{x}, \mathbf{x}) = \sum_{s=1}^{n} x_s\bar{x}_s = \sum_{s=1}^{n} |x_s|^2.$$

Nous obtenons ainsi un nombre réel qui est positif pour tout vecteur x non nul et nul pour le vecteur nul (0, 0, . . ., 0). *La racine carrée (valeur arithmétique) du nombre réel* (x, x) *est la norme ou la longueur du vecteur* x. En désignant cette norme par le symbole $\|\mathbf{x}\|$, on peut écrire:

$$\|\mathbf{x}\|^2 = (\mathbf{x}, \mathbf{x}) = \sum_{s=1}^{n} |x_s|^2; \quad \|\mathbf{x}\| = \sqrt{(\mathbf{x}, \mathbf{x})} = \sqrt{\sum_{s=1}^{n} |x_s|^2}.$$

L'égalité $\|\mathbf{x}\| = 0$ équivaut à ce que x soit nul. Soient trois vecteurs perpendiculaires x, y et z, c'est-à-dire

$$(\mathbf{x}, \mathbf{y}) = 0; \qquad (\mathbf{x}, \mathbf{z}) = 0; \qquad (\mathbf{y}, \mathbf{z}) = 0.$$

En appliquant la loi de distributivité d'un produit scalaire et en considérant les égalités précédentes, on obtient :

$$(x+y+z,\ x+y+z)=(x,x)+(y,y)+(z,z)$$

ou

$$\|x+y+z\|^2=\|x\|^2+\|y\|^2+\|z\|^2.$$

Cette formule exprime le *théorème de Pythagore*. Elle est vraie pour un nombre quelconque de composantes. Il est important que les vecteurs considérés soient orthogonaux deux à deux. On va montrer que *si les vecteurs* $x^{(1)}, x^{(2)}, \ldots, x^{(l)}$ *sont orthogonaux deux à deux et non nuls, ils sont linéairement indépendants*. En effet, supposons que :

$$\sum_{s=1}^{l} C_s x^{(s)}=0$$

et montrons que tous les C_s doivent être nuls. Multiplions scalairement les deux membres de cette égalité par $x^{(k)}$, où k est un des nombres 1, 2, ..., l :

$$\sum_{s=1}^{l} C_s (x^{(s)}, x^{(k)})=0.$$

Puisque les vecteurs $x^{(s)}$ sont orthogonaux deux à deux, $(x^{(s)}, x^{(k)}) = 0$ pour $s\neq k$, et la dernière formule devient : $C_k\,(x^{(k)}, x^{(k)}) = 0$, c'est-à-dire que $C_k\,\|x^{(k)}\|^2 = 0$, d'où il résulte que $C_k = 0$ puisque $\|x^{(k)}\|^2>0$, et cela quel que soit k.

I-2-7. Interprétation géométrique des systèmes homogènes. Soit le système homogène :

$$\left.\begin{aligned} a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n&=0,\\ a_{21}x_1+a_{22}x_2+\ldots+a_{2n}x_n&=0,\\ \ldots\ldots\ldots\ldots\ldots\ldots&\ldots\\ a_{n1}x_1+a_{n2}x_2+\ldots+a_{nn}x_n&=0. \end{aligned}\right\} \tag{25}$$

Introduisons les vecteurs :

$$a^{(1)}(\bar a_{11}, \bar a_{12}, \ldots, \bar a_{1n});\ \ldots;\quad a^{(n)}(\bar a_{n1}, \bar a_{n2}, \ldots, \bar a_{nn}). \tag{26}$$

Le système (25) peut alors être mis sous la forme condensée suivante :

$$(x, a^{(1)})=0;\ \ldots;\quad (x, a^{(n)})=0, \tag{27}$$

c'est-à-dire qu'il faut trouver un vecteur x perpendiculaire à tous les vecteurs $a^{(i)}$. Si le déterminant $|a_{ik}|$ n'est pas nul, le déterminant $|\bar a_{ik}|$, qui est son conjugué complexe, est aussi non nul. Dans ce cas, les vecteurs $a^{(j)}$ sont linéairement indépendants, et le système (27) n'a qu'une solution nulle, c'est-à-dire qu'il n'existe pas de vecteur (excepté le vecteur nul) qui soit simultanément perpendiculaire, dans un espace à n dimensions, à n vecteurs linéairement indépendants.

Considérons maintenant le cas où le déterminant du système (25) est nul et soit k le rang du système. Si l'on forme le tableau des éléments conjugués complexes, les déterminants seront conjugués complexes des déterminants des a_{ik} ; le rang du tableau conjugué complexe sera aussi k. Il y a ainsi k vecteurs linéairement indépendants parmi les vecteurs $a^{(j)}$ et les autres vecteurs sont des combinaisons linéaires de ces vecteurs. Sans restreindre la généralité, on peut supposer que ces vecteurs linéairement indépendants sont :

$$a^{(1)}, \ldots, a^{(k)}, \tag{28}$$

et l'on a pour les autres des expressions de la forme :

$$a^{(k+s)} = \beta_1^{(k+s)} a^{(1)} + \ldots + \beta_k^{(k+s)} a^{(k)} \qquad (k+s = k+1, k+2, \ldots, n),$$

où les $\beta_p^{(q)}$ sont des coefficients numériques. Il en résulte que si x est perpendiculaire aux vecteurs (28), il est perpendiculaire à tous les vecteurs $a^{(j)}$. En effet :

$$(x, a^{(k+s)}) = \sum_{i=1}^{k} (x, \overline{\beta}_i^{(k+s)} a^{(i)})$$

et toute la somme est nulle puisque chacune des composantes s'annule par hypothèse. Il suffit donc de résoudre les k premières équations du système. En considérant, comme toujours, que le déterminant d'ordre k, non nul, se trouve dans l'angle supérieur gauche, on obtient $(n-k)$ solutions $x^{(1)}, \ldots, x^{(n-k)}$ linéairement indépendantes pour le vecteur x cherché, en utilisant le procédé indiqué en [I-2-5] ; chaque solution est une combinaison linéaire de ces $(n-k)$ vecteurs. On peut dire que, dans ce cas, les vecteurs définis par la formule :

$$y = C_1 a^{(1)} + \ldots + C_k a^{(k)},$$

où les C_i sont des constantes arbitraires, engendrent un espace L_k à k dimensions, qui est un sous-espace de l'espace à n dimensions. De même, les vecteurs $x^{(1)}, x^{(2)}, \ldots, x^{(n-k)}$ engendrent un sous-espace M_{n-k} à $(n-k)$ dimensions. Le *sous-espace M_{n-k} est orthogonal au sous-espace L_k* en ce sens que tout vecteur de M_{n-k} est orthogonal à tout vecteur de L_k et réciproquement. Le sous-espace M_{n-k} est composé de vecteurs qui vérifient le système (27), c'est-à-dire qui sont orthogonaux à $a^{(1)}, a^{(2)}, \ldots, a^{(k)}$. Il est aisé de voir que les n vecteurs $a^{(1)}, \ldots, a^{(k)}, x^{(1)}, \ldots, x^{(n-k)}$ sont linéairement indépendants.

En effet, on suppose qu'ils sont liés par la relation:

$$(c_1\mathbf{a}^{(1)} + \ldots + c_k\mathbf{a}^{(k)}) + (d_1\mathbf{x}^{(1)} + \ldots + d_{n-k}\mathbf{x}^{(n-k)}) = 0. \qquad (29)$$

Le premier terme représente un vecteur $\mathbf{a}$ de L_k et le second un vecteur $\mathbf{x}$ de M_{n-k}, et l'on a $\mathbf{a} + \mathbf{x} = 0$ ou $\mathbf{a} = -\mathbf{x}$. Mais les vecteurs $\mathbf{a}$ et $\mathbf{x}$ sont orthogonaux, c'est-à-dire que le vecteur $\mathbf{a}$ est orthogonal à lui-même, donc $(\mathbf{a}, \mathbf{a}) = 0$ ou $\| \mathbf{a} \| = 0$, d'où il résulte que le vecteur $\mathbf{a}$ est un vecteur nul. On peut en dire autant du vecteur $\mathbf{x}$. Ainsi:

$$c_1\mathbf{a}^{(1)} + \ldots + c_k\mathbf{a}^{(k)} = 0 \text{ et } d_1\mathbf{x}^{(1)} + \ldots + d_{n-k}\mathbf{x}^{(n-k)} = 0.$$

Mais les vecteurs $\mathbf{a}^{(1)}, \ldots, \mathbf{a}^{(k)}$ sont par hypothèse linéairement indépendants et, par conséquent, toutes les constantes c_s doivent être nulles; il en est de même pour les d_s. Ainsi, tous les coefficients de la relation (29) doivent être nuls, c'est-à-dire que les vecteurs $\mathbf{a}^{(1)}, \ldots, \mathbf{a}^{(k)}, \mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(n-k)}$ sont linéairement indépendants.

Tout vecteur $\mathbf{x}$ peut être mis de façon unique sous la forme:

$$\mathbf{x} = (\gamma_1\mathbf{a}^{(1)} + \ldots + \gamma_k\mathbf{a}^{(k)}) + (\delta_1\mathbf{x}^{(1)} + \ldots + \delta_{n-k}\mathbf{x}^{(n-k)}),$$

le premier terme donnant un vecteur appartenant à L_k et le second un vecteur appartenant à M_{n-k}. Les vecteurs qui forment M_{n-k} sont, comme on l'a déjà signalé, toutes les solutions possibles du système (27) et, quel que soit le système complet de solutions linéairement indépendantes choisi, il y a $(n - k)$ solutions, c'est-à-dire autant que la dimension de M_{n-k}. L'étude que l'on vient de faire d'un système homogène donne un résultat très important:

Si l'on a un sous-espace L_k à k dimensions ($k < n$), les vecteurs orthogonaux à ce sous-espace forment un sous-espace M_{n-k} de dimension ($n - k$) et tout vecteur $\mathbf{x}$ *de R_n peut être mis sous forme d'une somme* $\mathbf{x} = \mathbf{y} + \mathbf{z}$, *où* $\mathbf{y}$ *appartient à L_k et* $\mathbf{z}$ *à M_{n-k}.*

Montrons que *cette représentation est unique*. Soit encore une autre représentation: $\mathbf{x} = \mathbf{u} + \mathbf{v}$, où $\mathbf{u}$ appartient à L_k et $\mathbf{v}$ à M_{n-k}. Il faut montrer que $\mathbf{u} = \mathbf{y}$ et $\mathbf{v} = \mathbf{z}$. On a $\mathbf{y} + \mathbf{z} = \mathbf{u} + \mathbf{v}$, d'où $\mathbf{y} - \mathbf{u} = \mathbf{v} - \mathbf{z}$. La différence $\mathbf{y} - \mathbf{u}$ appartient à L_k et la différence $\mathbf{v} - \mathbf{z}$ appartient à M_{n-k}, d'où il résulte que le vecteur $\mathbf{y} - \mathbf{u}$ est orthogonal à lui-même, c'est-à-dire que $(\mathbf{y} - \mathbf{u}, \mathbf{y} - \mathbf{u}) = 0$ ou $\| \mathbf{y} - \mathbf{u} \| = 0$, d'où $\mathbf{y} - \mathbf{u} = 0$ et $\mathbf{y} = \mathbf{u}$. Il résulte de $\mathbf{y} - \mathbf{u} = \mathbf{v} - \mathbf{z}$ que $\mathbf{v} = \mathbf{z}$. Le vecteur $\mathbf{y}$ qui se trouve dans la représentation $\mathbf{x} = \mathbf{y} + \mathbf{z}$ est la *projection de* $\mathbf{x}$ *sur le sous-espace L_k*. Dans cette représentation les vecteurs $\mathbf{y}$ et $\mathbf{z}$ sont orthogonaux et le théorème de Pythagore donne: $\| \mathbf{x} \|^2 = \| \mathbf{y} \|^2 + \| \mathbf{z} \|^2$, d'où $\| \mathbf{y} \| \leqslant \| \mathbf{x} \|$ et on a l'égalité seulement dans le cas où $\mathbf{z}$ est le vecteur nul, c'est-à-dire lorsque $\mathbf{x}$ appartient à L_k, de sorte que $\mathbf{y} = \mathbf{x}$. De façon analogue, $\| \mathbf{z} \| \leqslant \| \mathbf{x} \|$, et on n'a l'égalité que dans le cas où $\mathbf{x}$ est orthogonal à L_k, c'est-à-dire que $\mathbf{z} = \mathbf{x}$. Les sous-espaces L_k et M_{n-k} sont des *sous-espaces orthogonaux supplémentaires*. Si $k = n$, L_n coïncide avec R_n, et M_0 se réduit au vecteur nul.

Soit un espace réel à trois dimensions et soit $k = 2$, de sorte que $n - k = 3 - 2 = 1$. Le sous-espace L_2 est un plan P passant par le point O et M_1 est une droite passant par O et perpendiculaire à P. Tout vecteur peut être représenté de façon unique comme la somme de deux vecteurs dont l'un se trouve dans le plan P et l'autre sur la droite K. Nous venons de donner une interprétation géométrique de la solution d'un système homogène dans le cas où le nombre d'équations est égal au nombre d'inconnues. On peut faire le même raisonnement dans le cas général où le nombre de vecteurs $\mathbf{a}^{(s)}$ n'est pas obligatoirement égal à n. La même remarque est valable pour le paragraphe suivant.

1-2-8. Cas d'un système non homogène. Soit le système non homogène :

$$\left.\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n = b_n. \end{array}\right\} \tag{30}$$

On peut l'interpréter comme le problème de la détermination d'un vecteur $\mathbf{x}(x_1, \ldots, x_n)$ pour des vecteurs donnés (26) dans le système :

$$(\mathbf{x}, \mathbf{a}^{(1)}) = b_1; \quad \ldots; \qquad (\mathbf{x}, \mathbf{a}^{(n)}) = b_n. \tag{31}$$

Si le déterminant du système n'est pas nul, le théorème de Cramer donne une solution définie. Supposons que le déterminant du système soit nul et que le rang du tableau de ses coefficients soit égal à k, le déterminant d'ordre k, non nul, se trouvant, comme toujours, dans le coin supérieur gauche. Outre le système (30), écrivons un système d'équations homogènes dont les coefficients sont obtenus à partir des coefficients du système donné en remplaçant les lignes par les colonnes et tous les nombres par leurs conjugués complexes. Ce nouveau système est de la forme :

$$\left.\begin{array}{l} \overline{a}_{11}y_1 + \overline{a}_{21}y_2 + \ldots + \overline{a}_{n1}y_n = 0, \\ \overline{a}_{12}y_1 + \overline{a}_{22}y_2 + \ldots + \overline{a}_{n2}y_n = 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \overline{a}_{1n}y_1 + \overline{a}_{2n}y_2 + \ldots + \overline{a}_{nn}y_n = 0. \end{array}\right\} \tag{32}$$

Le tableau de ses coefficients a, comme auparavant, le rang k, et le déterminant d'ordre k qui se trouve dans l'angle supérieur gauche est, comme auparavant, différent de zéro. On dit que ce système homogène est *associé* au système (30). Comme on l'a vu ci-dessus, sa solution générale est une combinaison linéaire de $(n-k)$ solutions (de vecteurs) que l'on peut obtenir, par exemple, en résolvant suivant le théorème de Cramer les k premières équations par rapport à $y_1, \ldots, y_k$, en supposant que les autres y_{k+s} sont nuls, sauf un qui est égal à l'unité. On arrive ainsi pour $y_{k+1}=1$ au système :

$$\left.\begin{array}{l}\bar{a}_{11}y_1+\bar{a}_{21}y_2+\ldots+\bar{a}_{k1}y_k=-\bar{a}_{k+1,\,1},\\ \bar{a}_{12}y_1+\bar{a}_{22}y_2+\ldots+\bar{a}_{k2}y_k=-\bar{a}_{k+1,\,2},\\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\ \bar{a}_{1k}y_1+\bar{a}_{2k}y_2+\ldots+\bar{a}_{kk}y_k=-\bar{a}_{k+1,\,k}.\end{array}\right\}$$

En résolvant ce système et en passant aux conjugués complexes, on obtient :

$$\bar{y}_m=-\frac{\Delta'_m}{\Delta'}\quad(m=1,2,\ldots,k) \tag{33}$$

$$\bar{y}_{k+1}=1\,;\quad \bar{y}_{k+2}=\bar{y}_{k+3}=\ldots=\bar{y}_n=0,$$

où

$$\Delta'=\begin{vmatrix}a_{11} & a_{21} & \ldots & a_{k1}\\ a_{12} & a_{22} & \ldots & a_{k2}\\ \ldots & \ldots & \ldots & \ldots\\ a_{1k} & a_{2k} & \ldots & a_{kk}\end{vmatrix}\neq 0,$$

et on obtient Δ'_m à partir de Δ' en remplaçant les éléments de la colonne d'indice m par $a_{k+1,1}, \ldots, a_{k+1,k}$. La condition pour que le vecteur **b** de composantes $(b_1, \ldots, b_n)$ soit perpendiculaire au vecteur **y**, qu'on vient d'obtenir en résolvant le système (32), est :

$$(\mathbf{b},\mathbf{y})=-\sum_{m=1}^{k}\frac{\Delta'_m}{\Delta'}b_m+b_{k+1}=0$$

ou :

$$-\sum_{m=1}^{k}\Delta'_m b_m+\Delta' b_{k+1}=0. \tag{34}$$

En remplaçant les lignes par les colonnes dans le déterminant Δ'_m, puis en faisant passer la ligne d'indice m à la dernière place au moyen de $(k - m)$ permutations des dernières lignes, nous obtenons :

$$-\Delta'_m = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1k} \\ \cdot & \cdot & \cdot & \cdot \\ a_{m-1,1} & a_{m-1,2} & \dots & a_{m-1,k} \\ a_{m+1,1} & a_{m+1,2} & \dots & a_{m+1,k} \\ \cdot & \cdot & \cdot & \cdot \\ a_{k1} & a_{k2} & \dots & a_{kk} \\ a_{k+1,1} & a_{k+1,2} & \dots & a_{k+1,k} \end{vmatrix} \cdot (-1)^{k+1+m}.$$

C'est précisément le complémentaire algébrique de l'élément b_m dans le déterminant caractéristique :

$$\Delta_{k+1} = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1k} & b_1 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{k1} & a_{k2} & \dots & a_{kk} & b_k \\ a_{k+1,1} & a_{k+1,2} & \dots & a_{k+1,k} & b_{k+1} \end{vmatrix},$$

et ce déterminant caractéristique est nul d'après la condition (34). De façon analogue, pour $y_{k+s} = 1$, nous obtenons la condition $\Delta_{k+s} = 0$. On arrive ainsi au résultat suivant : *si le déterminant du système* (30) *est nul, pour que ce système ait une solution, il faut et il suffit que le vecteur* $(b_1, \dots, b_n)$ *soit orthogonal à tous les vecteurs qui donnent la solution du système homogène associé* (32).

La solution générale du système (30) sera la somme d'une solution particulière de ce système et de la solution générale du système homogène correspondant que l'on obtient à partir de (30) en remplaçant tous les b_j par zéro. Cette solution générale du système homogène contient $(n - k)$ constantes arbitraires.

Indiquons encore une autre interprétation géométrique du théorème fondamental de résolution des systèmes, importante pour la suite. Soient n formes linéaires à n variables indépendantes :

$$\begin{gathered} y_1 = a_{11}x_1 + \dots + a_{1n}x_n, \\ \dots\dots\dots\dots \\ y_n = a_{n1}x_1 + \dots + a_{nn}x_n. \end{gathered}$$

On suppose que les x_s peuvent prendre des valeurs complexes arbitraires et soient $(y_1, \dots, y_n)$ les composantes d'un vecteur. Si le déterminant $| a_{ik} |$ n'est pas nul, on obtient des valeurs définies de x_k pour des valeurs quelconques de y_k et les formules précédentes donnent tout l'espace à n dimensions des y. Supposons que le tableau $\| a_{ik} \|$ soit de rang $r < n$. Sans restreindre la généralité, on peut considérer que le déterminant d'ordre r, qui se trouve dans le coin supérieur gauche, n'est pas nul. En vertu du théorème fondamental de résolution des systèmes, l'ensemble des valeurs $(y_1, \ldots, y_n)$, obtenues au moyen des formules précédentes, possède la propriété suivante: les valeurs $y_1, \ldots, y_r$ peuvent être arbitraires, mais si on les fixe, les valeurs $y_{r+1}, \ldots, y_n$ seront parfaitement définies, à savoir elles s'obtiennent en écrivant que le déterminant caractéristique est nul. Dans le langage géométrique cela signifie que les formules précédentes définissent un sous-espace à r dimensions, engendré par les vecteurs obtenus en posant égal à l'unité l'un des y_s $(s = 1, 2, \ldots, r)$ et les autres nuls. Donc, si le tableau $\| a_{ik} \|$ est de rang r, les formules précédentes donnent un ensemble de valeurs $(y_1, \ldots, y_n)$, qui engendrent un sous-espace à r dimensions.

Nous avons étudié le cas où le nombre de formes linéaires est égal au nombre de variables x_s. On peut prendre le cas général:

$$\begin{aligned} y_1 &= a_{11}x_1 + \ldots + a_{1n}x_n, \\ &\ldots\ldots\ldots\ldots\ldots \\ y_m &= a_{m1}x_1 + \ldots + a_{mn}x_n. \end{aligned}$$

Dans ce cas, pour des x_s arbitraires, ces formules définissent dans un espace à m dimensions un sous-espace de dimension égale au rang du tableau $\| a_{ik} \|$. La démonstration est la même que la précédente.

I-2-9. Déterminant de Gram. Inégalité d'Hadamard. Soient m vecteurs:

$$\mathbf{x}^{(s)}\,(x_1^{(s)}, x_2^{(s)}, \ldots, x_n^{(s)}) \qquad (s = 1, 2, \ldots, m).$$

Formons le déterminant d'ordre m à partir des produits scalaires $(\mathbf{x}^{(i)}, \mathbf{x}^{(k)})$ et introduisons la notation particulière:

$$G(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \ldots, \mathbf{x}^{(m)}) = |(\mathbf{x}^{(i)}, \mathbf{x}^{(k)})|_1^m =$$

$$= \begin{vmatrix} (\mathbf{x}^{(1)}, \mathbf{x}^{(1)}) & (\mathbf{x}^{(1)}, \mathbf{x}^{(2)}) & \ldots & (\mathbf{x}^{(1)}, \mathbf{x}^{(m)}) \\ (\mathbf{x}^{(2)}, \mathbf{x}^{(1)}) & (\mathbf{x}^{(2)}, \mathbf{x}^{(2)}) & \ldots & (\mathbf{x}^{(2)}, \mathbf{x}^{(m)}) \\ \ldots & \ldots & \ldots & \ldots \\ (\mathbf{x}^{(m)}, \mathbf{x}^{(1)}) & (\mathbf{x}^{(m)}, \mathbf{x}^{(2)}) & \ldots & (\mathbf{x}^{(m)}, \mathbf{x}^{(m)}) \end{vmatrix}. \tag{35}$$

Ce déterminant s'appelle *déterminant de Gram des vecteurs*:

$$\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \ldots, \mathbf{x}^{(m)}.$$

Considérons séparément les différents cas: $m = n$, $m < n$ et $m > n$. Le terme général du déterminant est de la forme:

$$(\mathbf{x}^{(i)}, \mathbf{x}^{(k)}) = \sum_{s=1}^{n} x_s^{(i)} \overline{x_s^{(k)}}.$$

Pour $m = n$, le déterminant (35) est égal au produit des déterminants:

$$\begin{vmatrix} x_1^{(1)} & x_2^{(1)} & \ldots & x_n^{(1)} \\ x_1^{(2)} & x_2^{(2)} & \ldots & x_n^{(2)} \\ \ldots & \ldots & \ldots & \ldots \\ x_1^{(n)} & x_2^{(n)} & \ldots & x_n^{(n)} \end{vmatrix} \cdot \begin{vmatrix} \overline{x_1^{(1)}} & \overline{x_1^{(2)}} & \ldots & \overline{x_1^{(n)}} \\ \overline{x_2^{(1)}} & \overline{x_2^{(2)}} & \ldots & \overline{x_2^{(n)}} \\ \ldots & \ldots & \ldots & \ldots \\ \overline{x_n^{(1)}} & \overline{x_n^{(2)}} & \ldots & \overline{x_n^{(n)}} \end{vmatrix}.$$

En considérant que les valeurs du déterminant sont invariables lorsqu'on remplace les lignes par les colonnes, on peut affirmer que le second déterminant est le conjugué complexe du premier et, par conséquent, la valeur du déterminant de Gram (35) est égale, pour $m = n$, au carré du module du déterminant $|x_k^{(i)}|_1^n$, formé par les composantes $x_k^{(i)}$ des vecteurs $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$. Ainsi, le déterminant (35) est positif si les vecteurs indiqués sont linéairement indépendants et nul s'ils sont linéairement dépendants [I-2-5]. Pour $m \neq n$, on a deux tableaux rectangulaires:

$$\left\| \begin{matrix} x_1^{(1)} & x_2^{(1)} & \ldots & x_n^{(1)} \\ x_1^{(2)} & x_2^{(2)} & \ldots & x_n^{(2)} \\ \cdot & \cdot & \cdot & \cdot \\ x_1^{(m)} & x_2^{(m)} & \ldots & x_n^{(m)} \end{matrix} \right\| \quad (36_1) \quad \text{et} \quad \left\| \begin{matrix} \overline{x_1^{(1)}} & \overline{x_1^{(2)}} & \ldots & \overline{x_1^{(m)}} \\ \overline{x_2^{(1)}} & \overline{x_2^{(2)}} & \ldots & \overline{x_2^{(m)}} \\ \cdot & \cdot & \cdot & \cdot \\ \overline{x_n^{(1)}} & \overline{x_n^{(2)}} & \ldots & \overline{x_n^{(m)}} \end{matrix} \right\|, \quad (36_2)$$

et le tableau associé au déterminant (35) est le produit des deux derniers tableaux [I-1-7]. Le déterminant (35) est nul pour $m > n$, d'après le théorème démontré en [I-1-7]. Mais, dans ce cas, les vecteurs $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \ldots, \mathbf{x}^{(m)}$ sont linéairement dépendants [I-2-5]. Pour $m < n$, d'après le théorème de [I-1-7]:

$$G(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \ldots, \mathbf{x}^{(m)}) = \\ = \sum_{r_1 < r_2 < \ldots < r_m} X\begin{pmatrix} 1, 2, \ldots, m \\ r_1, r_2, \ldots, r_m \end{pmatrix} Y\begin{pmatrix} r_1, r_2, \ldots, r_m \\ 1, 2, \ldots, m \end{pmatrix},$$

où $X\begin{pmatrix} 1, 2, \ldots, m \\ r_1, r_2, \ldots, r_m \end{pmatrix}$ désigne les mineurs du tableau (36_1) et $Y\begin{pmatrix} r_1, r_2, \ldots, r_m \\ 1, 2, \ldots, m \end{pmatrix}$ les mineurs du tableau (36_2). Comme ci-dessus, $Y\begin{pmatrix} r_1, r_2, \ldots, r_m \\ 1, 2, \ldots, m \end{pmatrix}$ est le conjugué complexe de $X\begin{pmatrix} 1, 2, \ldots, m \\ r_1, r_2, \ldots, r_m \end{pmatrix}$ et la dernière formule s'écrit sous la forme:

$$G(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \ldots, \mathbf{x}^{(m)}) = \sum_{r_1 < r_2 < \ldots < r_m} \left| X\begin{pmatrix} 1, 2, \ldots, m \\ r_1, r_2, \ldots, r_m \end{pmatrix} \right|^2. \quad (37)$$

Si les vecteurs $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \ldots, \mathbf{x}^{(m)}$ sont linéairement indépendants, le tableau (36_1) est de rang m [I-2-5], et il y a au moins un terme positif parmi les termes non négatifs du second membre de l'égalité (37). Si les vecteurs sont linéairement dépendants, le tableau (36_1) est de rang inférieur à m, tous les déterminants d'ordre m, contenus dans ce tableau, sont nuls et il résulte de (37) que, dans ce cas, $G(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \ldots, \mathbf{x}^{(m)}) = 0$. L'étude des trois cas: $m = n$, $m > n$ et $m < n$ conduit au théorème général suivant:

Théorème. *Le déterminant de Gram* $G(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \ldots, \mathbf{x}^{(m)})$ *est positif si les vecteurs* $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \ldots, \mathbf{x}^{(m)}$ *sont linéairement indépendants et nul s'ils sont linéairement dépendants.*

Donnons maintenant une formule pour le déterminant de Gram. Soit $\mathbf{x}$ un vecteur quelconque de R_n et soit la décomposition $\mathbf{x} = \mathbf{y} + \mathbf{z}$, où $\mathbf{y}$ appartient au sous-espace défini par les vecteurs $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \ldots, \mathbf{x}^{(m)}$ et où $\mathbf{z}$ est orthogonal à ce sous-espace. La formule que nous voulons démontrer est de la forme:

$$G(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \ldots, \mathbf{x}^{(m)}, \mathbf{x}) = \|\mathbf{z}\|^2 G(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \ldots, \mathbf{x}^{(m)}). \quad (38)$$

En tenant compte des égalités: $(x^{(s)}, x) = (x^{(s)}, y)$; $(x, x^{(s)}) = (y, x^{(s)})$ qui résultent de ce que z est orthogonal à tous les $x^{(s)}$, et de la formule $(x, x) = (y, y) + (z, z)$ [I-2-6] nous pouvons écrire:

$$G(x^{(1)}, x^{(2)}, \ldots, x^{(m)}, x) = \begin{vmatrix} (x^{(1)}, x^{(1)}) & (x^{(1)}, x^{(2)}) & \ldots & (x^{(1)}, x^{(m)}) & (x^{(1)}, y) \\ (x^{(2)}, x^{(1)}) & (x^{(2)}, x^{(2)}) & \ldots & (x^{(2)}, x^{(m)}) & (x^{(2)}, y) \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ (x^{(m)}, x^{(1)}) & (x^{(m)}, x^{(2)}) & \ldots & (x^{(m)}, x^{(m)}) & (x^{(m)}, y) \\ (y, x^{(1)}) & (y, x^{(2)}) & \ldots & (y, x^{(m)}) & (y, y) + (z, z) \end{vmatrix}.$$

En mettant les éléments de la dernière ligne sous la forme:

$$(y, x^{(1)}) + 0, (y, x^{(2)}) + 0, \ldots, (y, x^{(m)}) + 0, (y, y) + (z, z)$$

et en représentant le déterminant sous forme d'une somme de deux déterminants, nous pouvons écrire d'après la propriété IV de [I-1-3]:

$$G(x^{(1)}, x^{(2)}, \ldots, x^{(m)}, x) = G(x^{(1)}, x^{(2)}, \ldots, x^{(m)}, y) + \begin{vmatrix} (x^{(1)}, x^{(1)}) & (x^{(1)}, x^{(2)}) & \ldots & (x^{(1)}, x^{(m)}) & (x^{(1)}, y) \\ (x^{(2)}, x^{(1)}) & (x^{(2)}, x^{(2)}) & \ldots & (x^{(2)}, x^{(m)}) & (x^{(2)}, y) \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ (x^{(m)}, x^{(1)}) & (x^{(m)}, x^{(2)}) & \ldots & (x^{(m)}, x^{(m)}) & (x^{(m)}, y) \\ 0 & 0 & \ldots & 0 & \| z \|^2 \end{vmatrix}. \qquad (39)$$

Le vecteur y appartient au sous-espace déterminé par les vecteurs $x^{(1)}, x^{(2)}, \ldots, x^{(m)}$, c'est-à-dire qu'il s'exprime linéairement en fonction des $x^{(s)}$ et, d'après le théorème démontré:

$$G(x^{(1)}, x^{(2)}, \ldots, x^{(m)}, y) = 0.$$

En développant le déterminant de la formule (39) suivant les éléments de la dernière ligne, nous en déduisons la formule (38).

Il résulte de cette formule que:

$$G(x^{(1)}, x^{(2)}, \ldots, x^{(m)}, x) \leqslant \| x \|^2 G(x^{(1)}, x^{(2)}, \ldots, x^{(m)}). \qquad (40)$$

Si les vecteurs $x^{(s)}$ sont linéairement dépendants:

$$G(x^{(1)}, x^{(2)}, \ldots, x^{(m)}, x) = G(x^{(1)}, x^{(2)}, \ldots, x^{(m)}) = 0.$$

S'ils sont linéairement indépendants, l'inégalité (40) se transforme en égalité dans le cas et seulement dans le cas où $y = 0$, c'est-à-dire lorsque x est orthogonal à tous les $x^{(s)}$.

Si nous appliquons plusieurs fois l'inégalité (40) au déterminant de Gram initial $G(x^{(1)}, x^{(2)}, \ldots, x^{(m)})$, nous en déduisons la majoration:

$$G(x^{(1)}, x^{(2)}, \ldots, x^{(m)}) \leqslant \| x^{(1)} \|^2 \| x^{(2)} \|^2 \ldots \| x^{(m)} \|^2. \qquad (41)$$

Il faut, pour cela, considérer que $G(x^{(1)}) = \| x^{(1)} \|^2$.

On a l'égalité dans (41) dans le cas et seulement dans le cas où les vecteurs sont orthogonaux deux à deux. On suppose en outre qu'aucun des vecteurs n'est nul. L'inégalité obtenue donne facilement une majoration pour n'importe quel déterminant. Soit Δ un déterminant d'ordre n d'éléments a_{ik}. Considérons les éléments de la ligne d'indice i de ce déterminant $(a_{i1}, a_{i2}, \ldots, a_{in})$ comme les composantes d'un vecteur $x^{(i)}$ de R_n. Formons un nouveau déterminant dont les éléments sont $\bar{a}_{ik}$ conjugués complexes des a_{ik}. Il est évidemment égal à $\bar{\Delta}$.

Si l'on multiplie Δ par $\overline{\Delta}$ d'après la règle de la ligne par la ligne, on obtient le déterminant de Gram $G(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \ldots, \mathbf{x}^{(n)})$; la valeur de ce déterminant est égale à $\Delta\overline{\Delta}$ d'après le théorème de la multiplication des déterminants, c'est-à-dire à $|\Delta|^2$. En appliquant l'inégalité (41), on obtient la majoration suivante du module d'un déterminant, que l'on doit à Hadamard :

$$|\Delta|^2 \leqslant \sum_{k=1}^{n} |a_{1k}|^2 \cdot \sum_{k=1}^{n} |a_{2k}|^2 \ldots \sum_{k=1}^{n} |a_{nk}|^2. \tag{42}$$

Si le déterminant Δ a des éléments réels, on peut écrire :

$$\Delta^2 \leqslant \sum_{k=1}^{n} a_{1k}^2 \cdot \sum_{k=1}^{n} a_{2k}^2 \ldots \sum_{k=1}^{n} a_{nk}^2. \tag{43}$$

Si l'on a la majoration $|a_{ik}| \leqslant M$ $(i, k = 1, 2, \ldots, n)$ pour les éléments du déterminant Δ, on a évidemment :

$$\sum_{k=1}^{n} |a_{ik}|^2 \leqslant nM^2,$$

et (42) donne la majoration :

$$|\Delta| \leqslant n^{\frac{n}{2}} M^n. \tag{44}$$

D'après ce qui vient d'être exposé, on a l'égalité dans (42) dans le cas et seulement dans le cas où les vecteurs $\mathbf{x}^{(i)}$ sont orthogonaux deux à deux.

On peut obtenir d'autres valeurs pour le déterminant de Gram en généralisant l'inégalité (40).

Supposons que chacune des lettres X, Y, Z désigne une suite de vecteurs de R_n. L'inégalité est de la forme :

$$G(X, Y, Z)\, G(X) \leqslant G(X, Z)\, G(X, Y). \tag{45}$$

On n'exclut pas, dans cette inégalité, le cas d'une suite vide de vecteurs, c'est-à-dire d'une suite qui ne contient aucun vecteur. Si W est une suite vide, il faut poser $G(W) = 1$.

A partir de cette inégalité on peut obtenir la valeur suivante du déterminant de Gram :

$$G(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \ldots, \mathbf{x}^{(m)}) \leqslant \left[\prod_{k=1}^{m} G(\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(k-1)}, \mathbf{x}^{(k+1)}, \ldots, \mathbf{x}^{(m)})\right]^{\frac{1}{m-1}},$$

où Π est le signe du produit. L'application répétée de cette majoration donne de nouvelles majorations dans lesquelles le déterminant de Gram contient un nombre plus petit de vecteurs. Pour toutes ces majorations l'égalité est vraie dans le cas et seulement dans le cas où les vecteurs sont orthogonaux deux à deux.

I-2-10. Systèmes d'équations différentielles linéaires à coefficients constants. Appliquons les résultats obtenus au problème de l'intégration des systèmes d'équations différentielles linéaires à coefficients constants. Ecrivons ce type de système sous la forme :

$$\left.\begin{array}{l} x_1' = a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n, \\ x_2' = a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ x_n' = a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n, \end{array}\right\} \qquad (46)$$

où les x_j sont des fonctions cherchées de t, les x_j' leurs dérivées et les a_{ik} des constantes données. Cherchons une solution sous la forme :

$$x_1 = b_1e^{\lambda t}; \quad x_2 = b_2e^{\lambda t}; \ \ldots; \quad x_n = b_ne^{\lambda t}. \qquad (47)$$

En substituant dans le système (46) et en simplifiant par $e^{\lambda t}$, nous avons, pour déterminer les constantes $b_1, \ldots, b_n$, le système d'équations :

$$\left.\begin{array}{l} (a_{11} - \lambda)\, b_1 + a_{12}b_2 + \ldots + a_{1n}b_n = 0, \\ a_{21}b_1 + (a_{22} - \lambda)\, b_2 + \ldots + a_{2n}b_n = 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{n1}b_1 + a_{n2}b_2 + \ldots + (a_{nn} - \lambda)\, b_n = 0. \end{array}\right\} \qquad (48)$$

Puisque nous devons obtenir une solution non triviale pour les inconnues b_j, le déterminant de ce système doit être nul, c'est-à-dire que l'on obtient l'équation suivante pour la constante λ :

$$\begin{vmatrix} a_{11} - \lambda & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} - \lambda & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} - \lambda \end{vmatrix} = 0. \qquad (49)$$

Cette équation s'appelle *équation séculaire*. On la rencontre dans l'étude des oscillations des systèmes mécaniques sans frottement dans le cas particulier où le tableau des coefficients a_{ik} est symétrique, c'est-à-dire lorsque $a_{ik} = a_{ki}$ et que tous les coefficients sont réels ; on étudiera cette question par la suite pour les petites oscillations. Pour le moment, étudions le système dans le cas général. L'équation (49) est une équation algébrique de degré n dont le terme de degré le plus élevé est $(-\lambda)^n$. Si cette équation a n racines distinctes :

$$\lambda = \lambda_1, \ \ldots, \ \lambda = \lambda_n,$$

on a, en faisant $\lambda = \lambda_j$ dans les coefficients du système (48), n équations homogènes à n inconnues $b_1, \ldots, b_n$ de déterminant nul et l'on peut obtenir une solution non triviale pour ces inconnues. D'après les formules (47), on a ainsi n solutions linéairement indépendantes du système (46) et leur combinaison linéaire donne l'intégrale générale du système. Si l'équation séculaire (49) a des racines multiples, la solution du problème sera plus compliquée : à toute racine de l'équation (49) d'ordre de multiplicité k devront correspondre k solutions linéairement indépendantes du système (46) ; l'une de ces solutions a la forme (47) et les autres contiennent en plus un facteur : un polynôme de t. Mais dans ce cas, contrairement à celui d'une équation à coefficients constants, il peut arriver (tome II, [II-1-15]) que pas seulement une mais plusieurs solutions (ou même toutes), correspondant à la racine multiple, soient de la forme (47).

Revenons à l'équation séculaire (49) qui joue un rôle important dans ce problème. La résolution de cette équation, même approchée, présente, pour des n grands, des difficultés pratiques parce que l'inconnue cherchée λ n'est pas dans une colonne ou une ligne quelconques, mais figure dans les termes diagonaux. Le développement du premier membre de cette équation nécessite d'importants calculs que l'on a déjà indiqués [I-1-4]. Donnons un exemple de transformation de l'équation (49), qui la met sous une forme plus commode, λ ne figurant que dans une colonne.

Formons la combinaison linéaire des quantités cherchées :

$$\xi=\alpha_{01}x_1+\alpha_{02}x_2+\ldots+\alpha_{0n}x_n, \tag{50}$$

où les α_{0j} sont des coefficients numériques pris de façon arbitraire. On différentie l'équation (50) n fois par rapport à t en remplaçant chaque fois les dérivées x_j' du second membre par leurs expressions extraites du système (46). Nous obtenons ainsi $(n+1)$ équations de la forme :

$$\left.\begin{aligned}
&\xi=\alpha_{01}x_1+\alpha_{02}x_2+\ldots+\alpha_{0n}x_n,\\
&\xi'=\alpha_{11}x_1+\alpha_{12}x_2+\ldots+\alpha_{1n}x_n,\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&\xi^{(n-1)}=\alpha_{n-1,1}x_1+\alpha_{n-1,2}x_2+\ldots+\alpha_{n-1,n}x_n,\\
&\xi^{(n)}=\alpha_{n1}x_1+\alpha_{n2}x_2+\ldots+\alpha_{nn}x_n.
\end{aligned}\right\} \tag{51}$$

Supposons que le déterminant constitué par les coefficients α_{ik} qui figurent dans les n premières équations soit non nul. Les n premières équations donnent l'expression de x_j en fonction de $\xi, \xi', \ldots, \xi^{(n-1)}$, et si l'on porte ces valeurs dans la dernière équation, on a une équation d'ordre n en ξ. On peut directement éliminer x_j des $(n+1)$ équations (51) en utilisant les déterminants. En effet, écrivons ces équations de la façon suivante :

$$\left.\begin{aligned}
&\xi x_0+\alpha_{01}x_1+\alpha_{02}x_2+\ldots+\alpha_{0n}x_n=0,\\
&\xi' x_0+\alpha_{11}x_1+\alpha_{12}x_2+\ldots+\alpha_{1n}x_n=0,\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&\xi^{(n)}x_0+\alpha_{n1}x_1+\alpha_{n2}x_2+\ldots+\alpha_{nn}x_n=0,
\end{aligned}\right\}$$

où $x_0=-1$, et on considère ces $(n+1)$ équations comme des équations homogènes par rapport aux $x_0, x_1, \ldots, x_n$.

Le déterminant de ce système doit être nul, ce qui donne le résultat suivant :

$$\begin{vmatrix}
\xi & \alpha_{01} & \alpha_{02} & \ldots & \alpha_{0n}\\
\xi' & \alpha_{11} & \alpha_{12} & \ldots & \alpha_{1n}\\
\ldots & \ldots & \ldots & \ldots & \ldots\\
\xi^{(n)} & \alpha_{n1} & \alpha_{n2} & \ldots & \alpha_{nn}
\end{vmatrix}=0. \tag{52}$$

Cherchons la solution de cette équation sous la forme:

$$\xi = e^{\lambda t}.$$

La portant dans la première colonne du déterminant (52) et sortant le facteur $e^{\lambda t}$ de la première colonne, nous avons, après simplification, l'équation suivante pour déterminer λ:

$$\begin{vmatrix} 1 & \alpha_{01} & \alpha_{02} & \dots & \alpha_{0n} \\ \lambda & \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ \dots & \dots & \dots & \dots & \dots \\ \lambda^n & \alpha_{n1} & \alpha_{n2} & \dots & \alpha_{nn} \end{vmatrix} = 0. \tag{53}$$

Il est facile de démontrer que l'équation (53) a, par hypothèse, les mêmes racines que l'équation (49). En effet, soit $\lambda = \lambda_0$ une racine de l'équation (53). On a alors pour (52) une solution de la forme:

$$\xi = Ce^{\lambda_0 t}, \tag{54}$$

où C est une constante arbitraire. Les n premières équations du système (51) donnent des solutions de x_j de la forme (47) pour $\lambda = \lambda_0$, c'est-à-dire que $\lambda = \lambda_0$ est aussi une racine de l'équation (49). Réciproquement, si $\lambda = \lambda_0$ est une racine de l'équation (49), on a pour x_j une solution de la forme (47) où les b_j sont des constantes non toutes nulles. En portant ces expressions de x_j dans la première équation du système (51), on obtient pour ξ une solution de la forme (54); cette solution est certainement non triviale. En effet, dans le cas contraire on aurait $\xi = \xi' = \dots = \xi^{(n-1)} = 0$, d'où il résulte en vertu des n premières équations du système (51) que: $x_1 = x_2 = \dots = x_n = 0$.

Ainsi, toute racine $\lambda = \lambda_0$ de l'équation (49) est effectivement une racine de l'équation (53). L'équation (53) a donc, compte tenu des hypothèses, les mêmes racines que l'équation (49).

Les calculs sont très simples si l'on prend la formule (50) sous la forme $\xi = x_1$. Dans ce cas, l'équation (53) est de la forme:

$$\begin{vmatrix} 1 & 1 & 0 & \dots & 0 \\ \lambda & \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ \dots & \dots & \dots & \dots & \dots \\ \lambda^n & \alpha_{n1} & \alpha_{n2} & \dots & \alpha_{nn} \end{vmatrix} = 0.$$

Soit à la place du système (46) un système d'équations du deuxième ordre de la forme:

$$\left.\begin{aligned} x_1'' &= a_{11}x_1 + a_{11}x_2 + \dots + a_{1n}x_n, \\ x_2'' &= a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n, \\ &\dots\dots\dots\dots\dots\dots\dots \\ x_n'' &= a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n. \end{aligned}\right\} \tag{55}$$

On rencontre souvent ces systèmes en mécanique. Si l'on cherche leur solution sous la forme:

$$x_j = b_j \cos(\lambda t + \varphi),$$

on obtient pour λ une équation de la forme:

$$\begin{vmatrix} a_{11}+\lambda^2 & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22}+\lambda^2 & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn}+\lambda^2 \end{vmatrix} = 0. \tag{56}$$

Les constantes b_j sont déterminées à partir d'un système analogue au système (48) et φ est arbitraire.

Enfin, pour les systèmes de la forme:

$$\left.\begin{aligned} x_1'' &= a_{11}x_1 + \dots + a_{1n}x_n + c_{11}x_1' + \dots + c_{1n}x_n', \\ x_2'' &= a_{21}x_1 + \dots + a_{2n}x_n + c_{21}x_1' + \dots + c_{2n}x_n', \\ &\dots\dots\dots\dots\dots\dots\dots\dots \\ x_n'' &= a_{n1}x_1 + \dots + a_{nn}x_n + c_{n1}x_1' - \dots + c_{nn}x_n' \end{aligned}\right\} \tag{57}$$

qui contiennent en outre les dérivées premières, on est conduit, en cherchant des solutions de la forme (47), à une équation séculaire de la forme:

$$\begin{vmatrix} a_{11}+c_{11}\lambda-\lambda^2 & a_{12}+c_{12}\lambda & \dots & a_{1n}+c_{1n}\lambda \\ a_{21}+c_{21}\lambda & a_{22}+c_{22}\lambda-\lambda^2 & \dots & a_{2n}+c_{2n}\lambda \\ \dots & \dots & \dots & \dots \\ a_{n1}+c_{n1}\lambda & a_{n2}+c_{n2}\lambda & \dots & a_{nn}+c_{nn}\lambda-\lambda^2 \end{vmatrix} = 0. \tag{58}$$

En introduisant les inconnues auxiliaires:

$$x_{n+1} = x_1'; \quad x_{n+2} = x_2'; \quad \dots; \quad x_{2n} = x_n', \tag{59}$$

on ramène le système (57) à $2n$ équations du premier ordre, dont n équations s'obtiennent à partir de (57) en faisant:

$$x_j'' = x_{n+j}' \quad \text{et} \quad x_j' = x_{n+j} \quad (j = 1, 2, \dots, n),$$

et les n autres sont les équations (59).

I-2-11. Déterminants fonctionnels. Soient n fonctions des n variables $x_1, x_2, \dots, x_n$:

$$\varphi_1(x_1, x_2, \dots, x_n); \quad \varphi_2(x_1, x_2, \dots, x_n); \quad \dots; \quad \varphi_n(x_1, x_2, \dots, x_n). \tag{60}$$

On appelle *déterminant fonctionnel* de ces fonctions, par rapport aux variables x_s, un déterminant d'ordre n dont les éléments sont: $a_{ik} = \frac{\partial \varphi_i}{\partial x_k}$. On utilise pour désigner le déterminant fonctionnel la notation :

$$\frac{D(\varphi_1, \ldots, \varphi_n)}{D(x_1, \ldots, x_n)} = \begin{vmatrix} \frac{\partial \varphi_1}{\partial x_1} & \frac{\partial \varphi_1}{\partial x_2} & \cdots & \frac{\partial \varphi_1}{\partial x_n} \\ \frac{\partial \varphi_2}{\partial x_1} & \frac{\partial \varphi_2}{\partial x_2} & \cdots & \frac{\partial \varphi_2}{\partial x_n} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial \varphi_n}{\partial x_1} & \frac{\partial \varphi_n}{\partial x_2} & \cdots & \frac{\partial \varphi_n}{\partial x_n} \end{vmatrix}. \quad (61)$$

On a déjà rencontré ce type de déterminant quand on a fait un changement de variables dans les intégrales multiples (tome II, [III-1-7] et [III-2-12]). Si l'on fait un changement de variables dans le plan :

$$x = \varphi(u, v); \qquad y = \psi(u, v), \quad (62)$$

le point (u, v) ayant pour homologue le point (x, y), alors la valeur absolue du déterminant fonctionnel

$$\frac{D(\varphi, \psi)}{D(u, v)} \quad (63)$$

est le coefficient de variation de l'aire au point donné (u, v) de la transformation ponctuelle (62). Nous supposons de plus que dans le domaine où l'on applique la transformation ponctuelle (62), les dérivées partielles des fonctions (62) par rapport à u et v sont continues et que le déterminant (63) n'est pas nul. De même si l'on fait une transformation ponctuelle dans un espace à trois dimensions :

$$x = \varphi(q_1, q_2, q_3); \quad y = \psi(q_1, q_2, q_3); \quad z = \omega(q_1, q_2, q_3),$$

où le point de coordonnées (q_1, q_2, q_3) a pour homologue le point (x, y, z) et où le volume (V_1) devient le volume (V), la formule de changement de variables dans une intégrale triple a la forme (tome II, [III-1-7]) :

$$\iiint_V f(x, y, z)\, dx\, dy\, dz = \iiint_{(V_1)} f[\varphi, \psi, \omega]\, |D|\, dq_1\, dq_2\, dq_3,$$

où :

$$D = \frac{D(\varphi, \psi, \omega)}{D(q_1, q_2, q_3)},$$

et $|D|$ donne le coefficient de variation du volume au voisinage de (q_1, q_2, q_3) lorsqu'on passe de ce point au point (x, y, z).

Nous aurions pu considérer de la même façon une fonction d'une variable indépendante $u = f(x)$ comme une transformation ponctuelle sur l'axe OX, où le point d'abscisse x a pour homologue le point d'abscisse u. La valeur absolue de la dérivée $|f'(x)|$ caractérise alors, évidemment, les variations de dimension linéaire en ce voisinage. Ce qu'on vient de dire reste vrai dans le cas d'une transformation ponctuelle dans un espace à n dimensions lorsqu'on fait un changement de variables dans une intégrale multiple d'ordre n (tome II, [III-4-10]).

L'analogie, dans le cas d'un espace à deux et trois dimensions, entre le déterminant fonctionnel et la dérivée a pour conséquence une certaine analogie entre leurs propriétés formelles.

Soit un système de fonctions:

$$\varphi_1(y_1, \ldots, y_n), \ldots, \varphi_n(y_1, \ldots, y_n),$$

et soient $y_1, \ldots, y_n$ non pas des variables indépendantes mais des fonctions de $x_1, \ldots, x_n$, de telle sorte que les fonctions φ_i soient des fonctions des variables x_i. Nous pouvons former trois déterminants fonctionnels:

$$\frac{D(\varphi_1, \ldots, \varphi_n)}{D(y_1, \ldots, y_n)}; \quad \frac{D(\varphi_1, \ldots, \varphi_n)}{D(x_1, \ldots, x_n)}; \quad \frac{D(y_1, \ldots, y_n)}{D(x_1, \ldots, x_n)}.$$

Les éléments de ces déterminants fonctionnels sont respectivement:

$$\frac{\partial \varphi_i}{\partial y_k}; \quad \frac{\partial \varphi_i}{\partial x_k}; \quad \frac{\partial y_i}{\partial x_k}.$$

Mais on a compte tenu de la règle de différentiation des fonctions de fonctions:

$$\frac{\partial \varphi_i}{\partial x_k} = \frac{\partial \varphi_i}{\partial y_1} \cdot \frac{\partial y_1}{\partial x_k} + \ldots + \frac{\partial \varphi_i}{\partial y_n} \cdot \frac{\partial y_n}{\partial x_k},$$

et si l'on fait le produit des déterminants, on obtient une égalité qui exprime la première propriété des déterminants fonctionnels:

$$\frac{D(\varphi_1, \ldots, \varphi_n)}{D(x_1, \ldots, x_n)} = \frac{D(\varphi_1, \ldots, \varphi_n)}{D(y_1, \ldots, y_n)} \cdot \frac{D(y_1, \ldots, y_n)}{D(x_1, \ldots, x_n)}. \tag{64}$$

Cette égalité est analogue à la règle de différentiation des fonctions de fonctions d'une variable indépendante.

On peut considérer un système de fonctions φ_i comme une transformation des variables x_i en des nouvelles variables φ_i:

$$\varphi_i = \varphi_i(x_1, \ldots, x_n) \quad (i = 1, 2, \ldots, n). \tag{65}$$

Etudions d'abord le cas particulier de la transformation identique:

$$\varphi_1 = x_1; \quad \varphi_2 = x_2; \quad \ldots; \quad \varphi_n = x_n.$$

Son déterminant fonctionnel est :

$$\begin{vmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \end{vmatrix} = 1.$$

Supposons que les équations (65) soient résolues par rapport à x_i, de telle sorte que les x_i soient exprimés en fonction des φ_i :

$$x_i = x_i(\varphi_1, \dots, \varphi_n) \quad (i = 1, 2, \dots, n). \tag{66}$$

Il est naturel d'appeler la transformation (66) inverse de la transformation (65). Si l'on porte les expressions (66) dans le second membre des égalités (65), on obtient les identités $\varphi_1 = \varphi_1; \dots; \varphi_n = \varphi_n$, autrement dit, on obtient la transformation identique. Appliquons, dans ce cas particulier, la formule (64) ; il faut alors considérer que $y_i = x_i$ et que $x_i = \varphi_i$, et l'on obtient dans le premier membre de la formule (64) le déterminant fonctionnel de la transformation identique :

$$\frac{D(\varphi_1, \dots, \varphi_n)}{D(\varphi_1, \dots, \varphi_n)} = \frac{D(\varphi_1, \dots, \varphi_n)}{D(x_1, \dots, x_n)} \frac{D(x_1, \dots, x_n)}{D(\varphi_1, \dots, \varphi_n)}$$

ou

$$\frac{D(\varphi_1, \dots, \varphi_n)}{D(x_1, \dots, x_n)} \frac{D(x_1, \dots, x_n)}{D(\varphi_1, \dots, \varphi_n)} = 1, \tag{67}$$

c'est-à-dire que *le produit des déterminants fonctionnels d'une transformation et de la transformation inverse est égal à l'unité*. Cette propriété est analogue à celle de la dérivée d'une fonction inverse dans le cas d'une variable indépendante.

Dégageons à présent le sens de la condition qui veut que le déterminant fonctionnel

$$\frac{D(\varphi_1, \varphi_2, \dots, \varphi_n)}{D(x_1, x_2, \dots, x_n)} \tag{68}$$

des fonctions $\varphi_1(x_1, \dots, x_n)$; $\varphi_2(x_1, \dots, x_n)$; ...; $\varphi_n(x_1, \dots, x_n)$ par rapport aux variables x_s soit identiquement nul. Supposons que ces fonctions soient liées par la relation fonctionnelle :

$$F(\varphi_1, \dots, \varphi_n) = 0. \tag{69}$$

Cette égalité est une identité par rapport aux variables indépendantes x_j. En différentiant par rapport à toutes les variables indépendantes, on obtient n identités :

$$\left.\begin{array}{l} \dfrac{\partial F}{\partial \varphi_1} \dfrac{\partial \varphi_1}{\partial x_1} + \dots + \dfrac{\partial F}{\partial \varphi_n} \dfrac{\partial \varphi_n}{\partial x_1} = 0, \\ \dots\dots\dots\dots\dots\dots \\ \dfrac{\partial F}{\partial \varphi_1} \dfrac{\partial \varphi_1}{\partial x_n} + \dots + \dfrac{\partial F}{\partial \varphi_n} \dfrac{\partial \varphi_n}{\partial x_n} = 0. \end{array}\right\} \tag{70}$$

On peut étudier ces identités comme des équations linéaires par rapport aux n quantités :

$$\frac{\partial F}{\partial \varphi_1}, \ldots, \frac{\partial F}{\partial \varphi_n},$$

de plus, ces quantités ne peuvent pas être simultanément identiquement nulles, car dans le cas contraire F ne contiendrait aucune des fonctions φ_i. Ainsi, le déterminant du système homogène (70) doit être nul, ce qui implique que le déterminant fonctionnel (68) soit nul. Il résulte de la relation fonctionnelle (69) que le déterminant fonctionnel (68) est identiquement nul. On peut démontrer l'inverse, c'est-à-dire : *le fait que le déterminant fonctionnel* (68) *soit identiquement nul est la condition nécessaire et suffisante pour que les fonctions* $\varphi_i(x_1, \ldots, x_n)$ *soient liées* *.

Soient, à titre d'exemple, trois fonctions de trois variables indépendantes :

$$\varphi_1 = x_1^2 + x_2^2 + x_3^2; \quad \varphi_2 = x_1 + x_2 + x_3; \quad \varphi_3 = x_1x_2 + x_1x_3 + x_2x_3. \tag{71}$$

Il est facile de vérifier qu'elles sont liées par la relation suivante :

$$\varphi_2^2 - \varphi_1 - 2\varphi_3 = 0.$$

Formons le déterminant fonctionnel du système de fonctions (71) :

$$\frac{D(\varphi_1, \varphi_2, \varphi_3)}{D(x_1, x_2, x_3)} = \begin{vmatrix} 2x_1 & 2x_2 & 2x_3 \\ 1 & 1 & 1 \\ x_2 + x_3 & x_1 + x_3 & x_1 + x_2 \end{vmatrix}.$$

Le lecteur démontrera que ce déterminant est identiquement nul.

1-2-12. Fonctions implicites. Nous avons démontré (tome I) le théorème d'existence d'une fonction implicite définie par une équation. Nous allons à présent étendre ce théorème au cas d'un système d'équations. Soit $x = x_0$, $y = y_0$ la solution de l'équation :

$$F(x, y) = 0 \tag{72}$$

et supposons que $F(x, y)$ et ses dérivées partielles du premier ordre soient continues pour $x = x_0$, $y = y_0$ et pour toutes valeurs de x, y suffisamment voisines, et que la dérivée partielle $F'_y(x, y)$ ne soit pas nulle pour $x = x_0$, $y = y_0$. L'équation (72) définit alors, pour des x suffisamment voisins de x_0, une fonction unique $y(x)$, continue, ayant une dérivée et vérifiant la condition $y(x_0) = y_0$. On peut démontrer de la même façon que l'équation :

$$F(x, y, z) = 0,$$

qui a pour solution $x = x_0$, $y = y_0$, $z = z_0$ à condition que la fonction $F(x, y, z)$ soit continue ainsi que ses dérivées partielles du premier ordre au voisinage des valeurs indiquées et que $F'_z(x_0, y_0, z_0) \neq 0$, définit une fonction unique $z(x, y)$, continue au voisinage de $x = x_0$, $y = y_0$, possédant des dérivées par rapport à x et à y et vérifiant la condition $z(x_0, y_0) = z_0$. Considérons maintenant un système de deux équations :

$$\varphi(x, y, z) = 0; \qquad \psi(x, y, z) = 0. \tag{73}$$

Supposons que ce système ait la solution $x = x_0$, $y = y_0$, $z = z_0$, *que les fonctions* $\varphi(x, y, z)$, $\psi(x, y, z)$ *et leurs dérivées partielles soient continues au voisinage*

---

* Le raisonnement concernant le système (70) est purement formel et n'est pas une démonstration rigoureuse.

des valeurs indiquées, et que le déterminant fonctionnel:

$$\frac{D(\varphi, \psi)}{D(y, z)} = \begin{vmatrix} \frac{\partial \varphi}{\partial y} & \frac{\partial \varphi}{\partial z} \\ \frac{\partial \psi}{\partial y} & \frac{\partial \psi}{\partial z} \end{vmatrix} = \frac{\partial \varphi}{\partial y}\frac{\partial \psi}{\partial z} - \frac{\partial \varphi}{\partial z}\frac{\partial \psi}{\partial y} \tag{74}$$

ne soit pas nul pour ces valeurs des variables. Le système (73) définit alors pour des x suffisamment voisins de x_0, un système unique de fonctions $y(x)$, $z(x)$ continues, ayant des dérivées du premier ordre et vérifiant la condition: $y(x_0) = y_0$, $z(x_0) = z_0$.

Puisque l'expression (74) n'est pas nulle pour $x = x_0$, $y = y_0$, $z = z_0$, on a au moins une dérivée partielle $\frac{\partial \psi}{\partial y}$ ou $\frac{\partial \psi}{\partial z}$ non nulle. Supposons, par exemple, que $\frac{\partial \psi}{\partial z}$ ne soit pas nulle pour les valeurs des variables indiquées. La seconde équation (73) définit de façon unique la fonction $z(x, y)$ d'après le théorème précédent. En portant cette valeur de la fonction dans la première équation du système, on obtient une équation où x et y sont les variables:

$$\varphi[x, y, z(x, y)] = 0. \tag{75}$$

Pour démontrer le théorème, il reste à montrer que la dérivée partielle du premier membre de l'équation (75) par rapport à y n'est pas nulle pour $x = x_0$, $y = y_0$. Cette dérivée est:

$$\left(\frac{\partial \varphi}{\partial y}\right) = \frac{\partial \varphi}{\partial y} + \frac{\partial \varphi}{\partial z} \cdot \frac{\partial z}{\partial y}, \tag{76}$$

où $\left(\frac{\partial \varphi}{\partial y}\right)$ est la dérivée totale de $\varphi(x, y, z)$ par rapport à y. La fonction $z(x, y)$ est la solution de la seconde équation (73), ce qui conduit à l'identité:

$$\psi[x, y, z(x, y)] \equiv 0.$$

En différentiant cette identité par rapport à y, on obtient:

$$\frac{\partial \psi}{\partial y} + \frac{\partial \psi}{\partial z} \cdot \frac{\partial z}{\partial y} \equiv 0. \tag{77}$$

Multiplions les deux membres de (76) par $\frac{\partial \psi}{\partial z}$ et additionnons-les à l'identité (77) après avoir multiplié ses deux membres par $-\frac{\partial \varphi}{\partial t}$. Nous obtenons ainsi:

$$\frac{\partial \psi}{\partial z} \cdot \left(\frac{\partial \varphi}{\partial y}\right) = \frac{D(\varphi, \psi)}{D(y, z)}.$$

Pour $x = x_0$, $y = y_0$, la fonction $z(x, y)$ devient z_0, alors $\frac{\partial \psi}{\partial z}$ et (74) ne sont pas nuls pour les valeurs des variables que l'on a données, et c'est pourquoi $\left(\frac{\partial \varphi}{\partial y}\right)$ n'est pas nul. Par conséquent, l'équation (75) définit une fonction unique $y(x)$. En portant cette valeur dans $z(x, y)$, nous obtenons z, fonction de x. Cette démonstration reste vraie dans le cas de plusieurs variables indépendantes.

Dans le cas général, le théorème des fonctions implicites s'énonce ainsi: *soit un système de n équations:*

$$F_1(x_1, \ldots, x_m, y_1, \ldots, y_n)=0; \quad \ldots; \quad F_n(x_1, \ldots, x_m, y_1, \ldots, y_n)=0, \tag{78}$$

qui a la solution:

$$x_k=x_k^{(0)}, \quad y_l=y_l^{(0)} \qquad \begin{pmatrix} k=1, \ldots, m \\ l=1, \ldots, n \end{pmatrix}; \tag{79}$$

soient F_l des fonctions continues ayant des dérivées partielles continues du premier ordre au voisinage des valeurs (79), *et soit enfin le déterminant fonctionnel*

$$\frac{D(F_1, \ldots, F_n)}{D(y_1, \ldots, y_n)}=\begin{vmatrix} \frac{\partial F_1}{\partial y_1} & \frac{\partial F_1}{\partial y_2} & \cdots & \frac{\partial F_1}{\partial y_n} \\ \frac{\partial F_2}{\partial y_1} & \frac{\partial F_2}{\partial y_2} & \cdots & \frac{\partial F_2}{\partial y_n} \\ \cdot & \cdot & \cdot & \cdot \\ \frac{\partial F_n}{\partial y_1} & \frac{\partial F_n}{\partial y_2} & \cdots & \frac{\partial F_n}{\partial y_n} \end{vmatrix} \tag{80}$$

non nul pour les valeurs (79). *Les équations* (78) *définissent alors pour des x_k suffisamment voisins de $x_k^{(0)}$ un système unique de fonctions $y_l(x_1, \ldots, x_n)$ continues, ayant des dérivées du premier ordre et vérifiant la condition $y_l(x_1^{(0)}, \ldots, x_k^{(0)}) = y_l^{(0)}$.*

Supposons que ce théorème soit vrai pour un système de $(n-1)$ équations (il est effectivement vrai pour $n=1$ et $n=2$). Démontrons qu'il est vrai aussi pour un système de n équations. En développant le déterminant (80) par rapport aux éléments de la première colonne, on peut affirmer qu'au moins un des complémentaires algébriques correspondant à ces éléments n'est pas nul pour les valeurs (79), puisque, par hypothèse, le déterminant lui-même n'est pas nul pour ces valeurs. En numérotant les fonctions F_l, on peut supposer que le complémentaire algébrique de l'élément $\frac{\partial F_1}{\partial y_1}$ n'est pas nul. Ce complémentaire algébrique est le déterminant fonctionnel des fonctions $F_2, \ldots, F_n$, par rapport aux variables $y_2, \ldots, y_n$. D'après le théorème appliquable aux systèmes de $(n-1)$ équations, les équations

$$F_2(x_1, \ldots, x_m, y_1, \ldots, y_n)=0; \quad \ldots; \quad F_n(x_1, \ldots, x_m, y_1, \ldots, y_n)=0 \tag{81}$$

définissent de façon unique les fonctions:

$$y_2=\varphi_2(x_1, \ldots, x_m, y_1); \quad \ldots; \quad y_n=\varphi_n(x_1, \ldots, x_m, y_1). \tag{82}$$

En portant ces fonctions dans la première équation du système (78), nous obtenons une équation qui définit y_1:

$$F_1(x_1, \ldots, x_m, y_1, \varphi_2, \ldots, \varphi_n)=0. \tag{83}$$

Il nous reste à vérifier que la dérivée totale du premier membre de cette équation par rapport à y_1 n'est pas nulle pour les valeurs (79). Cette dérivée est donnée par la formule:

$$\left(\frac{\partial F_1}{\partial y_1}\right)=\frac{\partial F_1}{\partial y_1}+\sum_{s=2}^{n}\frac{\partial F_1}{\partial \varphi_s}\cdot\frac{\partial \varphi_s}{\partial y_1}. \tag{84}$$

En portant les fonctions (82) dans le premier membre des équations (81), nous obtenons des identités que l'on différentie par rapport à y_1:

$$\frac{\partial F_l}{\partial y_1}+\sum_{s=2}^{n}\frac{\partial F_l}{\partial \varphi_s}\cdot\frac{\partial \varphi_s}{\partial y_1}=0 \qquad (l=2, \ldots, n). \tag{85}$$

Désignons par $A_1, A_2, \ldots, A_n$ les complémentaires algébriques correspondant aux éléments de la première colonne du déterminant (80). Multiplions (84) par A_1, (85) par A_l et retranchons les dernières identités de la première. Nous obtenons ainsi l'égalité:

$$A_1\left(\frac{\partial F_1}{\partial y_1}\right)=\sum_{l=1}^{n}\frac{\partial F_l}{\partial y_1}A_l+\sum_{s=2}^{n}\left[\sum_{l=1}^{n}\frac{\partial F_l}{\partial \varphi_s}A_l\right]\frac{\partial \varphi_s}{\partial y_1}.$$

La première somme du second membre donne le déterminant (80) qu'on désigne, pour simplifier, par la lettre D. La sommation sur l dans le second membre donne la somme de produits des éléments d'une colonne de D avec un indice différent de l'unité par les complémentaires algébriques des éléments correspondants de la première colonne, c'est-à-dire que cette somme est nulle. Il faut remarquer que la différentiation par rapport à φ_s revient à différentier par rapport à y_s. On peut donc écrire la formule précédente sous la forme:

$$A_1\left(\frac{\partial F_1}{\partial y_1}\right)=D.$$

Or, A_1 et D sont non nuls pour les valeurs (79), et, par conséquent, il en est de même en ce qui concerne la dérivée du premier membre de l'équation (83) par rapport à y_1 et cette équation donne la fonction unique $y_1(x_1, x_2, \ldots, x_m)$; en portant cette fonction dans la fonction (82), nous obtenons le résultat final.

Un cas particulier du théorème des fonctions implicites est *le théorème d'inversion des systèmes de fonctions. Soient données les équations*:

$$y_k=f_k(x_1, \ldots, x_n) \qquad (k=1, 2, \ldots, n). \tag{86}$$

Supposons que les fonctions f_k soient continues ainsi que leurs dérivées du premier ordre au voisinage des valeurs $x_k = x_k^{(0)}$ $(k = 1, 2, \ldots, n)$ et que pour ces valeurs le déterminant fonctionnel

$$\frac{D(f_1, \ldots, f_n)}{D(x_1, \ldots, x_n)} \tag{87}$$

ne soit pas nul. Les équations (86) définissent alors de façon unique $x_k(y_1, \ldots, y_n)$ comme des fonctions des $y_1, \ldots, y_n$ au voisinage des valeurs $y_k^{(0)} = f_k(x_1^{(0)}, \ldots, x_n^{(0)})$, en outre, ces fonctions sont continues, ont des dérivées du premier ordre et $x_k(y_1^{(0)}, \ldots, y_n^{(0)}) = x_k^{(0)}$.

Il suffit, pour démontrer ce théorème, de considérer les équations

$$f_k(x_1, \ldots, x_n)-y_k=0 \qquad (s=1, 2, \ldots, n)$$

et d'utiliser le théorème des fonctions implicites, x_k jouant ici le rôle de y_l.

Si les f_k sont des fonctions linéaires homogènes des variables x_k, le système (86) est de la forme:

$$y_k=a_{k1}x_1+a_{k2}x_2+\ldots+a_{kn}x_n.$$

Le déterminant (87) se réduit dans ce cas au déterminant $|a_{kl}|$ des coefficients a_{kl} et le théorème de Cramer permet d'obtenir la solution unique du système.

Chapitre II

TRANSFORMATIONS LINÉAIRES ET FORMES QUADRATIQUES

II-1. Transformations linéaires

II-1-1. Transformation des coordonnées dans un espace à trois dimensions. On appelle transformation linéaire de n variables une transformation de la forme:

$$\left.\begin{aligned} x'_1 &= a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n, \\ x'_2 &= a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n, \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ x'_n &= a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n. \end{aligned}\right\} \qquad (1)$$

On peut interpréter cette transformation comme le passage d'un vecteur $(x_1, \ldots, x_n)$ dans un espace à n dimensions à un autre vecteur $(x'_1, \ldots, x'_n)$. On peut encore considérer $(x_1, \ldots, x_n)$ comme les coordonnées d'un point dans un espace à n dimensions et la transformation (1) comme le passage d'un point à un autre.

On peut interpréter (1) encore d'une autre façon: $(x_1, \ldots, x_n)$ et $(x'_1, \ldots, x'_n)$ sont les composantes d'un seul et même vecteur (coordonnées d'un seul et même point) mais pour un choix différent des axes de coordonnées. Les formules (1) donnent alors les formules de transformation des composantes (des coordonnées) lors du passage d'un système de coordonnées à un autre. On a déjà rencontré plus d'une fois des formules de la forme (1) pour $n = 2$ et $n = 3$.

La première partie de ce chapitre sera consacrée à l'étude détaillée de transformations linéaires du type (1). On étudie d'abord le cas d'un espace réel à trois dimensions, puis on passe au cas général d'un espace à n dimensions et de composantes complexes. Commençons l'étude de l'espace à trois dimensions dans le cas très simple où à la transformation (1) correspond le passage d'axes orthogonaux à d'autres axes orthogonaux. En repérant les vecteurs par rapport à l'origine des coordonnées, nous pouvons considérer, évidemment, (x_1, x_2, x_3) soit comme les composantes d'un vecteur, soit comme les coordonnées de son extrémité.

La géométrie analytique nous indique que les formules de transformation des coordonnées orthogonales rectilignes sont de la forme:

$$\left.\begin{aligned} x_1' &= a_{11}x_1 + a_{12}x_2 + a_{13}x_3, \\ x_2' &= a_{21}x_1 + a_{22}x_2 + a_{23}x_3, \\ x_3' &= a_{31}x_1 + a_{32}x_2 + a_{33}x_3, \end{aligned}\right\} \tag{2}$$

où les a_{ik} sont les cosinus des angles formés par les nouveaux axes avec les anciens; les a_{ik} sont définis par le tableau suivant:

	X_1	X_2	X_3
X_1'	a_{11}	a_{12}	a_{13}
X_2'	a_{21}	a_{22}	a_{23}
X_3'	a_{31}	a_{32}	a_{33}

(3)

Comme on le sait, le tableau des coefficients (3) possède, dans ce cas, la propriété suivante: la somme des carrés des éléments de chaque ligne et de chaque colonne est égale à l'unité, et la somme des produits des éléments correspondants de deux lignes ou de deux colonnes différentes est nulle. La valeur du déterminant $|a_{ik}|$ est évidemment égale [I-1-5] au volume d'un parallélépipède rectangle dont les arêtes sont égales à l'unité et dirigées suivant les nouveaux axes de coordonnées, c'est-à-dire que cette valeur est égale à l'unité si l'orientation des axes est la même, et à (-1) dans le cas contraire. La transformation donnant le passage inverse de (x_1', x_2', x_3') à (x_1, x_2, x_3) est évidemment de la forme:

$$\left.\begin{aligned} x_1 &= a_{11}x_1' + a_{21}x_2' + a_{31}x_3', \\ x_2 &= a_{12}x_1' + a_{22}x_2' + a_{32}x_3', \\ x_3 &= a_{13}x_1' + a_{23}x_2' + a_{33}x_3'. \end{aligned}\right\} \tag{2$_1$}$$

Autrement dit, on obtient la transformation inverse de (2) en échangeant simplement les lignes et les colonnes du tableau des coefficients. Le déterminant de cette transformation inverse est évidemment égal au déterminant de la transformation (2).

Montrons maintenant que la propriété énoncée ci-dessus pour les coefficients de la transformation (2) peut être obtenue à partir d'une condition qui résulte directement de la nature géométrique de la question: déterminons toutes les transformations réelles de la forme (2) telles que:

$$x_1'^2 + x_2'^2 + x_3'^2 = x_1^2 + x_2^2 + x_3^2. \tag{4}$$

Ce problème est posé de façon que l'on puisse généraliser les transformations précédentes au cas d'un espace ayant un nombre quelconque de dimensions. Montrons que les transformations qu'on veut obtenir coïncident avec celles étudiées ci-dessus, c'est-à-dire que (4) conduit aux relations données ci-dessus pour les coefficients a_{ik}. En portant les expressions (2) dans le premier membre de la relation (4), en supprimant les parenthèses, en égalant à l'unité les coefficients des variables élevées au carré et en annulant les coefficients des termes rectangles, on obtient six relations de la forme :

$$a_{1k}a_{1l}+a_{2k}a_{2l}+a_{3k}a_{3l}=\delta_{kl} \quad (k, l=1, 2, 3), \tag{5}$$

où :

$$\delta_{kl}=0 \quad \text{pour } k\neq l \text{ et } \delta_{kk}=1, \tag{6}$$

c'est-à-dire que la somme des carrés des coefficients de chaque colonne est égale à l'unité et que la somme des produits des coefficients correspondants des différentes colonnes est nulle. C'est la *condition d'orthogonalité suivant les colonnes*. Ainsi, les coefficients de chaque colonne sont égaux aux cosinus directeurs d'une direction, et les directions correspondant à différentes colonnes sont orthogonales. Il en résulte que la transformation (2) coïncide dans ce cas avec celle étudiée ci-dessus, et la propriété d'orthogonalité est vraie non seulement pour les colonnes mais aussi pour les lignes.

On peut interpréter les formules (2) non plus comme la transformation des coordonnées dans un espace immobile, mais comme la transformation d'un espace lorsque les axes de coordonnées ne varient pas. On suppose d'abord que le déterminant de la transformation est égal à $(+1)$, c'est-à-dire que les deux systèmes d'axes de coordonnées ont même orientation. On peut faire tourner l'espace comme un solide autour de l'origine en même temps que les axes (X_1', X_2', X_3') de telle sorte que ces axes coïncident avec les axes (X_1, X_2, X_3) que nous supposons immobiles lors de cette rotation et auxquels on rapporte les coordonnées de tout point avant et après la rotation. Si un point M a les coordonnées (x_1, x_2, x_3) avant la rotation, il occupe, après la rotation, la position M' et il a les nouvelles coordonnées (x_1', x_2', x_3'). Etant donné que le point M se déplace en même temps que les axes de coordonnées (X_1', X_2', X_3'), les coordonnées (x_1', x_2', x_3') du point M' par rapport aux axes (X_1, X_2, X_3) avec lesquels coïncident les axes (X_1', X_2', X_3') après la rotation coïncident avec les coordonnées du point M par rapport aux axes (X_1', X_2', X_3') avant la rotation. On voit, donc, que les formules (2) représentent dans le cas où le déterminant est égal à $(+1)$ la transformation des coordonnées d'un point quelconque après la rotation de l'espace.

Supposons maintenant que le déterminant $|a_{ik}|$ soit égal à (-1). Etudions à la place de la transformation (2) la tranformation :

$$x_i''=-a_{i1}x_1-a_{i2}x_2-a_{i3}x_3 \qquad (i=1, 2, 3).$$

Ses coefficients ont également les propriétés (5), mais le déterminant des coefficients est ici égal à $(+1)$, c'est-à-dire qu'à cette transformation correspond une rotation de l'espace autour de l'origine. Pour obtenir les coordonnées (x_1', x_2', x_3'), on doit effectuer encore une transformation:

$$x_1' = -x_1''; \quad x_2' = -x_2''; \quad x_3' = -x_3'';$$

un tel changement de signe de toutes les coordonnées représente une symétrie par rapport à l'origine. Ainsi, lorsque le déterminant est égal à (-1), la transformation (2) correspond à une rotation de l'espace autour de l'origine suivie d'une symétrie par rapport à l'origine.

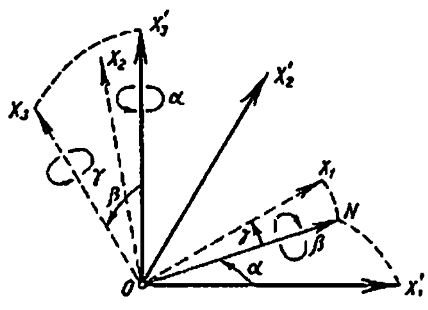


Fig. 1.

Nous avons vu ci-dessus que les neuf coefficients a_{ik} doivent vérifier les six relations (5). Il en résulte qu'ils dépendent de trois paramètres indépendants. L'un des choix possibles de ces paramètres dans le cas d'une rotation de l'espace autour de l'origine est le suivant.

Soient deux systèmes d'axes de coordonnées: l'un X_1', X_2', X_3', fixe et auquel on rapporte toutes les coordonnées, et l'autre X_1, X_2, X_3, lié invariablement à l'espace en rotation. Pour définir une rotation, il faut trouver trois paramètres qui définissent la position du second système d'axes par rapport au premier. Soit ON (fig. 1) l'intersection du plan $X_1'OX_2'$ avec le plan X_1OX_2. On prend sur cette droite une direction déterminée, et soit α l'angle $X_1'ON$ compté à partir de OX_1'. Soient encore les angles $\beta = X_3'OX_3$ et $\gamma = NOX_1$. Ces trois angles caractérisent entièrement la position du second système d'axes par rapport au premier, c'est-à-dire qu'ils caractérisent la rotation considérée. On la désigne par $\{\alpha, \beta, \gamma\}$. Il résulte de ce qui précède que le mouvement effectué est le résultat de trois mouvements successifs: 1) d'une rotation autour de l'axe X_3' d'un angle α; 2) d'une rotation autour de l'axe X_1' d'un angle β; 3) d'une rotation autour du nouvel axe X_3 d'un angle γ. Ces trois angles sont les angles d'Euler. Les limites de variation de ces angles sont: $0 \leqslant \alpha < 2\pi$; $0 \leqslant \beta < \pi$; $0 \leqslant \gamma < 2\pi$.

Si $\beta = 0$, le mouvement $\{\alpha, \beta, \gamma\}$ se ramène à une simple rotation autour de l'axe X_3 d'un angle $\alpha + \gamma$, et dans ce sens, on a pour δ quelconque:

$$\{\alpha, 0, \gamma\} = \{\alpha + \delta, 0, \gamma - \delta\}. \tag{7}$$

Ce qui veut dire que dans certains cas les paramètres $\{\alpha, \beta, \gamma\}$ ne correspondent pas univoquement à une rotation. A des valeurs différentes des paramètres correspondent les mêmes rotations. Il est facile d'obtenir les formules qui expriment les coefficients a_{ik} par les fonctions trigonométriques des angles α, β, γ (cf. [III-1-11]). On utilisera par la suite d'autres paramètres caractérisant la rotation de l'espace autour de l'origine et on reviendra sur les angles d'Euler.

II-1-2. Transformations linéaires générales d'un espace réel à trois dimensions. Etudions maintenant les transformations linéaires réelles du type (2). à coefficients quelconques. Supposons toujours que le déterminant de la transformation ne soit pas nul:

$$|a_{ik}| \neq 0. \tag{7}$$

Dans ce cas, on dit généralement que la transformation est une *transformation propre*. Si elle ne vérifie pas les conditions (5), elle est liée à une déformation de l'espace (tome II, [IV-2-7]). On notera que la transformation (2) est caractérisée par le tableau de ses coefficients à partir duquel on peut définir la loi de transformation d'un vecteur quelconque ayant les composantes (x_1, x_2, x_3) en un nouveau vecteur de composantes (x_1', x_2', x_3'). On désigne le tableau des coefficients par une lettre:

$$A = \begin{Vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{Vmatrix}, \tag{8}$$

et, comme précédemment, on met ce tableau entre les doubles traits verticaux pour le distinguer du déterminant qui lui est associé. On l'appelle *matrice*. On désigne le déterminant du tableau (8) par le symbole $D(A)$ qui représente un nombre bien déterminé. On écrit symboliquement la transformation (2) sous la forme:

$$\mathbf{x}' = A\mathbf{x}, \tag{9}$$

où $\mathbf{x}'$ est un vecteur de composantes (x_1', x_2', x_3') et $\mathbf{x}$ un vecteur de composantes (x_1, x_2, x_3).

La transformation qui laisse chaque vecteur invariable est la *transformation identique*. A cette transformation correspond le tableau suivant:

$$\begin{Vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{Vmatrix}, \tag{10}$$

qui est la *matrice unité* et qu'on désigne par le symbole I.

En considérant $D(A) \neq 0$, on peut résoudre les équations (2) par rapport à (x_1, x_2, x_3) et l'on obtient les formules:

$$\left.\begin{aligned} x_1 &= \frac{A_{11}}{D(A)} x_1' + \frac{A_{21}}{D(A)} x_2' + \frac{A_{31}}{D(A)} x_3', \\ x_2 &= \frac{A_{12}}{D(A)} x_1' + \frac{A_{22}}{D(A)} x_2' + \frac{A_{32}}{D(A)} x_3', \\ x_3 &= \frac{A_{13}}{D(A)} x_1' + \frac{A_{23}}{D(A)} x_2' + \frac{A_{33}}{D(A)} x_3', \end{aligned}\right\} \tag{11}$$

où les A_{ik} sont les complémentaires algébriques des éléments a_{ik} dans le déterminant $D(A)$. On dit que cette transformation linéaire est la transformation inverse de (2) et si la matrice de cette dernière est A, on désigne la matrice de la transformation (11) par A^{-1}. Introduisons maintenant une notion importante pour la suite: la notion de produit de deux transformations ou de produit de deux matrices. Soient deux transformations linéaires de (x_1, x_2, x_3) en (x_1', x_2', x_3'):

$$\left.\begin{aligned} x_1' &= a_{11}x_1 + a_{12}x_2 + a_{13}x_3, \\ x_2' &= a_{21}x_1 + a_{22}x_2 + a_{23}x_3, \\ x_3' &= a_{31}x_1 + a_{32}x_2 + a_{33}x_3 \end{aligned}\right\} \text{ ou } \mathbf{x}' = A\mathbf{x}, \tag{12}$$

puis de (x_1', x_2', x_3') en (x_1'', x_2'', x_3''):

$$\left.\begin{aligned} x_1'' &= b_{11}x_1' + b_{12}x_2' + b_{13}x_3', \\ x_2'' &= b_{21}x_1' + b_{22}x_2' + b_{23}x_3', \\ x_3'' &= b_{31}x_1' + b_{32}x_2' + b_{33}x_3' \end{aligned}\right\} \text{ ou } \mathbf{x}'' = B\mathbf{x}'. \tag{13}$$

On peut remplacer ce passage successif de (x_1, x_2, x_3) à (x_1', x_2', x_3') puis de (x_1', x_2', x_3') à (x_1'', x_2'', x_3'') par le passage direct de (x_1, x_2, x_3) à (x_1'', x_2'', x_3'') qui sera aussi une transformation linéaire:

$$x_k'' = c_{k1}x_1 + c_{k2}x_2 + c_{k3}x_3 \quad (k = 1, 2, 3). \tag{14}$$

On appelle cette dernière transformation linéaire produit des transformations (12) et (13), et il est indispensable de noter l'ordre dans lequel les transformations ont été effectuées. On obtient les formules (14) en substituant les expressions (12) dans le second membre de la formule (13). On obtient directement l'expression des éléments c_{ik} du tableau du produit des transformations à partir des éléments des tableaux des transformations que l'on multiplie:

$$c_{ik} = \sum_{s=1}^{3} b_{is}a_{sk} \qquad (i, k = 1, 2, 3). \tag{15}$$

On écrit habituellement la transformation (14) de la façon suivante:

$$\mathbf{x}'' = BA\mathbf{x}. \tag{16}$$

La matrice C, dont les éléments sont c_{ik} obtenus d'après les formules (15), est le produit des matrices A et B; on écrit cela sous la forme suivante :

$$C = BA, \tag{17}$$

et il faut lire de droite à gauche l'ordre dans lequel ont été effectuées les transformations. En tenant compte du théorème de la multiplication des déterminants et des formules (15), on peut écrire l'égalité évidente pour les déterminants associés aux transformations :

$$D(C) = D(B)\, D(A), \tag{18}$$

c'est-à-dire que *le déterminant d'un produit de transformations est égal au produit des déterminants de ces transformations.* Il est facile de démontrer les relations suivantes qui ont un sens géométrique simple :

$$AA^{-1} = A^{-1}A = I. \tag{19}$$

De plus, il faut remarquer qu'il résulte du procédé même par lequel a été formée la transformation inverse que la transformation inverse de A^{-1} est la transformation A. En effet, on obtient à nouveau les formules (2), en résolvant le système (11) par rapport à x'_k. On peut écrire cela de la façon suivante :

$$(A^{-1})^{-1} = A. \tag{20}$$

La notion de produit de transformations peut être étendue au cas d'un nombre quelconque de facteurs. Le résultat des transformations successives de matrices A, B et C sera une nouvelle transformation de matrice D :

$$D = CBA. \tag{21}$$

Si les matrices A, B et C ont des éléments a_{ik}, b_{ik} et c_{ik}, la matrice D, qui est leur produit, aura des éléments donnés par les formules :

$$d_{ik} = \sum_{p,\, q=1}^{3} c_{iq} b_{qp} a_{pk}. \tag{22}$$

En effet, pour les éléments de la matrice $E = BA$, on a les formules :

$$e_{ik} = \sum_{p=1}^{3} b_{ip} a_{pk},$$

et enfin, pour la matrice CE, on aura, d'après (15), les formules :

$$d_{ik} = \sum_{q=1}^{3} c_{iq} e_{qk},$$

c'est-à-dire précisément les formules (22). Par la suite, on désigne souvent les éléments de la matrice A par le symbole:

$$\{A\}_{ik}.$$

En général, le produit de matrices n'est pas commutatif, c'est-à-dire que $BA \neq AB$, mais il est facile de voir qu'il est associatif; on peut regrouper les facteurs:

$$C(BA) = (CB)A. \tag{23}$$

On doit multiplier à gauche la matrice A du premier membre par B, puis multiplier le résultat obtenu par la matrice C. Dans le second membre, on doit d'abord multiplier la matrice B par C, puis multiplier la matrice A par la matrice que l'on a obtenue par le produit précédent. Il est facile de voir que dans les deux cas, les éléments de matrice obtenus dans le produit final s'expriment au moyen des formules (22). Cela a déjà été démontré pour le premier membre. Quant au second membre, on a, en effectuant successivement les multiplications indiquées:

$$\{CB\}_{ik} = \sum_{q=1}^{3} \{C\}_{iq}\{B\}_{qk}$$

et

$$\{(CB)A\}_{ik} = \sum_{p=1}^{3} \{CB\}_{ip}\{A\}_{pk} = \sum_{p,\,q=1}^{3} \{C\}_{iq}\{B\}_{qp}\{A\}_{pk},$$

ce qui coïncide, évidemment, avec (22), compte tenu des notations choisies.

Il faut noter encore un type important de transformations linéaires:

$$x_1' = k_1x_1; \qquad x_2' = k_2x_2; \qquad x_3' = k_3x_3 \tag{24}$$

qui se ramène à une dilatation le long des axes de coordonnées, les coefficients numériques k_1, k_2, k_3 caractérisant la valeur de cette dilatation (ou contraction). La matrice de la transformation qu'on a indiquée est, évidemment, de la forme:

$$\begin{Vmatrix} k_1 & 0 & 0 \\ 0 & k_2 & 0 \\ 0 & 0 & k_3 \end{Vmatrix},$$

c'est-à-dire que tous les éléments situés en dehors de la diagonale principale sont nuls. Cette matrice est une *matrice diagonale* et on la désigne par le symbole:

$$[k_1,\ k_2,\ k_3].$$

En particulier, si les nombres sont identiques, c'est-à-dire si $k_1 = k_2 = k_3 = k$, la transformation se ramène à la multiplication de toutes les composantes du vecteur par un seul et même nombre k, et ce sera, évidemment, une homothétie ayant pour centre l'origine des coordonnées. Tout vecteur qui ne change pas de sens (on considère que $k > 0$) ne varie que par sa longueur, celle-ci étant multipliée par k. On désigne cette transformation de la façon suivante :

$$\mathbf{x}' = k\mathbf{x},$$

c'est-à-dire qu'on considère le nombre k comme un cas particulier de matrice : la matrice diagonale ayant des éléments égaux sur la diagonale principale au nombre k :

$$\begin{Vmatrix} k & 0 & 0 \\ 0 & k & 0 \\ 0 & 0 & k \end{Vmatrix}. \tag{25}$$

Il est facile de voir qu'en utilisant (15) le produit de ces matrices se ramène à une multiplication ordinaire de nombres, c'est-à-dire :

$$[k, k, k] \cdot [l, l, l] = [kl, kl, kl].$$

Il est en général aisé de vérifier que, pour les matrices diagonales, on a une loi de multiplication simple :

$$[k_1, k_2, k_3]\,[l_1, l_2, l_3] = [k_1l_1, k_2l_2, k_3l_3], \tag{26}$$

c'est-à-dire que deux dilatations successives le long des axes de coordonnées sont équivalentes à une dilatation unique dont les coefficients sont égaux au produit des coefficients correspondants des dilatations composantes. Il resulte directement de la formule (26) que le produit de deux matrices diagonales ne change pas lorsqu'on intervertit ses facteurs. En représentant un nombre quelconque sous forme de matrice diagonale (25) et en utilisant la formule (15), on voit facilement que le produit kA se ramène à la multiplication de tous les éléments de la matrice A par le nombre k. Ce produit ne dépend pas de l'ordre des facteurs, c'est-à-dire que :

$$\{kA\}_{ik} = \{Ak\}_{ik} = k\,\{A\}_{ik}. \tag{27}$$

On a considéré la transformation linéaire fondamentale (2) comme une déformation de l'espace, le vecteur de composantes (x_1, x_2, x_3) devenant un nouveau vecteur de composantes (x'_1, x'_2, x'_3). On peut bien entendu, comme ci-dessus, interpréter cette transformation comme une transformation ponctuelle où le point de coordonnées (x_1, x_2, x_3) devient le point de coordonnées (x'_1, x'_2, x'_3).

On peut utiliser n'importe quel système d'axes pour définir les composantes d'un vecteur, c'est-à-dire qu'on peut prendre n'importe quel ensemble de trois vecteurs non coplanaires **i**, **j**, **k** comme vecteurs de base. On a alors pour tout vecteur **x** une représentation unique de la forme (tome II, [IV-1-3]):

$$\mathbf{x} = x_1\mathbf{i} + x_2\mathbf{j} + x_3\mathbf{k}. \tag{28}$$

Les nombres x_1, x_2, x_3 sont appelés composantes du vecteur **x** dans le système d'axes choisi, défini par les vecteurs **i**, **j** et **k**. Il faut étudier l'influence du choix des vecteurs de base sur la forme de la transformation linéaire.

Si, dans un système de coordonnées défini par les vecteurs de base **i**, **j** et **k**, une transformation linéaire est de la forme (12), quelle est la forme de cette même transformation linéaire par rapport à un autre système de coordonnées défini par les vecteurs $\mathbf{i}_1$, $\mathbf{j}_1$, $\mathbf{k}_1$? Supposons que les nouveaux vecteurs de base soient exprimés en fonction des anciens par les formules:

$$\left.\begin{aligned} \mathbf{i}_1 &= t_{11}\mathbf{i} + t_{12}\mathbf{j} + t_{13}\mathbf{k}, \\ \mathbf{j}_1 &= t_{21}\mathbf{i} + t_{22}\mathbf{j} + t_{23}\mathbf{k}, \\ \mathbf{k}_1 &= t_{31}\mathbf{i} + t_{32}\mathbf{j} + t_{33}\mathbf{k}. \end{aligned}\right\} \tag{29}$$

Notons que le déterminant formé des coefficients t_{ik} ne doit pas être nul. Dans le cas contraire les vecteurs $\mathbf{i}_1$, $\mathbf{j}_1$ et $\mathbf{k}_1$ seraient linéairement dépendants, donc coplanaires. Dans le nouveau système de coordonnées, le vecteur défini par la formule (28) aura de nouvelles composantes:

$$y_1\mathbf{i}_1 + y_2\mathbf{j}_1 + y_3\mathbf{k}_1.$$

Etablissons tout d'abord une formule exprimant les nouvelles composantes du vecteur en fonction des anciennes. En utilisant les expressions (29) des nouveaux vecteurs de base, on a évidemment:

$$\sum_{s=1}^{3} y_s (t_{s1}\mathbf{i} + t_{s2}\mathbf{j} + t_{s3}\mathbf{k}) = x_1\mathbf{i} + x_2\mathbf{j} + x_3\mathbf{k}.$$

En comparant les coefficients de **i**, **j** et **k**, on obtient des formules exprimant les anciennes composantes en fonction des nouvelles:

$$\left.\begin{aligned} x_1 &= t_{11}y_1 + t_{21}y_2 + t_{31}y_3, \\ x_2 &= t_{12}y_1 + t_{22}y_2 + t_{32}y_3, \\ x_3 &= t_{13}y_1 + t_{23}y_2 + t_{33}y_3. \end{aligned}\right\} \tag{30}$$

Le tableau de cette transformation diffère de celui de la transformation (29) par le fait que les lignes sont remplacées par les colonnes. En effet, dans chaque ligne du tableau (29) le premier indice ne varie pas, et dans le tableau (30) c'est le second indice. En désignant le tableau de la transformation (29) par T, le tableau de la transformation (30) sera représenté par $T^{(*)}$ appelé transposée de T. On peut alors écrire les formules (30) sous la forme abrégée suivante:

$$(x_1, x_2, x_3) = T^{(*)}(y_1, y_2, y_3), \tag{31}$$

où (x_1, x_2, x_3) désigne les trois composantes du vecteur par rapport aux anciens vecteurs de base et (y_1, y_2, y_3) celles par rapport aux nouveaux vecteurs de base. L'expression des nouvelles composantes en fon tion des anciennes est de la forme:

$$(y_1, y_2, y_3) = T^{(*)-1}(x_1, x_2, x_3),$$

où $T^{(*)-1}$ est la transformation linéaire inverse de $T^{(*)}$. D'ordinaire on l'appelle *contragrédiente* de T. On désigne pour simplifier le tableau qui lui correspond par la lettre:

$$U = T^{(*)-1}. \tag{32}$$

Donc, lorsqu'on change les vecteurs de base d'après la formule (29), les composantes de tout vecteur subissent une transformation linéaire de tableau U associé et défini par la formule (32). Ainsi, les deux vecteurs $\mathbf{x}$ (x_1, x_2, x_3) et $\mathbf{x}'$ (x_1', x_2', x_3') qui apparaissent dans la transformation (9) auront après transformation des vecteurs de base d'autres composantes définies en fonction des anciennes par les formules:

$$(y_1, y_2, y_3) = U(x_1, x_2, x_3); \; (y_1', y_2', y_3') = U(x_1', x_2', x_3'). \tag{33}$$

Il nous faut donc déterminer une relation linéaire entre les composantes (y_1, y_2, y_3) et (y_1', y_2', y_3'). On peut effectuer le passage du vecteur (y_1, y_2, y_3) au nouveau vecteur (y_1', y_2', y_3') de la façon suivante: on passe d'abord du vecteur (y_1, y_2, y_3) au vecteur (x_1, x_2, x_3), ce qui, d'après (33), se fait à l'aide du tableau U^{-1}, ensuite du vecteur (x_1, x_2, x_3) au vecteur (x_1', x_2', x_3') à l'aide du tableau A de la transformation (9) et, enfin, du vecteur (x_1', x_2', x_3') au vecteur (y_1', y_2', y_3') à l'aide du tableau de transformation U. On aura finalement une transformation linéaire de la forme:

$$\mathbf{y}' = UAU^{-1}\mathbf{y}. \tag{34}$$

On dit que cette transformation est semblable à la transformation (9) et que sa matrice UAU^{-1} est semblable à la matrice A.

Énonçons le résultat définitif.

Si les formules (33) *représentent une transformation linéaire des composantes du vecteur due au changement des vecteurs de base, toute transformation linéaire de la forme*

$$\mathbf{x}' = A\mathbf{x}$$

devient, dans le nouveau système de coordonnées:

$$y' = UAU^{-1}y.$$

II-1-3. **Vecteurs affines covariants et contravariants.** Supposons que la transformation linéaire (9) exprime simplement le passage d'un système de coordonnées cartésiennes à un autre, c'est-à-dire que ses coefficients soient les cosinus directeurs définis par le tableau (3). Dans ce cas, comme on l'a vu [II-1-1], le tableau transposé $A^{(*)}$ coïncide avec le tableau inverse A^{-1} et, par conséquent, le tableau contragrédient $A^{(*)-1}$ coïncide avec le tableau initial A, c'est-à-dire:

$$A^{(*)} = A^{-1}; \qquad A^{(*)-1} = A. \tag{35}$$

Si l'on considère un vecteur qui ne varie ni en longueur ni en direction, on peut affirmer que ses composantes se transforment suivant les mêmes formules (9) que les coordonnées, c'est-à-dire:

$$\left.\begin{aligned} x_1' &= a_{11}x_1 + a_{12}x_2 + a_{13}x_3, \\ x_2' &= a_{21}x_1 + a_{22}x_2 + a_{23}x_3, \\ x_3' &= a_{31}x_1 + a_{32}x_2 + a_{33}x_3. \end{aligned}\right\} \tag{36}$$

Ainsi, un vecteur est caractérisé par trois nombres dans tout système donné de coordonnées cartésiennes. Lorsqu'on passe d'un système cartésien à un autre, ces trois nombres (composantes du vecteur) se transforment suivant les mêmes formules (36) que les coordonnées. Supposons maintenant qu'on tienne compte non seulement du passage d'un système cartésien à un autre, mais aussi de toutes les transformations linéaires possibles des coordonnées à un déterminant non nul, ce qui correspond, comme on l'a vu, à un choix arbitraire de trois vecteurs non coplanaires comme vecteurs de base. Etudions, en même temps que le tableau A de la transformation (36), le tableau contragrédient $V = A^{(*)-1}$. Ils seront dans le cas général différents, et l'on peut ainsi avoir deux définitions du vecteur lors d'une transformation linéaire quelconque des coordonnées. On peut d'abord déterminer un vecteur comme un ensemble de trois nombres qui se transforment, lorsqu'on passe d'un système de coordonnées à un autre, selon les mêmes formules que les coordonnées elles-mêmes, c'est-à-dire d'après les formules:

$$(x_1', x_2', x_3') = A(x_1, x_2, x_3). \tag{37}$$

C'est un *vecteur affine contravariant*, et la transformation linéaire générale (36) est parfois appelée transformation affine. On peut aussi définir un vecteur de façon que ses composantes, pour toute transformation linéaire (36), subissent la transformation contragrédiente:

$$(x_1', x_2', x_3') = V(x_1, x_2, x_3). \tag{38}$$

C'est un *vecteur affine covariant.*

Dans les deux cas, lorsqu'on connaît les composantes d'un vecteur dans un système quelconque de coordonnées, on a par là même ses composantes dans tous les autres systèmes de coordonnées obtenus au moyen d'une transformation affine quelconque. Donnons des exemples de vecteurs des deux types. Tout d'abord, le rayon vecteur unissant deux points donnés de l'espace est, évidemment, un vecteur contravariant puisque ses composantes (les différences des coordonnées de ses extrémités) se transforment suivant les mêmes formules linéaires que les coordonnées elles-mêmes. Supposons que les coordonnées du point (x_1, x_2, x_3) soient des fonctions d'un paramètre t et définissons le vecteur vitesse de composantes:

$$\left(\frac{dx_1}{dt}, \frac{dx_2}{dt}, \frac{dx_3}{dt}\right).$$

En différentiant les formules fondamentales (36) par rapport à t, on vérifie que le vecteur vitesse est aussi un vecteur contravariant.

Donnons maintenant un exemple de vecteur covariant. Soit $f(x_1, x_2, x_3)$ une fonction du point dans l'espace; définissons dans un système de coordonnées quelconque le vecteur appelé *gradient* de cette fonction de composantes:

$$\frac{\partial f}{\partial x_1}, \quad \frac{\partial f}{\partial x_2}, \quad \frac{\partial f}{\partial x_3}.$$

On aura, d'après la règle de différentiation des fonctions de fonction et d'après les formules (36):

$$\frac{\partial f}{\partial x_s} = a_{1s}\frac{\partial f}{\partial x_1'} + a_{2s}\frac{\partial f}{\partial x_2'} + a_{3s}\frac{\partial f}{\partial x_3'} \qquad (s = 1, 2, 3),$$

c'est-à-dire que les composantes du gradient dans le système d'axes (x_1, x_2, x_3) s'expriment au moyen des composantes du gradient dans le système d'axes (x_1', x_2', x_3') à l'aide d'une transformation linéaire de tableau $A^{(*)}$ et, par conséquent, les composantes dans le système d'axes (x_1', x_2', x_3') s'expriment au moyen des composantes du gradient dans le système d'axes (x_1, x_2, x_3) à l'aide d'une transformation linéaire de tableau $A^{(*)-1} = V$, c'est-à-dire que le gradient de la fonction est un vecteur covariant.

Il est facile d'exprimer les formules (37) et (38) au moyen des dérivées partielles des nouvelles coordonnées par rapport aux anciennes et inversement. Introduisons des notations quelque peu différentes des précédentes qui sont couramment utilisées dans la théorie des vecteurs. On désigne les composantes des vecteurs contravariants au moyen d'un indice placé en haut, et pour les composantes des vecteurs covariants, on place l'indice en bas. Toutefois, on met l'indice en haut, pour désigner les coordonnées.

On peut représenter les coefficients de la transformation (36) sous forme de dérivées partielles:

$$a_{ik} = \frac{\partial x'^{(i)}}{\partial x^{(k)}}. \tag{39}$$

Les éléments de la matrice contragrédiente V sont:

$$V_{ik} = \frac{A_{ik}}{D(A)},$$

et la matrice $(A^{-1})^{(*)}$ a les mêmes éléments, c'est-à-dire:

$$(A^{(*)})^{-1} = (A^{-1})^{(*)},$$

on peut donc passer d'abord à la matrice inverse, puis échanger les lignes et les colonnes. Lors du passage à la matrice inverse, le coefficient c_{ik} est $\dfrac{\partial x^{(i)}}{\partial x'^{(k)}}$, et après transposition on obtient pour les éléments de la matrice V les expressions:

$$V_{ik} = \frac{\partial x^{(k)}}{\partial x'^{(i)}}. \tag{40}$$

Soient $u^{(s)}$ les composantes d'un vecteur contravariant dans le système d'axes $x^{(k)}$, et $u'^{(s)}$ dans le système $x'^{(s)}$. On a par définition:

$$u'^{(i)} = \sum_{s=1}^{3} \frac{\partial x'^{(i)}}{\partial x^{(s)}} u^{(s)} \qquad (i = 1, 2, 3). \tag{41}$$

De même, par définition on obtient pour un vecteur covariant :

$$u'_i = \sum_{s=1}^{3} \frac{\partial x^{(s)}}{\partial x'^{(i)}} u_s. \tag{42}$$

Notons qu'on peut utiliser ces formules pour définir les composantes d'un vecteur non seulement lors d'une transformation linéaire des coordonnées, mais aussi dans le cas très général d'une transformation où les coordonnées s'expriment les unes en fonction des autres au moyen de fonctions quelconques, non linéaires en général.

Il existe encore un moyen pour déterminer si un vecteur est covariant, un vecteur contravariant étant alors défini simplement comme un vecteur dont les composantes se transforment au moyen des mêmes formules que les coordonnées. Soit un vecteur contravariant $u^{(s)}$ et un vecteur covariant v_s.

Formons la somme des produits:

$$u^{(1)}v_1 + u^{(2)}v_2 + u^{(3)}v_3. \tag{43}$$

Elle reste invariante, c'est-à-dire que c'est un scalaire, si $u^{(s)}$ et v_s varient respectivement suivant les formules (41) et (42).

En effet, on a d'après la règle de différentiation des fonctions de fonction:

$$\sum_{s=1}^{3} u'^{(s)}v'_s = \sum_{s=1}^{3} \left[\sum_{h=1}^{3} \frac{\partial x'^{(s)}}{\partial x^{(h)}} u^{(h)}\right]\left[\sum_{l=1}^{3} \frac{\partial x^{(l)}}{\partial x'^{(s)}} v_l\right] =$$
$$= u^{(1)}v_1 + u^{(2)}v_2 + u^{(3)}v_3.$$

De sorte qu'ayant déterminé, par la méthode indiquée ci-dessus, qu'un vecteur est contravariant, on peut trouver la loi de variation des composantes d'un vecteur covariant en imposant que l'expression (43) soit invariante. En utilisant exactement les mêmes calculs que ceux du paragraphe précédent, on trouve que si l'expression (43) reste invariante, les composantes v_s doivent subir une transformation linéaire contragrédiente de celle des composantes $u^{(s)}$. Le lecteur peut démontrer que pour toute transformation des coordonnées (non obligatoirement linéaire), le vecteur vitesse est contravariant et que le gradient d'une fonction est un vecteur covariant.

On a défini ci-dessus, de façon purement formelle, les vecteurs contravariants et covariants à l'aide des formules qui permettent de passer d'un système de coordonnées à un autre. Soit un vecteur **x** dont on connaît la grandeur et la direction. Un système de vecteurs de base étant défini, on obtient les composantes du vecteur au moyen de la formule (28). Ces composantes sont des *composantes contravariantes* et on peut mettre la formule (28) sous la forme:

$$\mathbf{x} = x^{(1)}\mathbf{i} + x^{(2)}\mathbf{j} + x^{(3)}\mathbf{k}. \tag{44}$$

La *composante covariante* du vecteur **x** par rapport au vecteur de base **i** est la valeur de la projection orthogonale de **x** sur **i** multipliée par la longueur de **i**, et il en est de même pour les deux autres vecteurs de base. On a ainsi trois composantes covariantes (x_1, x_2, x_3) pour chaque système de vecteurs de base. On peut montrer qu'elles se transforment comme les composantes d'un vecteur covariant lorsque l'on passe d'un système de vecteurs de base à un autre. Dans ce cas, on démontre que l'expression $x^{(1)}x_1 + x^{(2)}x_2 + x^{(3)}x_3$ donne le carré de la longueur du vecteur **x** et qu'elle reste donc invariante par rapport à toute transformation des vecteurs de base.

II-1-4. Notion de tenseur. Généralisons maintenant la notion de vecteur. Etudions d'abord les transformations linéaires des coordonnées. Soit un tableau de neuf nombres donnés dans un système de coordonnées:

$$b_{ik} \ (i, k = 1, 2, 3).$$

Formons l'expression :

$$\sum_{i, k=1}^{3} b_{ik} u^{(i)} v^{(k)}, \tag{45}$$

où $u^{(i)}$ et $v^{(k)}$ sont les composantes de deux vecteurs contravariants. En passant à de nouvelles coordonnées, on peut exprimer dans l'expression (45) $u^{(i)}$ et $v^{(k)}$ en fonction des nouvelles composantes $u'^{(i)}$ et $v'^{(k)}$ et on a ainsi l'expression (45) sous la forme:

$$\sum_{i, k=1}^{3} b_{ik} u^{(i)} v^{(k)} = \sum_{i, k=1}^{3} b'_{ik} u'^{(i)} v'^{(k)}. \tag{46}$$

Dans le nouveau système de coordonnées on a un tableau de neuf nombres avec pour éléments b'_{ik}. Ce tableau, défini dans n'importe quel système de coordonnées en imposant que l'expression (45) soit invariante, est un *tenseur covariant du deuxième ordre*. De même, si l'on prend deux vecteurs covariants de composantes u_i et v_k et si l'on forme l'expression:

$$\sum_{i, k=1}^{3} b^{(i, k)} u_i v_k, \tag{47}$$

en se donnant un tableau de neuf nombres $b^{(i, k)}$ dans un système quelconque de coordonnées et en imposant que l'expression (47) soit invariante, on obtient un tableau de neuf nombres dans n'importe quel autre système de coordonnées. On définit ainsi un *tenseur contravariant du deuxième ordre*. Enfin, en prenant un vecteur contravariant de composantes $u^{(i)}$ et un vecteur covariant de composantes v_k et en formant l'expression:

$$\sum_{i, k=1}^{3} b_i^{(k)} u^{(i)} v_k, \tag{48}$$

on arrive de la même façon à la notion de *tenseur mixte du deuxième ordre*.

Montrons maintenant la façon dont on peut, avec les coefficients de la transformation linéaire des coordonnées (36), écrire les formules qui expriment les composantes d'un tenseur dans un nouveau système de coordonnées en fonction de ses composantes dans l'ancien système. Considérons d'abord le cas d'un tenseur covariant du deuxième ordre. Les composantes $u^{(i)}$ et $v^{(k)}$ des vecteurs contravariants dans l'ancien système de coordonnées s'expriment avec les composantes $u'^{(i)}$ et $v'^{(k)}$ dans le nouveau système de coordonnées au moyen d'une transformation linéaire de tableau A^{-1}. En désignant les éléments de ce tableau par $\{A^{-1}\}_{ik}$, on a:

$$u^{(i)} = \sum_{k=1}^{3} \{A^{-1}\}_{ik} u'^{(k)} ; \quad v^{(i)} = \sum_{k=1}^{3} \{A^{-1}\}_{ik} v'^{(k)}.$$

En substituant dans l'expression (45) et en définissant le coefficient du produit $u'^{(i)} v'^{(k)}$, on obtient l'expression de la composante b'_{ik} du tenseur dans le nouveau système de coordonnées:

$$b'_{ik} = \sum_{p,\,q=1}^{3} b_{pq}\{A^{-1}\}_{pi}\{A^{-1}\}_{qk}. \tag{49}$$

Dans le cas d'un tenseur contravariant du deuxième ordre, on exprime exactement de la même façon les composantes des vecteurs covariants u_i et v_k en fonction des nouvelles composantes. D'après la définition du vecteur covariant, u'_i s'exprime en fonction de u_i au moyen du tableau $A^{(*)-1}$ et, par conséquent, u_i s'exprime en fonction de u'_i au moyen du tableau $A^{(*)}$ transposé du tableau A; il en est de même pour v_i, c'est-à-dire:

$$u_i = \sum_{k=1}^{3} \{A\}_{ki}\, u'_k\,; \quad v_i = \sum_{k=1}^{3} \{A\}_{ki}\, v'_k.$$

En portant ces valeurs dans l'expression (47), on obtient les formules de transformation des composantes d'un tenseur contravariant du deuxième ordre:

$$b'^{(i,\,k)} = \sum_{p,\,q=1}^{3} b^{(p,\,q)}\{A\}_{ip}\{A\}_{kq}. \tag{50}$$

On a exactement de la même façon, pour les composantes d'un tenseur mixte du deuxième ordre, la formule de transformation suivante:

$$b'^{(k)}_{i} = \sum_{p,\,q=1}^{3} b^{(q)}_{p}\{A^{-1}\}_{pi}\{A\}_{kq}. \tag{51}$$

Si l'on exprime les coefficients de la transformation linéaire au moyen des dérivées partielles:

$$\frac{\partial x'^{(i)}}{\partial x^{(k)}} \quad \text{et} \quad \frac{\partial x^{(i)}}{\partial x'^{(k)}}$$

et si l'on porte ces expressions dans les formules précédentes, on a les formules de transformation des tenseurs du deuxième ordre dans le cas d'une transformation de coordonnées quelconque. On peut définir de façon tout à fait analogue la notion de tenseur d'ordre supérieur à deux, mais nous ne nous y arrêterons pas.

On s'est trouvé jusqu'ici toujours en présence d'un tableau représentant une transformation linéaire d'un espace à trois dimensions rapporté à des axes de coordonnées. Soit B ce tableau, et supposons que l'on ait effectué une transformation affine des coordonnées suivant les formules:

$$(y_1,\ y_2,\ y_3) = A\,(x_1,\ x_2,\ x_3),$$

où A est un tableau qui a un déterminant non nul. Comme on l'a démontré plus haut, la transformation de l'espace aura, dans les nouvelles coordonnées, le tableau:

$$ABA^{-1}.$$

Cette transformation du tableau B coïncide, comme on le constate facilement, avec la transformation indiquée plus haut pour un tenseur mixte du deuxième ordre. En effet, si l'on applique les règles de multiplication des tableaux, on a les formules suivantes:

$$\{BA^{-1}\}_{qi} = \sum_{p=1}^{3} \{B\}_{qp} \{A^{-1}\}_{pi},$$

et

$$\{A(BA^{-1})\}_{ki} = \sum_{q=1}^{3} \{A\}_{kq} \{BA^{-1}\}_{qi} = \sum_{p,\, q=1}^{3} \{B\}_{qp} \{A^{-1}\}_{pi} \{A\}_{kq}.$$

En désignant $\{B\}_{ik}$ par $b_k^{(i)}$, on obtient des formules de la forme (51)

Examinons quelques tenseurs particuliers. Supposons qu'un tenseur covariant ait dans un système de coordonnées la propriété suivante:

$$b_{ik} = b_{ki} \qquad (i,\ k = 1,\ 2,\ 3). \tag{52}$$

Il est facile de constater qu'il aura la même propriété dans n'importe quel autre système de coordonnées. En effet, d'après (49):

$$b'_{ki} = \sum_{p,\, q=1}^{3} b_{pq} \{A^{-1}\}_{pk} \{A^{-1}\}_{qi}$$

ou compte tenu de (52):

$$b'_{ki} = \sum_{p,\, q=1}^{3} b_{qp} \{A^{-1}\}_{pk} \{A^{-1}\}_{qi},$$

ou, si l'on change les indices de sommation:

$$b'_{ki} = \sum_{p,\, q=1}^{3} b_{pq} \{A^{-1}\}_{qk} \{A^{-1}\}_{pi},$$

d'où l'on voit que b'_{ki} coïncide effectivement avec b'_{ik}. Ce tenseur est un *tenseur covariant symétrique*. La définition d'un *tenseur contravariant symétrique* est analogue. Si dans un système de coordonnées on a $b_{ik} = -b_{ki}$ ou $b^{(i,k)} = -b^{(k,i)}$, il en sera de même dans n'importe quel autre système de coordonnées, et les tenseurs correspondants sont des *tenseurs antisymétriques*. Pour un tenseur mixte, la relation $b_i^{(k)} = b_k^{(i)}$ n'est plus vérifiée lors d'une transformation des coordonnées. Passons maintenant à l'étude de quelques cas particuliers de tenseurs.

II-1-5. Exemples de tenseurs orthogonaux affines. Dans les exemples suivants nous nous limiterons aux transformations linéaires déjà étudiées [II-1-1] qui correspondent au passage d'un système cartésien à un autre. Ces transformations sont appelées habituellement transformations orthogonales dans un espace à trois dimensions. Pour elles, la transformation contragrédiente $A^{(*)-1}$ coïncide avec A, comme on l'a déjà vu, et il n'y a plus de différence entre vecteur covariant et vecteur contravariant. On n'aura pour ces transformations de coordonnées qu'une seule notion de tenseur du deuxième ordre. En désignant le tableau des coefficients de la transformation orthogonale des coordonnées, comme ci-dessus, par $\{A\}_{ik}$, on obtient la formule de transformation:

$$b'_{ik} = \sum_{p,\, q=1}^{3} b_{pq} \{A\}_{ip} \{A\}_{kq}. \tag{53}$$

qui résulte des formules du paragraphe précédent. On considère les éléments de chaque colonne du tableau $\| b_{ik} \|$ comme les composantes d'un vecteur. On a ainsi trois vecteurs:

$$\mathfrak{b}^{(1)}(b_{11}, b_{21}, b_{31}); \qquad \mathfrak{b}^{(2)}(b_{12}, b_{22}, b_{32}); \qquad \mathfrak{b}^{(3)}(b_{13}, b_{23}, b_{33}).$$

On supposera que le premier vecteur correspond à l'axe x_1, le second à l'axe x_2 et le troisième à l'axe x_3. On associe à toute direction (n) le vecteur $\mathfrak{b}^{(n)}$ au moyen de la formule:

$$\mathfrak{b}^{(n)} = \cos(n, x_1)\, \mathfrak{b}^{(1)} + \cos(n, x_2)\, \mathfrak{b}^{(2)} + \cos(n, x_3)\, \mathfrak{b}^{(3)}. \tag{54}$$

Prenons maintenant des axes cartésiens quelconques (x_1', x_2', x_3') à la place des anciens (x_1, x_2, x_3), et déterminons d'après la formule (54) des vecteurs correspondant aux directions des nouveaux axes de coordonnées:

$$\mathfrak{b}'^{(k)} = \cos(x_k', x_1)\, \mathfrak{b}^{(1)} + \cos(x_k', x_2)\, \mathfrak{b}^{(2)} + \cos(x_k', x_3)\, \mathfrak{b}^{(3)}. \tag{55}$$

Si l'on étudie les projections de ces vecteurs sur de nouveaux axes de coordonnées x_1', x_2', x_3', on a un tableau de neuf nombres $\| b_{ik}' \|$ analogue au tableau $\| b_{ik} \|$. Démontrons que les éléments du nouveau tableau s'expriment en fonction des éléments du tableau b_{il} d'après les formules de transformation des composantes des tenseurs du deuxième ordre. Soit, en effet, l'élément b_{12}'. Par définition, c'est la composante du vecteur $\mathfrak{b}'^{(2)}$ sur le nouvel axe x_1'. La formule (55) donne:

$$\mathfrak{b}'^{(2)} = \cos(x_2', x_1)\, \mathfrak{b}^{(1)} + \cos(x_2', x_2)\, \mathfrak{b}^{(2)} + \cos(x_2', x_3)\, \mathfrak{b}^{(3)}, \tag{56}$$

d'où l'on voit que $\mathfrak{b}'^{(2)}$ est une fonction linéaire des vecteurs $\mathfrak{b}^{(l)}$ et il suffit, pour obtenir b_{12}', de remplacer dans le second membre de la formule (56) les vecteurs $\mathfrak{b}^{(l)}$ par leurs projections sur l'axe x_1', c'est-à-dire de remplacer ces vecteurs par les expressions suivantes: $b_{1l} \cos(x_1', x_1) + b_{2l} \cos(x_1', x_2) + b_{3l} \cos(x_1', x_3)$ $(l = 1, 2, 3)$. Il faut remarquer en outre que, d'après le tableau (2), $\cos(x_i', x_k) = a_{ik} = \{A\}_{ik}$.

En portant ces expressions dans le second membre de la formule (56), on a:

$$b_{12}' = \sum_{p,\, q=1}^{3} b_{pq} \{A\}_{1p} \{A\}_{2q},$$

ce qui est analogue à la formule (53). On peut, donc, affirmer que si trois vecteurs $\mathfrak{b}^{(1)}, \mathfrak{b}^{(2)}, \mathfrak{b}^{(3)}$ sont perpendiculaires et si d'après la formule (54) un vecteur est défini quelle que soit la direction (n), le tableau de neuf nombres, qui donnent les projections des vecteurs $\mathfrak{b}'^{(k)}$ $(k = 1, 2, 3)$ sur l'axe $x'^{(k)}$ dans un système de coordonnées cartésiennes quelconque, définit un tenseur orthogonal affine du deuxième ordre, c'est-à-dire un tenseur du deuxième ordre défini pour toutes les transformations orthogonales possibles.

Il faut remarquer que lorsqu'on dit que $\mathfrak{b}^{(1)}$ correspond au sens de l'axe x_1, cela ne signifie pas que $\mathfrak{b}^{(1)}$ doit avoir la même direction que l'axe x_1. Seule la formule (54) est très importante: elle associe à une direction (n) un vecteur $\mathfrak{b}^{(n)}$ dont la direction ne coïncide pas en général avec (n).

Donnons maintenant deux exemples de tenseurs orthogonaux affines du deuxième ordre. Le premier est le tenseur des contraintes en élasticité. Étudions un corps élastique déformé; on fait passer par un point fixe M de ce corps une surface infiniment petite $d\sigma$ ayant pour normale (n). On montre dans la théorie de l'élasticité que l'action de la partie du corps élastique, qui se trouve du côté de la normale définie par son sens, sur cette surface est égale au produit d'un vecteur $\mathfrak{b}^{(n)}$, dépendant du sens de la normale (n), par la surface $d\sigma$. En étudiant les conditions d'équilibre d'un tétraèdre infiniment petit extrait du corps élastique, on obtient l'égalité (54) et la contrainte est bien un tenseur du deuxième ordre. Ce tenseur sera caractérisé dans n'importe quel système de coordonnées cartésiennes par un tableau de neuf nombres $\| b_{ik} \|$, et ce tenseur, comme on le démontre en élasticité, est symétrique, c'est-à-dire que $b_{ik} = b_{ki}$. Autrement dit, la projection sur l'axe x_i de la contrainte agissant sur une surface perpendiculaire à l'axe x_k est égale à la projection sur l'axe x_k de la contrainte agissant sur une surface perpendiculaire à l'axe x_i.

Passons maintenant à un autre exemple de tenseur. Soit un champ vectoriel $\mathbf{C}(M)$. Si l'on prend un système de coordonnées cartésiennes (x_1, x_2, x_3) et qu'on prenne les dérivées des composantes du champ (c_1, c_2, c_3) par rapport aux coordonnées, on obtient un tableau de neuf grandeurs:

$$\left\| \begin{matrix} \frac{\partial c_1}{\partial x_1} & \frac{\partial c_1}{\partial x_2} & \frac{\partial c_1}{\partial x_3} \\ \frac{\partial c_2}{\partial x_1} & \frac{\partial c_2}{\partial x_2} & \frac{\partial c_2}{\partial x_3} \\ \frac{\partial c_3}{\partial x_1} & \frac{\partial c_3}{\partial x_2} & \frac{\partial c_3}{\partial x_3} \end{matrix} \right\| . \tag{57}$$

On définit, pour une direction quelconque (n), le vecteur correspondant à cette direction comme étant la dérivée $\frac{\partial \mathbf{c}}{\partial n}$, de sorte que les éléments du tableau (57) qui se trouvent dans la k-ième colonne, donnent les composantes du vecteur correspondant à la direction de l'axe x_k. On a pour toute direction (n) la formule (tome II, [IV-2-2]):

$$\frac{\partial c_i}{\partial n} = \cos(n, x_1) \frac{\partial c_i}{\partial x_1} + \cos(n, x_2) \frac{\partial c_i}{\partial x_2} + \cos(n, x_3) \frac{\partial c_i}{\partial x_3} \quad (i = 1, 2, 3),$$

c'est-à-dire que le tableau qui vient d'être défini est un tenseur du deuxième ordre. Ce tenseur ne sera en général ni symétrique ni antisymétrique. Mais il est facile de le mettre sous forme de la somme d'un tenseur symétrique et d'un tenseur antisymétrique, si l'on entend par somme de deux tableaux la somme des éléments correspondants de ces deux tableaux.

Faisons d'abord quelques remarques générales. Les formules (53) étant linéaires, il en résulte que si $\| b_{ik} \|$ et $\| c_{ik} \|$ sont deux tenseurs, la somme $\| b_{ik} + c_{ik} \|$ est également un tenseur. De plus, cette formule reste vraie si l'on intervertit les indices, c'est-à-dire que:

$$b'_{ki} = \sum_{p, q=1}^{3} b_{qp} \{A\}_{ip} \{A\}_{kq},$$

de sorte que si des tableaux définis pour tous les axes donnent un tenseur, les tableaux transposés donnent aussi un tenseur. Soit un tenseur $\| b_{ik} \|$. On peut le mettre sous forme de la somme:

$$\| b_{ik} \| = \left\| \frac{b_{ik} + b_{ki}}{2} \right\| + \left\| \frac{b_{ik} - b_{ki}}{2} \right\| .$$

Le premier terme est évidemment un tenseur symétrique et le second un tenseur antisymétrique.

En appliquant ce procédé au tenseur défini par le tableau (57), sa partie symétrique est :

$$\left\| \begin{matrix} \frac{\partial c_1}{\partial x_1}, & \frac{1}{2}\left(\frac{\partial c_1}{\partial x_2}+\frac{\partial c_2}{\partial x_1}\right), & \frac{1}{2}\left(\frac{\partial c_1}{\partial x_3}+\frac{\partial c_3}{\partial x_1}\right) \\ \frac{1}{2}\left(\frac{\partial c_1}{\partial x_2}+\frac{\partial c_2}{\partial x_1}\right), & \frac{\partial c_2}{\partial x_2}, & \frac{1}{2}\left(\frac{\partial c_2}{\partial x_3}+\frac{\partial c_3}{\partial x_2}\right) \\ \frac{1}{2}\left(\frac{\partial c_1}{\partial x_3}+\frac{\partial c_3}{\partial x_1}\right), & \frac{1}{2}\left(\frac{\partial c_2}{\partial x_3}+\frac{\partial c_3}{\partial x_2}\right), & \frac{\partial c_3}{\partial x_3} \end{matrix} \right\|. \tag{58}$$

S'il y a une déformation du milieu continu et si $\overline{MC}$ est le vecteur déplacement, c'est-à-dire le vecteur dont se déplace le point M du milieu, le tableau (58) définit le *tenseur des déformations*. La partie antisymétrique du tenseur est :

$$\left\| \begin{matrix} 0, & \frac{1}{2}\left(\frac{\partial c_1}{\partial x_2}-\frac{\partial c_2}{\partial x_1}\right), & \frac{1}{2}\left(\frac{\partial c_1}{\partial x_3}-\frac{\partial c_3}{\partial x_1}\right) \\ \frac{1}{2}\left(\frac{\partial c_2}{\partial x_1}-\frac{\partial c_1}{\partial x_2}\right), & 0, & \frac{1}{2}\left(\frac{\partial c_2}{\partial x_3}-\frac{\partial c_3}{\partial x_2}\right) \\ \frac{1}{2}\left(\frac{\partial c_3}{\partial x_1}-\frac{\partial c_1}{\partial x_3}\right), & \frac{1}{2}\left(\frac{\partial c_3}{\partial x_2}-\frac{\partial c_2}{\partial x_3}\right), & 0 \end{matrix} \right\|. \tag{59}$$

On a déjà divisé un tenseur en deux parties dans le cas particulier d'une déformation linéaire homogène (tome II, [IV-2-7]) et l'on a vu que, dans ce cas, une rotation du milieu continu (sans déformation) autour d'un axe correspondait à la partie antisymétrique.

II-1-6. Cas d'un espace complexe à n dimensions. Considérons maintenant le cas général d'un espace à n dimensions. On a déjà dit qu'un vecteur dans cet espace est une suite de n nombres réels ou complexes [I-2-5] :

$$\mathbf{x}\,(x_1,\ x_2,\ \ldots,\ x_n).$$

Ces nombres sont les composantes du vecteur $\mathbf{x}$. On suppose que l'espace soit rapporté à des vecteurs de base déterminés :

$$\mathbf{a}^{(1)}(1,\ 0,\ \ldots,\ 0)\ ;\ \mathbf{a}^{(2)}(0,\ 1,\ \ldots,\ 0)\ ;\ \ldots;\ \mathbf{a}^{(n)}(0,\ 0,\ \ldots,\ 1),$$

de sorte que :

$$\mathbf{x}=x_1\mathbf{a}^{(1)}+x_2\mathbf{a}^{(2)}+\ldots+x_n\mathbf{a}^{(n)}. \tag{60}$$

On a déjà défini [I-2-5] la condition pour que deux vecteurs soient égaux et les opérations élémentaires qu'on effectue sur eux.

On appelle **transformation linéaire** dans un espace à n dimensions le passage d'un vecteur $\mathbf{x}\,(x_1,\ x_2,\ \ldots,\ x_n)$ à un vecteur $\mathbf{y}\,(y_1,\ y_2,\ \ldots\ \ldots,\ y_n)$ suivant les formules :

$$y_i=a_{i1}x_1+a_{i2}x_2+\ldots+a_{in}x_n \quad (i=1,\ 2,\ \ldots,\ n), \tag{61}$$

ou :

$$\mathbf{y}=A\mathbf{x}, \tag{62}$$

où A est le tableau ou la matrice $\| a_{ik} \|_1^n$ de la transformation. Si son déterminant $D(A)$ n'est pas nul, on dit que la transformation (62) est une transformation non singulière et que la matrice A est une matrice (un tableau) non singulière. Dans ce cas, on obtient en résolvant l'équation (61) par rapport à x_i la transformation inverse de (61) ou (62):

$$\mathbf{x} = A^{-1}\mathbf{y}, \tag{63}$$

où le tableau A^{-1} a pour éléments:

$$\{A^{-1}\}_{ik} = \frac{A_{ki}}{D(A)}; \tag{64}$$

on désigne le déterminant du tableau A par $D(A)$ et les complémentaires algébriques des éléments a_{ik} par A_{ik}.

Ensuite, d'une façon identique à la précédente [II-1-2], on définit le produit de deux transformations comme l'application successive de deux transformations:

$$\mathbf{y} = A\mathbf{x}; \qquad \mathbf{z} = B\mathbf{y};$$

ce qui est équivalent à une transformation linéaire:

$$\mathbf{z} = BA\mathbf{x},$$

que l'on appelle *produit* des transformations A et B et dont le tableau est défini par:

$$\{BA\}_{ik} = \sum_{s=1}^{n} \{B\}_{is} \{A\}_{sk}. \tag{65}$$

Ce produit dépend généralement de l'ordre des facteurs, c'est-à-dire que (sauf exception) on a:

$$BA \neq AB.$$

Il est facile d'étendre la définition du produit au cas d'un nombre quelconque de facteurs, de plus, le produit est associatif, c'est-à-dire qu'on peut regrouper les facteurs:

$$(CB)A = C(BA). \tag{66}$$

La transformation inverse vérifie les relations:

$$AA^{-1} = A^{-1}A = I; \qquad (A^{-1})^{-1} = A, \tag{67}$$

où l'on désigne par le symbole I la matrice unité dont les éléments situés sur la diagonale principale sont égaux à un, et dont tous les autres éléments sont nuls. La transformation identique $y_i = x_i$ $(i = 1, 2, \ldots, n)$ correspond à cette matrice.

On définit la matrice diagonale d'ordre n:

$$[k_1, k_2, \ldots, k_n] = \begin{Vmatrix} k_1 & 0 & 0 & \ldots & 0 \\ 0 & k_2 & 0 & \ldots & 0 \\ 0 & 0 & k_3 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \ldots & k_n \end{Vmatrix}. \qquad (68)$$

Il lui correspond la transformation $y_i = k_i x$ $(i = 1, 2, \ldots, n)$.

Le produit de matrices diagonales ne dépend pas de l'ordre des facteurs et il est défini par la formule:

$$[k_1, k_2, \ldots, k_n][l_1, l_2, \ldots, l_n] = \\ = [l_1, l_2, \ldots, l_n][k_1, k_2, \ldots, k_n] = \\ = [k_1 l_1, k_2 l_2, \ldots, k_n l_n].$$

Dans le cas particulier où $k_1 = k_2 = \ldots = k_n = k$, on obtient la matrice:

$$[k, k, \ldots, k] = \begin{Vmatrix} k & 0 & 0 & \ldots & 0 \\ 0 & k & 0 & \ldots & 0 \\ 0 & 0 & k & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \ldots & k \end{Vmatrix}, \qquad (69)$$

à laquelle correspond le produit de toutes les composantes du vecteur par le nombre k. On considère, d'après ce qui a été dit au début de ce paragraphe, que la matrice (69) est simplement un nombre k, c'est-à-dire que l'on considère le nombre k comme un cas particulier de matrice. On voit facilement que si l'on utilise la formule (65), le produit du nombre k, pris pour matrice (69), par une matrice A quelconque ne dépend pas de l'ordre des facteurs et se ramène au produit de tous les éléments de la matrice A par le nombre k:

$$\{[k, k, \ldots, k] A\}_{ik} = \{kA\}_{ik} = k\{A\}_{ik}. \qquad (70)$$

Supposons maintenant qu'on ait pris comme vecteurs de base non pas les vecteurs $\mathbf{a}^{(k)}$, mais les vecteurs $\mathbf{b}^{(k)}$ qui s'expriment en fonction de $\mathbf{a}^{(k)}$ d'après les formules:

$$\left.\begin{array}{l} \mathbf{b}^{(1)} = t_{11}\mathbf{a}^{(1)} + t_{12}\mathbf{a}^{(2)} + \ldots + t_{1n}\mathbf{a}^{(n)}, \\ \mathbf{b}^{(2)} = t_{21}\mathbf{a}^{(1)} + t_{22}\mathbf{a}^{(2)} + \ldots + t_{2n}\mathbf{a}^{(n)}, \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ \mathbf{b}^{(n)} = t_{n1}\mathbf{a}^{(1)} + t_{n2}\mathbf{a}^{(2)} + \ldots + t_{nn}\mathbf{a}^{(n)}, \end{array}\right\} \qquad (71)$$

où le déterminant constitué par les éléments t_{ih} n'est pas nul. Réciproquement, les vecteurs $\mathbf{a}^{(k)}$ s'expriment linéairement en fonction des vecteurs $\mathbf{b}^{(k)}$, et toute combinaison linéaire des vecteurs $\mathbf{a}^{(k)}$ est en même temps une combinaison linéaire des vecteurs $\mathbf{b}^{(k)}$ et inversement. Autrement dit, les vecteurs $\mathbf{b}^{(k)}$, en tant que vecteurs de base, engendrent le même espace que les vecteurs $\mathbf{a}^{(k)}$. Si un vecteur $\mathbf{x}$ a dans un système de coordonnées défini par les vecteurs de base $\mathbf{a}^{(k)}$ des composantes $(x_1, \ldots, x_n)$, il aura, dans un système de coordonnées défini par les vecteurs de base $\mathbf{b}_k$ d'autres composantes $(x'_1, \ldots, x'_n)$ qui s'expriment en fonction des anciennes au moyen d'une transformation linéaire contragrédiente de la transformation (71), ce qu'on peut écrire:

$$(x'_1, \ldots, x'_n) = T^{(*)-1}(x_1, \ldots, x_n), \tag{72}$$

où le tableau $T^{(*)}$ est le transposé du tableau T qui correspond à la transformation (71).

Si l'on a une transformation de l'espace définie dans le premier système de coordonnées par la formule (62), cette transformation s'exprime dans le nouveau système de coordonnées par la formule:

$$\mathbf{y}' = UAU^{-1}\mathbf{x}', \tag{73}$$

où $U = T^{(*)-1}$.

La matrice

$$UAU^{-1}$$

est appelée matrice *semblable* à A.

Dans l'analyse précédente les notions de vecteur et de matrice étaient fondamentales. Il faut remarquer que l'on considère aussi parfois le vecteur $\mathbf{x}$ $(x_1, \ldots, x_n)$ comme une matrice dont l'une des colonnes, peu importe laquelle, est $(x_1, \ldots, x_n)$, et dont les autres éléments sont nuls. Supposons, par exemple, que l'on place les composantes du vecteur dans la première colonne. On a ainsi le vecteur sous forme d'une matrice:

$$\begin{Vmatrix} x_1 & 0 \ldots 0 \\ x_2 & 0 \ldots 0 \\ \cdots & \cdots \\ x_n & 0 \ldots 0 \end{Vmatrix}.$$

On désigne parfois une matrice, dont une seule colonne contient des éléments différents de zéro, par le symbole:

$$\begin{pmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ x_n \end{pmatrix} = \begin{Vmatrix} x_1 & 0 \ldots 0 \\ x_2 & 0 \ldots 0 \\ \cdots & \cdots \\ x_n & 0 \ldots 0 \end{Vmatrix}. \tag{74}$$

Montrons à présent que la transformation linéaire (62) peut être écrite sous forme du produit de la matrice (74) par la matrice de la transformation A. En effet, si l'on multiplie la matrice (74) par la matrice A selon la règle (65) et si l'on tient compte de ce que seuls les éléments de la première colonne de la matrice (74) ne sont pas nuls, on obtient le produit sous la forme d'une matrice dont seuls les éléments de la première colonne ne sont pas nuls, et ces éléments sont :

$$y_i = a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{in}x_n,$$

c'est-à-dire qu'ils donnent la transformation linéaire (62). On peut ainsi écrire cette transformation sous la forme :

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = A \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \tag{75}$$

où l'on a dans le second membre un produit de deux matrices.

On notera en conclusion de ce paragraphe les lois générales auxquelles obéissent les opérations sur les vecteurs dans un espace à n dimensions :

$$\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}; \qquad (\mathbf{x} + \mathbf{y}) + \mathbf{z} = \mathbf{x} + (\mathbf{y} + \mathbf{z}).$$

Si $\mathbf{x}$ et $\mathbf{y}$ sont deux vecteurs quelconques, le vecteur $\mathbf{z} = \mathbf{y} - \mathbf{x}$ ayant pour composantes $(y_k - x_k)$ est le seul vecteur qui vérifie la condition $\mathbf{x} + \mathbf{z} = \mathbf{y}$.

Soient a et b deux nombres quelconques. On a :

$$(a+b)\,\mathbf{x} = a\mathbf{x} + b\mathbf{x}; \quad a\,(b\mathbf{x}) = (ab)\,\mathbf{x}; \quad a\,(\mathbf{x}+\mathbf{y}) = a\mathbf{x} + a\mathbf{y}.$$

Pour l'unité on a $1\mathbf{x} = \mathbf{x}$ et $0\mathbf{x} = 0$, où le zéro qui se trouve dans le second membre désigne un vecteur dont toutes les composantes sont nulles.

II-1-7. Eléments de calcul matriciel. Dans les formules du paragraphe précédent, la matrice apparaissait comme un nouveau symbole sur lequel on pouvait agir comme sur des nombres ordinaires, d'où l'idée de construire une nouvelle algèbre qui s'appliquerait aux symboles matriciels. Autrement dit, on considérera la matrice comme un nouvel être, en quelque sorte un nombre *hypercomplexe*. De même qu'à l'aide de deux nombres réels, on a pu construire des nombres d'une nature nouvelle, à savoir les nombres complexes $a + ib$, on peut maintenant, avec n^2 nombres complexes a_{ik} disposés sous forme d'un tableau carré, arriver à la notion d'un nouvel être : la *matrice*. Il existe toutefois une différence essentielle : on a vu que l'on pouvait effectuer avec les nombres complexes toutes les opérations algébriques formelles que l'on effectue sur les nombres réels. On obtient pour les matrices une algèbre qui diffère notablement de celle des nombres complexes. Cette différence est due surtout à la *non-commutativité du produit*, c'est-à-dire au fait que le résultat d'un produit dépend de l'ordre des facteurs. Etablissons maintenant les règles fondamentales de l'algèbre des matrices, et l'on se servira souvent des résultats obtenus ci-dessus en considérant la matrice comme le tableau d'une transformation linéaire.

On étudiera dans la suite, sauf si l'on précise le contraire, des matrices carrées d'ordre n. Si A est une matrice de ce type, ses éléments seront désignés par $\{A\}_{ik}$.

Deux matrices A et B sont égales si et seulement si:

$$\{A\}_{ik}=\{B\}_{ik} \quad (i,\ k=1,\ 2,\ \ldots,\ n), \tag{76}$$

c'est-à-dire si tous leurs éléments correspondants sont identiques.

L'addition des matrices est définie par la formule:

$$\{A+B\}_{ik}=\{A\}_{ik}+\{B\}_{ik}, \tag{77}$$

c'est-à-dire qu'elle se ramène à l'addition des éléments.

Le produit est défini par la formule:

$$\{BA\}_{ik}=\sum_{s=1}^{n}\{B\}_{is}\{A\}_{sk}. \tag{78}$$

Comme on l'a vu plus haut, on a généralement $BA \neq AB$, mais d'après l'associativité [II-1-2]:

$$(CB)\,A=C\,(BA). \tag{79}$$

Le déterminant du produit est égal au produit des déterminants des matrices:

$$D\,(BA)=D\,(B)\cdot D\,(A). \tag{80}$$

Il est évident qu'on a aussi la loi de distributivité:

$$(A+B)\,C=AC+BC \quad \text{et} \quad C\,(A+B)=CA+CB. \tag{81}$$

Il faut remarquer encore une particularité du produit: le produit des matrices peut être nul, c'est-à-dire qu'il peut devenir une matrice dont tous les éléments sont nuls, même si tous les éléments des matrices initiales ne sont pas nuls. Soit par exemple le produit de deux matrices identiques du second ordre:

$$\begin{Vmatrix}0 & 0\\ 1 & 0\end{Vmatrix}\cdot\begin{Vmatrix}0 & 0\\ 1 & 0\end{Vmatrix}=\begin{Vmatrix}0 & 0\\ 0 & 0\end{Vmatrix}.$$

Comme on l'a fait dans le paragraphe précédent, on peut introduire la notion de matrice inverse A^{-1}, si A est une matrice non singulière, c'est-à-dire si $D(A) \neq 0$. Si $C = BA$ et si R_A, R_B, R_C sont les rangs des matrices A, B, C, $R_C \leqslant R_A$, comme on l'a vu [I-1-7]. Si B est une matrice non singulière, $A = B^{-1}C$, et l'on peut affirmer comme ci-dessus que $R_A \leqslant R_C$ et, par conséquent, $R_C = R_A$, c'est-à-dire que *lorsqu'on multiplie (à droite ou à gauche) la matrice A par la matrice non singulière B, son rang ne varie pas.* Pour la matrice unité I, on a la relation:

$$BI = IB = B, \tag{82}$$

où B est une matrice quelconque.

Il est aisé de voir que la matrice A^{-1} est la solution unique des équations:

$$AX = I \quad \text{et} \quad XA = I, \tag{83}$$

où I est la matrice unité. En effet, si l'on multiplie à gauche, par exemple, la première égalité par A^{-1} et si l'on tient compte de (79) et de (67), on obtient $X = A^{-1}$, et il en est de même pour la seconde équation. Il faut remarquer que si $D(A) = 0$, les équations (83) n'ont pas de solution, c'est-à-dire que la matrice A n'a pas d'inverse. En effet, on a comme conséquence des équations (83) : $D(A)\,D(X) = 1$, ce qui contredit l'hypothèse $D(A) = 0$.

On se rappelle aussi la notion de matrice diagonale introduite dans le paragraphe précédent et le fait qu'on peut considérer tout nombre k comme un cas particulier de matrice. Il est facile d'introduire la notion de puissance entière positive d'une matrice:

$$A^p = AA \ldots A.$$

Les puissances entières négatives d'une matrice s'introduisent comme puissances entières positives de la matrice inverse:

$$A^{-p} = (A^{-1})^p. \tag{84}$$

On a évidemment:

$$A^{-p} = (A^p)^{-1}, \quad \text{c'est-à-dire } A^{-p}A^p = A^pA^{-p} = I. \tag{85}$$

Le symbole du quotient de deux matrices

$$\frac{A}{B}$$

n'a pas de sens déterminé. On peut l'interpréter de deux façons: soit comme le produit AB^{-1}, soit comme le produit $B^{-1}A$, ces deux produits étant en général différents, et le symbole du quotient a un sens déterminé seulement dans les cas particuliers où ces produits coïncident.

La notion fondamentale de matrices semblables a été introduite dans le paragraphe précédent. Ecrivons quelques formules évidentes:

$$(CBA)^{-1} = A^{-1}B^{-1}C^{-1}, \tag{86}$$

$$CBAC^{-1} = (CBC^{-1})(CAC^{-1}). \tag{87}$$

Si l'on désigne par $A^{(*)}$ la matrice transposée de A, on a la formule:

$$(CBA)^{(*)} = A^{(*)}B^{(*)}C^{(*)}, \tag{88}$$

qu'il est facile de vérifier en utilisant la définition du produit. Introduisons deux notations. On désigne par $\overline{A}$ la matrice dont les éléments sont les nombres complexes conjugués des éléments de la matrice A, c'est-à-dire:

$$\{\overline{A}\}_{ik} = \overline{\{A\}_{ik}}, \tag{89}$$

et l'on désigne par le symbole $\overline{\alpha}$ le nombre complexe conjugué de α, et par $\widetilde{A}$ la matrice obtenue à partir de A lorsqu'on remplace les lignes par les colonnes et tous les éléments par leurs conjugués complexes, c'est-à-dire:

$$\{\widetilde{A}\}_{ik} = \overline{\{A\}_{ki}}. \tag{90}$$

La matrice $\widetilde{A}$ est la conjuguée hermitienne de la matrice A. Il est facile de vérifier que:

$$\widetilde{CBA} = \widetilde{A}\widetilde{B}\widetilde{C}. \tag{91}$$

Le lecteur peut vérifier également la formule élémentaire suivante:

$$(A^{(*)})^{-1} = A^{-1(*)},$$

c'est-à-dire qu'on peut intervertir le signe de la matrice inverse et de la transposition, ce qu'on a déjà fait ci-dessus [II-1-1].

Donnons encore une formule utile pour la suite. Il résulte de (67) que:

$$D(A)\,D(A^{-1}) = 1,$$

c'est-à-dire que:

$$D(A^{-1}) = D(A)^{-1}. \tag{92}$$

Autrement dit, le déterminant d'une matrice inverse est l'inverse du déterminant de la matrice initiale.

Soit encore la notion de *matrice quasi diagonale* qui est une généralisation des matrices diagonales. On va expliquer cette notion en prenant d'abord un cas particulier. Soit une matrice du septième ordre de la forme :

$$\left\| \begin{matrix} b_{11} & b_{12} & b_{13} & 0 & 0 & 0 & 0 \\ b_{21} & b_{22} & b_{23} & 0 & 0 & 0 & 0 \\ b_{31} & b_{32} & b_{33} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & c_{11} & c_{12} & 0 & 0 \\ 0 & 0 & 0 & c_{21} & c_{22} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & d_{11} & d_{12} \\ 0 & 0 & 0 & 0 & 0 & d_{21} & d_{22} \end{matrix} \right\| .$$

On désigne par B la matrice du troisième ordre d'éléments b_{ik}, et par C et D les matrices du second ordre d'éléments c_{ik} et d_{ik}. La matrice du septième ordre est *quasi diagonale* de structure $\{3, 2, 2\}$, et on la désigne par le symbole :

$$[B, C, D].$$

Supposons que la diagonale principale de la matrice d'ordre n, composée des éléments a_{ii}, soit divisée en m parties ; la première partie est formée des k_1 premiers éléments, la seconde des k_2 éléments suivants, etc., de sorte que $k_1 + \ldots + k_m = n$. On peut considérer les k_1 premiers éléments comme la diagonale principale d'une matrice X_1 d'ordre k_1, les k_2 éléments suivants comme la diagonale principale d'une matrice X_2 d'ordre k_2, etc. Supposons que tous les éléments de la matrice A n'appartenant pas aux matrices X_s soient nuls. La matrice A est alors appelée *matrice quasi diagonale de structure* $\{k_1, \ldots, k_m\}$, et on la désigne par :

$$A = [X_1, X_2, \ldots, X_m].$$

Les calculs sur des matrices quasi diagonales de même structure sont très simples. On donnera les formules correspondantes sans les démontrer. La démonstration peut être faite de façon élémentaire en partant de la définition de ces opérations. Pour l'addition de matrices quasi diagonales de même structure, on a :

$$[X_1\ X_2, \ldots, X_m] + [Y_1, Y_2, \ldots, Y_m] =$$
$$= [X_1 + Y_1,\ X_2 + Y_2, \ldots, X_m - Y_m]. \quad (93)$$

et l'identité des structures conduit à ce que l'ordre de chaque matrice X_k coïncide avec celui de la matrice correspondante Y_k. De même pour le produit et l'élévation à une puissance, on a :

$$[Y_1, Y_2, \ldots, Y_m]\,[X_1, X_2, \ldots, X_m] =$$
$$= [Y_1X_1,\ Y_2X_2, \ldots, Y_mX_m]. \quad (94)$$

$$[X_1, X_2, \ldots, X_m]^p = [X_1^p, X_2^p, \ldots, X_m^p], \quad (95)$$

où p est un entier positif ou négatif quelconque ; si p est un entier négatif, il faut bien entendu que les déterminants $D(X_\lambda)$ ne soient pas nuls.

La règle d'une telle transformation d'une matrice $[X_1, X_2, \ldots, X_m]$ est exprimée à l'aide d'une matrice de même structure par la formule :

$$[Y_1, \ldots, Y_m][X_1, \ldots, X_m][Y_1, \ldots, Y_m]^{-1} =$$
$$= [Y_1X_1Y_1^{-1}, \ldots, Y_mX_mY_m^{-1}]. \quad (96)$$

Il faut noter la signification géométrique des transformations linéaires que l'on obtient à partir des matrices quasi diagonales. Pour simplifier l'étude, considérons le cas de la matrice quasi diagonale du septième ordre de structure $\{3, 2, 2\}$. On étudie la transformation linéaire correspondant à cette matrice. Si dans le vecteur initial $(x_1, \ldots, x_7)$ on a $x_4 = x_5 = x_6 = x_7 = 0$, dans le vecteur transformé on a évidemment : $y_4 = y_5 = y_6 = y_7 = 0$, c'est-à-dire que tous les vecteurs appartenant au sous-espace formé par les trois premiers vecteurs de base appartiennent encore à ce sous-espace après la transformation qui s'exprime elle-même au moyen de la matrice du troisième ordre B. Il en est de même pour le sous-espace formé par les deux vecteurs de base suivants et, enfin, pour le sous-espace formé par les deux derniers vecteurs de base.

On appelle *sous-espace engendré par les vecteurs* $\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(l)}$ l'ensemble des vecteurs définis par la formule :

$$c_1\mathbf{x}^{(1)} + \ldots + c_l\mathbf{x}^{(l)},$$

où $c_1, \ldots, c_l$ sont des constantes arbitraires.

II-1-8. Valeurs propres des matrices et réduction des matrices à la forme canonique. Les matrices semblables ne sont évidemment pas égales au sens de (76), mais elles sont équivalentes du point de vue géométrique car elles définissent la même transformation linéaire de l'espace, mais dans des systèmes de coordonnées différents. Cherchons maintenant les invariants de ces matrices, c'est-à-dire des expressions qui auraient une même valeur pour toutes les matrices semblables. Il n'est pas difficile de former un de ces invariants : le déterminant de la matrice. Soit A une matrice et UAU^{-1} une matrice semblable, U étant une matrice quelconque dont le déterminant n'est pas nul. On a d'après (80) et (92) :

$$D(UAU^{-1}) = D(U)\,D(A)\,D(U^{-1}) = D(U)\,D(A)\,D(U)^{-1} = D(A).$$

Pour obtenir d'autres invariants, formons un polynôme $\varphi(\lambda)$ de degré n par rapport au paramètre λ, égal au déterminant de la matrice, obtenue à partir de la matrice A en retranchant λ de tous les termes diagonaux, c'est-à-dire que nous supposons:

$$\varphi(\lambda)=\begin{vmatrix} a_{11}-\lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22}-\lambda & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn}-\lambda \end{vmatrix}, \tag{97}$$

où les a_{ik} sont les éléments de la matrice A. On peut l'écrire de la façon suivante:

$$\varphi(\lambda)=D(A-\lambda)=D(A-\lambda I), \tag{98}$$

puisque, par hypothèse, λ ou λI est une matrice diagonale dont tous les éléments de la diagonale principale sont égaux à λ. En portant UAU^{-1} au lieu de A et en tenant compte de ce que toute matrice commute avec le nombre λ et, par conséquent, $U\lambda U^{-1}=\lambda$, on obtient:

$$D(UAU^{-1}-\lambda)=D[U(A-\lambda)U^{-1}]=D(A-\lambda),$$

c'est-à-dire:

$$D(UAU^{-1}-\lambda)=D(A-\lambda). \tag{99}$$

On voit donc que le polynôme (97) formé pour la matrice UAU^{-1} coïncide avec le même polynôme formé pour la matrice A. Autrement dit, tous les coefficients du polynôme (97) sont des invariants pour des matrices semblables. Le coefficient du terme de plus haut degré de ce polynôme est égal à $(-1)^n$. Il faut noter deux coefficients particuliers de ce polynôme: le terme constant et le coefficient de $(-1)^{n-1}\lambda^{n-1}$. Le premier est égal au déterminant et on a déjà vu cet invariant plus haut. Quant au coefficient de $(-1)^{n-1}\lambda^{n-1}$, on voit en utilisant les résultats de [I-1-5] qu'il coïncide avec la somme des éléments diagonaux. Cette somme est la *trace de la matrice* et on la note de la façon suivante:

$$\mathrm{Tr}(A)=\{A\}_{11}+\{A\}_{22}+\dots+\{A\}_{nn}=a_{11}+a_{22}+\dots+a_{nn}.$$

Ainsi, les matrices semblables ont leur déterminant et leur trace identiques. Ecrivons maintenant l'équation:

$$D(A-\lambda)=0, \tag{100}$$

qui est l'*équation caractéristique de la matrice* A et ses racines sont les *valeurs caractéristiques ou valeurs propres de la matrice* A. D'après ce qui précède, on peut dire que des matrices semblables ont des valeurs propres identiques. On a déjà vu une équation de la forme (100).

Posons maintenant le problème suivant: peut-on trouver une matrice V qui, intervenant dans la transformation $A\to V^{-1}AV$, soit telle que la matrice obtenue soit diagonale? Autrement dit, peut-on, du point de vue des transformations linéaires dans l'espace, choisir des axes de coordonnées par rapport auxquels la transformation linéaire caractérisée dans le premier système de coordonnées par la matrice A se ramène simplement dans les nouveaux axes à une transformation de la forme $y_k = \lambda_k x_k$. Il faut remarquer qu'on écrit une matrice semblable de la forme $V^{-1}AV$ à la place de la matrice précédente de la forme UAU^{-1}, ce qui est sans importance.

Ecrivons la condition précédente sous la forme:

$$V^{-1}AV = [\lambda_1, \lambda_2, \ldots, \lambda_n], \tag{101}$$

on cherche à définir les éléments de la matrice V et les nombres λ_k. On peut exprimer cette condition, en multipliant à gauche les deux membres par V, sous la forme:

$$AV = V[\lambda_1, \lambda_2, \ldots, \lambda_n]. \tag{102}$$

Déterminons à partir de la formule (65) les éléments des deux membres de l'égalité écrite dont les indices sont i et k. On obtient ainsi n^2 équations:

$$\sum_{s=1}^{n} a_{is}v_{sk} = v_{ik}\lambda_k,$$

où les a_{ik} et les v_{ik} sont les éléments des matrices A et V.

En fixant le second indice k et en faisant $i = 1, 2, \ldots, n$, on obtient n équations ne contenant que les éléments $v_{1k}, \ldots, v_{nk}$ de la colonne d'indice k de la matrice V et le nombre λ_k:

$$\sum_{s=1}^{n} a_{is}v_{sk} = \lambda_k v_{ik} \quad (i = 1, 2, \ldots, n). \tag{103}$$

Si l'on considère les éléments $(v_{1k}, \ldots, v_{nk})$ comme les composantes d'un vecteur $\mathbf{v}^{(k)}$, on peut écrire l'égalité précédente sous forme de l'égalité vectorielle:

$$A\mathbf{v}^{(k)} = \lambda_k \mathbf{v}^{(k)}. \tag{104}$$

On voit ainsi qu'en cherchant la matrice V qui met la matrice A sous forme diagonale, on est conduit à chercher des vecteurs $\mathbf{v}^{(k)}$ qui, après application de la transformation linéaire définie par la matrice A, redonne ces vecteurs à un facteur numérique près; ceci est l'analogie algébrique du fait que dans la mécanique quantique moderne, la mécanique des matrices d'Heisenberg est équivalente à la mécanique ondulatoire de Schrödinger. Dans la première, on ramène une matrice (infinie) à la forme diagonale, alors qu'en mécanique ondulatoire, on cherche des vecteurs (dans un espace à une infinité de dimensions) tels qu'une transformation linéaire les redonne à un facteur numérique près. On a parlé ci-dessus d'analogie algébrique car en se limitant à un espace à un nombre fini de dimensions, on réduit ces problèmes à des problèmes de nature algébrique. Dans les cas plus compliqués d'espace à une infinité de dimensions, on sort du domaine de l'algèbre ordinaire et on a besoin de l'analyse mathématique. On élucidera ces questions par la suite et l'on remarquera que dans les applications à la physique du cas d'un nombre fini de dimensions, il suffit de s'en tenir aux matrices A d'un type particulier (matrices hermitiennes pour lesquelles $a_{kl} = \bar{a}_{lk}$) et la matrice U doit être aussi d'un type déterminé (matrice unitaire dont on donnera la définition ci-dessous). On étudie ici le problème général d'une matrice finie quelconque et l'on se borne à énoncer les résultats finals sans en donner la démonstration complète. On résoudra entièrement les problèmes intéressants dans les applications.

Passons à la résolution du système (103) ou du système (104). On peut l'écrire sous la forme développée suivante:

$$\left.\begin{array}{l} (a_{11}-\lambda_k)\,v_{1k}+a_{12}v_{2k}+\ldots+a_{1n}v_{nk}=0, \\ a_{21}v_{1k}+(a_{22}-\lambda_k)\,v_{2k}+\ldots+a_{2n}v_{nk}=0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{n1}v_{1k}+a_{n2}v_{2k}+\ldots+(a_{nn}-\lambda_k)\,v_{nk}=0. \end{array}\right\} \qquad (105)$$

Pour obtenir une solution non triviale pour $(v_{1k}, \ldots, v_{nk})$, il faut et il suffit que le déterminant de ce système soit nul, c'est-à-dire qu'il faut et il suffit que le nombre λ_k soit racine de l'équation caractéristique. Etudions en détail seulement le cas où cette équation a des racines distinctes. Soient $\lambda_1, \ldots, \lambda_n$ ces racines. En remplaçant λ_k dans les coefficients du système (105) par la première racine λ_1, on peut définir ainsi les éléments de la première colonne du tableau V, mais on ne cherche pas le nombre de valeurs de v_{l1}. Choisissons une solution quelconque du système, mais qui ne soit pas triviale. De même, en remplaçant λ_k par λ_2 dans les coefficients du système (105), on peut déterminer les éléments de la seconde colonne de la matrice V, etc., jusqu'à la n-ième colonne. Les égalités (105) sont équivalentes à (102) et pour passer à l'égalité principale (101), il suffit que V ait une matrice inverse V^{-1}, c'est-à-dire que le déterminant de V ne soit pas nul. Faisons la démonstration par l'absurde et supposons qu'il soit nul. Comme on sait [I-2-5], cela est équivalent à supposer qu'il existe entre les vecteurs $\mathbf{v}^{(k)}$ définis par les colonnes de la matrice V une relation linéaire:

$$C_1\mathbf{v}^{(1)}+\ldots+C_n\mathbf{v}^{(n)}=0,$$

où tous les coefficients C_k ne sont pas nuls. On applique aux deux membres de cette égalité $(n-1)$ fois la transformation définie par la matrice A. En utilisant (104), on obtient n égalités:

$$C_1\mathbf{v}^{(1)} + \ldots + C_n\mathbf{v}^{(n)} = 0,$$
$$\lambda_1 C_1\mathbf{v}^{(1)} + \ldots + \lambda_n C_n\mathbf{v}^{(n)} = 0,$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$\lambda_1^{n-1} C_1\mathbf{v}^{(1)} + \ldots + \lambda_n^{n-1} C_n\mathbf{v}^{(n)} = 0.$$

En tenant compte du fait que tous les vecteurs $C_k\mathbf{v}^{(k)}$ ne sont pas nuls, on peut affirmer que le déterminant de ce système doit être nul:

$$\begin{vmatrix} 1 & 1 & \ldots & 1 \\ \lambda_1 & \lambda_2 & \ldots & \lambda_n \\ \ldots & \ldots & \ldots & \ldots \\ \lambda_1^{n-1} & \lambda_2^{n-1} & \ldots & \lambda_n^{n-1} \end{vmatrix} = 0,$$

où les nombres λ_k sont différents par hypothèse. Mais la dernière égalité contredit le fait que le déterminant de Vandermonde de nombres inégaux n'est pas nul. On a ainsi démontré qu'il est possible de mettre une matrice sous la forme d'une matrice diagonale au moyen d'une similitude dans le cas où toutes les valeurs propres de la matrice sont différentes. Dans le cas où il y a des valeurs égales parmi les valeurs propres, il peut arriver qu'on ne puisse mettre la matrice sous la forme d'une matrice diagonale au moyen d'une similitude. Cependant il existe dans ce cas aussi la représentation la plus simple ou, comme on dit, canonique de la matrice. Cette représentation canonique, dans le cas où la matrice se réduit à une forme diagonale, est de la forme:

$$[\lambda_1, \lambda_2, \ldots, \lambda_n],$$

où les λ_k sont les valeurs propres de la matrice. On établit au moins un résultat dans le cas général *. Soit $\lambda = a$ une racine multiple d'ordre k de l'équation (100). Supposons que pour tous les déterminants d'ordre $(n - 1)$ du tableau situé dans le premier membre de l'équation (100), le nombre $\lambda = a$ soit une racine multiple d'ordre k_1, mais pas plus élevé, c'est-à-dire que tous ces déterminants sont divisibles par $(\lambda - a)^{k_1}$, mais l'un au moins n'est pas divisible par $(\lambda - a)^{k_1+1}$. Supposons alors que pour tous les déterminants de ce tableau d'ordre $(n - 2)$ $\lambda = a$ soit racine d'ordre au plus k_2, etc., et, enfin, que tous les déterminants d'ordre $(n - m)$ aient cette racine d'ordre k_m et qu'au moins un déterminant d'ordre $(n - m - 1)$ ne soit pas nul pour $\lambda = a$. Il en sera de même, évidemment, pour les

---

* On donnera sa démonstration en appendice spécial dans la seconde partie de ce tome.

déterminants d'ordre moins élevé. On peut démontrer que les nombres k_s décroissent : $k > k_1 > k_2 > \ldots > k_m$. On introduit les nombres entiers positifs suivants :

$$l_1 = k - k_1; \quad l_2 = k_1 - k_2; \ \ldots; \quad l_{m+1} = k_m,$$

et l'on a alors évidemment : $l_1 + l_2 + \ldots + l_{m+1} = k$. Les binômes

$$(\lambda - a)^{l_1}; \quad (\lambda - a)^{l_2}; \ \ldots; \ (\lambda - a)^{l_{m+1}}$$

sont des *diviseurs élémentaires de la matrice A* correspondant à la racine $\lambda = a$. On peut ainsi définir des diviseurs élémentaires pour toutes les valeurs propres de la matrice A, et l'on obtient alors un *ensemble de diviseurs élémentaires* :

$$(\lambda - \lambda_1)^{\rho_1}; \quad (\lambda - \lambda_2)^{\rho_2}; \ \ldots; \quad (\lambda - \lambda_p)^{\rho_p}. \tag{106}$$

où :

$$\rho_1 + \rho_2 + \ldots + \rho_p = n \tag{107}$$

et il peut y avoir des nombres identiques parmi les λ_k.

On a vu précédemment que les valeurs propres ne changent pas dans une similitude. L'ensemble des diviseurs élémentaires d'une matrice a la même propriété. Soient de nouvelles matrices simples $I_\rho(a)$. On entend par là une matrice d'ordre ρ dont la diagonale principale est constituée partout par le nombre a, la diagonale située sous celle-ci par l'unité, et dont les autres éléments sont tous nuls :

$$I_\rho(a) = \begin{vmatrix} a & 0 & 0 & \ldots & 0 & 0 \\ 1 & a & 0 & \ldots & 0 & 0 \\ 0 & 1 & a & \ldots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \ldots & a & 0 \\ 0 & 0 & 0 & \ldots & 1 & a \end{vmatrix}. \tag{108}$$

Le résultat suivant est essentiel pour la question de la représentation des matrices sous la forme canonique : si la matrice A possède des diviseurs élémentaires (106), il existe une matrice U, ayant un déterminant non nul, telle que :

$$UAU^{-1} = [I_{\rho_1}(\lambda_1), \ I_{\rho_2}(\lambda_2), \ \ldots, \ I_{\rho_p}(\lambda_p)]. \tag{109}$$

Il faut remarquer que si l'on connaît toutes les valeurs propres de la matrice A, on trouve la matrice U au moyen d'opérations algébriques élémentaires. Si $\rho = 1$, on entend par $I_\rho(a)$ simplement le nombre a. Il peut arriver même que lorsque certaines valeurs propres sont identiques, tous les diviseurs élémentaires (106) soient premiers, c'est-à-dire qu'ils soient de la forme:

$$(\lambda-\lambda_1);\ (\lambda-\lambda_2);\ \ldots;\ (\lambda-\lambda_n).$$

Dans ce cas, la matrice quasi diagonale

$$[I_{\rho_1}(\lambda_1),\ I_{\rho_2}(\lambda_2),\ \ldots,\ I_{\rho_p}(\lambda_p)]$$

devient simplement la matrice diagonale $[\lambda_1, \ldots, \lambda_n]$, et la matrice se met sous la forme diagonale.

Il est aisé de démontrer que pour mettre une matrice sous la forme diagonale, il faut et il suffit que le rang du tableau des coefficients dans le système (105) soit égal à $(n-\mu_k)$, où μ_k est l'ordre de multiplicité de la racine λ_k de l'équation séculaire. Lorsque cette condition est vérifiée, le système (105) définit μ_k vecteurs linéairement indépendants $(v_{1k}, v_{2k}, \ldots, v_{nk})$ [I-2-7].

Il faut noter que la matrice U dans la formule (109) n'est pas définie de façon unique. En particulier, si d est la valeur du déterminant de la matrice U, on peut remplacer dans la formule (109):

$$U \text{ par } \frac{1}{\sqrt[n]{d}}U \text{ et } U^{-1} \text{ par } \sqrt[n]{d}\,U,$$

c'est-à-dire qu'on peut considérer que le déterminant de la matrice U dans la formule (109) est égal à l'unité. On se bornera à ces remarques sur le problème général de la mise d'une matrice sous la forme canonique. Nous reviendrons sur cette question dans un appendice spécial inséré dans la seconde partie du troisième tome. Comme nous l'avons déjà mentionné plus haut, par la suite, nous considérerons en détail ce problème pour les matrices de type spécial.

II-1-9. Transformations unitaires et transformations orthogonales. On utilise, dans ce paragraphe et dans les suivants, les notions de produit scalaire et de norme (longueur) d'un vecteur; ces notions ont été déjà considérées [I-2-6]. Rappelons que le carré de la norme (longueur) est défini par la formule:

$$\|\mathbf{x}\|^2=(\mathbf{x},\ \mathbf{x})=\sum_{s=1}^{n}|x_s|^2, \tag{110}$$

ou, dans le cas de composantes réelles,

$$\|\mathbf{x}\|^2=\sum_{s=1}^{n}x_s^2.$$

Cette définition de la norme est liée au choix des vecteurs de base, c'est-à-dire des axes de coordonnées. Nous appellerons un système de coordonnées, auquel on adjoint la définition de la norme donnée ci-dessus, système normal ou cartésien. Outre la longueur d'un vecteur, on définit le *produit scalaire de deux vecteurs* à l'aide de la formule suivante :

$$(\mathrm{x},\ \mathrm{y}) = x_1\overline{y}_1 + x_2\overline{y}_2 + \ldots + x_n\overline{y}_n. \tag{111}$$

Dans le cas de vecteurs réels, cette formule a une forme plus symétrique :

$$(\mathrm{x},\ \mathrm{y}) = x_1y_1 + x_2y_2 + \ldots + x_ny_n.$$

Il résulte de (111) que lorsqu'on change l'ordre des vecteurs, la valeur d'un produit scalaire se transforme en son conjugué complexe, c'est-à-dire que :

$$(\mathrm{y},\ \mathrm{x}) = \overline{(\mathrm{x},\ \mathrm{y})}. \tag{112}$$

On dit que deux vecteurs sont *perpendiculaires* ou *orthogonaux* si leur produit scalaire est nul.

Par la suite, on considérera toujours, à moins qu'on ne précise le contraire, qu'il s'agit de systèmes de coordonnées cartésiennes. Les transformations linéaires correspondant au passage d'un système à un autre prennent alors une valeur particulière. On sait qu'à tout passage d'un système de coordonnées à un autre correspond une transformation linéaire des composantes. Soit une transformation de ce type :

$$(y_1,\ \ldots,\ y_n) = U(x_1,\ \ldots,\ x_n), \tag{113}$$

où le premier système de coordonnées est cartésien. Pour que le nouveau système soit cartésien, il faut et il suffit que la longueur du vecteur soit exprimée dans le nouveau système également par la somme des carrés des modules des composantes, c'est-à-dire :

$$|y_1|^2 + \ldots + |y_n|^2 = |x_1|^2 + \ldots + |x_n|^2. \tag{114}$$

Montrons que la valeur du produit scalaire est alors exprimée dans le nouveau système de coordonnées par une formule analogue à la formule (111). Supposons, en effet, que l'on ait dans le premier système de coordonnées deux vecteurs :

$$\mathrm{x}(x_1,\ \ldots,\ x_n) \text{ et } \mathrm{x}'(x_1',\ \ldots,\ x_n'),$$

les vecteurs suivants leur correspondent dans le nouveau système :

$$\mathrm{y}(y_1,\ \ldots,\ y_n) \quad \text{et} \quad \mathrm{y}'(y_1',\ \ldots,\ y_n').$$

Construisons deux nouveaux vecteurs $\mathrm{z} = \mathrm{x} + \mathrm{x}'$ et $\mathrm{u} = \mathrm{x} + i\mathrm{x}'$ qui aient pour composantes $(x_k + x_k')$ et $(x_k + ix_k')$. En considérant que la condition (114) est vérifiée, on a :

$$\sum_{k=1}^{n} (y_k + y_k')(\overline{y}_k - \overline{y}_k') = \sum_{k=1}^{n} (x_k + x_k')(\overline{x}_k + \overline{x}_k'),$$

d'où l'on obtient finalement d'après (114):

$$\sum_{k=1}^{n} (y_k\bar{y}'_k + y'_k\bar{y}_k) = \sum_{k=1}^{n} (x_k\bar{x}'_k + x'_k\bar{x}_k), \tag{115$_1$}$$

puisque:

$$\sum_{k=1}^{n} |y_k|^2 = \sum_{k=1}^{n} |x_k|^2 \text{ et } \sum_{k=1}^{n} |y'_k|^2 = \sum_{k=1}^{n} |x'_k|^2.$$

De même:

$$\sum_{k=1}^{n} (y_k + iy'_k)(\bar{y}_k - i\bar{y}'_k) = \sum_{k=1}^{n} (x_k + ix'_k)(\bar{x}_k - i\bar{x}'_k),$$

et de là:

$$\sum_{k=1}^{n} (y'_k\bar{y}_k - y_k\bar{y}'_k) = \sum_{k=1}^{n} (x'_k\bar{x}_k - x_k\bar{x}'_k). \tag{115$_2$}$$

Les égalités (115$_1$) et (115$_2$) donnent:

$$\sum_{k=1}^{n} y_k\bar{y}'_k = \sum_{k=1}^{n} x_k\bar{x}'_k, \tag{116}$$

c'est-à-dire que le produit scalaire s'exprime effectivement au moyen de la formule précédente. Ainsi, si la transformation (113) vérifie la condition (114), elle vérifie aussi la condition (116), c'est-à-dire qu'elle ne change pas la valeur du produit scalaire. Réciproquement (114) résulte de la condition (116) si l'on suppose que $x'_k = x_k$ dans (116), puisque le produit scalaire de deux vecteurs identiques est, évidemment, le carré de la longueur du vecteur. On dit habituellement que les transformations linéaires qui vérifient la condition (114) ou la condition (116) sont les *transformations unitaires*.

Si l'on considère un espace réel et des matrices réelles de transformations linéaires, la condition (114) se réduit tout simplement à la condition:

$$y_1^2 + y_2^2 + \ldots + y_n^2 = x_1^2 + x_2^2 + \ldots + x_n^2, \tag{117}$$

et les transformations réelles correspondantes sont des transformations *orthogonales*. Ce sont évidemment des cas particuliers des transformations unitaires.

Etudions maintenant les propriétés fondamentales des transformations unitaires. Ecrivons la condition (114) sous forme explicite pour la transformation (113), en désignant par u_{ik} les éléments de la matrice U:

$$\sum_{k=1}^{n} |u_{k1}x_1 + \ldots + u_{kn}x_n|^2 = \sum_{k=1}^{n} |x_k|^2$$

ou

$$\sum_{k=1}^{n}(u_{k1}x_1+\ldots+u_{kn}x_n)(\overline{u}_{k1}\overline{x}_1+\ldots+\overline{u}_{kn}\overline{x}_n)=\sum_{k=1}^{n}x_k\overline{x}_k. \quad (118)$$

On ouvre les parenthèses dans le premier membre de la formule et l'on identifie à l'unité les coefficients de $x_p\overline{x}_p$, et ceux de $x_p\overline{x}_q$ $(p\neq q)$ à zéro. On a alors la condition nécessaire et suffisante pour les éléments d'une transformation unitaire sous la forme :

$$\left.\begin{array}{l}\sum_{k=1}^{n}|u_{kp}|^2=1 \quad (p=1,\,2\,\ldots,\,n),\\ \sum_{k=1}^{n}u_{kp}\overline{u}_{kq}=0 \quad (p\neq q),\end{array}\right\} \quad (119)$$

c'est-à-dire que la somme des carrés des modules des éléments de chaque colonne doit être égale à l'unité, et la somme des produits des éléments d'une colonne par les quantités conjuguées complexes des éléments correspondants d'une autre colonne doit être nulle. On écrit parfois ces conditions sous la forme :

$$\sum_{k=1}^{n}u_{kp}\overline{u}_{kq}=\delta_{pq}, \quad (120)$$

où les δ_{pq} sont les éléments de la matrice unité, c'est-à-dire :

$$\delta_{pq}=\left\{\begin{array}{l}0\ (p\neq q)\\ 1\ (p=q)\end{array}\right. . \quad (121)$$

On a appliqué ci-dessus la méthode des coefficients indéterminés pour l'identité (118), ce qui est suffisant pour que l'identité soit vérifiée. On démontre facilement que, en donnant à x_k des valeurs particulières, l'identité des coefficients pour des termes semblables est une condition nécessaire.

On prend un déterminant $D(A)$ et un autre déterminant $D(\overline{A})$ composé des éléments conjugués complexes. En les multipliant colonne par colonne [I-1-6], on obtient d'après (119) le déterminant de la matrice unité, c'est-à-dire l'unité. D'autre part, il est évident que les deux déterminants s'expriment au moyen de nombres complexes conjugués, et il résulte directement de ce qu'on vient de dire que :

$$|D(A)|^2=1,$$

c'est-à-dire que *le carré du module du déterminant de la matrice unitaire est égal à l'unité.* Autrement dit, le déterminant de la matrice unitaire a le module égal à l'unité, c'est-à-dire que c'est un nombre complexe $e^{i\varphi}$, où φ est réel.

Introduisons la matrice $U^{(*)}$ transposée de U. Les conditions (119) qui sont les conditions d'orthogonalité suivant les colonnes, peuvent être écrites sous forme d'égalité matricielle :

$$\overline{U}^{(*)}U = I, \tag{122}$$

ce qui est équivalent à :

$$U^{-1} = \overline{U}^{(*)} = \tilde{U}, \tag{123}$$

c'est-à-dire que si une matrice est unitaire, sa matrice inverse coïncide avec la matrice conjuguée hermitienne.

La transformation U^{-1}, inverse de U, définit le passage du vecteur y au vecteur x. Elle vérifie évidemment aussi la condition (114), c'est-à-dire que si U est une matrice unitaire, la matrice inverse U^{-1} est aussi une matrice unitaire. Autrement dit, d'après (123), la matrice $\tilde{U}$ est une matrice unitaire et ses colonnes vérifient la condition d'orthogonalité. Mais les colonnes de $\tilde{U}$ sont les lignes de $\overline{U}$. On peut donc affirmer que dans une matrice unitaire, non seulement les colonnes mais aussi les lignes vérifient les conditions d'orthogonalité, c'est-à-dire que l'on a, à côté de (120), les formules :

$$\sum_{k=1}^{n} u_{pk}\overline{u}_{qk} = \delta_{pq}. \tag{124}$$

De la même façon, si des matrices U_1 et U_2 vérifient la condition (114), leur produit U_2U_1 vérifie évidemment aussi cette condition, c'est-à-dire que le produit de deux matrices unitaires est aussi une matrice unitaire.

On peut mettre la définition de la matrice unitaire sous deux formes distinctes :

$$|U\mathbf{x}|^2 = |\mathbf{x}|^2 \quad \text{ou} \quad (U\mathbf{x}, U\mathbf{x}') = (\mathbf{x}, \mathbf{x}'), \tag{125$_1$}$$

x et x′ étant dans la deuxième égalité des vecteurs quelconques.

Etudions à présent le cas où une matrice unitaire a des éléments réels. On dit alors que c'est une matrice orthogonale et que la transformation correspondante est orthogonale. On a dans ce cas, à la place des formules (120) et (124), les formules :

$$\sum_{k=1}^{n} u_{kp}u_{kq} = \delta_{pq}; \quad \sum_{k=1}^{n} u_{pk}u_{qk} = \delta_{pq}. \tag{125$_2$}$$

De plus, le déterminant de la transformation doit être, évidemment, un nombre réel et c'est pourquoi il ne peut être égal qu'à ± 1. Ces transformations orthogonales réelles dans un espace à n dimensions sont tout à fait analogues aux transformations dans un espace à trois dimensions [II-1-4]. En outre, dans ce cas réel, $\tilde{U}$ coïncide avec $U^{(*)}$, c'est-à-dire qu'on obtient la transformation inverse U^{-1} à partir de U en remplaçant les lignes par les colonnes.

Il faut remarquer encore que tout nombre complexe $e^{i\varphi}$, où φ est réel, que l'on considère comme une matrice $[e^{i\varphi}, e^{i\varphi}, \ldots, e^{i\varphi}]$, est une matrice unitaire, et si U est une matrice unitaire, le produit $e^{i\varphi}U$ est aussi une matrice unitaire. On a montré ci-dessus le sens du produit d'un nombre par une matrice [II-1-6].

II-1-10. Inégalité de Cauchy-Bouniakovsky. Etablissons dans ce paragraphe une inégalité qu'on utilisera par la suite. Soient des nombres réels quelconques $\alpha_1, \alpha_2, \ldots, \alpha_m$ et $\beta_1, \beta_2, \ldots, \beta_m$, on a :

$$\left(\sum_{k=1}^{m} \alpha_k\beta_k\right)^2 \leqslant \sum_{k=1}^{m} \alpha_k^2 \cdot \sum_{k=1}^{m} \beta_k^2, \tag{126}$$

l'égalité ayant lieu si et seulement si α_k et β_k sont proportionnels :

$$\frac{\beta_1}{\alpha_1} = \frac{\beta_2}{\alpha_2} = \ldots = \frac{\beta_m}{\alpha_m}. \tag{127}$$

Soit ξ un nombre réel quelconque. On forme la somme :

$$S = \sum_{k=1}^{m} (\xi\alpha_k - \beta_k)^2,$$

qui n'est évidemment pas négative. L'égalité a lieu si et seulement si :

$$\frac{\beta_1}{\alpha_1} = \frac{\beta_2}{\alpha_2} = \ldots = \frac{\beta_m}{\alpha_m} = \xi,$$

et dans ce cas, on a évidemment :

$$\left(\sum_{k=1}^{m} \alpha_k\beta_k\right)^2 = \sum_{k=1}^{m} \alpha_k^2 \cdot \sum_{k=1}^{m} \beta_k^2.$$

En général, en ouvrant les parenthèses dans l'expression de S, on obtient un trinôme du second degré :

$$S = A\xi^2 - 2B\xi + C,$$

où

$$A = \sum_{k=1}^{m} \alpha_k^2 ; \quad B = \sum_{k=1}^{m} \alpha_k\beta_k ; \quad C = \sum_{k=1}^{m} \beta_k^2.$$

Ce trinôme reste non négatif pour tous les ξ réels, d'où il résulte que $AC - B^2 \geqslant 0$, c'est-à-dire que $B^2 \leqslant AC$, soit l'inégalité (126).

Si $AC - B^2 = 0$, le trinôme s'annule pour un certain ξ réel et l'on a déjà vu que la condition (127) est alors vérifiée. Réciproquement, lorsque cette condition est vérifiée, on a l'égalité dans (126). Supposons à présent que α_k et β_k sont des nombres complexes. En tenant compte de ce que :

$$\left|\sum_{k=1}^{m} \alpha_k\beta_k\right| \leqslant \sum_{k=1}^{m} |\alpha_k|\,|\beta_k|,$$

en appliquant l'inégalité (126) à la dernière somme, constituée de termes positifs, on obtient :

$$\left|\sum_{k=1}^{m} \alpha_k\beta_k\right|^2 \leqslant \sum_{k=1}^{m} |\alpha_k|^2 \cdot \sum_{k=1}^{m} |\beta_k|^2. \qquad (126_1)$$

Il est facile de montrer que dans ce cas, l'égalité a lieu pour α_k et β_k conjugués complexes, si et seulement si $|\alpha_k|$ et $|\beta_k|$ sont proportionnels et que tous les produits $\alpha_k\beta_k$ aient un argument identique. Comme on l'a déjà vu (tome II, [VI-1-8]), l'inégalité (126) peut être appliquée non seulement à des sommes, mais aussi à des intégrales. Si $f_1(x)$ et $f_2(x)$ sont deux fonctions réelles dans l'intervalle $a \leqslant x \leqslant b$, l'inégalité pour les intégrales est de la forme :

$$\left[\int_a^b f_1(x)\, f_2(x)\, dx\right]^2 \leqslant \int_a^b f_1^2(x)\, dx \int_a^b f_2^2(x)\, dx \qquad (126_2)$$

En effet, on écrit l'expression :

$$\int_a^b [\xi f_1(x) - f_2(x)]^2\, dx = \xi^2 \int_a^b f_1^2(x)\, dx - 2\xi \int_a^b f_1(x)\, f_2(x)\, dx + \\ + \int_a^b f_2^2(x)\, dx,$$

où ξ est un nombre réel quelconque. Il résulte de la forme du premier membre que cette expression ne peut être négative pour aucun ξ réel. Mais si le trinôme $A\xi^2 - 2\xi B + C$ n'est pas négatif pour tous les ξ réels, $AC - B^2 \geqslant 0$. Si l'on applique ce résultat au trinôme précédent, on obtient l'inégalité (126_2). Cette inégalité a été démontrée pour les intégrales par Bouniakovsky et pour les sommes par Cauchy.

II-1-11. Propriétés du produit scalaire et de la norme. On étudie à présent certaines propriétés des produits scalaires et des normes. En appliquant l'inégalité (126_1) et en tenant compte de ce que $|\bar{y}_k| = |y_k|$, on peut écrire :

$$|(\mathbf{x}, \mathbf{y})|^2 = \left|\sum_{k=1}^{n} x_k\bar{y}_k\right|^2 \leqslant \sum_{k=1}^{n} |x_k|^2 \cdot \sum_{k=1}^{n} |y_k|^2,$$

c'est-à-dire:

$$|(\mathbf{x}, \mathbf{y})| \leqslant \|\mathbf{x}\| \cdot \|\mathbf{y}\|. \tag{128}$$

Démontrons l'inégalité du triangle:

$$\|\mathbf{x}+\mathbf{y}\| \leqslant \|\mathbf{x}\| + \|\mathbf{y}\|. \tag{129}$$

On a:

$$\|\mathbf{x}+\mathbf{y}\|^2 = (\mathbf{x}+\mathbf{y},\ \mathbf{x}+\mathbf{y}) = (\mathbf{x}, \mathbf{x}) + (\mathbf{y}, \mathbf{y}) + (\mathbf{x}, \mathbf{y}) + (\mathbf{y}, \mathbf{x}),$$

ou, en tenant compte de (128), on obtient:

$$\|\mathbf{x}+\mathbf{y}\|^2 \leqslant \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 + 2\|\mathbf{x}\| \cdot \|\mathbf{y}\| = (\|\mathbf{x}\| + \|\mathbf{y}\|)^2,$$

d'où il résulte (129).

Etudions enfin l'influence du choix d'un système de coordonnées sur la métrique de l'espace, c'est-à-dire sur l'expression du carré de la longueur d'un vecteur. Supposons qu'à la place d'un système cartésien on prenne un nouveau système de coordonnées et prenons pour vecteurs de base des vecteurs indépendants:

$$\mathbf{z}^{(1)}, \mathbf{z}^{(2)}, \ldots, \mathbf{z}^{(n)}.$$

On a pour tout vecteur:

$$\mathbf{x} = z_1\mathbf{z}^{(1)} + \ldots + z_n\mathbf{z}^{(n)},$$

où les z_k sont les composantes du vecteur dans le nouveau système de coordonnées.

Le carré de la longueur de ce vecteur s'exprime au moyen du produit scalaire du vecteur par lui-même, c'est-à-dire:

$$|\mathbf{x}|^2 = (z_1\mathbf{z}^{(1)} + \ldots + z_n\mathbf{z}^{(n)},\ z_1\mathbf{z}^{(1)} + \ldots + z_n\mathbf{z}^{(n)}).$$

En ouvrant les parenthèses d'après les formules vues ci-dessus, on a l'expression suivante pour le carré de la longueur du vecteur:

$$|\mathbf{x}|^2 = \sum_{i,\,k=1}^{n} \alpha_{ik} z_i \bar{z}_k, \tag{130}$$

où les coefficients α_{ik} sont définis par les formules:

$$\alpha_{ik} = (\mathbf{z}^{(i)}, \mathbf{z}^{(k)}).$$

Quand on intervertit les indices, on obtient le conjugué complexe, c'est-à-dire:

$$\alpha_{ki} = \bar{\alpha}_{ik}. \tag{131}$$

Une somme de la forme (130) ayant des coefficients qui vérifient la condition (131) est une *forme d'Hermite*. Il est évident que toute expression de la forme (130) avec la condition (131) n'aura pour tous les z_h complexes possibles que des valeurs réelles, puisque pour $i \neq k$ deux termes de la somme (130) sont conjugués complexes, et les coefficients α_{kk} seront réels dans les termes de la forme $\alpha_{kk}|z_k|^2$ d'après la condition (131). De plus, étant donné la construction de la forme hermitienne (130), on peut affirmer que la somme (130) sera non négative et s'annulera seulement si tous les z_k sont nuls. La formule (130) définit la métrique de l'espace dans le nouveau système de coordonnées.

La métrique (130) coïncide avec la métrique (110) dans le système cartésien correspondant si $\alpha_{ik} = 0$ pour $i \neq k$ et $\alpha_{kk} = 1$, ou $(\mathbf{z}^{(i)}, \mathbf{z}^{(k)}) = 0$ pour $i \neq k$ et $(\mathbf{z}^{(k)}, \mathbf{z}^{(k)}) = 1$, autrement dit, si les vecteurs $\mathbf{z}^{(k)}$, pris comme vecteurs de base, sont des vecteurs unités orthogonaux (dont la longueur est égale à l'unité).

Dans la suite tout système de vecteurs orthogonaux et unités $\mathbf{z}^{(k)}$ $(k = 1, 2, \ldots, l)$ sera dit système *orthonormé*.

Notons encore que si la formule (113) définit une transformation unitaire pour les composantes d'un vecteur, alors la transformation correspondante qui permet de passer des anciens vecteurs de base aux nouveaux est donnée par le tableau:

$$U^{(*)-1},$$

contragrédient de U. Dans ce cas, d'après (123), ce tableau coïncide avec le tableau associé à $\overline{U}$ et pour les transformations réelles orthogonales il coïncide tout simplement avec U.

II-1-12. Orthogonalisation des vecteurs. Soient m vecteurs quelconques linéairement indépendants $\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(m)}$. L'ensemble des vecteurs de la forme:

$$C_1\mathbf{x}^{(1)} + \ldots + C_m\mathbf{x}^{(m)},$$

où les C_k sont des coefficients arbitraires, définit tout l'espace, si $m = n$, et un certain sous-espace R_m de dimension m, si $m < n$. Montrons qu'on peut toujours construire un système orthonormé de vecteurs $\mathbf{z}^{(k)}$ $(k = 1, 2, \ldots, m)$ qui forme le même sous-espace R_m que les vecteurs $\mathbf{x}^{(k)}$. Autrement dit, les $\mathbf{z}^{(k)}$ doivent s'exprimer linéairement en fonction des $\mathbf{x}^{(k)}$ et, réciproquement, les $\mathbf{x}^{(k)}$ doivent s'exprimer en fonction des $\mathbf{z}^{(k)}$. On peut construire ces vecteurs de la manière suivante:

$$\left.\begin{aligned} &\mathbf{y}^{(1)} = \mathbf{x}^{(1)},\\ &\mathbf{y}^{(2)} = \mathbf{x}^{(2)} - (\mathbf{x}^{(2)}, \mathbf{z}^{(1)})\,\mathbf{z}^{(1)},\\ &\mathbf{y}^{(3)} = \mathbf{x}^{(3)} - (\mathbf{x}^{(3)}, \mathbf{z}^{(1)})\,\mathbf{z}^{(1)} - (\mathbf{x}^{(3)}, \mathbf{z}^{(2)})\,\mathbf{z}^{(2)},\\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \end{aligned}\right\} \qquad (132)$$

où:

$$\mathbf{z}^{(1)} = \frac{\mathbf{y}^{(1)}}{|\mathbf{y}^{(1)}|}; \quad \mathbf{z}^{(2)} = \frac{\mathbf{y}^{(2)}}{|\mathbf{y}^{(2)}|}; \ldots; \quad \mathbf{z}^{(m)} = \frac{\mathbf{y}^{(m)}}{|\mathbf{y}^{(m)}|}. \qquad (133)$$

On obtient le vecteur $z^{(1)}$ à partir de $y^{(1)}$ en le divisant par la longueur de $y^{(1)}$ et, par conséquent, la longueur de $z^{(1)}$ est égale à l'unité. Ensuite, le vecteur $y^{(2)}$ est construit suivant la formule indiquée ci-dessus. Il résulte de sa définition qu'il est orthogonal à $z^{(1)}$:

$$(y^{(2)}, z^{(1)}) = (x^{(2)}, z^{(1)}) - (x^{(2)}, z^{(1)})\ (z^{(1)}, z^{(1)}) = 0.$$

En divisant le vecteur $y^{(2)}$ par sa longueur, on obtient le vecteur $z^{(2)}$. On construit ensuite le vecteur $y^{(3)}$ suivant la formule indiquée ci-dessus. Il résulte de cette formule que le vecteur est orthogonal à $z^{(1)}$ et $z^{(2)}$.

En effet, d'après l'orthogonalité de $z^{(1)}$ et de $z^{(2)}$, on obtient par exemple:

$$(y^{(3)}, z^{(2)}) = (x^{(3)}, z^{(2)}) - (x^{(3)}, z^{(2)})\ (z^{(2)}, z^{(2)}) = 0.$$

En divisant le vecteur $y^{(3)}$ par sa longueur, on obtient le vecteur $z^{(3)}$, etc.

Tous les vecteurs ainsi construits s'expriment linéairement en fonction des $x^{(k)}$. Réciproquement, les $x^{(k)}$ s'expriment en fonction des $z^{(k)}$. Pour cela il suffit de résoudre progressivement les égalités précédentes par rapport à $x^{(1)}$, $x^{(2)}$, etc.

Il faut également remarquer qu'aucun des vecteurs $y^{(k)}$ ainsi construits ne devient nul. En effet, si à une certaine étape du calcul on avait obtenu un vecteur $y^{(k)}$ nul, étant donné qu'il s'exprime linéairement en fonction des $x^{(s)}$ et que le coefficient de $x^{(k)}$, dans cette expression linéaire, est égal à l'unité, les vecteurs $x^{(s)}$ auraient été linéairement dépendants entre eux, ce qui est contraire à l'hypothèse d'après laquelle ces vecteurs sont linéairement indépendants.

Soulignons que si nous avons un ensemble de vecteurs deux à deux orthogonaux, non nuls, ces vecteurs sont alors linéairement indépendants.

Si $m = n$, les $z^{(k)}$ donnent un système orthonormé complet. Si $m < n$, il faut, pour obtenir un système complet de coordonnées cartésiennes, que l'on construise, en plus des vecteurs $z^{(k)}$, $(n - m)$ vecteurs orthogonaux entre eux et aux vecteurs $z^{(k)}$. Ces nouveaux vecteurs unités doivent donc former un sous-espace R'_{n-m} de dimension $(n - m)$, orthogonal au sous-espace R_m [I-2-5]. Les nouveaux vecteurs cherchés u doivent vérifier le système d'équations:

$$(u, x^{(1)}) = 0, \ \ldots, (u, x^{(m)}) = 0.$$

On a un système de m équations homogènes à n inconnues, le rang de ce système étant égal à m car les vecteurs $x^{(k)}$ sont linéairement indépendants [I-2-5]. Ce système a $(n - m)$ solutions linéairement indépendantes, c'est-à-dire qu'on obtient $(n - m)$ vecteurs linéairement indépendants. En appliquant à ces vecteurs le procédé d'orthogonalisation indiqué plus haut et en leur donnant une longueur unité, on obtient avec $z^{(1)}, \ldots, z^{(m)}$ un système orthonormé complet. Le sous-espace R_m engendré par le système orthonormé de vecteurs $z^{(k)}$ peut être formé dans un autre système semblable. En effet, il suffit pour cela d'appliquer au système de vecteurs $z^{(k)}$ une transformation unitaire. On voit ainsi que l'orthogonalisation d'un système de vecteurs peut se faire de différentes façons. L'exemple donné ci-dessus en est une.

II-2. Formes quadratiques

II-2-1. Transformation des formes quadratiques en somme de carrés. Considérons dans un espace une surface du deuxième ordre dont le centre coïncide avec l'origine des coordonnées:

$$Ax^2+By^2+Cz^2+2Dxy+2Exz+2Fyz+G=0.$$

On peut toujours choisir de nouveaux axes de coordonnées (x', y', z') de telle sorte que l'équation transformée ne comprenne que les termes contenant les carrés des coordonnées, c'est-à-dire que l'équation transformée soit de la forme:

$$\lambda_1 x'^2+\lambda_2 y'^2+\lambda_3 z'^2+G=0.$$

Le problème se ramène à la détermination d'une transformation orthogonale, entre les variables (x', y', z') et (x, y, z), telle que, dans le premier membre de l'équation, l'ensemble des termes du second degré, par rapport aux coordonnées, se réduise à une somme de carrés. Posons le problème analogue dans le cas d'un espace réel à n dimensions. Soit donnée une forme quadratique réelle de n variables:

$$\varphi(x_1, \ldots, x_n)=\sum_{i,\,k=1}^{n} a_{ik}x_ix_k, \tag{134}$$

les a_{ik} étant des coefficients réels vérifiant la condition:

$$a_{ki}=a_{ik}. \tag{135}$$

Nous avons pour l'exemple précédent: $x=x_1$, $y=x_2$, $z=x_3$ et $a_{11}=A$; $a_{22}=B$; $a_{33}=C$; $a_{12}=a_{21}=D$; $a_{13}=a_{31}=E$; $a_{23}=a_{32}=F$.

Nous appelons *matrice de la forme quadratique* (134) la matrice d'éléments a_{ik}. Cette matrice est symétrique, c'est-à-dire qu'elle coïncide avec sa transposée.

Supposons que nous ayons transformé la forme (134) en introduisant à la place des x_k de nouvelles variables x'_k, et que cette transformation s'écrive sous la forme:

$$(x_1, \ldots, x_n)=B(x'_1, \ldots, x'_n), \tag{136}$$

où B est la matrice d'éléments b_{ik}. Portant l'expression (136) dans la formule (134), nous avons :

$$\varphi = \sum_{i,k=1}^{n} a_{ik} (b_{i1}x'_1 + \ldots + b_{in}x'_n)(b_{k1}x'_1 + \ldots + b_{kn}x'_n).$$

En ouvrant les parenthèses, nous trouvons pour coefficient de $x'_p x'_q$ $(p \neq q)$:

$$\sum_{i,k=1}^{n} a_{ik} (b_{ip}b_{kq} + b_{ik}b_{kp}).$$

Tenant compte de (135), on voit facilement que la moitié de l'expression ci-dessus est simplement égale à la somme :

$$\sum_{i=1}^{n} b_{ip} \sum_{k=1}^{n} a_{ik}b_{kq}.$$

En considérant d'une façon analogue le cas $p = q$, nous voyons que, dans les nouvelles coordonnées, la forme quadratique est:

$$\varphi = \sum_{i,k=1}^{n} c_{ik}x'_i x'_k, \tag{137}$$

où

$$c_{ik} = c_{ki} = \sum_{i=1}^{n} b_{ti} \sum_{s=1}^{n} a_{ts}b_{sk}.$$

La sommation sur s donne $\{AB\}_{tk}$. Si dans b_{ti} nous désignons par t le numéro de la colonne et par i celui de la ligne, b_{ti} est alors un élément de la matrice transposée $\{B^{(*)}\}_{it}$, d'où :

$$c_{ik} = c_{ki} = \sum_{t=1}^{n} \{B^{(*)}\}_{it} \{AB\}_{tk},$$

c'est-à-dire que la forme (137) a pour matrice celle déterminée à partir de la matrice A de la forme des anciennes coordonnées et de la matrice B de la transformation (136), de la façon suivante :

$$C = B^{(*)}AB. \tag{138}$$

Si la transformation (136) est orthogonale, la transposée $B^{(*)}$ de la matrice orthogonale B coïncide alors avec la matrice inverse B^{-1}, et dans ce cas la formule (138) est remplacée par la formule :

$$C = B^{-1}AB. \tag{139}$$

Ainsi, le problème de la formation d'une transformation orthogonale (136), réduisant la forme quadratique (134) à une somme de carrés, est équivalent à celui de la construction d'une matrice orthogonale B, telle que la matrice (139) soit une matrice diagonale $[\lambda_1, \ldots, \lambda_n]$, car une matrice de cette forme, qui se réduit à une somme de carrés, est précisément une matrice diagonale, de plus, ses éléments λ_k ne sont autres que les coefficients des carrés $x_k'^2$. Ainsi, nous avons donc, comme précédemment :

$$B^{-1}AB = [\lambda_1, \ldots, \lambda_n]$$

ou

$$AB = B[\lambda_1, \ldots, \lambda_n]. \tag{140}$$

Remarquons que dans ce cas, la matrice A n'est pas une matrice arbitraire, mais une matrice symétrique réelle et B doit être une matrice orthogonale. Nous allons procéder de la même façon que dans [II-1-8] lorsque nous y avons considéré le cas général. Écrivons l'équation (140) sous la forme :

$$\sum_{s=1}^{n} a_{is}b_{sk} = \lambda_k b_{ik}. \tag{141}$$

Nous avons n équations pour les éléments de la k-ième colonne de la matrice B. En introduisant le vecteur $\mathbf{x}^{(k)}$ de composantes $(b_{1k}, \ldots, b_{nk})$, nous pouvons écrire la dernière équation de la façon suivante :

$$A\mathbf{x}^{(k)} = \lambda_k \mathbf{x}^{(k)}. \tag{142}$$

En faisant passer le second membre de (141) dans le premier, nous en déduisons un système de n équations homogènes pour déterminer $b_{1k}, \ldots, b_{nk}$:

$$\left.\begin{array}{l}
(a_{11}-\lambda_k)\,b_{1k} + a_{12}b_{2k} + \ldots + a_{1n}b_{nk} = 0, \\
a_{21}b_{1k} + (a_{22}-\lambda_k)\,b_{2k} + \ldots + a_{2n}b_{nk} = 0, \\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
a_{n1}b_{1k} + a_{n2}b_{2k} + \ldots + (a_{nn}-\lambda_k)\,b_{nk} = 0.
\end{array}\right| \tag{143}$$

Le déterminant de ce système doit être nul et nous trouvons l'équation algébrique de degré n pour les nombres λ_k :

$$\begin{vmatrix}
a_{11}-\lambda & a_{12} & \ldots & a_{1n} \\
a_{21} & a_{22}-\lambda & \ldots & a_{2n} \\
\ldots & \ldots & \ldots & \ldots \\
a_{n1} & a_{n2} & \ldots & a_{nn}-\lambda
\end{vmatrix} = 0. \tag{144}$$

Comme nous le savons, c'est précisément l'équation caractéristique de la matrice A.

Démontrons d'abord que pour une matrice symétrique réelle A l'équation (144) a toutes ses racines réelles. Au préalable donnons un mode nouveau d'écriture des formes quadratiques. Soit $\mathbf{x}$ un vecteur de composantes $(x_1, \ldots, x_n)$, réelles ou complexes, et soit A une matrice d'éléments a_{ik} arbitraires. Formons le produit scalaire:

$$(A\mathbf{x}, \mathbf{x}) = \sum_{i=1}^{n} \bar{x}_i (a_{i1}x_1 + \ldots + a_{in}x_n).$$

Nous voyons qu'il peut être mis sous la forme:

$$(A\mathbf{x}, \mathbf{x}) = \sum_{i,\,k=1}^{n} a_{ik}\bar{x}_i x_k. \tag{145}$$

Si la condition

$$a_{ki} = \bar{a}_{ik} \quad (a_{kk} \text{ réels}) \tag{146}$$

est vérifiée, c'est alors une forme hermitienne à valeurs obligatoirement réelles. Le cas d'une matrice symétrique réelle A est un cas particulier des conditions (146). Si, de plus, les composantes du vecteur $\mathbf{x}$ sont réelles, la formule (145) nous donne alors la forme quadratique (134).

Démontrons maintenant que les racines de l'équation (144) sont réelles. Soit λ_k une racine de cette équation. Le système (143) nous donne dans ce cas les composantes du vecteur $\mathbf{x}^{(k)}$ (réel ou complexe) qui satisfait à l'équation (142). Multiplions scalairement à droite par $\mathbf{x}^{(k)}$ les deux membres de cette équation. Nous trouvons:

$$|\mathbf{x}^{(k)}|^2\lambda_k = (A\mathbf{x}^{(k)}, \mathbf{x}^{(k)}).$$

Comme nous l'avons vu, l'expression du second membre est un nombre réel, donc λ_k est aussi un nombre réel. Nous avons ainsi démontré que les racines de l'équation (144) sont réelles non seulement pour les matrices symétriques réelles, mais aussi pour celles dont les éléments vérifient les conditions (146). Ces matrices sont appelées *matrices hermitiennes*.

Dans le cas considéré, les coefficients du système (143) sont des nombres réels et nous pouvons supposer que les composantes de $\mathbf{x}^{(k)}$ sont réelles. Montrons à présent que si λ_p et λ_q sont deux racines distinctes de l'équation (144), les vecteurs correspondants $\mathbf{x}^{(p)}$ et $\mathbf{x}^{(q)}$, qui satisfont à l'équation (142), sont alors réciproquement orthogonaux. Nous avons par hypothèse:

$$A\mathbf{x}^{(p)} = \lambda_p\mathbf{x}^{(p)}; \quad A\mathbf{x}^{(q)} = \lambda_q\mathbf{x}^{(q)}.$$

Multipliant scalairement la première équation par $\mathbf{x}^{(q)}$ et la seconde par $\mathbf{x}^{(p)}$, puis en les retranchant, nous obtenons:

$$(A\mathbf{x}^{(p)}, \mathbf{x}^{(q)}) - (\mathbf{x}^{(p)}, A\mathbf{x}^{(q)}) = (\lambda_p - \lambda_q)(\mathbf{x}^{(p)}, \mathbf{x}^{(q)}). \tag{147}$$

Montrons maintenant que pour tout couple de vecteurs x et y (réels ou complexes) nous avons la formule:

$$(A\mathbf{x}, \mathbf{y}) = (\mathbf{x}, A\mathbf{y}), \tag{148}$$

si les éléments de la matrice A satisfont aux conditions (146). En effet, le premier membre de la formule (148) donne:

$$(A\mathbf{x}, \mathbf{y}) = \sum_{k=1}^{n} (a_{k1}x_1 + \ldots + a_{kn}x_n)\bar{y}_k = \sum_{i,\,k=1}^{n} a_{ki}x_i\bar{y}_k,$$

ou en vertu de (146):

$$(A\mathbf{x}, \mathbf{y}) = \sum_{i,\,k=1}^{n} \bar{a}_{ik}x_i\bar{y}_k.$$

Le second membre de la formule (148) nous donne le même résultat. Cette formule est donc aussi valable pour les matrices symétriques réelles qui sont un cas particulier des matrices hermitiennes. En vertu de (148), le premier membre de (147) est nul et puisque $\lambda_p \neq \lambda_q$ nous avons $(\mathbf{x}^{(p)}, \mathbf{x}^{(q)}) = 0$, c'est-à-dire que les vecteurs $\mathbf{x}^{(p)}$ et $\mathbf{x}^{(q)}$ sont réellement orthogonaux. Dans le cas considéré, ces vecteurs sont réels et la condition de leur orthogonalité se ramène à ce que la somme des produits de leurs composantes est nulle.

Si l'équation (144) a des racines distinctes, nous avons alors n vecteurs réels $\mathbf{x}^{(k)}$ réciproquement orthogonaux. L'équation (142) est linéaire, homogène en $\mathbf{x}^{(k)}$ et nous pouvons donc multiplier la solution de cette équation par une constante arbitraire de telle sorte que l'on puisse supposer que les vecteurs $\mathbf{x}^{(k)}$ mentionnés plus haut sont de longueur unité.

Les composantes de ces vecteurs forment les colonnes de la matrice B. Autrement dit, *cette matrice satisfait à la condition d'orthonormalité suivant les colonnes et est une matrice orthogonale. Ainsi*, le problème de la réduction d'une forme quadratique à une somme de carrés par une transformation orthogonale, ou, ce qui revient au même, de la réduction d'une matrice A à la forme diagonale, est résolu si l'on suppose que l'équation (144) ait des racines distinctes. Les nombres λ_k sont appelés *valeurs propres de la matrice A* et les vecteurs $\mathbf{x}^{(k)}$ *vecteurs propres de cette matrice.*

II-2-2. Cas des racines multiples de l'équation caractéristique. Considérons à présent le cas général, l'équation (144) ayant des racines égales. Soit $\lambda = \lambda_1$ une racine quelconque de l'équation (144) et construisons la solution de l'équation (142) qui lui correspond. On obtient un certain vecteur réel $\mathbf{x}^{(1)}$ de longueur unité. Adjoignons-lui $(n - 1)$ vecteurs réels unités de telle sorte qu'ils forment avec lui un système orthonormé complet [II-1-12]. Comme nous le savons, le passage des vecteurs de base à de nouveaux vecteurs de base est donné par une transformation orthogonale sur les composantes des vecteurs, la matrice A se transformant en la matrice semblable $A_1 = B_1^{-1}AB_1$. La nouvelle équation qui lui correspond :

$$A_1\mathrm{x} = \lambda\mathrm{x} \tag{149}$$

a pour solution, obtenue avec la valeur propre $\lambda = \lambda_1$ (les valeurs propres ne changent pas lors d'une similitude qui fait passer d'une matrice à une matrice semblable), le vecteur $\mathrm{x}^{(1)}$ que nous avons pris pour premier vecteur de base et qui a pour composantes (1, 0, ..., 0). Portant cette solution dans l'équation (149), nous avons:

$$A_1(1, 0, \ldots, 0) = (\lambda_1, 0, \ldots, 0),$$

il en découle immédiatement pour les éléments de la première colonne:

$$\{A_1\}_{11} = \lambda_1; \quad \{A_1\}_{21} = \{A_1\}_{31} = \ldots = \{A_1\}_{n1} = 0. \tag{150}$$

Montrons à présent que la matrice réelle A_1 est également symétrique, c'est-à-dire qu'elle coïncide avec sa transposée. En effet :

$$A_1^{(*)} = (B_1^{-1}AB_1)^{(*)} = B_1^{(*)}A^{(*)}B_1^{(*)-1}.$$

Mais en vertu de l'orthogonalité de la matrice B_1:

$$B_1^{(*)} = B_1^{-1} \quad \text{et} \quad B_1^{(*)-1} = B_1,$$

il en découle:

$$A_1^{(*)} = A_1.$$

Tenant compte des formules (150) et de ce que la matrice A_1 est symétrique, nous pouvons écrire :

$$\{A_1\}_{11} = \lambda_1; \quad \{A_1\}_{21} = \{A_1\}_{12} = \ldots = \{A_1\}_{n1} = \{A_1\}_{1n} = 0,$$

c'est-à-dire que tous les éléments de la première ligne et de la première colonne de la matrice A_1 sont nuls excepté peut-être l'élément $\{A_1\}_{11} = \lambda_1$, la matrice A_1 a alors la forme :

$$A_1 = \begin{Vmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & a_{22}^{(1)} & \ldots & a_{2n}^{(1)} \\ \cdot & \cdot & \cdot\ \cdot\ \cdot & \cdot \\ 0 & a_{n2}^{(1)} & \ldots & a_{nn}^{(1)} \end{Vmatrix},$$

dans laquelle nous désignons par $a_{ik}^{(1)}$ les éléments de A_1.

Avec les nouvelles coordonnées la forme quadratique φ s'écrit :

$$\varphi = \lambda_1 y_1'^2 + \sum_{i,\,k=2}^{n} a_{ik}^{(1)} y_i' y_k'.$$

Nous avons ainsi fait apparaître un carré et nous sommes ramenés à la considération d'une forme quadratique à $(n-1)$ variables :

$$\sum_{i,\,k=2}^{n} a_{ik}^{(1)} y_i y_k$$

ou encore, ce qui est analogue, à la considération de la matrice C_1 d'ordre $(n-1)$, sous-matrice de A_1, qui lui correspond. Nous pouvons ici raisonner de la même manière que plus haut et, dans le sous-espace à $(n-1)$ dimensions engendré par les $(n-1)$ derniers vecteurs de base, trouver un vecteur unité $\mathbf{x}^{(2)}$ qui soit solution de l'équation :

$$C_1\mathbf{x}^{(2)} = \lambda_2\mathbf{x}^{(2)}.$$

Il est évident que ce vecteur est orthogonal au vecteur $\mathbf{x}^{(1)}$. Il résulte de cette deuxième transformation que le vecteur de base $\mathbf{x}^{(1)}$ est conservé et que les autres vecteurs de base ont pour homologues d'autres vecteurs de base orthogonaux unités dont le premier est $\mathbf{x}^{(2)}$. Dans ces nouvelles variables, la forme quadratique φ s'écrit :

$$\varphi = \lambda_1 y_1''^2 + \lambda_2 y_2''^2 + \sum_{i,\,k=3}^{n} a_{i,k}^{(2)} y_i'' y_k''.$$

En continuant ainsi, nous réduisons la forme quadratique à une somme des carrés, c'est-à-dire que la matrice qui lui correspond est mise sous forme diagonale. Ceci est le résultat de l'application de plusieurs transformations orthogonales, ce qui équivaut évidemment à l'application d'une seule transformation orthogonale B qui en est le produit.

La matrice diagonale finale

$$B^{-1}AB = [\lambda_1, \ldots, \lambda_n] \tag{151}$$

est semblable à la matrice initiale A et, par conséquent, son équation caractéristique

$$\begin{vmatrix} \lambda_1-\lambda & 0 & \ldots & 0 \\ 0 & \lambda_2-\lambda & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \ldots & \lambda_n-\lambda \end{vmatrix} = 0$$

coïncide avec l'équation (144), en d'autres termes, les coefficients λ_k des termes carrés de la forme quadratique réduite

$$\varphi = \lambda_1 x_1'^2 + \ldots + \lambda_n x_n'^2 \tag{152}$$

sont les racines de l'équation (144), de plus, chaque racine multiple est répétée autant de fois que sa multiplicité l'indique.

Chaque colonne de la transformation orthogonale finale B donne, comme nous le savons, un vecteur, solution de l'équation (142); de plus, de la loi même de construction il découle que la valeur λ_k coïncide dans la forme quadratique (152) avec le coefficient de la variable correspondante. Explicitons cette relation. En vertu de (136), la transformation orthogonale B, qui satisfait à la condition (140), transforme les variables $(x_1', \ldots, x_n')$ en les variables $(x_1, \ldots, x_n)$.

La transformation inverse B^{-1} est la transposée de B, c'est-à-dire que nous avons:

$$x_k' = b_{1k}x_1 + \ldots + b_{nk}x_n \qquad (k = 1, 2, \ldots, n) \tag{153}$$

et le vecteur $\mathbf{x}^{(k)}$ de composantes $(b_{1k}, \ldots, b_{nk})$ est solution de l'équation (142) pour $\lambda = \lambda_k$.

Montrons enfin que nous avons trouvé *toutes les solutions de l'équation* (142). Il découle des raisonnements précédents que la valeur λ_k est une racine de l'équation (144). Soit λ une racine de cette équation et, pour fixer les idées, supposons qu'elle soit de multiplicité trois. On peut évidemment prendre $\lambda = \lambda_1 = \lambda_2 = \lambda_3$. Les raisonnements précédents donnent pour l'équation

$$A\mathbf{x} = \lambda_1 \mathbf{x} \tag{154}$$

les trois solutions:

$$\mathbf{x}^{(1)}(b_{11}, \ldots, b_{n1}); \quad \mathbf{x}^{(2)}(b_{12}, \ldots, b_{n2}); \quad \mathbf{x}^{(3)}(b_{13}, \ldots, b_{n3}).$$

Montrons que toute solution de l'équation (154) est une combinaison linéaire de ces dernières. En effet, s'il n'en était pas ainsi, il existerait une solution $\mathbf{y}$ de cette équation, linéairement indépendante de $\mathbf{x}^{(1)}$, $\mathbf{x}^{(2)}$ et $\mathbf{x}^{(3)}$. Le vecteur $\mathbf{y}$ peut être aussi complexe, mais dans ce cas, ses parties réelle et imaginaire doivent séparément satisfaire à l'équation (154), puisque cette équation a ses coefficients réels. Il est évident qu'au moins une de ces parties est un vecteur linéairement indépendant de $\mathbf{x}^{(k)}$ $(k = 1, 2, 3)$. Nous pouvons donc supposer que le vecteur $\mathbf{y}$, dont il est question, est réel. Comme nous l'avons démontré précédemment, il doit être orthogonal à tous les vecteurs $\mathbf{x}^{(k)}$ pour $k > 3$, puisque ces derniers correspondent aux λ_k distincts de λ_1. Il en résulte que le vecteur $\mathbf{y}$ et l'ensemble des vecteurs $\mathbf{x}^{(k)}$ sont linéairement indépendants, c'est-à-dire que nous avons $(n+1)$ vecteurs linéairement indépendants, ce qui est impossible. Donc, pour chaque racine de l'équation (144), de multiplicité m, l'équation (154) a exactement m solutions réelles linéairement indépendantes.

Substituant à λ_k, dans les coefficients du système (143), une racine quelconque $\lambda = \lambda_0$, de multiplicité m, nous obtenons un système homogène ayant m solutions linéairement indépendantes, c'est-à-dire que le rang de ce système doit être $(n - m)$. Autrement dit, ce système se réduit à $(n - m)$ équations. Considérons une solution quelconque de ce système et multiplions-la par un facteur tel que la somme des quantités élevées au carré qui entrent dans cette solution soit égale à l'unité. Nous obtenons ainsi un vecteur qui correspond à la racine $\lambda = \lambda_0$ de l'équation. Pour trouver le vecteur suivant ajoutons aux $(n - m)$ équations de notre système l'équation exprimant l'orthogonalité du vecteur inconnu à celui déjà construit. Pour déterminer les composantes du vecteur cherché, nous avons un système homogène de $(n - m + 1)$ équations. En considérant une solution quelconque de ce système et en la normant de nouveau à l'unité (la longueur du vecteur est l'unité) passons à la détermination du troisième vecteur correspondant à la même racine $\lambda = \lambda_0$ de l'équation. Pour cela ajoutons aux $(n - m)$ équations initiales de ce système deux autres équations qui expriment la condition d'orthogonalité du vecteur inconnu et des deux vecteurs déjà connus, etc., jusqu'à ce que l'on construise un système orthonormé composé de m vecteurs qui correspondent à la racine $\lambda = \lambda_0$ de l'équation, de multiplicité m. Il en découle immédiatement que la construction des solutions fondamentales de l'équation (142) possède un certain degré d'arbitraire. Si toutes les racines de l'équation sont simples, cet arbitraire ne conduit qu'à ce que toutes les composantes du vecteur $x^{(h)}$ peuvent être multipliées par (-1). Supposons maintenant que l'équation (144) ait une racine de multiplicité m. Dans ce cas, les m vecteurs orthogonaux unités correspondants, solution de l'équation (142), engendrent un sous-espace R_m de dimension m. Il est évident que nous pouvons choisir d'une manière quelconque dans ce sous-espace des vecteurs de base unités orthogonaux qui soient aussi solutions de l'équation (142) pour $\lambda = \lambda_0$, c'est-à-dire que nous pouvons passer d'un système orthonormé à un autre, au moyen d'une transformation orthogonale du sous-espace R_m. Tout ce qui vient d'être dit reste également valable pour toute autre racine multiple de l'équation (144).

Pour illustrer ce qui précède reprenons le problème avec lequel nous avons débuté dans [II-2-1], à savoir celui de la réduction de l'équation d'une surface du second ordre à ses axes de symétrie. Supposons que cette surface soit un ellipsoïde. Le cas des racines distinctes (144) correspond à un ellipsoïde de demi-axes distincts. Dans ce cas, l'unique arbitraire dans le choix des axes de coordonnées finals réside dans le fait que l'on peut changer l'orientation de ces axes. Si l'équation (144) qui, dans le cas considéré, est une équation du troisième degré, a deux racines égales, alors l'ellipsoïde est un ellipsoïde de révolution et les deux axes de symétrie peuvent être choisis d'une manière arbitraire dans le plan qui passe par le centre et qui est perpendiculaire à l'axe de révolution pourvu qu'ils soient réciproquement orthogonaux, c'est-à-dire que dans ce cas le choix des axes finals résdoi encore dans une transformation orthogonale arbitraire dans le plan mentionné ci-dessus. Enfin, si l'équation (144) a trois racines égales, l'ellipsoïde est alors une sphère et l'équation ne contient pas de termes rectangulaires. Dans ce cas nous pouvons, en général, choisir d'une manière arbitraire les axes de coordonnées rectangulaires rectilignes dans l'espace.

II-2-3. Exemples. Donnons deux exemples numériques.

1. Rapporter aux axes de symétrie la surface d'équation:

$$x_1^2+5x_2^2+x_3^2+2x_1x_2+6x_1x_3+2x_2x_3=5.$$

La forme quadratique correspondante est de la forme:

$$\varphi=x_1^2+x_1x_2+3x_1x_3+x_2x_1+5x_2^2+x_2x_3+3x_3x_1+x_3x_2+x_3^2.$$

L'équation caractéristique de sa matrice est:

$$\begin{vmatrix} 1-\lambda & 1 & 3 \\ 1 & 5-\lambda & 1 \\ 3 & 1 & 1-\lambda \end{vmatrix}=0,$$

d'où en développant par rapport aux éléments de la première ligne on a:

$$(1-\lambda)[(5-\lambda)(1-\lambda)-1]-(1-\lambda-3)+3[1-3(5-\lambda)]=0$$

ou

$$\lambda^3-7\lambda^2+36=0.$$

Les racines de cette équation, comme il est facile de montrer, sont:

$$\lambda_1=-2; \quad \lambda_2=3; \quad \lambda_3=6,$$

et l'équation de notre surface rapportée aux axes de symétrie est:

$$-2x_1'^2+3x_2'^2+6x_3'^2=5.$$

Définissons à présent les éléments de la matrice orthogonale:

$$B=\begin{Vmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{Vmatrix}.$$

Nous avons pour eux le système:

$$\left.\begin{aligned} (1-\lambda)\,b_{1h}+b_{2h}+3b_{3h}&=0, \\ b_{1h}+(5-\lambda)\,b_{2h}+b_{3h}&=0, \\ 3b_{1h}+b_{2h}+(1-\lambda)\,b_{3h}&=0. \end{aligned}\right\} \qquad (155)$$

Faisons d'abord $\lambda=\lambda_1=-2$, ce qui nous donne les deux équations:

$$\left.\begin{aligned} 3b_{11}+b_{21}+3b_{31}&=0, \\ b_{11}+7b_{21}+b_{31}&=0. \end{aligned}\right\}$$

La solution de ce système est :

$$b_{11} = -k_1;\quad b_{21} = 0;\quad b_{31} = k_1,$$

où k_1 est un nombre arbitraire. Choisissons-le de façon que la somme des carrés des composantes de la solution soit égale à l'unité. Nous avons donc :

$$b_{11} = \frac{1}{\sqrt{2}};\quad b_{21} = 0;\quad b_{31} = -\frac{1}{\sqrt{2}},$$

de plus, on peut prendre aussi comme solution cette dernière changée de signe.

Faisons à présent $\lambda = \lambda_2 = 3$ dans les coefficients du système (155). Nous obtenons un système dans lequel la troisième équation est la différence de deux premières et nous avons ainsi les deux équations :

$$-2b_{12} + b_{22} + 3b_{32} = 0,$$
$$b_{12} + 2b_{22} + b_{32} = 0.$$

On trouve facilement la solution normée à l'unité de ce système :

$$b_{12} = \frac{1}{\sqrt{3}};\quad b_{22} = -\frac{1}{\sqrt{3}};\quad b_{32} = \frac{1}{\sqrt{3}}.$$

Substituons enfin la troisième racine dans les coefficients du système (155). Nous avons encore un système dans lequel une des équations est la conséquence des deux autres. Résolvant les deux équations qui restent et normant à l'unité la solution obtenue, nous avons :

$$b_{13} = \frac{1}{\sqrt{6}};\quad b_{23} = \frac{2}{\sqrt{6}};\quad b_{33} = \frac{1}{\sqrt{6}}.$$

Dans le cas donné, les formules de transformation des variables sont de la forme :

$$x_1' = \frac{1}{\sqrt{2}} x_1 \qquad\qquad - \frac{1}{\sqrt{2}} x_3,$$
$$x_2' = \frac{1}{\sqrt{3}} x_1 - \frac{1}{\sqrt{3}} x_2 + \frac{1}{\sqrt{3}} x_3,$$
$$x_3' = \frac{1}{\sqrt{6}} x_1 + \frac{2}{\sqrt{6}} x_2 + \frac{1}{\sqrt{6}} x_3.$$

2. Rapporter aux axes de symétrie la surface d'équation :

$$2x_1^2 + 6x_2^2 + 2x_3^2 + 8x_1x_3 = 1.$$

Dans ce cas, la forme quadratique s'écrit sous la forme :

$$\varphi = 2x_1^2 + 0x_1x_2 + 4x_1x_3 + 0x_2x_1 + 6x_2^2 + 0x_2x_3 + 4x_3x_1 + 0x_3x_2 + 2x_3^2.$$

l'équation caractéristique de sa matrice est :

$$\begin{vmatrix} 2-\lambda & 0 & 4 \\ 0 & 6-\lambda & 0 \\ 4 & 0 & 2-\lambda \end{vmatrix} = 0.$$

Développant ce déterminant, nous avons l'équation :

$$\lambda^3 - 10\lambda^2 + 12\lambda + 72 = 0.$$

Ses racines sont $\lambda_1=-2$; $\lambda_2=\lambda_3=6$, c'est-à-dire que cette équation a une racine double.

Passons à la détermination des coefficients de la transformation orthogonale. Nous avons le système :

$$\left.\begin{aligned}(2-\lambda)\,b_{1k}+4b_{3k}&=0,\\ (6-\lambda)\,b_{2k}&=0,\\ 4b_{1k}+(2-\lambda)\,b_{3k}&=0.\end{aligned}\right\} \qquad (155_1)$$

Faisant $\lambda=-2$, nous obtenons, comme il est facile de le voir, la solution normée à l'unité :

$$b_{11}=\frac{1}{\sqrt{2}}\,; \quad b_{21}=0\,; \quad b_{31}=-\frac{1}{\sqrt{2}}\,.$$

Substituons maintenant la racine double $\lambda=6$ dans les coefficients du système (155) pour laquelle nous devons obtenir deux solutions linéairement indépendantes et réciproquement orthogonales. Pour la substitution en question le système se réduit à l'équation :

$$-b_{12}+b_{32}=0. \qquad (155_2)$$

Considérons la solution de cette équation normée à l'unité :

$$b_{12}=\frac{1}{\sqrt{2}}\,; \quad b_{22}=0\,; \quad b_{32}=\frac{1}{\sqrt{2}}\,.$$

Pour trouver la deuxième solution remarquons qu'elle doit satisfaire simultanément à l'équation (155_2) et à la condition d'orthogonalité avec la solution déjà trouvée. Ainsi, pour la déterminer nous avons les deux équations :

$$\left.\begin{aligned}-b_{13}+b_{33}&=0,\\ \frac{1}{\sqrt{2}}\,b_{13}+\frac{1}{\sqrt{2}}\,b_{33}&=0\end{aligned}\right\}$$

ou $b_{13}=b_{33}=0$, d'où la solution normée à l'unité est $b_{13}=0$; $b_{23}=1$; $b_{33}=0$.

La transformation orthogonale est donc de la forme :

$$\begin{aligned}x_1'&=\frac{1}{\sqrt{2}}\,x_1-\frac{1}{\sqrt{2}}\,x_3,\\ x_2'&=\frac{1}{\sqrt{2}}\,x_1+\frac{1}{\sqrt{2}}\,x_3,\\ x_3'&=\qquad\quad x_2\end{aligned}$$

et l'équation de la surface rapportée aux axes de symétrie :

$$-2x_1'^2+6\,(x_2'^2+x_3'^2)=1.$$

II-2-4. Classification des formes quadratiques. Le problème de la réduction d'une forme quadratique à une somme de carrés peut être posé sous une forme encore plus générale que nous ne l'avons fait jusqu'ici et n'exigeant pas que la transformation linéaire entre les nouvelles et les anciennes variables soit forcément orthogonale. Autrement dit, nous pouvons poser le problème de la façon suivante:

mettre la forme quadratique réelle (134) sous la forme :

$$\varphi=\mu_1X_1^2+\mu_2X_2^2+\ldots+\mu_nX_n^2, \tag{156}$$

où les X_k sont n formes linéaires réelles linéairement indépendantes des variables x_k. Dans cet énoncé du problème, les coefficients μ_k ne sont pas des nombres déterminés, comme cela était jusqu'à présent, mais nous pouvons tout de même donner une proposition concernant ces coefficients, à savoir : le nombre des coefficients non nuls est constamment égal au rang de la matrice des coefficients a_{ik} de la forme quadratique. Autrement dit, pour toute réduction d'une forme quadratique à une somme de carrés de formes linéaires linéairement indépendantes, le nombre des carrés est égal au rang de la matrice mentionnée. De plus, selon *le théorème d'inertie des formes quadratiques*, pour toute transformation d'une forme quadratique réelle en la forme (156), où les formes linéaires X_k sont également réelles, *le nombre des coefficients positifs* μ_k *(et le nombre des coefficients négatifs* μ_k*) est constamment le même.* Ces considérations seront démontrées un peu plus loin.

Le problème général de la réduction d'une forme quadratique à la forme (156), que l'on vient de poser, se résout d'une manière très simple en faisant apparaître des expressions linéaires élevées au carré. Montrons-le sur un exemple :

$$\varphi=x_1^2+4x_2^2+x_3^2+2x_1x_2-6x_1x_3+8x_2x_3.$$

Ajoutant aux termes $(x_1^2+2x_1x_2-6x_1x_3)$ l'expression $(x_2^2+9x_3^2-6x_2x_3)$, nous obtenons un carré parfait et nous pouvons écrire φ sous la forme :

$$\varphi=(x_1+x_2-3x_3)^2+3x_2^2-8x_3^2+14x_2x_3.$$

De la même façon, isolant encore un carré, nous arrivons finalement à représenter la forme quadratique sous la forme (156) :

$$\varphi=(x_1+x_2-3x_3)^2-2\left(2x_3-\frac{7}{4}x_2\right)^2+\frac{73}{8}(x_2)^2.$$

Les formes linéaires entre parenthèses sont évidemment linéairement indépendantes.

Dans le cas où les carrés des variables n'apparaissent pas dans l'expression de φ, il faut conduire les calculs autrement. Soit :

$$\varphi=ax_1x_2+Px_1+Qx_2+R,$$

où a est un coefficient numérique non nul, P et Q sont des formes linéaires des variables, indépendantes de x_1 et x_2, R une forme quadratique également indépendante de x_1 et x_2. Nous pouvons écrire :

$$\varphi=a\left(x_1+\frac{Q}{a}\right)\left(x_2+\frac{P}{a}\right)+R-\frac{PQ}{a^2}.$$

Si l'on pose :

$$X_1=\frac{1}{2}\left(x_1+x_2-\frac{P+Q}{a}\right); \quad X_2=\frac{1}{2}\left(x_1-x_2-\frac{P-Q}{a}\right)$$

et

$$\varphi_1 = R - \frac{PQ}{a^2},$$

nous avons alors :

$$\varphi = aX_1^2 - aX_2^2 + \varphi_1,$$

où φ_1 est une forme quadratique indépendante de x_1 et x_2. En faisant apparaître deux carrés, nous nous libérons de deux variables.

La réduction d'une forme quadratique à la forme (156) donne la possibilité d'une classification naturelle de telles formes. Considérons plusieurs cas.

I. Supposons que dans la formule (156) tous les coefficients μ_k soient positifs. Dans ce cas, la forme est dite *définie positive*. Il est facile de montrer qu'elle a des valeurs positives pour toutes les valeurs réelles des x_k et ne peut s'annuler que si tous les x_k sont nuls. En effet, pour que le second membre de la formule (156) s'annule, il faut et il suffit, puisque tous les μ_k sont positifs, que toutes les formes linéaires x_k soient nulles. Nous obtenons ainsi pour les x_k un système de n équations homogènes à déterminant non nul (les formes sont linéairement indépendantes) et, par suite, ce système n'a que la solution nulle.

II. Si tous les coefficients μ_k sont négatifs, la forme est dite *définie négative*. Comme plus haut, on peut montrer qu'elle ne prend que des valeurs négatives pour tous les x_k réels, de plus elle s'annule si et seulement si tous les x_k sont nuls.

III. Considérons à présent le cas où plusieurs coefficients μ_k sont nuls et tous les autres de signe déterminé, positifs par exemple. Dans ce cas, nous avons pour la forme φ la représentation suivante :

$$\varphi = \mu_1 X_1^2 + \ldots + \mu_m X_m^2 \quad (m < n), \tag{156$_1$}$$

où tous les μ_k sont positifs. Ici les valeurs de notre forme ne peuvent être négatives pour aucune valeur de x_k mais peuvent être nulles même si les valeurs des x_k ne sont pas nulles. En effet, pour que la forme s'annule, nous devons écrire un système de m équations homogènes en x_k :

$$X_1 = X_2 = \ldots = X_m = 0,$$

et comme $m < n$, alors ce système a probablement des solutions non nulles. De même, si dans la formule (156$_1$) tous les coefficients μ_k sont négatifs, la forme quadratique ne peut prendre des valeurs positives, mais peut s'annuler même pour des valeurs non nulles des x_k. Dans le cas considéré la forme est dite *à signe constant* ou *semi-définie* soit positive, soit négative.

IV. Enfin si parmi les coefficients μ_k de la formule (156) certains sont positifs, d'autres négatifs, il est facile de voir alors que la forme quadratique peut prendre des valeurs tant positives que négatives pour toutes les valeurs réelles de x_k. Dans ce cas, on dit qu'elle est *indéterminée*.

La classification précédente des formes quadratiques réelles a une application directe dans le problème de détermination de maximum ou de minimum des fonctions de plusieurs variables. Soit une fonction de n variables indépendantes $x_1, \ldots, x_n$:

$$\psi(x_1, \ldots, x_n).$$

Supposons que pour les valeurs $x_1 = \ldots = x_n = 0$ les conditions nécessaires de maximum ou de minimum soient vérifiées, c'est-à-dire que toutes les dérivées partielles de la fonction ψ par rapport aux variables indépendantes s'annulent. Développant cette fonction en série de Maclaurin, nous avons:

$$\psi(x_1, \ldots, x_n) - \psi(0, \ldots, 0) = \varphi(x_1, \ldots, x_n) + \omega,$$

où $\varphi(x_1, \ldots, x_n)$ désigne une forme quadratique des variables x_k et ω l'ensemble de termes en x_k d'ordre supérieur à deux. Si la forme quadratique φ est définie positive, nous avons alors un minimum de la fonction au point $x_1 = \ldots = x_n = 0$. Si elle est définie négative, nous avons un maximum. Si elle est indéterminée, nous n'avons ni de minimum, ni de maximum et si, enfin, φ est une forme semi-définie, nous sommes en présence alors d'un cas douteux. Ce résultat est un complément logique de celui que nous avions donné (tome I, [V-2-3]) dans le cas des fonctions de deux variables indépendantes.

Démontrons la proposition énoncée au début du paragraphe. Soit une forme quadratique:

$$\varphi = \sum_{i,\, k=1}^{n} a_{ik} x_i x_k \qquad (a_{ik} = a_{ki}),$$

où r est le rang du tableau de ses coefficients. Considérons le système des n formes linéaires:

$$\frac{1}{2}\frac{\partial\varphi}{\partial x_s} = \sum_{l=1}^{n} a_{sl} x_l \qquad (s = 1, 2, \ldots, n). \tag{157}$$

En calculant ces dérivées partielles premières, nous avons utilisé les conditions $a_{ik} = a_{ki}$. Il est évident que le nombre r est le rang du système de formes (157) au sens de [I-2-4].

Supposons que l'on puisse mettre φ sous forme d'une somme de carrés de m formes linéairement indépendantes:

$$y_s = \beta_{s1} x_1 + \beta_{s2} x_2 + \ldots + \beta_{sn} x_n \qquad (s = 1, 2, \ldots, m), \tag{158}$$

c'est-à-dire

$$\varphi = \mu_1 y_1^2 + \mu_2 y_2^2 + \ldots + \mu_m y_m^2, \tag{159}$$

où les μ_s ne sont pas nuls. Nous devons démontrer que $m = r$. Tenant compte de l'expression (159), formons les formes linéaires (157) :

$$\frac{1}{2}\frac{\partial\varphi}{\partial x_s} = \mu_1\beta_{1s}y_1 + \mu_2\beta_{2s}y_2 + \ldots + \mu_m\beta_{ms}y_m \qquad (s = 1, 2, \ldots, n). \qquad (157_1)$$

Les variables y_s peuvent prendre des valeurs arbitraires puisque les formes (158) sont linéairement indépendantes. C'est pourquoi dans la définition de la dépendance linéaire des formes (157_1), les y_s peuvent être considérés comme des variables indépendantes et le nombre maximum de formes linéairement indépendantes dans le système (157_1) doit coïncider avec le rang du tableau des coefficients $\mu_k\beta_{ki}$, où le numéro de la colonne k prend toutes les valeurs possibles: $k = 1, 2, \ldots, m$, et le numéro de la ligne i les valeurs $i = 1, 2, \ldots, n$. Les éléments de chaque colonne de ce tableau ont un facteur commun μ_k, non nul, et, par suite, le rang du tableau $\mu_k\beta_{ki}$ coïncide avec celui du tableau β_{ki}. Puisque le système des m formes (158) est un système de formes linéairement indépendantes, ce rang coïncide avec m, c'est-à-dire que le maximum de formes linéairement indépendantes dans le système (157_1) ou (157) est m. D'autre part, ce nombre coïncide par hypothèse avec r, il en découle donc que $m = r$.

Montrons maintenant que pour toute représentation de φ par une formule de la forme (159), où les y_s sont des formes réelles linéairement indépendantes, le nombre des coefficients positifs et négatifs μ_s est toujours le même. Démontrons-le par l'absurde. Supposons que nous ayons deux représentations de φ au moyen de formules de la forme (159) et supposons que le nombre des coefficients positifs dans ces représentations soit différent:

$$\left.\begin{aligned} \varphi &= \lambda_1 y_1^2 + \ldots + \lambda_p y_p^2 - \lambda_{p+1} y_{p+1}^2 - \ldots - \lambda_m y_m^2, \\ \varphi &= \lambda_1' y_1'^2 + \ldots + \lambda_q' y_q'^2 - \lambda_{q+1}' y_{q+1}'^2 - \ldots - \lambda_m' y_m'^2. \end{aligned}\right\} \qquad (160)$$

Dans ces formules λ_s et λ_s' sont supposés positifs. Les formes $y_1, \ldots, y_m$ sont linéairement indépendantes, il en est de même pour les formes $y_1', \ldots, y_m'$. Puisque $p \neq q$, on peut toujours prendre, par exemple, $p < q$. Montrons que cela conduit à une absurdité. Adjoignons les formes $y_{m+1}, \ldots, y_n$ aux formes $y_1, \ldots, y_m$ de façon à obtenir un système complet de formes linéairement indépendantes [I-2-4]. Ecrivons le système d'équations linéaires homogènes pour les $x_1, x_2, \ldots, x_n$:

$$y_1 = 0; \ldots; y_p = 0; \; y_{q+1}' = 0; \ldots; y_m' = 0; \; y_{m+1} = 0; \ldots; y_n = 0. \quad (161)$$

Le nombre de ces équations homogènes est $p + (m - q) + (n - m) = n - (q - p)$ et, comme $q - p > 0$, ce nombre est inférieur à n. Donc, le système homogène envisagé a des solutions réelles non nulles. Considérons une solution quelconque: $x_s = x_s^{(0)}$ $(s = 1, 2, \ldots, n)$. Pour ces valeurs de x_s, en vertu de (161), nous avons:

$$\varphi = -\lambda_{p+1}y_{p+1}^2 - \ldots - \lambda_m y_m^2 = \lambda_1' y_1'^2 + \ldots + \lambda_q' y_q'^2.$$

On voit donc que pour $x_s = x_s^{(0)}$ la forme quadratique φ s'annule et, par conséquent, les $x_s = x_s^{(0)}$, outre les équations (161), doivent satisfaire aux équations :

$$y_{p+1} = 0; \ldots; y_m = 0.$$

Nous obtenons finalement que les $x_s = x_s^{(0)}$ doivent annuler toutes les formes du système complet de formes linéairement indépendantes: $y_1, y_2, \ldots, y_n$. Mais cela n'est pas possible puisque dans le système homogène en $x_1, x_2, \ldots, x_n$

$$y_1 = 0; \; y_2 = 0; \ldots; \; y_n = 0$$

le déterminant n'est pas nul, les formes y_s étant linéairement indépendantes. Nous aboutissons à une contradiction, ce qui démontre précisément le théorème d'inertie.

II-2-5. **Formule de Jacobi.** Donnons, sans démonstration, la formule de Jacobi qui sous une forme commode réduit une forme quadratique à une somme de carrés.

Pour cela introduisons certaines notations. Soient:

$$A_i(x)=\sum_{k=1}^{n} a_{ik}x_k \quad (i=1,\ 2,\ \ldots,\ n),$$

$$\Delta_0=1\ ;\ \Delta_1=a_{11}\ ;\ \Delta_k=\begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1k} \\ a_{21} & a_{22} & \ldots & a_{2k} \\ \cdot & \cdot & \cdot & \cdot \\ a_{k1} & a_{k2} & \ldots & a_{kk} \end{vmatrix},$$

$$(k=2,\ 3,\ \ldots,\ n)$$

$$X_1=A_1(x)\ ;\ X_k=\begin{vmatrix} a_{11} & \ldots & a_{1k-1} & A_1(x) \\ a_{21} & \ldots & a_{2k-1} & A_2(x) \\ \cdot & \cdot & \cdot & \cdot \\ a_{k1} & \ldots & a_{kk-1} & A_k(x) \end{vmatrix}.$$

Si le rang du tableau des éléments a_{ik} est r et si les déterminants Δ_1, $\Delta_2, \ldots, \Delta_r$ ne sont pas nuls, la formule de Jacobi est de la forme:

$$\varphi=\sum_{i,\ k=1}^{n} a_{ik}x_ix_k=\sum_{k=1}^{r}\frac{X_k^2}{\Delta_k\Delta_{k-1}}, \tag{162}$$

et les formes linéaires X_k ($k = 1,\ 2,\ \ldots,\ r$) sont linéairement indépendantes. La dernière formule permet de déterminer, d'après les signes de Δ_k, le type auquel se rapporte la forme φ par rapport au théorème d'inertie.

En particulier, si tous les déterminants Δ_1, Δ_2, $\ldots$, Δ_n sont positifs (de plus $r = n$), il découle de (162) que φ est définie positive. On peut aussi démontrer la proposition réciproque: si φ est une forme définie positive, alors tous les déterminants mentionnés sont positifs. Dans les applications de la formule (162) on peut évidemment numéroter les variables x_s suivant un ordre quelconque. Avec un changement de numérotation les déterminants Δ_k indiqués plus haut changent évidemment et chaque mineur principal de la matrice $\|a_{ik}\|_1^n$ peut être un déterminant de la suite Δ_k pour une numérotation déterminée des variables x_s. D'après ce qui vient d'être dit, tous les mineurs principaux d'une forme définie positive φ sont positifs, mais il suffit, en outre, de s'assurer que les déterminants

$$\Delta_s'(s=1,\ 2,\ \ldots,\ n)$$

sont positifs. On peut montrer que pour qu'une forme soit positive, il faut et il suffit que tous les mineurs principaux soient non négatifs, c'est-à-dire qu'ils soient plus grands ou égaux à zéro. Ici, il ne suffit pas de déterminer le signe des déterminants Δ_s, mais il faut encore définir le signe de tous les mineurs principaux.

Pour qu'une forme φ soit définie négative, il faut et il suffit que les inégalités $(-1)^k\Delta_k > 0$ ($k=1,2,\ \ldots\ ,\ n$) soient vérifiées. Pour que la forme φ

soit négative, il faut et il suffit que les mineurs principaux aient le signe $(-1)^k$, où k est l'ordre du mineur, ou soient nuls.

La démonstration des propositions qu'on vient d'énoncer se trouve dans le livre de F. R. Gantmakher « Théorie des matrices » (1953).

II-2-6. Réduction simultanée de deux formes quadratiques à une somme de carrés. Soient deux formes quadratiques:

$$\varphi_1 = \sum_{i,k=1}^{n} a_{ik}x_ix_k; \quad \varphi_2 = \sum_{i,k=1}^{n} b_{ik}x_ix_k,$$

avec φ_1 définie positive et donc réductible à une somme de n carrés positifs. On demande de trouver une transformation linéaire (non obligatoirement orthogonale) telle que les deux formes soient réduites à une somme de carrés.

Introduisons d'abord de nouvelles variables y_k de façon que la forme φ_1 se réduise à une somme de carrés. On peut y arriver, par exemple, par un procédé élémentaire indiqué dans [II-2-5]. Avec les nouvelles variables nous avons les représentations suivantes des formes quadratiques:

$$\varphi_1 = \sum_{k=1}^{n} \mu_k y_k^2; \quad \varphi_2 = \sum_{i,k=1}^{n} b'_{ik}y_iy_k.$$

Par hypothèse, tous les nombres μ_k sont positifs et nous pouvons introduire les nouvelles variables réelles $z_k = \sqrt{\mu_k}\, y_k$. Nous obtenons alors des formules de la forme:

$$\varphi_1 = \sum_{k=1}^{n} z_k^2; \quad \varphi_2 = \sum_{i,k=1}^{n} b''_{ik}z_iz_k.$$

Faisons une transformation orthogonale de façon que les variables z_k soient remplacées par les variables z'_k et qui réduit la forme φ_2 à une somme de carrés.

Alors φ_1 reste égale à une somme de carrés puisque la transformation est orthogonale et les deux formes se mettent finalement sous forme d'une somme de carrés:

$$\varphi_1 = \sum_{k=1}^{n} z_k'^2; \quad \varphi_2 = \sum_{k=1}^{n} \lambda_k z_k'^2.$$

Les nombres λ_k sont appelés *valeurs propres de la forme φ_2 par rapport à la forme φ_1.*

Etablissons à présent que l'équation à laquelle doivent satisfaire les nombres λ_k est semblable à l'équation (144) de [II-2-1]. Pour cela introduisons la notion de *discriminant* d'une forme quadratique, à savoir: *on appelle discriminant d'une forme quadratique le déterminant constitué par les coefficients de la forme.*

Supposons que nous transformions la forme φ, de matrice des coefficients A, de manière qu'elle s'exprime dans les nouvelles variables à l'aide de la transformation :

$$(x_1, \ldots, x_n) = B(x'_1, \ldots, x'_n).$$

Comme on le sait de [II-2-1], la matrice de la nouvelle forme est :

$$C = B^{(*)}AB,$$

et son déterminant se calcule d'après la formule:

$$D(C) = D(B^{(*)})\, D(A)\, D(B).$$

Il est clair que les déterminants $D(B^{(*)})$ et $D(B)$ sont égaux puisque les tableaux correspondants s'obtiennent l'un à partir de l'autre en échangeant les lignes et les colonnes. Nous avons donc :

$$D(C) = D(A)\, D(B)^2,$$

c'est-à-dire que par une transformation linéaire sur les variables d'une forme quadratique le discriminant de la forme est multiplié par le carré du déterminant de la transformation qui permet de passer des nouvelles variables aux anciennes.

Revenons maintenant aux formes quadratiques φ_1 et φ_2 et formons la forme quadratique :

$$\omega = \varphi_2 - \lambda\varphi_1 = \sum_{i,\,k=1}^{n} (b_{ik} - \lambda a_{ik})\, x_i x_k,$$

dont les coefficients dépendent du paramètre λ.

Le résultat de la transformation qui permet de passer aux nouvelles variables est que cette forme s'écrit :

$$\omega = \sum_{k=1}^{n} (\lambda_k - \lambda)\, z_k'^2,$$

et son discriminant par rapport aux nouvelles variables s'exprime par le produit :

$$(\lambda_1 - \lambda)(\lambda_2 - \lambda) \ldots (\lambda_n - \lambda), \tag{163}$$

et, par rapport aux anciennes variables, ce discriminant est égal au déterminant d'éléments $(b_{ik} - \lambda a_{ik})$. Comme nous l'avons montré, ces deux discriminants ne diffèrent que par un facteur, le carré du déterminant de la transformation, indépendant de λ et non nul. Il en découle immédiatement que les deux discriminants ont les mêmes racines par rapport au paramètre λ. Tenant compte de

(163), nous voyons que les nombres λ_h sont les racines de l'équation :

$$\begin{vmatrix} b_{11}-\lambda a_{11} & b_{12}-\lambda a_{12} & \dots & b_{1n}-\lambda a_{1n} \\ b_{21}-\lambda a_{21} & b_{22}-\lambda a_{22} & \dots & b_{2n}-\lambda a_{2n} \\ \dots & \dots & \dots & \dots \\ b_{n1}-\lambda a_{n1} & b_{n2}-\lambda a_{n2} & \dots & b_{nn}-\lambda a_{nn} \end{vmatrix} = 0.$$

II-2-7. Petites oscillations. Nous avons vu (tome II, [I-2-6]) que le mouvement d'un système mécanique à n degrés de liberté, dont les liaisons sont indépendantes du temps et qui se trouve sous l'action de forces dérivant d'un potentiel, est défini par un système d'équations différentielles de la forme :

$$\frac{d}{dt}\left(\frac{\partial T}{\partial q'_k}\right)-\frac{\partial T}{\partial q_k}=\frac{\partial U}{\partial q_k} \quad (k=1,\ 2,\ \dots,\ n), \tag{164}$$

où T est l'énergie cinétique du système et U une fonction donnée (la fonction de force) des q_k que nous considérons indépendante de t. Comme il l'a été dit ci-dessus (tome II, [I-2-6]), T est une forme quadratique des dérivées q'_k de q_k par rapport au temps :

$$T=\sum_{i,\ k=1}^{n} a_{ik}q'_iq'_k \qquad (a_{ki}=a_{ik}), \tag{165}$$

en outre les coefficients sont des fonctions données des q_k. Supposons que pour $q_k=0$ les dérivées partielles s'annulent :

$$\frac{\partial U}{\partial q_k}=0 \text{ pour } q_1=\dots=q_n=0 \quad (k=1,2,\ \dots,\ n). \tag{166}$$

Le système d'équations différentielles (164) a la solution évidente $q_1=\dots$ $\dots=q_n=0$ à laquelle correspond une position d'équilibre du système. La fonction U n'est définie qu'à une constante près et nous pouvons toujours supposer qu'elle s'annule pour $q_1=\dots=q_n=0$. En vertu de la condition (166) on peut donc affirmer que le développement de la fonction U suivant les puissances des q_k ne commence que par les termes du second ordre. Supposons que la forme quadratique qui s'obtient des termes du second ordre soit définie négative, il en découle que U a un maximum pour $q_1=\dots=q_n=0$ ou, ce qui revient au même, que l'énergie potentielle $(-U)$ a un minimum. Comme nous l'avons démontré (tome II, [I-2-6]), cette position d'équilibre $q_1=\dots$ $\dots=q_n=0$ est stable et pour de petites perturbations initiales le système effectue de petites oscillations autour de la position d'équilibre de sorte que les q_k restent petits pendant toute la durée du mouvement. C'est pourquoi pendant l'étude de ces petites oscillations nous pouvons supposer que U se réduit aux termes du second ordre, c'est-à-dire qu'elle est de la forme :

$$-U=\sum_{i,\ k=1}^{n} b_{ik}q_iq_k \quad (b_{ki}=b_{ik}). \tag{167}$$

De la même façon nous pouvons prendre approximativement $q_k=0$ dans les coefficients a_{ik} de l'expression (165), après quoi ces coefficients deviennent des nombres donnés. Après toutes ces hypothèses simplificatrices le système

(164) se ramène à un système de n équations linéaires à coefficients constants :

$$\left.\begin{array}{l} a_{11}q_1'' + a_{12}q_2'' + \ldots + a_{1n}q_n'' + b_{11}q_1 + b_{12}q_2 + \ldots + b_{1n}q_n = 0, \\ a_{21}q_1'' + a_{22}q_2'' + \ldots + a_{2n}q_n'' + b_{21}q_1 + b_{22}q_2 + \ldots + b_{2n}q_n = 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{n1}q_1'' + a_{n2}q_2'' + \ldots + a_{nn}q_n'' + b_{n1}q_1 + b_{n2}q_2 + \ldots + b_{nn}q_n = 0. \end{array}\right\} \quad (168)$$

Si nous cherchons la solution de ce système sous forme d'oscillations harmoniques de mêmes fréquence et phase initiale, mais d'amplitudes distinctes :

$$q_k = A_k \cos(\lambda t + \varphi) \qquad (k = 1, 2, \ldots, n), \qquad (169)$$

portant alors dans le système (168) nous avons un système d'équations pour déterminer les A_k et λ :

$$\left.\begin{array}{l} (b_{11} - \lambda^2 a_{11}) A_1 + (b_{12} - \lambda^2 a_{12}) A_2 + \ldots + (b_{1n} - \lambda^2 a_{1n}) A_n = 0, \\ (b_{21} - \lambda^2 a_{21}) A_1 + (b_{22} - \lambda^2 a_{22}) A_2 + \ldots + (b_{2n} - \lambda^2 a_{2n}) A_n = 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ (b_{n1} - \lambda^2 a_{n1}) A_1 + (b_{n2} - \lambda^2 a_{n2}) A_2 + \ldots + (b_{nn} - \lambda^2 a_{nn}) A_n = 0. \end{array}\right\} \quad (170)$$

Pour que ce système ait une solution en A_k non nulle, nous devons annuler son déterminant :

$$\begin{vmatrix} b_{11} - \lambda^2 a_{11} & b_{12} - \lambda^2 a_{12} & \ldots & b_{1n} - \lambda^2 a_{1n} \\ b_{21} - \lambda^2 a_{21} & b_{22} - \lambda^2 a_{22} & \ldots & b_{2n} - \lambda^2 a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ b_{n1} - \lambda^2 a_{n1} & b_{n2} - \lambda^2 a_{n2} & \ldots & b_{nn} - \lambda^2 a_{nn} \end{vmatrix} = 0. \qquad (171)$$

En considérant une racine quelconque de cette équation et en la portant dans les coefficients du système (170) nous trouvons pour les A_k des solutions, une ou plusieurs, que nous pouvons ensuite multiplier par une constante arbitraire. La formule (169) contient encore une constante arbitraire φ.

Nous obtenons une solution encore plus précise du problème en appliquant la théorie des formes quadratiques. Avant tout remarquons que la forme quadratique (165), par rapport aux variables q_k', qui exprime l'énergie cinétique en cas de mouvement est, de par la nature même du problème, une forme définie positive. En outre, dans le cas donné, d'après les conditions du problème, la forme (167) est aussi définie positive. Comme nous l'avons vu, on peut introduire à la place des variables q_k de nouvelles variables p_k liées aux anciennes par une transformation linéaire à coefficients constants telles que dans les nouvelles variables les formes T et $(-U)$ soient réduites simultanément à des sommes de carrés, de plus, la forme T doit être une somme de carrés de coefficients égaux à l'unité. Notons, en outre, que la relation linéaire entre les p_k et les q_k conduit à la même relation entre les p_k' et les q_k'. Nous avons donc :

$$T = \sum_{s=1}^{n} p_s'^2 ; \quad -U = \sum_{s=1}^{n} \lambda_s^2 p_s^2, \qquad (172)$$

où tous les coefficients des p_s^2 sont positifs, ce qui nous permet de les désigner par un nombre élevé au carré. Nous pouvons écrire à la place du système (168) l'équation de Lagrange (164) dans les nouvelles coordonnées :

$$\frac{d}{dt}\left[\frac{\partial T}{\partial p_k'}\right] = \frac{\partial U}{\partial p_k}.$$

En y substituant (172) nous avons un système très simple :

$$p_k'' + \lambda_k^2 p_k = 0 \qquad (k = 1, 2, \ldots, n).$$

Les solutions de ce système sont :

$$p_k = C_k \cos(\lambda_k t + \psi_k) \qquad (k = 1, 2, \ldots, n), \tag{173}$$

où C_k et ψ_k sont des constantes arbitraires. Les coordonnées généralisées p_k sont dites coordonnées principales du système mécanique.

Les coordonnées fondamentales q_k s'expriment linéairement par les principales à l'aide de coefficients constants. Il découle immédiatement des résultats de [II-2-6] que les nombres λ_k doivent être racines des équations (170). Notons qu'il peut y avoir aussi des racines égales, mais là aussi les formes (169) donnent la solution générale du problème des petites oscillations dans le cas considéré.

II-2-8. Propriétés extrémales des valeurs propres d'une forme quadratique. Considérons sous un autre angle le problème de la réduction d'une forme quadratique réelle à une somme de carrés. Pour plus de commodité, limitons-nous au cas de trois variables :

$$\varphi = \sum_{i,\,k=1}^{3} a_{ik} x_i x_k = \sum_{k=1}^{3} \lambda_k x_k'^2, \tag{174}$$

où les x_k sont données par une transformation orthogonale sur les x_k' :

$$\left.\begin{aligned} x_1 &= b_{11}x_1' + b_{12}x_2' + b_{13}x_3', \\ x_2 &= b_{21}x_1' + b_{22}x_2' + b_{23}x_3', \\ x_3 &= b_{31}x_1' + b_{32}x_2' + b_{33}x_3'. \end{aligned}\right\} \tag{175}$$

Pour fixer les idées, supposons que les nombres λ_k soient ordonnés par ordre de grandeurs décroissantes, c'est-à-dire :

$$\lambda_1 > \lambda_2 > \lambda_3. \tag{176}$$

Notre problème consiste à déterminer les nombres λ_k et les coefficients b_{ik} d'après les valeurs de la forme φ sur la sphère unité K, c'est-à-dire sur la sphère ayant son centre à l'origine des coordonnées et de rayon unité :

$$x_1^2 + x_2^2 + x_3^2 = 1 \quad \text{ou} \quad x_1'^2 + x_2'^2 + x_3'^2 = 1. \tag{177}$$

Chaque point de cette sphère caractérise une direction dans l'espace, définie par un vecteur unité issu de l'origine des coordonnées et passant par le point envisagé. Nous pouvons écrire la formule (174) sous la forme :

$$\varphi = \lambda_1 (x_1'^2 + x_2'^2 + x_3'^2) + (\lambda_2 - \lambda_1) x_2'^2 + (\lambda_3 - \lambda_1) x_3'^2,$$

d'où l'on voit que, sur la sphère unité K considérée, nous avons :

$$\varphi = \lambda_1 + (\lambda_2 - \lambda_1) x_2'^2 + (\lambda_3 - \lambda_1) x_3'^2.$$

Il en découle immédiatement que λ_1 est un maximum de φ sur K. Il est clair que ce maximum est atteint au point

$$x_1' = 1 \,; \qquad x_2' = x_3' = 0,$$

ou, en vertu de (175), dans les anciennes coordonnées, on voit que c'est un point de la sphère K de coordonnées :

$$x_1 = b_{11} \,; \qquad x_2 = b_{21} \,; \qquad x_3 = b_{31}.$$

Ce point définit un vecteur qui correspond à la première colonne de la transformation orthogonale (175), c'est-à-dire un vecteur solution de l'équation

$$Ax = \lambda x \tag{178}$$

pour $\lambda = \lambda_1$. Ainsi, la plus grande valeur propre de la forme quadratique (174) est égale au maximum de la valeur de cette forme sur la sphère unité et le vecteur propre $x^{(1)}$ correspondant, solution de l'équation (178), est un vecteur issu de l'origine et qui passe par le point de la sphère unité où le maximum en question est atteint.

Passons maintenant à la détermination de la deuxième valeur propre et du vecteur propre correspondant. Faisons dans la formule $x_1' = 0$. Un plan passant par l'origine et perpendiculaire au vecteur $x^{(1)}$ correspond à cette équation. L'intersection de ce plan et de la sphère unité est la circonférence:

$$x_2'^2 + x_3'^2 = 1.$$

Nous avons sur cette circonférence:

$$\varphi = \lambda_2 x_2'^2 + \lambda_3 x_3'^2,$$

d'où l'on voit immédiatement que λ_2 est le maximum de φ sur la sphère unité à condition que les vecteurs correspondants soient perpendiculaires au vecteur déjà déterminé $x^{(1)}$. De la même façon que ci-dessus, démontrons que le vecteur propre $x^{(2)}$, c'est-à-dire la solution de l'équation (178) pour $\lambda = \lambda_2$, est un vecteur qui passe par le point où le maximum est atteint.

Connaissant deux vecteurs, nous déterminons le troisième $x^{(3)}$ comme étant perpendiculaire aux deux premiers et la valeur propre λ_3 est la valeur de la forme φ au point de la sphère unité, intersection de cette dernière avec le vecteur $x^{(3)}$.

Si, par exemple, nous avions $\lambda_1 = \lambda_2$, alors, lors de la détermination du premier maximum de la forme φ sur la sphère unité, nous n'aurions pas eu un point, mais toute une circonférence sur laquelle ce maximum est atteint.

Le raisonnement précédent se généralise facilement au cas d'un nombre quelconque fini de dimensions. Pour cette raison, nous ne donnerons dans ce cas général que le résultat analogue au précédent. Soit une forme quadratique réelle de n variables:

$$\varphi = \sum_{i,\,k=1}^{n} a_{ik} x_i x_k. \tag{179}$$

Un vecteur unité dans un espace réel à n dimensions est représenté par un ensemble de nombres réels dont la somme des carrés est égale à l'unité. Nous dirons que les extrémités de ces vecteurs se trouvent sur la sphère unité et il est clair que l'équation de cette sphère unité est:

$$x_1^2 + x_2^2 + \ldots + x_n^2 = 1. \tag{180}$$

La plus grande valeur propre de la forme φ est égale au maximum de cette forme sur la sphère unité (180) et le vecteur propre correspondant est défini par le vecteur $x^{(1)}$ issu de l'origine et passant par le point de la sphère où φ atteint son maximum. Pour déterminer la valeur propre, la deuxième en grandeur, considérons les vecteurs unités perpendiculaires au vecteur déjà connu $x^{(1)}$. Il existe parmi eux un vecteur $x^{(2)}$ qui donne la valeur la plus grande de la forme φ. Ce deuxième maximum λ_2 est précisément la deuxième valeur propre de la forme et le vecteur $x^{(2)}$ le vecteur propre correspondant. Considérons maintenant des vecteurs unités perpendiculaires à $x^{(1)}$ et $x^{(2)}$, ce qui revient à ajouter à (180) les deux conditions:

$$(\mathbf{x}^{(1)}, \mathbf{x}) = 0 \text{ et } (\mathbf{x}^{(2)}, \mathbf{x}) = 0.$$

Il existe parmi eux un vecteur pour lequel φ prend de nouveau sa valeur la plus grande. Cette valeur est précisément la troisième valeur propre en grandeur de la forme λ_3 et le vecteur correspondant le troisième vecteur propre, etc.

Nous aurions pu ordonner les valeurs propres d'une forme quadratique non pas dans l'ordre décroissant, mais dans celui croissant de sorte que la première valeur propre eût été la plus petite et la suivante la seconde dans l'ordre croissant, etc. Alors, nous aurions obtenu des problèmes entièrement analogues aux précédents, mais partout où il était question de plus grande valeur, nous aurions dû dire « plus petite valeur ».

Tous les raisonnements précédents se généralisent également au cas de la réduction simultanée de deux formes quadratiques à une somme de carrés. Soient deux formes quadratiques:

$$\varphi = \sum_{i,\,k=1}^{n} a_{ik} x_i x_k\,; \qquad \psi = \sum_{i,\,k=1}^{n} b_{ik} x_i x_k.$$

qui se réduisent à une somme de carrés

$$\varphi = \sum_{k=1}^{n} x_k'^2\,; \qquad \psi = \sum_{k=1}^{n} \lambda_k x_k'^2,$$

à l'aide de la transformation linéaire:

$$(x_1, \ldots, x_n) = B\,(x_1', \ldots, x_n'),$$

nous supposons en outre que les nombres λ_k sont ordonnés dans l'ordre décroissant.

Or, λ_1 est la plus grande valeur de ψ si $\varphi = 1$, de plus, cette plus grande valeur est atteinte précisément pour

$$x_1 = b_{11}\,; \quad x_2 = b_{21}\,; \;\ldots\,; \; x_n = b_{n1}.$$

On définit de la même manière les valeurs propres suivantes.

II-2-9. Matrices et formes hermitiennes. Dans les paragraphes précédents, nous avons considéré des matrices symétriques réelles et nous avons remarqué qu'elles sont un cas particulier des matrices hermitiennes dont les éléments sont des nombres complexes vérifiant la relation:

$$a_{ki} = \bar{a}_{ik}. \qquad (181)$$

Pour $i = k$ cette relation montre que les éléments diagonaux a_{kk} sont réels.

On peut énoncer d'une autre manière la définition de la matrice hermitienne: une matrice hermitienne ne change pas si l'on échange entre elles les lignes et les colonnes et si l'on remplace tous les éléments par leur conjugué complexe, c'est-à-dire avec les notations de [II 1-7]:

$$\tilde{A}^* = A \text{ ou } \tilde{A} = A. \qquad (182)$$

Comme nous le savons, la matrice $\tilde{A}$ est dite conjuguée hermitienne de A. C'est pour cette raison que les matrices hermitiennes sont encore dites auto-adjointes.

Dans [II-2-1] nous avons montré qu'une matrice hermitienne A satisfait à la relation:

$$(A\mathbf{x}, \mathbf{y}) = (\mathbf{x}, A\mathbf{y}) \tag{183}$$

quels que soient les vecteurs $\mathbf{x}$ et $\mathbf{y}$.

Cette relation, ainsi que les deux précédentes, peut servir de définition pour les matrices hermitiennes.

Donnons encore une propriété des matrices hermitiennes.

Soient A une matrice hermitienne et U une matrice unitaire quelconque. Il est facile de montrer que $U^{-1}AU$, tout comme A, est une matrice hermitienne. Par hypothèse, $\overline{A}^{(*)} = A$. Il faut démontrer que $U^{-1}AU$ possède cette propriété. Nous avons de [II-1-7]:

$$\overline{(U^{-1}AU)}^{(*)} = \overline{U}^{(*)}\overline{A}^{(*)}\overline{U}^{(*)-1}$$

ou, tenant compte de la condition sur A et de ce que U est une matrice unitaire, il en découle $\overline{U}^{(*)} = U^{-1}$, nous en déduisons:

$$\overline{(U^{-1}AU)}^{(*)} = U^{-1}AU,$$

ce qu'il fallait démontrer.

On peut dire que pour une transformation unitaire des coordonnées, agissant sur les composantes d'un vecteur à l'aide de la formule:

$$(x_1, \ldots, x_n) = U(x_1', \ldots, x_n'),$$

la matrice hermitienne A, en tant qu'opérateur d'une transformation linéaire d'un espace, est dans les nouvelles coordonnées de la forme $U^{-1}AU$ et, par conséquent, la proposition qui vient d'être démontrée peut être énoncée comme suit: les transformations unitaires agissant dans un espace conservent l'hermicité des matrices considérées comme opérateur.

Posons à présent le problème de la réduction d'une matrice hermitienne à la forme diagonale à l'aide d'une transformation unitaire:

$$U^{-1}AU = [\lambda_1, \ldots, \lambda_n]. \tag{184}$$

Tout comme précédemment pour les matrices symétriques réelles, le problème est équivalent à celui de la résolution d'une équation de la forme:

$$A\mathbf{x} = \lambda\mathbf{x}, \tag{185}$$

où λ et l'un des nombres λ_k et les composantes du vecteur $\mathbf{x}$ sont les éléments de la colonne correspondante de la matrice U.

Les nombres λ_k et les vecteurs $\mathbf{x}^{(k)}$ qui leur correspondent sont appelés valeurs propres et vecteurs propres de la matrice A.

Comme nous le savons, les valeurs propres doivent être nécessairement les racines de l'équation:

$$\begin{vmatrix} a_{11}-\lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22}-\lambda & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn}-\lambda \end{vmatrix} = 0. \tag{186}$$

Soient $\lambda = \lambda_1$ une racine quelconque de cette équation et $\mathbf{x}^{(1)}$ la solution de l'équation (185) pour $\lambda = \lambda_1$.

Cette équation est linéaire et homogène de sorte que sa solution peut être multipliée par une constante quelconque, ce qui nous permet de supposer que la longueur du vecteur $\mathbf{x}^{(1)}$ est égale à l'unité. Prenons ce vecteur pour premier vecteur de base d'un nouveau système de coordonnées et complétons-le d'une manière quelconque par $(n-1)$ vecteurs unités de façon à obtenir un système orthonormé de vecteurs. Prenons ces vecteurs pour nouveaux vecteurs de base et soit U_1 la transformation unitaire qui correspond au passage des anciens vecteurs de base aux nouveaux. Dans le nouveau système de coordonnées, la matrice hermitienne A se transforme en une nouvelle matrice hermitienne $A_1 = U_1^{-1} A U_1$, en outre l'équation correspondante

$$A_1\mathbf{x} = \lambda\mathbf{x}$$

doit pour $\lambda = \lambda_1$ avoir pour solution le vecteur de composantes $(1, 0, \dots, 0)$. Tout comme dans [II-2-2], ceci nous montre que tous les éléments de la première ligne et de la première colonne de la matrice A_1 s'annuleront, excepté l'élément qui se trouve à l'intersection de la première ligne et de la première colonne, qui est égal à λ_1.

Puisque la matrice A_1 est hermitienne, il en découle immédiatement que l'élément λ_1 doit être un nombre réel, d'où, entre autres, il s'ensuit que toutes les racines de l'équation (186) sont réelles, ce que nous savions déjà. Ainsi, la matrice A_1 est de la forme:

$$\begin{Vmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & a_{22}^{(1)} & \dots & a_{2n}^{(1)} \\ \dots & \dots & \dots & \dots \\ 0 & a_{n2}^{(1)} & \dots & a_{nn}^{(1)} \end{Vmatrix},$$

c'est-à-dire elle est une matrice quasi diagonale de la forme:

$$[\lambda_1, C_1],$$

où par C_1 nous désignons la matrice hermitienne d'ordre $(n-1)$ et d'éléments $a_{ik}^{(1)}$. En répétant le raisonnement précédent, nous pouvons à l'aide d'une transformation unitaire U_2 sur les vecteurs de base, à l'exception du premier, réduire la matrice C_1 à une forme telle que tous les éléments de sa première ligne et de sa première colonne soient nuls, excepté l'élément qui se trouve à leur intersection.

Nous pouvons considérer la transformation unitaire envisagée comme une transformation unitaire sur l'espace à n dimensions. Ce sera une matrice unitaire quasi diagonale de la forme:

$$[1,\ U_2].$$

Le résultat de cette transformation est que notre matrice hermitienne se transforme en une matrice hermitienne de la forme:

$$[1,\ U_2]^{-1}\,[\lambda_1,\ C_1]\,[1,\ U_2] = [\lambda_1,\ U_2^{-1}C_1U_2],$$

et sous forme développée cette matrice est:

$$\begin{vmatrix} \lambda_1 & 0 & 0 & \dots & 0 \\ 0 & \lambda_2 & 0 & \dots & 0 \\ 0 & 0 & a_{33}^{(2)} & \dots & a_{3n}^{(2)} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & a_{n3}^{(2)} & \dots & a_{nn}^{(2)} \end{vmatrix}.$$

Continuant à raisonner ainsi nous ramenons, en fin de compte, la matrice hermitienne à la forme diagonale ; en outre, la transformation unitaire générale U qui intervient dans la formule (184) est le produit de transformations unitaires réalisées pendant le processus indiqué plus haut de la réduction successive de la matrice à la forme diagonale.

Considérons maintenant l'équation (185). Dans [II-2-2] nous avons montré que les solutions qui correspondent à des valeurs distinctes de λ sont nécessairement orthogonales.

De même que dans [II-2-2] nous pouvons montrer que les vecteurs représentés par les colonnes de la matrice U avec les valeurs correspondantes de λ donnent toutes les solutions de l'équation (185). Il est seulement nécessaire d'avoir présente en mémoire une règle importante concernant les racines multiples de l'équation (186). Par exemple, si l'équation (186) a une racine $\lambda = \lambda_1$, de multiplicité m, pour $\lambda = \lambda_1$ l'équation (185) a alors m solutions linéairement indépendantes $\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(m)}$. Il est clair que toute leur combinaison linéaire à coefficients arbitraires est également solution de l'équation (185), c'est-à-dire que l'équation

$$A\mathbf{x} = \lambda_1\mathbf{x}$$

a un ensemble de solutions qui n'est autre que le sous-espace engendré par les vecteurs $\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(m)}$, solutions données sous forme de somme :

$$x = C_1 \mathbf{x}^{(1)} + \ldots + C_m \mathbf{x}^{(m)},$$

à coefficients arbitraires $C_1, \ldots, C_m$. Nous pouvons choisir d'une manière arbitraire dans ce sous-espace un système orthonormé de m vecteurs dont les composantes sont précisément les colonnes de la matrice U qui correspondent à la valeur propre $\lambda = \lambda_1$. Nous avons donc ici le même degré d'arbitraire dans le choix de la matrice U que nous avions dans [II-2-2] pour B. En outre, nous pouvons multiplier les composantes de chaque vecteur $\mathbf{x}^{(i)}$, résultat de la résolution de l'équation (185), par une quantité de module égal à l'unité, c'est-à-dire par une quantité de la forme $e^{i\varphi}$ (facteur de phase). De plus, la longueur du vecteur reste égale à l'unité et la condition d'orthogonalité de ce vecteur avec les autres vecteurs du système complet de solutions de l'équation (185) est conservée. Enfin, nous pouvons dans la matrice U changer d'une manière arbitraire l'ordre des colonnes. Il est évident que cette transformation peu importante se ramène à un changement de numérotation des vecteurs de base dans le nouveau système de coordonnées et qu'elle n'entraîne que la permutation des nombres λ_k dans la forme diagonale de la matrice. Dans la suite, on supposera toujours que ces nombres sont ordonnés dans l'ordre croissant.

Passons maintenant à la considération des formes hermitiennes. Nous dirons qu'une matrice hermitienne A correspond à une *forme hermitienne* :

$$(A\mathbf{x}, \mathbf{x}) = \sum_{i,\,k=1}^{n} a_{ik} \bar{x}_i x_k, \tag{187}$$

où $x_1, \ldots, x_n$ sont les composantes du vecteur $\mathbf{x}$, le nombre $(A\mathbf{x}, \mathbf{x})$ étant réel [II-2-1].

Supposons maintenant que nous ayons transformé notre espace à l'aide d'une transformation unitaire ; les anciennes composantes d'un vecteur s'expriment à l'aide des nouvelles d'après la formule : $\mathbf{x} = U\mathbf{x}'$. Avec les nouvelles coordonnées, la forme hermitienne (187) est :

$$(AU\mathbf{x}', U\mathbf{x}').$$

Tenant compte de la propriété (125_1) des transformations unitaires, nous pouvons multiplier à gauche par la matrice unitaire U^{-1} les deux vecteurs qui engendrent le produit scalaire, et nous obtenons ainsi, dans les nouvelles coordonnées, l'expression suivante pour la forme hermitienne :

$$(U^{-1}AU\mathbf{x}', \mathbf{x}'). \tag{188}$$

En particulier, si une transformation unitaire U réduit une matrice A à la forme diagonale, c'est-à-dire que (184) est vérifiée, alors dans les nouvelles coordonnées notre forme hermitienne ne contient que les termes contenant les produits $\bar{x}'_i \cdot x'_i$ et notre forme hermitienne se réduit à la somme des modules au carré:

$$(U^{-1}AU\mathbf{x}',\ \mathbf{x}') = \lambda_1 \bar{x}'_1 x'_1 + \lambda_2 \bar{x}'_2 x'_2 + \ldots + \lambda_n \bar{x}'_n x'_n.$$

Donc, ici, tout comme dans [II-2-1], le problème de la réduction d'une matrice A à la forme diagonale est équivalent à celui de la réduction de la forme hermitienne correspondante à la somme des modules au carré.

Au lieu des formes hermitiennes, on considère parfois des formes dites *bilinéaires* définies de la manière suivante:

$$(A\mathbf{y},\ \mathbf{x}) = \sum_{i,\,k=1}^{n} a_{ik} \bar{x}_i y_k.$$

Si une transformation unitaire agit dans un espace de telle sorte que les nouvelles composantes soient exprimées à l'aide des anciennes d'après les formules précédentes, nous avons alors dans les nouvelles coordonnées:

$$(A\mathbf{y},\ \mathbf{x}) = (AU\mathbf{y}',\ U\mathbf{x}')$$

ou, en vertu d'une propriété des transformations unitaires:

$$(U^{-1}AU\mathbf{y}',\ \mathbf{x}').$$

Enfin, si U réduit A à la forme diagonale, alors dans les coordonnées correspondantes, la forme bilinéaire se ramène à la forme simple suivante:

$$\sum_{k=1}^{n} \lambda_k \bar{x}'_k y'_k.$$

Remarquons que toute matrice diagonale à éléments réels est une matrice hermitienne, donc, la matrice $U^{-1}[\lambda_1, \ldots, \lambda_n]\, U$, où U est une matrice unitaire quelconque, est aussi hermitienne. Précédemment, nous avons vu que inversement toute matrice hermitienne peut être mise sous cette forme.

On classe les formes hermitiennes, de même que les formes quadratiques réelles [II-2-4], d'après le signe des valeurs propres λ_k. Par exemple, si tous les λ_k sont positifs, la forme hermitienne est dite définie positive. Sa propriété caractéristique est que les valeurs qu'elle prend sont positives pour tous les x_k et qu'elle ne peut s'annuler que pour $x_1 = \ldots = x_n = 0$. On définit d'une façon analogue les formes hermitiennes semi-définies et celles indéterminées. L'étude est semblable à celle des formes quadratiques réelles et est basée sur la formule:

$$(A\mathbf{x},\ \mathbf{x}) = \lambda_1 \bar{x}'_1 x'_1 + \ldots + \lambda_n \bar{x}'_n x'_n.$$

La formule (183) est vraie pour les matrices hermitiennes. Si A est une matrice quelconque et $\tilde{A} = \overline{A}^{(*)}$ la matrice conjuguée, nous avons

$$(Ax, y) = (x, \tilde{A}y) \qquad (183_1)$$

au lieu de (183).

Si les a_{ik} sont les éléments de la matrice A, ceux de la matrice $\tilde{A}$ sont alors $\{\tilde{A}\}_{ik} = \bar{a}_{ki}$ et la formule (183_1) est vérifiée directement par substitution de la même manière que la formule (183).

II-2-10. Matrices hermitiennes commutatives. Soient A et B deux matrices hermitiennes. Cherchons les conditions pour lesquelles le produit BA est une matrice hermitienne. Formons la matrice conjuguée du produit BA :

$$(\overline{BA})^{(*)} = \overline{A}^{(*)}\overline{B}^{(*)}$$

ou, en vertu du fait que les matrices A et B sont hermitiennes :

$$(\overline{BA})^{(*)} = AB.$$

Pour que BA soit une matrice hermitienne, il faut et il suffit que AB coïncide avec BA, c'est-à-dire que *les matrices soient commutatives*. Supposons que les matrices hermitiennes A et B se réduisent à la forme diagonale à l'aide d'une transformation unitaire U :

$$A = U^{-1}[\lambda_1, \ldots, \lambda_n]\, U; \quad B = U^{-1}[\mu_1, \ldots, \mu_n]\, U.$$

Il est facile de voir que dans ce cas elles sont commutatives :

$$AB = BA = U^{-1}[\lambda_1\mu_1, \ldots, \lambda_n\mu_n]\, U.$$

Inversement, montrons maintenant que si deux matrices hermitiennes sont commutatives, on peut alors, à l'aide d'une même transformation unitaire, les réduire simultanément à la forme diagonale, c'est-à-dire que la commutativité des matrices hermitiennes n'est pas seulement une condition nécessaire, mais aussi suffisante permettant de les réduire simultanément à la forme diagonale à l'aide d'une transformation unitaire. Ainsi, supposons que $AB = BA$. Remarquons que les matrices qui leur sont semblables sont aussi commutatives. En effet :

$$(C^{-1}AC)(C^{-1}BC) = C^{-1}ABC = C^{-1}BAC,$$

une expression analogue s'obtient pour le produit :

$$(C^{-1}BC)(C^{-1}AC).$$

Supposons que l'on prenne pour C une transformation unitaire réduisant A à la forme diagonale, et soumette la matrice B à la même transformation. Les nouvelles matrices sont aussi commutatives et, au cours de la démonstration de notre proposition, nous pouvons simplement supposer que la matrice A a déjà la forme diagonale, c'est-à-dire que les éléments a_{ik} satisfont à la condition:

$$a_{ik}=0 \text{ pour } i\neq k. \tag{189}$$

Désignons par b_{ik} les éléments de la matrice B et écrivons la condition de commutativité des matrices:

$$\sum_{s=1}^{n} a_{is}b_{sk}=\sum_{s=1}^{n} b_{is}a_{sk} \quad (i,\ k=1,\ 2,\ \ldots,\ n).$$

En vertu de (189), ces conditions sont de la forme:

$$(a_{ii}-a_{kk})\,b_{ik}=0 \quad (i,\ k=1,\ 2,\ \ldots,\ n). \tag{190}$$

Si tous les nombres a_{ii} sont distincts, il découle immédiatement des égalités (190) que $b_{ik}=0$ pour $i\neq k$, c'est-à-dire que la matrice B a également la forme diagonale et notre proposition est démontrée.

Considérons maintenant le cas général où parmi les nombres a_{ii} certains sont égaux. Pour fixer les idées, supposons que ces nombres se divisent en deux groupes de nombres égaux:

$$a_{11}=\ldots=a_{mm};\ a_{m+1,\,m+1}=\ldots=a_{nn}.$$

Il découle immédiatement de la formule (190) que, dans ce cas, les éléments b_{ik} peuvent être distincts de zéro seulement si les deux indices i et k sont supérieurs à m ou bien s'ils sont tous les deux inférieurs à m. Ainsi, dans le cas donné, la matrice B a la forme quasi diagonale:

$$B=[B_1,\ B_2],$$

où B_1 et B_2 sont des matrices hermitiennes respectivement d'ordre m et $(n-m)$. Sous forme développée nous pouvons écrire la matrice B comme suit:

$$\begin{Vmatrix} b_{11} \ldots b_{m} & 0 & \ldots & 0 \\ \cdots\cdots\cdots & \cdots & \cdots & \cdots \\ b_{m1} \ldots b_{mm} & 0 & \ldots & 0 \\ 0 \ldots 0 & b_{m+1,\,m+1} & \ldots & b_{m+1,\,n} \\ \cdots\cdots\cdots & \cdots & \cdots & \cdots \\ 0 \ldots 0 & b_{n,\,m+1} & \ldots & b_{nn} \end{Vmatrix}.$$

Nous pouvons, sans changer la forme diagonale de A, faire subir au sous-espace, engendré par les m premiers vecteurs de base, une transformation unitaire arbitraire, et la même chose reste valable pour le sous-espace engendré par les $(n-m)$ derniers vecteurs de base. Choisissons les transformations unitaires V_1 et V_2 de façon que les matrices B_1 et B_2 s'écrivent sous la forme diagonale. D'une manière générale, nous avons une transformation unitaire de tout l'espace à n dimensions, qui a la forme quasi diagonale:

$$[V_1 V_2].$$

En vertu de ce qui vient d'être dit, dans les nouvelles coordonnées la matrice A conserve la forme diagonale et la matrice B prend la forme:

$$[V_1, V_2]^{-1}[B_1, B_2][V_1, V_2] = [V_1^{-1}B_1V_1, V_2^{-1}B_2V_2],$$

c'est-à-dire qu'elle a aussi la forme diagonale, et notre proposition est démontrée.

Si à présent nous construisons pour nos matrices commutatives les équations:

$$A\mathbf{x} = \lambda\mathbf{x}; \qquad B\mathbf{x} = \mu\mathbf{x}, \tag{191}$$

il découle immédiatement de ce qui a été dit ci-dessus que nous pouvons construire un même système de n solutions linéairement indépendantes pour les deux équations. Ces solutions sont les colonnes de la matrice unitaire U qui réduit les deux matrices à la forme diagonale. Autrement dit, nous pouvons pour deux matrices hermitiennes commutatives construire un même système complet de n vecteurs propres linéairement indépendants. En ce qui concerne les valeurs propres, c'est-à-dire les valeurs des paramètres λ et μ, elles sont en général toutes distinctes. Notons que, de ce qui précède, il ne découle pas encore que tout vecteur propre de la matrice A soit précisément un vecteur propre de la matrice B. Si toutes les valeurs propres de A et B sont distinctes, telles qu'à chaque valeur λ_k et μ_k corresponde, à un facteur numérique près, un seul vecteur, il en est alors évidemment ainsi. Mais cela n'a pas lieu, en général, si parmi les valeurs propres il y en a d'égales. Soient $\mathbf{x}^{(k)}$ un système complet de vecteurs propres des matrices A et B et λ_k et μ_k les valeurs propres correspondantes. Supposons, par exemple, que $\lambda_1 = \lambda_2$, mais $\mu_1 \neq \mu_2$. En outre, les vecteurs $C_1\mathbf{x}^{(1)} + C_2\mathbf{x}^{(2)}$, pour un choix arbitraire des constantes C_1 et C_2, sont des vecteurs propres de A, mais non pas de B.

Tous les raisonnements précédents se transposent facilement au cas de plusieurs matrices: si nous avons plusieurs matrices hermitiennes $A_1, \ldots, A_l$, pour qu'on puisse les ramener simultanément à la forme diagonale, à l'aide d'une transformation unitaire, il faut et il suffit qu'elles soient commutatives deux à deux, c'est-à-dire que $A_iA_k = A_kA_i$ pour tous i et k de 1 jusqu'à l.

II-2-11. Réduction des matrices unitaires à la forme diagonale. En relation avec leur réduction à la forme diagonale, les matrices unitaires possèdent une propriété tout à fait analogue à celle des matrices hermitiennes: si V est une matrice unitaire, on peut toujours trouver une matrice unitaire U telle que la matrice

$$U^{-1}VU$$

soit une matrice diagonale. Nous pouvons écrire notre problème sous la forme suivante:

$$VU = U[\lambda_1, \ldots, \lambda_k], \tag{192}$$

où U est la matrice unitaire et λ_k les nombres que l'on cherche.

Comme précédemment pour la matrice hermitienne, aux colonnes de la matrice U correspondent des vecteurs $\mathbf{x}^{(k)}$ et ces vecteurs doivent être des solutions de l'équation:

$$V\mathbf{x} = \lambda\mathbf{x}, \tag{193}$$

où les λ coïncident avec les nombres λ_k. D'où il découle immédiatement que ces nombres λ doivent être racines de l'équation caractéristique:

$$\begin{vmatrix} v_{11}-\lambda & v_{12} & \ldots & v_{1n} \\ v_{21} & v_{22}-\lambda & \ldots & v_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ v_{n1} & v_{n2} & \ldots & v_{nn}-\lambda \end{vmatrix} = 0, \tag{194}$$

où les v_{ik} sont les éléments de la matrice V.

Remarquons d'abord que si les matrices V_1 et U_1 sont unitaires, la matrice $U_1^{-1}V_1U_1$ l'est aussi. En effet, du fait que U_1 est unitaire, il découle que U_1^{-1} est unitaire, et le produit de matrices unitaires est aussi une matrice unitaire.

Considérons une racine $\lambda = \lambda_1$ de l'équation (194); la substituant à λ dans l'équation (193), définissons le vecteur unité $\mathbf{x}^{(1)}$ satisfaisant à cette équation, prenons ce vecteur pour nouveau vecteur de base et adjoignons-lui $(n-1)$ vecteurs unités de façon à obtenir n vecteurs unités orthogonaux. Le passage des anciens vecteurs de base aux nouveaux est équivalent à une certaine transformation unitaire U_1 et la matrice unitaire V se transforme en la matrice semblable:

$$V_1 = U_1^{-1}VU_1.$$

L'équation correspondante

$$V_1\mathbf{x} = \lambda\mathbf{x}$$

a, pour $\lambda = \lambda_1$, comme solution un vecteur de composantes $(1, 0, \ldots, 0)$, d'où il s'ensuit immédiatement, comme précédemment, que les éléments de la première colonne de la matrice V_1 sont tous nuls, excepté le premier qui est égal à λ_1. Puisque dans une matrice unitaire la somme des carrés des modules des éléments de chaque colonne est égale à l'unité, nous pouvons affirmer que le nombre λ_1 a module égal à l'unité. Rappelons maintenant que dans une matrice unitaire V_1 la somme des carrés des modules des éléments de chaque ligne est aussi égale à l'unité. Mais, comme on vient de le montrer, le premier élément de la première ligne λ_1 est égal en module à l'unité et, par conséquent, les autres éléments de cette ligne doivent être nuls. Ainsi, le résultat de la première transformation unitaire réduit la matrice unitaire à une forme telle que dans la première ligne et la première colonne tous les éléments sont nuls, excepté le premier:

$$\begin{Vmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & v_{22}^{(1)} & \dots & v_{2n}^{(1)} \\ \cdot & \cdot & \cdot & \cdot \\ 0 & v_{n2}^{(1)} & \dots & v_{nn}^{(1)} \end{Vmatrix},$$

ce que nous avons vu plus haut pour les matrices hermitiennes. Les éléments $v_{ik}^{(1)}$ forment une matrice unitaire d'ordre $(n-1)$. Appliquant encore une fois une transformation unitaire, nous pouvons obtenir des zéros pour la première ligne et la première colonne de cette matrice, excepté pour le premier élément dont le module est égal à l'unité. Finalement, le résultat de deux transformations unitaires ramène la matrice unitaire à la forme:

$$\begin{Vmatrix} \lambda_1 & 0 & 0 & \dots & 0 \\ 0 & \lambda_2 & 0 & \dots & 0 \\ 0 & 0 & v_{33}^{(2)} & \dots & v_{3n}^{(2)} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & v_{n3}^{(2)} & \dots & v_{nn}^{(2)} \end{Vmatrix}.$$

En continuant ce processus nous ramenons la matrice unitaire à l'aide d'une transformation unitaire à la forme diagonale. Notons qu'il découle immédiatement des considérations précédentes que *toutes les valeurs propres d'une matrice unitaire ont le module égal à l'unité.*

De même que dans [II-2-10] on peut montrer que si plusieurs matrices unitaires sont deux à deux commutatives, on peut les mettre sous forme diagonale à l'aide d'une même transformation unitaire.

Notons encore le fait suivant. Soit U une matrice unitaire qui réduit une matrice quelconque A à la forme diagonale, c'est-à-dire soit $U^{-1}AU$ une matrice diagonale. Comme on le sait, le module du déterminant de U est égal à l'unité, et nous pouvons choisir un nombre réel ω de telle sorte que le déterminant de la matrice unitaire $e^{i\omega}U$ soit égal à l'unité. La matrice unitaire $e^{i\omega}U$ réduit aussi A

à la forme diagonale, car:

$$(e^{i\omega}U)^{-1} A (e^{i\omega}U) = e^{i\omega}e^{-i\omega}U^{-1}AU = U^{-1}AU.$$

Ainsi, on peut toujours supposer que le déterminant d'une matrice unitaire U réduisant une matrice quelconque à la forme diagonale est égal à l'unité.

Exemple. Considérons l'exemple de la réduction à la forme diagonale d'une matrice orthogonale réelle du troisième ordre:

$$V = \left\| \begin{matrix} v_{11} & v_{12} & v_{13} \\ v_{21} & v_{22} & v_{23} \\ v_{31} & v_{32} & v_{33} \end{matrix} \right\|. \tag{195}$$

Supposons que le déterminant de cette matrice soit égal à $(+1)$ de sorte que cette matrice correspond à un mouvement de l'espace à trois dimensions autour de l'origine. L'équation caractéristique de la matrice (195) a, par hypothèse, un terme constant égal à l'unité, car ce dernier coïncide, évidemment, avec le déterminant de la matrice. Nous avons vu, d'autre part, que toutes les racines de l'équation caractéristique doivent avoir le module égal à l'unité. Le terme de plus haut degré de l'équation caractéristique est $(-\lambda)^3 = -\lambda^3$ et, par conséquent, le terme constant de l'équation, c'est-à-dire l'unité, est tout simplement égal au produit des racines de cette équation. Etant donné que cette équation a des coefficients réels, seuls deux cas sont possibles: ou bien cette équation a une racine égale à l'unité et les deux autres sont conjuguées complexes de module égal à l'unité, c'est-à-dire que les deux autres racines sont de la forme $e^{\pm i\varphi}$, ou bien l'équation a l'unité pour une racine et les deux autres sont égales à (-1). Le second cas est un cas particulier du premier pour $\varphi = \pi$.

A la valeur propre $\lambda = 1$ correspond un vecteur réel $\mathbf{x}^{(1)}$ qui doit être solution de l'équation:

$$V\mathbf{x}^{(1)} = \mathbf{x}^{(1)}. \tag{196}$$

Autrement dit, ce vecteur reste fixe lors de la rotation de l'espace définie par la matrice V. Ce vecteur, correspondant à la valeur réelle $\lambda = 1$, est un vecteur réel et il définit naturellement l'axe de rotation de l'espace (la rotation de l'espace autour de l'origine des coordonnées est équivalente à une rotation autour d'un axe passant par l'origine des coordonnées). Pour déterminer les composantes de ce vecteur $\mathbf{x}^{(1)}$ en fonction des éléments de V, écrivons l'équation (196) sous la forme:

$$V^{-1}\mathbf{x}^{(1)} = \mathbf{x}^{(1)}$$

ou, la matrice V étant réelle et unitaire, nous pouvons écrire:

$$V^{(*)}\mathbf{x}^{(1)} = \mathbf{x}^{(1)}.$$

Retranchant de (196) cette dernière égalité, nous avons:

$$(V - V^{(*)})\, \mathbf{x}^{(1)} = 0.$$

Ecrivons cette égalité sous forme développée: désignons par (u_{11}, u_{21}, u_{31}) les composantes du vecteur $\mathbf{x}^{(1)}$. Nous obtenons le système d'équations:

$$\left.\begin{aligned} (v_{12}-v_{21})\, u_{21} + (v_{13}-v_{31})\, u_{31} &= 0, \\ (v_{21}-v_{12})\, u_{11} + (v_{23}-v_{32})\, u_{31} &= 0, \\ (v_{31}-v_{13})\, u_{11} + (v_{32}-v_{23})\, u_{21} &= 0, \end{aligned}\right\}$$

d'où il découle immédiatement les formules qui déterminent la direction de l'axe de rotation :

$$u_{11} : u_{21} : u_{31} = (v_{23} - v_{32}) : (v_{31} - v_{13}) : (v_{12} - v_{21}).$$

Les deux autres vecteurs propres $\mathbf{x}^{(2)}$ et $\mathbf{x}^{(3)}$ doivent satisfaire aux équations :

$$V\mathbf{x}^{(2)} = e^{i\varphi}\mathbf{x}^{(2)} \quad \text{et} \quad V\mathbf{x}^{(3)} = e^{-i\varphi}\mathbf{x}^{(3)}, \tag{197}$$

et sont des vecteurs à composantes complexes. Nous pouvons déterminer φ de ce que la somme des racines de l'équation caractéristique est égale à la somme des termes de la diagonale, c'est-à-dire à la trace de la matrice V :

$$1 + e^{-i\varphi} + e^{i\varphi} = 1 + 2\cos\varphi = v_{11} + v_{22} + v_{33},$$

on peut en outre supposer que φ se trouve entre 0 et π.

Il découle des équations (197) que les composantes des vecteurs $\mathbf{x}^{(2)}$ et $\mathbf{x}^{(3)}$ sont conjuguées complexes, puisque les valeurs de λ dans les équations (197) sont des nombres imaginaires conjugués. Formons une nouvelle matrice unitaire :

$$U_0 = \left\| \begin{matrix} 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & \frac{i}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & -\frac{i}{\sqrt{2}} \end{matrix} \right\|. \tag{198}$$

Il est aisé de voir directement que les éléments des colonnes de la matrice $W = UU_0$ sont égaux aux composantes des vecteurs :

$$\mathbf{x}^{(1)}; \qquad \frac{\mathbf{x}^{(2)} + \mathbf{x}^{(3)}}{\sqrt{2}}; \qquad i\,\frac{\mathbf{x}^{(2)} - \mathbf{x}^{(3)}}{\sqrt{2}},$$

c'est-à-dire qu'ils sont réels. En outre, la matrice W, en tant que produit de deux matrices unitaires, est aussi une matrice unitaire, c'est-à-dire que W est une matrice orthogonale. Appliquons maintenant à la matrice V une similitude en ayant recours à une matrice unitaire réelle W. Nous trouvons :

$$W^{-1}VW = U_0^{-1}U^{-1}VUU_0 = U_0^{-1}\,[1,\ e^{i\varphi},\ e^{-i\varphi}]\,U_0.$$

Faisant le produit des matrices, nous obtenons :

$$W^{-1}VW = \left\| \begin{matrix} 1 & 0 & 0 \\ 0 & \cos\varphi & -\sin\varphi \\ 0 & \sin\varphi & \cos\varphi \end{matrix} \right\|. \tag{199}$$

Nous pouvons toujours supposer que le déterminant de la matrice orthogonale W est égal à $(+1)$, car dans le cas contraire nous aurions pu multiplier cette matrice par (-1), ce qui par ailleurs conserve la relation (199). Ainsi, à la matrice W correspond aussi un mouvement de l'espace à trois dimensions. La matrice (199), obtenue comme résultat de la transformation des coordonnées $\mathbf{x} = W\mathbf{x}'$, est semblable à la matrice V et définit dans les nouvelles coordonnées la même transformation que la matrice initiale V définissait dans les anciennes coordonnées. Il découle immédiatement de la forme même de la matrice (199) qu'à cette matrice (199) correspond une rotation autour du nouvel axe $\mathbf{x}'^{(1)}$ d'angle φ et notre transformation se ramène à prendre pour axe $\mathbf{x}'^{(1)}$, mentionné plus haut, l'axe de rotation représenté par le vecteur $\mathbf{x}^{(1)}$.

Il découle de ce qui précède encore une propriété importante, à savoir: toutes les matrices réelles, auxquelles correspond une rotation de l'espace d'un angle donné φ, peuvent être mises à l'aide d'une similitude (différente pour chaque matrice distincte) sous la même forme (199) et, par conséquent, toutes ces matrices sont semblables entre elles.

Les matrices qui correspondent aux différents angles de rotation ne peuvent pas être semblables entre elles, étant donné que les valeurs propres de ces matrices 1, $e^{i\varphi}$ et $e^{-i\varphi}$ sont certainement différentes les unes des autres pour des valeurs distinctes de l'angle φ. Toutes ces propriétés ont une signification géométrique très simple.

II-2-12. Matrices de projection. Considérons maintenant un cas particulier des matrices hermitiennes. Soit R_m un sous-espace de dimension m engendré par m vecteurs linéairement indépendants $\mathbf{y}^{(1)}, \ldots, \mathbf{y}^{(m)}$. Ce sous-espace R_m est l'ensemble des vecteurs de la forme:

$$C_1\mathbf{y}^{(1)}+\ldots+C_m\mathbf{y}^{(m)},$$

où C_k sont des coefficients numériques arbitraires. Orthogonalisant les vecteurs $\mathbf{y}^{(k)}$, nous pouvons construire un système orthonormé de vecteurs $\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(m)}$ qui engendre le même sous-espace R_m. Nous pouvons ensuite le compléter en un système orthonormé complet de n vecteurs en construisant un autre système orthonormé $\mathbf{x}^{(m+1)}, \ldots, \mathbf{x}^{(n)}$.

Ces derniers vecteurs engendrent un sous-espace R'_{n-m} de dimension $(n-m)$, en outre les deux sous-espaces R_m et R'_{n-m} sont orthogonaux en ce sens que chaque vecteur du sous-espace R_m est orthogonal à tout vecteur du sous-espace R'_{n-m} [I-2-7]. Développant un vecteur quelconque $\mathbf{x}$ suivant les vecteurs de base $\mathbf{x}^{(k)}$:

$$\mathbf{x}=x_1\mathbf{x}^{(1)}+\ldots+x_n\mathbf{x}^{(n)}, \tag{200}$$

nous pouvons le représenter sous la forme d'une somme de deux vecteurs:

$$\mathbf{x}=[x_1\mathbf{x}^{(1)}+\ldots+x_m\mathbf{x}^{(m)}]+[x_{m+1}\mathbf{x}^{(m+1)}+\ldots+x_n\mathbf{x}^{(n)}]=\mathbf{u}+\mathbf{v}, \tag{201}$$

dont l'un appartient à R_m et l'autre à R'_{n-m}. Il est facile de voir que ce développement de tout vecteur $\mathbf{x}$ en deux composantes est unique. En effet, supposons qu'en plus du développement (201) nous en ayons un second $\mathbf{x}=\mathbf{u}'+\mathbf{v}'$ vérifiant la propriété mentionnée ci-dessus. Il vient $\mathbf{u}+\mathbf{v}=\mathbf{u}'+\mathbf{v}'$ ou $\mathbf{u}-\mathbf{u}'=$ $=\mathbf{v}'-\mathbf{v}$.

Le vecteur du premier membre appartient à R_m, tandis que celui du second membre appartient à R'_{n-m}, donc, $\mathbf{u}-\mathbf{u}'$ et $\mathbf{v}-\mathbf{v}'$ doivent être orthogonaux.

Mais il est évident que tout vecteur orthogonal à lui-même est un vecteur nul [I-2-7], donc, $\mathbf{u}-\mathbf{u}'=0$, c'est-à-dire que $\mathbf{u}$ coïncide avec $\mathbf{u}'$ et $\mathbf{v}$ avec $\mathbf{v}'$, autrement dit, les vecteurs $\mathbf{u}$ et $\mathbf{v}$ sont déterminés d'une manière unique par le vecteur $\mathbf{x}$. Le vecteur $\mathbf{u}$ est appelé *projection du vecteur* $\mathbf{x}$ *sur le sous-espace* R_m. La matrice qui réalise le passage du vecteur $\mathbf{x}$ au vecteur $\mathbf{u}$ est appelée *matrice de projection sur le sous-espace* R_m, nous la notons P_{R_m}. La forme de cette matrice dépend évidemment du choix des axes de coordonnées.

Si pour vecteurs de base nous choisissons les $\mathbf{x}^{(k)}$, le vecteur $\mathbf{x}$ est alors représenté par la formule (201) et le vecteur $\mathbf{u}$ est donné par la formule:

$$\mathbf{u}=x_1\mathbf{x}^{(1)}+\ldots+x_m\mathbf{x}^{(m)},$$

et, dans le cas considéré, l'opération de projection se ramène tout simplement à conserver les m premières composantes et à annuler les autres composantes. Il est évident que la matrice de projection correspondante est une matrice diagonale de la forme: $P_{R_m} = [1, 1, \ldots, 1, 0, 0, \ldots, 0]$, où l'unité occupe les m premières places et le zéro les autres. Si nous avions numéroté les vecteurs de base d'une autre façon, nous aurions obtenu alors un ordre différent de succession des coefficients, mais comme précédemment nous aurions une matrice diagonale composée d'unités et de zéros. Dans le cas général, pour un choix arbitraire des axes de coordonnées rectangulaires, la matrice de projection a la forme:

$$P_{R_m} = U^{-1}[1, \ldots, 1, 0, \ldots, 0]\,U, \tag{202}$$

où U est une matrice unitaire et où les valeurs propres de P_{R_m} sont soit nulles, soit égales à l'unité. Inversement, chaque matrice hermitienne de la forme mentionnée est une matrice de projection sur un sous-espace dont la dimension est égale au nombre des valeurs propres de P_{R_m} égales à l'unité.

On peut encore définir la matrice de projection d'une autre façon: *la matrice de projection est une matrice hermitienne telle qu'elle vérifie la relation:*

$$P^2 = P. \tag{203}$$

En effet, tenant compte de ce que $1^2 = 1$ et $0^2 = 0$, il est facile de voir que les matrices de la forme (202) vérifient la relation (203). Inversement, si une matrice hermitienne vérifie la relation (203) et si nous la représentons sous la forme: $P = U^{-1}[\lambda_1, \ldots, \lambda_n]\,U$, alors en vertu de (203) $U^{-1}[\lambda_1^2, \ldots, \lambda_n^2]\,U = U^{-1}[\lambda_1, \ldots, \lambda_n]\,U$, c'est-à-dire que $\lambda_k^2 = \lambda_k$ $(k = 1, 2, \ldots, n)$, d'où il découle immédiatement que $\lambda_k = 1$ ou 0. Si toutes les valeurs propres d'une matrice sont égales à l'unité, cette matrice est alors la matrice unité et la transformation identique lui correspond, c'est-à-dire, en d'autres termes, la projection d'un vecteur sur tout l'espace (le vecteur restant inchangé). Excluant ce cas trivial, une matrice de projection a au moins une valeur propre nulle; par suite, le déterminant de cette matrice, égal au produit des valeurs propres, est nul, et la matrice inverse P^{-1} n'a pas de sens. Remarquons encore qu'il découle directement de sa définition que la matrice de projection P_{R_m} conserve les vecteurs qui appartiennent au sous-espace R_m et diminue la longueur des vecteurs qui n'appartiennent pas à R_m.

Considérons certaines opérations sur les matrices de projection. Soient deux matrices de projection P_R et P_S, telles que leur produit soit nul, c'est-à-dire une matrice dont tous les éléments sont nuls:

$$P_S P_R = 0. \tag{204}$$

Considérons un vecteur $\mathbf{x}$ du sous-espace R, donc $P_R\mathbf{x} = \mathbf{x}$. La formule (204) nous donne:

$$P_S\mathbf{x} = 0.$$

Il en découle que $\mathbf{x}$ est orthogonal à tout vecteur du sous-espace S. En effet, dans le cas contraire nous aurions pu trouver dans le sous-espace S un vecteur unité $\mathbf{y}$ qui ne soit pas orthogonal à $\mathbf{x}$ et, le prenant pour premier vecteur de base, nous aurions eu pour première composante du vecteur $\mathbf{x}$ une quantité non nulle, qui lors de la projection de $\mathbf{x}$ sur S resterait inchangée. Ainsi, nous voyons que si la condition (204) est satisfaite, tout vecteur de R est alors orthogonal à tout vecteur de S et inversement. Mais alors, nous avons aussi:

$$P_R P_S = 0. \tag{205}$$

En effet, pour tout vecteur $\mathbf{y}$, le vecteur $P_S\mathbf{y}$ appartient à S, donc est orthogonal à chaque vecteur de R, c'est-à-dire que pour tout vecteur $\mathbf{y}$ nous avons: $P_R P_S\mathbf{y} = 0$, ce qui est équivalent à (205). Inversement, si deux sous-espaces R et S sont orthogonaux au sens mentionné, on a alors (204) et (205).

Considérons maintenant la somme de deux matrices de projection:

$$P = P_R + P_S \tag{206}$$

et supposons que les conditions (204) et (205) soient satisfaites. Montrons que la matrice (206) qui est une matrice hermitienne est aussi une matrice de projection. Pour s'en convaincre, vérifions que son carré coïncide avec elle-même:

$$P^2 = (P_R + P_S)(P_R + P_S) = P_R^2 + P_R P_S + P_S P_R + P_S^2,$$

d'où en vertu de la condition imposée et de ce que P_R et P_S sont des matrices de projection, nous avons: $P^2 = P_R + P_S = P$.

Dans le cas considéré, il est aisé de montrer qu'à la matrice P correspond l'opération de projection sur le sous-espace $(R + S)$, réunion des sous-espaces R et S en ce sens que le sous-espace $(R + S)$ est engendré par l'ensemble de tous les vecteurs qui engendrent les sous-espaces R et S, c'est-à-dire que si les vecteurs $\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(p)}$ engendrent R et les vecteurs $\mathbf{y}^{(1)}, \ldots, \mathbf{y}^{(q)}$ S, le sous-espace $(R+S)$ est l'ensemble des vecteurs:

$$C_1\mathbf{x}^{(1)} + \ldots + C_p\mathbf{x}^{(p)} + D_1\mathbf{y}^{(1)} + \ldots + D_q\mathbf{y}^{(q)},$$

où C_k et D_k sont des constantes arbitraires. La propriété précédente se généralise au cas d'un nombre quelconque fini de termes:

$$P = P_{S_1} + \ldots + P_{S_m}. \tag{207}$$

Si les sous-espaces S_k sont deux à deux orthogonaux, c'est-à-dire si chaque vecteur de S_i est orthogonal à tout vecteur de S_j, pour i et j différents, la somme (207) représente alors une matrice de projection sur le sous-espace $(S_1 + \ldots + S_m)$ engendré par tous les vecteurs qui ont servi à engendrer les sous-espaces S_k. Dans un cas particulier cette somme peut être égale à la matrice unité $I = P_{S_1} + \ldots + P_{S_m}$ et dans ce cas on dit généralement que l'on a le développement de l'unité par des matrices de projection ou encore le développement de l'unité tout simplement.

Considérons encore le produit de deux matrices de projection:

$$P = P_S P_R. \tag{208}$$

Pour que ce produit soit une matrice de projection, il faut avant tout qu'il soit une matrice hermitienne, et, pour cela, comme on le sait de [II-2-10], il faut que les matrices soient commutatives:

$$P_R P_S = P_S P_R. \tag{209}$$

Montrons que cette condition est suffisante, c'est-à-dire que dans le cas considéré le carré de la matrice P^2 coïncide avec la matrice P:

$$P^2 = P_S P_R P_S P_R$$

ou, en changeant les matrices de place, en vertu de la condition (209):

$$P^2 = P_S^2 P_R^2 = P_S P_R,$$

ce qu'il fallait démontrer. Il est facile de vérifier qu'avec la condition de commutativité (209), à la matrice (208) correspond l'opération de projection sur un sous-espace engendré par les vecteurs communs aux deux ensembles de vecteurs qui engendrent R et S.

Notons encore un résultat sans en donner la démonstration qui ne présente d'ailleurs aucune difficulté, à savoir: si le sous-espace S est une partie du sous-espace R, alors la différence:

$$P = P_R - P_S \tag{210}$$

est aussi une matrice de projection. Si les $x^{(h)}$ sont les vecteurs fondamentaux qui engendrent S, pour obtenir les vecteurs fondamentaux qui engendrent R, nous devons ajouter aux vecteurs envisagés ci-dessus encore un ou plusieurs vecteurs linéairement indépendants. Ces derniers vecteurs engendront un sous-espace T et la matrice (210), dans le cas considéré, est la matrice de projection sur ce sous-espace. A l'aide des matrices de projection on peut énoncer d'une manière unique le problème de la réduction d'une matrice hermitienne à la forme diagonale, même dans le cas des valeurs propres multiples.

Supposons par exemple, que nous ayons une matrice hermitienne:

$$A = U[\lambda_1, \ldots, \lambda_n] U^{-1},$$

où U est une matrice unitaire, et pour fixer les idées, supposons que les nombres λ_k se divisent en deux groupes identiques entre eux, et, de plus, que les m premiers nombres soient égaux à μ et les $(n-m)$ autres à ν:

$$A = U[\mu, \ldots, \mu, \nu, \ldots, \nu] U^{-1}.$$

Nous pouvons écrire la matrice sous la forme:

$$A = \mu U[1, \ldots, 1, 0, \ldots, 0] U^{-1} + \nu U[0, \ldots, 0, 1, \ldots, 1] U^{-1}.$$

Introduisons les matrices de projection:

$$P_R = U[1, \ldots, 1, 0, \ldots, 0] U^{-1}; \quad P_S = U[0, \ldots, 0, 1, \ldots, 1] U^{-1}.$$

Les sous-espaces correspondants R et S sont évidemment orthogonaux et la somme des matrices de projection coïncide avec la matrice unité. Ainsi, nous avons dans le cas donné $A = \mu P_R + \nu P_S$, de plus, $\lambda_1 = \ldots = \lambda_m = \mu$ et $\lambda_{m+1} = \ldots = \lambda_n = \nu$.

Dans le cas général, le problème de la réduction d'une matrice hermitienne à la forme diagonale se ramène au développement suivant de la matrice unité:

$$I = P_{S_1} + \ldots + P_{S_m} \tag{211}$$

de manière que la matrice A puisse être représentée sous la forme:

$$A = \mu_1 P_{S_1} + \ldots + \mu_m P_{S_m}, \tag{212}$$

où μ_k sont les différentes valeurs propres de notre matrice. Ainsi, à chaque matrice hermitienne correspond un développement déterminé de l'unité (211) tel que cette matrice se représente sous la forme (212).

Il est facile de traduire tous les résultats précédents dans le langage des formes hermitiennes. A toute matrice de projection P_R d'éléments p_{ik} correspond une forme hermitienne:

$$P_R(\mathbf{x}) = (P_R\mathbf{x}, \mathbf{x}) = \sum_{i,k=1}^{n} p_{ik}\bar{x}_i x_k, \tag{213}$$

dite parfois *forme singulière* (Einzelform). Le symbole $P_R(\mathbf{x})$ est l'écriture abrégée de $(P_R\mathbf{x}, \mathbf{x})$.

Si le sous-espace correspondant R est de dimension l, et si nous prenons pour les l premiers vecteurs de base l vecteurs orthogonaux unités du sous-espace R, alors dans un tel système de coordonnées notre forme (213) s'écrit:

$$(P_R\mathbf{x}', \mathbf{x}') = \bar{x}_1' x_1' + \bar{x}_2' x_2' + \ldots + \bar{x}_l' x_l'.$$

Remarquons ensuite que si les matrices P_{S_k} sont celles d'un développement de l'unité, d'après (211), prenant alors pour vecteurs de base des vecteurs orthogonaux unités dans chaque sous-espace S_k, nous avons:

$$\sum_{k=1}^{m} P_{S_k}(x') = \sum_{i=1}^{n} \overline{x}_i' x_i',$$

et, par conséquent, pour tout choix des axes de coordonnées la somme

$$\sum_{k=1}^{m} P_{S_k}(x)$$

exprime le carré de la longueur du vecteur. Ainsi, nous pouvons dire que *le problème de la réduction d'une forme hermitienne A à la somme de carrés est équivalent aux deux égalités*:

$$(Ax, x) = \sum_{k=1}^{m} \mu_k P_{S_k}(x), \tag{214}$$

$$|x|^2 = \sum_{k=1}^{m} P_{S_k}(x). \tag{215}$$

L'introduction des matrices de projection permet ainsi d'énoncer le problème de la réduction d'une matrice hermitienne à la forme diagonale sans choix particulier des axes de coordonnées. Ceci donne la possibilité de transposer les résultats précédents, avec les modifications nécessaires, au cas d'espaces de dimensions infinies, ce qui constitue le problème fondamental de l'appareil mathématique de la mécanique quantique moderne. Il n'en sera question que beaucoup plus loin. Cette extension au cas de dimensions infinies sort du domaine de l'algèbre et est liée essentiellement à l'introduction de l'appareil d'analyse.

II-2-13. Fonctions de matrices. Une matrice peut jouer le rôle d'argument pour une fonction. Ici nous nous limiterons à la considération des fonctions les plus élémentaires, à savoir: *polynôme et fraction rationnelle d'une matrice*. Une étude plus détaillée de la théorie des fonctions de matrices sera donnée comme conséquence de la théorie des fonctions d'une variable complexe. Un polynôme $f(A)$ de degré m de la matrice variable A est de la forme:

$$f(A) = c_0 + c_1 A + \ldots + c_m A^m, \tag{216}$$

où les c_k sont des coefficients numériques. Dans le cas donné, la valeur de la fonction est aussi une matrice dont les éléments s'expriment d'après les formules:

$$\{f(A)\}_{ik} = c_0 \delta_{ik} + c_1 \{A\}_{ik} + \ldots + c_m \{A^m\}_{ik},$$

où:

$$\delta_{ik} = 0 \quad \text{pour } i \neq k \quad \text{et } \delta_{ii} = 1.$$

On peut considérer des polynômes de plusieurs matrices mais il faut alors se souvenir de la non-commutativité de ces matrices pour la multiplication. La forme générale d'un polynôme du deuxième degré dépendant de deux matrices variables A et B est:

$$f(A, B) = c_0 + c_1A + c_2B + c_3A^2 + c_4B^2 + c_5AB + c_6BA.$$

Remplaçons dans la formule (216) la matrice A par une matrice semblable quelconque $U^{-1}AU$. Tenant compte de $(U^{-1}AU)^k = U^{-1}A^kU$, nous avons:

$$f(U^{-1}AU) = c_0 + c_1U^{-1}AU + \ldots + c_mU^{-1}A^mU =$$
$$= U^{-1}(c_0 + c_1A + \ldots + c_mA^m)\,U,$$

c'est-à-dire:

$$f(U^{-1}AU) = U^{-1}f(A)\,U. \tag{217}$$

Nous avons une formule analogue dans le cas des polynômes de plusieurs matrices:

$$f(U^{-1}AU, U^{-1}BU) = U^{-1}f(A, B)\,U. \tag{218}$$

Etudions plus à fond le cas des matrices hermitiennes. Si A est une matrice hermitienne, on voit directement de la définition que toute puissance positive entière A^k, ainsi que le produit cA, où c est une constante réelle, sont aussi des matrices hermitiennes. En outre, la somme finie de matrices hermitiennes est aussi une matrice hermitienne. Il en découle immédiatement que si, dans la formule (216), A est une matrice hermitienne et les coefficients c_k sont des nombres réels, la valeur de la fonction $f(A)$ est aussi une matrice hermitienne. Il est évident que la matrice hermitienne $f(A)$ est commutative avec la matrice A et on peut les réduire simultanément à la forme diagonale à l'aide d'une transformation unitaire. Remarquons d'abord que si nous substituons à A une matrice diagonale $[\lambda_1, \ldots, \lambda_n]$ dans la fonction (216), nous avons encore comme résultat une matrice diagonale:

$$\sum_{k=0}^{m} c_k[\lambda_1^k, \ldots, \lambda_n^k] = [f(\lambda_1), \ldots, f(\lambda_n)], \tag{219}$$

où $f(\lambda_k)$ est la valeur numérique de notre polynôme lorsqu'on remplace A par le nombre λ_k.

Supposons maintenant que V soit une transformation unitaire réduisant la matrice A à la forme diagonale:

$$A = V[\lambda_1, \ldots, \lambda_n]\,V^{-1}.$$

D'après (217) et (219) nous avons:

$$f(A) = V[f(\lambda_1), \ldots, f(\lambda_n)]\,V^{-1},$$

c'est-à-dire que V réduit aussi $f(A)$ à la forme diagonale, en outre $f(\lambda_k)$ sont les valeurs propres de cette dernière matrice.

Passons maintenant à la considération des fractions rationnelles. Soient $f_1(A)$ et $f_2(A)$ deux polynômes de la matrice A. Considérons la fraction :

$$\frac{f_1(A)}{f_2(A)}. \tag{220}$$

Comme nous l'avons vu précédemment, le rapport de deux matrices n'a pas, en général, une valeur déterminée [II-1-7], mais dans le cas donné, comme il est facile de le montrer, nous obtenons pour le rapport (220) une valeur déterminée si le déterminant de la matrice $f_2(A)$ n'est pas nul. La fraction (220) peut s'écrire de deux manières :

$$f_1(A)\,f_2(A)^{-1} \quad \text{ou} \quad f_2(A)^{-1}\,f_1(A).$$

Montrons que ces deux produits sont identiques :

$$f_1(A)\,f_2(A)^{-1} = f_2(A)^{-1}\,f_1(A),$$

ou, ce qui est équivalent :

$$f_2(A)\,f_1(A) = f_1(A)\,f_2(A). \tag{221}$$

Puisque nos polynômes ne dépendent que de la matrice A, ils sont alors commutatifs, c'est-à-dire que (221) est réellement satisfait et le rapport (220) a une valeur déterminée. Il est encore facile de vérifier que les fractions rationnelles d'une même matrice se multiplient entre elles de la même façon que les fractions ordinaires. En effet :

$$\frac{f_1(A)}{f_2(A)}\frac{f_3(A)}{f_4(A)} = f_1(A)\,f_2(A)^{-1}\,f_3(A)\,f_4(A)^{-1}$$

ou, tenant compte de la commutativité :

$$\frac{f_1(A)}{f_2(A)}\frac{f_3(A)}{f_4(A)} = f_1(A)\,f_3(A)\cdot[f_2(A)\,f_4(A)]^{-1} = \frac{f_1(A)\,f_3(A)}{f_2(A)\,f_4(A)}.$$

Considérons l'exemple d'une fraction rationnelle de la forme :

$$U = \frac{1+iA}{1-iA}, \tag{222}$$

où A est une matrice hermitienne, c'est-à-dire que $\bar{A}^{(*)} = A$. Il est facile de montrer que la matrice U est unitaire, c'est-à-dire que

$$\bar{U}^{(*)} = U^{-1}. \tag{223}$$

En effet, nous avons :

$$\bar{U} = \frac{1-i\bar{A}}{1+i\bar{A}} = (1-i\bar{A})\,(1+i\bar{A})^{-1},$$

d'où, en passant à la matrice transposée, nous obtenons [II-1-7] :

$$\bar{U}^{(*)} = (1+i\bar{A})^{(*)-1}\,(1-i\bar{A})^{(*)} = (1+i\bar{A}^{(*)})^{-1}\,(1-i\bar{A}^{(*)})$$

ou en vertu de $\bar{A}^{(*)} = A$:

$$\bar{U}^{(*)} = (1 + iA)^{-1}(1 - iA) = \frac{1 - iA}{1 + iA} = U^{-1},$$

c'est-à-dire que (223) est vérifié et que U est une matrice unitaire réelle.

Nous pouvons écrire la formule (222) sous la forme:

$$U(1 - iA) = (1 + iA),$$

de plus, d'après (222), U est commutative avec A et par suite:

$$A = i\frac{U - 1}{U + 1}. \qquad (224)$$

Comme précédemment, on peut montrer que si U est une matrice unitaire et le déterminant de la matrice $U + 1$ non nul, la matrice A donnée par la formule (224) est aussi une matrice hermitienne. Ainsi, toute matrice unitaire pour laquelle $D(U + 1) \neq 0$ peut être représentée par une matrice hermitienne A d'après la formule (222).

II-2-14. Espace à une infinité de dimensions. Passons à présent à la notion d'espace à une infinité de dimensions. Préalablement il nous faut introduire la notion de limite d'une variable complexe. Supposons qu'une variable complexe $z = x + yi$ prenne une suite de valeurs:

$$z_1 = x_1 + y_1 i;\quad z_2 = x_2 + y_2 i;\quad \ldots;\quad z_n = x_n + y_n i;\quad \ldots \qquad (225)$$

On dit que le nombre complexe $\alpha = a + bi$ est la limite de la suite (225), si le module de la différence $(\alpha - z_n)$ tend vers zéro, lorsque n croît indéfiniment, c'est-à-dire que $|\alpha - z_n| \to 0$ quand $n \to \infty$, et on écrit $\alpha = \lim z_n$ ou $z_n \to \alpha$. Mais:

$$|\alpha - z_n| = |(a - x_n) + (b - y_n)i| = \sqrt{(a - x_n)^2 + (b - y_n)^2}.$$

Vu que les 2 termes sous la racine carrée sont non négatifs, la condition $|\alpha - z_n| \to 0$ équivaut à $x_n \to a$ et $y_n \to b$. Ainsi:

$$x_n + y_n i \to a + bi \qquad (226)$$

équivaut à $x_n \to a$ et $y_n \to b$. Considérons une série à termes complexes:

$$\sum_{k=1}^{\infty} (a_k + b_k i). \qquad (227)$$

Elle est dite convergente si la somme des n premiers termes

$$S_n = \sum_{k=1}^{n} (a_k + b_k i) = (a_1 + a_2 + \ldots + a_n) + (b_1 + b_2 + \ldots + b_n)i$$

tend vers une limite: $S_n \to a+bi$ lorsque n croît indéfiniment et cette limite $(a+bi)$ est appelée somme de la série. De la définition de la limite il s'ensuit que la convergence de la série (227) équivaut à la convergence des séries:

$$a = \sum_{k=1}^{\infty} a_k \quad \text{et} \quad b = \sum_{k=1}^{\infty} b_k, \tag{228}$$

constituées des parties réelles et imaginaires des termes de la série (227).

Supposons que la série

$$\sum_{k=1}^{\infty} |a_k + ib_k| = \sum_{k=1}^{\infty} \sqrt{a_k^2 + b_k^2} \tag{229}$$

des modules des termes de la série (227) soit convergente. Des inégalités évidentes:

$$|a_k| \leqslant \sqrt{a_k^2 + b_k^2} \quad \text{et} \quad |b_k| \leqslant \sqrt{a_k^2 + b_k^2}, \tag{230}$$

on conclut que les séries (228) sont convergentes et même absolument convergentes et, par suite, que la série (227) est aussi convergente, c'est-à-dire que si la série (229) converge, la série (227) converge *a fortiori*. Dans ce cas, la série (227) est dite *absolument convergente*. Appliquant le critère général de Cauchy, nous pouvons énoncer la condition nécessaire et suffisante de convergence absolue de la façon suivante: pour tout ε positif petit il existe un N tel que:

$$\sum_{k=n}^{n+p} |a_k + ib_k| < \varepsilon, \tag{231}$$

pour $n > N$ et p un nombre entier positif arbitraire.

Appliquons maintenant ce qui vient d'être dit ci-dessus à quelques cas particuliers importants pour la suite. Considérons une série de la forme:

$$\sum_{k=1}^{\infty} \alpha_k \beta_k, \tag{232}$$

où α_k et β_k sont des nombres complexes tels que les séries

$$\sum_{k=1}^{\infty} |\alpha_k|^2 \quad \text{et} \quad \sum_{k=1}^{\infty} |\beta_k|^2 \tag{233}$$

soient convergentes. Appliquons l'inégalité démontrée dans [II-1-10]:

$$\left(\sum_{k=n}^{n+p} |\alpha_k \beta_k|\right)^2 \leqslant \sum_{k=n}^{n+p} |\alpha_k|^2 \sum_{k=n}^{n+p} |\beta_k|^2.$$

Tenant compte de la convergence des séries (233), nous trouvons que la somme :

$$\sum_{k=n}^{n+p} |\alpha_k \beta_k|$$

est aussi petite que l'on veut pour des n grands et des p arbitraires, c'est-à-dire que la convergence des séries (233) assure la convergence absolue de la série (232).

Considérons à présent la série :

$$\sum_{k=1}^{\infty} |\alpha_k + \beta_k|^2 = \sum_{k=1}^{\infty} (\alpha_k + \beta_k)(\bar{\alpha}_k + \bar{\beta}_k) \tag{234}$$

et supposons comme précédemment que les séries (233) soient convergentes. La série (234) peut être représentée sous la forme d'une somme de quatre séries :

$$\sum_{k=1}^{\infty} |\alpha_k|^2; \quad \sum_{k=1}^{\infty} |\beta_k|^2; \quad \sum_{k=1}^{\infty} \alpha_k \bar{\beta}_k; \quad \sum_{k=1}^{\infty} \bar{\alpha}_k \beta_k.$$

Les deux premières séries convergent par hypothèse, quant à la convergence des deux autres, elle découle de la proposition démontrée ci-dessus, à savoir que la convergence des séries (233) assure aussi la convergence de la série (234).

Passons maintenant à la considération d'un *espace à une infinité de dimensions*. Nous appelons *vecteur dans un tel espace* une suite infinie de nombres complexes :

$$\mathbf{x}(x_1, x_2, \ldots),$$

en outre, nous supposons toujours que ces nombres satisfont à la condition de convergence de la série :

$$\sum_{k=1}^{\infty} |x_k|^2. \tag{235}$$

L'ensemble de tous ces vecteurs est appelé *espace de Hilbert*, du nom de celui qui le premier étudia un tel espace. Pour plus de commodité cet espace sera noté l_2.

Comme précédemment, nous introduisons pour les vecteurs de l'espace l_2 les principales opérations : multiplication d'un vecteur par un nombre et addition de vecteurs. Si les x_k sont les composantes d'un vecteur $\mathbf{x}$, celles du vecteur $c\mathbf{x}$, où c est un nombre complexe, sont les cx_k. Si les x_k et les y_k sont les composantes des vecteurs $\mathbf{x}$ et $\mathbf{y}$, celles du vecteur $(\mathbf{x} + \mathbf{y})$ sont les $(x_k + y_k)$. La différence $\mathbf{x} - \mathbf{y}$ est égale à la somme de $\mathbf{x}$ et de $(-1)\,\mathbf{y}$ (cf. [I-2-5]). De la convergence de la série (235) découle la convergence de la série $\sum_{k=1}^{\infty} |cx_k|^2$. De même, si les séries

$$\sum_{k=1}^{\infty} |x_k|^2 \quad \text{et} \quad \sum_{k=1}^{\infty} |y_k|^2$$

sont convergentes, il découle de ce qui précède que la série

$$\sum_{k=1}^{\infty} |x_k + y_k|^2$$

converge aussi, c'est-à-dire que les suites de nombres $(cx_1, cx_2, \ldots)$ et $(x_1 + y_1, x_2 + y_2, \ldots)$ définissent les vecteurs $\mathbf{cx}$ et $\mathbf{x} + \mathbf{y}$, si $\mathbf{x}$ et $\mathbf{y}$ appartiennent à l_2. Le vecteur nul est par définition un vecteur dont toutes les composantes sont nulles. Il sera noté dans les égalités vectorielles par le chiffre zéro. Les opérations sur les vecteurs vérifient les règles habituelles (cf. [I-2-5]):

$$\mathbf{x}+\mathbf{y}=\mathbf{y}+\mathbf{x}; \qquad (\mathbf{x}+\mathbf{y})+\mathbf{z}=\mathbf{x}+(\mathbf{y}+\mathbf{z}),$$

$$(a+b)\mathbf{x}=a\mathbf{x}+b\mathbf{x}; \quad a(\mathbf{x}+\mathbf{y})=a\mathbf{x}+a\mathbf{y}; \quad a(b\mathbf{x})=(ab)\mathbf{x}.$$

De même, d'après ce qui a été dit ci-dessus, nous pouvons définir le produit scalaire des vecteurs $\mathbf{x}$ et $\mathbf{y}$:

$$(\mathbf{x}, \mathbf{y}) = \sum_{k=1}^{\infty} x_k \bar{y}_k,$$

d'où découlent les formules [I-2-6]:

$$(\mathbf{y}, \mathbf{x}) = \overline{(\mathbf{x}, \mathbf{y})}; \quad (a\mathbf{x}, \mathbf{y}) = a(\mathbf{x}, \mathbf{y}); \quad (\mathbf{x}, a\mathbf{y}) = \bar{a}(\mathbf{x}, \mathbf{y}),$$

$$(\mathbf{x}, \mathbf{y}, \mathbf{z}) = (\mathbf{x}, \mathbf{z}) + (\mathbf{y}, \mathbf{z}); \qquad (\mathbf{z}, \mathbf{x}+\mathbf{y}) = (\mathbf{z}, \mathbf{x}) + (\mathbf{z}, \mathbf{y}).$$

La somme

$$(\mathbf{x}, \mathbf{x}) = \sum_{k=1}^{\infty} |x_k|^2 \tag{236}$$

définit le carré de la *norme* (longueur) du vecteur $\mathbf{x}$. Introduisons la notation suivante pour désigner la norme d'un vecteur :

$$\sum_{k=1}^{\infty} |x_k|^2 = \|\mathbf{x}\|^2, \tag{237}$$

$\|\mathbf{x}\|$ désignant la norme du vecteur $\mathbf{x}$.

On a pour le produit scalaire l'inégalité [II-1-11] :

$$|(\mathbf{x}, \mathbf{y})| \leqslant \|\mathbf{x}\| \cdot \|\mathbf{y}\| \tag{238}$$

et on établit l'inégalité du triangle de la même façon que dans [II-1-11] :

$$\|\mathbf{x}+\mathbf{y}\| \leqslant \|\mathbf{x}\| + \|\mathbf{y}\|. \tag{239}$$

La norme de tout vecteur est positive, excepté pour le vecteur nul dont la norme est nulle. Deux vecteurs u et v sont dits *réciproquement orthogonaux* ou tout simplement *orthogonaux* si leur produit scalaire est nul, c'est-à-dire si $(\mathbf{u}, \mathbf{v}) = 0$ et $(\mathbf{v}, \mathbf{u}) = 0$, l'une de ces égalités étant conséquence de l'autre.

Si les vecteurs $\mathbf{x}^{(k)}$ $(k = 1, 2, \ldots, m)$ sont deux à deux orthogonaux, c'est-à-dire si $(\mathbf{x}^{(i)}, \mathbf{x}^{(k)}) = 0$ pour $i \neq k$, alors:

$$(\mathbf{x}^{(1)} + \ldots + \mathbf{x}^{(m)}, \mathbf{x}^{(1)} + \ldots + \mathbf{x}^{(m)}) = (\mathbf{x}^{(1)}, \mathbf{x}^{(1)}) + \ldots + (\mathbf{x}^{(m)}, \mathbf{x}^{(m)})$$

ou, ce qui revient au même :

$$\|\mathbf{x}^{(1)} + \ldots + \mathbf{x}^{(m)}\|^2 = \|\mathbf{x}^{(1)}\|^2 + \ldots + \|\mathbf{x}^{(m)}\|^2, \tag{240}$$

c'est-à-dire que *le carré de la norme d'une somme de vecteurs deux à deux orthogonaux est égal à la somme des carrés des normes des termes.* Il est naturel d'appeler cette proposition *théorème de Pythagore.* Il découle immédiatement de la définition de la norme que si c est un nombre complexe, on a alors pour la norme du vecteur $c\mathbf{x}$

$$\|c\mathbf{x}\| = |c| \cdot \|\mathbf{x}\|.$$

On dit que les vecteurs

$$\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \ldots, \mathbf{x}^{(m)} \tag{241}$$

forment un système orthonormé si ces vecteurs sont deux à deux orthogonaux et si la norme de chacun d'entre eux est égale à l'unité, c'est-à-dire si :

$$(\mathbf{x}^{(i)}, \mathbf{x}^{(k)}) = \delta_{ik},$$

où $\delta_{ik} = 0$ pour $i \neq k$ et $\delta_{ii} = 1$. Notons que la formule (240) donne alors :

$$\|\alpha_1 \mathbf{x}^{(1)} + \ldots + \alpha_m \mathbf{x}^{(m)}\|^2 = |\alpha_1|^2 + \ldots + |\alpha_m|^2,$$

où α_k sont des nombres complexes quelconques. Les systèmes orthonormés peuvent être aussi composés d'un ensemble infini de vecteurs. Comme exemple, donnons les vecteurs de base de l'espace l_2:

$$\mathbf{a}^{(1)}(1, 0, 0, 0, \ldots); \ \mathbf{a}^{(2)}(0, 1, 0, 0, \ldots); \ \mathbf{a}^{(3)}(0, 0, 1, 0, \ldots); \ \ldots$$

Nous avons pour les composantes d'un vecteur quelconque $\mathbf{y}(y_1, y_2, \ldots)$:

$$y_k = (\mathbf{y}, \mathbf{a}^{(k)}).$$

Revenons au système orthonormé fini (241). Le produit scalaire $(\mathbf{y}, \mathbf{x}^{(k)})$ est fréquemment appelé *coefficient de Fourier* du vecteur $\mathbf{y}$ par rapport au système orthonormé (241) ou *projection* de $\mathbf{y}$ sur l'axe $\mathbf{x}^{(k)}$. La somme

$$\sum_{k=1}^{m}(\mathbf{y},\ \mathbf{x}^{(k)})\,\mathbf{x}^{(k)}$$

ne coïncide pas avec $\mathbf{y}$. Représentons $\mathbf{y}$ sous la forme :

$$\mathbf{y}=\sum_{k=1}^{m}(\mathbf{y},\ \mathbf{x}^{(k)})\,\mathbf{x}^{(k)}+\mathbf{u}. \tag{242}$$

Multipliant scalairement les deux membres de cette égalité par $\mathbf{x}^{(l)}$ et tenant compte de ce que le système (241) est orthonormé, nous obtenons :

$$(\mathbf{y},\ \mathbf{x}^{(l)})=(\mathbf{y},\ \mathbf{x}^{(l)})+(\mathbf{u},\ \mathbf{x}^{(l)}),$$

c'est-à-dire $(\mathbf{u},\ \mathbf{x}^{(l)})=0$ $(l=1,\ 2,\ \ldots,\ m)$. Autrement dit, le vecteur $\mathbf{u}$ est orthogonal à tous les vecteurs $\mathbf{x}^{(k)}$ $(k=1,\ 2,\ \ldots,\ m)$. Nous pouvons ainsi appliquer le théorème de Pythagore au second membre de l'égalité (242) :

$$\|\mathbf{y}\|^2=\sum_{k=1}^{m}|(\mathbf{y},\ \mathbf{x}^{(k)})|^2+\|\mathbf{u}\|^2. \tag{243}$$

Il en découle immédiatement l'inégalité :

$$\sum_{k=1}^{m}|(\mathbf{y},\ \mathbf{x}^{(k)})|^2\leqslant\|\mathbf{y}\|^2, \tag{244}$$

dite *inégalité de Bessel*. On a le signe d'égalité dans cette inégalité seulement si $\mathbf{u}=0$, c'est-à-dire si $\mathbf{u}$ est le vecteur nul. Cela découle directement de (243).

Pour transposer ces derniers résultats au cas des systèmes orthonormés infinis il nous faut introduire la notion de limite d'une suite de vecteurs et considérer les séries infinies dont les termes sont des vecteurs appartenant à l_2.

II-2-15. Convergence des vecteurs. Soit une suite infinie de vecteurs $\mathbf{v}^{(k)}$ $(k=1,\ 2,\ \ldots)$. On dit que cette suite *tend vers le vecteur* $\mathbf{v}$ ou que le vecteur $\mathbf{v}$ est la *limite* de cette suite, si pour $k\to\infty$

$$\|\mathbf{v}-\mathbf{v}^{(k)}\|\to 0,\quad \text{c'est-à-dire}\quad \|\mathbf{v}-\mathbf{v}^{(k)}\|^2\to 0. \tag{245}$$

Désignant par $v_1^{(k)}$, $v_2^{(k)}$, ... les composantes de $\mathbf{v}^{(k)}$ et par v_1, v_2, ... les composantes de $\mathbf{v}$, nous pouvons écrire la condition (245) sous la forme développée :

$$\lim_{k\to\infty}\sum_{k=1}^{\infty}[|v_1-v_1^{(k)}|^2+|v_2-v_2^{(k)}|^2+\ldots]=0. \tag{246}$$

La somme de termes non négatifs tend vers zéro et l'on peut en dire autant de chaque terme, c'est-à-dire que de (246) il découle:

$$\lim_{k\to\infty} |v_m - v_m^{(k)}| = 0 \qquad (m = 1, 2, \ldots), \tag{247}$$

et chaque composante $v_m^{(k)}$ tend vers la composante correspondante v_m. Plus précisément, les parties réelle et imaginaire de $v_m^{(k)}$ tendent vers les parties réelle et imaginaire de v_m [II-2-14]. Notons que (247) découle de (246), mais que l'inverse n'est pas vrai, c'est-à-dire que (246) ne découle pas de (247). Prenons pour exemple le vecteur $\mathbf{v}^{(k)}$ de composantes $(0, \ldots, 0, 1, 0, \ldots)$, l'unité occupant la place de rang k. Pour k croissant indéfiniment chacune des composantes devient nulle, c'est-à-dire que pour tout entier m nous avons: $v_m^{(k)} \to 0$, et donc $v_m = 0$ $(m = 1, 2, \ldots)$, mais en même temps la somme (246) reste constamment égale à l'unité.

Si la suite $\mathbf{v}^{(k)}$ tend vers $\mathbf{v}$, on écrit que $\mathbf{v}^{(k)} \Rightarrow \mathbf{v}$. Considérons un exemple de convergence. Soit $\mathbf{v}\,(v_1, v_2, \ldots)$ un vecteur. Définissons les vecteurs $\mathbf{v}^{(k)}$ de la façon suivante: l'ensemble des vecteurs $\mathbf{v}^{(k)}$ a les k premières composantes égales à celles de $\mathbf{v}$, les autres composantes sont nulles, c'est-à-dire:

$$\mathbf{v}^{(k)}\,(v_1, v_2, \ldots, v_k, 0, 0, \ldots).$$

Il est aisé de montrer que $\mathbf{v}^{(k)} \Rightarrow \mathbf{v}$. En effet, dans le cas considéré:

$$\|\mathbf{v} - \mathbf{v}^{(k)}\|^2 = \sum_{n=k+1}^{\infty} |v_n|^2,$$

et d'après la convergence de la série de terme général $|v_n|^2$, la somme écrite tend vers zéro lorsque k croît indéfiniment. Donnons quelques règles simples liées à la notion de limite. Si $\mathbf{u}^{(k)} \Rightarrow \mathbf{u}$ et $\mathbf{v}^{(k)} \Rightarrow \mathbf{v}$,

$$\mathbf{u}^{(k)} + \mathbf{v}^{(k)} \Rightarrow \mathbf{u} + \mathbf{v} \quad \text{et} \quad (\mathbf{u}^{(k)}, \mathbf{v}^{(k)}) \to (\mathbf{u}, \mathbf{v}).$$

Remarquons qu'un produit scalaire est un nombre complexe et dans la dernière formule nous écrivons $\to$ et non $\Rightarrow$. Cette formule exprime la continuité du produit scalaire. D'après (234), nous avons:

$$\|(\mathbf{u} + \mathbf{v}) - (\mathbf{u}^{(k)} + \mathbf{v}^{(k)})\| = \|(\mathbf{u} - \mathbf{u}^{(k)}) + (\mathbf{v} - \mathbf{v}^{(k)})\| \leqslant$$
$$\leqslant \|\mathbf{u} - \mathbf{u}^{(k)}\| + \|\mathbf{v} - \mathbf{v}^{(k)}\|$$

et en vertu de la définition de la limite $\|\mathbf{u} - \mathbf{u}^{(k)}\| \to 0$ et $\|\mathbf{v} - \mathbf{v}^{(k)}\| \to 0$. De l'inégalité écrite, il decoule aussi que:

$$\|(\mathbf{u} + \mathbf{v}) - (\mathbf{u}^{(k)} + \mathbf{v}^{(k)})\| \to 0,$$

c'est-à-dire que réellement $\mathbf{u}^{(k)} + \mathbf{v}^{(k)} \Rightarrow \mathbf{u} + \mathbf{v}$. En vertu de la définition de la limite, il découle ensuite que:

$$\mathbf{u}^{(k)} = \mathbf{u} + \mathbf{s}^{(k)}; \quad \mathbf{v}^{(k)} = \mathbf{v} + \mathbf{t}^{(k)},$$

où $\|s^{(k)}\| \to 0$ et $\|t^{(k)}\| \to 0$. Nous avons pour le produit scalaire:

$$(u^{(k)}, v^{(k)}) = (u + s^{(k)},\ v + t^{(k)}) =$$
$$= (u, v) + (u, t^{(k)}) + (s^{(k)}, v) + (s^{(k)}, t^{(k)}),$$

d'où:

$$|(u, v) - (u^{(k)}, v^{(k)})| \leqslant |(u, t^{(k)})| + |(s^{(k)}, v)| + |(s^{(k)}, t^{(k)})|,$$

ou d'après (238):

$$|(u, v) - (u^{(k)}, v^{(k)})| \leqslant \|u\| \cdot \|t^{(k)}\| + \|s^{(k)}\| \cdot \|v\| + \|s^{(k)}\| \cdot \|t^{(k)}\|.$$

Le second membre tend vers zéro et par suite:

$$|(u, v) - (u^{(k)}, v^{(k)})| \to 0, \text{ c'est-à-dire que } (u^{(k)}, v^{(k)}) \to (u, v).$$

En particulier, $(u^{(k)}, u^{(k)}) \to (u, u)$, c'est-à-dire que $\|u^{(k)}\|^2 \to \|u\|^2$ ou $\|u^{(k)}\| \to \|u\|$.

On démontre facilement que si la suite des nombres c_k a c pour limite, $c_k u^{(k)} \Rightarrow cu$.

La condition nécessaire et suffisante d'existence de la limite donnée par le critère de Cauchy est alors vérifiée. Formulons ce critère dans le cas donné. Soit une suite de vecteurs:

$$v^{(k)} \qquad (k = 1, 2, \ldots). \tag{248}$$

Pour que cette suite ait une limite, il faut et il suffit que pour ε positif arbitraire il existe N tel que:

$$\|v^{(n)} - v^{(m)}\| \leqslant \varepsilon \tag{249}$$

quand n et $m > N$.

Montrons d'abord que cette condition est nécessaire. Supposons que la suite (248) ait v pour limite. Nous pouvons écrire:

$$v^{(n)} - v^{(m)} = (v^{(n)} - v) + (v - v^{(m)})$$

et d'après l'inégalité du triangle:

$$\|v^{(n)} - v^{(m)}\| \leqslant \|v^{(n)} - v\| + \|v - v^{(m)}\|.$$

Il découle immédiatement de la définition de la limite que les deux termes du second membre tendent vers zéro lorsque n et m croissent et, par conséquent, il en est de même pour le premier membre, c'est-à-dire que la condition (249) doit nécessairement être vérifiée. Démontrons à présent que la condition (249) est suffisante. Supposons que cette condition soit vérifiée et démontrons que la suite (248) tend vers une limite. Sous forme développée la condition (249) peut encore s'écrire:

$$\sum_{s=1}^{\infty} |v_s^{(n)} - v_s^{(m)}|^2 \leqslant \varepsilon^2 \quad \text{pour} \quad n \text{ et } m > N, \tag{250}$$

où $v_s^{(j)}$ sont les composantes de $\mathbf{v}^{(j)}$. Il en découle immédiatement que pour tout s nous avons: $|v_s^{(n)} - v_s^{(m)}| \leqslant \varepsilon$ pour n et $m > N$ ou, en développant suivant les parties réelle et imaginaire:

$$v_s^{(n)} = a_s^{(n)} + ib_s^{(n)},$$

nous pouvons écrire :

$$|a_s^{(n)} - a_s^{(m)}| < \varepsilon \text{ et } |b_s^{(n)} - b_s^{(m)}| < \varepsilon \text{ pour } n \text{ et } m > N.$$

Appliquant le critère de Cauchy (tome I, [I-2-7]) on peut affirmer que $a_s^{(n)}$ et $b_s^{(n)}$ ont des limites pour $n \to \infty$. Nous les noterons a_s et b_s, et nous pouvons écrire que $v_s^{(n)}$ a pour limite le nombre complexe $a_s + ib_s$. Montrons d'abord que la série

$$\sum_{s=1}^{\infty} |v_s|^2$$

converge, c'est-à-dire que les v_s sont les composantes d'un vecteur. Fixons un nombre fini de premiers termes dans la somme (250) et prenons dans cette somme finie la limite en n, nous avons:

$$\sum_{s=1}^{M} |v_s - v_s^{(m)}|^2 \leqslant \varepsilon^2,$$

où M est un nombre entier arbitraire. A la limite, quand dans cette dernière inégalité $M \to \infty$, nous trouvons :

$$\sum_{s=1}^{\infty} |v_s - v_s^{(m)}|^2 \leqslant \varepsilon^2, \tag{251}$$

d'où il découle que les nombres $v_s - v_s^{(m)}$ sont les composantes d'un certain vecteur. Il en est de même pour les nombres $v_s^{(m)}$ et, par conséquent, pour leur somme, c'est-à-dire les nombres v_s. Ainsi, ces nombres sont les composantes d'un vecteur $\mathbf{v}$ et l'inégalité (251) peut s'écrire sous la forme:

$$\|\mathbf{v} - \mathbf{v}^{(m)}\| \leqslant \varepsilon,$$

pour $m > N$, c'est-à-dire $\mathbf{v}^{(m)} \to \mathbf{v}$. Par conséquent, la suite (248) a bien une limite. Chaque composante v_s du vecteur $\mathbf{v}$ étant définie comme étant la limite de $v_s^{(m)}$, il en découle immédiatement que cette limite est unique. Considérons maintenant une somme infinie de vecteurs:

$$\mathbf{u}^{(1)} + \mathbf{u}^{(2)} + \ldots \tag{252}$$

Elle est dite convergente si la somme des n premiers termes

$$\mathbf{s}^{(n)} = \mathbf{u}^{(1)} + \ldots + \mathbf{u}^{(n)}$$

a une limite au sens indiqué ci-dessus pour $n \to \infty$. En vertu du critère de Cauchy, la condition nécessaire et suffisante de convergence est donnée par l'inégalité:

$$\| s^{(n+p)} - s^{(n)} \| = \| u^{(n+1)} + \ldots + u^{(n+p)} \| < \varepsilon, \qquad (253)$$

satisfaite pour $n > N$ et p arbitraire.

Notons que de la continuité du produit scalaire [II-2-15] il découle que si la série (252) converge, s étant sa somme, z un élément quelconque de l_2, alors:

$$(s, z) = \sum_{k=1}^{\infty} (u^{(k)}, z) \quad \text{et} \quad (z, s) = \sum_{k=1}^{\infty} (z, u^{(k)}), \qquad (254)$$

c'est-à-dire en bref que les produits scalaires (s, z) et (z, s) peuvent être définis terme à terme.

Etablissons maintenant la condition nécessaire et suffisante de convergence d'une série (252) constituée de vecteurs $u^{(k)}$ deux à deux orthogonaux. En vertu du critère de Cauchy, nous devons former l'expression (253) dont le carré, conformément au théorème de Pythagore, est égal à:

$$\| u^{(n+1)} \|^2 + \ldots + \| u^{(n+m)} \|^2.$$

Il en découle immédiatement que *pour la convergence d'une série, constituée de vecteurs deux à deux orthogonaux, il faut et il suffit que la série des carrés des normes des termes de la série converge*. Ce résultat peut encore s'énoncer autrement. Soit $x^{(k)}$ ($k = 1, 2, \ldots$) un système orthonormé infini. Formons la série:

$$\sum_{k=1}^{\infty} \alpha_k x^{(k)}, \qquad (255)$$

où les α_k sont des nombres complexes. Il découle directement de ce qui vient d'être dit que *la condition nécessaire et suffisante de convergence de la série* (255) *est la convergence de la série*:

$$\sum_{k=1}^{\infty} |\alpha_k|^2.$$

En outre, il en découle immédiatement que le changement de l'ordre des termes de la série (255) n'influe pas sur sa convergence. On peut de plus montrer que la somme de cette série ne change pas. Supposons que la série (255) converge et désignons par s sa somme:

$$s = \sum_{k=1}^{\infty} \alpha_k x^{(k)}. \qquad (255_1)$$

Multipliant par $\mathbf{x}^{(l)}$ les deux membres, tenant compte de ce que le système des $\mathbf{x}^{(k)}$ est orthonormé, nous trouvons: $\alpha_l = (\mathbf{s}, \mathbf{x}^{(l)})$ ($l = 1, 2, \ldots$). Remplaçant α_k par $(\mathbf{s}, \mathbf{x}^{(k)})$ et multipliant scalairement par $\mathbf{s}$ les deux membres de (255_1), nous trouvons:

$$\|\mathbf{s}\|^2 = \sum_{k=1}^{\infty} |(\mathbf{s}, \mathbf{x}^{(k)})|^2, \tag{256}$$

c'est-à-dire que *les coefficients de la série convergente* (255_1) *sont les coefficients de Fourier de sa somme et qu'on a pour cette somme la formule de fermeture* (256) *par rapport au système orthonormé des* $\mathbf{x}^{(k)}$.

Revenons à la série (252). Supposons que ses termes soient deux à deux orthogonaux et qu'elle converge. Nous avons donc $\mathbf{u}^{(k)} = \alpha_k \mathbf{x}^{(k)}$, où $\alpha_k = \|\mathbf{u}^{(k)}\|$ sont des nombres et les $\mathbf{x}^{(k)}$ forment un système orthonormé. Le résultat obtenu ci-dessus peut se mettre sous la forme :

$$\|\mathbf{s}\|^2 = \sum_{k=1}^{\infty} \|\mathbf{u}^{(k)}\|^2,$$

c'est-à-dire que le théorème de Pythagore est vrai aussi pour les séries convergentes dont les termes sont deux à deux orthogonaux.

II-2-16. Systèmes orthonormés. Soit un système orthonormé infini :

$$\mathbf{x}^{(k)} \quad (k = 1, 2, \ldots), \tag{257}$$

c'est-à-dire :

$$(\mathbf{x}^{(i)}, \mathbf{x}^{(k)}) = \delta_{ik}, \tag{257_1}$$

et $\mathbf{y}$ un vecteur quelconque de l_2. Formons « la série de Fourier » de $\mathbf{y}$ par rapport au système (257) :

$$\sum_{k=1}^{\infty} (\mathbf{y}, \mathbf{x}^{(k)})\, \mathbf{x}^{(k)}. \tag{258}$$

Comme nous l'avons vu dans [II-2-14], pour tout m fini :

$$\sum_{k=1}^{m} |(\mathbf{y}, \mathbf{x}^{(k)})|^2 \leqslant \|\mathbf{y}\|^2,$$

et à la limite pour $m \to \infty$ nous avons:

$$\sum_{k=1}^{\infty} |(\mathbf{y}, \mathbf{x}^{(k)})|^2 \leqslant \|\mathbf{y}\|^2, \tag{259}$$

d'où la convergence de la série du premier membre de cette inégalité (inégalité de Bessel) et, en vertu de ce qui a été dit dans [II-2-15], la convergence de la série (258).

Posons :

$$y = \sum_{k=1}^{\infty} (y, x^{(k)}) x^{(k)} + u. \tag{260}$$

De même que dans [II-2-14], on peut montrer que u est orthogonal à tous les $x^{(k)}$ et le théorème de Pythagore donne :

$$\|y\|^2 = \sum_{k=1}^{\infty} |(y, x^{(k)})|^2 + \|u\|^2.$$

Nous pouvons affirmer [II-2-14] que *le signe d'égalité dans l'inégalité* (259) *équivaut à un vecteur* u *nul dans la formule* (260).

Un système orthonormé (257) est dit *complet* s'il n'existe pas de vecteur, distinct du vecteur nul, qui soit orthogonal à tous les $x^{(k)}$; il est dit *fermé* si pour tout vecteur y de l_2 l'équation de fermeture

$$\|y\|^2 = \sum_{k=1}^{\infty} |(y, x^{(k)})|^2 \tag{261}$$

est satisfaite. Si le système (257) est complet, le vecteur u de la formule (260) est alors vecteur nul et nous avons le développement :

$$y = \sum_{k=1}^{\infty} (y, x^{(k)}) x^{(k)}, \tag{262}$$

d'où, comme nous l'avons vu dans [II-2-15], l'équation de fermeture (261). Inversement, supposons le système fermé et le vecteur u orthogonal à tous les $x^{(k)}$, c'est-à-dire $(u, x^{(k)}) = 0$ $(k = 1, 2, \ldots)$. Il vient de la formule (261) que $\|u\|^2 = 0$ lorsqu'on remplace y par u, c'est-à-dire que u est le vecteur nul. Ainsi, *la complétion et la fermeture sont des notions équivalentes*. Et il découle de ce qui précède que *la complétion (fermeture) d'un système* $x^{(k)}$ *est la possibilité de représenter tout vecteur* y *sous forme de série* (262).

Supposons que le système (257) soit fermé et établissons pour tout couple de vecteurs y et z de l_2 l'équation généralisée de fermeture. Il est aisé de voir que pour les vecteurs y + z et y + iz les coefficients de Fourier sont respectivement $(y, x^{(k)}) + (z, x^{(k)})$ et $(y, x^{(k)}) + i(z, x^{(k)})$. Appliquons à y + z et y + iz l'équation de fermeture :

$$\sum_{k=1}^{\infty} [(y, x^{(k)}) + (z, x^{(k)})] [\overline{(y, x^{(k)})} + \overline{(z, x^{(k)})}] = (y + z, y + z),$$

$$\sum_{k=1}^{\infty} [(y, x^{(k)}) + i(z, x^{(k)})] [\overline{(y, x^{(k)})} - i\overline{(z, x^{(k)})}] = (y + iz, y + iz).$$

Tenant compte de l'équation de fermeture pour y et z, nous avons :

$$\sum_{k=1}^{\infty}(\mathbf{y},\ \mathbf{x}^{(k)})\overline{(\mathbf{z},\ \mathbf{x}^{(k)})}+\sum_{k=1}^{\infty}(\mathbf{z},\ \mathbf{x}^{(k)})\overline{(\mathbf{y},\ \mathbf{x}^{(k)})}=(\mathbf{y},\ \mathbf{z})+(\mathbf{z},\ \mathbf{y}).$$

$$\sum_{k=1}^{\infty}(\mathbf{y},\ \mathbf{x}^{(k)})\overline{(\mathbf{z},\ \mathbf{x}^{(k)})}-\sum_{k=1}^{\infty}(\mathbf{z},\ \mathbf{x}^{(k)})\overline{(\mathbf{y},\ \mathbf{x}^{(k)})}=(\mathbf{y},\ \mathbf{z})-(\mathbf{z},\ \mathbf{y}),$$

d'où l'équation généralisée de fermeture pour le système orthonormé complet (257):

$$(\mathbf{y},\ \mathbf{z})=\sum_{k=1}^{\infty}(\mathbf{y},\ \mathbf{x}^{(k)})\overline{(\mathbf{z},\ \mathbf{x}^{(k)})}. \tag{261$_1$}$$

Si y coïncide avec z, cette formule coïncide avec la formule (261). Désignons par $x_s^{(k)}$ les composantes des vecteurs $\mathbf{x}^{(k)}$ $(k,\ s=1,\ 2,\ldots)$ du système orthonormé (257) et écrivons-les sous forme d'une matrice infinie:

$$\left\Vert\begin{matrix} x_1^{(1)} & x_1^{(2)} & x_1^{(3)}\ldots \\ x_2^{(1)} & x_2^{(2)} & x_2^{(3)}\ldots \\ x_3^{(1)} & x_3^{(2)} & x_3^{(3)}\ldots \\ \ldots & \ldots & \ldots \end{matrix}\right\Vert. \tag{263}$$

De la condition d'orthonormalité (257$_1$) il découle les formules:

$$\sum_{s=1}^{\infty}x_s^{(i)}\overline{x_s^{(k)}}=\delta_{ik}\qquad(i,\ k=1,\ 2,\ \ldots),$$

c'est-à-dire *l'orthonormalité des colonnes de la matrice* (263). Trouvons maintenant pour quelles conditions le système orthonormé (257) est complet.

Etablissons d'abord la condition nécessaire de complétion. Supposons que le système (257) soit complet et appliquons aux vecteurs $\mathbf{a}^{(k)}$ $(0,\ 0,\ \ldots,\ 1,\ 0,\ \ldots)$, que nous avons introduits dans [II-2-14], les équations de fermeture (261) et (261$_1$). Tenant compte de ce que $(\mathbf{a}^{(k)},\ \mathbf{x}^{(l)})=\overline{x_k^{(l)}}$ et que les $\mathbf{a}^{(k)}$ forment un système orthonormé, nous avons:

$$\sum_{s=1}^{\infty}x_i^{(s)}\overline{x_k^{(s)}}=\delta_{ik},$$

c'est-à-dire que pour que le système orthonormé des $\mathbf{x}^{(k)}$ soit complet, il est nécessaire que les lignes de la matrice (263) soient orthonormées. Montrons que pour la complétion il suffit d'exiger que les lignes soient normées:

$$\sum_{s=1}^{\infty}|x_k^{(s)}|^2=1. \tag{264}$$

Supposons que cette condition soit satisfaite et montrons que le système des $\mathbf{x}^{(k)}$ est complet. La formule (264) est, en vertu de $(\mathbf{a}^{(k)}, \mathbf{x}^{(s)}) = \overline{x_k^{(s)}}$ et de $\| \mathbf{a}^{(k)} \| = 1$, l'équation de fermeture pour les vecteurs $\mathbf{a}^{(k)}$ et, par conséquent, nous avons:

$$\mathbf{a}^{(k)} = \sum_{s=1}^{\infty} (\mathbf{a}^{(k)}, \mathbf{x}^{(s)}) \mathbf{x}^{(s)}. \tag{265}$$

En outre, les coefficients de Fourier de tout vecteur $\mathbf{y}$ $(y_1, y_2, \ldots)$ par rapport au système orthonormé des $\mathbf{a}^{(k)}$ coïncident avec les composantes $y^{(k)}$ $(k = 1, 2, \ldots)$ et l'équation de fermeture pour $\mathbf{y}$ coïncide tout simplement avec la définition de la norme de $\mathbf{y}$ (236), c'est-à-dire que le système des $\mathbf{a}^{(k)}$ est complet. Soit $\mathbf{v}$ un vecteur orthogonal à tous les $\mathbf{x}^{(k)}$. Montrons que $\mathbf{v}$ est un vecteur nul. De (265) il découle que $(\mathbf{a}^{(k)}, \mathbf{v}) = 0$ $(k = 1, 2, \ldots)$. Puisque les $\mathbf{a}^{(k)}$ forment un système complet, il en découle que $\mathbf{v}$ est le vecteur nul. Ainsi, la complétion du système orthonormé des $\mathbf{x}^{(k)}$ est démontrée à condition que (264) soit satisfait. Enonçons le résultat obtenu: *pour qu'un système orthonormé des* $\mathbf{x}^{(k)}$ *soit complet (fermé), il faut et il suffit (les lignes de la matrice* (263) *étant normées) que la formule* (264) *soit vérifiée.* Si cette condition est satisfaite, on a l'orthogonalité des lignes de la matrice (263).

II-2-17. Transformations linéaires à ensemble infini de variables. Considérons succinctement les transformations linéaires à ensemble infini de variables:

$$\left.\begin{array}{l} x_1' = a_{11}x_1 + a_{12}x_2 + \ldots, \\ x_2' = a_{21}x_1 + a_{22}x_2 + \ldots, \\ \ldots\ldots\ldots\ldots\ldots\ldots \end{array}\right\} \tag{266}$$

ou

$$\mathbf{x}' = A\mathbf{x}, \tag{267}$$

où A est une matrice infinie d'éléments a_{ik}. Posons d'abord comme condition que les séries infinies des seconds membres des égalités (266) sont convergentes pour tout vecteur $\mathbf{x}$ de l'espace l_2. Comme nous le savons, cette condition est satisfaite si les séries

$$\sum_{k=1}^{\infty} |a_{ik}|^2 \qquad (i = 1, 2, \ldots)$$

sont convergentes pour chaque i. On peut montrer que cette condition suffisante est aussi nécessaire. Si cette condition n'est pas satisfaite en effet, les séries des seconds membres des égalités (266) ne sont pas convergentes dans tout l'espace l_2, mais seulement dans un sous-ensemble de celui-ci.

Ainsi, il est naturel de poser comme condition que les nombres x'_k, résultats de la transformation (266), soient aussi les composantes d'un vecteur de l'espace l_2 si les x_k sont les composantes d'un vecteur quelconque, c'est-à-dire que la série

$$\sum_{k=1}^{\infty} |x'_k|^2$$

soit convergente dès que la série

$$\sum_{k=1}^{\infty} |x_k|^2$$

l'est.

Si la matrice A satisfait aux conditions énoncées ci-dessus, la transformation correspondante A est appelée *transformation bornée*. Ce terme signifie que pour une telle transformation on peut montrer l'existence d'un nombre positif M tel que:

$$\|\mathbf{x}'\|^2 \leqslant M \|\mathbf{x}\|^2 \tag{268}$$

ou sous forme développée:

$$\sum_{k=1}^{\infty} |x'_k|^2 \leqslant M \sum_{k=1}^{\infty} |x_k|^2. \tag{269}$$

Arrêtons-nous à un cas particulier des transformations linéaires. Considérons la transformation linéaire:

$$\left.\begin{array}{l} x'_1 = u_{11}x_1 + u_{12}x_2 + \ldots, \\ x'_2 = u_{21}x_1 + u_{22}x_2 + \ldots, \\ \cdots\cdots\cdots\cdots\cdots \end{array}\right\} \tag{270}$$

supposons, en outre, que les séries

$$\sum_{k=1}^{\infty} |u_{ik}|^2$$

soient convergentes pour tout i. Introduisons les vecteurs $\mathbf{u}^{(k)}$ de composantes: $\bar{u}_{k1}$, $\bar{u}_{k2}$, ..., et supposons que les coefficients u_{ik} soient tels que les vecteurs $\mathbf{u}^{(k)}$ forment un système orthonormé complet. Comme nous l'avons montré ci-dessus, on a alors orthogonalité et normalité du tableau des u_{ik} suivant les lignes et les colonnes, c'est-à-dire que:

$$\sum_{s=1}^{\infty} u_{sp}\bar{u}_{sq} = \delta_{pq}, \quad \sum_{s=1}^{\infty} u_{ps}\bar{u}_{qs} = \delta_{pq}. \tag{271}$$

Dans ce cas, la transformation correspondante (270) est dite *unitaire*.

Nous pouvons écrire les égalités (270) sous la forme:

$$\left.\begin{aligned}(\mathbf{x},\ \mathbf{u}^{(1)}) &= x_1',\\ (\mathbf{x},\ \mathbf{u}^{(2)}) &= x_2',\\ \dots&\dots\end{aligned}\right\} \tag{272}$$

La formule de fermeture donne:

$$\sum_{k=1}^{\infty} |x_k'|^2 = \|\mathbf{x}\|^2 = \sum_{k=1}^{\infty} |x_k|^2,$$

c'est-à-dire que de même que dans le cas d'un nombre fini de dimensions, il n'y a pas d'altération de la longueur des vecteurs par la transformation unitaire et nous pouvons prendre $M = 1$ dans la formule (268).

Le système (270) ou, ce qui revient au même, le système (272) peut être facilement résolu par rapport aux x_k si les nombres donnés x_k' sont tels que la série des carrés de leurs modules converge. Tenant compte de ce que les vecteurs $\mathbf{u}^{(k)}$ $(\bar{u}_{k1}, \bar{u}_{k2}, \dots)$ forment, par hypothèse, un système orthonormé complet, nous obtenons en vertu de (272):

$$\mathbf{x} = x_1'\mathbf{u}^{(1)} + x_2'\mathbf{u}^{(2)} + \dots \tag{273}$$

ou:

$$\left.\begin{aligned}x_1 &= \bar{u}_{11}x_1' + \bar{u}_{21}x_2' + \dots,\\ x_2 &= \bar{u}_{12}x_1' + \bar{u}_{22}x_2' + \dots,\\ \dots&\dots\dots\end{aligned}\right\} \tag{274}$$

Ces formules montrent que la transformation inverse d'une transformation unitaire s'obtient en échangeant les lignes et les colonnes et en remplaçant tous les coefficients par leurs conjugués complexes, c'est-à-dire qu'il y a analogie complète ici avec le cas d'un nombre fini de dimensions.

Dans le cas général, et même pour les matrices bornées, le problème de la matrice inverse et de la réduction d'une matrice à la forme diagonale présente de grandes difficultés et conduit à des résultats qui n'ont pas d'analogues exacts dans le cas des espaces à un nombre fini de dimensions. L'étude détaillée de la théorie des matrices infinies est donnée dans le tome V. Nous nous bornerons ici à exposer quelques résultats seulement. Précisons la condition nécessaire et suffisante sur les a_{ik} pour que la formule (266) soit une transformation bornée. Elle s'énonce ainsi: *il existe un nombre positif M tel que pour tout entier positif l et pour tous nombres complexes x_s $(s = 1, 2, \dots)$ l'inégalité suivante soit vérifiée*:

$$\left|\sum_{i,k=1}^{l} a_{ik}x_i\bar{x}_k\right| \leqslant M \sum_{k=1}^{l} |x_k|^2.$$

On démontre aussi la condition suffisante suivante pour qu'une transformation (266) soit bornée: *il existe un nombre positif p (indépendant de i et k) tel que les inégalités suivantes soient vérifiées*:

$$\sum_{i=1}^{\infty} |a_{ik}| \leqslant p \qquad \sum_{k=1}^{\infty} |a_{ik}| \leqslant p$$
$$(k=1,\ 2,\ \ldots). \qquad (i=1,\ 2,\ \ldots).$$

Si la matrice A définit une transformation bornée (266), il existe alors une matrice unique $\tilde{A}$ définissant également une transformation bornée telle que pour tous x et y l'égalité

$$(A\mathrm{x},\ \mathrm{y}) = (\mathrm{x},\ \tilde{A}\mathrm{y})$$

soit satisfaite et les éléments de la matrice $\tilde{A}$ s'expriment à l'aide des éléments de la matrice A suivant les formules $\tilde{a}_{ik} = \bar{a}_{ki}$. Si $\tilde{A}$ coïncide avec A, c'est-à-dire si $a_{ik} = \bar{a}_{ki}$, la transformation bornée (266) (ou la matrice A) est dite *auto-adjointe*.

Pour les transformations bornées on a la formule:

$$(A\mathrm{x},\ \mathrm{y}) = \sum_{n=1}^{\infty} \left(\sum_{m=1}^{\infty} a_{nm}x_m\right) \bar{y}_n = \sum_{m=1}^{\infty} x_m \left(\sum_{n=1}^{\infty} a_{nm}\bar{y}_n\right) =$$
$$= \lim_{\substack{k\to\infty \\ l\to\infty}} \sum_{m=1}^{k} \sum_{n=1}^{l} a_{nm}x_m\bar{y}_n.$$

Notons un cas particulier important des applications ou opérateurs bornés, à savoir celui où la série double

$$\sum_{n,\ m=1}^{\infty} |a_{nm}|^2 \tag{275}$$

converge. Alors la série double

$$\sum_{n,\ m=1}^{\infty} a_{nm}x_m\bar{y}_n$$

converge absolument pour tout choix des vecteurs x $(x_1,\ x_2,\ \ldots)$ et y $(y_1,\ y_2,\ \ldots)$. Si en plus de la convergence de la série (275) nous avons $a_{ik} = \bar{a}_{ki}$, il est alors possible de réduire la forme hermitienne à une somme de carrés à l'aide d'une transformation unitaire:

$$\sum_{n,\ m=1}^{\infty} a_{nm}x_m\bar{x}_n = \sum_{k=1}^{\infty} \lambda_k z_k \bar{z}_k,$$

où les λ_k sont des nombres réels et le vecteur z $(z_1,\ z_2,\ \ldots)$ s'obtient du vecteur x $(x_1,\ x_2,\ \ldots)$ en lui appliquant une transformation unitaire: $z = Ux$. Dans ce cas, $\lambda_k \to 0$ pour $k \to \infty$ (il peut arriver que $\lambda_k = 0$ pour tous les k suffisamment grands). Si A et B sont deux matrices infinies engendrant des transformations bornées, leur produit donne une transformation également bornée dont les coefficients se calculent d'après les formules ordinaires:

$$\{BA\}_{ik} = \sum_{s=1}^{\infty} \{B\}_{is}\{A\}_{sk}.$$

Remarquons encore que si une suite de vecteurs $\mathbf{x}^{(k)}$ a pour limite le vecteur $\mathbf{x}$, c'est-à-dire si $\mathbf{x}^{(k)} \Rightarrow \mathbf{x}$, alors $A\mathbf{x}^{(k)} \Rightarrow A\mathbf{x}$, si A est une matrice engendrant une transformation bornée.

Un rôle très important dans les applications à la physique mathématique revient aussi aux transformations linéaires non bornées. Leur étude sera donnée dans le tome V.

II-2-18. Espace fonctionnel L_2. Nous avons considéré l'espace l_2 dans lequel un vecteur quelconque est défini par un ensemble infini de composantes numérotées par les nombres entiers: x_1 pour la première composante, x_2 pour la deuxième, etc.

Passons maintenant à la considération d'une famille $L(\mathscr{E})$ de fonctions à valeurs complexes $f(x) = f_1(x) + if_2(x)$ définies sur un ensemble mesurable $\mathscr{E}$ et telles que les fonctions réelles $f_1(x)$ et $f_2(x)$ soient mesurables sur $\mathscr{E}$ et appartiennent à $L_2(\mathscr{E})$ (tome II, [VI-1-8, 9, 10]), d'où il découle que $|f(x)|$ est une fonction mesurable sur $\mathscr{E}$ et appartient à $L_2(\mathscr{E})$. Si les parties réelles et imaginaires de deux fonctions sont respectivement équivalentes, on dit alors que ces fonctions sont *équivalentes* et qu'elles s'identifient en tant qu'éléments de l'espace fonctionnel $L_2(\mathscr{E})$, c'est-à-dire qu'un élément quelconque de cet espace est l'ensemble de toutes les fonctions équivalentes et une fonction quelconque de cet ensemble un représentant de cet ensemble. L'élément nul de $L_2(\mathscr{E})$ est la fonction équivalente à la fonction nulle sur $\mathscr{E}$, par exemple la fonction identiquement nulle sur $\mathscr{E}$. Dans la suite nous écrirons L_2 au lieu de $L_2(\mathscr{E})$.

Les propriétés de l'espace fonctionnel L_2 sont analogues à celles de l_2. Les éléments de L_2 peuvent être multipliés par les nombres complexes et additionnés (tome II, [VI-1-8]). Le produit scalaire de deux éléments $f(x)$ et $g(x)$ est donné par la formule:

$$(f, g) = \int_{\mathscr{E}} f(x)\,\overline{g(x)}\,dx \tag{276}$$

et le carré de la norme de l'élément $f(x)$ par la formule:

$$\|f\|^2 = (f, f) = \int_{\mathscr{E}} |f(x)|^2\,dx. \tag{277}$$

La définition de l'orthogonalité est la même que celle donnée pour l_2 et les résultats de [II-2-14] sont valables sans aucune modification. La convergence d'une suite d'éléments $f_n(x)$ vers l'élément $f(x)$ est définie par la formule :

$$\lim_{n\to\infty}\int_{\mathscr{B}}|f(x)-f_n(x)|^2\,dx=0, \tag{278}$$

les définitions et les résultats de [II-2-15] et [II-2-16] restent valables sans aucune modification en tenant compte de (tome II, [VI-1-8, 9, 10]). Remarquons que si a_k et b_k sont les coefficients de Fourier des éléments $f(x)$ et $g(x)$ par rapport à un système orthonormé complet $\varphi_k(x)$ $(k = 1, 2, \ldots)$, c'est-à-dire si :

$$a_k=\int_{\mathscr{B}}f(x)\,\overline{\varphi_k(x)}\,dx\,; \qquad b_k=\int_{\mathscr{B}}g(x)\,\overline{\varphi_k(x)}\,dx, \tag{279}$$

on a alors, de même que dans [II-2-16], l'équation généralisée de fermeture :

$$\int_{\mathscr{B}}f(x)\,\overline{g(x)}\,dx=\sum_{k=1}^{\infty}a_k\bar{b}_k. \tag{280}$$

Il découle de ce qui vient d'être dit que la convergence dans L_2 de la série

$$\psi(x)=\psi_1(x)+\psi_2(x)+\ldots$$

n'est autre que la convergence en moyenne :

$$\lim_{n\to\infty}\int_{\mathscr{B}}|f(x)-s_n(x)|^2\,dx=0,$$

où $s_n(x)$ est la somme des n premiers termes de la série.

II-2-19. Lien entre les espaces l_2 et L_2. Soit, comme précédemment,

$$\varphi_k(x) \qquad (k=1, 2, \ldots) \tag{281}$$

un système orthonormé complet dans L_2. A tout élément de L_2 correspond une suite infinie a_k des coefficients de Fourier dont la somme des carrés des modules converge, et réciproquement, à toute suite infinie de nombres complexes c_k $(k = 1, 2, \ldots)$, dont la somme des carrés des modules converge, correspond un élément déterminé de L_2 pour lequel les c_k sont les coefficients de Fourier par rapport au système (281) (tome II, [VI-1-10]). Ainsi, ce système établit une correspondance biunivoque entre les éléments de L_2 et de l_2 : à tout élément de L_2 correspond un élément déterminé de l_2 et inversement (tome II, [VI-1-10]). L'addition, la multiplication par un nombre, le produit scalaire (en vertu de (280)) et la norme sont de même conservés. Si une suite d'éléments $f_n(x)$ de L_2 converge vers un élément $f(x)$, c'est-à-dire si $\| f - f_n \| \to 0$ pour $n \to \infty$, il en est de même pour les éléments correspondants de l_2, ce qui découle de l'égalité des normes:

$$\sum_{k=1}^{\infty} |a_k - a_k^{(n)}|^2 = \int_{\mathscr{B}} |f(x) - f_n(x)|^2 \, dx,$$

où a_k et $a_k^{(n)}$ sont les composantes des éléments de l_2 correspondant respectivement à $f(x)$ et $f_n(x)$.

Pour établir la correspondance entre les espaces L_2 et l_2 nous sommes partis d'un système orthonormé complet déterminé (281). Si l'on considère un autre système semblable:

$$\psi_k(x) \qquad (k = 1, 2, \ldots), \tag{282}$$

il est évident que la loi de correspondance est autre. Chaque fonction du système (282) se développe en série de Fourier suivant les $\varphi_k(x)$:

$$\psi_k(x) = \sum_{i=1}^{\infty} u_{ik}\varphi_i(x) \qquad (k = 1, 2, \ldots), \tag{283}$$

en outre, la série du second membre converge dans L_2. Tenant compte de l'équation généralisée de fermeture, nous avons:

$$\sum_{s=1}^{\infty} u_{si}\overline{u}_{sk} = \delta_{ik} \qquad (i, k = 1, 2, \ldots),$$

c'est-à-dire que la matrice U d'éléments u_{ik} est orthonormée suivant les lignes. Tenant compte de ce que

$$u_{ik} = \int_{\mathscr{B}} \psi_k(x) \overline{\varphi_i(x)} \, dx$$

et que le système (281) est complet, nous avons:

$$\sum_{i=1}^{\infty} |u_{ik}|^2 = \int_{\mathscr{B}} |\psi_k(x)|^2 \, dx = 1,$$

d'où il découle que le tableau U est celui d'une transformation unitaire. Il est aisé de démontrer que, inversement, tout tableau d'une transformation unitaire U conduit, conformément à (283), à un système orthonormé complet (282), s'il en est ainsi pour le système (281). Donnons, à titre d'exemple, pour l'espace L_2 de fonctions d'une variable indépendante sur l'intervalle ouvert $(-\pi, \pi)$, le système de fonctions (tome II, [VI-3-2]):

$$\varphi_k(x) = \frac{1}{\sqrt{2\pi}} e^{ikx} \qquad (k = 1, 2, \ldots). \tag{284}$$

On démontre aisément que ce système est orthonormé. Notons que la numérotation en k ne se fait pas de 1 à ∞, mais de $(-\infty)$ à $(+\infty)$. Ce changement de numérotation n'est pas important. D'après les résultats de (tome II, [VI-2-6]) on peut montrer que le système (284) est complet.

II-2-20. Opérateurs linéaires dans L_2. Supposons qu'il existe une loi bien définie d'après laquelle à chaque fonction $f(x)$ de $L_2(\mathscr{E})$ correspond une autre fonction $F(x)$ également de $L_2(\mathscr{E})$:

$$F(x) = A[f(x)], \tag{285}$$

où A désigne symboliquement cette loi de correspondance. Nous avons ici, pour ainsi dire, une généralisation de la notion de fonction: le rôle de la variable indépendante n'est plus celui d'un nombre arbitraire appartenant à un ensemble quelconque, par exemple, à un intervalle, mais celui d'une fonction arbitraire de $L_2(\mathscr{E})$ et la valeur de la fonction $f(x)$ est aussi une certaine fonction de $L_2(\mathscr{E})$. La correspondance établie à l'aide de la formule (285) est appelée d'ordinaire application ou *opérateur fonctionnel*. L'opérateur A est dit *linéaire*, si

$$\left.\begin{aligned} A[f(x)+g(x)] &= A[f(x)] + A[g(x)]; \\ A[\alpha f(x)] &= \alpha A[f(x)], \end{aligned}\right\} \tag{286}$$

où α est un nombre complexe arbitraire, et *borné* s'il existe un nombre positif M tel que:

$$\| Af \| \leqslant M \| f \|, \tag{287}$$

pour toute $f(x)$ de $L_2(\mathscr{E})$.

Si la suite $f_n(x)$ converge vers $f(x)$ dans L_2 et si A est un opérateur linéaire borné, alors $Af_n(x)$ converge vers $Af(x)$ dans L_2. Cela découle directement de l'inégalité:

$$\| A[f(x) - f_n(x)] \| \leqslant M \| f - f_n \|,$$

c'est-à-dire de

$$\| Af - Af_n \| \leqslant M \| f - f_n \|.$$

On peut énoncer donc, pour le cas considéré, les définitions des transformations hermitienne et unitaire. Un opérateur linéaire borné A est dit *hermitien* (ou *auto-adjoint*) si:

$$(Af, g) = (f, Ag), \tag{288}$$

pour toutes $f(x)$ et $g(x)$ de L_2. L'opérateur linéaire U est dit *unitaire* s'il transforme biunivoquement L_2 en lui-même:

$$U[f(x)] = F(x), \tag{289}$$

et n'altère pas la norme:

$$\| Uf(x) \| = \| f \|.$$

La biunivocité de la transformation (289) qui vient d'être mentionnée se ramène à ce qui suit: non seulement à chaque élément $f(x)$ de L_2 correspond un élément déterminé $F(x)$, mais à chaque $F(x)$ de L_2 correspond une image réciproque $f(x)$ bien déterminée. Il en découle que U a un opérateur inverse U^{-1} qui donne $f(x)$ à l'aide de $F(x)$. On voit aisément que U^{-1} est aussi un opérateur unitaire. Pour un opérateur unitaire l'inégalité (287) peut être remplacée par une égalité en posant $M = 1$. On peut montrer aisément que la transformation unitaire n'altère pas ni la norme, ni le produit scalaire, c'est-à-dire que:

$$(Uf, Ug) = (f, g)$$

pour toutes $f(x)$ et $g(x)$ de L_2 et que U transforme chaque système orthonormé complet de L_2 en un système analogue. Définissons encore la notion d'opérateur transposé ou adjoint. L'opérateur A^* est dit *transposé* de l'opérateur linéaire borné A si pour toutes $f(x)$ et $g(x)$ de L_2 on a l'égalité:

$$(Af, g) = (f, A^*g).$$

On peut montrer que pour chaque opérateur linéaire borné A il existe un opérateur transposé A^* unique. Considérons l'espace L_2 des fonctions $f(x)$ d'une variable indépendante définies sur un intervalle ouvert (a, b), fini ou infini. Dans un tel L_2 les opérateurs linéaires sont souvent définis par la formule:

$$F(x) = \int_a^b K(x, t) f(t)\, dt, \tag{290}$$

où $K(x, t)$ est une fonction définie (mesurable) dans le carré $a \leqslant x \leqslant b$, $a \leqslant t \leqslant b$. Les opérateurs de la forme (290) sont souvent appelés *opérateurs intégraux* et la fonction $K(x, t)$ *noyau de l'opérateur*. Si

$$\int_a^b \int_a^b |K(x, t)|^2\, dx\, dt < +\infty,$$

ou s'il existe un nombre p tel que $\int_a^b |K(x, t)|\, dt \leqslant p$ et

$\int_a^b |K(x, t)|dx \leqslant p$ pour tous les t, x de (a, b), alors la formule (290) définit un opérateur linéaire borné. On peut montrer que l'opérateur transposé d'un opérateur borné (290) est aussi un opérateur intégral de noyau $K^*(x, t) = \overline{K(t, x)}$. Un exemple de transformation unitaire dans L_2 sur l'intervalle $(-\infty, +\infty)$ est la transformation de Fourier:

$$F(x) = Uf(x) = \lim_{\substack{n \to \infty \\ m \to \infty}} \frac{1}{\sqrt{2\pi}} \int_{-n}^{m} e^{-ixt} f(t)\, dt,$$

où lim est la limite de la fonction en x obtenue comme résultat de l'intégration dans L_2 sur l'intervalle $(-\infty, +\infty)$, c'est-à-dire que si nous posons:

$$\varphi_{n,m}(x) = \frac{1}{\sqrt{2\pi}} \int_{-n}^{m} e^{-ixt} f(t)\, dt,$$

alors:

$$\lim_{\substack{n \to \infty \\ m \to \infty}} \int_{-\infty}^{+\infty} |F(x) - \varphi_{n,m}(x)|^2\, dx = 0.$$

Si $f(x)$, outre qu'elle appartient à L_2 sur l'intervalle $(-\infty, +\infty)$, est sommable sur cet intervalle, la transformation de Fourier peut s'écrire sous la forme habituelle:

$$F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-ixt} f(t)\, dt.$$

La conservation de la norme pour la transformation de Fourier s'écrit comme suit:

$$\int_{-\infty}^{+\infty} |F(x)|^2\, dx = \int_{-\infty}^{+\infty} |f(x)|^2\, dx.$$

Voyons à présent un exemple particulier. Soit dans l'espace L_2 sur l'intervalle ouvert $(-\pi, \pi)$ l'opérateur linéaire de multiplication par la variable indépendante:

$$A[f(x)] = xf(x). \tag{291}$$

On a évidemment:

$$\|Af\|^2 = \int_{-\pi}^{\pi} |x|^2 |f(x)|^2\, dx \leqslant \pi^2 \|f\|^2,$$

c'est-à-dire que pour l'opérateur (291) on peut prendre $M = \pi$ dans la formule (287). Construisons une transformation linéaire exprimant l'opérateur (291) dans l_2 si l'on prend pour base le système orthonormé complet (284). Soient c_k les coefficients de Fourier de $f(x)$ par rapport au système (284):

$$c_k = \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} e^{-ikx} f(x)\, dx \qquad (k = 0, \pm 1, \pm 2, \ldots)$$

et c'_k les coefficients de Fourier de $xf(x)$:

$$c'_k = \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} e^{-ikx} x f(x)\, dx \qquad (k = 0, \pm 1, \pm 2, \ldots).$$

Il nous faut construire une transformation linéaire (matrice infinie) qui exprime c'_k en fonction de c_k.

Calculons les coefficients de Fourier de la fonction $\frac{1}{\sqrt{2\pi}} e^{ikx} x$:

$$b_m = \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} e^{i(k-m)x} x\, dx \qquad (m = 0, \pm 1, \pm 2, \ldots).$$

Intégrant par parties, nous avons:

$$b_m = \frac{(-1)^{k-m}}{i(k-m)} \quad \text{si} \quad m \neq k; \; b_k = 0.$$

Ecrivant l'expression de c'_k sous la forme:

$$c'_k = \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} \overline{e^{ikx} x} f(x)\, dx$$

et appliquant l'équation généralisée de fermeture, nous trouvons:

$$c'_k = i \sum_{m=-\infty}^{+\infty}{}' \frac{(-1)^{k-m}}{k-m} c_m \qquad (k = 0, \pm 1, \pm 2, \ldots), \quad (292)$$

où dans Σ' le terme correspondant à $m = k$ est exclu. La formule (292) donne la transformation linéaire dans l_2 qui correspond à l'opérateur (291) dans L_2 si l'on prend les fonctions (284) pour fonctions de coordonnées dans l'espace L_2. On montre aisément que l'opérateur (291) est auto-adjoint.

Donnons le raisonnement général pour un opérateur linéaire borné auto-adjoint quelconque A (nous écrirons les coefficients de Fourier sous forme de produit scalaire et nous tiendrons compte de la définition (288) de l'opérateur auto-adjoint). Introduisons les coefficients de Fourier c_k et c'_k de $f(x)$ et $Af(x)$ par rapport au système orthonormé complet $\varphi_k(x)$:

$$c_k = (f, \varphi_k); \qquad c'_k = (Af, \varphi_k) = (f, A\varphi_k). \tag{293}$$

Introduisons aussi les coefficients de Fourier des fonctions $A\varphi_m(x)$:

$$a_{km} = (A\varphi_m, \varphi_k) = (\varphi_m, A\varphi_k), \tag{294}$$

il en découle que $a_{mk} = \overline{a_{km}}$, car par définition:

$$a_{mk} = (A\varphi_k, \varphi_m).$$

Tenant compte de l'équation généralisée de fermeture, de la formule pour les c'_k et de (294), nous avons:

$$c'_k = \sum_{m=1}^{\infty} (f, \varphi_m)\overline{(A\varphi_k, \varphi_m)} = \sum_{m=1}^{\infty} (f, \varphi_m)(\varphi_m, A\varphi_k),$$

c'est-à-dire que:

$$c'_k = \sum_{m=1}^{\infty} a_{km}c_m \qquad (k = 1, 2, \ldots).$$

Cette transformation représente l'opérateur A dans l_2 si les $\varphi_k(x)$ sont pris pour fonctions de coordonnées dans L_2.

Nous avons considéré dans les paragraphes précédents trois cas où l'opérateur linéaire est donné sur L_2 entier et est borné. Si nous considérons, par exemple, l'opérateur de dérivation $A[f(x)] = f'(x)$, il n'est pas donné sur L_2 entier, car toute fonction de L_2 n'a pas nécessairement de dérivée. De plus, cet opérateur n'est pas borné sur l'ensemble des fonctions sur lequel il est donné.

Un exposé plus détaillé et plus rigoureux des questions considérées dans ces derniers paragraphes sera donné dans le tome V.

Chapitre III

ÉLÉMENTS DE LA THÉORIE DES GROUPES ET REPRÉSENTATIONS LINÉAIRES DES GROUPES

III-1. Eléments de la théorie des groupes

III-1-1. Groupes des transformations linéaires. On considère l'ensemble de toutes les transformations unitaires dans un espace à n dimensions. Toutes ces transformations ont un déterminant non nul de sorte qu'à chaque transformation unitaire $U\mathbf{x}$, entièrement caractérisée par la matrice correspondante U, correspond une transformation inverse entièrement définie $U^{-1}\mathbf{x}$ qui est également unitaire [II-1-9]. De plus, si $U_1\mathbf{x}$ et $U_2\mathbf{x}$ sont deux transformations unitaires, leur produit $U_2U_1\mathbf{x}$ est également une transformation unitaire. Toutes ces propriétés de l'ensemble des transformations unitaires s'expriment en disant que *l'ensemble des transformations unitaires forme un groupe.*

En général, l'ensemble de certaines transformations linéaires dont le déterminant n'est pas nul forme un groupe si les conditions suivantes sont vérifiées: 1) *si une certaine transformation appartient à l'ensemble, la transformation inverse appartient également à l'ensemble*; 2) *le produit de deux transformations appartenant à l'ensemble (quel que soit l'ordre des facteurs) appartient également à l'ensemble et les facteurs peuvent être identiques.*

En tenant compte de ce que le produit de toute transformation par son inverse est la transformation identique, on peut affirmer qu'*un groupe contient forcément la transformation identique, c'est-à-dire la matrice unité.*

En général, la transformation linéaire est entièrement définie par sa matrice et dans tout ce qui précède, de même que dans ce qui suit, on peut parler de *groupe des transformations linéaires* ou de *groupe des matrices.*

Donnons encore des exemples de groupes des transformations linéaires. Il est facile de voir que l'ensemble de toutes les transformations réelles orthogonales forme un groupe. On sait que ces transformations réelles orthogonales ont un déterminant égal à (± 1). Si l'on prend l'ensemble des transformations réelles orthogonales dont le déterminant est égal à $(+1)$, elles forment également un groupe. Mais si l'on prend l'ensemble des transformations réelles orthogonales dont le déterminant est égal à (-1), elles ne forment pas un groupe car le produit de deux matrices dont le déterminant est (-1) donne une matrice dont le déterminant est égal à $(+1)$.

En particulier, si l'on considère le groupe des transformations réelles orthogonales à trois variables, ce sera un groupe constitué de rotations de l'espace autour de l'origine et de transformations qui s'obtiennent à la suite d'une telle rotation et d'une symétrie par rapport à l'origine. Si l'on prend le groupe des transformations linéaires orthogonales à trois variables, à déterminant égal à $(+1)$, ce sera le groupe des rotations de l'espace autour de l'origine.

Dans tous les cas considérés, le groupe contient un nombre infini de transformations, en particulier le groupe des rotations de l'espace à trois dimensions autour de l'origine dépend de trois paramètres réels arbitraires: les angles d'Euler dont il a déjà été question.

Comme autre exemple, nous considérons les rotations de l'espace d'un angle φ autour de l'axe Z. Les formules correspondantes sont de la forme:

$$\left.\begin{aligned} x' &= x\cos\varphi - y\sin\varphi, \\ y' &= x\sin\varphi + y\cos\varphi. \end{aligned}\right\} \tag{1}$$

Pour les valeurs possibles du paramètre réel φ dans l'intervalle ouvert $(0,\ 2\pi)$ on obtient, de façon évidente, un groupe contenant un nombre infini de transformations et dépendant d'un paramètre réel. On introduit la notation suivante pour la matrice de transformation:

$$Z_\varphi = \begin{Vmatrix} \cos\varphi, & -\sin\varphi \\ \sin\varphi, & \cos\varphi \end{Vmatrix}. \tag{2}$$

Il est évident que le produit de deux rotations d'angles φ_1 et φ_2 donne une rotation d'un angle $(\varphi_1+\varphi_2)$:

$$Z_{\varphi_2}Z_{\varphi_1} = Z_{\varphi_2+\varphi_1}, \tag{3}$$

et de même:

$$Z_{\varphi_1}Z_{\varphi_2} = Z_{\varphi_1+\varphi_2}.$$

On voit ainsi que, dans ce cas, toutes les transformations du groupe ou, comme on dit, tous les éléments du groupe commutent deux à deux. Un tel groupe est un *groupe abélien*. De plus, dans ce dernier exemple la multiplication de deux éléments du groupe se réduit simplement à l'addition des valeurs du paramètre φ qui correspondent aux matrices à multiplier.

On peut généraliser quelque peu ce dernier groupe en considérant non seulement les rotations du plan XY autour de l'origine, mais aussi les symétries par rapport à l'axe Y. Il est évident que l'ordre dans lequel ont lieu les transformations ne joue aucun rôle. Il agit sur le résultat, mais ce sont toujours des transformations réelles orthogonales à deux variables que l'on obtient. La forme générale de la matrice correspondante est

$$\{\varphi, d\} = \begin{Vmatrix} d\cos\varphi, & -d\sin\varphi \\ \sin\varphi, & \cos\varphi \end{Vmatrix}, \tag{4}$$

où φ est le paramètre précédent et d un nombre égal à ± 1. Si $d = +1$ on a une rotation simple du plan XY autour de l'origine, et si $d = -1$ on obtient une rotation suivie de la symétrie déjà indiquée. Il est facile de vérifier que le produit de matrices (4) obéit à la règle suivante :

$$\{\varphi_2, d_2\}\{\varphi_1, d_1\} = \{\varphi_1 + d_1\varphi_2, d_1 d_2\}. \tag{5}$$

Dans ce cas, le produit peut dépendre de l'ordre des facteurs, c'est-à-dire que le groupe n'est pas abélien. De même, il est évident que le groupe des transformations réelles orthogonales de l'espace à trois dimensions ou même le groupe des rotations de l'espace à trois dimensions autour de l'origine ne sont pas abéliens.

Nous avons donné jusque-là des exemples de groupes contenant un nombre infini de transformations (éléments) et les matrices correspondantes contenaient des paramètres réels arbitraires. Nous donnons maintenant quelques exemples de groupes contenant un nombre fini d'éléments. Soit m un entier positif quelconque. Considérons l'ensemble des rotations du plan XY autour de l'origine d'angles :

$$0, \frac{2\pi}{m}, \frac{4\pi}{m}, \ldots, \frac{2(m-1)\pi}{m}.$$

Nous avons ici en tout m transformations dont les matrices sont :

$$Z_{\frac{2k\pi}{m}} = \begin{Vmatrix} \cos\frac{2k\pi}{m}, & -\sin\frac{2k\pi}{m} \\ \sin\frac{2k\pi}{m}, & \cos\frac{2k\pi}{m} \end{Vmatrix} \qquad (k = 0, 1, \ldots, m-1).$$

Ces transformations forment évidemment un groupe et les éléments de ce groupe sont des puissances entières positives d'une même transformation, à savoir :

$$Z_{\frac{2k\pi}{m}} = \left(Z_{\frac{2\pi}{m}}\right)^k \qquad (k = 0, 1, \ldots, m-1). \tag{6}$$

Un tel groupe fini formé des puissances d'une transformation quelconque est dit *groupe cyclique*.

Si on prend un certain angle φ_0, non multiple de π, les transformations (matrices)

$$Z_{\varphi_0}^k = Z_{k\varphi_0} \qquad (k = 0, \pm 1, \pm 2, \ldots) \tag{7}$$

forment également un groupe. Mais ce groupe contient une infinité d'éléments, car il n'existe aucune valeur entière de k telle que $Z^k_{\varphi_0}$ coïncide avec $Z^0_{\varphi_0} = I$. Le groupe (7) est un groupe infini, mais ses matrices ne contiennent aucun paramètre variant continûment. Dans ce cas, on dit que *le nombre d'éléments du groupe est dénombrable*, c'est-à-dire que l'on peut repérer chaque élément du groupe à l'aide d'un indice égal à un entier, de sorte qu'aux différents éléments correspondent des indices différents et que tout nombre entier est indice d'un élément quelconque. Il n'est pas possible de le faire dans le cas de groupes contenant des paramètres variant continûment.

III-1-2. Groupes des polyèdres réguliers. Donnons encore des exemples de groupes finis. On les forme au moyen de rotations de l'espace à trois dimensions autour de l'origine. De telles rotations, dans un système de coordonnées définies, s'expriment au moyen de transformations linéaires des coordonnées. On notera que lorsqu'on parle de rotation de l'espace autour de l'origine on sous-entend simplement le résultat final du passage d'un état initial à un état transformé. La manière dont on passe de l'un à l'autre n'entre pas en ligne de compte. En effet, toute transformation linéaire définit les coordonnées du point transformé, mais, bien entendu, rien n'indique comment la transformation s'effectue. La manière dont s'effectue la transformation n'entre pas dans les raisonnements que l'on fait.

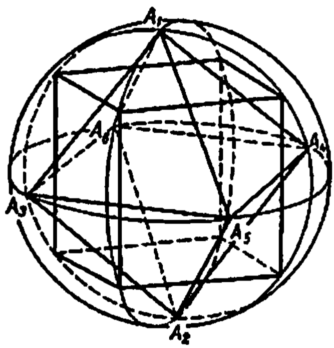


Fig. 2

Soit la sphère de centre à l'origine des coordonnées et de rayon unité. On inscrit dans cette sphère un polyèdre régulier, octaèdre, par exemple (fig. 2). La surface de ce polyèdre est formée de huit triangles équilatéraux. On considère maintenant l'ensemble des rotations de l'espace à trois dimensions autour de l'origine qui amène l'octaèdre considéré à se superposer à lui-même. Il est facile de voir que l'ensemble de ces rotations forme un groupe et que ce groupe contient un nombre fini d'éléments. Comptons le nombre d'éléments de ce groupe. Prenons un axe quelconque de l'octaèdre joignant deux sommets opposés. L'octaèdre coïncide avec lui-même si on le fait tourner autour de cet axe d'un angle 0, $\frac{\pi}{2}$, π, $\frac{3\pi}{2}$. La rotation d'angle 0 correspond évidemment à la transformation identique, c'est-à-dire à la matrice unité. On désigne les quatre rotations ainsi indiquées autour de l'axe considéré par:

$$S_0 = I, \ S_1, \ S_2, \ S_3. \tag{8}$$

Soit A l'un des sommets de l'octaèdre situé sur l'axe considéré. Introduisons les transformations linéaires:

$$T_1, \ T_2, \ T_3, \ T_4, \ T_5,$$

pour lesquelles l'octaèdre coïncide avec lui-même, et son sommet A coïncide avec l'un des cinq autres sommets de l'octaèdre. A côté des quatre rotations (8) on forme les 20 rotations de l'espace autour de l'origine de type:

$$T_kS_0, \ T_kS_1, \ T_kS_2, \ T_kS_3 \qquad (k = 1, 2, 3, 4, 5). \tag{9}$$

Il est facile de vérifier que les 24 rotations (8) et (9) sont différentes. C'est évident d'un point de vue géométrique et cela peut être démontré de la manière suivante: soit

$$T_pS_q = T_{p_1}S_{q_1}. \tag{10}$$

Les transformations S_l correspondent à des rotations autour de l'axe passant par le sommet A et lors de ces transformations A est fixe. Les transformations T_p et T_{p_1} pour différentes valeurs des indices p et p_1 transforment le sommet A en sommets différents et, par suite, il résulte de l'égalité (10) que les indices p et p_1 doivent être identiques, mais alors il est évident à partir de cette égalité, en multipliant à gauche par $T_p^{-1} = T_{p_1}^{-1}$, que les indices q et q_1 doivent aussi être identiques, c'est-à-dire que l'égalité (10) ne peut avoir lieu que lorsque les deux membres sont formés des mêmes facteurs. Par suite, les expressions (8) et (9) donnent 24 rotations différentes qui amènent l'octaèdre sur lui-même. Montrons maintenant que ceci épuise toutes les rotations qui ont cette propriété. En effet, soit V une rotation quelconque qui transforme l'octaèdre en lui-même. Supposons que le sommet A soit amené sur un autre sommet A_j quelconque et soit T_j une des transformations T_k qui transforme également A en A_j. Constituons la transformation $T_j^{-1}V$. Alors l'octaèdre se transforme en lui-même et le sommet A reste fixe. Par conséquent, le sommet opposé reste aussi fixe et la transformation ainsi formée est une des rotations S_l autour de l'axe passant par le sommet A, c'est-à-dire que $T_j^{-1}V = S_l$, d'où $V = T_jS_l$. Autrement dit, toute rotation qui transforme l'octaèdre en lui-même est contenue dans les 24 rotations obtenues. On trouve finalement que *le groupe des rotations dans lequel l'octaèdre se transforme en lui-même contient 24 éléments.*

Il est évident que l'on peut inscrire dans la sphère de rayon unité un cube de façon que les rayons partant du centre des faces de l'octaèdre aient leurs extrémités aux sommets du cube. Il en résulte que le groupe des rotations du cube sera le même que celui de l'octaèdre. Supposons que l'on ait choisi autrement la position de l'octaèdre, à savoir que la nouvelle position de l'octaèdre s'obtienne à partir de la position initiale au moyen d'une rotation réalisée par une certaine matrice U. Si V est une rotation quelconque qui fait repasser l'octaèdre précédent à sa position initiale, il est évident que UVU^{-1} donne une rotation au cours de laquelle le nouvel octaèdre se transforme en lui-même et inversement. Ainsi, si le groupe des rotations de l'octaèdre précédent est constitué des matrices V_k $(k = 1, 2, \ldots, 24)$, le groupe des rotations du nouvel octaèdre est simplement constitué des matrices semblables UV_kU^{-1}. Autrement dit, on obtient un groupe associé. En général, si l'ensemble de certaines matrices V_k forme un groupe, l'ensemble des matrices semblables UV_kU^{-1}, quelle que soit la matrice donnée U, forme également un groupe. Il est facile de le démontrer directement à partir de la définition des groupes. On dit que le second groupe est associé au premier.

Soit un tétraèdre. Sa surface est formée de quatre triangles équilatéraux et il présente quatre sommets. Un axe quelconque joint un des sommets A du tétraèdre au centre de la face opposée. Le tétraèdre coïncide avec lui-même si on le fait tourner autour de cet axe d'un angle $0, \frac{2\pi}{3}, \frac{4\pi}{3}$. Soit S_0, S_1, S_2 ces rotations. On introduit ensuite trois transformations linéaires T_1, T_2, T_3 pour lesquelles le tétraèdre coïncide avec lui-même et son sommet A avec l'un des trois autres sommets. Outre les rotations S_0, S_1 et S_2 on peut former les 9 rotations T_kS_0, T_kS_1, T_kS_2 $(k = 1, 2, 3)$. Les 12 rotations obtenues sont différentes et transforment le tétraèdre en lui-même.

Soit maintenant un icosaèdre, sa surface est formée de 20 triangles équilatéraux et il présente 12 sommets. Soit comme ci-dessus un axe quelconque qui joint un des sommets A de l'icosaèdre avec le sommet opposé. L'icosaèdre coïncide avec lui-même s'il fait une rotation autour de cet axe d'un angle égal à $\frac{2k\pi}{5}$ $(k = 0, 1, 2, 3, 4)$. Soient S_k ces rotations. On a de plus 11 rotations T_l $(l = 1, 2, \ldots, 11)$ qui amènent l'icosaèdre à coïncider avec lui-même, le sommet A se transformant en l'un des autres sommets. Le groupe complet des rotations qui transforment l'icosaèdre en lui-même est formé des 5 rotations S_k et des 55 rotations T_lS_k. Au total 60 rotations. Il en sera de même pour le groupe du dodécaèdre dont la surface est formée de 12 pentagones réguliers et qui présente 20 sommets. Pour le vérifier il suffit de disposer le dodécaèdre par rapport à l'icosaèdre comme on l'a fait ci-dessus pour le cube par rapport à l'octaèdre.

Considérons encore un groupe formé de rotations de l'espace à trois dimensions. Soit dans le plan XY un polygone régulier à n côtés dont le centre coïncide avec l'origine des coordonnées. Soit un axe quelconque qui joint le sommet A du polygone au côté opposé (si n est pair) ou avec le milieu du côté opposé (si n est impair). Si on fait tourner le plan XY autour de cet axe d'un angle nul ou π, le polygone coïncide avec lui-même. La première rotation est la transformation identique I et on désigne la seconde par S.

De plus on a les rotations T_k autour de l'axe Z d'un angle $\frac{2k\pi}{n}$ $(k = 1, 2, \ldots, n-1)$ pour lesquelles le polygone coïncide également avec lui-même, et son sommet A se transforme en l'un des autres sommets. Pour $k = 0$, on obtient la transformation identique $T_0 = I$. Le groupe complet des transformations qui transforment un polygone à n côtés en lui-même est formé de $2n$ transformations: T_k et T_kS $(k = 0, 1, 2, \ldots, n-1)$.

Le polygone à n côtés que nous venons de considérer, dont l'aire est comptée deux fois (aires intérieure et extérieure), est appelé dièdre et le groupe qu'il engendre groupe diédral.

III-1-3. Transformations de Lorentz. Tous les exemples de groupes des transformations linéaires que l'on a vus jusqu'ici étaient formés soit de transformations unitaires, soit de rotations de l'espace à trois dimensions (cas particulier de transformations unitaires). Etudions maintenant un autre groupe des transformations linéaires dont les éléments ne sont pas des transformations unitaires. Ce groupe joue un rôle important en relativité, en électrodynamique et dans la partie de la mécanique quantique liée à la relativité.

Considérons quatre variables x_1, x_2, x_3, x_4, dont les trois premières sont les coordonnées spatiales d'un point et la dernière le temps. En vertu de la condition fondamentale du principe de la relativité restreinte selon laquelle une certaine vitesse c (vitesse de la lumière) est invariante dans le cas d'un mouvement relatif, on cherche les transformations linéaires des quatre variables indiquées ci-dessus pour lesquelles l'expression

$$x_1^2 + x_2^2 + x_3^2 - c^2x_4^2$$

reste inchangée, c'est-à-dire qu'on cherche les transformations linéaires qui expriment les nouvelles variables x'_k en fonction des anciennes x_k et telles que l'on ait l'identité:

$$x_1'^2 + x_2'^2 + x_3'^2 - c^2x_4'^2 = x_1^2 + x_2^2 + x_3^2 - c^2x_4^2.$$

On considère d'abord le cas où les coordonnées x_2 et x_3 restent inchangées et seules les variables x_1 et x_4 figurent dans la transformation linéaire. On doit ainsi trouver des transformations linéaires

$$x'_1 = a_{11}x_1 + a_{14}x_4, \quad x'_4 = a_{41}x_1 + a_{44}x_4, \tag{11}$$

telles que :

$$x_1'^2 - c^2 x_4'^2 = x_1^2 - c^2 x_4^2. \tag{12}$$

Au lieu de x_4 on introduit une nouvelle variable imaginaire pure y_1 telle que :

$$y_1 = icx_4.$$

Les transformations linéaires cherchées doivent être de la forme :

$$x_1' = \alpha_{11}x_1 + \alpha_{12}y_1, \quad y_1' = \alpha_{21}x_1 + \alpha_{22}y_1, \tag{13}$$

où :

$$\alpha_{11} = a_{11}; \quad \alpha_{12} = \frac{a_{14}}{ic}; \quad \alpha_{21} = ica_{41}; \quad \alpha_{22} = a_{44},$$

et la condition (12) s'écrit alors :

$$x_1'^2 + y_1'^2 = x_1^2 + y_1^2. \tag{14}$$

Les coefficients α_{11} et α_{22} doivent être réels tandis que α_{12} et α_{21} doivent être imaginaires purs. C'est pourquoi on pose $\alpha_{12} = i\beta_{12}$ et $\alpha_{21} = i\beta_{21}$. La condition (14) est équivalente à la condition d'orthogonalité de la transformation (13) et, par suite, la somme des carrés des éléments de chacune des lignes et des colonnes doit être égale à l'unité. Ce qui donne $\beta_{12}^2 = \beta_{21}^2 = \alpha_{11}^2 - 1 = \alpha_{22}^2 - 1$ et $\alpha_{11}^2 = \alpha_{22}^2$. Posons $\alpha_{22} = \alpha$ et $\beta_{12} = \alpha\beta$. Supposons les coefficients α_{11} et α_{22} positifs, ce qui correspond à des directions de x_1 et x_4 inchangées. On obtient ainsi au lieu de (13), compte tenu des relations précédentes :

$$x_1' = \alpha x_1 + i\alpha\beta y_1, \quad y_1' = \alpha_{21}x_1 + \alpha y_1.$$

La condition d'orthogonalité des lignes

$$\alpha\alpha_{21} + i\alpha^2\beta = 0$$

donne $\alpha_{21} = -i\alpha\beta$, c'est-à-dire que β_{12} et β_{21} doivent être de signe opposé. Enfin, la condition

$$\alpha_{11}^2 + \alpha_{12}^2 = 1$$

donne :

$$\alpha^2 - \alpha^2\beta^2 = 1, \text{ soit } \alpha = \frac{1}{\sqrt{1-\beta^2}} \quad (\beta^2 < 1),$$

et finalement on obtient les formules suivantes :

$$x_1' = \frac{x_1 + i\beta y_1}{\sqrt{1-\beta^2}}; \qquad y_1' = \frac{-i\beta x_1 + y_1}{\sqrt{1-\beta^2}},$$

ou, en repassant de la variable $y_1 = icx_4$ à la variable x_4, on a :

$$x_1' = \frac{x_1 - \beta c x_4}{\sqrt{1-\beta^2}}; \qquad x_4' = \frac{-\frac{\beta}{c} x_1 + x_4}{\sqrt{1-\beta^2}}. \tag{15}$$

Il résulte de ces égalités que le système de coordonnées qui correspond aux variables à indice ′ se déplace par rapport au système de coordonnées initiales à la vitesse :

$$v = \beta c, \tag{16}$$

dans la *direction* de l'axe x_1. *En effet, si l'on considère* x_1' *constant, on obtient :*

$$dx_1 - \beta c\, dx_4 = 0, \text{ c'est-à-dire } \frac{dx_1}{dx_4} = \beta c.$$

En remplaçant β par v (formule 16), x_1 par x et x par t, on obtient la formule habituelle de la *transformation de Lorentz* à deux variables :

$$x' = \frac{x - vt}{\sqrt{1-\frac{v^2}{c^2}}}; \qquad t' = \frac{-\frac{v}{c^2} x + t}{\sqrt{1-\frac{v^2}{c^2}}}. \tag{17}$$

Dans le cas limite où $c \to \infty$ on obtient les formules habituelles du mouvement relatif de la mécanique classique :

$$x' = x - vt; \qquad t' = t.$$

Il est facile de vérifier que les transformations de Lorentz (17) qui dépendent d'un paramètre réel v forment un groupe. En résolvant les équations (17) par rapport à x et t, on obtient la transformation inverse de (17). Montrons que c'est aussi une transformation de Lorentz qui s'obtient à partir de la transformation (17) en remplaçant v par $(-v)$. Résolvant en effet les équations (17), on a :

$$\left(1 - \frac{v^2}{c^2}\right) x = \sqrt{1 - \frac{v^2}{c^2}} (x' + vt');$$

$$\left(1 - \frac{v^2}{c^2}\right) t = \sqrt{1 - \frac{v^2}{c^2}} \left(\frac{v}{c^2} x' + t'\right),$$

d'où il résulte que :

$$x = \frac{x' + vt'}{\sqrt{1-\frac{v^2}{c^2}}}; \qquad t = \frac{\frac{v}{c^2} x' + t'}{\sqrt{1-\frac{v^2}{c^2}}}.$$

Soient alors deux transformations de Lorentz L_1 et L_2 qui correspondent aux valeurs du paramètre $v = v_1$ et $v = v_2$. Formons leur produit $L_2 L_1$ et montrons que c'est aussi une transformation de Lorentz. Il faut pour cela former le produit des deux matrices:

$$\begin{Vmatrix} \dfrac{1}{\sqrt{1-\beta_2^2}} & -\dfrac{\beta_2 c}{\sqrt{1-\beta_2^2}} \\ -\dfrac{\frac{\beta_2}{c}}{\sqrt{1-\beta_2^2}} & \dfrac{1}{\sqrt{1-\beta_2^2}} \end{Vmatrix} \cdot \begin{Vmatrix} \dfrac{1}{\sqrt{1-\beta_1^2}} & -\dfrac{\beta_1 c}{\sqrt{1-\beta_1^2}} \\ -\dfrac{\frac{\beta_1}{c}}{\sqrt{1-\beta_1^2}} & \dfrac{1}{\sqrt{1-\beta_1^2}} \end{Vmatrix},$$

où :

$$\beta_1 = \frac{v_1}{c}; \qquad \beta_2 = \frac{v_2}{c}.$$

En appliquant les règles habituelles de multiplication des matrices, on obtient pour le produit la matrice suivante:

$$\frac{1+\beta_1\beta_2}{\sqrt{1-\beta_2^2}\sqrt{1-\beta_1^2}} \begin{Vmatrix} 1 & -\dfrac{\beta_1 c+\beta_2 c}{1+\beta_1\beta_2} \\ -\dfrac{\frac{\beta_1}{c}+\frac{\beta_2}{c}}{1+\beta_1\beta_2} & 1 \end{Vmatrix}. \tag{18}$$

Introduisons alors la nouvelle quantité:

$$v_3 = \frac{v_1+v_2}{1+\frac{v_1 v_2}{c^2}}. \tag{19}$$

Il est facile de vérifier que l'on a l'identité:

$$\frac{1+\frac{v_1 v_2}{c^2}}{\sqrt{1-\frac{v_2^2}{c^2}}\sqrt{1-\frac{v_1^2}{c^2}}} = \frac{1}{\sqrt{1-\frac{v_3^2}{c^2}}},$$

et la matrice (18) peut être mise sous la forme suivante:

$$\begin{Vmatrix} \dfrac{1}{\sqrt{1-\beta_3^2}} & -\dfrac{\beta_3 c}{\sqrt{1-\beta_3^2}} \\ -\dfrac{\frac{\beta_3}{c}}{\sqrt{1-\beta_3^2}} & \dfrac{1}{\sqrt{1-\beta_3^2}} \end{Vmatrix} \qquad \left(\beta_3 = \frac{v_3}{c}\right),$$

c'est-à-dire qu'il lui correspond aussi une transformation de Lorentz de paramètre $v = v_3$. Ainsi, la formule (19) donne la *règle d'addition des vitesses dans la théorie de la relativité restreinte*. Si dans la formule (19) on pose $v_1 = c$, il est facile de vérifier que la vitesse résultante

v_3 est telle que $v_3 = c$, c'est-à-dire que *la vitesse c ne varie pas si l'on compose deux mouvements.*

Pour obtenir les formules (15) on a fixé de façon déterminée les signes des coefficients de la transformation linéaire (11) en supposant les coefficients a_{11} et a_{44} positifs. On peut remplacer cette condition par une autre, à savoir que le coefficient a_{44} est positif ainsi que le déterminant :

$$a_{11}a_{44} - a_{14}a_{41}. \tag{20}$$

Il est facile de voir qu'il en résulte que le coefficient a_{11} est positif et inversement. En effet, le déterminant de la transformation (17) est égal à $(+1)$, c'est-à-dire que si $a_{11} > 0$ le déterminant (20) est positif. Si l'on prend $a_{11} = -\alpha$ et $a_{44} = \alpha$, où $\alpha > 0$, on obtient une transformation dont le déterminant est égal à (-1). Pour que le coefficient a_{11} soit positif, il faut que pour x_1 constant et $x_4 \to \infty$ on ait $x'_4 \to \infty$. On peut dire que ceci correspond au fait que le sens du temps n'est pas changé. Ainsi, les formules ne donnent pas toutes les transformations linéaires qui vérifient la condition (12), mais seulement celles pour lesquelles le déterminant (20) est positif et qui ne changent pas le sens du temps.

Considérons maintenant la *transformation de Lorentz générale* dans le cas de quatre variables x_k $(k = 1, 2, 3, 4)$ vérifiant la condition :

$$x_1'^2 + x_2'^2 + x_3'^2 - c^2x_4'^2 = x_1^2 + x_2^2 + x_3^2 - c^2x_4^2. \tag{21}$$

On considère x_k $(k = 1, 2, 3)$ et x'_k $(k = 1, 2, 3)$ comme des coordonnées cartésiennes de deux espaces à trois dimensions différents R et R'. Montrons qu'en choisissant de façon adéquate les axes de coordonnées dans ces deux espaces on peut réduire la transformation de Lorentz générale au cas particulier examiné ci-dessus. Soit T la transformation de Lorentz générale et S la transformation particulière examinée ci-dessus. La proposition équivaut à mettre T sous la forme :

$$T = VSU, \tag{22}$$

où U et V sont des transformations réelles orthogonales qui correspondent aux transformations de coordonnées dans les espaces R et R'.

On introduit, comme ci-dessus, quatre nouvelles variables :

$$y_1 = x_1\,;\ \ y_2 = x_2\,;\ \ y_3 = x_3\,;\ \ y_4 = icx_4,$$

et de façon analogue :

$$y'_1 = x'_1\,;\ \ y'_2 = x'_2\,;\ \ y'_3 = x'_3\,;\ \ y'_4 = icx'_4.$$

Au lieu de la condition (21) on obtient pour les nouvelles variables la condition d'orthogonalité habituelle :

$$y_1'^2 + y_2'^2 + y_3'^2 + y_4'^2 = y_1^2 + y_2^2 + y_3^2 + y_4^2. \tag{23}$$

La transformation linéaire cherchée sera de la forme:

$$y'_k = \alpha_{k1}y_1 + \alpha_{k2}y_2 + \alpha_{k3}y_3 + \alpha_{k4}y_4 \qquad (k = 1, 2, 3, 4). \tag{24}$$

Tenant compte de ce que y_4 et y'_4 doivent être imaginaires purs on peut affirmer que les coefficients α_{k1}, α_{k2}, α_{k3} pour $k = 1, 2, 3$ ainsi que le coefficient α_{44} doivent être réels, tandis que les coefficients α_{41}, α_{42}, α_{43}, α_{k4} pour $k = 1, 2, 3$ doivent être imaginaires purs. Le changement des axes de coordonnées dans l'espace R' équivaut à la transformation réelle orthogonale des variables y'_1, y'_2, y'_3. Soient les coefficients:

$$\alpha_{14} = i\beta_{14}; \quad \alpha_{24} = i\beta_{24}; \quad \alpha_{34} = i\beta_{34}.$$

Trois nombres réels β_{14}, β_{24}, β_{34} définissent un vecteur et si on prend la direction de ce vecteur comme nouveau premier axe de l'espace R', la transformation orthogonale correspondante fait que les coefficients α_{24} et α_{34} sont nuls. Pour s'en convaincre il suffit de remarquer que, compte tenu des formules (24), la transformation orthogonale des variables y'_1, y'_2, y'_3 se réduit à la même transformation pour β_{14}, β_{24}, β_{34}. Ainsi, nous supposons que cette transformation des coordonnées dans l'espace R' soit déjà effectuée de sorte que l'on a $\alpha_{24} = \alpha_{34} = 0$. La condition (23) montre que les coefficients de la transformation (24) doivent vérifier les conditions habituelles des transformations orthogonales. En tenant compte de ce que les coefficients indiqués ci-dessus sont nuls, on obtient, en considérant la deuxième et la troisième ligne, les conditions suivantes:

$$\alpha_{k1}^2 + \alpha_{k2}^2 + \alpha_{k3}^2 = 1 \qquad (k = 2, 3),$$

$$\alpha_{21}\alpha_{31} + \alpha_{22}\alpha_{32} + \alpha_{23}\alpha_{33} = 0,$$

où tous les coefficients sont réels. Compte tenu des conditions écrites, les deux vecteurs de composantes $(\alpha_{21}, \alpha_{22}, \alpha_{23})$ et $(\alpha_{31}, \alpha_{32}, \alpha_{33})$ auront une longueur unité et seront orthogonaux. Si dans l'espace R on choisit ces deux vecteurs comme vecteurs de base, dirigés suivant les axes X_2 et X_3, alors les deux sommes:

$$\alpha_{k1}y_1 + \alpha_{k2}y_2 + \alpha_{k3}y_3 \quad (k = 2, 3),$$

exprimant le produit scalaire des deux vecteurs considérés ci-dessus par le vecteur variable (y_1, y_2, y_3), s'expriment simplement en fonction de y_2 et y_3, c'est-à-dire que pour un tel choix des axes de coordonnées on a:

$$\alpha_{22} = \alpha_{33} = 1; \quad \alpha_{21} = \alpha_{23} = \alpha_{31} = \alpha_{32} = 0.$$

Ainsi, avec le choix d'axes qui a été fait dans les deux espaces on obtient finalement la matrice de la transformation (24) de la forme :

$$\begin{Vmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} & \alpha_{14} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \alpha_{41} & \alpha_{42} & \alpha_{43} & \alpha_{44} \end{Vmatrix}. \tag{25}$$

Cette matrice s'obtient en multipliant la matrice initiale par deux transformations orthogonales ne se rapportant qu'aux trois premières variables mais que l'on peut bien sûr considérer comme des transformations orthogonales à quatre variables, la quatrième variable restant constante. En tenant compte du fait que le produit de deux transformations orthogonales doit être également orthogonal, on peut affirmer que les éléments de la matrice (25) doivent aussi vérifier la condition d'orthogonalité. En écrivant cette condition pour la première ligne avec la deuxième et la troisième on obtient :

$$\alpha_{12} = \alpha_{13} = 0.$$

De même la condition d'orthogonalité pour la quatrième ligne avec la deuxième et la troisième donne :

$$\alpha_{43} = \alpha_{42} = 0.$$

En fin de compte on arrive à la matrice suivante :

$$\begin{Vmatrix} \alpha_{11} & 0 & 0 & \alpha_{14} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \alpha_{41} & 0 & 0 & \alpha_{44} \end{Vmatrix},$$

c'est-à-dire que dans ce cas on a une transformation linéaire :

$$y_1' = \alpha_{11}y_1 + \alpha_{14}y_4,$$
$$y_4' = \alpha_{41}y_1 + \alpha_{44}y_4,$$

qui doit vérifier la condition :

$$y_1'^2 + y_4'^2 = y_1^2 + y_4^2.$$

C'est justement d'une transformation de ce type qu'il a été question ci-dessus et c'est elle qui nous a conduit à la transformation de Lorentz spéciale du type (15). Ainsi, la formule (22) peut être considérée comme vérifiée. On notera que le choix du signe lors de la détermination de la transformation S sera soumis à la même règle que ci-dessus, si on impose que la transformation de Lorentz générale T ne change pas le sens du temps et qu'elle ait un déterminant positif. Les transformations orthogonales U et V peuvent toujours être considérées comme des rotations d'un espace à trois dimensions, de sorte que leur déterminant est positif et qu'elles ne touchent pas la quatrième variable. Il faut donc que la transformation S ait un déterminant positif et que cette transformation ne change pas le sens du temps, c'est-à-dire qu'avec l'hypothèse faite à propos de la transformation générale T on obtienne, pour la transformation particulière, la condition qui a servi d'hypothèse au calcul des formules obtenues. Les transformations de Lorentz générales qui vérifient les deux conditions imposées ci-dessus s'appellent d'ordinaire *transformations de Lorentz positives*. Des raisonnements précédents il résulte que les matrices qui leur correspondent s'obtiennent à partir de la formule (22), où S est la transformation de Lorentz spéciale du type (15) et U et V des matrices de rotation d'un espace à trois dimensions. On peut montrer que les transformations de Lorentz positives forment un groupe de même que les transformations (15).

Les raisonnements précédents montrent que la matrice de la transformation de Lorentz la plus générale définie simplement par la condition (21) peut être représentée par la formule (22), où U et V sont des rotations et S une transformation de Lorentz générale à deux variables. Si c'est une transformation positive, il résulte des formules (15) que $D(S) = 1$ et le déterminant de toute transformation de Lorentz positive est, lui aussi, égal à l'unité, puisque les déterminants de U et V sont égaux à l'unité et que, de plus, les matrices U, S et V doivent être considérées comme des matrices du quatrième ordre. Dans le cas général de transformation de Lorentz du deuxième ordre, il est facile de montrer que le déterminant peut être égal à (± 1) et, par suite, la transformation de Lorentz générale a également un déterminant égal à (± 1).

III-1-4. Permutations. Jusqu'à présent il a été question d'exemples de groupes dont les éléments étaient des transformations linéaires. La notion de groupe n'est pas liée forcément aux transformations linéaires et on peut construire des groupes pour des opérations d'autres espèces. Passons maintenant à l'examen d'opérations dont il a déjà été question [I-1-2]: les *permutations*. Donnons d'abord quelques notions et faits fondamentaux concernant les permutations.

Soient n objets que l'on numérote comme en [I-1-2], c'est-à-dire qu'on considère simplement comme des nombres entiers $1, 2, \ldots, n$. A l'aide de ces mêmes nombres on peut former $n!$ permutations. Soit l'une quelconque d'entre elles:

$$p_1, p_2, \ldots, p_n. \tag{26}$$

Les nombres p_k dans leur ensemble donnent tous les nombres entiers de 1 à n et dans la permutation (26) ils sont disposés dans un ordre défini. Comparons la permutation (26) avec la permutation fondamentale 1, 2, ..., n:

$$\begin{pmatrix} 1 & 2 & \dots & n \\ p_1 & p_2 & \dots & p_n \end{pmatrix} (P). \tag{27}$$

Le passage de la permutation fondamentale à la permutation (26) s'effectue en remplaçant 1 par p_1, 2 par p_2 et ainsi de suite. Désignons cette opération par la lettre P qu'on appellera dans la suite permutation. Définissons maintenant la notion de *permutation inverse* P^{-1}. Ce sera une opération qui transforme (26) en la suite fondamentale, c'est-à-dire une opération qui remplace p_1 par 1, p_2 par 2 et ainsi de suite. Expliquons ceci sur un exemple particulier. Prenons $n = 5$, et considérons la permutation:

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 2 & 5 & 1 & 4 \end{pmatrix} (P).$$

La permutation inverse sera alors de la forme:

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 2 & 1 & 5 & 3 \end{pmatrix} (P^{-1}).$$

Il est facile de voir que

$$(P^{-1})^{-1} = P. \tag{28}$$

Introduisons maintenant la notion de *produit de permutations*. Soient P_1 et P_2 deux permutations quelconques. Appelons produit des permutations P_2P_1 la permutation obtenue par application d'abord de la permutation P_1 puis de la permutation P_2. Par exemple, si l'on a les deux permutations:

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 1 & 4 & 3 & 2 \end{pmatrix} (P_2) \text{ et } \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 1 & 5 & 2 & 4 \end{pmatrix} (P_1),$$

leur produit P_2P_1 sera la permutation:

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 5 & 2 & 1 & 3 \end{pmatrix} (P_2P_1).$$

Il est alors évident que la permutation inverse P^{-1} est définie par la condition:

$$P^{-1}P = PP^{-1} = I, \tag{29}$$

où I désigne la permutation unité pour laquelle chaque élément est remplacé par lui-même.

Par suite, en appliquant plusieurs fois des permutations, on peut former un produit de plusieurs permutations $P_3P_2P_1$. Il est facile de voir qu'un tel produit est associatif, c'est-à-dire :

$$P_3(P_2P_1) = (P_3P_2)P_1. \tag{30}$$

En effet, en effectuant la permutation P_1 on peut ensuite successivement effectuer P_2 et P_3 ou, au lieu de cela, on peut remplacer cette application successive de P_2 et P_3 par l'application de la permutation (P_3P_2) qui est équivalente à l'application successive de P_2 et P_3. On notera enfin que la permutation unité vérifie de façon évidente la condition :

$$IP = PI = P, \tag{31}$$

où P est une permutation quelconque. Le produit de permutations ne sera pas en général commutatif, c'est-à-dire que les produits P_2P_1 et P_1P_2 seront des permutations différentes. Nous proposons de le vérifier sur l'exemple précédent.

Les notions fondamentales de produit, de permutation inverse et de permutation unité ont été ainsi introduites exactement de la même manière que pour les transformations linéaires (matrices).

On peut maintenant étendre cette analogie et établir la notion de groupe. L'ensemble des permutations forme un groupe si les deux conditions suivantes sont vérifiées : 1) la permutation inverse d'une permutation appartenant à l'ensemble appartient aussi à l'ensemble et 2) le produit de deux permutations appartenant à l'ensemble (quel que soit l'ordre des facteurs) appartient aussi à l'ensemble. Comme dans le cas des transformations linéaires, la permutation unité doit forcément appartenir au groupe.

L'ensemble des $n!$ permutations forme, de façon évidente, un groupe. Passons maintenant à l'établissement d'un autre groupe qui n'est formé que d'une partie du groupe précédent. On notera pour cela que toute permutation peut être réalisée au moyen de certaines transpositions [I-1-2]. Pour une permutation donnée, le nombre de transpositions peut être différent, mais, comme on l'a déjà démontré, il sera toujours pour une permutation donnée soit pair, soit impair. Les permutations constituées d'un nombre pair de transpositions forment par elles-mêmes un groupe. Le groupe formé par toutes les permutations est appelé *groupe symétrique* et le groupe constitué des permutations paires, c'est-à-dire des permutations se ramenant à un nombre pair de transpositions, groupe *alterné*.

Considérons maintenant des permutations d'un type particulier. Soient $l_1, l_2, \ldots, l_m$ m nombres différents choisis parmi les n premiers nombres. Supposons une permutation formée en remplaçant l_1 par l_2, l_2 par l_3 et ainsi de suite, l_{m-1} par l_m et enfin l_m par l_1. Une permutation de ce type ou *cycle* est désignée par le symbole $(l_1, l_2, \ldots, l_m)$. En effectuant une permutation circulaire des nombres à l'intérieur des parenthèses on obtient ainsi les cycles:

$$(l_2, l_3, \ldots, l_m, l_1), \qquad (l_3, l_4, \ldots, l_m, l_1, l_2), \ldots,$$

qui représentent des permutations identiques à $(l_1, l_2, \ldots, l_m)$. Si $m = 1$, on a le cycle (l_1) équivalent à la permutation unité et cela n'a pas de sens de l'examiner. Le cycle formé de deux nombres (l_1, l_2) est équivalent à la transposition des éléments l_1 et l_2.

Si on a deux cycles sans éléments communs, leur produit ne dépend pas de l'ordre des facteurs.

Soit par exemple $n = 5$. Considérons le produit des deux cycles sans éléments communs $(1, 3)\,(2, 4, 5)$ et $(2, 4, 5)\,(1, 3)$.

Ces deux produits donnent évidemment une seule et même permutation:

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 4 & 1 & 5 & 2 \end{pmatrix}.$$

On peut représenter toute permutation P sous la forme d'un *produit de cycles* sans éléments communs. Pour le montrer considérons l'élément 1 que l'on prend pour premier élément du cycle. Pour le deuxième élément on prend celui qui est obtenu à partir de 1 au moyen de P. Soit l_2 cet élément. Pour troisième élément on prend celui qui est obtenu à partir de l_2 au moyen de P et ainsi de suite, jusqu'à ce qu'on arrive enfin à un élément qui se transforme en 1 au moyen de P. Ce sera le dernier élément du cycle formé. Il est facile de voir que ce cycle ne peut pas contenir d'éléments identiques. Le cycle ainsi engendré n'épuise pas, en général, tous les n éléments. Parmi les éléments restants on prend l'un quelconque comme premier élément d'un nouveau cycle et on engendre un second cycle comme ci-dessus, etc.

Comme exemple prenons la permutation pour $n = 6$:

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 6 & 4 & 1 & 2 & 5 \end{pmatrix}.$$

En appliquant la méthode précédente on peut la représenter sous la forme d'un produit de cycles:

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 6 & 4 & 1 & 2 & 5 \end{pmatrix} = (1, 3, 4)\,(2, 6, 5).$$

L'ordre des facteurs du second membre n'a pas d'importance.

Il est facile de voir que le produit de deux transpositions peut être représenté sous la forme d'un produit de cycles à trois termes. Si des cycles à deux termes n'ont pas d'élément commun, on peut facilement vérifier que

$$(l_3, l_4)\,(l_1, l_2) = (l_1, l_3, l_4)\,(l_1, l_2, l_4).$$

En présence d'éléments communs on a

$$(l_1, l_3)(l_1, l_2) = (l_1, l_2, l_3).$$

Ainsi, toute permutation d'un groupe alterné peut être représentée sous la forme d'un produit de cycles à trois termes.

On notera encore que lors d'une permutation, on peut écrire dans la première ligne les nombres non obligatoirement dans leur suite naturelle mais dans n'importe quel ordre. Il est important que sous chaque nombre figure le nombre en lequel il est transformé par la permutation considérée. Donnons deux exemples d'une même permutation:

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 2 & 5 & 1 & 4 \end{pmatrix} = \begin{pmatrix} 3 & 1 & 5 & 4 & 2 \\ 5 & 3 & 4 & 1 & 2 \end{pmatrix}.$$

Soit une permutation quelconque:

$$P = \begin{pmatrix} a_1 & a_2 & \dots & a_n \\ b_1 & b_2 & \dots & b_n \end{pmatrix}.$$

La permutation inverse peut être mise sous la forme

$$P^{-1} = \begin{pmatrix} b_1 & b_2 & \dots & b_n \\ a_1 & a_2 & \dots & a_n \end{pmatrix}.$$

Soient deux permutations et écrivons la seconde deux fois:

$$P = \begin{pmatrix} 1 & 2 & \dots & n \\ c_1 & c_2 & \dots & c_n \end{pmatrix}; \qquad Q = \begin{pmatrix} 1 & 2 & \dots & n \\ d_1 & d_2 & \dots & d_n \end{pmatrix} = \begin{pmatrix} c_1 & c_2 & \dots & c_n \\ f_1 & f_2 & \dots & f_n \end{pmatrix}.$$

On a:

$$PQ^{-1} = \begin{pmatrix} 1 & 2 & \dots & n \\ c_1 & c_2 & \dots & c_n \end{pmatrix} \begin{pmatrix} d_1 & d_2 & \dots & d_n \\ 1 & 2 & \dots & n \end{pmatrix} = \begin{pmatrix} d_1 & d_2 & \dots & d_n \\ c_1 & c_2 & \dots & c_n \end{pmatrix},$$

et par suite:

$$QPQ^{-1} = \begin{pmatrix} c_1 & c_2 & \dots & c_n \\ f_1 & f_2 & \dots & f_n \end{pmatrix} \begin{pmatrix} d_1 & d_2 & \dots & d_n \\ c_1 & c_2 & \dots & c_n \end{pmatrix} = \begin{pmatrix} d_1 & d_2 & \dots & d_n \\ f_1 & f_2 & \dots & f_n \end{pmatrix}.$$

De ce qui a été écrit résulte la règle suivante: pour obtenir la permutation QPQ^{-1}, il faut dans les deux lignes de la permutation

$$P = \begin{pmatrix} 1 & 2 & \dots & n \\ c_1 & c_2 & \dots & c_n \end{pmatrix}$$

effectuer la permutation Q.

III-1-5. Groupes abstraits. Dans la définition des groupes nous pouvons ne pas tenir compte de la signification concrète des opérations qui agissent sur l'ensemble constituant le groupe et qui dans les paragraphes précédents étaient des transformations linéaires ou des permutations. On arrive ainsi à la notion de *groupe abstrait*.

Un groupe abstrait est un ensemble de certains symboles, tels que pour ces symboles on ait défini la multiplication, en ce sens que l'on se donne une règle définie grâce à laquelle à partir de deux éléments P et Q (identiques ou non) de l'ensemble on obtient un troisième élément qui appartient aussi à l'ensemble, que l'on appelle produit de ces éléments et que l'on désigne pas QP. En outre trois conditions doivent être vérifiées:

1. *La multiplication doit être associative*, c'est-à-dire que: $(RQ)\,P = R\,(QP)$, d'où il résulte, en général, que dans tout produit on peut, sans changer l'ordre des facteurs, réunir deux facteurs quelconques en un même groupe.

2. *Il doit exister dans l'ensemble un élément E et un seul tel que, multiplié à droite ou à gauche par n'importe quel autre élément du groupe, on obtienne le même élément*, c'est-à-dire:

$$EP = PE = P. \tag{32}$$

On appellera cet élément *élément unité*.

3. *A tout élément P de l'ensemble correspond dans ce même ensemble un élément Q et un seul tel que l'on ait*

$$QP = PQ = E \qquad (Q = P^{-1}). \tag{33}$$

Il résulte de (32) que si $P = E$, on a $EE = E$, c'est-à-dire que, d'après la définition de l'élément inverse, l'inverse de E sera l'élément E lui-même ($E^{-1} = E$).

On peut imposer des conditions qui définissent un groupe abstrait sous une forme plus restreinte et dont les autres conditions résultent comme conséquences nécessaires formelles, mais nous ne le ferons pas ici. Nous nous limiterons aux faits les plus simples et les plus fondamentaux liés à la notion de groupe abstrait. Un examen plus détaillé de la théorie des groupes fournirait, par lui-même, matière à un livre entier. Notre but est simplement de donner au lecteur les notions fondamentales et par là de faciliter son étude des livres de physique où la notion de groupe et les propriétés fondamentales des groupes sont souvent utilisées. Dans la suite, au lieu de E on écrira parfois I. L'élément Q défini par les relations (33) est appelé l'*inverse* de P et on le désigne par P^{-1}. Il est évident que l'on a la relation (28), car il résulte de (33) que P est l'inverse de Q.

Ayant établi la notion de groupe abstrait, donnons maintenant l'explication de certaines notions nouvelles et les démonstrations de certaines propriétés des groupes abstraits. On notera d'abord que le nombre d'éléments dans un groupe, comme on l'a déjà vu, peut être aussi bien fini qu'infini. Considérons le produit d'éléments du groupe

$$RQP.$$

C'est aussi un élément du groupe. L'élément inverse s'obtient exactement comme dans le cas du groupe des transformations linéaires, c'est-à-dire que l'on a

$$(RQP)^{-1} = P^{-1}Q^{-1}R^{-1}.$$

Il est facile de le vérifier en effectuant la multiplication et en utilisant l'associativité. Soit P un élément quelconque du groupe. Ses puissances entières positives

$$P^0 = I,\ P^1,\ P^2,\ \ldots$$

sont aussi des éléments du groupe. S'il existe un entier positif m tel que $P^m = I$, on dit que cet élément est *d'ordre fini*, l'*ordre d'un élément* étant la plus petite valeur de l'entier positif m pour lequel $P^m = I$. Alors les éléments

$$I,\ P,\ P^2,\ \ldots,\ P^{m-1}$$

sont tous distincts. En effet, de la condition $P^k = P^l$ $(k < l)$ il découle directement que $P^{l-k} = I$. Dans un groupe fini tous les éléments sont évidemment d'ordre fini.

Désignons les éléments du groupe par P_α. Si le groupe est fini, on peut considérer que l'indice α parcourt un ensemble fini d'entiers positifs. Si le groupe est infini, il peut parcourir toutes les valeurs entières [III-1-1], il peut varier continûment et même il peut être équivalent à plusieurs indices qui varient continûment. Soit U un élément quelconque donné du groupe. Formons tous les produits possibles UP_α. Il est aisé de voir que lorsque l'indice α varie d'une façon adéquate, le produit précédent donne à nouveau tous les éléments du groupe et ceci une seule fois.

En effet, si on multiplie à gauche l'égalité

$$UP_{\alpha_1} = UP_{\alpha_2}$$

par U^{-1}, on obtient $P_{\alpha_1} = P_{\alpha_2}$, c'est-à-dire que pour des valeurs différentes de α le produit UP_α est différent. Montrons maintenant que ce produit se transforme en n'importe quel élément du groupe. L'égalité $UP_\alpha = P_{\alpha_0}$ en effet est équivalente à $P_\alpha = U^{-1}P_{\alpha_0}$, c'est-à-dire que le produit UP_α donne l'élément P_{α_0} lorsque le facteur P_α est égal à l'élément $U^{-1}P_{\alpha_0}$ du groupe. On obtiendrait le même résultat en multipliant l'élément donné par U non à gauche, mais à droite. Ainsi, on est conduit au résultat suivant : *si P_α parcourt tous* *les éléments du groupe et si U est un élément quelconque donné du groupe, le produit UP_α (ou $P_\alpha U$) parcourt aussi tous les éléments du groupe et ceci une seule fois.*

Comme un cas particulier de groupe, considérons un groupe constitué de six éléments (groupe du sixième ordre) et désignons ces éléments par:

$$E,\ A,\ B,\ C,\ D,\ F.$$

La loi de multiplication est définie au moyen du tableau suivant:

	E	A	B	C	D	F
E	E	A	B	C	D	F
A	A	E	D	F	B	C
B	B	F	E	D	C	A
C	C	D	F	E	A	B
D	D	C	A	B	F	E
F	F	B	C	A	E	D

(34)

Pour définir le produit, ce tableau doit être utilisé de la manière suivante. Si on veut, par exemple, former le produit DB, il faut repérer l'élément B dans la première ligne et l'élément D dans la première colonne, alors l'intersection de cette ligne et de cette colonne donne l'élément A comme produit DB. Il est alors aisé de vérifier que toutes les conditions que nous avons indiquées dans la définition du groupe abstrait sont vérifiées et que l'élément E joue bien le rôle d'élément unité.

Dans les paragraphes précédents ont été donnés des exemples de réalisation concrète de la notion abstraite de groupe. Dans un cas, le rôle d'élément était joué par la transformation linéaire (sa matrice) et la multiplication de deux éléments se réduisait à l'application successive de deux transformations linéaires, c'est-à-dire à la multiplication des matrices qui leur correspondent. Dans l'autre cas, le rôle d'élément était joué par une permutation et la multiplication de deux éléments était l'application successive de deux permutations. Donnons encore quelques exemples concrets de réalisation de groupes.

Supposons que les éléments sont tous les nombres complexes possibles et que la multiplication de deux éléments se ramène à l'addition des nombres complexes correspondants. Dans ce cas le rôle d'élément unité est joué par le nombre 0 et l'élément inverse du nombre complexe α est le nombre $(-\alpha)$. Au lieu des nombres complexes on aurait pu prendre pour éléments tous les vecteurs possibles $\mathbf{x}\,(x_1, x_2, \ldots, x_n)$ de l'espace complexe à n dimensions R_n et définir la multiplication des éléments comme l'addition des vecteurs correspondants. Dans ce cas, le rôle d'élément unité serait tenu par le vecteur nul. Autrement dit, les éléments du groupe seraient des vecteurs de R_n et la loi de formation du groupe l'addition des vecteurs. On notera que, dans ces deux derniers exemples, le résultat de la multiplication de deux éléments du groupe ne dépend pas de l'ordre des facteurs, c'est-à-dire que les éléments du groupe commutent par deux. De tels groupes sont des *groupes abéliens* [cf. II-2-14]. L'exemple le plus simple de groupe abélien est le *groupe cyclique* formé d'un seul élément E et des puissances successives d'un certain élément P. Si m est le plus petit entier positif pour lequel $P^m = E$, le groupe cyclique est formé de m éléments : $E, P, P^2, \ldots, P^{m-1}$. Si un tel nombre positif m n'existe pas, le groupe cyclique est infini : $E, P, P^2, \ldots$.

III-1-6. Sous-groupe. Soit un groupe G. Supposons qu'un certain ensemble H d'éléments, ne comprenant qu'une partie des éléments de G, forme également un groupe par rapport à la même règle de multiplication. Dans ce cas, le groupe H est appelé *sous-groupe* de G. Il est aisé de voir que l'ensemble formé du seul élément unité de G est toujours un sous-groupe. C'est le sous-groupe trivial. Dans la suite, lorsqu'on parlera de sous-groupes, on aura en vue des sous-groupes non triviaux.

On désigne par H_α les éléments du sous-groupe H. Soit G_1 un élément du groupe G tel que G_1 n'appartienne pas à H. Les produits G_1H_α, comme on l'a déjà vu, donnent différents éléments du groupe G n'appartenant pas à H. En effet, dans le cas contraire on aurait pour certaines valeurs α_1 et α_2 de l'indice α : $G_1H_{\alpha_1} = H_{\alpha_2}$ d'où $G_1 = H_{\alpha_2}H_{\alpha_1}^{-1}$, c'est-à-dire que G_1 devrait appartenir à H, ce qui est contraire à l'hypothèse. Supposons maintenant que G_1 et G_2 sont deux éléments différents communs au groupe G et n'appartenant pas au sous-groupe H. Montrons que les ensembles d'éléments G_1H_α et G_2H_α ou bien n'ont pas d'éléments communs, ou bien coïncident, c'est-à-dire qu'ils sont formés des mêmes éléments. En effet, si pour certaines valeurs de l'indice α on a $G_2H_{\alpha_2} = G_1H_{\alpha_1}$, il en résulte que $G_2 = G_1H_{\alpha_1}H_{\alpha_2}^{-1} = G_1H_{\alpha_3}$, c'est-à-dire que G_2 appartient à l'ensemble d'éléments G_1H_α et de même G_1 appartient à l'ensemble d'éléments G_2H_α. Il en résulte que les produits G_1H_α et G_2H_α forment un seul et même ensemble d'éléments.

Considérons tous les éléments H_α du sous-groupe H. Ils n'épuisent pas tous les éléments du groupe G. Examinons un certain élément G_1 qui n'appartient pas à H et formons tous les produits possibles G_1H_α. Ces produits, comme on l'a vu, sont tous distincts entre eux et différents de H_α.

Les éléments H_α et G_1H_α eux aussi peuvent ne pas épuiser tout le groupe. Prenons un élément quelconque G_2 du groupe n'appartenant ni aux H_α ni aux G_1H_α et formons tous les produits possibles G_2H_α. Les éléments G_2H_α, on l'a vu ci-dessus, seront tous distincts entre eux et différents des éléments H_α et G_1H_α. Si les éléments H_α, G_1H_α et G_2H_α n'épuisent pas tous les éléments du groupe G, on prend un autre élément G_3, n'appartenant pas aux trois ensembles d'éléments déjà indiqués, et on forme les produits G_3H_α. On obtient ainsi de nouveaux éléments du groupe, etc. On suppose qu'au moyen d'un nombre fini d'opérations de ce type on épuise tous les éléments du groupe G et que, pour cela, on a besoin de $(m-1)$ éléments G_h. Tous les éléments du groupe G pourront alors être représentés sous la forme

$$H_\alpha, \ G_1H_\alpha, \ G_2H_\alpha, \ \ldots, \ G_{m-1}H_\alpha, \tag{35}$$

où l'indice α prend toutes les valeurs qui correspondent au sous-groupe H. Si l'on pose $G'_k = G_kH_{\alpha_0}$, où α_0 est donné d'une façon quelconque, l'ensemble d'éléments G'_kH_α, comme cela a déjà été montré, coïncide avec l'ensemble G_kH_α. Autrement dit, dans chaque ensemble G_hH_α $(G_0 = I)$, tout élément de l'ensemble peut jouer le rôle de G_h. Il en résulte que pour un sous-groupe donné H_α, la partition des éléments du groupe G en ensembles du type (35) est parfaitement définie. Les ensembles G_kH_α sont appelés *ensembles conjugués par rapport au sous-groupe* H_α.

Dans le cas considéré (35) le sous-groupe H est appelé *sous-groupe fini* et, plus exactement, *sous-groupe d'indice* m. Si le groupe G est fini, l'indice du sous-groupe H sera égal évidemment au rapport de l'ordre du groupe G entier à l'ordre du sous-groupe H, *l'ordre d'un groupe fini étant défini par le nombre d'éléments qu'il contient*. On notera que dans les ensembles (35) seul le premier ensemble forme un sous-groupe. Chacun des autres ensembles G_kH_α ne contient pas l'élément unité et, par conséquent, ne peut pas former un sous-groupe.

Pour construire l'ensemble (35) on a multiplié à gauche les éléments H_α du sous-groupe H par les éléments G_k du groupe G. On aurait pu effectuer cette multiplication également à droite. En introduisant au lieu de G_k une autre notation G'_k, on arriverait ainsi à une représentation des éléments du groupe G

$$H_\alpha, \ H_\alpha G'_1, \ H_\alpha G'_2, \ \ldots, \ H_\alpha G'_{m-1}, \tag{36}$$

au lieu de (35); alors l'indice m du sous-groupe ne change pas. L'ensemble d'éléments G_kH_α s'appelle parfois ensemble conjugué à gauche et $H_\alpha G'_k$ ensemble conjugué à droite.

On notera avant tout que si α prend toutes les valeurs correspondant au sous-groupe H, les éléments H_α^{-1} donnent tous les éléments de H. Ceci résulte directement de ce que l'élément inverse d'un élément appartenant à H appartient aussi à H. Démontrons maintenant que les indices des ensembles conjugués à droite et à gauche sont égaux. Prenons dans (35) deux ensembles quelconques différents G_pH_α et G_qH_α (p différent de q) et pour le premier posons, par exemple, $G_p = E$. Écrivons les éléments inverses:

$$(G_pH_\alpha)^{-1} = H_\alpha^{-1}G_p^{-1} \quad \text{et} \quad (G_qH_\alpha)^{-1} = H_\alpha^{-1}G_q^{-1}.$$

Tenant compte de la remarque déjà faite, on peut écrire ces ensembles d'éléments sous la forme $H_\alpha G_p^{-1}$, $H_\alpha G_q^{-1}$. Il est aisé de voir qu'ils ne doivent pas avoir d'éléments communs. En effet, s'ils en avaient, on aurait:

$$H_{\alpha_1} G_p^{-1} = H_{\alpha_2} G_q^{-1},$$

d'où :

$$G_p^{-1} G_q = H_{\alpha_1}^{-1} H_{\alpha_2} = H_{\alpha_3} \text{ soit } G_q = G_p H_{\alpha_3},$$

et il apparaîtrait alors que G_q appartient à l'ensemble $G_p H_{\alpha_3}$, ce qui est impossible. Ici les ensembles

$$H_\alpha,\ H_\alpha G_1^{-1},\ H_\alpha G_2^{-1},\ \ldots,\ H_\alpha G_{m-1}^{-1}$$

sont conjugués à droite, de sorte que dans (36) on peut simplement prendre $G_s' = G_s^{-1}$.

Etudions quelques exemples de sous-groupes. Soit G l'ensemble des transformations réelles orthogonales à trois variables et H l'ensemble des transformations réelles orthogonales à trois variables à déterminant égal à $(+1)$. Toute transformation réelle orthogonale est soit une rotation et appartient à H, soit le produit d'une rotation par une symétrie par rapport à l'origine et s'exprime par les formules

$$x' = -x; \quad y' = -y; \quad z' = -z. \quad (S) \tag{37}$$

Dans ce cas, le groupe G peut être mis sous la forme:

$$H_\alpha,\ SH_\alpha \tag{38}$$

ou bien :

$$H_\alpha,\ H_\alpha S, \tag{39}$$

où H_α désigne l'ensemble de tous les éléments du groupe H. Dans ce cas H_α est un sous-groupe d'indice 2.

Soit G le groupe symétrique des permutations de n éléments, H le groupe alterné constitué de permutations paires. Soit de plus S une permutation impaire quelconque, par exemple, la permutation formée par le cycle (1, 2), c'est-à-dire qui se réduit à la transposition des éléments 1 et 2. On peut évidemment représenter ici le groupe G sous la forme (38) ou (39). Dans les deux cas, la multiplication à gauche conduit au même résultat que la multiplication à droite.

Le groupe alterné est donc un sous-groupe symétrique d'indice 2.

Examinons encore le groupe fini de l'octaèdre régulier dont il a déjà été question [III-1-2]. Soit A un sommet quelconque de l'octaèdre et l l'axe passant par ce sommet. Soient S_0, S_1, S_2, S_3 les rotations d'angles 0, $\frac{\pi}{2}$, π et $\frac{3\pi}{2}$ autour de cet axe. Ces rotations forment un sous-groupe du groupe complet des rotations de l'octaèdre. Soit T_k ($k = 1, 2, 3, 4, 5$) les rotations qui transposent le sommet A aux cinq autres sommets de l'octaèdre. On peut représenter le groupe complet de l'octaèdre sous la forme :

$$S_\alpha, \; T_1S_\alpha, \; T_2S_\alpha, \; T_3S_\alpha, \; T_4S_\alpha, \; T_5S_\alpha,$$

et donc le sous-groupe S_α est un sous-groupe d'indice 6.

Soient G_s, G_s^{-1} ($s = 1, 2, \ldots, k$) des éléments du groupe G. Considérons l'ensemble de tous les éléments du groupe G que l'on peut représenter sous la forme d'un produit des éléments G_s, G_s^{-1} ($s = 1, 2, \ldots, k$).

Cet ensemble d'éléments forme, de façon évidente, un groupe qui est un sous-groupe de G ou qui coïncide avec G. On dit que ce sous-groupe est engendré par l'ensemble d'éléments G_s, G_s^{-1} ($s = 1, 2, \ldots, k$).

III-1-7. Classes et sous-groupe distingué. Soient U et V des éléments d'un groupe. L'élément $W = VUV^{-1}$ est *l'élément conjugué de U*. Il est aisé de voir qu'inversement, U est le conjugué de W. En effet, $U = V^{-1}WV$. Deux éléments U_1 et U_2 conjugués avec un même élément W :

$$U_1 = V_1WV_1^{-1}; \quad U_2 = V_2WV_2^{-1},$$

sont conjugués entre eux :

$$U_2 = V_2V_1^{-1}U_1(V_2V_1^{-1})^{-1}.$$

L'ensemble de tous les éléments conjugués d'un groupe forme ce que l'on appelle une *classe du groupe*. Une classe est parfaitement définie si l'on donne l'un de ses éléments U. En effet, si l'on se donne U, on obtient toute la classe au moyen de la formule $G_aUG_a^{-1}$, où G_a parcourt tous les éléments du groupe. Ainsi on peut diviser tout le groupe en classes. En tenant compte de la propriété fondamentale de l'élément unité que l'on a déjà donnée [III-1-5], on a

$$G_aIG_a^{-1} = I,$$

c'est-à-dire que l'élément unité forme à lui seul une classe.

Si l'élément U est d'ordre m, c'est-à-dire si m est le plus petit entier positif pour lequel $U^m = I$, tout élément conjugué $G_aUG_a^{-1}$ est du même ordre m, ce qui résulte directement de l'égalité

$$(G_aUG_a^{-1})^m = G_aU^mG_a^{-1} = I.$$

Autrement dit, tous les éléments d'une même classe ont un même ordre.

Notons que lorsque G_a parcourt tous les éléments du groupe G, le produit $G_aUG_a^{-1}$ peut donner plusieurs fois les éléments de la classe. Par exemple, si $U = I$, on a déjà vu que le produit obtenu donne toujours I.

Comme exemple, prenons à nouveau le groupe des rotations de l'octaèdre. Soit U la rotation d'angle $\frac{\pi}{2}$ autour d'un axe A_pA_q de l'octaèdre. Si la rotation T_k appartenant au groupe considéré transforme l'axe l en l_1, le sommet A_p en A_r et le sommet A_q en A_s, alors l'élément du groupe $T_kUT_k^{-1}$ donne une rotation d'angle $\frac{\pi}{2}$ autour de l'axe A_rA_s. Si par exemple T_k transforme A_p en A_q, le produit indiqué donnera une rotation d'angle $\frac{\pi}{2}$ autour de l'axe A_qA_p, autrement dit, une rotation d'angle $\frac{3\pi}{2}$ autour de l'axe A_pA_q. Si T_k transforme l'axe A_pA_q en lui-même, c'est-à-dire s'il y a rotation autour de cet axe, le produit $T_kUT_k^{-1}$ coïncide avec U. Dans ce cas, la classe des éléments conjugués de U représente l'ensemble des rotations d'angle $\frac{\pi}{2}$ autour des axes de l'octaèdre.

Il en est de même si on prend le groupe des rotations de l'espace à trois dimensions autour de l'origine, car, comme on le sait, tout élément U de ce groupe se présente comme une rotation d'un angle φ autour d'un certain axe. Dans ce cas la classe des éléments conjugués de U sera l'ensemble des rotations d'angle φ autour de tous les axes possibles passant par l'origine.

Une autre notion importante liée à la notion de classe est celle de *sous-groupe distingué* dont nous parlerons maintenant. Soient G un groupe et H un de ses sous-groupes. Soit G_1 un élément donné de G. Examinons l'ensemble d'éléments de ce groupe représentés par le produit

$$G_1H_\alpha G_1^{-1}, \tag{40}$$

où H_α désigne un élément variable du sous-groupe H, c'est-à-dire où H_α parcourt tous les éléments du sous-groupe H. Il est aisé de voir que les produits (40) forment aussi un sous-groupe. En effet, si l'on prend, par exemple, le produit de deux éléments appartenant à l'ensemble (40), il appartient également à cet ensemble

$$(G_1H_{\alpha_2}G_1^{-1})\,(G_1H_{\alpha_1}G_1^{-1}) = G_1H_{\alpha_2}H_{\alpha_1}G_1^{-1} = G_1H_{\alpha_3}G_1^{-1},$$

et l'on a de même toutes les conditions pour avoir un groupe.

Le sous-groupe (40) est *sous-groupe associé à* H et si G_1 appartient au sous-groupe H, le sous-groupe (40) est formé d'éléments appartenant aussi à H et il est aisé de voir qu'il coïncide tout simplement avec H.

Tout élément H_{α_0} du sous-groupe H peut être obtenu dans ce cas au moyen de la formule (40) si l'on prend

$$H_\alpha = G_1^{-1}H_{\alpha_0}G_1.$$

Si l'élément G_1 n'appartient pas au sous-groupe H, le sous-groupe (40) peut être différent du sous-groupe H.

Le sous-groupe H est appelé *sous-groupe distingué* du groupe G si, quel que soit le choix de l'élément G_1 de G, le sous-groupe (40) coïncide avec le sous-groupe H. Plus loin nous donnerons des exemples de sous-groupes distingués de groupes, mais pour l'instant passons à l'examen de quelques notions nouvelles, liées à celle de sous-groupe distingué.

Soit H un sous-groupe distingué du groupe G. Pour simplifier l'écriture, supposons que ce sous-groupe ait un indice fini m. Dans ce cas, tous les éléments du groupe G peuvent être représentés par :

$$H_\alpha,\ G_1H_\alpha,\ G_2H_\alpha,\ \ldots,\ G_{m-1}H_\alpha, \tag{41}$$

où H_α est un élément variable du sous-groupe H. Comme H est sous-groupe distingué, l'ensemble des éléments $G_kH_\alpha G_k^{-1}$ coïncide avec l'ensemble des éléments H_α, c'est-à-dire que l'ensemble des éléments G_kH_α coïncide avec l'ensemble des éléments $H_\alpha G_k$.

Ainsi, si *H est le sous-groupe distingué*, la partition des éléments du groupe en ensembles conjugués selon (41) coïncide avec la partition des éléments en ensembles conjugués selon la formule :

$$H_\alpha,\ H_\alpha G_1,\ H_\alpha G_2,\ \ldots,\ H_\alpha G_{m-1}. \tag{42}$$

Autrement dit, dans ce cas *les ensembles conjugués à droite coïncident avec les ensembles conjugués à gauche.*

Si H_{α_0} est un élément quelconque du sous-groupe distingué, quel que soit G_0 appartenant à G, l'élément $G_0H_{\alpha_0}G_0^{-1}$ appartient aussi au sous-groupe distingué, c'est-à-dire que si un élément appartient au sous-groupe distingué, toute la classe à laquelle il appartient dans le groupe appartient également au sous-groupe distingué. Il est facile de montrer qu'inversement, si un sous-groupe, qui contient un élément, contient toute la classe à laquelle cet élément appartient dans le groupe, ce sous-groupe est un sous-groupe distingué.

Passons maintenant à l'examen des ensembles conjugués selon les formules (41) ou (42), où les éléments H_α forment un sous-groupe distingué H. Soient $G_lH_\alpha G_kH_{\alpha'}$ les produits des éléments d'un ensemble conjugué G_lH_α par les éléments de l'ensemble conjugué $G_kH_{\alpha'}$.

On peut mettre l'ensemble de ces produits sous la forme :

$$G_l\,(H_\alpha G_k)\,H_{\alpha'}, \qquad G_lG_kH_\alpha H_{\alpha'}.$$

Les éléments H_α et $H_{\alpha'}$ appartiennent au sous-groupe distingué H et on peut en dire autant de leur produit. De sorte que les produits précédents peuvent être mis sous la forme :

$$G_lG_kH_\alpha.$$

Tous les éléments de ce type appartiennent à un seul et même ensemble conjugué (41), à savoir l'ensemble conjugué auquel appartient l'élément $G_l G_h$. Il est aussi aisé de montrer que l'on obtient ainsi tous les éléments de cet ensemble conjugué. En bref, si un sous-groupe est un sous-groupe distingué, en multipliant l'un des ensembles conjugués par l'autre, on obtient également un ensemble conjugué.

Considérons d'une part chacun des ensembles conjugués comme un élément nouveau et d'autre part le premier des ensembles conjugués de la formule (41), (38), sera pris comme élément unité. Le résultat précédent sur la multiplication des ensembles conjugués nous permet de multiplier les nouveaux éléments introduits, et de voir que cette règle de multiplication vérifie toutes les conditions que doivent remplir des éléments pour former un groupe, c'est-à-dire que les éléments que l'on a ainsi introduits, avec la règle indiquée de multiplication, forment un groupe dans lequel le premier des ensembles conjugués joue le rôle d'élément unité. Ce nouveau groupe dont l'ordre est égal à l'indice du sous-groupe distingué H est appelé groupe quotient de G par H.

Tout groupe G a deux sous-groupes distingués triviaux : l'un est formé du seul élément unité et l'autre coïncide avec le groupe tout entier.

Dans la suite, on considérera le sous-groupe distingué comme différent des deux sous-groupes distingués triviaux indiqués. Il peut arriver qu'un groupe n'ait pas de sous-groupes distingués.

Dans ce cas on l'appelle *groupe simple*.

III-1-8. Exemples. 1. Soit le groupe G des transformations réelles orthogonales dans l'espace à trois dimensions. Soit H le sous-groupe des déplacements, c'est-à-dire l'ensemble des transformations orthogonales à déterminant égal à $(+1)$. Soit de plus S la symétrie par rapport à l'origine des coordonnées définie par la formule (37). Si H_a est un élément variable de H, le groupe G peut être mis sous la forme

$$H_a, \; SH_a \quad \text{ou} \quad H_a, \; H_aS. \tag{43}$$

Si G_1 est une transformation quelconque de G, alors $G_1 H_a G_1^{-1}$ a un déterminant égal à $(+1)$, c'est-à-dire qu'il appartient à H et H est un sous-groupe distingué d'indice 2. Considérons le groupe quotient de G par H. Le premier des ensembles (43) correspond à l'élément unité E de ce groupe. Le produit de deux éléments du second ensemble, c'est-à-dire de deux transformations orthogonales à déterminant égal à (-1), donne une transformation orthogonale à déterminant égal à $(+1)$ qui appartient au premier ensemble. Si K est un élément qui correspond au second ensemble, il résulte de ce qui a été dit que $K^2 = E$. Ainsi le groupe quotient de G par H est formé des deux éléments E et K et on a $K^2 = E$, c'est-à-dire un groupe cyclique d'ordre 2. Ceci aura lieu en général pour les sous-groupes distingués d'indice 2.

2. Le groupe alterné sera un sous-groupe distingué d'indice 2 du groupe symétrique des permutations.

Ecrivons les éléments du groupe symétrique à trois éléments et désignons chacun d'eux par une lettre en utilisant les notations de [III-1-4] :

$$E\,; \quad A = (2,3)\,; \quad B = (1,2)\,; \quad C = (1,3)\,; \quad D = (1,3,2)\,: \quad F = (1,2,3).$$

Le groupe alterné formé des permutations E, D et F est un groupe cyclique du troisième ordre ($F = D^2$ et $D = F^2$) et on a $D^3 = F^3 = E$. Tout le groupe symétrique est formé de trois classes: la classe I — E, la classe II — A, B et C, la classe III — D et F.

Le groupe alterné est formé aussi de trois classes; la classe I — E, la classe II — D et la classe III — F. Il est facile de vérifier que la loi de multiplication des éléments du groupe symétrique considéré coïncide avec la loi donnée dans le tableau (34) du paragraphe [III-1-5].

Pour $n = 4$, ce groupe comprend 12 éléments qui se décomposent en quatre classes:

I E;

II $A_1 = (1, 2)(3, 4)$; $A_2 = (1, 3)(2, 4)$; $A_3 = (1, 4)(2, 3)$;

III $B_1 = (1, 2, 3)$; $B_2 = (2, 1, 4)$; $B_3 = (3, 4, 1)$; $B_4 = (4, 3, 2)$;

IV $C_1 = (1, 2, 4)$; $C_2 = (2, 1, 3)$; $C_3 = (3, 4, 2)$; $C_4 = (4, 3, 1)$.

La classe II contient trois éléments du deuxième ordre, les classes III et IV contiennent chacune quatre éléments du troisième ordre. Le produit de deux éléments de la classe II donne, comme il est aisé de le vérifier, de nouveau un élément de la classe II et puisque tous les éléments du deuxième ordre sont dans la classe II, on peut affirmer que ces trois éléments forment avec l'élément unité un sous-groupe distingué du groupe alterné considéré. Il est d'ordre 4 et d'indice 3. Il est aisé de vérifier que les éléments B_i de la classe III entrent tous dans un des ensembles conjugués par rapport au sous-groupe distingué et les éléments C_i dans un autre ensemble conjugué et, de plus, que le produit de deux éléments de la classe III donne un élément de la classe IV et le produit de deux éléments de la classe IV un élément de la classe III. Dans le groupe quotient l'élément unité E correspond au sous-groupe distingué indiqué. Soient A et B deux autres éléments du groupe quotient. Il résulte de ce qui a été déjà dit que $A^2 = B$ et $B^2 = A$ et il est alors évident que le groupe quotient formé des éléments E, A et A^2, tel que $A^3 = E$, est un groupe cyclique du troisième ordre.

On notera que les éléments E (1, 2, 3) et (2, 1, 3) de notre groupe alterné initial forment un sous-groupe cyclique du troisième ordre, mais que ce sous-groupe n'est pas un diviseur élémentaire.

Si on numérote les sommets du tétraèdre suivant un ordre quelconque, il est aisé de vérifier directement que le groupe alterné indiqué ci-dessus pour $n = 4$ correspond aux rotations pour lesquelles le tétraèdre se transforme en lui-même. Toute permutation définit le passage de l'un à l'autre des sommets. Les permutations de la classe III correspondent à des rotations d'angle $\frac{2\pi}{3}$ autour d'un des axes du tétraèdre et les permutations de la classe IV à des rotations du même angle autour des mêmes axes mais dans le sens négatif. Ainsi, les permutations (1, 2, 3) et (2, 1, 3) correspondent à des rotations autour de l'axe passant par le sommet qui porte le numéro 4. Les permutations de la classe II correspondent à des rotations du tétraèdre telles qu'aucun des sommets ne reste inchangé.

On peut montrer que pour $n > 4$ ce groupe alterné est simple.

3. Si on a un groupe abélien G et un sous-groupe quelconque H, quel que soit le choix des éléments H_α de H et G_i de G, on a $G_iH_\alpha = H_\alpha G_i$, c'est-à-dire que $G_iH_\alpha G_i^{-1} = H_\alpha$, d'où on voit que H est un sous-groupe distingué, c'est-à-dire que tout sous-groupe d'un groupe abélien est sous-groupe distingué de celui-ci. Comme exemple étudions le groupe G d'addition des vecteurs dans l'espace R_n dont il a déjà été question [II-2-18].

Comme sous-groupe H prenons les vecteurs appartenant à un certain sous-espace L_k de R_n ($0 < k < n$). Les ensembles conjugués s'obtiennent en ajoutant à un vecteur $\mathbf{x}$ de R_n tous les vecteurs du sous-espace L_k.

Si $\mathbf{x}$ appartient à L_k, l'ensemble conjugué coïncide avec le sous-groupe H. Introduisons les vecteurs de base $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \ldots, \mathbf{x}^{(k)}$ dans L_k et les vecteurs de base $\mathbf{x}^{(k+1)}, \ldots, \mathbf{x}^{(n)}$ dans le sous-espace supplémentaire M_{n-k}. Tout ensemble conjugué d'éléments sera formé, d'après ce qui a été dit ci-dessus, des vecteurs

$$c_1\mathbf{x}^{(1)}+c_2\mathbf{x}^{(2)}+\ldots+c_k\mathbf{x}^{(k)}+c_{k+1}\mathbf{x}^{(k+1)}+\ldots+c_n\mathbf{x}^{(n)},$$

où $c_{k+1}, \ldots, c_n$ ont des valeurs fixées, alors que $c_1, c_2, \ldots, c_k$ peuvent être pris arbitraires.

Ainsi tout ensemble conjugué peut être comparé à un vecteur déterminé de M_{n-k} et inversement à tout vecteur de M_{n-k} correspond un ensemble conjugué bien défini. A l'addition de deux vecteurs quelconques de deux ensembles conjugués correspond l'addition de vecteurs correspondant à ces ensembles dans M_{n-k}. Autrement dit, les éléments du groupe quotient peuvent être considérés comme vecteurs de M_{n-k} d'après la règle de formation des groupes précédente (addition de vecteurs).

Dans cet exemple, l'ordre du sous-groupe distingué B et son indice sont infinis.

III-1-9. Groupes isomorphes et homomorphes. On dit que deux groupes A et B sont *isomorphes* si on peut établir entre leurs éléments une correspondance telle qu'à tout élément de A corresponde un élément défini de B et, inversement, qu'à tout élément de B corresponde un élément défini de A (correspondance biunivoque), cette correspondance étant telle qu'au produit de deux éléments quelconques de A corresponde le produit des éléments correspondants de B. Si A et B sont des groupes abstraits isomorphes, ils ont une structure entièrement identique, c'est-à-dire qu'en fait ils ne diffèrent pas l'un de l'autre.

Passons maintenant à une nouvelle notion plus générale que la notion de groupe isomorphe. On dit qu'un groupe B est *homomorphe* à A si à chaque élément de A correspond un élément déterminé de B, et à chaque élément de B correspond au moins un élément de A, cette relation étant telle qu'au produit de deux éléments de A corresponde le produit des éléments correspondants de B. Dans ce cas, contrairement au cas des groupes isomorphes, la correspondance n'est pas biunivoque, c'est-à-dire qu'un seul et même élément du groupe B peut correspondre à plusieurs éléments différents du groupe A. Si A et B sont homomorphes et si à chaque élément de B correspond un élément déterminé de A, ces deux groupes sont aussi isomorphes. Remarquons de plus que si les éléments A_1 et A_2 de A correspondent aux éléments B_1 et B_2 de B, de la définition même de l'élément A_2A_1 de A il résulte qu'il correspond à l'élément B_2B_1 de B.

Soit A_0 l'élément unité de A et B_0 l'élément correspondant de B. Il est aisé de montrer que B_0 est l'élément unité. En effet, pour tout A_1 de A on a l'égalité:

$$A_0A_1=A_1A_0=A_1,$$

qui donne l'égalité correspondante entre éléments de B:

$$B_0B_1=B_1B_0=B_1,$$

et d'après la définition de l'homomorphisme, B_1 doit être considéré comme un élément quelconque de B. Cette dernière égalité montre que B_0 est l'élément unité de B. Ainsi, dans les groupes isomorphes et homomorphes à l'élément unité de A correspond l'élément unité de B. Prenons maintenant deux éléments inverses l'un de l'autre A_1 et A_1^{-1} de A. Soient B_1 et B_2 les éléments correspondants de B. L'égalité $A_1A_1^{-1} = A_1^{-1}A_1 = A_0$, où A_0 est l'élément unité, donne, d'après la définition des groupes homomorphes, $B_1B_2 = B_2B_1 = B_0$, où, d'après ce qui précède, B_0 est l'élément unité et donc on a $B_2 = B_1^{-1}$, c'est-à-dire qu'à des éléments inverses pris dans A correspondent des éléments inverses dans B.

Supposons que les groupes soient homomorphes et non isomorphes. Considérons l'ensemble des éléments C_α du groupe A auxquels correspond l'élément unité B_0 de B. Si C_α correspond à B_0, d'après ce qui précède, C_α^{-1} correspond à $B_0^{-1} = B_0$ et à chaque produit $C_{\alpha_2}C_{\alpha_1}$ correspond aussi $B_0B_0 = B_0$, c'est-à-dire que l'ensemble des éléments C_α du groupe A auxquels correspond l'élément unité de B forme un sous-groupe C du groupe A.

Montrons que ce sous-groupe est un sous-groupe distingué. En effet, soient A_1 un élément quelconque du groupe A et B_1 l'élément correspondant de B. A tout élément du type $A_1C_\alpha A_1^{-1}$ correspond on B un élément $B_1B_0B_1^{-1}$ ou, si l'on tient compte de la propriété fondamentale de l'élément unité, on peut dire qu'à tout élément du type $A_1C_\alpha A_1^{-1}$ correspond l'élément unité de B, c'est-à-dire que tout élément de cette forme est l'un des éléments de C_α, qu'il appartient au sous-groupe C et, par conséquent, que ce sous-groupe C est un sous-groupe distingué. Considérons maintenant la partition du groupe A en ensembles conjugués

$$C_\alpha, \quad A_1C_\alpha, \quad A_2C_\alpha, \ldots \tag{44}$$

Soit B_k l'élément correspondant à A_k. On prend deux éléments $A_kC_{\alpha_1}$ et $A_kC_{\alpha_2}$ appartenant à un même ensemble conjugué. Il leur correspondra les éléments B_kB_0 et B_kB_0, c'est-à-dire un même élément B_k de B.

Aux éléments A_kC_α et A_lC_α d'ensembles conjugués différents correspondent des éléments B_k et B_l. Montrons que ces derniers éléments sont différents. En effet, s'ils étaient identiques, à l'élément $A_k^{-1}A_l$ correspondrait l'élément unité B_0 de B, c'est-à-dire que l'élément $A_k^{-1}A_l$ devrait être l'un des éléments C_α et que l'on aurait $A_k^{-1}A_l = C_{\alpha_0}$ ou encore $A_l = A_kC_{\alpha_0}$, ce qui est impossible en raison de (44). Ainsi, *si le groupe B est homomorphe au groupe A, l'ensemble des éléments C de A qui correspondent à l'élément unité de B forme un sous-groupe distingué et tout ensemble conjugué avec ce sous-groupe distingué constitue l'ensemble de tous les éléments de A auxquels correspond un seul et même élément de B.* De la définition des groupes homomorphes, il résulte qu'au produit de deux éléments quelconques appartenant à des ensembles conjugués différents (ou identiques) correspond le produit des éléments correspondants du groupe B, autrement dit, qu'à tout ensemble conjugué de A correspond un élément de B, qu'à différents ensembles conjugués correspondent des éléments différents de B et cette correspondance montre que le groupe B est isomorphe au groupe quotient C_a dans A.

Comme exemple, prenons à nouveau le groupe des transformations réelles orthogonales dans l'espace à trois dimensions et associons à chaque transformation un nombre égal au déterminant de la transformation, en définissant le produit des nombres de la façon habituelle. Dans ce cas, le groupe sera homomorphe au groupe formé des deux éléments $(+1)$ et (-1) et la multiplication pour ces deux éléments est définie de la façon habituelle comme la multiplication des nombres. Le rôle d'élément unité sera joué par $(+1)$. Pour cet exemple, le sous-groupe distingué sera le groupe des rotations.

Si le groupe B est homomorphe, mais non isomorphe, au groupe A, l'ensemble des éléments du groupe A auxquels correspond l'élément unité de B est appelé *noyau d'homomorphisme*. On a vu que le noyau d'homomorphisme est un sous-groupe distingué du groupe A.

III-1-10. Exemples. 1. Soit le groupe G des transformations réelles orthogonales dans l'espace à trois dimensions. Associons à chacun de ses éléments un nombre égal au déterminant de cette transformation et définissons l'opération de groupe pour ces nombres comme étant la multiplication ordinaire. Alors le groupe G', formé des nombres $(+1)$ et (-1) avec la multiplication habituelle de ces nombres, sera homomorphe au groupe G. L'élément unité $(+1)$ du groupe G' correspond aux rotations de l'espace à trois dimensions de G. Ces rotations forment un sous-groupe distingué et le groupe quotient est un groupe cyclique d'ordre 2 [III-1-8].

2. Soit dans le plan XY le triangle équilatéral dont les sommets sont

$$(1,0)\,;\quad (\cos 120°,\ \sin 120°)\,;\qquad (\cos 240°,\ \sin 240°)$$

et formons le groupe G constitué des rotations du plan d'angles 0°, 120°, 240° autour de l'origine, pour lesquelles le triangle se transforme en lui-même et des symétries du plan par rapport à l'axe X suivies d'une rotation d'angles 0°, 120°, 240°. Ce sera le groupe diédral pour $n = 3$.

Ecrivons les matrices correspondant aux éléments de ce groupe:

$$E=\begin{Vmatrix}1 & 0\\ 0 & 1\end{Vmatrix};\quad A=\begin{Vmatrix}1 & 0\\ 0 & -1\end{Vmatrix};\quad B=\begin{Vmatrix}-\frac{1}{2} & \frac{1}{2}\sqrt{3}\\ \frac{1}{2}\sqrt{3} & \frac{1}{2}\end{Vmatrix},$$

$$C=\begin{Vmatrix}-\frac{1}{2} & -\frac{1}{2}\sqrt{3}\\ -\frac{1}{2}\sqrt{3} & \frac{1}{2}\end{Vmatrix};\quad D=\begin{Vmatrix}-\frac{1}{2} & \frac{1}{2}\sqrt{3}\\ -\frac{1}{2}\sqrt{3} & -\frac{1}{2}\end{Vmatrix};$$

$$F=\begin{Vmatrix}-\frac{1}{2} & -\frac{1}{2}\sqrt{3}\\ \frac{1}{2}\sqrt{3} & -\frac{1}{2}\end{Vmatrix}.$$

Si on prend les formules de multiplication données par le tableau (34) [III-1-5], on voit qu'elles correspondent à la multiplication des matrices formant le groupe. On a vu ci-dessus [III-1-8] que ceci correspond aussi au groupe symétrique des permutations de trois éléments

$$E;\ A=(2, 3);\quad B=(1, 2);\quad C=(1, 3);\quad D=(1, 3, 2);\quad F=(1, 2, 3). \qquad (45)$$

Ainsi, si l'on considère que les éléments de ces deux groupes, désignés par une seule et même lettre, correspondent, ces deux groupes seront isomorphes. Les permutations du groupe (45) correspondent aux permutations des sommets du triangle considéré, si l'on les numérote d'une façon adéquate.

Comme il a déjà été vu au paragraphe [III-1-8] le groupe du tétraèdre est isomorphe au groupe alterné des permutations pour $n = 4$.

3. On peut donner une méthode générale de construction des groupes des permutations homomorphes au groupe G donné. Soit H un sous-groupe quelconque de G d'indice fini n. Donnons-nous les ensembles conjugués d'éléments qui lui correspondent:

$$H,\ HS_1,\ HS_2,\ \ldots,\ HS_{n-1}.$$

Si l'on multiplie à droite chacun de ces ensembles par un certain élément S de G, il y aura simplement une certaine permutation de l'ordre de ces ensembles et l'on considère que cette permutation correspond à l'élément S de G. Il est facile de montrer qu'on obtient ainsi un groupe des permutations G', homomorphe au groupe G.

Pour que l'élément S de G corresponde à l'élément unité de G', il faut et il suffit que lors de la multiplication à droite par S tout ensemble conjugué se transforme en lui-même, c'est-à-dire que

$$H_\alpha S = H_\beta \text{ et } H_\alpha S_k S = H_\beta S_k \qquad (k=1, 2, \ldots, n-1),$$

où H_α est un élément quelconque de H et H_β appartient aussi à H. Les égalités qui ont été écrites peuvent être mises sous la forme

$$S = H_\alpha^{-1} H_\beta;\qquad S = (S_k^{-1} H_\alpha S_k)^{-1}(S_k^{-1} H_\beta S_k),$$

et il en résulte que pour que l'élément S corresponde à l'élément unité de G', il faut et il suffit que S appartienne simultanément à H et à tous les sous-groupes associés $S_k^{-1} H S_k$.

Si H est un sous-groupe distingué de G, la condition indiquée se réduit à l'appartenance de S à H et le groupe G' est, dans ce cas, isomorphe au groupe quotient. Si H se réduit à un seul élément unité, le groupe G est isomorphe au groupe des permutations G' qui s'obtient, si l'on multiplie à droite les éléments du groupe G:

$$E,\ S_1,\ S_2,\ \ldots,\ S_n$$

par un élément quelconque S de G, ce qui donne une certaine permutation des éléments de G. Dans la suite, on examinera en détail la construction de groupes des transformations linéaires isomorphes à un groupe donné.

III-1-11. Projection stéréographique. Après avoir donné l'exposé de la théorie générale des groupes passons à l'examen d'un exemple particulier de correspondance entre groupes, jouant un rôle important en physique. Etudions d'abord la *projection stéréographique* qui donne une certaine loi de correspondance entre les points d'une sphère et d'un plan.

Soit l'espace à trois dimensions repéré par les axes de coordonnées XYZ et la sphère C de centre à l'origine et de rayon unité. Soit S le point de la sphère de coordonnées $(0, 0, -1)$ et M un point variable de la sphère (fig. 3). La droite SM coupe le plan XY en un point P et on a ainsi une loi parfaitement définie de correspondance entre les points de la sphère C et les points du plan XY. Le point de la sphère S de coordonnées $(0, 0, -1)$ correspond à un point du plan infiniment éloigné. Cette correspondance que nous venons d'établir donne la projection stéréographique de la sphère sur le plan.

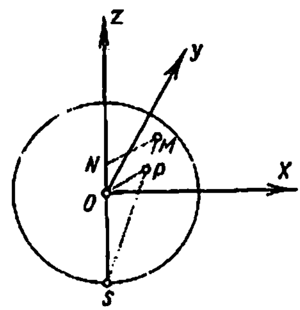


Fig. 3

Cherchons maintenant les formules qui donnent la projection stéréographique. Soit MN la perpendiculaire à l'axe Z issue du point M. On a, à partir de la similitude des triangles et en tenant compte de ce que $SO = 1$:

$$NM = (1 + ON)\,OP. \tag{46}$$

En désignant par (x, y, z) les coordonnées du point M et par (α, β) les coordonnées de P, on peut écrire:

$$NM = (1 + z)\,OP,$$

ou en projetant les segments parallèles OP et NM sur les axes X et Y:

$$x = (1 + z)\,\alpha; \quad y = (1 + z)\,\beta. \tag{47}$$

L'équation $x^2 + y^2 + z^2 = 1$ nous donne une équation du second degré pour z:

$$(\alpha^2 + \beta^2)(1 + z)^2 + z^2 - 1,$$

et en la résolvant on obtient

$$z = \frac{\pm 1 - (\alpha^2 + \beta^2)}{1 + (\alpha^2 + \beta^2)}.$$

Mais pour tous les points (α, β) situés à une distance finie nous devons avoir $z > -1$ et, par conséquent, dans la formule précédente on doit prendre $(+1)$. En utilisant les formules (47) on obtient finalement les expressions de (x, y, z) en fonction de (α, β):

$$x = \frac{2\alpha}{1 + \alpha^2 + \beta^2}; \quad y = \frac{2\beta}{1 + \alpha^2 + \beta^2}; \quad z = \frac{1 - (\alpha^2 + \beta^2)}{1 + \alpha^2 + \beta^2}. \tag{48}$$

Au lieu des deux coordonnées réelles α et β dans le plan, on peut introduire la coordonnée complexe $\zeta = \alpha + i\beta$. En désignant comme toujours par $\bar{\zeta}$ le nombre complexe conjugué de ζ, on peut écrire les formules précédentes sous la forme :

$$x+iy=\frac{2\zeta}{1+\zeta\bar{\zeta}}\,;\quad x-iy=\frac{2\bar{\zeta}}{1+\zeta\bar{\zeta}}\,;\quad z=\frac{1-\zeta\bar{\zeta}}{1+\zeta\bar{\zeta}}\,. \tag{49}$$

Représentons la variable complexe ζ sous la forme du rapport de deux variables complexes ξ et η

$$\zeta=\frac{\eta}{\xi}\,. \tag{50}$$

Les couples de valeurs ξ et η diffèrent en général par un facteur commun, c'est-à-dire que des couples du type $k\xi$, $k\eta$ et ξ, η donnent une seule et même valeur de ζ, c'est-à-dire un seul et même point du plan, et le couple de valeurs $\eta\neq 0$, $\xi=0$ donne un point du plan infiniment éloigné. Les nombres complexes ξ et η sont les coordonnées complexes homogènes dans le plan. Les formules (49) peuvent être mises sous la forme :

$$x=\frac{\bar{\xi}\eta+\xi\bar{\eta}}{\xi\bar{\xi}+\eta\bar{\eta}}\,;\quad y=\frac{1}{i}\,\frac{\bar{\xi}\eta-\xi\bar{\eta}}{\xi\bar{\xi}+\eta\bar{\eta}}\,;\quad z=\frac{\xi\bar{\xi}-\eta\bar{\eta}}{\xi\bar{\xi}+\eta\bar{\eta}}\,, \tag{51}$$

en utilisant (50) et en séparant les parties réelles et imaginaires.

Pour toutes valeurs complexes de ξ et η ces dernières formules donnent pour (x, y, z) des valeurs réelles qui vérifient la relation :

$$x^2+y^2+z^2-1=0, \tag{51$_1$}$$

comme on pouvait d'ailleurs s'y attendre, car le point (x, y, z) se trouve sur la sphère unité.

III-1-12. Groupe unitaire et groupe des déplacements. Soit maintenant une transformation unitaire des variables (ξ, η) :

$$\xi'=a\xi+b\eta,\quad \eta'=c\xi+d\eta. \tag{52}$$

Comme cette transformation est unitaire, on doit avoir :

$$\xi'\bar{\xi}'+\eta'\bar{\eta}'=\xi\bar{\xi}+\eta\bar{\eta}. \tag{53}$$

Les nouvelles variables (ξ', η') donnent un nouveau point sur la sphère :

$$x'=\frac{\bar{\xi}'\eta'+\xi'\bar{\eta}'}{\xi'\bar{\xi}'+\eta'\bar{\eta}'}\,;\quad y'=\frac{1}{i}\,\frac{\bar{\xi}'\eta'-\xi'\bar{\eta}'}{\xi'\bar{\xi}'+\eta'\bar{\eta}'}\,;\quad z'=\frac{\xi'\bar{\xi}'-\eta'\bar{\eta}'}{\xi'\bar{\xi}'+\eta'\bar{\eta}'}\,. \tag{54}$$

Le déterminant de la transformation unitaire (52) est, comme on le sait, égal en module à l'unité et s'exprime, par conséquent, à l'aide d'un terme de la forme $e^{i\varphi}$. En multipliant tous les coefficients de la transformation (52) par $e^{-i\frac{\varphi}{2}}$ on obtient une transformation unitaire à déterminant égal à 1. Mais alors ξ' et η' sont également multipliés par $e^{-i\frac{\varphi}{2}}$. Ce facteur complémentaire n'agit nullement sur la valeur de ζ. On peut ainsi se limiter à l'étude des transformations unitaires (52) à déterminant égal à l'unité, c'est-à-dire pour lesquelles :

$$ad - bc = 1. \tag{55}$$

Même alors deux transformations dont les coefficients ont des signes différents nous donnent des valeurs de ξ' et η' qui diffèrent par le signe et on arrive pour ces deux transformations au même point ζ'.

Si l'on remplace dans les formules (54) ξ' et η' par leurs expressions (52) et si l'on tient compte de la condition (53), on voit, en utilisant (51), que les variables (x', y', z') s'expriment sous forme de polynômes linéaires homogènes de (x, y, z). Compte tenu de (53), les dénominateurs de (51) et de (54) sont identiques et les variables (x, y, z) subissent la même transformation linéaire que :

$$u = \bar{\xi}\eta + \xi\bar{\eta}; \qquad v = \frac{1}{i}(\bar{\xi}\eta - \xi\bar{\eta}); \qquad w = \xi\bar{\xi} - \eta\bar{\eta} \tag{56}$$

lors de la transformation unitaire (52). On trouvera, dans la suite, l'expression exacte de cette transformation linéaire.

Etablissons avant tout la forme générale des transformations unitaires (52) à déterminant égal à l'unité. Les conditions habituelles d'une transformation unitaire donnent [II-1-9] :

$$a\bar{c} + b\bar{d} = 0; \qquad c\bar{c} + d\bar{d} = 1.$$

En multipliant la condition (55) par $\bar{c}$ et en utilisant la première des conditions écrites, on obtient :

$$-bd\bar{d} - bc\bar{c} = \bar{c},$$

d'où, compte tenu de la seconde condition, $\bar{c} = -b$ ou $c = -\bar{b}$. De façon tout à fait analogue on peut montrer que $d = \bar{a}$. On peut ainsi écrire toutes les transformations unitaires à déterminant égal à l'unité, sous la forme :

$$\left.\begin{aligned} \xi' &= a\xi + b\eta, \\ \eta' &= -\bar{b}\xi + \bar{a}\eta, \end{aligned}\right\} \tag{57}$$

où a et b sont des nombres complexes quelconques qui vérifient la condition :

$$a\bar{a} + b\bar{b} = 1. \tag{58}$$

Formons maintenant les expressions (56) avec les nouvelles variables :

$$u' + iv' = 2\bar{\xi}'\eta' ; \quad u' - iv' = 2\xi'\bar{\eta}' ; \quad w' = \xi'\bar{\xi}' - \eta'\bar{\eta}',$$

ou, en tenant compte de (57) :

$$u' + iv' = \bar{a}^2 2\bar{\xi}\eta - \bar{b}^2 2\xi\bar{\eta} - 2\overline{ab}\,(\xi\bar{\xi} - \eta\bar{\eta}),$$

$$u' - iv' = -b^2 2\bar{\xi}\eta + a^2 2\bar{\xi}\eta - 2ab\,(\xi\bar{\xi} - \eta\bar{\eta}),$$

$$w' = \bar{a}b 2\bar{\xi}\eta + a\bar{b} 2\xi\bar{\eta} + (a\bar{a} - b\bar{b})(\xi\bar{\xi} - \eta\bar{\eta}).$$

En faisant le changement de variables :

$$2\bar{\xi}\eta = u + iv ; \quad 2\xi\bar{\eta} = u - iv; \quad \bar{\xi}\xi - \eta\bar{\eta} = w,$$

et ajoutant et retranchant les deux premières équations, on obtient l'expression de (u', v', w') en fonction de (u, v, w) ou, ce qui revient au même, l'expression de (x', y', z') en fonction de (x, y, z) :

$$\left.\begin{aligned} x' &= \frac{1}{2}(a^2 + \bar{a}^2 - b^2 - \bar{b}^2)\,x + \\ &\quad + \frac{i}{2}(\bar{a}^2 + \bar{b}^2 - a^2 - b^2)\,y - (ab + \bar{a}\bar{b})\,z, \\ y' &= \frac{i}{2}(a^2 + \bar{b}^2 - \bar{a}^2 - b^2)\,x + \\ &\quad + \frac{1}{2}(a^2 + \bar{a}^2 + b^2 + \bar{b}^2)\,y + i\,(\bar{a}\bar{b} - ab)\,z, \\ z' &= (a\bar{b} + \bar{a}b)\,x + i\,(a\bar{b} - \bar{a}b)\,y + (a\bar{a} - b\bar{b})\,z. \end{aligned}\right\} \tag{59}$$

A toute transformation unitaire (57) correspond une certaine transformation du plan XY, ce qui, à son tour, donne une certaine transformation de la sphère, compte tenu de la correspondance par la projection stéréographique.

En liaison avec ceci, (59) est une transformation réelle qui fait que l'équation :

$$x^2 + y^2 + z^2 = 1$$

se transforme en l'équation :

$$x'^2 + y'^2 + z'^2 = 1.$$

Mais la transformation linéaire homogène (59) ne change pas le second membre 1, et, par conséquent, elle ne doit pas changer le premier membre de l'équation, c'est-à-dire que l'on a

$$x'^2 + y'^2 + z'^2 = x^2 + y^2 + z^2.$$

Tout ceci peut être obtenu directement à partir de la forme de la transformation (59). Ainsi, les formules (59) donnent des transformations réelles orthogonales à trois variables. Montrons maintenant que le déterminant de la transformation (59) est toujours égal à (+1). Ce déterminant est une fonction continue des parties réelles et imaginaires des variables complexes a et b qui doivent vérifier la relation (58). Mais la valeur du déterminant peut seulement être égale à (+1) ou à (—1) et, compte tenu de la continuité mentionnée ci-dessus, elle doit être toujours égale à (+1) ou à (—1). Mais si $a = 1$ et $b = 0$, les formules (59) donnent la transformation identique dont le déterminant est égal à (+1), c'est-à-dire que le déterminant de la transformation (59) sera toujours égal à (+1). Ainsi les transformations linéaires (59) représentent des rotations de l'espace autour de l'origine.

Démontrons maintenant que toute rotation de l'espace peut être mise sous la forme (59). Si on pose:

$$a = e^{-\frac{i}{2}\varphi}; \quad \bar{a} = e^{\frac{i}{2}\varphi}; \quad b = \bar{b} = 0,$$

c'est-à-dire si l'on prend la matrice de la transformation unitaire

$$A_\varphi = \begin{Vmatrix} e^{-\frac{i}{2}\varphi} & 0 \\ 0 & e^{\frac{i}{2}\varphi} \end{Vmatrix}, \tag{60}$$

les formules (59) donnent alors;

$$\left.\begin{aligned} x' &= x\cos\varphi - y\sin\varphi, \\ y' &= x\sin\varphi + y\cos\varphi, \\ z' &= z, \end{aligned}\right\} \tag{61}$$

c'est-à-dire que l'on obtient une rotation d'angle φ autour de l'axe Z.

Si l'on prend maintenant:

$$a = \bar{a} = \cos\frac{\psi}{2}; \quad b = -i\sin\frac{\psi}{2}; \quad \bar{b} = i\sin\frac{\psi}{2},$$

c'est-à-dire que la matrice de la transformation unitaire est définie de la façon suivante:

$$B_\psi = \begin{Vmatrix} \cos\frac{\psi}{2} & -i\sin\frac{\psi}{2} \\ -i\sin\frac{\psi}{2} & \cos\frac{\psi}{2} \end{Vmatrix}, \tag{62}$$

les formules (59) donnent

$$\left.\begin{aligned} x' &= x, \\ y' &= y\cos\psi - z\sin\psi, \\ z' &= y\sin\psi + z\cos\psi. \end{aligned}\right\} \tag{63}$$

Ce sera une rotation d'angle ψ autour de l'axe X.

Mais on a vu [II-1-1] que toute rotation d'angles d'Euler $\{\alpha, \beta, \gamma\}$ peut être obtenue par une rotation d'angle α autour de l'axe Z suivie d'une rotation d'angle β autour du nouvel axe X et enfin d'une rotation d'angle γ autour du nouvel axe Z.

On désigne par Z_φ la matrice du troisième ordre qui correspond à la transformation (61) et par X_ψ la matrice de la transformation (63).

Une rotation d'angle α autour de l'axe Z sera réalisée par la matrice Z_α. Alors le nouvel axe X s'obtient à partir de l'ancien axe X au moyen de la même matrice, et la rotation d'angle β autour du nouvel axe X sera réalisée au moyen de la matrice $Z_\alpha X_\beta Z_\alpha^{-1}$ et les deux premières rotations au moyen de la matrice :

$$Z_\alpha X_\beta Z_\alpha^{-1} Z_\alpha = Z_\alpha X_\beta.$$

Comme ci-dessus une rotation d'angle γ autour du nouvel axe Z sera donnée par la matrice :

$$(Z_\alpha X_\beta)\, Z_\gamma\, (Z_\alpha X_\beta)^{-1},$$

et finalement la rotation $\{\alpha, \beta, \gamma\}$ sera donnée par la matrice :

$$(Z_\alpha X_\beta)\, Z_\gamma\, (Z_\alpha X_\beta)^{-1}\, (Z_\alpha X_\beta)$$

soit :

$$Z_\alpha X_\beta Z_\gamma. \tag{64}$$

Dans les raisonnements précédents on a utilisé le fait évident que si Z_φ est une matrice donnant une rotation d'angle φ autour d'un certain axe l passant par l'origine et si la matrice M transforme l'axe l en l_1, une rotation d'angle φ autour de l_1 sera définie par la matrice semblable :

$$M Z_\varphi M^{-1}.$$

Remarquons que si A_1 et A_2 sont deux transformations unitaires (57) auxquelles correspondent des transformations orthogonales (59) U_1 et U_2, au produit $A_2 A_1$ correspondra également le produit $U_2 U_1$. Ainsi une rotation de l'espace $\{\alpha, \beta, \gamma\}$ s'obtient, d'après (64), à partir d'une matrice unitaire qui est le produit de trois matrices unitaires :

$$\left\| \begin{matrix} e^{-i\frac{\alpha}{2}} & 0 \\ 0 & e^{i\frac{\alpha}{2}} \end{matrix} \right\| \cdot \left\| \begin{matrix} \cos\frac{\beta}{2} & -i\sin\frac{\beta}{2} \\ -i\sin\frac{\beta}{2} & \cos\frac{\beta}{2} \end{matrix} \right\| \cdot \left\| \begin{matrix} e^{-i\frac{\gamma}{2}} & 0 \\ 0 & e^{i\frac{\gamma}{2}} \end{matrix} \right\| . \tag{65}$$

Donc à toute transformation unitaire correspond une certaine rotation de l'espace à trois dimensions et l'on obtient ainsi toutes les rotations. Au produit de deux transformations unitaires correspondra le produit des deux rotations. On peut dire que les formules (59) définissent un homomorphisme du groupe des transformations unitaires à déterminant égal à 1 au groupe des rotations de l'espace à trois dimensions.

Considérons maintenant les transformations unitaires qui donnent la transformation identique, c'est-à-dire l'élément unité du groupe des rotations. La troisième des formules (59) donne alors:

$$a\bar{b}=0; \qquad a\bar{a}-b\bar{b}=1,$$

d'où $|a|=1$ et $b=0$. Soit $a=e^{i\vartheta}$. La première des formules (59) donne alors:

$$\frac{1}{2}(e^{i2\vartheta}+e^{-i2\vartheta})=1.$$

Il en résulte que $\vartheta=0$ ou π, c'est-à-dire $a=\pm 1$.

On obtient ainsi deux transformations unitaires dont les matrices sont:

$$E=\begin{Vmatrix}1 & 0\\ 0 & 1\end{Vmatrix}; \quad S=\begin{Vmatrix}-1 & 0\\ 0 & -1\end{Vmatrix}=-E,$$

et auxquelles correspond l'élément unité du groupe des rotations.

Supposons maintenant que deux matrices unitaires U et V donnent une seule et même rotation. De plus, $V^{-1}U$ donne la transformation identique dans le groupe des rotations de l'espace, c'est-à-dire que $V^{-1}U=E$ ou $(-E)$, de sorte que $U=V$ ou $U=-V$. Remarquons que si l'on a le signe moins devant la matrice, c'est qu'il faut changer de signe tous les éléments de la matrice. Les raisonnements précédents montrent que les transformations unitaires (57) ne donnent la même rotation de l'espace que si elles diffèrent par le signe. Inversement, si elles ne diffèrent que par le signe, comme on l'a déjà dit, il résulte de la formule (59) qu'elles donnent une seule et même rotation de l'espace. Finalement on peut dire que *le groupe des rotations de l'espace sera homomorphe au groupe des transformations unitaires* (57) *à déterminant égal à* 1 *et des rotations analogues s'obtiennent lorsque et seulement lorsque les matrices unitaires ne diffèrent que par le signe.*

Les matrices E et $(-E)$ forment le sous-groupe distingué H du groupe G des transformations unitaires (57) à déterminant égal à 1. Tout ensemble conjugué par rapport à ce sous-groupe distingué est formé de deux éléments G_1 et $(-G_1)$, où G_1 est n'importe quel élément du groupe. De ce qui a été dit ci-dessus il résulte que *le groupe des rotations est isomorphe au groupe quotient de G par H.*

Les formules (59) contiennent deux paramètres complexes a et b qui doivent vérifier la relation (58). Chacun des paramètres complexes contient deux paramètres réels:

$$a = a_1 + ia_2; \quad b = b_1 + ib_2,$$

et la relation (58) est équivalente à:

$$a_1^2 + a_2^2 + b_1^2 + b_2^2 = 1.$$

Ainsi les formules (59) contiennent quatre paramètres réels qui doivent vérifier une même relation, c'est-à-dire que les formules (59) contiennent trois paramètres réels indépendants, comme cela doit être pour le groupe des rotations. Les paramètres a et b sont les *paramètres de Cayley-Klein*. Il est facile d'obtenir leur expression en fonction des angles d'Euler. En effet, en multipliant les trois matrices unitaires (65), on obtient, comme on a vu ci-dessus, une matrice unitaire qui correspond à la rotation d'angles d'Euler $\{\alpha, \beta, \gamma\}$. Par multiplication on obtient pour les paramètres a et b les expressions suivantes:

$$a = e^{-i\frac{\alpha+\gamma}{2}} \cos\frac{\beta}{2}; \qquad b = -ie^{i\frac{\gamma-\alpha}{2}} \sin\frac{\beta}{2}. \tag{66}$$

Si on ajoute 2π à α ou à γ, a et b changent de signe, mais la rotation, en fait, reste la même. On en a déjà parlé.

III-1-13. Groupe linéaire général et groupe de Lorentz. On vient d'établir la relation qui existe entre le groupe unitaire à deux variables et le groupe des rotations de l'espace à trois dimensions. On peut établir de même la relation entre le groupe linéaire général à deux variables à déterminant égal à 1 et le groupe de Lorentz.

Introduisons quatre variables x_1, x_2, x_3, x_0. Revenant aux formules (51) qui donnent la projection stéréographique, on peut faire

$$x = \frac{x_1}{x_0}; \qquad y = \frac{x_2}{x_0}; \qquad z = \frac{x_3}{x_0}. \tag{67}$$

Ceci donne:

$$\frac{x_1}{x_0} = \frac{\bar{\xi}\eta + \xi\bar{\eta}}{\xi\bar{\xi} + \eta\bar{\eta}}; \quad \frac{x_2}{x_0} = \frac{1}{i}\,\frac{\bar{\xi}\eta - \xi\bar{\eta}}{\xi\bar{\xi} + \eta\bar{\eta}}; \quad \frac{x_3}{x_0} = \frac{\xi\bar{\xi} - \eta\bar{\eta}}{\xi\bar{\xi} + \eta\bar{\eta}}.$$

Ces formules définissent x_k à un facteur arbitraire près, et on peut écrire:

$$\left.\begin{aligned} x_0 &= \xi\bar{\xi} + \eta\bar{\eta}; \quad x_1 = \bar{\xi}\eta + \xi\bar{\eta}; \\ x_2 &= \frac{1}{i}(\bar{\xi}\eta - \xi\bar{\eta}); \quad x_3 = \xi\bar{\xi} - \eta\bar{\eta}. \end{aligned}\right\} \tag{68}$$

Les variables précédentes vérifiaient la relation (51_1), par conséquent, compte tenu de (67), les nouvelles variables définies par les formules (68), quelles que soient les valeurs complexes de ξ

et de η, vérifieront la relation

$$x_1^2+x_2^2+x_3^2-x_0^2=0. \tag{69}$$

Lorsque la transformation était unitaire pour ξ et η, l'expression $(\xi\bar{\xi}+\eta\bar{\eta})$ restait invariante, c'est-à-dire que d'après (68) la variable x_0 restait invariante. Cette variable exprime maintenant le temps et on obtient ainsi la rotation de l'espace à trois dimensions.

Abandonnons le fait que la transformation est unitaire et considérons le groupe général des transformations linéaires

$$\xi'=a\xi+b\eta;\quad \eta'=c\xi+d\eta=0. \tag{70}$$

Dans la suite nous procéderons de façon analogue au cas des transformations unitaires. On forme les expressions:

$$\left.\begin{aligned} x_1+ix_2&=2\bar{\xi}\eta; & x_1-ix_2&=2\xi\bar{\eta};\\ x_0+x_3&=2\xi\bar{\xi}; & x_0-x_3&=2\eta\bar{\eta}. \end{aligned}\right\} \tag{71}$$

Pour les nouvelles variables ξ' et η' on obtient les nouvelles valeurs de x'_k:

$$\begin{aligned} x'_1+ix'_2&=2\bar{\xi}'\eta'; & x'_1-ix'_2&=2\xi'\bar{\eta}';\\ x'_0+x'_3&=2\xi'\bar{\xi}'; & x'_0-x'_3&=2\eta'\bar{\eta}'. \end{aligned}$$

En portant les expressions (70) et (71), on a:

$$\left.\begin{aligned} x'_1+ix'_2&=\bar{a}d\,(x_1+ix_2)+\bar{b}c\,(x_1-ix_2)+\\ &\quad+\bar{a}c\,(x_0+x_3)+\bar{b}d\,(x_0-x_3),\\ x'_1-ix'_2&=b\bar{c}\,(x_1+ix_2)+a\bar{d}\,(x_1-ix_2)+\\ &\quad+a\bar{c}\,(x_0+x_3)+b\bar{d}\,(x_0-x_3),\\ x'_0+x'_3&=\bar{a}b\,(x_1+ix_2)+a\bar{b}\,(x_1-ix_2)+\\ &\quad+a\bar{a}\,(x_0+x_3)+b\bar{b}\,(x_0-x_3),\\ x'_0-x'_3&=\bar{c}d\,(x_1+ix_2)+c\bar{d}\,(x_1-ix_2)+\\ &\quad+c\bar{c}\,(x_0+x_3)+d\bar{d}\,(x_0-x_3), \end{aligned}\right\} \tag{72}$$

d'où l'on obtient les expressions linéaires à coefficients réels de x'_k en fonction de x_k que nous n'écrirons pas. On notera simplement que si l'on ajoute les deux dernières équations (72), dans l'expression de x'_0 le coefficient de x_0 sera égal à $\frac{1}{2}\,(a\bar{a}+b\bar{b}+c\bar{c}+d\bar{d})$, c'est-à-dire positif.

Les nouvelles variables x'_k, comme les variables initiales, vérifient l'équation

$$x_1'^2 + x_2'^2 + x_3'^2 - x_0'^2 = 0. \tag{73}$$

Si dans le premier membre de cette équation on remplace les x'_k par leurs expressions en fonction de x_k, on doit obtenir l'équation (69). Mais le premier membre de l'équation (73) peut alors différer du premier membre de l'équation (69) par un facteur constant, c'est-à-dire que dans ce cas on a

$$x_1'^2 + x_2'^2 + x_3'^2 - x_0'^2 = k\,(x_1^2 + x_2^2 + x_3^2 - x_0^2),$$

où k est une certaine constante. En utilisant les formules précédentes et en tenant compte de ce que

$$x_1'^2 + x_2'^2 + x_3'^2 - x_0'^2 = (x'_1 + ix'_2)\,(x'_1 - ix'_2) - (x'_0 + x'_3)\,(x'_0 - x'_3),$$

il est facile de montrer que $k = (ad - bc)\,(\bar{a}\bar{d} - \bar{b}\bar{c}) = |\,ad - bc\,|^2$. Si l'on veut obtenir $k = 1$, c'est-à-dire la transformation de Lorentz :

$$x_1'^2 + x_2'^2 + x_3'^2 - x_0'^2 = x_1^2 + x_2^2 + x_3^2 - x_0^2, \tag{74}$$

il faut prendre les transformations linéaires (70) avec un déterminant de module égal à l'unité, c'est-à-dire qui s'exprime par un nombre de la forme $e^{i\varphi}$. En multipliant, comme on l'a déjà fait, tous les coefficients de la transformation (70) par $e^{-i\frac{\varphi}{2}}$, on ne change pas les valeurs de x'_1, x'_2, x'_3 définies par les formules (68), si on remplace ξ et η par ξ' et η', car ces formules contiennent le produit de l'une des quantités (ξ', η') par l'une des quantités $(\bar{\xi}', \bar{\eta}')$ et d'autre part on rend le déterminant égal à l'unité.

Ainsi considérons les transformations (70) à déterminant égal à l'unité:

$$ad - bc = 1. \tag{75}$$

Comme dans le paragraphe précédent on peut montrer que la transformation linéaire qui exprime x'_k en fonction de x_k possède un déterminant égal à $(+1)$. Rappelons de plus que le coefficient de x_0 dans l'expression de x'_0 est positif, c'est-à-dire que cette transformation a un déterminant égal à $(+1)$ et ne change pas le sens du temps, c'est-à-dire que les transformations (72) sont des transformations de Lorentz positives.

Ainsi les transformations linéaires conformes à la condition (75) donnent les transformations de Lorentz positives déjà définies au paragraphe [III-1-3].

Comme au paragraphe précédent, cherchons maintenant s'il est Possible d'obtenir à partir des formules (72) n'importe quelle transformation de Lorentz positive. Remarquons avant tout que comme au paragraphe précédent, le produit de deux transformations linéaires (70) correspond au produit des transformations de Lorentz correspondantes, c'est-à-dire, d'une façon plus précise, que si A et B sont deux transformations linéaires (70) donnant, d'après (72), les transformations de Lorentz T_1 et T_2, la transformation de Lorentz T_2T_1 correspondra à la transformation linéaire BA. Comme on l'a vu [III-1-3], toute transformation de Lorentz positive peut être mise sous la forme:

$$T = VSU,$$

où U et V sont de simples rotations d'un espace à trois dimensions et S une transformation de Lorentz positive à deux variables. D'après les résultats du paragraphe précédent on peut obtenir toute rotation au moyen d'une transformation unitaire de la forme (70), à déterminant égal à l'unité. Ainsi il nous reste à montrer que l'on peut obtenir toute transformation de Lorentz positive S à deux variables d'après les formules (72) pour un choix correspondant de la transformation linéaire (70). En comparant (74) avec (21) de [III-1-3] on voit que l'on a maintenant $c = 1$, de sorte que les formules (17), qui donnent les transformations de Lorentz positives à deux variables, se mettent sous la forme:

$$\left.\begin{aligned} x_3' &= \frac{-vx_0 + x_3}{\sqrt{1-v^2}}; \quad & x_0' &= \frac{x_0 - vx_3}{\sqrt{1-v^2}}; \\ x_1' &= x_1; & x_2' &= x_2. \end{aligned}\right\} \qquad (76)$$

Introduisons la quantité:

$$u = \frac{1}{\sqrt{1-v^2}} > 1,$$

et considérons une transformation linéaire (70) de la forme:

$$\xi' = l\xi; \qquad \eta' = \frac{1}{l}\eta,$$

où l est une constante réelle. Son déterminant est évidemment égal à l'unité. Dans ce cas $a = l$, $d = \frac{1}{l}$ et $b = c = 0$. En reportant ceci dans les formules (72), on obtient justement la transformation (76) si l vérifie les conditions:

$$\frac{l^2}{2} + \frac{1}{2l^2} = u; \qquad \frac{l^2}{2} - \frac{1}{2l^2} = -vu.$$

Ceci nous donne directement $l^2 = u \pm \sqrt{u^2 - 1}$. La deuxième condition montre que si v est positif, il faut prendre la racine de l^2 qui est plus petite que l'unité et si v est négatif, la racine plus grande que l'unité et cette deuxième condition est alors vérifiée. En prenant la racine on obtient pour l deux valeurs qui ne diffèrent que par le signe. On peut ainsi finalement affirmer que *le groupe des transformations linéaires* (70) *à déterminant égal à l'unité est homomorphe au groupe des transformations de Lorentz positives et que cet homomorphisme est réalisé par les formules* (72). Comme au paragraphe précédent, cet homomorphisme ne sera pas un isomorphisme, c'est-à-dire que différentes transformations (70) peuvent conduire à une seule et même transformation de Lorentz. Des formules (72) il résulte directement que la transformation identique dans le groupe de Lorentz s'obtient à partir des deux transformations linéaires de matrices:

$$E=\left\|\begin{matrix}1 & 0\\ 0 & 1\end{matrix}\right\|; \quad S=\left\|\begin{matrix}-1 & 0\\ 0 & -1\end{matrix}\right\|=-E,$$

et, de même que dans le paragraphe précédent, on peut montrer que *toute transformation du groupe de Lorentz ne peut être obtenue qu'à l'aide de deux transformations linéaires* (70), *dont les coefficients ne diffèrent que par le signe.*

De même qu'au paragraphe [III-1-12], les éléments E et $(-E)$ forment un sous-groupe distingué H du groupe des transformations linéaires à déterminant égal à 1 et *le groupe des transformations de Lorentz positives est isomorphe au groupe quotient de G par H.*

Les transformations linéaires (70) contiennent quatre coefficients complexes liés par la condition (75). Ainsi, les formules (72) contiennent trois paramètres complexes arbitraires ou, autrement dit, six paramètres réels arbitraires.

III-2. Représentations linéaires des groupes

III-2-1. Représentation des groupes au moyen de transformations linéaires. Soit G un groupe dont les éléments sont G_α. On suppose qu'à chaque élément G_α est associée une matrice définie A_α et que toutes les matrices A_α sont du même ordre et à déterminants non nuls. On suppose de plus que cette correspondance est telle qu'à tout produit $G_{\alpha_2}G_{\alpha_1}$ corresponde la matrice $A_{\alpha_2}A_{\alpha_1}$, produit de A_{α_2} et A_{α_1}. On dit alors que *les matrices A_α ou bien les transformations linéaires correspondantes donnent une représentation linéaire du groupe G.* Soient G_0 l'élément unité du groupe et A_0 la matrice correspondante. Comme $G_0G_\alpha = G_\alpha$, on doit avoir $A_0A_\alpha = A_\alpha$, d'où, en multipliant à droite par A_α^{-1}, on a $A_0 = I$, c'est-à-dire qu'à l'élément unité doit correspondre la matrice unité. Soient G_{α_1} et G_{α_2} deux éléments inverses et A_{α_1} et A_{α_2} les matrices correspondantes. De l'égalité $G_{\alpha_2}\cdot G_{\alpha_1} = G_0$ il résulte que $A_{\alpha_2}A_{\alpha_1} = I$, c'est-à-dire qu'à des éléments inverses correspondent des matrices inverses. Il découle de ce qui précède que *les matrices A_α* (ou les transformations linéaires correspondantes) *forment un groupe* A, *homomorphe au groupe* G. Si aux divers éléments de G correspondent des matrices différentes, A sera non seulement homomorphe mais isomorphe à G. Dans ce cas on dit que l'on a *une représentation linéaire biunivoque du groupe* G.

S'il n'en est pas ainsi, l'ensemble des éléments de G auxquels correspond la matrice unité de A forme un sous-groupe distingué du groupe G, et le groupe A sera isomorphe au groupe quotient du sous-groupe distingué [III-1-6].

Si le groupe fondamental G est lui-même un groupe des transformations linéaires, on voit qu'il donne l'une de ses représentations linéaires possibles.

Faisons une remarque au sujet de la définition donnée des représentations linéaires.

Supposons que l'on sache qu'à tout élément G_α correspond une matrice définie A_α et que le produit des éléments correspond au produit des matrices, mais que l'on ignore si les déterminants des matrices A_α ne sont pas nuls. Montrons que si l'un des déterminants $D(A_{\alpha_0})$ est nul, tous les déterminants $D(A_\alpha)$ sont nuls. En effet, l'ensemble des matrices $A_{\alpha_0}A_\alpha$, lorsque α varie, contient toutes les matrices qui correspondent aux éléments du groupe [III-1-5]. Mais $D(A_{\alpha_0}A_\alpha) = D(A_{\alpha_0})D(A_\alpha)$ et le produit est nul car le premier facteur par hypothèse est nul. Ainsi, ayant la loi de correspondance qui fait qu'à un produit correspond un produit, il nous suffit de vérifier que l'un des déterminants $D(A_\alpha)$ n'est pas nul; par exemple, il suffit de vérifier qu'à l'élément unité de G correspond la matrice unité de A.

Soit X une matrice du même ordre que les matrices A_α à déterminant non nul. On a:

$$(XA_{\alpha_2}X^{-1})(XA_{\alpha_1}X^{-1}) = XA_{\alpha_2}A_{\alpha_1}X^{-1},$$

et, par conséquent, les matrices $XA_\alpha X^{-1}$ donnent aussi une représentation linéaire du groupe G. On dit que *ces deux représentations semblables sont équivalentes*. Supposons que l'ordre des matrices A_α soit n et que $(x_1, \ldots, x_n)$ soient les composantes d'un vecteur de l'espace à n dimensions sur lequel agissent les transformations A_α, de sorte que le groupe A soit:

$$\mathbf{x}' = A_\alpha \mathbf{x}. \tag{77}$$

La représentation linéaire équivalente:

$$\mathbf{y}' = XA_\alpha X^{-1}\mathbf{y}, \tag{78}$$

comme on le sait [II-1-6], signifie que, dans l'espace considéré, on a introduit de nouveaux axes de coordonnées et que les nouvelles composantes s'expriment en fonction des anciennes au moyen des formules:

$$(y_1, \ldots, y_n) = X(x_1, \ldots, x_n). \tag{79}$$

Avec ces nouveaux axes, les transformations linéaires de l'espace s'expriment au moyen des formules (78), c'est-à-dire que les représentations linéaires équivalentes peuvent être obtenues par simple changement d'axes de coordonnées d'après les formules (79). Les variables $(x_1, \ldots, x_n)$ qui figurent dans les formules (77) sont les *êtres de la représentation linéaire*. Le passage à la représentation linéaire équivalente revient donc à un changement des êtres des représentations linéaires par d'autres au moyen d'une certaine transformation linéaire (79) à déterminant non nul.

Soit une représentation linéaire du groupe G au moyen des matrices A_α d'ordre n et une autre représentation linéaire du même groupe au moyen des matrices B_α d'ordre m. On forme les matrices quasi diagonales d'ordre $(n + m)$:

$$[A_\alpha, B_\alpha] = \begin{Vmatrix} A_\alpha & 0 \\ 0 & B_\alpha \end{Vmatrix}. \tag{80}$$

D'après la règle de multiplication des matrices quasi diagonales on a:

$$[A_{\alpha_2}, B_{\alpha_2}][A_{\alpha_1}, B_{\alpha_1}] = [A_{\alpha_2}A_{\alpha_1}, B_{\alpha_2}B_{\alpha_1}].$$

Ainsi, les matrices (80) donnent aussi une représentation linéaire du groupe G. En général, ayant plusieurs représentations du groupe G au moyen des matrices $A_\alpha, B_\alpha, C_\alpha$, on peut former une nouvelle représentation en utilisant les matrices quasi diagonales:

$$D_\alpha = [A_\alpha, B_\alpha, C_\alpha] = \begin{Vmatrix} A_\alpha & 0 & 0 \\ 0 & B_\alpha & 0 \\ 0 & 0 & C_\alpha \end{Vmatrix}. \tag{81}$$

Remarquons maintenant que si l'on passe à une représentation équivalente à l'aide des matrices $XD_\alpha X^{-1}$, le caractère quasi diagonal des matrices n'est en général pas conservé et, d'après l'aspect de cette nouvelle représentation, il ne sera plus possible de dire immédiatement si elle est formeé d'autres représentations à un nombre de dimensions moindre selon la loi (81) à une représentation équivalente près. Si la représentation linéaire D_α a une forme purement quasi diagonale (80), elle se décompose en plusieurs représentations linéaires A_α et B_α à un nombre moindre de dimensions, c'est-à-dire à matrices d'ordre plus petit. Dans ce cas la représentation linéaire est dite *réduite*. Si une représentation linéaire E_α n'est pas quasi diagonale, mais si une représentation qui lui est équivalente $XE_\alpha X^{-1}$ est quasi diagonale, la représentation E_α est dite *réductible*. Enfin, si non seulement la représentation elle-même mais toutes les représentations qui lui sont équivalentes ne sont pas quasi diagonales, c'est-à-dire réductibles, on dit que cette représentation est *irréductible*.

Indiquons quelques conditions qui font que l'on peut affirmer que la représentation est réductible. Soit la représentation linéaire formée des matrices A_α d'ordre n qui donnent des transformations linéaires par rapport aux variables $(x_1, x_2, \ldots, x_n)$. On suppose que toutes les matrices A_α sont unitaires et que le sous-espace R' engendré par les k premiers vecteurs de base se transforme en lui-même au moyen de la transformation A_α, c'est-à-dire que pour $x_{k+1} = x_{k+2} = \ldots = x_n = 0$, on a $x'_{k+1} = x'_{k+2} = \ldots = x'_n = 0$. Autrement dit, toutes les matrices A_α sont de la forme :

$$\left\| \begin{matrix} A'_\alpha & N_\alpha \\ 0 & A''_\alpha \end{matrix} \right\|, \tag{82}$$

où A'_α est une matrice d'ordre k, A''_α une matrice d'ordre $(n - k)$ et dans le coin en bas à gauche on a $(n - k)$ lignes et k colonnes formées de zéros. Examinons l'espace R'', engendré par les $(n - k)$ derniers vecteurs de base. Il sera constitué de vecteurs orthogonaux à tous les vecteurs du sous-espace R' déjà mentionné. Comme chaque transformation A_α transforme le sous-espace R' en lui-même et, compte tenu de ce que cette transformation est unitaire et conserve l'orthogonalité des vecteurs, tout vecteur du sous-espace R'' doit après la transformation A_α se transformer en un vecteur appartenant aussi à ce sous-espace. Autrement dit, pour $x_1 = \ldots = x_k = 0$, on a $x'_1 = \ldots = x'_k = 0$. Il en résulte que dans les matrices (82) les éléments situés dans le coin en haut à droit, constitué de k lignes et $(n - k)$ colonnes, doivent aussi être tous nuls, c'est-à-dire que les matrices de la représentation linéaire considérée seront :

$$\left\| \begin{matrix} A'_\alpha & 0 \\ 0 & A''_\alpha \end{matrix} \right\| = [A'_\alpha, A''_\alpha],$$

et que, par conséquent, la transformation sera réduite. Supposons maintenant que toutes les transformations unitaires A_α laissent invariant un certain sous-espace R_1 à k dimensions $(k < n)$, où n est l'ordre des matrices A_α. On transforme les axes de coordonnées de façon que le sous-espace R_1 soit engendré par les k premiers vecteurs de base, ce qui revient à passer à la représentation linéaire équivalente et peut être effectué par transformation unitaire. Après une telle transformation, compte tenu de ce qui précède, on obtient une représentation réduite. On a ainsi le théorème suivant.

Théorème. *Si une représentation linéaire d'un groupe est formée de matrices unitaires et si ces matrices laissent invariant un certain sous-espace, cette représentation est une représentation réductible.*

La question de la réductibilité ou de l'irréductibilité des représentations est étroitement liée à celle de la transformation des matrices A_α en matrices semblables $XA_\alpha X^{-1}$. On notera quelques cas particuliers de passage aux représentations équivalentes obtenues pour un choix particulier de la matrice X. Construisons une matrice X de manière que le nombre un soit le deuxième élément de la première ligne et le nombre zéro les autres; le nombre un soit le premier élément de la deuxième ligne et le nombre zéro les autres éléments; à partir de la troisième ligne la diagonale principale est constituée de nombre un et les autres éléments sont des zéros, c'est-à-dire :

$$X = \begin{Vmatrix} 0 & 1 & 0 & 0 & \ldots & 0 \\ 1 & 0 & 0 & 0 & \ldots & 0 \\ 0 & 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 0 & 1 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot\cdot\cdot & \cdot \\ 0 & 0 & 0 & 0 & \ldots & 1 \end{Vmatrix}.$$

En développant par rapport à la dernière ligne on voit que $D(X) = -1$. En appliquant les règles habituelles de multiplication des matrices, il est facile de vérifier que si Y est une matrice, la matrice semblable XYX^{-1} s'obtient à partir de Y en permutant la première et la deuxième ligne, la première et la deuxième colonne. De même toute permutation de lignes accompagnée d'une permutation correspondante de colonnes est équivalente au passage à une certaine matrice semblable par une transformation X qui ne dépend pas de la matrice Y. Donc, *une même permutation de lignes et de colonnes dans toutes les matrices A_α qui donnent une certaine représentation linéaire d'un groupe équivaut à passer à la représentation équivalente.*

Si l'on peut répartir les entiers 1, 2, . . ., n en deux classes, de sorte qu'à l'intersection de toute ligne de chaque matrice A_α, de numéro appartenant à l'une des classes, et de n'importe quelle colonne, de numéro appartenant à l'autre classe, on ait des zéros, une telle représentation est réductible. En effet, pour en effectuer la réduction, c'est-à-dire pour la ramener à la forme réduite, il suffit de permuter lignes et colonnes de façon que les lignes et les colonnes de l'une des classes soient en haut et à gauche, tandis que les lignes et les colonnes de l'autre classe soient en bas et à droite.

En conclusion à ce paragraphe notons encore le cas d'une représentation linéaire du groupe G d'ordre 1, c'est-à-dire où toutes les matrices A_α sont des matrices du premier ordre, autrement dit des nombres ordinaires. Dans ce cas, à chaque élément G_α du groupe correspond une transformation $x' = m_\alpha x$ ou, plus simplement, le nombre m_α, et au produit G_2G_1 correspond le produit ordinaire des nombres m_2m_1.

III-2-2. Théorèmes fondamentaux. Soit un groupe fini G contenant m éléments $G_1, \ldots, G_m$ et soient $A_1, \ldots, A_m$ des matrices d'un certain ordre n qui donnent une représentation linéaire de ce groupe. On désigne par $\mathbf{x}\,(x_1, \ldots, x_n)$ les êtres de cette représentation. Soit l'expression:

$$\varphi\,(x_1, \ldots, x_n) = \sum_{s=1}^{m} |A_s\mathbf{x}|^2. \tag{83}$$

Sous forme développée on a:

$$\varphi = \sum_{s=1}^{m} \sum_{i=1}^{n} (a_{i1}^{(s)}x_1 + \ldots + a_{in}^{(s)}x_n)\,(\overline{a}_{i1}^{(s)}\overline{x}_1 + \ldots + \overline{a}_{in}^{(s)}\overline{x}_n), \tag{84}$$

où $a_{ik}^{(s)}$ désigne les éléments de la matrice A_s. Il est facile de vérifier que l'expression (84) est une forme d'Hermite, c'est-à-dire que les coefficients de $\overline{x}_px_q$ et $x_p\overline{x}_q$ sont des nombres complexes conjugués. De plus, il résulte de (83) que cette forme d'Hermite représente la somme des carrés des longueurs de certains vecteurs, c'est-à-dire que c'est une forme d'Hermite définie positive [II-2-9]. Autrement dit, en effectuant une certaine transformation unitaire:

$$\mathbf{y} = U\mathbf{x},$$

qui transforme cette forme en somme de carrés:

$$\varphi = \sum_{j=1}^{n} \lambda_j \overline{y}_j y_j,$$

on a tous les coefficients λ_j positifs. En faisant encore la transformation $z_j = \sqrt{\lambda_j}y_j$, la forme d'Hermite φ dans les nouvelles variables sera une somme de carrés parfaits:

$$\varphi = \overline{z}_1z_1 + \ldots + \overline{z}_nz_n. \tag{85}$$

On effectue sur les variables $(x, \ldots, x_n)$ la transformation:

$$\mathbf{x}' = A_k\mathbf{x}, \tag{86}$$

qui appartient à la représentation linéaire du groupe. Il est facile de voir qu'alors la forme d'Hermite φ ne change pas. En effet,

$$\varphi\,(x_1', \ldots, x_n') = \sum_{s=1}^{m} |A_sA_k\mathbf{x}|^2.$$

Mais on sait [III-1-5] que l'ensemble des transformations (matrices)

$$A_1A_k,\ A_2A_k,\ \ldots,\ A_mA_k$$

coïncide avec l'ensemble des matrices:

$$A_1,\ A_2,\ \ldots,\ A_m;$$

c'est pourquoi si l'on exprime la transformation (86) au moyen des nouvelles variables $(z_1, \ldots, z_n)$ liées aux anciennes par les formules de la forme:

$$(z_1, \ldots, z_n) = B_0(x_1, \ldots, x_n),$$

où B_0 est une certaine matrice, alors au lieu du groupe A_k on obtient le groupe des matrices semblables $B_0A_kB_0^{-1}$ et toutes les transformations de ce groupe ne changent pas l'expression (85), c'est-à-dire qu'elles ne changent pas la somme des carrés des modules et par suite sont toutes des transformations unitaires. On a ainsi montré, dans le cas de groupes finis, que toute représentation linéaire est équivalente à une certaine représentation unitaire, c'est-à-dire à une représentation formée de transformations unitaires. Si on ajoute certaines conditions, cette propriété se conserve même pour les représentations linéaires de groupes infinis dépendant de paramètres (et dans la suite, lorsqu'on parlera de représentation linéaire d'un groupe, on sous-entendra toujours représentation unitaire). On a ainsi le théorème suivant:

Théorème I. *Toute représentation linéaire d'un groupe (fini) a une représentation unitaire équivalente.*

Cherchons maintenant la condition nécessaire et suffisante pour qu'une représentation linéaire soit réductible. Introduisons auparavant un nouveau terme: on appelle la matrice diagonale $[k, \ldots, k]$ contenant sur sa diagonale des éléments identiques matrice *multiple* de la matrice unité. *Une telle matrice peut être représentée sous la forme kI.* On a déjà vu que du point de vue des opérations algébriques elle est équivalente au nombre k.

Supposons maintenant que l'on ait une représentation linéaire réductible d'un groupe. Une telle représentation est réalisée par exemple au moyen d'une matrice de la forme:

$$D_a = X[A_a,\ B_a,\ C_a]\,X^{-1},$$

où X est une matrice et où la matrice entre crochets est quasi diagonale. On forme la matrice:

$$Y = X[kI,\ lI,\ mI]\,X^{-1},$$

où le terme central, quasi diagonal, a la même construction que dans les matrices D_a. Il est aisé de voir que la matrice Y commute avec toutes les matrices D_a. En effet,

$$D_a Y = X[A_a k, B_a l, C_a m] X^{-1},$$

et de même :

$$Y D_a = X[kA_a, lB_a, mC_a] X^{-1}.$$

Mais dans la multiplication de toute matrice par un nombre l'ordre des facteurs ne joue aucun rôle. De plus, si les nombres k, l et m sont différents, ce que l'on suppose, la matrice Y n'est pas un multiple de la matrice unité. En effet, elle a de façon évidente des valeurs propres différentes k, l et m. On a donc le théorème suivant :

Théorème II. *Si une représentation linéaire est réductible, il existe alors une matrice qui n'est pas un multiple de la matrice unité et qui commute avec toutes les matrices appartenant à cette représentation linéaire réductible.* On peut montrer que la réciproque est vraie, c'est-à-dire que :

Théorème III. *S'il existe une matrice Y qui n'est pas un multiple de la matrice unité et qui commute avec toutes les matrices D_a d'une représentation linéaire, cette représentation linéaire est alors réductible.*

Ainsi, par hypothèse, on a pour tout indice a :

$$D_a Y = Y D_a. \tag{87}$$

Soit Z une matrice dont le déterminant n'est pas nul et telle que toutes les matrices ZD_aZ^{-1} sont unitaires: $ZD_aZ^{-1} = U_a$. On récrira l'égalité précédente sous la forme :

$$Z^{-1}U_aZY = YZ^{-1}U_aZ.$$

En multipliant à gauche par Z et à droite par Z^{-1}, on obtient :

$$U_a(ZYZ^{-1}) = (ZYZ^{-1})U_a,$$

c'est-à-dire que la matrice ZYZ^{-1} commute avec toutes les matrices de la représentation unitaire. Cette matrice n'est évidemment pas un multiple de la matrice unité, car si $ZYZ^{-1} = kI$, alors $Y = kI$. Il suffit de démontrer que la représentation linéaire équivalente U_a est réductible. Ainsi, la démonstration du théorème se réduit au cas d'une transformation linéaire unitaire. Pour simplifier l'écriture supposons que la représentation linéaire formée des matrices D_a est unitaire.

Soit λ_1 une valeur propre de la matrice Y. La matrice $\lambda_1 = \lambda_1 I$ commute [II-1-6] avec toute matrice et, par suite, la matrice $Y - \lambda_1 I$ de même que Y vérifie les conditions (87), c'est-à-dire qu'elle commute avec toutes les matrices D_a. Il est aisé de voir qu'au moins l'une des valeurs propres de la matrice $Y_1 = Y - \lambda_1 I$ est nulle. En effet, l'équation caractéristique de la matrice Y_1 est:

$$D(Y_1-\lambda I)=D[Y-(\lambda+\lambda_1)I]=0,$$

c'est-à-dire qu'on l'obtient à partir de l'équation caractéristique de Y en remplaçant λ par $(\lambda+\lambda_1)$ et comme parmi les valeurs propres de la matrice Y il y avait la valeur propre λ_1, alors parmi les valeurs propres de Y_1 il y en aura au moins une qui sera nulle. Par conséquent, le déterminant de la matrice Y_1, qui est égal au produit de ses valeurs propres, sera aussi nul. On peut ainsi, en démontrant le théorème, supposer que toutes les matrices D_a sont unitaires et que le déterminant de la matrice Y qui figure dans la formule (87) est nul.

Considérons l'ensemble des vecteurs de composantes:

$$\left.\begin{aligned} x_1 &= y_{11}u_1+y_{12}u_2+\ldots+y_{1n}u_n, \\ x_2 &= y_{21}u_1+y_{22}u_2+\ldots+y_{2n}u_n, \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ x_n &= y_{n1}u_1+y_{n2}u_2+\ldots+y_{nn}u_n, \end{aligned}\right\} \tag{88}$$

où les u_s prennent n'importe quelle valeur et où les y_{ik} sont les éléments de la matrice Y. Comme le déterminant de la matrice Y est nul, le rang du tableau $\| y_{ik} \|$ est inférieur à n. Supposons qu'il soit égal à un nombre $r<n$. Dans ce cas, on sait [I-2-9] que les formules (88) définissent un sous-espace R' de dimension r.

Considérons le premier membre de l'égalité:

$$D_aY\mathbf{u}=YD_a\mathbf{u}. \tag{89}$$

Le vecteur $Y\mathbf{u}$ a les composantes (88) et par suite $D_aY\mathbf{u}$ est le résultat de l'application de la transformation D_a à un vecteur arbitraire du sous-espace R'. Dans le second membre de la formule (89) on a le résultat de l'application de la transformation Y au vecteur $D_a\mathbf{u}$, c'est-à-dire que les composantes du second membre s'expriment aussi au moyen des formules (88), où simplement $u_1, \ldots, u_n$ sont remplacés par les composantes du vecteur $D_a\mathbf{u}$, c'est-à-dire que le second membre de la formule (89) représente un vecteur appartenant au sous-espace R'. Ainsi, en comparant le premier membre au second, nous voyons que l'application de la transformation D_a à tout vecteur du sous-espace R' donne un vecteur appartenant aussi à ce sous-espace. Mais on a vu [III-2-1] que si des transformations unitaires laissent invariant un sous-espace, elles forment alors une représentation réductible. Ainsi le théorème est démontré.

Les théorèmes II et III montrent que *la condition nécessaire et suffisante pour qu'une représentation linéaire soit irréductible est qu'il n'existe pas de matrice dont la forme soit différente de kI et qui commute avec toutes les matrices de la représentation linéaire.*

Il résulte directement du théorème I que dans le théorème de [III-2-1] il n'est pas nécessaire de mentionner que la représentation est unitaire et on peut affirmer que si toutes les matrices d'une représentation laissent invariant un sous-espace, alors une telle représentation est réductible. La proposition réciproque est évidente.

III-2-3. Groupes abéliens et représentation du premier ordre. Un groupe G est un groupe *abélien* si tous ses éléments commutent deux à deux, c'est-à-dire si pour tout indice on a $G_{a_2}G_{a_1} = G_{a_1}G_{a_2}$ [III-1-5]. Soient A_{a_1} et A_{a_2} des matrices correspondant à G_{a_1} et G_{a_2} dans une représentation linéaire. Le produit de A_{a_2} et de A_{a_1} correspond à $G_{a_2}G_{a_1}$ et le produit de A_{a_1} et de A_{a_2} à $G_{a_1}G_{a_2}$. Mais ces produits coïncident et, par suite, on doit avoir:

$$A_{a_2}A_{a_1} = A_{a_1}A_{a_2},$$

c'est-à-dire que *les matrices formant une représentation linéaire d'un groupe abélien commutent deux à deux.*

Supposons la représentation unitaire, c'est-à-dire que toutes les matrices soient unitaires. Alors on sait qu'il existe une transformation unitaire U telle que toutes les matrices UA_aU^{-1} sont diagonales [II-2-11], c'est-à-dire que dans ce cas une représentation linéaire équivalente est formée de matrices diagonales:

$$UA_aU^{-1} = [k_1^{(a)}, \ldots, k_n^{(a)}].$$

On voit ainsi que dans ce cas la représentation linéaire se décompose en n représentations du premier ordre:

$$B_a^{(s)} = k_s^{(a)} \qquad (s = 1, 2, \ldots, n).$$

Ainsi, *toute représentation unitaire d'un groupe abélien est équivalente à un ensemble de représentations du premier ordre et le passage à la représentation équivalente s'effectue au moyen d'une transformation unitaire.*

Considérons maintenant une série d'exemples de représentation de groupes abéliens et quelques exemples de représentations linéaires du premier ordre de groupes non abéliens.

E x e m p l e 1. Considérons le groupe cyclique (abélien) d'ordre m formé des éléments:

$$S^0 = I, \; S, \; S^2, \; \ldots, \; S^{m-1} \quad (S^m = I). \tag{90}$$

Si à l'élément S correspond la transformation linéaire $x' = \omega x$ ou, ce qui revient au même, le nombre ω, aux éléments de (90) correspondront les nombres:

$$1, \omega, \omega^2, \ldots, \omega^{m-1}.$$

Puisque $S^m = I$, on doit avoir $\omega^m = 1$, c'est-à-dire:

$$\omega = e^{\frac{2k\pi i}{m}},$$

où k est un entier que l'on peut prendre égal à l'un des nombres $0, 1, 2, \ldots, m-1$.

Voyons plus en détail le cas où $m = 2$. Alors:

$$I, \ S \text{ et } S^2 = I,$$

c'est-à-dire que $S = S^{-1}$. Si $k = 0$, aux deux éléments I et S correspond une même transformation identique $x' = x$ ou le nombre 1; si $k = 1$, à l'élément I correspond la transformation $x' = x$ et à l'élément S la transformation $x' = -x$ ou, plus simplement, à l'élément I le nombre 1 et à l'élément S le nombre (-1). Dans les applications physiques on rencontre un cas important, celui d'un groupe formé de la transformation identique de l'espace à trois dimensions et la transformation de symétrie par rapport à l'origine:

$$x' = -x; \quad y' = -y; \quad z' = -z \, (S).$$

Il est évident que l'on a $m = 2$. Les deux représentations indiquées ci-dessus peuvent être appelées représentation identiqué et alternée de la symétrie par rapport à l'origine.

Exemple 2. Soit le groupe des rotations autour de l'axe Z. Les matrices de ce groupe sont de la forme:

$$Z_\varphi = \begin{Vmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{Vmatrix} \tag{91}$$

et de plus, on l'a déjà vu, elles vérifient la relation évidente:

$$Z_{\varphi_2} Z_{\varphi_1} = Z_{\varphi_1} Z_{\varphi_2} = Z_{\varphi_1 + \varphi_2}.$$

La fonction $e^{l\varphi}$ vérifie la même relation.

Mais il faut remarquer que si $\varphi = 2\pi$, la rotation équivaut à la transformation identique et, par suite, on doit avoir $e^{2\pi l} = 1$, c'est-à-dire que le nombre l doit être de la forme $l = ki$, où k est un entier quelconque. On a ainsi une infinité de représentations linéaires du groupe des rotations et à la matrice (91) correspond le nombre $e^{\varphi ki}$.

En donnant à k toutes les valeurs possibles:

$$k = 0, \ \pm 1, \ \pm 2, \ \ldots,$$

on obtient une infinité de représentations linéaires du groupe des rotations.

Exemple 3. Soit maintenant le groupe formé des $n!$ permutations de n éléments. On peut associer à chaque permutation le nombre $(+1)$, de sorte que l'on obtienne une représentation symétrique du groupe des permutations. Autrement dit, on peut associer à chaque permutation de la classe I, formée d'un nombre pair de transpositions, le nombre $(+1)$, et à chaque permutation de la classe II le nombre (-1). Ainsi, on obtient la représentation antisymétrique du groupe des permutations. Dans cette représentation, à toute permutation du sous-groupe alterné correspond le nombre $(+1)$ et aux autres permutations le nombre (-1). Ainsi, on peut montrer que ces deux cas épuisent toutes les possibilités de représentations linéaires du premier ordre pour le groupe des permutations. Ce groupe a d'autres représentations d'ordre supérieur à 1.

Exemple 4. Considérons maintenant le groupe de toutes les transformations réelles orthogonales dans le plan, c'est-à-dire le groupe formé des rotations du plan autour de l'origine jointes à la symétrie par rapport à l'axe Y. On a déjà vu [III-1-1] que les matrices de ce groupe sont de la forme:

$$\{\varphi, d\} = \left\| \begin{matrix} d\cos\varphi & -d\sin\varphi \\ \sin\varphi & \cos\varphi \end{matrix} \right\|, \tag{92}$$

où $d = 1$ pour une rotation pure et $d = -1$ pour une rotation associée à une symétrie. En plus de la représentation linéaire du premier ordre pour laquelle à chaque matrice (92) correspond le nombre $(+1)$, on peut construire une représentation linéaire du premier ordre dans laquelle on associe à la matrice (92) le nombre $(+1)$ si $d = 1$ et le nombre (-1) si $d = -1$. Ceci donne effectivement une représentation linéaire car le produit de deux matrices de la forme (92) correspond à une rotation pure si d a le même signe et à une rotation avec une symétrie si d a des signes différents dans les matrices que l'on multiplie.

III-2-4. Représentations linéaires d'un groupe unitaire à deux variables. Considérons les représentations linéaires du groupe unitaire à deux variables. On sait que ce groupe est de la forme:

$$\left.\begin{aligned} x'_1 &= ax_1 + bx_2, \\ x'_2 &= -\bar{b}x_1 + \bar{a}x_2, \end{aligned}\right\} \tag{93}$$

où les nombres complexes a et b doivent vérifier la condition:

$$a\bar{a} + b\bar{b} = 1. \tag{94}$$

On construit $(m+1)$ quantités:

$$\xi_0 = x_1^m;\ \xi_1 = x_1^{m-1}x_2;\ \ldots;\ \xi_m = x_2^m. \tag{95}$$

Si nous prenons $\xi'_k = x_1'^{m-k}x_2'^k$ et reportons les expressions de x'_1 et x'_2 tirées de (93), il est évident que chaque ξ'_k s'exprime linéairement en fonction des ξ_k, et par suite à toute transformation du groupe (93) correspond une transformation linéaire des variables ξ en fonction des variables ξ'_k. Il est évident que le produit des transformations correspond à un autre produit de transformations, et on a ainsi une représentation linéaire du groupe (93) d'ordre $(m+1)$. Mais il s'avère que cette représentation n'est pas unitaire. Pour construire une représentation unitaire il suffit d'introduire dans chaque variable (95) un facteur constant supplémentaire et, au lieu des formules (95), on définit les variables au moyen des formules:

$$\eta_k = \frac{x_1^{m-k} x_2^k}{\sqrt{(m-k)!\,k!}} \qquad (k = 0,\ 1,\ \ldots,\ m), \tag{96$_1$}$$

et de même:

$$\eta'_k = \frac{x_1'^{m-k} x_2'^k}{\sqrt{(m-k)!\,k!}} \qquad (k = 0,\ 1,\ \ldots,\ m), \tag{96$_2$}$$

où $0! = 1$.

Vérifions que pour une telle définition des variables la représentation considérée est unitaire, c'est-à-dire que:

$$\sum_{k=0}^{m} \eta'_k \overline{\eta'_k} = \sum_{k=0}^{m} \eta_k \overline{\eta_k}. \tag{97}$$

En effet, en appliquant la formule du binôme de Newton, on a:

$$m! \cdot \sum_{k=1}^{m} \eta'_k \overline{\eta'_k} = m! \sum_{k=0}^{m} \frac{x_1'^{m-k} \overline{x_1'^{m-k}} x_2'^k \overline{x_2'^k}}{(m-k!)\,k!} = (x'_1 \overline{x'_1} + x'_2 \overline{x'_2})^m,$$

et de même:

$$m! \cdot \sum_{k=0}^{m} \eta_k \overline{\eta_k} = (x_1 \overline{x_1} + x_2 \overline{x_2})^m.$$

Mais, compte tenu du fait que la transformation (93) est unitaire:

$$x'_1 \overline{x'_1} + x'_2 \overline{x'_2} = x_1 \overline{x_1} + x_2 \overline{x_2},$$

et, par conséquent, on a la relation (97).

Obtenons maintenant les formules qui donnent, sous forme explicite, les coefficients de la représentation unitaire du groupe (93). Dans ce but modifions quelque peu les notations précédentes, posons:

$$\eta_l = \frac{x_1^{j+l} x_2^{j-l}}{\sqrt{(j+l)!\,(j-l)!}} \, (l = -j,\ -j+1,\ \ldots,\ j-1,\ j). \tag{98}$$

Avec les notations précédentes $m=2j$ et le nombre j est entier si m est pair et *demi-entier* si m est impair. Si par exemple $m=5$, les formules (98) donnent les six variables suivantes:

$$\eta_{-\frac{5}{2}}=\frac{x_2^5}{\sqrt{5!}};\quad \eta_{-\frac{3}{2}}=\frac{x_1x_2^4}{\sqrt{1!4!}};\quad \eta_{-\frac{1}{2}}=\frac{x_1^2x_2^3}{\sqrt{2!3!}};$$

$$\eta_{\frac{1}{2}}=\frac{x_1^3x_2^2}{\sqrt{3!2!}};\quad \eta_{\frac{3}{2}}=\frac{x_1^4x_2}{\sqrt{4!1!}};\quad \eta_{\frac{5}{2}}=\frac{x_1^5}{\sqrt{5!}}.$$

Dans ce cas les variables ne sont pas numérotées au moyen des six premiers entiers, mais par des nombres fractionnaires qui diffèrent l'un de l'autre d'une unité et vont de $\left(-\frac{5}{2}\right)$ à $\left(+\frac{5}{2}\right)$. Si par exemple $m=4$, compte tenu des formules (98), on a les cinq variables:

$$\eta_{-2}=\frac{x_2^4}{\sqrt{4!}};\quad \eta_{-1}=\frac{x_1x_2^3}{\sqrt{1!3!}};\quad \eta_0=\frac{x_1^2x_2^2}{\sqrt{2!2!}};$$

$$\eta_1=\frac{x_1^3x_2}{\sqrt{3!1!}};\quad \eta_2=\frac{x_1^4}{\sqrt{4!}}.$$

Dans ce cas la numérotation des variables s'effectue au moyen d'entiers variant de (-2) à $(+2)$. Pour tout $m=2j$ fixé, on a une numérotation identique des lignes et des colonnes des matrices qui donnent une représentation linéaire d'ordre $(2j+1)$ du groupe (93).

Passons maintenant à la détermination des éléments de ces matrices. On a:

$$\eta'_l=\frac{x_1'^{j+l}x_2'^{j-l}}{\sqrt{(j+l)!(j-l)!}}=\frac{(ax_1+bx_2)^{j+l}(-\bar{b}x_1+\bar{a}x_2)^{j-l}}{\sqrt{(j+l)!(j-l)!}},$$

et il nous faut représenter le troisième membre sous forme d'une combinaison linéaire des quantités η_l. En appliquant la formule du binôme de Newton, on a:

$$\eta'=\sum_{k=0}^{j+l}\sum_{k'=0}^{j-l}(-1)^{j-l-k'}\frac{\sqrt{(j+l)!(j-l)!}}{k!k'!(j+l-k)!(j-l-k')!}\times$$
$$\times\,\bar{a}^{k'}a^{j+l-k}\bar{b}^{j-l-k'}b^kx_1^{2j-k-k'}x_2^{k+k'}.$$

Si on considère que $p!=\infty$, lorsque p est un entier négatif, on peut dans la formule précédente effectuer la sommation sur k et k' de $(-\infty)$ à $(+\infty)$, car les composantes supplémentaires contiennent au dénominateur un facteur infini et, par conséquent, ces termes s'annulent. Introduisons au lieu de k' une nouvelle variable de sommation $s=j-k-k'$ sur laquelle on sommera de $(-\infty)$ à $(+\infty)$ suivant les valeurs entières ou demi-entières selon que j est entier ou demi-entier. On obtient ainsi:

$$\eta_l' = \sum_k \sum_s (-1)^{k+s-l} \frac{\sqrt{(j+l)!\,(j-l)!}}{k!\,(j-k-s)!\,(j+l-k)!\,(k+s-l)!} \times$$
$$\times \bar{a}^{j-k-s} a^{j+l-k} \bar{b}^{k+s-l} b^k x_1^{j+s} x_2^{j-s}.$$

Mais, compte tenu de (98), on a:

$$x_1^{j+s} x_2^{j-s} = \sqrt{(j+s)!\,(j-s)!}\,\eta_s,$$

et finalement on obtient la relation linéaire cherchée sous la forme:

$$\eta_l' = \sum_k \sum_s (-1)^{k+s-l} \frac{\sqrt{(j+l)!\,(j-l)!\,(j+s)!\,(j-s)!}}{k!\,(j-k-s)!\,(j+l-k)!\,(k+s-l)!} \times$$
$$\times \bar{a}^{j-k-s} a^{j+l-k} \bar{b}^{k+s-l} b^k \eta_s.$$

Ainsi, pour j donné, les éléments de la matrice de la transformation linéaire d'ordre $(2j+1)$ correspondant à la transformation unitaire (93) de matrice

$$\begin{Vmatrix} a & b \\ -\bar{b} & \bar{a} \end{Vmatrix}$$

sont:

$$D_j \begin{Bmatrix} a, & b \\ -\bar{b}, & \bar{a} \end{Bmatrix}_{ls} = (-1)^{s-l} \sum_k (-1)^k \frac{\sqrt{(j+l)!\,(j-l)!\,(j+s)!\,(j-s)!}}{k!\,(j-k-s)!\,(j+l-k)!\,(k+s-l)!} \times$$
$$\times \bar{a}^{j-k-s} a^{j+l-k} \bar{b}^{k+s-l} b^k. \qquad (99)$$

Les indices l et s prennent toutes les valeurs:

$$l,\ s = -j,\ -j+1,\ \ldots,\ j-1,\ j,$$

et on se souviendra que si j est demi-entier, on a une numérotation des lignes et des colonnes de la matrice en nombre de demi-entiers. Tenant compte de ce que $p! = \infty$, si p est un entier négatif, on obtient les bornes suivantes de sommation dans les formules (99):

$$k \geqslant 0; \quad k \geqslant l-s; \quad k \leqslant j-s; \quad k \leqslant j+l. \qquad (100)$$

Il est possible de simplifier les formules (99) en passant à une représentation semblable. Soient A une matrice d'éléments a_{pq} et $S = [\delta_1, \ldots, \delta_n]$ une matrice diagonale.

En utilisant la règle habituelle de multiplication, il est aisé de vérifier que la matrice SAS^{-1} a les éléments suivants:

$$\{SAS^{-1}\}_{pq} = \delta_p a_{pq} \delta_q^{-1}.$$

Si l'on applique maintenant cette règle à la matrice

$$D_j \begin{Bmatrix} a & b \\ -\bar{b} & \bar{a} \end{Bmatrix}$$

et si de plus $\delta_k = (-1)^k$, alors dans les formules (99) le facteur $(-1)^{s-l}$ disparaît et dans la suite on pourra négliger ce facteur.

Démontrons maintenant que *la représentation linéaire du groupe unitaire* (93), *définie par la matrice d'éléments* (99), *est irréductible*. Auparavant deux lemmes sont à démontrer.

L e m m e I. *Si une matrice diagonale, dont tous les éléments diagonaux sont différents deux à deux, commute avec une matrice A, alors A est aussi une matrice diagonale.*

Par hypothèse on a:

$$A[\delta_1, \ldots, \delta_n] = [\delta_1, \ldots, \delta_n]A,$$

où les nombres δ_k sont différents deux à deux. Soient a_{pq} les éléments de la matrice A. En appliquant la règle de multiplication. on obtient à partir de la condition précédente:

$$a_{pq}\delta_q = \delta_p a_{pq} \quad \text{soit} \quad a_{pq}(\delta_q - \delta_p) = 0,$$

et, par suite, $a_{pq} = 0$ si $p \neq q$, c'est-à-dire que la matrice A est effectivement diagonale.

L e m m e II. *Si une matrice diagonale $[\delta_1, \ldots, \delta_n]$ commute avec une matrice A dont l'une des colonnes au moins ne contient aucun zéro, alors $\delta_1 = \ldots = \delta_n$.*

En permutant lignes et colonnes, c'est-à-dire en passant aux matrices semblables, on peut obtenir que la colonne qui ne contient pas de zéros soit la première. Alors la matrice diagonale reste comme auparavant diagonale et les matrices considérées vont commuter comme par le passé. Ainsi on peut considérer, en désignant par a_{pq} les éléments de la matrice A, que

$$a_{i1} \neq 0 \qquad (i = 1, 2, \ldots, n),$$

et de plus, par hypothèse, on a:

$$a_{i1}(\delta_1 - \delta_i) = 0 \qquad (i = 1, 2, \ldots, n),$$

d'où $\delta_1 = \ldots = \delta_n$. Le lemme est ainsi démontré.

Démontrons maintenant que la représentation linéaire définie par les matrices (99) est irréductible. Soit Y une matrice d'ordre $(2j + 1)$ qui commute avec toutes les matrices:

$$D_j \begin{Bmatrix} a & b \\ -\bar{b} & \bar{a} \end{Bmatrix}$$

obtenues pour différentes valeurs de a et b vérifiant la condition (94). Pour démontrer l'irréductibilité il faut montrer que Y doit être un multiple de la matrice unité. Considérons d'abord le cas où $b = 0$ et $a = e^{i\alpha}$. Ces nombres complexes vérifient évidemment la condition (94).

En utilisant les formules (99), on obtient tout d'abord que:

$$D_j\begin{Bmatrix} e^{i\alpha} & 0 \\ 0 & e^{-i\alpha}\end{Bmatrix}_{ls}=0 \quad \text{si} \quad l\neq s.$$

Les élements diagonaux sont dans ce cas:

$$D_j\begin{Bmatrix} e^{i\alpha} & 0 \\ 0 & e^{-i\alpha}\end{Bmatrix}_{ll}=e^{i2l\alpha}\,(l=-j,\ -j+1,\ \ldots,\ j-1,\ j)$$

et la matrice de la forme:

$$D_j\begin{Bmatrix} e^{i\alpha} & 0 \\ 0 & e^{-i\alpha}\end{Bmatrix}=\begin{Vmatrix} e^{-i2j\alpha} & 0 & 0 & \ldots & 0 \\ 0 & e^{-i2(j-1)\alpha} & 0 & \ldots & 0 \\ 0 & 0 & e^{-i2(j-2)\alpha} & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \ldots & e^{-2j\alpha}\end{Vmatrix}, \qquad (101)$$

c'est-à-dire que cette matrice diagonale possède des éléments différents sur la diagonale principale si l'on choisit convenablement α. En utilisant le premier lemme on peut affirmer que la matrice Y, qui doit commuter avec la matrice (101), doit être aussi diagonale, c'est-à-dire que

$$Y=[\delta_1,\ \ldots,\ \delta_n]. \qquad (102)$$

Supposons maintenant que les nombres a et b ne sont pas nuls et considérons ce que devient la première colonne de la matrice $D_j\begin{Bmatrix} a & b \\ -\bar{b} & \bar{a}\end{Bmatrix}$. Les éléments en sont définis par les formules (99) si l'on fait $s=-j$. Les inégalités (100) donnent alors:

$$k\geqslant 0\ ;\quad k\geqslant l+j;\quad k\leqslant 2j;\quad k\leqslant j+l$$
$$(l=-j,\ -j+1,\ \ldots,\ j-1,\ j),$$

d'où l'on voit que la somme qui figure dans la formule (99) se réduit alors à une seule composante non nulle qui s'obtient pour $k=j+l$. Ainsi, dans ce cas, la première colonne de la matrice

$$D_j\begin{Bmatrix} a & b \\ -\bar{b} & \bar{a}\end{Bmatrix}$$

ne contiendra effectivement pas de zéros. Mais, compte tenu du lemme II et de ce que la matrice diagonale (102) doit commuter avec cette matrice, tous les nombres δ_k sont identiques, c'est-à-dire que Y est un multiple de la matrice unité. Ainsi les matrices

$$D_j\begin{Bmatrix} a & b \\ -\bar{b} & \bar{a}\end{Bmatrix}$$

donnent effectivement une représentation linéaire irréductible du groupe unitaire (93). En donnant à j la série de valeurs :

$$j=0,\ \frac{1}{2},\ 1,\ \frac{3}{2},\ 2,\ \ldots$$

on obtient une infinité de représentations linéaires. Pour $j=0$, on obtient la représentation identique triviale qui à tout élément du groupe (93) fait correspondre le nombre unité. Examinons maintenant, pour $j>0$, àquelles transformations du groupe (93) correspond dans le groupe des représentations $D_j\left\{\begin{matrix} a & b \\ -\bar{b} & \bar{a}\end{matrix}\right\}$ la transformation identique définie par les égalités $\eta_l'=\eta_l$ ou, ce qui revient au même, par les égalités :

$$(ax_1+bx_2)^{j+l}(-\bar{b}x_1+\bar{a}x_2)^{j-l}=x_1^{j+l}x_2^{j-l}$$
$$(l=-j,\ \ -j+1,\ \ldots,\ j-1,\ j).$$

En supposant $j=l$, on obtient $(ax_1+bx_2)^{2j}=x_1^{2j}$, d'où il résulte que $b=0$ et l'identité précédente s'écrit :

$$a^{j+l}\bar{a}^{j-l}x_1^{j+l}x_2^{j-l}=x_1^{j+l}x_2^{j-l}\quad (l=-j,\ -j+1,\ \ldots,\ j-1,\ j),$$

d'où $a^{j+l}\bar{a}^{j-l}=1$. Mais $|a|=1$ lorsque $b=0$ et la dernière égalité s'écrit :

$$a^{2l}=1\quad (l=-j,\ \ -j+1,\ \ldots,\ j-1,\ j).$$

Si j est un demi-entier, il faut poser $l=\frac{1}{2}$, ce qui donne $a=1$. Si j est un entier, l'égalité $a^{2l}=1$ se réduit à $a^2=1$, d'où $a=\pm 1$.

Ainsi, si j est un demi-entier, la transformation identique du groupe $D_j\left\{\begin{matrix} a & b \\ -\bar{b} & \bar{a}\end{matrix}\right\}$ correspond à une et une seule transformation identique dans le groupe (93), c'est-à-dire que dans ce cas $D_j\left\{\begin{matrix} a & b \\ -\bar{b} & \bar{a}\end{matrix}\right\}$ est une représentation biunivoque du groupe (93). Si j est un entier, deux transformations de matrices

$$E=\begin{Vmatrix}1 & 0\\ 0 & 1\end{Vmatrix};\quad S=\begin{Vmatrix}-1 & 0\\ 0 & -1\end{Vmatrix}=-E$$

correspondent dans le groupe (93) à la transformation identique dans le groupe $D_j\left\{\begin{matrix} a & b \\ -\bar{b} & \bar{a}\end{matrix}\right\}$. Ces transformations forment un groupe cyclique du deuxième ordre et $D_j\left\{\begin{matrix} a & b \\ -\bar{b} & \bar{a}\end{matrix}\right\}$ est une représentation biunivoque du groupe quotient [III-1-7]. Autrement dit, à toute transformation de la représentation $D_j\left\{\begin{matrix} a & b \\ -\bar{b} & \bar{a}\end{matrix}\right\}$ pour j entier correspondent deux transformations du groupe (93), dont les nombres a et b ne diffèrent que par le signe.

III-2-5. Représentations linéaires du groupe des rotations. Les résultats précédents sont particulièrement importants, car le groupe unitaire (93) est étroitement lié au groupe des rotations de l'espace à trois dimensions et les résultats obtenus donnent des représentations linéaires irréductibles du groupe des rotations.

A toute transformation unitaire (93) correspond une rotation définie et le changement simultané du signe de a et b donne une transformation unitaire à laquelle correspond la même rotation. Les paramètres a et b sont liés aux angles d'Euler par la formule [III-1-12]:

$$a = e^{-\frac{1}{2}i(\gamma+\alpha)}\cos\frac{1}{2}\beta; \quad b = -ie^{\frac{1}{2}i(\gamma-\alpha)}\sin\frac{1}{2}\beta. \tag{103}$$

Considérons d'abord le cas où j est un entier. Dans ce cas les formules (99) montrent que si l'on change simultanément les signes de a et b, les composantes qui figurent dans le second membre ne changent pas, car la somme des exposants de a, $\bar{a}$, b et $\bar{b}$ est alors égale à un nombre pair $2j$. Ainsi aux deux transformations unitaires qui donnent la même rotation correspond la même matrice dans la représentation linéaire. Autrement dit, pour j entier à chaque rotation d'angles d'Euler $\{\alpha, \beta, \gamma\}$ correspond une matrice définie dans la représentation linéaire D_j. Au lieu de:

$$D_j\left\{\begin{matrix} a & b \\ -\bar{b} & \bar{a}\end{matrix}\right\}$$

désignons cette matrice par:

$$D_j\{\alpha, \beta, \gamma\}. \tag{104}$$

Si j est un demi-entier, en changeant simultanément le signe de a et de b toutes les expressions (99) changent de signe, c'est-à-dire qu'aux transformations unitaires qui conduisent à un même déplacement correspondent alors des matrices différentes, dont tous les éléments diffèrent par le signe. Dans ce cas à chaque rotation correspondent deux matrices qui diffèrent par le signe, c'est-à-dire que dans (104) il faut mettre deux signes devant D_j. Ainsi pour j entier, les matrices (104) donnent une représentation linéaire du groupe des rotations. Pour j demi-entier on n'obtient pas en général une représentation linéaire. Dans ce cas, on a une représentation linéaire à deux déterminations.

Pour obtenir les éléments des matrices (104), il suffit de remplacer dans (99) a et b par leurs valeurs tirées des formules (103). En négligeant le facteur $(-1)^{s-l}$, on obtient:

$$D_j\{\alpha, \beta, \gamma\}_{ls} = i^{s-l} \sum_k (-1)^k \frac{\sqrt{(j+l)!\,(j-l)!\,(j+s)!\,(j-s)!}}{k!\,(j-k-s)!\,(j+l-k)!\,(k+s-l)!} \times$$

$$\times e^{-il\alpha - is\gamma} \cos^{2j+l-2k-s} \frac{1}{2}\beta \sin^{2k+s-l} \frac{1}{2}\beta. \quad (105)$$

Si l'on passe à la représentation équivalente au moyen de la matrice:

$$X = \begin{Vmatrix} 0 & 0 \dots 0 & 1 \\ 0 & 0 \dots 1 & 0 \\ \cdot & \cdot \;\cdot\;\cdot\;\cdot & \cdot \\ 0 & 1 \dots 0 & 0 \\ 1 & 0 \dots 0 & 0 \end{Vmatrix},$$

tout se ramène à une permutation des lignes et des colonnes dans l'ordre inverse, c'est-à-dire que cela équivaut à remplacer l et s par $(-l)$ et $(-s)$. Ainsi au lieu des formules (105) on peut écrire:

$$D'_j\{\alpha, \beta, \gamma\}_{ls} = i^{l-s} \sum_k (-1)^k \frac{\sqrt{(j+l)!\,(j-l)!\,(j+s)!\,(j-s)!}}{k!\,(j-k+s)!\,(j-l-k)!\,(k-s+l)!} \times$$

$$\times e^{il\alpha + is\gamma} \cos^{2j-l-2k+s} \frac{1}{2}\beta \sin^{2k-s+l} \frac{1}{2}\beta, \quad (106)$$

et d'après les considérations de [III-2-4] on peut négliger le facteur i^{l-s}.

Examinons quelques cas particuliers très simples. Si $j = 0$, on a une représentation linéaire du premier ordre:

$$\eta' = \eta.$$

C'est la représentation identique triviale. Si $j = \frac{1}{2}$, on a $2j + 1 = 2$ et les quantités $\eta_{-\frac{1}{2}}$ et $\eta_{\frac{1}{2}}$ sont égales à x_2 et x_1, c'est-à-dire que le groupe unitaire (93) est sa propre représentation linéaire (à une permutation des lignes et des colonnes près).

Pour le groupe des déplacements on obtient une représentation à deux déterminations du deuxième ordre définie par les matrices:

$$D'_{\frac{1}{2}}\{\alpha, \beta, \gamma\} = \begin{Vmatrix} e^{-\frac{1}{2}i(\gamma+\alpha)} \cos\frac{1}{2}\beta & ie^{\frac{1}{2}i(\gamma-\alpha)} \sin\frac{1}{2}\beta \\ ie^{-\frac{1}{2}i(\gamma-\alpha)} \sin\frac{1}{2}\beta & e^{\frac{1}{2}i(\gamma+\alpha)} \cos\frac{1}{2}\beta \end{Vmatrix}.$$

Pour $j=1$ on a une représentation linéaire du troisième ordre:

$$D_1'\{\alpha,\beta,\gamma\}=\left\|\begin{matrix} e^{-i(\gamma+\alpha)}\dfrac{1+\cos\beta}{2} & -e^{-i\alpha}\dfrac{\sin\beta}{\sqrt{2}} & e^{i(\gamma-\alpha)}\dfrac{1-\cos\beta}{2} \\ e^{-i\gamma}\dfrac{\sin\beta}{\sqrt{2}} & \cos\beta & -e^{i\gamma}\dfrac{\sin\beta}{\sqrt{2}} \\ e^{-i(\gamma-\alpha)}\dfrac{1-\cos\beta}{2} & e^{i\alpha}\dfrac{\sin\beta}{\sqrt{2}} & e^{i(\gamma+\alpha)}\dfrac{1+\cos\beta}{2} \end{matrix}\right\|.$$

Les représentations linéaires $D_j'\{\alpha,\beta,\gamma\}$, pour j entier, donnent une représentation biunivoque du groupe des rotations. Ceci résulte directement de ce qu'à chaque $D_j\{\alpha,\beta,\gamma\}$ correspondent deux matrices du groupe (93) qui diffèrent simplement par les signes de a et b, et à ces matrices correspond, comme on l'a déjà vu, une même rotation. Si j est un demi-entier, à toute rotation correspondent deux matrices de la représentation $D_j\{\alpha,\beta,\gamma\}$ qui ne diffèrent que par le signe. En particulier, à la transformation identique du groupe des rotations correspondent, de $D_j\{\alpha,\beta,\gamma\}$, les matrices $\pm E$, où E est la matrice unité d'ordre $(2j+1)$. *Si on se limite aux transformations de $D_j\{\alpha,\beta,\gamma\}$ suffisamment proches de la transformation identique, les $D_j\{\alpha,\beta,\gamma\}$ sont des représentations univoques du groupe des rotations.* Alors on peut se limiter dans les formules générales (106) aux valeurs de α, β et γ suffisamment proches de zéro. Mais si on ajoute 2π à α ou γ, étant donné que s et l sont des demi-entiers, tous les éléments de la matrice $D_j\{\alpha,\beta,\gamma\}$ changent de signe et on obtient une seconde représentation de la rotation qui est en fait la même. Montrerons par la suite que *les représentations indiquées sont des représentations irréductibles isomorphes du groupe des rotations.*

Comme les représentations $D_j\{\alpha,\beta,\gamma\}$ sont toutes les représentations irréductibles du groupe des rotations, la matrice $D_1'\{\alpha,\beta,\gamma\}$ doit être semblable à la matrice $D\{\alpha,\beta,\gamma\}$ qui correspond aux rotations de l'espace d'angles d'Euler $\{\alpha,\beta,\gamma\}$. On a vu [III-1-12] que $D=Z_\alpha X_\beta Z_\gamma$ et en multipliant les matrices situées à gauche, on obtient:

$$D=\left|\begin{matrix} \cos\alpha\cos\gamma-\sin\alpha\cos\beta\sin\gamma & -\cos\alpha\sin\gamma-\sin\alpha\cos\beta\cos\gamma & \sin\alpha\sin\beta \\ \sin\alpha\cos\gamma+\cos\alpha\cos\beta\sin\gamma & -\sin\alpha\sin\gamma+\sin\alpha\cos\beta\cos\gamma & -\cos\alpha\cos\beta \\ \sin\beta\sin\gamma & \sin\beta\cos\gamma & \cos\beta \end{matrix}\right|,$$

et il est aisé de vérifier que:

$$AD_1'\{\alpha,\beta,\gamma\}A^{-1}=D\{\alpha,\beta,\gamma\},$$

où :

$$A = \begin{Vmatrix} 1 & 0 & 1 \\ i & 0 & -i \\ 0 & \sqrt{2}i & 0 \end{Vmatrix}.$$

III-2-6. Théorème du groupe simple des rotations. Démontrons maintenant que le groupe des rotations est un groupe simple, c'est-à-dire qu'il n'a pas de sous-groupes distingués [III-1-7]. S'il avait un tel sous-groupe distingué compte tenu de ce qui a été dit [III-1-12], il lui correspondrait un sous-groupe distingué du groupe G des transformations (57) à déterminant égal à 1 distinct du sous-groupe distingué H formé de E et de $(-E)$. Ainsi, il reste à montrer que le groupe G n'a pas de sous-groupes distingués autres que H, c'est-à-dire qu'il faut démontrer que si le diviseur élémentaire H_1 du groupe G contient la matrice A autre que E et $(-E)$, H_1 coïncide avec G. Remarquons avant tout que si H_1 contient une matrice B, compte tenu de la définition du sous-groupe distingué, H_1 contient toutes les matrices $U^{-1}BU$, où U est une matrice quelconque du groupe G. En choisissant convenablement la matrice U, on peut obtenir ainsi n'importe quelle matrice du groupe G ayant les mêmes valeurs propres que la matrice B. Par conséquent, pour établir que H_1 coïncide avec G, il suffit de montrer que H_1 contient des matrices ayant des valeurs propres quelconques parmi celles possibles. Ces valeurs doivent être de la forme $e^{i\omega}$ et $e^{-i\omega}$, où ω est un nombre réel, car la matrice est unitaire et de déterminant égal à l'unité.

D'après ce qui a été dit, on peut prendre à la place de la matrice A la matrice $U^{-1}AU$ et ainsi on peut supposer que A est une matrice diagonale.

On voit ainsi que H_1 contient la matrice $A = [e^{i\varphi}, e^{-i\varphi}]$, où φ est réel et $e^{i\varphi} \neq \pm 1$. Alors $A^{-1} = [e^{-i\varphi}, e^{i\varphi}]$. Prenons une matrice arbitraire du groupe G:

$$U = \begin{Vmatrix} x & y \\ -\bar{y} & \bar{x} \end{Vmatrix} \qquad (\bar{x}x + y\bar{y} = 1).$$

Alors :

$$U^{-1} = \begin{Vmatrix} \bar{x} & -y \\ \bar{y} & x \end{Vmatrix}.$$

Comme ce sous-groupe H_1 contient A et qu'il est un sous-groupe distingué, il doit contenir la matrice :

$$Y = A\,(UA^{-1}U^{-1}).$$

En multipliant les matrices et en tenant compte de ce que $x\bar{x}+y\bar{y}=1$, on obtient l'expression suivante de la trace s de la matrice Y:

$$s=2-4y\bar{y}\sin^2\varphi=2-4\rho^2\sin^2\varphi,$$

où $\rho=|y|$ peut prendre n'importe quelle valeur dans l'intervalle $0\leqslant\rho\leqslant 1$ et $\sin\varphi\neq 0$. Les valeurs propres $(e^{i\alpha}, e^{-i\alpha})$ de la matrice Y sont les racines de l'équation:

$$\lambda^2-s\lambda+1=0,\ \text{c'est-à-dire}\ \lambda^2+(4\rho^2\sin^2\varphi-2)\lambda+1=0.$$

Lorsque ρ varie de $\rho=0$ à $\rho=1$, les valeurs de α varient de $\alpha=0$ à $\alpha=2\varphi$. Introduisons la notation suivante:

$$U_\beta=[e^{i\beta},\ e^{-i\beta}].$$

De ce qui a été dit, il résulte que H_1 contient toutes les matrices H_α lorsque $0\leqslant\alpha\leqslant 2\varphi$. Maintenant il est aisé de montrer que H contient n'importe quelle matrice U_β $(\beta>0)$. En effet, choisissons un nombre n, entier positif, tel que l'inégalité $0<\frac{\beta}{n}<2\varphi$ soit vérifiée. Alors H_1 contient $U_{\frac{\beta}{n}}$ et donc aussi:

$$U^n_{\frac{\beta}{n}}=U_\beta.$$

Ainsi H_1 contient des matrices de valeurs propres quelconques et, compte tenu de ce qui a été dit, coïncide avec G. On a ainsi démontré que *le groupe des rotations est un groupe simple*.

Il en résulte directement que le groupe des rotations ne peut avoir de représentations homomorphes (non isomorphes). En effet, si l'on avait une représentation de ce type, à la transformation identique du groupe des représentations devraient correspondre dans le groupe des rotations des transformations engendrant un sous-groupe distingué, qui n'existe pas, comme on l'a déjà démontré.

III-2-7. Equation de Laplace et représentations linéaires du groupe des rotations. Etudions maintenant la relation qui existe entre les représentations linéaires des groupes et les équations différentielles. Cette relation est à la base de l'application des représentations linéaires aux problèmes de physique moderne. Commençons par le cas simple de l'équation de Laplace (tome II, [VII-3-1]). Ceci n'apportera rien de nouveau mais simplement servira à expliciter le problème général. Auparavant établissons certains faits généraux qui jouent un rôle important dans les questions de représentations linéaires des groupes et que l'on connaît déjà à partir de certains cas particuliers.

Soit un groupe G, dont on cherche les représentations linéaires et qui est le groupe des transformations linéaires d'ordre n:

$$x'_k = g_{k1}^{(\alpha)} x_1 + \ldots + g_{kn}^{(\alpha)} x_n \quad (k = 1, 2, \ldots, n), \tag{107}$$

où l'indice α, qui caractérise l'élément du groupe G, peut prendre toutes les valeurs au sein d'un ensemble fini ou infini. Supposons de plus qu'il existe m fonctions:

$$\varphi_s(x_1, \ldots, x_n) \quad (s = 1, 2, \ldots, m), \tag{108}$$

telles que si l'on change les variables indépendantes d'après les formules (107), ces fonctions subissent également une transformation linéaire:

$$\varphi_s(x'_1, \ldots, x'_n) = a_{s1}^{(\alpha)} \varphi_1(x_1, \ldots, x_n) + \ldots + a_{sm}^{(\alpha)} \varphi_m(x_1, \ldots, x_n) \tag{109}$$

$$s = (1, 2, \ldots, m).$$

On a ici une matrice A_α, à éléments $a_{ik}^{(\alpha)}$, qui correspond à la transformation (107) du groupe G. On considère deux transformations du groupe:

$$(x'_1, \ldots, x'_n) = G_{\alpha_1}(x_1, \ldots, x_n); \; (x''_1, \ldots, x''_n) = G_{\alpha_2}(x'_1, \ldots, x'_n);$$

$$G_{\alpha_3} = G_{\alpha_2} G_{\alpha_1}.$$

La transformation correspondante des fonctions (108) est:

$$\varphi_s(x'_1, \ldots, x'_n) = a_{s1}^{(\alpha_1)} \varphi_1(x_1, \ldots, x_n) + \ldots + a_{sm}^{(\alpha_1)} \varphi_m(x_1, \ldots, x_n) \tag{110_1}$$

et

$$\varphi_s(x''_1, \ldots, x''_n) = a_{s1}^{(\alpha_2)} \varphi_1(x'_1, \ldots, x'_n) + \ldots + a_{sm}^{(\alpha_2)} \varphi_m(x'_1, \ldots, x'_n). \tag{110_2}$$

En portant dans (110_2) au lieu de $\varphi_s(x'_1, \ldots, x'_n)$ leurs expressions (110_1) on a la dépendance de $\varphi_s(x''_1, \ldots, x''_n)$ par $\varphi_s(x_1, \ldots, x_n)$ qui donne la matrice A_{α_3}. On obtient ainsi:

$$A_{\alpha_3}\}_{ik} = \sum_{s=1}^{m} a_{is}^{(\alpha_2)} a_{sk}^{(\alpha_1)}, \text{ c'est-à-dire } A_{\alpha_3} = A_{\alpha_2} A_{\alpha_1},$$

et les formules (109) définissent une certaine représentation linéaire d'ordre m pour le groupe G. Dans les raisonnements précédents nous avons supposé que les fonctions φ_s sont linéairement indépendantes. Alors les transformations linéaires (109) sont définies de façon univoque et $D(A_\alpha) \neq 0$, car dans le cas contraire les $\varphi_s(x'_1, \ldots, x'_n)$ seraient liées par une relation linéaire.

Dans le cas particulier où l'on construisait les représentations linéaires du groupe unitaire, le rôle des fonctions φ_s était joué par les fonctions (96_1).

Supposons que G est le groupe des rotations de l'espace à trois dimensions, de sorte que $n = 3$, et supposons que les fonctions φ_s sont orthogonales et normées dans une boule K de centre à l'origine, c'est-à-dire que

$$\iiint_K \varphi_p(x_1, x_2, x_3)\overline{\varphi_q(x_1, x_2, x_3)}\,dx_1\,dx_2\,dx_3 = \delta_{pq}. \qquad (111)$$

Montrons que la représentation linéaire (109) du groupe des rotations est alors unitaire. En effet, à la suite de la rotation G_α, la boule K se transforme en elle-même et le déterminant de G_α est égal à l'unité. La condition (111) donne alors:

$$\iiint_K \varphi_p(x_1', x_2', x_3')\overline{\varphi_q(x_1', x_2', x_3')}\,dx_1'\,dx_2'\,dx_3' = \delta_{pq}$$

ou, compte tenu de (109):

$$\iiint_K \left[\sum_{i=1}^{m} a_{pi}^{(\alpha)}\varphi_i(x_1, x_2, x_3)\cdot\sum_{j=1}^{m}\overline{a_{qj}^{(\alpha)}\varphi_j(x_1, x_2, x_3)}\right]dx_1'\,dx_2'\,dx_3' = \delta_{pq}.$$

En passant aux variables (x_1, x_2, x_3), on doit, d'après la loi de changement de variables dans l'intégrale triple, remplacer $dx_1'\,dx_2'\,dx_3'$ par $dx_1\,dx_2\,dx_3$ et ensuite intégrer sur la même boule K. Compte tenu de la condition (111), on a:

$$\sum_{i=1}^{m} a_{pi}^{(\alpha)}\overline{a_{qi}^{(\alpha)}} = \delta_{pq} \qquad (p, q = 1, 2, \ldots, m),$$

où $\delta_{pq} = 0$ si $p \neq q$ et $\delta_{pp} = 1$, c'est-à-dire que chacune des matrices A est telle que ses lignes sont orthogonales de sorte que la matrice transposée est telle que ses colonnes sont orthogonales et, par suite [II-1-9], aussi ses lignes, c'est-à-dire que la matrice de base est orthogonale par ses lignes comme par ses colonnes, autrement dit, la matrice A_α est effectivement une matrice unitaire quel que soit l'indice α.

Soit maintenant l'équation de Laplace à deux variables:

$$\frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial y^2} = 0 \qquad (112)$$

ou, en utilisant les notations vectorielles:

$$\operatorname{div}\operatorname{grad} U = 0. \qquad (113)$$

Considérons un polynôme homogène en x et y de degré l:

$$\varphi_l(x, y) = a_0x^l + a_1x^{l-1}y + \ldots + a_kx^{l-k}y^k + \ldots + a_ly^l. \qquad (114)$$

Montrons qu'il existe deux polynômes linéairement indépendants de la forme (114) qui sont solution de l'équation (112) et que toute solution de l'équation (112), polynôme homogène de degré l, doit être une combinaison linéaire de ces deux polynômes à coefficients constants. En effet, les coefficients du polynôme (114) s'expriment au moyen des formules:

$$a_k = \frac{1}{(l-k)!\,k!}\,\frac{\partial^l \varphi_l(x, y)}{\partial x^{l-k}\,\partial y^k}.$$

Mais puisque ce polynôme doit vérifier l'équation (112), on peut remplacer la double différentiation par rapport à y par une double différentiation par rapport à x avec changement de signe, car l'équation (112) peut être mise sous la forme:

$$\frac{\partial^2 U}{\partial y^2} = -\frac{\partial^2 U}{\partial x^2}.$$

Ainsi, on obtient pour les coefficients a_k une expression de la forme:

$$a_k = \pm\frac{1}{(l-k)!\,k!}\,\frac{\partial^l \varphi_l}{\partial x^l} \quad \text{ou} \quad a_k = \pm\frac{1}{(l-k)!\,k!}\,\frac{\partial^l \varphi_l}{\partial x^{l-1}\,\partial y},$$

c'est-à-dire que tous les coefficients du polynôme (114) s'expriment en fonction des coefficients a_0 et a_1. Ce raisonnement montre qu'il n'existe pas plus de deux polynômes homogènes, linéairement indépendants, vérifiant l'équation (112). Montrons maintenant que ces deux polynômes distincts existent. Pour cela considérons le polynôme homogène:

$$\omega_l(x, y) = (x + iy)^l.$$

En ouvrant les parenthèses et en séparant les parties réelle et imaginaire, on obtient:

$$\omega_l(x, y) = \varphi_l(x, y) + i\psi_l(x, y),$$

où les polynômes $\varphi_l(x, y)$ et $\psi_l(x, y)$ sont des polynômes réels, homogènes, de degré l, linéairement indépendants. En différentiant $\omega_l(x, y)$ on obtient:

$$\frac{\partial^2 \omega_l(x, y)}{\partial x^2} = l(l-1)(x+iy)^{l-2};$$

$$\frac{\partial^2 \omega_l(x, y)}{\partial y^2} = -l(l-1)(x+iy)^{l-2},$$

où $\omega_l(x, y)$ vérifie l'équation (112). On peut en dire de même des parties réelle et imaginaire de cette expression, c'est-à-dire des polynômes $\varphi_l(x, y)$ et $\psi_l(x, y)$ qui donnent les deux solutions cherchées de l'équation (112). Introduisons les coordonnées polaires:

$$x = r\cos\varphi; \qquad y = r\sin\varphi,$$

d'où :

$$\omega_l(x, y) = r^l e^{il\varphi}.$$

Alors les polynômes φ_l et ψ_l prennent la forme simple :

$$\varphi_l(x, y) = r^l \cos l\varphi; \qquad \psi_l(x, y) = r^l \sin l\varphi.$$

Faisons tourner le plan XY d'un angle ϑ autour de l'origine :

$$x' = x\cos\vartheta - y\sin\vartheta, \qquad y' = x\sin\vartheta + y\cos\vartheta. \tag{115}$$

Il est facile de voir qu'alors l'équation (112) reste invariante, c'est-à-dire que dans les nouvelles variables l'équation a exactement la même forme :

$$\frac{\partial^2 U}{\partial x'^2} + \frac{\partial^2 U}{\partial y'^2} = 0. \tag{116}$$

Ceci peut être vérifié directement en utilisant les formules (115) et la règle de différentiation des fonctions de fonction et résulte directement de ce que le premier membre de l'équation (113) a une valeur définie ne dépendant pas du choix des axes et par suite conserve la même forme quel que soit le choix des coordonnées orthogonales rectilignes. Les polynômes $\varphi_l(x', y')$ et $\psi_l(x', y')$ doivent vérifier l'équation (116) et par suite l'équation (112), c'est pourquoi ils doivent s'exprimer linéairement en fonction de $\varphi_l(x, y)$ et $\psi_l(x, y)$. Ceci donne la représentation linéaire du groupe des rotations du plan.

Au lieu des deux polynômes indiqués on prendra deux autres polynômes qui en sont des combinaisons linéaires :

$$\varphi_l'(x, y) = \varphi_l(x, y) - i\psi_l(x, y); \qquad \psi_l'(x, y) = \varphi_l(x, y) + i\psi_l(x, y)$$

ou

$$\varphi_l'(x, y) = (x - iy)^l = r^l e^{-il\varphi}; \qquad \psi_l'(x, y) = (x + iy)^l = r^l e^{il\varphi}.$$

Ces polynômes subissent la transformation :

$$\varphi_l'(x', y') = r^l e^{-il(\varphi+\vartheta)} = e^{-il\vartheta}\varphi_l'(x, y),$$
$$\psi_l'(x', y') = r^l e^{il(\varphi+\vartheta)} = e^{il\vartheta}\psi_l'(x, y),$$

c'est-à-dire qu'à la transformation (115) correspond dans la représentation linéaire la matrice :

$$\begin{Vmatrix} e^{-il\vartheta} & 0 \\ 0 & e^{il\vartheta} \end{Vmatrix},$$

où l'angle ϑ peut prendre n'importe quelle valeur. De la forme de la matrice il résulte que cette représentation linéaire a une forme réduite. Elle donne deux représentations linéaires du premier ordre définies par les nombres $e^{-il\vartheta}$ et $e^{il\vartheta}$. Dans tous les raisonnements précédents l'entier l pouvait prendre n'importe quelle valeur. On a obtenu ainsi les mêmes représentations linéaires du groupe des rotations du plan que dans [III-2-5].

Passons maintenant à l'équation de Laplace à trois variables:

$$\frac{\partial^2 U}{\partial x^2}+\frac{\partial^2 U}{\partial y^2}+\frac{\partial^2 U}{\partial z^2}=0 \tag{117}$$

ou

$$\operatorname{div}\operatorname{grad} U=0. \tag{118}$$

Examinons aussi les polynômes homogènes de degré l à trois variables:

$$\varphi_l(x, y, z)=a_0 z^l+X_1(x, y)\, z^{l-1}+X_2(x, y)\, z^{l-2}+\ldots+ \\ +X_{l-1}(x, y)\, z+X_l(x, y), \tag{119}$$

où $X_k(x, y)$ est un polynôme homogène de degré k de ses arguments. Chacun de ces polynômes contient $(k+1)$ coefficients arbitraires de sorte qu'en général le polynôme homogène $\varphi_l(x, y, z)$ de degré l à trois variables contient le nombre suivant

$$1+2+3+\ldots+(l+1)=\frac{(l+1)(l+2)}{2}$$

de coefficients arbitraires.

En portant l'expression (119) dans l'équation (117), on obtient à gauche un polynôme homogène de degré $(l-2)$ et en annulant ses coefficients on obtient $\frac{(l-1)\,l}{2}$ équations homogènes avec $\frac{(l+1)(l+2)}{2}$ coefficients inconnus du polynôme $\varphi_l(x, y, z)$. On a:

$$\frac{(l+1)(l+2)}{2}-\frac{(l-1)\,l}{2}=2l+1,$$

et par conséquent au moins $(2l+1)$ coefficients du polynôme $\varphi_l(x, y, z)$ restent arbitraires, c'est-à-dire qu'il existe au moins $(2l+1)$ polynômes homogènes de degré l, linéairement indépendants, vérifiant l'équation (117). En utilisant la même méthode que dans le cas de deux variables, on peut montrer qu'il n'y en a pas plus de $(2l+1)$, c'est-à-dire qu'il y en a exactement $(2l+1)$. Désignons ces polynômes par

$$\psi_s^{(l)}(x, y, z) \qquad (s=1, 2, \ldots, 2l+1).$$

Si

$$(x', y', z')=U\cdot(x, y, z)$$

est une rotation quelconque de l'espace à trois dimensions autour de l'origine, l'équation (117) reste invariante et les polynômes $\psi_s^{(l)}(x, y, z)$ donnent une certaine représentation linéaire du groupe des rotations de l'espace à trois dimensions d'ordre $(2l+1)$.

Dans la suite on exposera en détail la théorie de ces polynômes harmoniques et on obtiendra leurs expressions explicites. On verra qu'on peut toujours les choisir de façon qu'ils soient orthogonaux et normés sur n'importe quelle boule de centre à l'origine. Alors la représentation linéaire du groupe des rotations qu'ils fournissent est unitaire. On peut montrer que c'est justement la représentation linéaire équivalente à la représentation $D_l\{\alpha, \beta, \gamma\}$ obtenue [III-2-5]. Nous reviendrons sur cette question.

III-2-8. Produit direct de matrices. Soient deux matrices:

$$A=\left\|\begin{matrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{matrix}\right\| \quad \text{et} \quad B=\left\|\begin{matrix} b_{11} & b_{12} & \dots & b_{1m} \\ b_{21} & b_{22} & \dots & b_{2m} \\ \cdot & \cdot & \cdot & \cdot \\ b_{m1} & b_{m2} & \dots & b_{mm} \end{matrix}\right\|, \tag{120}$$

la première d'ordre n, la seconde d'ordre m. On forme une nouvelle matrice C d'éléments $C_{ij;\,kl}$, en multipliant chaque élément de la matrice A par chacun des éléments de la matrice B:

$$\{C\}_{ij;\,kl}=c_{ij;\,kl}=a_{ik}b_{jl}. \tag{121}$$

Dans ce cas le premier indice est l'ensemble des deux nombres entiers (i, j) tandis que le second est l'ensemble des nombres entiers (k, l) avec:

$$i \text{ et } k=1, 2, \dots, n;$$

$$j \text{ et } l=1, 2, \dots, m.$$

En d'autres termes on désigne ici les lignes et les colonnes par l'ensemble de deux nombres entiers, le premier prenant les valeurs de 1 à n et le second de 1 à m. On peut bien entendu numéroter les lignes et les colonnes de la manière habituelle, c'est-à-dire au moyen de nombres entiers variant de 1 à nm, et à chaque couple de nombres (i, j) ou (k, l) correspond un nombre entier déterminé dans la nouvelle numérotation, et si ces couples sont identiques, les nombres entiers correspondants sont identiques. Cette numérotation au moyen d'entiers peut se faire de différentes façons. Lorsqu'on passe d'une numérotation à une autre, on permute simultanément lignes et colonnes, c'est-à-dire qu'on passe à la matrice semblable.

La matrice C est le *produit direct des matrices* A et B et on la désigne habituellement par:

$$C=A\times B, \tag{122}$$

L'ordre des facteurs dans ce produit d'un type nouveau ne joue aucun rôle.

Supposons par exemple que les deux matrices (120) soient des matrices du deuxième ordre. Leur produit direct est alors une matrice du quatrième ordre que l'on peut écrire par exemple sous la forme suivante :

$$C=\left\|\begin{matrix} a_{11}b_{11} & a_{11}b_{12} & a_{12}b_{11} & a_{12}b_{12} \\ a_{11}b_{21} & a_{11}b_{22} & a_{12}b_{21} & a_{12}b_{22} \\ a_{21}b_{11} & a_{21}b_{12} & a_{22}b_{11} & a_{22}b_{12} \\ a_{21}b_{21} & a_{21}b_{22} & a_{22}b_{21} & a_{22}b_{22} \end{matrix}\right\| = \left\|\begin{matrix} c_{11;\,11} & c_{11;\,12} & c_{11;\,21} & c_{11;\,22} \\ c_{12;\,11} & c_{12;\,12} & c_{12;\,21} & c_{12;\,22} \\ c_{21;\,11} & c_{21;\,12} & c_{21;\,21} & c_{21;\,22} \\ c_{22;\,11} & c_{22;\,12} & c_{22;\,21} & c_{22;\,22} \end{matrix}\right\|,$$

ou, autrement, en permutant simultanément les lignes et les colonnes.

Supposons que A et B soient deux matrices diagonales :

$$A=[\gamma_1, \ldots, \gamma_n]\,; \quad B=[\delta_1, \ldots, \delta_m].$$

Dans ce cas $a_{ik}=0$ et $b_{jl}=0$ si $i\neq k$ et $j\neq l$ et, par conséquent, compte tenu de (121), $c_{ij;\,kl}$ n'est pas nul dans le cas où le couple des nombres (i, j) coïncide avec le couple (k, l), c'est-à-dire si la matrice C est aussi diagonale. Sur sa diagonale principale on aura tous les produits possibles de γ_k par δ_l. Si tous les γ_k et δ_l sont égaux à l'unité, C sera également une matrice unité. On a donc le théorème suivant :

Théorème I. *Le produit direct de deux matrices diagonales est une matrice diagonale et le produit direct de deux matrices unité est une matrice unité.*

Démontrons le théorème suivant :

Théorème II. *Si $A^{(1)}$ et $A^{(2)}$ sont deux matrices du même ordre n et $B^{(1)}$ et $B^{(2)}$ deux matrices du même ordre m, on a*

$$(A^{(2)}\times B^{(2)})\,(A^{(1)}\times B^{(1)})=A^{(2)}A^{(1)}\times B^{(2)}B^{(1)}. \tag{123}$$

Remarquons que lorsqu'on écrit deux matrices du même ordre côte à côte sans aucun signe, il s'agit toujours du produit ordinaire de deux matrices. En désignant les éléments de la matrice par des lettres minuscules munies de deux indices, on a, compte tenu de la définition du produit direct :

$$\{A^{(t)}\times B^{(t)}\}_{ij;\,kl}=a_{ik}^{(t)}b_{jl}^{(t)} \qquad (t=1,\ 2),$$

et en utilisant la règle de la multiplication ordinaire des matrices, on obtient, pour les éléments du premier membre de l'égalité (123), la formule suivante

$$d_{ij;\,kl}=\sum_{p=1}^{n}\sum_{q=1}^{m} a_{ip}^{(2)}b_{jq}^{(2)}a_{pk}^{(1)}b_{ql}^{(1)}. \tag{124}$$

On montrera que les mêmes formules s'obtiennent pour les éléments du second membre. On a par définition du produit ordinaire de deux matrices:

$$\{A^{(2)}A^{(1)}\}_{ik} = \sum_{p=1}^{n} a_{ip}^{(2)}a_{pk}^{(1)}\,; \qquad \{B^{(2)}B^{(1)}\}_{jl} = \sum_{q=1}^{m} b_{jq}^{(2)}b_{ql}^{(1)},$$

et, par définition du produit direct :

$$d_{ij;\,kl} = \sum_{p=1}^{n} a_{ip}^{(2)}a_{pk}^{(1)}\cdot\sum_{q=1}^{m} b_{jq}^{(2)}b_{ql}^{(1)},$$

ce qui coïncide avec (124). Passons maintenant à la démonstration du dernier théorème sur le produit direct.

Théorème III. *Si les matrices A et B sont unitaires, leur produit direct $C = A \times B$ est aussi une matrice unitaire.*

En vertu du théorème on a :

$$\sum_{s=1}^{n} a_{sp}\bar{a}_{sq} = \delta_{pq}; \qquad \sum_{s=1}^{m} b_{sp}\bar{b}_{sq} = \delta_{pq}. \tag{125}$$

Vérifions pour la matrice C les conditions d'orthogonalité et de normalité par colonnes et désignons

$$\sum_{i=1}^{n}\sum_{j=1}^{m} c_{ij;\,p_1q_1}\bar{c}_{ij;\,p_2q_2} = \delta_{p_1q_1;\,p_2q_2},$$

c'est-à-dire que, compte tenu de (121) :

$$\delta_{p_1q_1;\,p_2q_2} = \sum_{i=1}^{n}\sum_{j=1}^{m} a_{ip_1}\bar{a}_{ip_2}b_{jq_1}\bar{b}_{jq_2} = \sum_{i=1}^{n} a_{ip_1}\bar{a}_{ip_2}\sum_{j=1}^{m} b_{jq_1}\bar{b}_{jq_2}. \tag{126}$$

Si les couples de nombres (p_1, q_1) et (p_2, q_2) sont différents, l'un au moins des facteurs du second membre de (126) sera nul et si ces couples coïncident, les deux facteurs seront égaux à l'unité, compte tenu de (125). Ainsi, $\delta_{p_1q_1;\,p_2q_2}$ est nul si les couples considérés ne coïncident pas et est égal à l'unité si les couples coïncident. Le théorème est ainsi démontré.

On peut évidemment multiplier le produit direct de deux matrices par une troisième au sens du produit direct. On obtient le produit direct de trois matrices:

$$A^{(1)} \times A^{(2)} \times A^{(3)}.$$

En conservant les notations précédentes, on a pour les éléments de cette nouvelle matrice l'expression suivante :

$$c_{ikl;\,i'k'l'} = a_{ii'}^{(1)}a_{kk'}^{(2)}a_{ll'}^{(3)}.$$

On forme de façon analogue le produit direct d'un nombre quelconque fini de matrices et ce produit direct est une matrice dont l'ordre est égal au produit des ordres des matrices facteurs. L'ordre des facteurs ne joue aucun rôle.

III-2-9. Composition de deux représentations linéaires de groupes. Soit un groupe G d'éléments G_α. Supposons construites deux représentations linéaires de ce groupe

$$x_i' = a_{i1}^{(\alpha)} x_1 + \ldots + a_{in}^{(\alpha)} x_n \quad (i = 1, 2, \ldots, n) \tag{127}$$

et

$$y_k' = b_{k1}^{(\alpha)} y_1 + \ldots + b_{km}^{(\alpha)} y_m \quad (k = 1, 2, \ldots, m), \tag{128}$$

où l'indice α peut prendre toutes les valeurs d'un ensemble fini ou infini. On désigne par $A^{(\alpha)}$ et $B^{(\alpha)}$ les matrices des transformations (127) et (128) et on forme leur produit direct

$$C^{(\alpha)} = A^{(\alpha)} \times B^{(\alpha)}. \tag{129}$$

Montrons que les matrices $C^{(\alpha)}$ donnent aussi une représentation linéaire du groupe G. En effet, à tout élément G_α du groupe G correspond une matrice $C^{(\alpha)}$; au produit $G_{\alpha_2} G_{\alpha_1} = G_{\alpha_3}$ correspond la matrice $C^{(\alpha_2)} C^{(\alpha_1)}$, qui est définie, compte tenu de (123), au moyen de la formule:

$$C^{(\alpha_2)} C^{(\alpha_1)} = (A^{(\alpha_2)} \times B^{(\alpha_2)}) (A^{(\alpha_1)} \times B^{(\alpha_1)}) = (A^{(\alpha_2)} A^{(\alpha_1)}) \times (B^{(\alpha_2)} B^{(\alpha_1)}).$$

Mais, étant donné que les matrices $A^{(\alpha)}$ et $B^{(\alpha)}$ donnent une représentation linéaire du groupe, on a :

$$A^{(\alpha_2)} A^{(\alpha_1)} = A^{(\alpha_3)} \quad \text{et} \quad B^{(\alpha_2)} B^{(\alpha_1)} = B^{(\alpha_3)},$$

et, par conséquent :

$$C^{(\alpha_2)} C^{(\alpha_1)} = A^{(\alpha_3)} \times B^{(\alpha_3)},$$

c'est-à-dire que, compte tenu de (129) :

$$C^{(\alpha_2)} C^{(\alpha_1)} = C^{(\alpha_3)}.$$

Ainsi, au produit des éléments G_α correspond notamment le produit des matrices correspondantes $C^{(\alpha)}$ et ces matrices donnent une nouvelle représentation linéaire du groupe G. Remarquons qu'à l'élément unité de G correspond le produit direct des matrices unité $A^{(\alpha)}$ et $B^{(\alpha)}$, c'est-à-dire la matrice unité $C^{(\alpha)}$.

On forme nm produits $x_i y_k$ et on soumet chacun des facteurs aux transformations (127) et (128). On a

$$x_i' y_k' = (a_{i1}^{(\alpha)} x_1 + \ldots + a_{in}^{(\alpha)} x_n) \cdot (b_{k1}^{(\alpha)} y_1 + \ldots + b_{km}^{(\alpha)} y_m)$$

ou, en ouvrant les parenthèses :

$$x'_i y'_k = \sum_{p=1}^{n} \sum_{q=1}^{m} c^{(\alpha)}_{ik;\,pq} x_p y_q,$$

où

$$c^{(\alpha)}_{ik;\,pq} = a^{(\alpha)}_{ip} b^{(\alpha)}_{kq},$$

c'est-à-dire que *si x_i et y_k sont des êtres des représentations linéaires définies par les matrices $A^{(\alpha)}$ et $B^{(\alpha)}$, alors les produits $x_i y_k$ sont des êtres de la représentation linéaire du même groupe définie par les matrices $C^{(\alpha)}$*. Si les matrices $A^{(\alpha)}$ et $B^{(\alpha)}$ donnent des représentations linéaires irréductibles, la matrice $C^{(\alpha)}$ ne donne pas forcément une représentation linéaire irréductible. On étudiera en détail le cas où G est le groupe des rotations de l'espace à trois dimensions et les matrices $A^{(\alpha)}$ et $B^{(\alpha)}$ les représentations linéaires irréductibles distinctes de ce groupe obtenues dans [III-2-5]. Montrons que dans ce cas le produit

$$D_{j_1}\{\alpha,\ \beta,\ \gamma\} \times D_{j_2}\{\alpha,\ \beta,\ \gamma\}$$

est réductible et déterminons à partir de quelles représentations irréductibles il est formé.

Comme exemple considérons l'équation de Schrödinger dans le cas de deux électrons soumis au champ d'un noyau positif. Cette équation est de la forme :

$$\left[-\frac{h^2}{8\pi^2 m} \sum_{s=1}^{2} \left(\frac{\partial^2}{\partial x_s^2} + \frac{\partial^2}{\partial y_s^2} + \frac{\partial^2}{\partial z_s^2}\right) + V\right] \psi = E\psi, \tag{130}$$

où :

$$V = \sum_{s=1}^{2} - \frac{e^2 e_0}{\sqrt{x_s^2 + y_s^2 + z_s^2}} + \frac{1}{2} \frac{e^2}{\sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2}}; \tag{131}$$

et les constantes ont les significations habituelles. Le deuxième terme de l'expression de V provient de l'interaction des électrons. Si on néglige, en première approximation, cette interaction, l'équation devient :

$$(H_1 + H_2)\,\psi = E\psi, \tag{132}$$

où :

$$H_s = -\frac{h^2}{8\pi^2 m} \left(\frac{\partial^2}{\partial x_s^2} + \frac{\partial^2}{\partial y_s^2} + \frac{\partial^2}{\partial z_s^2}\right) - \frac{e^2 e_0}{\sqrt{x_s^2 + y_s^2 + z_s^2}} \quad (s = 1,\ 2).$$

Supposons que les équations

$$H_1\psi = E\psi\,; \qquad H_2\psi = E\psi \tag{133}$$

ont les valeurs propres E_1 et E_2 et les fonctions propres correspondantes :

$$\psi_1(x_1,\ y_1,\ z_1) \quad \text{et} \quad \psi_2(x_2,\ y_2,\ z_2),$$

c'est-à-dire que :

$$H_1\psi_1 = E_1\psi_1, \qquad H_2\psi_2 = E_2\psi_2. \tag{134}$$

Si l'on porte dans l'équation (132) :

$$\psi = \psi_1(x_1, y_1, z_1)\cdot\psi_2(x_2, y_2, z_2),$$

on obtient, compte tenu de (134) :

$$(H_1 + H_2)\,\psi = \psi_2 H_1\psi_1 + \psi_1 H_2\psi_2 = (E_1 + E_2)\,\psi_1\psi_2 = (E_1 + E_2)\,\psi,$$

c'est-à-dire que l'équation (132) a pour fonction propre $\psi_1\psi_2$ qui correspond à la valeur $(E_1 + E_2)$. Le premier membre des équations (133) contient l'opérateur de Laplace et la distance d'un point à l'origine des coordonnées, par conséquent, ces premiers membres ne changent pas si on effectue une rotation de l'espace autour de l'origine. Il peut arriver que la valeur propre $E = E_1$ dans la première des équations (133) corresponde à plusieurs fonctions propres ψ_1. Toutes ces fonctions sont solutions de l'équation et donnent une certaine représentation linéaire du groupe des rotations, de même qu'en [III-2-5] les polynômes harmoniques homogènes donnaient aussi une représentation du groupe des rotations. On suppose que c'est une représentation $D_{j_1}\{\alpha, \beta, \gamma\}$. De même, les solutions de la seconde des équations (133), pour la valeur propre donnée $E = E_2$, fournissent une représentation $D_{j_2}\{\alpha, \beta, \gamma\}$ du groupe des rotations. Le produit $\psi_1\psi_2$, compte tenu de ce qui précède, donne une représentation linéaire du groupe des rotations qui coïncide avec le produit direct $D_{j_1}\times D_{j_2}$, et pour obtenir la caractéristique physique de la valeur propre correspondante $(E_1 + E_2)$ de l'équation (132), il est essentiel d'isoler dans cette représentation les représentations irréductibles en lesquelles on peut la décomposer. Ce fait joue un rôle essentiel dans la théorie des perturbations.

III-2-10. Produit direct de groupes et représentations linéaires. La notion de produit direct de matrices joue un rôle important dans une autre question que l'on examinera maintenant. Soient deux groupes G et H d'éléments G_α et H_β dont les indices α et β prennent en général des valeurs indépendantes l'une de l'autre des ensembles différents. On définit un nouveau groupe F dont les éléments sont des couples d'éléments de G et H :

$$F_{\alpha\beta} = (G_\alpha, H_\beta),$$

et le premier élément du couple est l'élément de G, le second l'élément de H. Appelons élément unité de ce groupe l'élément $F_{\alpha\beta}$ obtenu quand G_α et H_β sont les éléments unité de G et H et, de façon analogue, définissons les éléments inverses du groupe F. La loi de multiplication dans le groupe F est définie par la formule :

$$F_{\alpha_2\beta_2}F_{\alpha_1\beta_1} = (G_{\alpha_2}G_{\alpha_1}, H_{\beta_2}H_{\beta_1}).$$

Il est facile de vérifier que l'ensemble des éléments $F_{\alpha\beta}$ forme bien un groupe. Ce groupe F est le produit direct des groupes G et H. Soient une représentation linéaire quelconque du groupe G réalisée au moyen des matrices $A^{(\alpha)}$ et une représentation linéaire du groupe H réalisée par les matrices $B^{(\beta)}$. En utilisant la formule

(123), comme dans le paragraphe précédent, on peut montrer que *les produits directs*

$$C^{(\alpha,\beta)} = A^{(\alpha)} \times B^{(\beta)}$$

donnent une représentation linéaire du groupe F. De plus, si les représentations $A^{(\alpha)}$ et $B^{(\beta)}$ sont unitaires, la représentation $C^{(\alpha,\beta)}$ du groupe F est aussi unitaire [III-2-8].

Montrons maintenant que *si les représentations* $A^{(\alpha)}$ *et* $B^{(\beta)}$ *sont irréductibles, la représentation* $C^{(\alpha,\beta)}$ *du groupe* F *sera irréductible.* Soit n l'ordre des matrices $A^{(\alpha)}$ et m celui des matrices $B^{(\beta)}$. Les matrices $C^{(\alpha,\beta)}$ sont d'ordre nm. Supposons qu'il existe une matrice X d'ordre nm qui commute avec toutes les matrices $C^{(\alpha,\beta)}$. On désigne les éléments de la matrice par les lettres minuscules correspondantes. On a par suite, pour n'importe quel indice i, j, p et q, ainsi que pour n'importe quel α et β :

$$\sum_{l=1}^{m}\sum_{k=1}^{n} x_{ij;\,kl}a_{kp}^{(\alpha)}b_{lq}^{(\beta)} = \sum_{l=1}^{m}\sum_{k=1}^{n} a_{ik}^{(\alpha)}b_{jl}^{(\beta)}x_{kl;\,pq}, \tag{135}$$

où :

$$a_{kp}^{(\alpha)}b_{lq}^{(\beta)} = c_{kl;\,pq}^{(\alpha,\beta)} \quad \text{et} \quad a_{ik}^{(\alpha)}b_{jl}^{(\beta)} = c_{ij;\,kl}^{(\alpha,\beta)}.$$

Si l'on prend $G^{(\alpha)}$ comme élément unité du groupe G, $A^{(\alpha)}$ est la matrice unité, c'est-à-dire que $a_{kp}^{(\alpha)} = 0$ si $k \neq p$ et $a_{pp}^{(\alpha)} = 1$, et la formule (135) donne alors :

$$\sum_{l=1}^{m} x_{ij;\,pl}b_{lq}^{(\beta)} = \sum_{l=1}^{m} b_{jl}^{(\beta)}x_{il;\,pq}, \tag{136}$$

de même en posant que $B^{(\beta)}$ est l'élément unité du groupe B, on obtient :

$$\sum_{k=1}^{m} x_{ij;\,kq}a_{kp}^{(\alpha)} = \sum_{k=1}^{n} a_{ik}^{(\alpha)}x_{kj;\,pq}. \tag{137}$$

Si l'on prend nm éléments $x_{ij;\,kl}$ et si l'on fixe les indices i et k, on obtient m^2 éléments :

$$x_{ij;\,kl} \qquad (j,\ l = 1,\ 2,\ \ldots,\ m),$$

qui donnent une matrice d'ordre m. On désigne cette matrice par $X_1^{(i,k)}$. De même, en fixant dans $x_{ij;\,kl}$ les indices j et l, on obtient la matrice $X_2^{(j,l)}$ d'ordre n. Compte tenu de (136), toutes les matrices $X_1^{(i,k)}$ commutent avec toutes les matrices $B^{(\beta)}$ qui donnent une représentation irréductible du groupe B et, par suite, toutes les matrices $X_1^{(i,k)}$ sont des multiples de la matrice unité, c'est-à-dire que les éléments $x_{ij;\,kl}$, pour i et k fixés, ont la même valeur si $j = l$ et de plus $x_{ij;\,kl} = 0$ si $j \neq l$. On peut écrire ceci sous la forme suivante :

$$x_{ij;\,kl} = x_{i1;\,k1}\delta_{jl}. \tag{138$_1$}$$

De même, il résulte de l'examen de la matrice $X_2^{(j,\,l)}$ que :

$$x_{ij;\,kl} = x_{1j;\,1l}\delta_{ik}, \tag{138$_2$}$$

où :

$$\delta_{pq} = 0 \quad \text{si} \quad p \neq q \quad \text{et} \quad \delta_{pp} = 1.$$

De la comparaison de (138$_1$) et (138$_2$) il découle donc que les $x_{ij;\,kl}$ sont différents de zéro seulement si $i = k$ et $j = l$ et que dans ce cas tous les nombres $x_{ij;\,ij}$ sont identiques entre eux, c'est-à-dire que la matrice X qui commute avec toutes les matrices $C^{(\alpha,\,\beta)}$ est forcément un multiple de la matrice unité. Ainsi, la représentation linéaire du groupe F définie par le produit direct $A^{(\alpha)} \times B^{(\beta)}$ est irréductible. On peut montrer que *de cette façon on obtient toutes les représentations irréductibles du groupe F.*

Supposons que G et H soient des groupes des transformations linéaires avec un même nombre de variables et que deux matrices arbitraires G_α et H_β commutent deux à deux, c'est-à-dire que

$$G_\alpha H_\beta = H_\beta G_\alpha. \tag{139}$$

Dans les raisonnements précédents on a considéré l'élément du groupe F défini par le couple d'éléments (G_α, H_β) et on a construit la loi décrite ci-dessus de multiplication à l'intérieur du groupe F. On peut considérer que l'élément du groupe F est le simple produit des matrices (139) qui ne dépend pas de l'ordre. Ce nouveau groupe F est isomorphe à l'ancien groupe F. Si G_{α_0} et H_{β_0} sont des matrices unité, le produit $G_{\alpha_0}H_{\beta_0} = H_{\beta_0}G_{\alpha_0}$ est aussi une matrice unité. La matrice $G_\alpha^{-1}H_\beta^{-1} = H_\beta^{-1}G_\alpha^{-1}$ est évidemment l'inverse du produit $G_\alpha H_\beta$ et, compte tenu de (139), on a la loi de multiplication suivante :

$$G_{\alpha_2}H_{\beta_2} \cdot G_{\alpha_1}H_{\beta_1} = (G_{\alpha_2}G_{\alpha_1})(H_{\beta_2}H_{\beta_1}),$$

c'est-à-dire que toutes les propriétés supposées lors de la formation du groupe F sont vérifiées dans ce cas, de sorte que les produits (139) peuvent être considérés comme un élément variable du groupe F. Comme exemple particulier prenons le cas où G est le groupe des rotations de l'espace à trois dimensions et H le groupe du deuxième ordre formé de la transformation identique I et de la symétrie S par rapport à l'origine [III-1-6]. Dans ce cas, les conditions (139) sont vérifiées. Si G_α est une rotation quelconque de l'espace, il est évident que $G_\alpha S = SG_\alpha$. Le groupe F est ici le groupe de toutes les transformations réelles orthogonales de l'espace à trois dimensions. Pour le groupe H on avait [III-2-3] deux représentations linéaires du premier ordre: l'une identique, formée des nombres $(+1)$ et l'autre antisymétrique qui faisait correspondre le nombre $(+1)$ à la matrice I et (-1) à la matrice S. Soit maintenant une représentation linéaire $D_j\{\alpha, \beta, \gamma\}$ du groupe des rotations. On considère le produit direct des matrices de cette représentation avec les deux représentations du groupe de symétrie par rapport à l'origine. Dans un cas, on obtient une représentation linéaire du groupe complet des transformations orthogonales qui à toute rotation d'angles d'Euler $\{\alpha, \beta, \gamma\}$ fait correspondre une même matrice $D_j\{\alpha, \beta, \gamma\}$, que la rotation soit prise sous sa forme pure ou bien associée à une symétrie par rapport à l'origine. Cette représentation du groupe des transformations orthogonales est désignée par $D_j^+\{\alpha, \beta, \gamma\}$. Dans l'autre cas, à une rotation pure correspond la matrice $D_j\{\alpha, \beta, \gamma\}$ et à une rotation associée à une symétrie la matrice $-D_j\{\alpha, \beta, \gamma\}$. Une telle représentation du groupe des transformations orthogonales est représentée par $D_j^-\{\alpha, \beta, \gamma\}$.

Etudions encore un exemple de produit direct de deux groupes. Soient deux points: (x_1, y_1, z_1) et (x_2, y_2, z_2). Supposons que le groupe G soit le groupe des rotations de l'espace à trois dimensions. Les variables considérées subissent alors des transformations linéaires:

$$\begin{aligned} x_k' &= g_{11}x_k + g_{12}y_k + g_{13}z_k, \\ y_k' &= g_{21}x_k + g_{22}y_k + g_{23}z_k, \quad (k = 1, 2) \\ z_k' &= g_{31}x_k + g_{32}y_k + g_{33}z_k, \end{aligned} \tag{140}$$

où le tableau des g_{ik} est la matrice d'une rotation quelconque. Supposons de plus que H est le groupe formé de la transformation identique et de la transformation à laquelle correspond la permutation des numéros 1 et 2 des points. Cette dernière transformation est de la forme

$$\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \ (S). \tag{141}$$

Il est évident que $S^2 = I$ et, par suite, que le groupe H est formé de deux transformations I et S. Si G_α est une rotation, on a $G_\alpha S = SG_\alpha$, car il est indifférent de savoir si on numérote les points avant ou après la rotation. Dans ce cas, on obtient les mêmes représentations linéaires pour le groupe F que ci-dessus. Si au lieu de deux points on en avait pris n, le groupe H formé des changements de numérotation de ces points aurait pour ses éléments les transformations linéaires à n variables, et ce groupe H serait isomorphe au groupe des permutations de n éléments. Dans ce cas également, la rotation et la permutation des indices des points commutent et en prenant le produit direct de la matrice de la représentation linéaire du groupe des rotations et des matrices d'une représentation linéaire du groupe des permutations, on obtient une représentation linéaire du groupe F.

III-2-11. Décomposition du produit $D_j \times D_{j'}$ des représentations linéaires du groupe des rotations. Revenons à ce que nous avons dit en [III-2-9]. Nous avons vu que, dans le cas de l'équation de Schrödinger pour deux électrons, si l'on néglige l'interaction des électrons, les fonctions propres de l'équation de Schrödinger donnent une représentation linéaire du groupe des rotations qui s'obtient en faisant la composition de deux représentations linéaires du groupe des rotations. Les résultats du paragraphe précédent montrent qu'il est important de pouvoir décomposer une telle représentation linéaire en ses parties irréductibles. Dans ce paragraphe on s'occupera de ce problème. Soient deux représentations irréductibles $D_j\,\{\alpha, \beta, \gamma\}$ et $D_{j'}\,\{\alpha, \beta, \gamma\}$ du groupe des rotations. On forme leur produit $D_j \times D_{j'}$ qui donne aussi [III-2-9] une représentation linéaire du groupe des rotations. On cherche à isoler les parties irréductibles dont est formée cette représentation.

Les êtres de la représentation linéaire D_j d'ordre $(2j+1)$ sont les quantités:

$$U_m = \frac{u_1^{j+m} u_2^{j-m}}{\sqrt{(j+m)!\,(j-m)!}} \quad (m = -j,\ -j+1,\ \ldots,\ j-1,\ j), \tag{142}$$

ceux de la représentation linéaire $D_{j'}$ sont :

$$V_{m'} = \frac{v_1^{j'+m'} v_2^{j'-m'}}{\sqrt{(j'+m')!\,(j'-m')!}} \quad (m' = -j',\ -j'+1,\ \ldots,\ j'-1,\ j') \tag{143}$$

et (u_1, u_2) et (v_1, v_2) sont soumis aux mêmes transformations unitaires à déterminant égal à $(+1)$ [III-2-4]. Si l'on forme les $(2j+1)\,(2j'+1)$ quantités:

$$W_{mm'} = U_m V_{m'} = \frac{u_1^{j+m} u_2^{j-m} v_1^{j'+m'} v_2^{j'-m'}}{\sqrt{(j+m)!\,(j-m)!\,(j'+m')!\,(j'-m')!}} \tag{144}$$

$$\begin{pmatrix} m = -j,\ -j+1,\ \ldots,\ j-1,\ j \\ m' = -j',\ -j'+1,\ \ldots,\ j'-1,\ j' \end{pmatrix},$$

ces quantités sont les êtres de la représentation linéaire du groupe des rotations définie par le produit $D_j \times D_{j'}$.

Dans la suite, on considère j et j' soit entiers, soit demi-entiers, c'est-à-dire, à strictement parler, qu'on prend des représentations linéaires du groupe unitaire à deux variables et à déterminant égal à 1.

Soit k un entier (ou un demi-entier) vérifiant l'inégalité:

$$|j - j'| \leqslant k \leqslant j + j'. \tag{145}$$

Montrons que l'on peut composer à partir des quantités (144) $(2k+1)$ combinaisons linéaires qui donnent la représentation linéaire D_k du groupe des rotations.

On forme, pour démontrer cette proposition, une expression de la forme:

$$L = (u_1 v_2 - u_2 v_1)^l (u_1 x_1 + u_2 x_2)^{2j-l} (v_1 x_1 + v_2 x_2)^{2j'-l}, \tag{146}$$

où l est un entier fixé vérifiant les inégalités:

$$l \geqslant 0; \quad l \leqslant 2j; \quad l \leqslant 2j'. \tag{147}$$

Si les variables (u_1, u_2) et (v_1, v_2) subissent une même transformation linéaire:

$$u_1' = a_{11}u_1 + a_{12}u_2; \qquad v_1' = a_{11}v_1 + a_{12}v_2,$$
$$u_2' = a_{21}u_1 + a_{22}u_2; \qquad v_2' = a_{21}v_1 + a_{22}v_2,$$

à déterminant égal à $(+1)$, c'est-à-dire si $a_{11}a_{22} - a_{12}a_{21} = 1$, il est facile de voir que le premier facteur de l'expression (146) reste inchangé. En effet:

$$u_1'v_2' - u_2'v_1' = (a_{11}a_{22} - a_{12}a_{21})(u_1 v_2 - u_2 v_1).$$

L'expression (146) est un polynôme homogène de x_1 et x_2 de degré $2(j+j'-l)$. Il est donc formé de termes de la forme:

$$a_s x_1^s x_2^{2(j+j'-l)-s} \quad (s = 0, 1, \ldots, 2(j+j'-l)).$$

En introduisant la notation:

$$k = j + j' - l, \tag{148}$$

$$y_{m''} = \frac{x_1^{k+m''} x_2^{k-m''}}{\sqrt{(k+m'')!\,(k-m'')!}} \tag{149}$$

$$(m'' = -k, \quad -k+1, \ldots, k-1, k),$$

on peut écrire l'expression (146) sous la forme:

$$L = \sum_{m''=-k}^{+k} c_{m''} y_{m''}. \tag{150}$$

Les coefficients $c_{m''}$ dépendent des variables (u_1, u_2) (v_1, v_2).

Il résulte de l'expression (146) que $c_{m''}$ est un polynôme homogène de (u_1, u_2) de degré $2j$ et un polynôme homogène de (v_1, v_2) de degré $2j'$, c'est-à-dire que $c_{m''}$ sera constitué de termes de la forme

$$a'_{pq} u_1^p u_2^{2j-p} v_1^q v_2^{2j'-q},$$

ou, en tenant compte de (142) et (143), on peut affirmer que les $c_{m''}$ sont des combinaisons linéaires du produit :

$$c_{m''}=\sum_{m}\sum_{m'} d_{mm'}^{(m'')}U_mV_{m'} \qquad (m''=-k,\ -k+1,\ \ldots,\ k-1,\ k), \quad (151)$$

où les coefficients $d_{mm'}^{(m'')}$ ne contiennent pas u_k et v_k. Remarquons que dans l'expression (146) les variables u_1 et v_1 figurent soit en liaison avec le facteur x_1, soit dans le premier facteur de (146) et que ce premier facteur donne la somme des exposants de u_1 et v_1 égale à l. En tenant compte de ce que $y_{m''}$ contient $x_1^{k+m''}$, on peut affirmer que, dans les termes de (151) la somme des exposants de u_1 et v_1 est égale à $k+m''+l$ ou, compte tenu de (148), égale à $j+j'+m'$. Mais U_m contient u_1^{j+m} et $V_{m'}$ contient $v_1^{j'+m'}$, d'où il résulte que chacune des expressions (151) contient seulement les produits $U_mV_{m'}$ pour lesquels $m+m'=m''$. Montrons maintenant que les combinaisons linéaires (151) des quantités $U_mV_{m'}$ donnent justement une représentation linéaire du groupe des rotations équivalente à D_k.

Auparavant rappelons la définition des transformations contragrédientes. Si on a deux transformations linéaires

$$(x_1',\ \ldots,\ x_n')=A\,(x_1,\ \ldots,\ x_n) \quad \text{et} \quad (y_1',\ \ldots,\ y_n')=B\,(y_1,\ \ldots,\ y_n),$$

pour que l'égalité

$$x_1'y_1'+\ldots+x_n'y_n'=x_1y_1+\ldots+x_ny_n$$

soit vérifiée, il faut et il suffit que B soit contragrédient de A, c'est-à-dire que $B=A^{(*)-1}$ (cf. [II-1-2] et [II-2-9]). Supposons que les variables $(u_1,\ u_2)$ et $(v_1,\ v_2)$ soient soumises simultanément à une transformation unitaire A, à déterminant égal à $(+1)$, et que les variables x_1 et x_2 soient soumises lors de l'application de cette transformation $A^{(*)-1}$ à la transformation contragrédiente A. De la définition de la transformation contragrédiente, il résulte que les sommes

$$u_1x_1+u_2x_2 \quad \text{et} \quad v_1x_1+v_2x_2$$

restent inchangées. De plus, on a montré que le premier facteur de l'expression (146) reste inchangé si l'on effectue cette transformation de variables. Ainsi, la somme L, toute entière, reste inchangée, autrement dit, compte tenu de (150), les variables $c_{m''}$ subissent une transformation B contragrédiente de la transformation C à laquelle sont soumises les variables $y_{m''}$.

Introduisons les nouvelles variables :

$$z_{m''}=\frac{u_1^{k+m''}u_2^{k-m''}}{\sqrt{(k+m'')!\,(k-m'')!}}$$

$$(m''=-k,\ -k+1,\ \ldots,\ k-1,\ k).$$

En appliquant la formule du binôme de Newton, on peut écrire :

$$(u_1x_1+u_2x_2)^{2k}=(2k)!\sum_{m''=-k}^{+k} z_{m''}y_{m''}.$$

Le premier membre de cette expression reste invariant pour les transformations citées et, par suite, on peut en dire autant du second membre, c'est-à-dire que les variables $z_{m''}$ comme les variables $c_{m''}$ sont soumises à la transformation B contragrédiente de la transformation C. Mais on sait que les variables $z_{m''}$ donnent justement la représentation linéaire D_k du groupe des rotations si (u_1, u_2) sont les êtres d'un groupe unitaire à déterminant égal à $(+1)$. Notre proposition est donc démontrée.

On peut ainsi former à partir des variables (144) $(2k+1)$ combinaisons linéaires que l'on interprète comme étant les composantes d'un certain vecteur d'un espace à $(2j+1)$ $(2j'+1)$ dimensions et l'on a ainsi une représentation linéaire D_k du groupe des rotations. En tenant compte de la formule (148) et des inégalités (147), on voit que k peut prendre les valeurs suivantes :

$$k=j+j',\ j+j'-1,\ \ldots,\ |j-j'|. \tag{152}$$

Calculons maintenant combien de combinaisons linéaires des quantités (144) il faut former en tout. Pour fixer les idées, supposons que $j \geqslant j'$. Le nombre de combinaisons linéaires est :

$$(2j+2j'+1)+(2j+2j'-1)+\ldots+(2j-2j'+1).$$

C'est la somme d'une progression arithmétique dont le nombre de termes est :

$$\frac{(2j+2j'+1)-(2j-2j'+1)}{2}+1=2j'+1,$$

et le nombre total de combinaisons linéaires $(2j+1)$ $(2j'+1)$, c'est-à-dire qu'il est égal au nombre de quantités (144). On obtiendrait le même résultat si $j<j'$. En posant pour abréger

$$(2j+1)(2j'+1)=r,$$

on désigne les combinaisons linéaires des quantités (144) par :

$$w_1,\ w_2,\ \ldots,\ w_r, \tag{153}$$

et l'on considère que ces combinaisons linéaires sont disposées dans l'ordre donné par la représentation linéaire D_k, où k prend les valeurs (152). A la suite d'une transformation unitaire à déterminant égal à $(+1)$ sur les variables (u_1, u_2) et (v_1, v_2) on obtient de nouvelles valeurs $U'_mV'_m$ des variables (144) et de nouvelles valeurs w'_s $(s=1, 2, \ldots, r)$ des variables (153), lesquelles s'expriment, en outre, en fonction de w_s au moyen de la matrice quasi diagonale:

$$[D_{j+j'}, D_{j+j'-1}, \ldots, D_{|j-j'|}], \tag{154}$$

et chaque D_k correspond à la transformation unitaire qui agit sur (u_1, u_2) et (v_1, v_2). Plus loin nous montrerons que les formes linéaires (153) des quantités (144) sont linéairement indépendantes. Soit T la matrice de la transformation linéaire au moyen de laquelle les w_s s'expriment en fonction des variables (144). Le produit direct $D_j \times D_{j'}$ est la matrice de la transformation linéaire des variables (144) et, compte tenu de ce qui précède, on a:

$$[D_{j+j'}, D_{j+j'-1}, \ldots, D_{|j-j'|}] = T(D_j \times D_{j'})T^{-1}, \tag{155}$$

ce qui donne la décomposition du produit direct en parties irréductibles. La formule précédente s'écrit généralement:

$$D_j \times D_{j'} = D_{j+j'} + D_{j+j'-1} + \ldots + D_{|j-j'|}. \tag{156}$$

Rappelons que chaque D_k est définie au moyen d'une transformation unitaire et s'écrit: $D_k\begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix}$. Le résultat précédent peut se généraliser au cas de plusieurs facteurs. Ainsi, on peut écrire:

$$\begin{aligned} D_1 \times D_1 \times D_1 &= (D_2 + D_1 + D_0) \times D_1 = \\ &= D_3 + D_2 + D_1 + D_2 + D_1 + D_0 + D_1 = \\ &= D_3 + 2D_2 + 3D_1 + D_0. \end{aligned}$$

La matrice D_1 est une matrice du troisième ordre [III-2-4], le produit direct $D_1 \times D_1$ une matrice du neuvième ordre et enfin le produit direct $D_1 \times D_1 \times D_1$ une matrice du vingt-septième ordre. La formule précédente montre que cette matrice est équivalente, quel que soit le choix des transformations unitaires, à la matrice quasi diagonale:

$$[D_3, D_2, D_2, D_1, D_1, D_1, D_0].$$

L'ordre de cette dernière matrice est égal [III-2-4] à:

$$(2\cdot 3 + 1) + 2(2\cdot 2 + 1) + 3(2\cdot 1 + 1) + (2\cdot 0 + 1) = 27.$$

Démontrons maintenant que les w_s sont linéairement indépendants en tant que formes linéaires des quantités (144). Les quantités w_s, avec les notations précédentes, sont les $c_{m''}$, il faut seulement remarquer que pour obtenir $c_{m''}$, on peut prendre différentes valeurs de k ou, ce qui revient au même, différentes valeurs de l, de sorte qu'il est beaucoup plus normal d'écrire $c_{m''}^{(l)}$. On a vu déjà que tout $c_{m''}^{(l)}$ s'exprime seulement en fonction des quantités $U_m V_{m'}$.

pour lesquelles $m + m' = m''$. Il en résulte que les seuls $c_{m''}^{(l)}$ qui puissent être linéairement dépendants sont ceux pour lesquels les l sont différents, mais les m'' égaux. En ouvrant dans l'expression (146) les parenthèses dans les deux derniers facteurs et en regroupant les coefficients de $x_1^{k+m''}x_2^{k-m''}$, où k est défini par la formule (148), on peut obtenir à un facteur constant près l'expression de $c_{m''}^{(l)}$ en fonction de u_k et v_k. Ce sont des produits de $(u_1v_2 - u_2v_1)^l$ par un polynôme en u_1, u_2, v_1 et v_2 avec des coefficients positifs entiers. Il est aisé de voir que ces expressions ne peuvent être linéairement dépendantes pour les l distincts. Supposons par exemple que l'on ait une relation linéaire de la forme :

$$\alpha_1 c_{m''}^{(l_1)} + \alpha_2 c_{m''}^{(l_2)} + \alpha_3 c_{m''}^{(l_3)} = 0,$$

où $l_1 < l_2 < l_3$ et les α_k sont des constantes non nulles. Cette relation doit être vérifiée identiquement pour tous les u_1, u_2, v_1 et v_2. Posons par exemple $u_2 = v_1 = v_2 = 1$. Compte tenu de ce qui a déjà été dit de la forme de l'expression de $c_{m''}^{(l)}$, on obtient une relation de la forme :

$$\alpha_1 (u_1 - 1)^{l_1} p_1(u_1) + \alpha_2 (u_1 - 1)^{l_2} p_2(u_1) + \alpha_3 (u_1 - 1)^{l_3} p_3(u_1) = 0,$$

où les $p_k(u_1)$ sont des polynômes de u_1 à coefficients positifs entiers. En divisant la relation précédente par $(u_1 - 1)^{l_1}$ et en supposant ensuite $u_1 = 1$, on obtient $\alpha_1 = 0$, ce qui est contraire à nos raisonnements précédents et montre qu'il ne peut y avoir de relation linéaire.

En ouvrant les parenthèses dans l'expression (146), on peut construire effectivement l'expression de w_s en fonction des variables (144).

III-2-12. Orthogonalité. Les matrices qui forment des représentations unitaires irréductibles non équivalentes ont une certaine propriété que l'on désigne en général sous le terme d'*orthogonalité*. On l'utilise souvent dans l'application de la théorie des groupes en physique. Enonçons d'abord cette propriété.

Soit un groupe fini G d'ordre m dont les éléments sont

$$G_1, G_2, \ldots, G_m$$

et soient :

$$A^{(1)}, \ldots, A^{(m)} \text{ et } B^{(1)}, \ldots, B^{(m)}$$

deux systèmes de matrices donnant des représentations linéaires du groupe G. En désignant les éléments de ces matrices au moyen de lettres minuscules munies inférieurement de deux indices et en considérant que les deux représentations linéaires indiquées sont non équivalentes, irréductibles et formées de matrices unitaires, on a les égalités suivantes :

$$\sum_{s=1}^{m} a_{ij}^{(s)} \overline{b_{kl}^{(s)}} = 0, \qquad (157)$$

qui existent quels que soient les indices. On a des égalités analogues pour une représentation unitaire irréductible. Soit p l'ordre de la matrice $A^{(s)}$ qui donne une représentation unitaire irréductible. On a les formules suivantes:

$$\sum_{s=1}^{m} a_{ij}^{(s)} \overline{a_{kl}^{(s)}} = \frac{m}{p} \delta_{ik}\delta_{jl}, \tag{158}$$

c'est-à-dire que la somme du premier membre est nulle si les couples de nombres (i, j) et (k, l) sont différents et égale à $\frac{m}{p}$ si ces couples sont identiques.

La démonstration de l'orthogonalité utilise le théorème III de [III-2-2]. Rappelons auparavant la notion de produit dans le cas de matrices rectangulaires. Soient deux matrices C et D d'éléments

$$\{D\}_{ik} \begin{pmatrix} i=1, 2, \ldots n_1 \\ k=1, 2, \ldots, n_2 \end{pmatrix} \quad \text{et} \quad \{C\}_{jl} \begin{pmatrix} j=1, 2, \ldots, n_2 \\ l=1, 2, \ldots, n_3 \end{pmatrix},$$

en outre, le nombre n_2 de colonnes de la matrice D coïncide avec le nombre de lignes de la matrice C. Les éléments du produit DC sont définis par la formule:

$$\{DC\}_{ik} = \sum_{s=1}^{n_2} \{D\}_{is} \{C\}_{sk}.$$

La nouvelle matrice DC a évidemment n_1 lignes et n_3 colonnes.

Enonçons maintenant le théorème fondamental.

T h é o r è m e. *Si les matrices unitaires $A^{(s)}$ d'ordre p et les matrices unitaires $B^{(s)}$ d'ordre q donnent des représentations irréductibles non équivalentes du groupe G et si une matrice rectangulaire C à p lignes et q colonnes vérifie pour tout s les conditions*

$$A^{(s)}C = CB^{(s)} \qquad (s=1, 2, \ldots, m), \tag{159}$$

cette matrice C a tous ses éléments nuls.

Considérons d'abord le cas où $p = q$ lorsque C est aussi une matrice carrée. Si le déterminant de C n'est pas nul, $C^{(-1)}$ existe et il résulte de (159) que:

$$A^{(s)} = CB^{(s)}C^{-1},$$

c'est-à-dire que les deux représentations sont équivalentes, ce qui est contraire aux hypothèses du théorème. Ainsi, le déterminant de C doit être nul. Supposons que tous les éléments de C ne soient pas nuls. Nous désignons ces éléments par c_{ik}. On sait que les formes linéaires

$$c_{i1}x_1 + \ldots + c_{ip}x_p \quad (i=1, 2, \ldots, p)$$

définissent pour des x_s arbitraires un sous-espace dont le nombre de dimensions est égal au rang de la matrice C [I-2-7], c'est-à-dire que dans le cas considéré on a un sous-espace d'un nombre de dimensions $\geqslant 1$ et $< p$. Autrement dit, ce n'est pas tout l'espace à p dimensions mais un sous-espace R. On écrit (159) comme une transformation linéaire du vecteur de composantes $(x_1, \ldots, x_p)$:

$$A^{(s)}C(x_1, \ldots, x_p) = CB^{(s)}(x_1, \ldots, x_p) \quad (s=1, 2, \ldots, m).$$

Dans le premier membre $C(x_1, \ldots, x_p)$ est un vecteur arbitraire de R et tout le second membre est une transformation linéaire C du vecteur $B^{(s)}(x_1, \ldots, x_p)$ qui appartient aussi à R. Autrement dit, la transformation

$A^{(s)}$ de tout vecteur de R donne aussi un vecteur de R. Alors d'après [III-2-2] les $A^{(s)}$ donnent des représentations réductibles, ce qui est contraire aux hypothèses du théorème.

Si $p > q$, le rang de la matrice C est dans tous les cas inférieur à p et les formes linéaires:

$$c_{i1}x_1 + \ldots + c_{iq}x_q \quad (i = 1, 2, \ldots, p)$$

définissent dans l'espace à p dimensions un sous-espace R dont le nombre de dimensions est inférieur à p, de sorte que la démonstration précédente reste valable. Supposons enfin que $p < q$ et dans les conditions (159) passons aux matrices transposées. On a

$$B^{(s)*}C^* = C^*A^{(s)*}.$$

Dans ce cas l'ordre q des matrices $B^{(s)*}$ est plus grand que l'ordre p des matrices $A^{(s)*}$. On en déduit, comme ci-dessus, que les matrices unitaires $B^{(s)*}$ laissent un espace invariant, donc on peut choisir des vecteurs de base tels qu'elles se mettent sous la forme quasi diagonale. Alors les matrices $B^{(s)}$ sont aussi quasi diagonales, ce qui est contraire aux hypothèses du théorème. Le théorème est ainsi démontré.

On pourrait ne pas se soucier, dans les hypothèses du théorème, de ce que les matrices $A^{(s)}$ et $B^{(s)}$ sont unitaires. On sait qu'en passant aux représentations semblables, on peut toujours considérer $A^{(s)}$ et $B^{(s)}$ comme unitaires et le passage à ces représentations dans les relations (159) introduit au lieu de C une nouvelle matrice C_1 liée à C par une relation de la forme

$$C = D_1C_1D_2,$$

si C_1 est une matrice dont tous les éléments sont nuls, C aussi.

Passons maintenant à la démonstration des formules (157). Nous désignerons $A^{(s)}$ et $B^{(s)}$ par $A(G_s)$ et $B(G_s)$, où G_s est l'élément du groupe G auquel correspondent les matrices $A^{(s)}$ et $B^{(s)}$. Soit X une matrice quelconque ayant p lignes et q colonnes. Introduisons la matrice:

$$C = \sum_{s=1}^{m} A(G_s)XB(G_s)^{-1} \tag{160}$$

et montrons qu'elle vérifie les relations (159).

Soit G_t un élément donné quelconque du groupe G. On a:

$$A(G_t)C = \sum_{s=1}^{m} A(G_t)A(G_s)XB(G_s)^{-1}.$$

Mais, compte tenu de la définition de la représentation linéaire:

$$A(G_t)A(G_s) = A(G_tG_s) \quad \text{et} \quad B(G_t)B(G_s) = B(G_tG_s),$$

et donc:

$$A(G_t)C = \sum_{s=1}^{m} A(G_tG_s)XB(G_tG_s)^{-1}B(G_t).$$

L'élément G_s parcourt tout le groupe, on peut en dire autant du produit G_tG_s, de sorte que la formule précédente peut s'écrire sous la forme:

$$A(G_t)C = CB(G_t),$$

c'est-à-dire que la matrice C, définie par la formule (160), vérifie effectivement les relations (159) et que, par conséquent, cette matrice C a tous ses éléments nuls. Ainsi, quelle que soit la matrice X, on a :

$$\sum_{s=1}^{m} A(G_s)\,X B(G_s)^{-1}=0.$$

Supposons un élément donné quelconque $\{X\}_{jl}$ de la matrice X égal à l'unité et les autres éléments nuls. La formule écrite donne alors :

$$\sum_{s=1}^{m} \{A(G_s)\}_{ij}\,\{B(G_s)^{-1}\}_{lk}=0.$$

La matrice $B(G_s)$ s'obtient à partir de $B(G_s)^{-1}$ en échangeant les lignes et les colonnes et tous les éléments par leurs conjugués complexes car elle est unitaire, de sorte que la formule précédente avec les mêmes notations devient :

$$\sum_{s=1}^{m} a_{ij}^{(s)}\overline{b_{kl}^{(s)}}=0,$$

ce qui coïncide avec (157).

De même, en construisant la matrice :

$$D=\sum_{s=1}^{m} A(G_s)\,X A(G_s)^{-1},$$

où X est une matrice carrée d'ordre p, on peut démontrer que :

$$A(G_s)\,D=DA(G_s) \qquad (s=1,\ 2,\ \ldots,\ m),$$

et, compte tenu du théorème III [III-2-2], on peut affirmer que D est un multiple de la matrice unité, car :

$$\sum_{s=1}^{m} A(G_s)\,XA(G_s)^{-1}=cI,$$

où le nombre c dépend du choix de X. Supposons à nouveau $\{X\}_{jl}=1$, et les autres éléments de X nuls et désignons par c_{jl} les valeurs correspondantes du nombre c. On peut écrire :

$$\sum_{s=1}^{m} \{A(G_s)\}_{ij}\,\{A(G_s)^{-1}\}_{lk}=c_{jl}\delta_{ik}. \tag{161}$$

Pour déterminer c_{jl} on fait $i=k$ et on somme par rapport à i de 1 à p :

$$pc_{jl}=\sum_{s=1}^{m}\sum_{i=1}^{p} \{A(G_s)^{-1}\}_{li}\,\{A(G_s)\}_{ij}=\sum_{s=1}^{m}\{I\}_{lj}.$$

Si $l=j$, le second membre est égal à m et si $l\neq j$, il est nul. Il en résulte que $c_{jl}=\frac{m}{p}\,\delta_{jl}$ et, par suite, la formule (161) s'écrit :

$$\sum_{s=1}^{m} \{A(G_s)\}_{ij}\{A(G_s)^{-1}\}_{lk}=\frac{m}{p}\,\delta_{ik}\delta_{jl}. \tag{162}$$

ce qui coïncide avec (158), si on tient compte du fait que les matrices $A(G_s)$ sont unitaires.

Il est aisé de voir que la relation (157) est vraie, non seulement pour les représentations unitaires de groupes, mais aussi pour des représentations irréductibles non équivalentes. Soient $A'(G_s)$ et $B'(G_s)$ deux représentations d'ordre p et q et $A(G_s)$ et $B(G_s)$ des représentations unitaires équivalentes aux premières, telles que

$$A(G_s)=C_1A'(G_s)C_1^{-1}; \quad B(G_s)=C_2B'(G_s)C_2^{-1},$$

où C_1 et C_2 sont des matrices définies ne dépendant pas de s. $B(G_s)$ étant unitaire, on a:

$$B(G_s)^{-1}=\overline{B(G_s)^*}=\overline{(C_2^{-1})^*}\,\overline{B'(G_s)^*}\,\overline{C_2^*},$$

et la formule (157) peut s'écrire sous la forme:

$$\sum_{s=1}^{m} C_1A'(G_s)C_1^{-1}X\overline{(C_2^{-1})^*}\,\overline{B'(G_s)^*}\,\overline{C_2^*}=0,$$

d'où, en multipliant à gauche par C_1^{-1}, à droite par $\overline{(C_2^*)^{-1}}$ et on introduisant la matrice arbitraire $Y=C_1^{-1}X\overline{(C_2^{-1})^*}$ à p lignes et q colonnes:

$$\sum_{s=1}^{m} A'(G_s)Y\overline{B'(G_s)^*}=0.$$

Mais Y est arbitraire et l'on obtient comme ci-dessus:

$$\sum_{s=1}^{m} a_{ij}'^{(s)}\overline{b_{kl}'^{(s)}}=0.$$

Remarquons aussi que la formule (162) reste vraie pour n'importe quelle représentation et pas seulement pour les représentations unitaires. Ceci résulte de la démonstration et de ce que dans l'énoncé du théorème ci-dessus il n'est pas fait mention que les matrices $A^{(s)}$ et $B^{(s)}$ soient obligatoirement unitaires.

III-2-13. **Caractères.** Soient, comme ci-dessus, $A(G_s)$ et $B(G_s)$ deux représentations irréductibles, non équivalentes, d'ordre p et q, du groupe G d'éléments $G_1, G_2, \ldots, G_m$. On désigne par $X(G_s)$ et $X'(G_s)$ les traces des matrices dans ces représentations, c'est-à-dire la somme de leurs éléments diagonaux:

$$X(G_s)=\sum_{i=1}^{p}\{A(G_s)\}_{ii}; \quad X'(G_s)=\sum_{k=1}^{q}\{B(G_s)\}_{kk}.$$

Ces nombres sont les *caractères* des représentations. Des représentations équivalentes ont évidemment des caractères identiques [II-1-8] et l'on peut supposer les représentations citées unitaires. L'orthogonalité donne:

$$\sum_{s=1}^{m}\{A(G_s)\}_{ii}\overline{\{B(G_s)\}_{kk}}=0,$$

d'où en sommant par rapport à i et k, on obtient l'orthogonalité des caractères:

$$\sum_{s=1}^{m} X(G_s)\overline{X'(G_s)}=0. \qquad (163)$$

De même, la formule (158) donne :

$$\sum_{s=1}^{m} \{A(G_s)\}_{ii} \overline{\{A(G_s)\}_{kk}} = \frac{m}{p} \delta_{ik}$$

et, en sommant sur i et k, on obtient :

$$\sum_{s=1}^{m} X(G_s) \overline{X(G_s)} = m. \tag{164}$$

En utilisant ces formules, on peut démontrer les théorèmes suivants.

Théorème 1. *La condition nécessaire et suffisante pour que deux représentations irréductibles soient équivalentes est que tous leurs caractères soient égaux.*

On a déjà indiqué que les caractères des représentations équivalentes (réductibles ou irréductibles) sont égaux et, par là même, la condition nécessaire est établie. Supposons maintenant le contraire, à savoir que le système de caractères coïncide pour deux représentations irréductibles, c'est-à-dire que $X(G_s) = X'(G_s)$, $(s = 1, 2, \ldots, m)$ et démontrons que les représentations sont équivalentes. Compte tenu de (164) on a

$$\sum_{s=1}^{m} X(G_s) \overline{X'(G_s)} = m,$$

d'où il résulte que les représentations sont équivalentes, car s'il n'en était pas ainsi, on devrait avoir l'égalité (163). Remarquons que l'ordre des matrices dans des représentations équivalentes doit être identique. A chaque représentation irréductible on associe, dans un espace complexe à m dimensions R_m, les vecteurs de composantes:

$$\frac{1}{\sqrt{m}} X(G_1), \quad \frac{1}{\sqrt{m}} X(G_2), \ldots, \quad \frac{1}{\sqrt{m}} X(G_m).$$

Compte tenu de (164), ces vecteurs sont normés et, compte tenu de (163), les vecteurs correspondant à des représentations non équivalentes sont orthogonaux. Il en résulte qu'*il ne peut pas exister plus de m représentations irréductibles non équivalentes du groupe G d'ordre m*. Dans la suite, on précisera le nombre total de représentations irréductibles non équivalentes d'un groupe. Pour l'instant, désignons ce nombre par l. Soient $\omega^{(i)}$ $(i = 1, 2, \ldots, l)$ ces représentations irréductibles non équivalentes et:

$$X^{(i)}(G_1), \quad X^{(i)}(G_2), \ldots, X^{(i)}(G_m) \qquad (i = 1, 2, \ldots, l)$$

les caractères de ces représentations. Soit une certaine représentation ω de caractères :

$$X(G_1), \quad X(G_2), \ldots, X(G_m).$$

La représentation ω étant réductible, elle sera représentée par des matrices quasi diagonales formées des matrices de la représentation $\omega^{(i)}$. On aura ainsi pour les caractères :

$$X(G_s) = \sum_{i=1}^{l} a_i X^{(i)}(G_s), \tag{165}$$

où les a_i sont des entiers non négatifs qui indiquent combien de fois la représentation $\omega^{(i)}$ figure dans la composition de la représentation ω après qu'elle a été réduite.

On peut déterminer les formules qui donnent les coefficients a_i suivant les caractères de la représentation ω. Soit k l'un des nombres $1, 2, \ldots, l$. On multiplie les deux membres de (165) par $\overline{X^{(k)}(G_s)}$ et on somme sur s. En utilisant (163) et (164), on obtient:

$$\sum_{s=1}^{m} X(G_s)\overline{X^{(k)}(G_s)}=a_k m,$$

d'où :

$$a_k=\frac{1}{m}\sum_{s=1}^{m} X(G_s)\overline{X^{(k)}(G_s)}. \tag{166}$$

Cette formule donne pour tout a_k une valeur déterminée, d'où résulte le théorème suivant:

Théorème 2. *Toute représentation réductible se décompose en un ensemble unique de représentations irréductibles.*

En utilisant la formule (166), on généralise le théorème 1 au cas de représentations quelconques et non seulement irréductibles.

Théorème 3. *La condition nécessaire et suffisante pour que deux représentations soient équivalentes est que tous leurs caractères coïncident.*

On a déjà noté le fait que la condition est nécessaire en démontrant le théorème 1. Réciproquement, si les caractères $X(G_s)$ de deux représentations sont identiques, d'après les formules (166) on obtient des valeurs égales pour les nombres a_k, et par suite les deux représentations se réduisent à une matrice quasi diagonale formée de représentations irréductibles identiques. Passant alors, si cela est nécessaire, aux représentations équivalentes on peut supposer que les représentations irréductibles indiquées sont disposées dans une matrice quasi diagonale suivant un ordre identique, car une même permutation des lignes et des colonnes est équivalente au passage à une représentation équivalente.

Par là même les représentations ayant les mêmes caractères se réduisent aux mêmes matrices quasi diagonales, c'est-à-dire qu'elles sont équivalentes.

Passons maintenant à l'examen du nombre l de toutes les représentations irréductibles, non équivalentes, du groupe G. Les éléments de ce groupe se distribuent en classes. Dans une même classe figurent tous les éléments qui s'obtiennent à partir d'un G_t au moyen de la formule:

$$G_sG_tG_s^{-1} \qquad (s=1, 2, \ldots, m).$$

A tous ces éléments correspondent, dans toute représentation, des matrices semblables ayant des traces identiques. Soit r le nombre de classes du groupe G. D'après ce qui a été dit ci-dessus, toute représentation linéaire du groupe G n'a pas plus de r caractères différents, et chaque caractère correspond non à un élément isolé mais à l'ensemble des éléments qui figurent dans une classe. Supposons la classe C_1 formée de g_1 éléments, la classe C_2 de g_2 éléments et enfin la classe C_r de g_r éléments. Les composantes de la somme (163) sont identiques pour les éléments d'une même classe, donc désignant par $X(C_k)$, $X'(C_k)$ les caractères correspondant aux éléments figurant dans la classe C_k, pour deux représentations irréductibles non équivalentes, on peut écrire (163) sous la forme:

$$\sum_{k=1}^{r} X(C_k)\overline{X'(C_k)}\,g_k=0$$

et (164) sous la forme :

$$\sum_{h=1}^{r} X(C_h)\overline{X(C_h)}\,g_h = m.$$

Ainsi, on a pour les caractères $X^{(i)}(C_h)$ des représentations irréductibles non équivalentes $\omega^{(i)}$ $(i=1, 2, \ldots, l)$:

$$\begin{aligned} &\sum_{h=1}^{r} X^{(i_1)}(C_h)\overline{X^{(i_2)}(C_h)}\,g_h = 0 \quad \text{si} \quad i_1 \neq i_2, \\ &\sum_{h=1}^{r} X^{(i)}(C_h)\overline{X^{(i)}(C_h)}\,g_h = m. \end{aligned} \tag{167}$$

Introduisons dans l'espace R_r, à r dimensions, l vecteurs de composantes :

$$\sqrt{\frac{g_1}{m}}X^{(i)}(C_1),\ \sqrt{\frac{g_2}{m}}X^{(i)}(C_2), \ldots, \sqrt{\frac{g_r}{m}}X^{(i)}(C_r)$$
$$(i=1, 2, \ldots, l).$$

Les égalités précédentes montrent que ces vecteurs sont orthogonaux deux à deux et normés, et donc linéairement indépendants. Il en résulte que leur nombre l n'est pas supérieur au nombre de dimensions, c'est-à-dire que $l \leqslant r$. D'où:

Théorème 4. *Le nombre de représentations irréductibles non équivalentes d'un groupe n'est pas supérieur au nombre de classes de ce groupe.*

Au paragraphe suivant nous démontrerons que l'on a toujours $l = r$. On a déjà démontré que $l \leqslant r$; pour démontrer que $l = r$ il suffit de démontrer que $l \geqslant r$. La démonstration de cette inégalité est liée à l'introduction de notions et relations nouvelles entre les caractères, ce qui, d'ailleurs, présente un intérêt en soi.

Examinons encore une relation entre les caractères d'une représentation irréductible quelconque. Supposons la classe C_h formée des éléments $G_1^{(h)}$, $G_2^{(h)}, \ldots, G_{g_h}^{(h)}$. Si G_s est un élément quelconque du groupe, les éléments $G_sG_i^{(h)}G_s^{-1}$ $(i = 1, 2, \ldots, g_h)$ donnent aussi tous les éléments de la classe C_h, mais dans un autre ordre. Il en résulte que si l'on prend l'ensemble de tous les produits d'éléments figurant dans des classes C_p et C_q:

$$G_u^{(p)}G_v^{(q)} \quad (u=1, 2, \ldots, g_p; \quad v=1, 2, \ldots, g_q), \tag{168}$$

alors l'ensemble des éléments

$$G_sG_u^{(p)}G_v^{(q)}G_s^{-1} = (G_sG_u^{(p)}G_s^{-1})\,(G_sG_v^{(q)}G_s^{-1})$$

sera le même. Il en résulte que l'ensemble des éléments (168) a la propriété suivante: si un élément quelconque appartient à cet ensemble, la classe contenant cet élément appartient toute entière à ce même ensemble et chaque élément de cette classe figure dans l'ensemble des éléments (168) un nombre identique de fois. *Désignons par a_{pqh} le nombre entier, non négatif, qui indique combien* *de fois les éléments de la classe C_k figurent dans l'ensemble des éléments* (168). On écrit cela de façon conventionnelle comme suit:

$$C_pC_q=\sum_{k=1}^{r} a_{pqk}C_k \tag{169}$$

ou:

$$(G_1^{(p)}+G_2^{(p)}+\ldots+G_{g_p}^{(p)})(G_1^{(q)}+G_2^{(q)}+\ldots+G_{g_q}^{(q)})=$$
$$=\sum_{k=1}^{r} a_{pqk}(G_1^{(k)}+G_2^{(k)}+\ldots+G_{g_k}^{(k)}). \tag{170}$$

Soient $A(G_s)$ les matrices d'ordre n d'une représentation linéaire irréductible du groupe G. Formons la somme des matrices correspondant aux éléments de la classe C_k et désignons cette matrice par $A(C_k)$:

$$A(C_k)=\sum_{j=1}^{g_k} A(G_j^{(k)})$$

De ce que les éléments $G_sG_i^{(k)}G_s^{-1}$ pour $i=1, 2, \ldots, g_k$ et G_s quelconque de G donnent tout l'ensemble des éléments de la classe C_k, on voit que la matrice $A(C_k)$ commute avec toutes les matrices $A(G_s)$. Cette matrice $A(C_k)$ est donc un multiple de la matrice unité [III-2-2] et l'on peut écrire:

$$A(C_k)=b_kI \qquad (k=1, 2, \ldots, r), \tag{171}$$

où b_k sont des nombres quelconques. De la définition des nombres a_{pqk}, c'est-à-dire de la formule symbolique (170), on obtient la relation suivante entre les nombres b_k:

$$b_pb_q=\sum_{k=1}^{r} a_{pqk}b_k. \tag{172}$$

La trace de la matrice $A(C_k)$ est égale à la somme des traces des matrices $A(G_i^{(k)})$ $(i=1, 2, \ldots, g_k)$, c'est-à-dire à $g_kX(C_k)$. D'autre part, il résulte de (171) que la trace de $A(C_k)$ est égale à nb_k, c'est-à-dire que $nb_k=g_kX(C_k)$, d'où:

$$b_k=\frac{g_k}{n}X(C_k),$$

et la relation (172) donne le théorème suivant.

Théorème 5. *Les caractères de toute représentation irréductible formée à l'aide de matrices d'ordre n sont liés par la relation*:

$$g_pX(C_p)g_qX(C_q)=n\sum_{k=1}^{r} a_{pqk}g_kX(C_k). \tag{173}$$

Remarquons que parmi les classes C_k il est une, formée d'un élément unité E du groupe G. Dans toute représentation linéaire il lui correspond la matrice unité dont la trace est égale à l'ordre n. On désignera toujours cette classe par C_1, de sorte que $X(C_1)=n$ et que la formule précédente peut s'écrire sous la forme:

$$g_pX(C_p)g_qX(C_q)=X(C_1)\sum_{k=1}^{r} a_{pqk}g_kX(C_k). \tag{174}$$

Déterminons maintenant la valeur des constantes a_{pq1}. A chaque classe C_p correspond une classe $C_{p'}$ formée des éléments inverses de ceux de C_p. Ceci résulte de la définition des classes et de ce que la formule $G_s G_i G_s^{-1} = G_u$ donne $G_s G_i^{-1} G_s^{-1} = G_u^{-1}$.

La classe $C_{p'}$ peut coïncider avec C_p, c'est-à-dire qu'il peut arriver que $p' = p$. En tout cas, les classes C_p et $C_{p'}$ contiennent un même nombre d'éléments: $g_{p'} = g_p$. Si dans les formules (173) ou (174) on prend $q = p'$, la classe C_1 figure g_p fois dans le second membre et si l'on prend $q \neq p'$, le second membre ne contient pas C_1, c'est-à-dire que:

$$a_{pq1} = \begin{cases} 0 & \text{si} \quad q \neq p', \\ g_p & \text{si} \quad q = p'. \end{cases} \tag{175}$$

III-2-14. Représentations régulières des groupes. On a déjà donné une méthode qui permet de représenter n'importe quel groupe fini au moyen du groupe des permutations. Tout groupe des permutations peut être représenté comme un groupe des transformations.

En effet, si on a, par exemple, la permutation:

$$\begin{matrix} 1, & 2, & 3, & 4 \\ 2, & 4, & 3, & 1, \end{matrix}$$

on peut l'écrire sous forme d'une transformation linéaire qui fait passer x_1 en y_2, x_2 en y_4, x_3 en y_3 et x_4 en y_1:

$$\begin{aligned} y_1 &= 0x_1 + 0x_2 + 0x_3 + x_4, \\ y_2 &= x_1 + 0x_2 + 0x_3 + 0x_4, \\ y_3 &= 0x_1 + 0x_2 + x_3 + 0x_4, \\ y_4 &= 0x_1 + x_2 + 0x_3 + 0x_4. \end{aligned}$$

Considérons la représentation suivante du groupe G au moyen du groupe des permutations. On multiplie à droite les éléments $G_1, G_2, \ldots, G_m$ par un élément G_s. Ceci conduit à une permutation des éléments, c'est-à-dire, compte tenu de ce qui a été dit, à une certaine matrice P_s que l'on considère comme correspondant à l'élément G_s. Cette représentation est la représentation régulière du groupe G. Un des éléments, G_1, est l'élément unité du groupe. On le désigne par E. A cet élément correspond la matrice unité P_1 et, par conséquent, la trace de cette matrice est égale à m, c'est-à-dire que $X(E) = m$. En multipliant les éléments $G_1, G_2, \ldots, G_m$ par un élément quelconque G_s aucun des éléments G_k ne reste en place, c'est-à-dire que dans la matrice correspondante tous les éléments diagonaux sont nuls, et dans la représentation régulière $X(G_s) = 0$ si $G_s \neq E$.

Supposons que pendant la réduction de la représentation régulière elle contienne h_k fois la représentation $\omega^{(k)}$, dont il a déjà été question. On a alors, compte tenu de ce qui a déjà été dit:

$$\sum_{i=1}^{l} h_i X^{(i)}(G_s) = \begin{cases} 0 & \text{si} \quad G_s \neq E, \\ m & \text{si} \quad G_s = E. \end{cases} \tag{176}$$

En multipliant les deux membres de cette égalité par $\overline{X^{(k)}(G_s)}$ et on sommant sur s, on obtient, compte tenu de (163) et (164):

$$h_k m = m X^{(k)}(E),$$

mais $X^{(k)}(E)$ est égal à l'ordre des matrices dans la représentation $\omega^{(k)}$ que l'on a désigné par n_k et $\overline{X^{(k)}(E)} = X^{(k)}(E) = h_k$, d'où il résulte que $h_k = n_k$, et la formule (176) peut s'écrire:

$$\sum_{t=1}^{l} X^{(t)}(E)\, X^{(t)}(G_s) = \sum_{t=1}^{l} n_t X^{(t)}(G_s) = \begin{cases} 0 & \text{si } G_s \neq E, \\ m & \text{si } G_s = E. \end{cases} \tag{177}$$

On a donc le théorème suivant:

Théorème 6. *Une représentation régulière contient chaque représentation irréductible $\omega^{(k)}$ un nombre de fois égal à l'ordre n_k des matrices dans la représentation $\omega^{(k)}$ et les caractères de la représentation $\omega^{(k)}$ vérifient la formule (177).*

Ecrivons maintenant la formule (174) pour les représentations $\omega^{(t)}$:

$$g_p X^{(t)}(C_p)\, g_q X^{(t)}(C_q) = X^{(t)}(C_1) \sum_{h=1}^{r} a_{pqh} X^{(t)}(C_h)\, g_h,$$

et sommons sur t de $t=1$ à $t=l$:

$$g_p g_q \sum_{t=1}^{l} X^{(t)}(C_p)\, X^{(t)}(C_q) = \sum_{h=1}^{r} a_{pqh} \sum_{t=1}^{l} X^{(t)}(C_1)\, X^{(t)}(C_h)\, g_h.$$

En tenant compte de (177), on obtient:

$$g_p g_q \sum_{t=1}^{l} X^{(t)}(C_p)\, X^{(t)}(C_q) = a_{pq1} m,$$

c'est-à-dire en vertu de (175):

$$\sum_{t=1}^{l} X^{(t)}(C_p)\, X^{(t)}(C_q) = \begin{cases} 0 & \text{pour } q \neq p', \\ \dfrac{m}{g_{p'}} & \text{pour } q = p'. \end{cases} \tag{178}$$

Formons l équations linéaires homogènes par rapport à $x_1, x_2, \ldots, x_r$:

$$\sum_{q=1}^{r} x_q X^{(k)}(C_q) = 0 \qquad (k = 1, 2, \ldots, l) \tag{179}$$

et montrons que le système obtenu n'a d'autres solutions qu'une solution *triviale*.

En effet, multipliant les deux membres de (179) par $X^{(k)}(C_p)$ et sommant sur k, on obtient $x_{p'} = 0$, p' étant n'importe quel nombre de la suite $1, 2, \ldots, r$. Comme le système (179) n'a d'autres solutions qu'une solution *triviale*, le nombre d'équations du système n'est pas inférieur au nombre d'inconnues, c'est-à-dire que $l \geqslant r$. On a déjà démontré que $l \leqslant r$, d'où il résulte que $l = r$, c'est-à-dire:

Théorème 7. *Le nombre total de représentations irréductibles, non équivalentes, d'un groupe fini G est égal au nombre de classes de ce groupe.*

Examinons encore une conséquence du théorème 6. Une représentation régulière du groupe G est formée de matrices d'ordre m. D'autre part, en vertu du théorème 6, elle contient n_k fois chaque représentation $\omega^{(k)}$ formée de matrices d'ordre n_k.

Il en résulte l'égalité:

$$\sum_{k=1}^{r} n_k^2 = m, \tag{180}$$

que l'on peut formuler ainsi:

Théorème 8. *La somme des carrés des ordres des représentations irréductibles, non équivalentes, $\omega^{(k)}$ est égale à l'ordre du groupe G.*

III-2-15. Exemples de représentations de groupes finis. 1. Considérons le groupe abélien G formé des éléments $A_2^k A_1^i$, où $i = 0, 1, 2, \ldots, m-1$; $k = 0, 1, 2, \ldots, n-1$ et les éléments A_1 et A_2 commutent. On a $A_1^m = E$ et $A_2^n = E$ et si $i = k = 0$, on prend $A_2^0 A_1^0 = E$. Chaque élément de G forme une classe et toutes les représentations irréductibles du groupe sont du premier ordre. Soient α et β des valeurs quelconques des racines de degré m et n de l'unité. Associons le nombre $\beta^k\alpha^i$ à l'élément $A_2^k A_1^i$. Il est facile de vérifier qu'on obtient ainsi une représentation du groupe. En donnant à α et β toutes les valeurs possibles des racines indiquées de l'unité, on obtient en tout mn représentations différentes du premier ordre. Le nombre de classes (c'est-à-dire d'éléments) est aussi mn et par là même on a construit toutes les représentations irréductibles non équivalentes. On peut de la même façon construire les représentations dont le nombre d'éléments « générateurs » (c'est-à-dire d'éléments A_s) est supérieur à deux.

2. Passons au groupe diédral d'ordre n. Il est formé de $2n$ éléments:

$$E,\ A^i,\ T,\ TA^i \qquad (i = 1, 2, \ldots, n-1),$$

et:

$$A^n = E;\quad T^2 = E;\quad TAT^{-1} = A^{-1} \quad (T^{-1} = T). \tag{181}$$

La dernière des relations écrites est évidente si on considère le sens géométrique des rotations A et T. Il résulte de cette relation que $TA^iT^{-1} = A^{-i}$. On fait, tout d'abord, $n = 2m + 1$. Le groupe est alors formé de $(m + 2)$ classes. L'une contient E, m classes contiennent deux éléments, A^s et A^{-s} $(s = 1, 2, \ldots, m)$, de plus $A^{-s} = A^{2l+1-s}$, et une classe contient tous les éléments de la forme T et TA^i. Cela est facile à vérifier à partir des relations indiquées.

On a deux représentations du premier ordre. Dans l'une à chaque élément on associe le nombre 1. Dans l'autre, à l'élément A on associe 1 et à l'élément T (-1). Soit de plus $\varepsilon = \cos\dfrac{2\pi}{n} + i\sin\dfrac{2\pi}{n}$. On peut construire m représentations du deuxième ordre, en associant aux éléments A et T les matrices:

$$A \to \begin{Vmatrix} \varepsilon^s & 0 \\ 0 & \varepsilon^{-s} \end{Vmatrix};\qquad T \to \begin{Vmatrix} 0 & 1 \\ 1 & 0 \end{Vmatrix} \qquad (s = 1, 2, \ldots, m). \tag{182}$$

Les matrices écrites vérifient les relations (181) et, par là même, donnent une représentation du groupe, car toute relation entre les éléments A et T découle de (181). Chaque représentation est irréductible, car, dans le cas contraire, elle pourrait être ramenée à deux représentations du premier ordre et les matrices de la représentation considérée devraient commuter, ce qui n'est pas le cas, quel que soit s, comme il est facile de le vérifier.

Le fait que les représentations (182) ne soient pas équivalentes, pour des s différents, résulte de ce que pour différents s les matrices qui correspondent à l'élément A ont des ensembles de valeurs propres ε^s et ε^{-s} différents. Ainsi, on a construit les $(m + 2)$ représentations irréductibles non équivalentes. La formule (180), dans le cas considéré, se réduit à:

$$2\cdot 1^2 + m\cdot 2^2 = 4m + 2 = 2n.$$

Si $n=2m$, la représentation (182) correspond à la valeur $s=m$, elle est de la forme :

$$A \to \begin{Vmatrix} -1 & 0 \\ 0 & -1 \end{Vmatrix}; \qquad T \to \begin{Vmatrix} 0 & 1 \\ 1 & 0 \end{Vmatrix}$$

et se décompose en deux représentations du premier ordre :

$$A \to (-1); \quad T \to (+1) \quad \text{et} \quad A \to (-1); \quad T \to (-1).$$

Pour l'obtenir, il suffit d'appliquer la matrice S de manière que STS^{-1} soit diagonale, les valeurs propres de la matrice T sont alors égales à ± 1. Ainsi, si $n = 2m$, on a quatre représentations du premier ordre et $(m-1)$ représentations du deuxième ordre. La formule (180) devient alors:

$$4 \cdot 1^2 + (m-1) 2^2 = 4m = 2n.$$

3. Considérons les représentations du groupe du tétraèdre ou, ce qui revient au même, du groupe alterné qui lui est isomorphe pour $n = 4$ [III-1-8]. Le groupe est formé de quatre classes et son ordre est égal à 12. Il doit avoir quatre représentations irréductibles non équivalentes. Les ordres de ces représentations doivent vérifier l'égalité :

$$n_1^2 + n_2^2 + n_3^2 + n_4^2 = 12.$$

Cette équation a, à l'ordre des composantes dans le premier membre près, une seule solution en nombres entiers positifs :

$$n_1 = n_2 = n_3 = 1; \qquad n_4 = 3,$$

c'est-à-dire que le groupe a trois représentations du premier ordre et une représentation du troisième ordre. Dans les représentations du premier ordre, aux éléments qui figurent dans une seule et même classe correspond un seul et même nombre. Il est aisé de montrer que dans les trois représentations du premier ordre les nombres suivants correspondent aux classes:

$$\begin{array}{llll} \mathrm{I} \to 1; & \mathrm{II} \to 1; & \mathrm{III} \to 1; & \mathrm{IV} \to 1; \\ \mathrm{I} \to 1; & \mathrm{II} \to 1; & \mathrm{III} \to \varepsilon; & \mathrm{IV} \to \varepsilon^2; \\ \mathrm{I} \to 1; & \mathrm{II} \to 1; & \mathrm{III} \to \varepsilon^2; & \mathrm{IV} \to \varepsilon, \end{array}$$

où :

$$\varepsilon = \cos\frac{2\pi}{3} + i \sin\frac{2\pi}{3}.$$

La représentation irréductible du troisième ordre est donnée par le groupe du tétraèdre lui-même, c'est-à-dire le groupe des rotations de l'espace (matrices du troisième ordre) qui transforment le tétraèdre en lui-même. Si cette représentation était réductible, elle devrait donner trois représentations du premier ordre, ce qui est impossible, ne serait-ce que parce que le groupe du tétraèdre n'est pas un groupe abélien. La théorie exposée dans ces derniers paragraphes concerne les groupes finis. Pour la transposer au groupe des rotations on doit étudier plus en détail les groupes infinis dépendant de paramètres. Avant de passer à l'examen général de ces groupes on étudiera le problème des représentations linéaires du groupe de Lorentz. Ces représentations, à côté des représentations du groupe des rotations, serviront d'exemples de base pour l'étude des groupes infinis dépendant de paramètres.

III-2-16. Représentations linéaires des groupes à deux variables. Nous avons construit les représentations linéaires d'un groupe unitaire à deux variables [III-2-4], ce qui nous a conduit aux représentations linéaires du groupe des rotations. Nous pouvons de façon analogue construire les représentations linéaires des groupes à deux variables, à déterminant égal à l'unité:

$$x_1' = ax_1 + bx_2, \qquad ad - bc = 1. \tag{183}$$
$$x_2' = cx_1 + dx_2,$$

Ceci conduit, compte tenu de ce qui a été dit [III-1-13], aux représentations à une ou deux déterminations du groupe des transformations de Lorentz positives. Les résultats sont sensiblement différents des résultats de [III-2-4].

L'une des représentations linéaires possibles du groupe unitaire (93) est la représentation de ce groupe par lui-même, c'est-à-dire une représentation linéaire qui à chaque transformation (93) fait correspondre la même transformation. Il est facile de voir qu'une autre représentation linéaire est celle où à chaque transformation (93) on associe une transformation dont les coefficients sont les conjugués complexes:

$$y_1' = \bar{a}y_1 + \bar{b}y_2; \qquad y_2' = -by_1 + ay_2.$$

Mais cette représentation est équivalente à la précédente, ce qui résulte, comme il est facile de le voir, de la formule:

$$\begin{Vmatrix} 0 & 1 \\ -1 & 0 \end{Vmatrix} \begin{Vmatrix} a & b \\ -\bar{b} & \bar{a} \end{Vmatrix} = \begin{Vmatrix} \bar{a} & \bar{b} \\ -b & a \end{Vmatrix} \begin{Vmatrix} 0 & 1 \\ -1 & 0 \end{Vmatrix}.$$

La représentation conjuguée du groupe (183)

$$y_1' = \bar{a}y_1 + \bar{b}y_2; \qquad y_2' = \bar{c}y_1 + \bar{d}y_2 \tag{184}$$

n'est pas équivalente au groupe (183) lui-même. Pour le vérifier, il suffit de considérer le cas où $b = c = 0$. Alors la matrice de la transformation (183) a comme valeurs propres a et d et la matrice (184) $\bar{a}$ et $\bar{d}$. Il est évident que l'on peut choisir les nombres complexes a et d vérifiant la condition $ad = 1$, de telle sorte que l'ensemble des nombres $\bar{a}$ et $\bar{d}$ soit différent de l'ensemble des nombres a et d et que par suite les transformations correspondantes ne puissent être des similitudes. Ainsi on a dans ce cas déjà deux représentations non équivalentes du second ordre: le groupe (183) lui-même et le groupe (184). On verra plus loin que ces représentations sont irréductibles.

On peut de plus construire une représentation du groupe (183) exactement comme on l'a fait en [III-2-4]. Dans les formules (99), il suffit simplement de remplacer $\bar{a}$ par d et $\bar{b}$ par $(-c)$. Ceci donne une représentation d'ordre $(2j+1)$, où j est un entier non négatif ou un demi-entier:

$$D_j \begin{Bmatrix} a & b \\ c & d \end{Bmatrix}_{ls} = \sum_k \frac{\sqrt{(j+l)!\,(j-l)!\,(j+s)!\,(j-s)!}}{k!\,(j-k-s)!\,(j+l-k)!\,(k+s-l)!} \times$$

$$\times a^{j+l-k} b^k c^{k+s-l} d^{j-k-s} \left(j=0,\ \frac{1}{2},\ 1,\ \frac{3}{2},\ \ldots\right). \qquad (185)$$

Ici l et s prennent toutes les valeurs:

$$l \text{ et } s = -j,\ -j+1,\ \ldots,\ j-1,\ j,$$

et la sommation sur k est définie par les inégalités:

$$k>0; \quad k \geqslant l-s; \quad k \leqslant j-s; \quad k \leqslant j+l.$$

Dans les formules (185) posons $0! = 1$ et $0^0 = 1$. Si $j=0$, on obtient la représentation identique de l'unité. De (185) on déduit immédiatement d'autres représentations en remplaçant, dans les seconds membres de (185), les nombres a, b, c et d par leurs conjugués. On désigne les représentations correspondantes par:

$$\overline{D}_{j'} \begin{Bmatrix} a & b \\ c & d \end{Bmatrix} \left(j'=0,\ \frac{1}{2},\ 1,\ \frac{3}{2},\ \ldots\right). \qquad (186)$$

Formons alors la composition des représentations (185) et (186) [III-2-9]. Cela permet d'obtenir une nouvelle représentation d'ordre $(2j+1)(2j'+1)$ qu'on désigne par:

$$E_{j,j'} \begin{Bmatrix} a & b \\ c & d \end{Bmatrix}. \qquad (187)$$

Des formules (185), il est facile d'écrire les éléments des matrices correspondant à cette représentation. On prend deux représentations différentes (187) mais de même ordre:

$$E_{p,q} \begin{Bmatrix} a & b \\ c & d \end{Bmatrix} \text{ et } E_{p_1,q_1} \begin{Bmatrix} a & b \\ c & d \end{Bmatrix};$$

$$(2p+1)(2q+1) = (2p_1+1)(2q_1+1).$$

Montrons que ces deux représentations ne sont pas équivalentes. Supposons $b=c=0$. Alors la matrice (185) se réduit à une matrice diagonale d'éléments diagonaux:

$$D_j \begin{Bmatrix} a & 0 \\ 0 & d \end{Bmatrix}_{ll} = a^{j+l} d^{j-l} \quad (l=-j,\ -j+1,\ \ldots,\ j-1,\ j).$$

Le produit direct de deux matrices diagonales est une matrice diagonale et, par conséquent, les matrices $E_{p,q}$ et E_{p_1,q_1}, pour $b=c=0$, ont comme valeurs propres:

$$E_{p,q}: a^{p+l}d^{p-l}(\bar{a})^{q+m}(\bar{d})^{q-m}\begin{pmatrix} l=-p,\ -p+1,\ \ldots,\ p-1,\ p \\ m=-q,\ -q+1,\ \ldots,\ q-1,\ q \end{pmatrix}$$

$$E_{p_1,q_1}: a^{p_1+l_1}d^{p_1-l_1}(\bar{a})^{q_1+m_1}(\bar{d})^{q_1-m_1}\begin{pmatrix} l_1=-p_1,\ -p_1+1,\ \ldots \\ \ldots,\ p_1-1,\ p_1 \\ m_1=-q_1,\ -q_1+1,\ \ldots \\ \ldots,\ q_1-1,\ q_1 \end{pmatrix}$$

ou, en tenant compte de ce que $ad=1$:

$$E_{p,q}: a^{2l}(\bar{a})^{2m}; \qquad E_{p_1,q_1}: a^{2l_1}(\bar{a})^{2m_1}.$$

On peut prendre pour a n'importe quel nombre complexe différent de 0 et évidemment choisir ce nombre de façon que l'ensemble des valeurs propres de la matrice $E_{p,q}$ soit différent de l'ensemble des valeurs propres de la matrice E_{p_1,q_1}, ce qui montre que les représentations (187) ne sont pas équivalentes, si j et j' sont différents. Remarquons que si $j'=0$, la représentation (187) coïncide avec la représentation (185) et si $j=0$, elle coïncide avec la représentation que l'on obtient à partir de (185) pour $j=j'$ et lorsque l'on remplace a, b, c et d par les quantités conjuguées. Remarquons une particularité des représentations (187). Ces représentations ne sont pas équivalentes aux représentations unitaires. Si elles étaient équivalentes à des représentations unitaires, toutes les valeurs propres de n'importe quelle matrice de la représentation devraient avoir le module égal à l'unité. Mais on a vu que pour $b=c=0$ les valeurs propres de la représentation $E_{p,q}$ sont égales à $a^{2l}(\bar{a})^{2m}$ et peuvent avoir le module différent de l'unité. La seule exception est le cas de la représentation identique triviale $E_{0,0}$ pour laquelle à chaque élément du groupe (183) correspond l'unité.

On a vu [III-2-2] que si une représentation, pas nécessairement équivalente à une représentation unitaire, est réductible, c'est-à-dire équivalente à une représentation ayant des matrices quasi diagonales de même structure, il existe nécessairement une matrice non multiple de la matrice unité et qui commute avec toutes les matrices de la représentation. Ainsi, pour démontrer que n'importe quelle représentation (187) n'est pas réductible, il suffit de démontrer que si une matrice commute avec toutes les matrices de la représentation (187), cette matrice est un multiple de la matrice unité. Ceci peut être fait comme en [III-2-4]. Ainsi, les représentations (187) ne sont pas équivalentes deux à deux, et chacune d'elles est une représentation irréductible. On utilise souvent une définition de la réductibilité différente de celle donnée en [III-1-7], à savoir: une représentation est dite réductible si toutes ses transformations linéaires (qu'on suppose d'ordre n) laissent invariant un certain sous-espace L_k et que $0 < k < n$.

On a vu [III-1-7] que si une représentation est réductible dans ce dernier sens et si elle est formée de matrices unitaires, elle est aussi réductible au sens de la définition de [III-2-1], c'est-à-dire équivalente à une représentation quasi diagonale. Si la représentation n'est pas unitaire, alors, de l'invariance d'un certain sous-espace il ne découle pas que la représentation est réductible au sens de la définition de [III-2-1]. On peut montrer que toute représentation (187) du groupe est irréductible en ce sens, comme on l'a déjà démontré, et que de plus elle ne laisse invariant aucun sous-espace. En outre, on peut montrer que toute représentation linéaire du groupe (183) est soit équivalente à une représentation (187), soit équivalente à une représentation de formule réduite constituée de plusieurs représentations (187).

On a vu [III-2-9] que la composition de deux représentations linéaires du groupe est équivalente au produit des êtres des représentations linéaires qui entrent dans cette composition. On déduit de ceci que les êtres des représentations (187) sont les expressions:

$$\eta_{kk'} = \frac{x_1^{j+k} x_2^{j-k}}{\sqrt{(j+k)!\,(j-k)!}} \cdot \frac{y_1^{j'+k'} y_2^{j'-k'}}{\sqrt{(j'+k')!\,(j'-k')!}}$$

$$\begin{pmatrix} k = j,\ j-1,\ \ldots,\ -j+1,\ -j \\ k' = j',\ j'-1,\ \ldots,\ -j'+1,\ -j' \end{pmatrix},$$

où x_1 et x_2 subissent la transformation (183) et y_1 et y_2 la transformation (184).

On n'a parlé jusque-là que des représentations linéaires du groupe des transformations de Lorentz positives [III-1-13]. Ces transformations positives ne forment qu'une partie des transformations de Lorentz à déterminant égal à l'unité. Il existe en outre des transformations de Lorentz à déterminant égal à (—1). L'étude de la structure de ces ensembles plus généraux de transformations et l'extension des représentations linéaires du groupe des transformations de Lorentz positives au cas du groupe complet de Lorentz présentent certaines particularités par rapport au groupe des transformations orthogonales dans l'espace à trois dimensions. Remarquons que dans la définition du groupe complet de Lorentz on peut imposer comme condition que le sens du temps reste inchangé. Il faut alors ajouter au groupe de Lorentz considéré la réflexion:

$$x_1' = -x_1;\quad x_2' = -x_2;\quad x_3' = -x_3;\quad x_4' = x_4.$$

On trouvera l'étude de ces questions, par exemple dans les ouvrages de Cartan: « Théorie des spineurs » et de Van der Waerden: « Die gruppentheoretische Methode in der Quantenmechanik ».

III-2-17. Le groupe de Lorentz est un groupe simple. En utilisant une méthode analogue à celle de [III-2-6], démontrons maintenant que le groupe de Lorentz est un groupe simple. Pour cela, il suffit de montrer que le groupe G, formé des transformations (183), n'a d'autres sous-groupes distingués que les matrices E et $(-E)$. Supposons qu'existe un tel sous-groupe distingué H_1 comprenant la matrice:

$$A=\begin{Vmatrix} a & b \\ c & d \end{Vmatrix} \qquad (ad-bc=1),$$

différente de E et de $(-E)$. Nous devons démontrer que H_1 coïncide avec G. Si H_1 contient une matrice B, le sous-groupe H_1 contient toutes les matrices $U^{-1}BU$ (où U est une matrice quelconque de G). En tenant compte du résultat fondamental de la réduction des matrices à la forme canonique, ainsi que du fait que le déterminant de la matrice U, qui réduit une matrice quelconque à sa forme canonique, peut toujours être considéré égal à l'unité [II-1-8], on voit qu'il suffit de montrer que H_1 contient les matrices de valeurs propres quelconques distinctes t et t^{-1}, où t est un nombre complexe quelconque différent de 0 et de $(+1)$. Remarquons que le produit des valeurs propres des matrices du groupe G doit être égal à l'unité, que H_1 doit contenir les matrices E et $(-E)$ et que dans le cas de valeurs propres égales et de double diviseur élémentaire, H_1 doit contenir les matrices

$$\begin{Vmatrix} 1 & 0 \\ 1 & 1 \end{Vmatrix} \quad \text{et} \quad \begin{Vmatrix} -1 & 0 \\ 1 & -1 \end{Vmatrix}. \qquad (188)$$

Prenons une matrice variable du groupe G:

$$X=\begin{Vmatrix} x & y \\ z & x \end{Vmatrix} \qquad (x^2-yz=1)$$

et formons la matrice:

$$Y=A(XA^{-1}X^{-1}),$$

laquelle, d'après [III-2-6], doit appartenir à H_1. La trace s de la matrice Y est de la forme:

$$s=2+bz^2+cy^2-[(a-d)^2+2bc]\,yz.$$

Puisque A est différent de E et de $(-E)$, on n'aura pas simultanément $b=c=0$ et $a=d$. D'où il découle que s n'est pas une constante et qu'en faisant varier z et y on peut donner à s n'importe quelle valeur complexe. Les valeurs propres de la matrice Y sont définies à partir de l'équation du second degré:

$$\lambda^2 - s\lambda + 1 = 0.$$

Ainsi, on peut obtenir pour ces racines n'importe quelle valeur t et t^{-1}, par conséquent, H_1 contient toutes les matrices à valeurs propres différentes et déterminant égal à l'unité. De plus, H_1 contient E et $(-E)$ que l'on peut représenter sous la forme du produit:

$$-E = [t, t^{-1}] \cdot [-t^{-1}, -t],$$

chacun des facteurs appartenant à H_1. Les matrices (188) peuvent aisément être représentées sous la forme d'un produit de deux matrices à déterminants égaux à l'unité et valeurs propres distinctes, d'où il résulte que H_1 les contient également. En effet,

$$\begin{Vmatrix} 1 & 0 \\ 1 & 1 \end{Vmatrix} = \begin{Vmatrix} \frac{1}{\beta} & 0 \\ 0 & \beta \end{Vmatrix} \cdot \begin{Vmatrix} \beta & 0 \\ \frac{1}{\beta} & \frac{1}{\beta} \end{Vmatrix},$$

$$\begin{Vmatrix} -1 & 0 \\ 0 & -1 \end{Vmatrix} = \begin{Vmatrix} \frac{1}{\beta} & 0 \\ 0 & \beta \end{Vmatrix} \cdot \begin{Vmatrix} -\beta & 0 \\ \frac{1}{\beta} & -\frac{1}{\beta} \end{Vmatrix} \qquad (\beta \neq 0, \pm 1).$$

Ainsi, on a démontré que H_1 coïncide avec G, c'est-à-dire que G n'a pas de sous-groupes distingués différents de celui formé par E et $(-E)$ et, par là même, on a démontré que le groupe des transformations de Lorentz positives est un groupe simple. Il en résulte [III-2-6] que ce groupe ne peut avoir de représentations homomorphes ou non isomorphes.

III-3. Groupes continus

III-3-1. Groupes continus. Constantes de structure. Le groupe des rotations de l'espace à trois dimensions et celui des transformations de Lorentz positives constituent des exemples de groupes infinis dont les éléments dépendent de paramètres qui peuvent varier continûment. Pour le groupe des rotations, le paramètre est, par exemple, donné par les angles d'Euler. Dans les cas considérés, les groupes sont formés de transformations linéaires et la relation entre les groupes et les paramètres se réduit au fait que les éléments des matrices qui définissent les transformations linéaires sont fonctions de ces paramètres. Dans la suite nous ne considérerons que des groupes de transformations linéaires.

Supposons que les éléments a_{ik} des matrices des transformations linéaires qui constituent un groupe G soient fonctions de r paramètres réels $\alpha_1, \alpha_2, \ldots, \alpha_r$ et que de plus les conditions suivantes soient vérifiées: les a_{ik} sont des fonctions univoques des paramètres α_s pour toutes les valeurs de ces paramètres, suffisamment voisines de zéro, et aux valeurs nulles des paramètres $\alpha_1 = \alpha_2 = \ldots = \alpha_r = 0$ correspond l'élément unité du groupe G qui est caractérisé par les conditions $a_{ik} = 0$, si $i \neq k$ et $a_{ii} = 1$. Supposons de plus qu'à tout élément du groupe G suffisamment voisin de l'élément unité correspondent des valeurs définies des paramètres α_s suffisamment voisines de zéro. La proximité des éléments du groupe de l'élément unité est définie par le fait que les éléments a_{ik} des matrices correspondantes soient voisins de zéro si $i \neq k$ et de l'unité si $i = k$.

Ainsi, avec les hypothèses faites, on a une correspondance biunivoque entre les éléments du groupe G qui se trouvent dans un certain voisinage de l'élément unité et un certain voisinage de l'origine des coordonnées dans l'espace réel T_r à r dimensions. Dans la suite on aura non seulement une telle correspondance biunivoque locale mais une correspondance biunivoque globale telle qu'à chaque élément du groupe G corresponde un point déterminé de l'espace T_r appartenant à un domaine V de cet espace contenant l'origine et, inversement, qu'à tout point de V corresponde un élément défini du groupe G. Pour l'instant, la correspondance locale indiquée nous suffit. Désignons par les symboles G_α, G_β, G_γ, etc., les éléments du groupe G qui correspondent aux valeurs α_s, β_s, γ_s $(s = 1, 2, \ldots, r)$ des paramètres. Du point de vue local, il faut considérer que les paramètres sont suffisamment voisins de zéro et que les éléments du groupe sont suffisamment voisins de l'élément unité.

Considérons le produit d'éléments quelconques du groupe

$$G_\beta G_\alpha = G_\gamma.$$

Les paramètres γ_s qui caractérisent l'élément G_γ obtenu au moyen de la multiplication indiquée sont des fonctions univoques des paramètres α_s et β_s:

$$\gamma_s = \varphi_s(\beta_1, \beta_2, \ldots, \beta_r; \quad \alpha_1, \alpha_2, \ldots, \alpha_r). \tag{189}$$

Supposons que ce soient des fonctions continues ayant des dérivées continues jusqu'au quatrième ordre pour tous les α_s et β_s suffisamment voisins de zéro.

L'élément unité du groupe correspondant aux valeurs nulles des paramètres, il en résulte les égalités:

$$\left.\begin{aligned} &\varphi_s(\beta_1, \beta_2, \ldots, \beta_r; 0, 0, \ldots, 0) = \beta_s; \\ &\varphi_s(0, 0, \ldots, 0; \alpha_1, \alpha_2, \ldots, \alpha_r) = \alpha_s, \end{aligned}\right. \quad (s = 1, 2, \ldots, r), \tag{190}$$

d'où :

$$\left.\begin{aligned} \frac{\partial\varphi_l}{\partial\beta_k} &= \delta_{lk} \quad \text{si} \quad \alpha_s = 0\,; \\ \frac{\partial\varphi_l}{\partial\alpha_k} &= \delta_{lk} \quad \text{si} \quad \beta_s = 0 \end{aligned}\right\} \quad (s = 1, 2, \ldots, r). \tag{191}$$

Les paramètres $\tilde{\alpha}_s$ qui correspondent à l'élément inverse G_α^{-1} sont définis par les relations :

$$\varphi_s(\tilde{\alpha}_1, \tilde{\alpha}_2, \ldots, \tilde{\alpha}_r\,;\ \alpha_1, \alpha_2, \ldots, \alpha_r) = 0 \quad (s = 1, 2, \ldots, r), \tag{192}$$

et les égalités écrites ont lieu si l'on suppose tous les α_s et tous les $\tilde{\alpha}_s$ nuls. Le déterminant fonctionnel des premiers membres des équations (192) par rapport à $\tilde{\alpha}_s$ est égal, compte tenu de (191), à l'unité si α_s et $\tilde{\alpha}_s$ sont nuls. Ainsi, en vertu du théorème sur les fonctions implicites, les équations (192) définissent les $\tilde{\alpha}_s$ comme des fonctions continues pour tous les α_s suffisamment voisins de zéro et les $\tilde{\alpha}_s$ sont nuls si $\alpha_s = 0$. Développons les fonctions (189) en séries entières de α_s et de β_s en utilisant la formule de Maclaurin et poussons le développement jusqu'aux termes du troisième ordre. En tenant compte des formules (190) et (191), on obtient :

$$\gamma_s = \alpha_s + \beta_s + \sum_{i,k} a_{i,k}^{(s)}\alpha_i\beta_k + \sum_{i,k,l} a_{i,k,l}^{(s)}\alpha_i\alpha_k\beta_l + \\ + \sum_{i,k,l} b_{i,k,l}^{(s)}\alpha_i\beta_k\beta_l + \varepsilon^{(s)}, \tag{193}$$

où $a_{i,k}^{(s)}$, $a_{i,k,l}^{(s)}$ et $b_{i,k,l}^{(s)}$ sont des coefficients numériques, $\varepsilon^{(s)}$ est un infiniment petit au moins du quatrième ordre par rapport à α_s et β_s et la sommation sur i, k et l s'effectue de 1 jusqu'à r. Les nombres

$$C_{ik}^{(s)} = a_{ik}^{(s)} - a_{ki}^{(s)} \qquad (s, i, k = 1, 2, \ldots, r) \tag{194}$$

sont les *constantes de structure du groupe G* pour les paramètres choisis α_s.

Si on remplace α_s par d'autres paramètres α'_s :

$$\alpha_s = \omega_s(\alpha'_1, \alpha'_2, \ldots, \alpha'_r) \qquad (s = 1, 2, \ldots, r)$$

de sorte que $\omega_s(0, 0, \ldots, 0) = 0$, les équations écrites sont solubles d'une manière unique par rapport à α'_s pour tous les α_s suffisamment voisins de zéro et les fonctions ω_s ont un nombre suffisant de dérivées, les constantes de structure pour les nouveaux paramètres α'_s seront différentes.

Il résulte de la définition (194) que:

$$C_{ki}^{(s)} = -C_{ik}^{(s)}. \qquad (194_1)$$

De plus, d'après (192) et tenant compte de ce que le produit des éléments du groupe G est associatif, on peut aussi démontrer la relation suivante entre les constantes de structure:

$$\sum_{s=1}^{r} (C_{is}^{(t)} C_{jk}^{(s)} + C_{js}^{(t)} C_{ki}^{(s)} + C_{ks}^{(t)} C_{ij}^{(s)}) = 0 \quad (i, j, k, t = 1, 2, \ldots, r). \qquad (194_2)$$

Nous n'utiliserons pas cette relation et n'en donnerons pas la démonstration.

Revenons aux formules (193). Si α_s et β_s sont suffisamment voisins de zéro, les quantités γ_s sont voisines de zéro aussi. D'après la formule (191) et le théorème sur les fonctions implicites, on peut affirmer que les équations (193) dans un certain voisinage de l'origine des coordonnées de l'espace T_r peuvent être résolues par rapport à β_s:

$$\beta_s = \psi_s(\gamma_1, \gamma_2, \ldots, \gamma_r; \quad \alpha_1, \alpha_2, \ldots, \alpha_r) \quad (s = 1, 2, \ldots, r). \qquad (195)$$

Remarquons en passant que les conditions: $\beta_s = 0$ ($s = 1, 2, \ldots, r$) sont équivalentes aux conditions $\gamma_s = \alpha_s$ ($s = 1, 2, \ldots, r$). En utilisant les formules (193) et (195), on forme deux matrices carrées $S(\alpha_s)$ et $T(\alpha_s)$ d'ordre r dont les éléments $S_{ik}(\alpha_s)$ et $T_{ik}(\alpha_s)$ dépendent des paramètres α_s:

$$S_{ik}(\alpha_s) = \left(\frac{\partial \gamma_i}{\partial \beta_k}\right)_{\beta_s = 0}; \quad T_{ik}(\alpha_s) = \left(\frac{\partial \beta_i}{\partial \gamma_k}\right)_{\gamma_s = \alpha_s}$$
$$(s, i, k = 1, 2, \ldots, r). \qquad (196)$$

D'après la règle de différentiation des fonctions de fonction et en calculant la dérivée de γ_i par rapport à γ_k ou la dérivée de β_i par rapport à β_k, on obtient:

$$S(\alpha_s) T(\alpha_s) = E \text{ et } T(\alpha_s) S(\alpha_s) = E, \qquad (197)$$

où E est la matrice unité d'ordre r. Il résulte des formules (191) que $S(\alpha_s)$ pour les $\alpha_s = 0$ se transforme en la matrice unité. Et de (197) on déduit que $T(\alpha_s)$ a la même propriété. Les constantes de structure $C_{ik}^{(p)}$ s'expriment en fonction des éléments de ces matrices par

$$C_{ik}^{(p)} = \left(\frac{\partial S_{pk}(\alpha_s)}{\partial \alpha_i} - \frac{\partial S_{pi}(\alpha_s)}{\partial \alpha_k}\right)_{\alpha_s = 0} \qquad (198)$$

soit:

$$C_{ik}^{(p)} = \left(\frac{\partial T_{pi}(\alpha_s)}{\partial \alpha_k} - \frac{\partial T_{pk}(\alpha_s)}{\partial \alpha_i}\right)_{\alpha_s = 0}. \qquad (199)$$

En effet, compte tenu de (193) et (196), on obtient:

$$a_{ik}^{(p)} = \left(\frac{\partial\gamma_p}{\partial\alpha_i\,\partial\beta_k}\right)_{\alpha_s=\beta_s=0} = \left(\frac{\partial S_{pk}(\alpha_s)}{\partial\alpha_i}\right)_{\alpha_s=0}, \tag{200}$$

et en permutant les indices i et k, on peut écrire:

$$a_{ki}^{(p)} = \left(\frac{\partial S_{pi}(\alpha_s)}{\partial\alpha_k}\right)_{\alpha_s=0}, \tag{201}$$

d'où il résulte la formule (198). De plus, en tenant compte de (197), on a:

$$\sum_{j=1}^{r} S_{pj}(\alpha_s)\, T_{jk}(\alpha_s) = \delta_{pk}.$$

Différentions les deux membres par rapport à α_i et supposons ensuite que tous les α_s soient nuls. Les matrices $S(\alpha_s)$ et $T(\alpha_s)$ se transformant en la matrice unité si $\alpha_s = 0$ $(s = 1, 2, \ldots, r)$, on obtient:

$$\left(\frac{\partial S_{pk}(\alpha_s)}{\partial\alpha_i}\right)_{\alpha_s=0} + \left(\frac{\partial T_{pk}(\alpha_s)}{\partial\alpha_i}\right)_{\alpha_s=0} = 0,$$

c'est-à-dire que, compte tenu de (201):

$$a_{ik}^{(p)} = -\left(\frac{\partial T_{pk}(\alpha_s)}{\partial\alpha_i}\right)_{\alpha_s=0},$$

d'où il résulte la formule (199). Les formules (193) définissent l'opération fondamentale de groupe qui à partir des paramètres α_s et β_s des éléments G_α et G_β du groupe G donne les paramètres γ_s correspondant au produit $G_\beta G_\alpha$. On voit, à partir des expressions (193), que si α_s et β_s sont voisins de zéro, en première approximation l'opération de groupe se réduit à $\gamma_s = \alpha_s + \beta_s$, de sorte qu'en première approximation le groupe est abélien. Si le groupe est réellement abélien, on a

$$\varphi_s(\beta_1\beta_2, \ldots, \beta_r;\ \alpha_1\alpha_2, \ldots, \alpha_r) = \\ = \varphi_s(\alpha_1\alpha_2, \ldots, \alpha_r;\ \beta_1\beta_2, \ldots, \beta_r) \qquad (s = 1, 2, \ldots, r),$$

et dans les développements (193) $a_{ki}^{(s)} = a_{ik}^{(s)}$, c'est-à-dire que pour un groupe abélien toutes les constantes de structure sont nulles. Pour les groupes généraux les termes du second ordre dans le développement (193) donnent déjà un écart par rapport à la commutativité, ce qui se caractérise par la présence de constantes de structure non nulles. En utilisant le développement (193), il est aisé d'obtenir le développement des paramètres $\tilde{\alpha}_s$ correspondant à l'élément $G_{\tilde{\alpha}}$. Pour cela il suffit de faire $\gamma_s = 0$ dans les formules (193) et de remplacer β_s par $\tilde{\alpha}_s$. En utilisant les règles habituelles de différentiation des fonctions implicites, on obtient:

$$\tilde{\alpha}_s = -\alpha_s + \sum_{i,k} a_{ik}^{(s)} \alpha_i \alpha_k + \varepsilon_1^{(s)},$$

où $\varepsilon_1^{(s)}$ est un infiniment petit au moins du troisième ordre par rapport à $\alpha_1, \alpha_2, \ldots, \alpha_r$.

III-3-2. Transformations infinitésimales. Soit un groupe continu G de transformations linéaires d'ordre n, défini par les paramètres α_s ($s = 1, 2, \ldots, r$). Comme ci-dessus, désignons par G_α la matrice des transformations correspondant aux paramètres α_s et telle que la transformation linéaire est de la forme:

$$\mathbf{x} = G_\alpha \mathbf{u}, \tag{202}$$

où $\mathbf{u}$ est un vecteur quelconque de l'espace complexe R_n à n dimensions et $\mathbf{x}$ le vecteur transformé. On introduit la différentiation d'une matrice; si les éléments d'une matrice A sont des fonctions différentiables par rapport à un paramètre t, on appellera dérivée de la matrice A, par rapport à t, la matrice dont les éléments s'obtiennent en différentiant les éléments de la matrice A par rapport à t, c'est-à-dire

$$\left\{\frac{dA}{dt}\right\}_{ik} = \frac{d\{A\}_{ik}}{dt}.$$

Si les éléments de A dépendent de plusieurs variables, on aura des dérivées partielles.

De même, si les composantes d'un vecteur $\mathbf{z}$ ($z_1, z_2, \ldots, z_n$) de l'espace R_n sont des fonctions différentiables par rapport à t, le vecteur $\frac{d\mathbf{z}}{dt}$ est défini comme le vecteur dont les composantes sont $\frac{dz_i}{dt}$, c'est-à-dire que la différentiation d'un vecteur se réduit à la différentiation de ses composantes (tome II, [IV-2-1]).

On introduit les *transformations infinitésimales du groupe* G:

$$I_k = \left(\frac{\partial G_\alpha}{\partial \alpha_k}\right)_{\alpha_s=0} \quad (k = 1, 2, \ldots, r), \tag{203}$$

où I_k désigne une matrice d'ordre n dont les éléments sont des nombres.

Revenons à la formule (202) et supposons que $\mathbf{u}$ est un vecteur fixe, c'est-à-dire que ses composantes ne dépendent pas de α_s. Le vecteur transformé dépend en général de ces paramètres et on obtient alors des équations différentielles fondamentales pour ce vecteur.

Pour cela on applique aux deux membres de l'équation (202) l'opération linéaire définie par la matrice G_β :

$$G_\beta \mathbf{x} = G_\gamma \mathbf{u},$$

où $G_\gamma = G_\beta G_\alpha$, et les paramètres γ_s s'obtiennent en fonction de α_s et β_s, d'après l'opération fondamentale de groupe (193). On différentie les deux membres de cette dernière formule par rapport à β_p et on fait ensuite $\beta_s = 0$, c'est-à-dire $\gamma_s = \alpha_s$. D'après la définition (203), on obtient :

$$I_p \mathbf{x} = \sum_{j=1}^{r} \left[\frac{\partial (G_\gamma \mathbf{u})}{\partial \gamma_j}\right]_{\gamma_s = \alpha_s} \left(\frac{\partial \gamma_j}{\partial \beta_p}\right)_{\beta_s = 0}.$$

Le premier facteur sous le signe somme est égal à la dérivée du second membre de (202) par rapport à α_j et, en tenant compte des notations (196), on peut écrire la dernière formule sous la forme :

$$I_p \mathbf{x} = \sum_{j=1}^{r} S_{jp}(\alpha_s) \frac{\partial \mathbf{x}}{\partial \alpha_j} \qquad (p = 1, 2, \ldots, r).$$

Si l'on introduit les vecteurs :

$$X\left(\frac{\partial \mathbf{x}}{\partial \alpha_1}, \frac{\partial \mathbf{x}}{\partial \alpha_2}, \ldots, \frac{\partial \mathbf{x}}{\partial \alpha_r}\right) \quad \text{et} \quad Y(I_1 \mathbf{x}, I_2 \mathbf{x}, \ldots, I_r \mathbf{x}),$$

les formules précédentes peuvent s'écrire sous forme d'une transformation linéaire :

$$Y = S^*(\alpha_s) X,$$

où $S^*(\alpha_s)$ désigne la matrice transposée. En multipliant à gauche par $T^*(\alpha_s)$ et en tenant compte de (197), on obtient :

$$X = T^*(\alpha_s) Y,$$

ou en développant :

$$\frac{\partial \mathbf{x}}{\partial \alpha_p} = \sum_{j=1}^{r} T_{jp}(\alpha_s) I_j \mathbf{x} \qquad (p = 1, 2, \ldots, r). \tag{204}$$

La composante x_k du vecteur $\mathbf{x}$ défini par la formule (202) vérifie l'équation :

$$\frac{\partial x_k}{\partial \alpha_p} = \sum_{j=1}^{r} T_{jp}(\alpha_s) \sum_{l=1}^{n} \{I_j\}_{kl} x_l \quad \begin{pmatrix} k = 1, 2, \ldots, n \\ p = 1, 2, \ldots, r \end{pmatrix}, \tag{205}$$

où $\{I_j\}_{kl}$ sont les composantes de la matrice I_j. Il faut ajouter à l'équation (204) pour $\mathbf{x}$ la condition initiale qui résulte immédiatement de la formule (202) :

$$\mathbf{x}\,|_{\alpha_s = 0} = \mathbf{u}, \tag{206}$$

où u est un vecteur arbitraire donné. Remarquons que les $T_{jp}(\alpha_s)$ qui figurent dans les coefficients de l'équation (204) sont définis directement par l'opération de groupe (193). L'équation (204) donne les relations entre les I_j. Pour cela, il suffit d'écrire que la dérivée seconde de x par rapport à α_p et α_q ne dépend pas de l'ordre de la différentiation.

Il résulte de (204) que

$$\frac{\partial^2 x}{\partial\alpha_p\,\partial\alpha_q}=\sum_{j=1}^{r}\left(\frac{\partial T_{jp}(\alpha_s)}{\partial\alpha_q}I_j x+T_{jp}(\alpha_s)\,I_j\,\frac{\partial x}{\partial\alpha_q}\right)$$

ou, en remplaçant $\frac{\partial x}{\partial\alpha_q}$ par son expression (204), pour $p=q$, on obtient :

$$\frac{\partial^2 x}{\partial\alpha_p\,\partial\alpha_q}=\sum_{j=1}^{r}\frac{\partial T_{jp}(\alpha_s)}{\partial\alpha_q}I_j x+\sum_{j=1}^{r}\sum_{k=1}^{r}T_{jp}(\alpha_s)\,T_{kq}(\alpha_s)\,I_j I_k x.$$

En permutant p et q dans le second membre et en égalant le second membre ainsi obtenu à celui de la formule précédente, on a :

$$\Big[\sum_{j=1}^{r}\Big(\frac{\partial T_{jp}(\alpha_s)}{\partial\alpha_q}-\frac{\partial T_{jq}(\alpha_s)}{\partial\alpha_p}\Big)I_j+\sum_{j=1}^{r}\sum_{k=1}^{r}\Big(T_{jp}(\alpha_s)\,T_{kq}(\alpha_s)-$$

$$-T_{jq}(\alpha_s)\,T_{kp}(\alpha_s)\Big)\,I_j I_k\Big]x=0. \qquad (207)$$

Supposons dans cette relation tous les α_s nuls. D'après (199) et le fait que $T(\alpha_s)=E$, si tous les α_s sont nuls, on obtient :

$$\Big[\sum_{j=1}^{r}C_{pq}^{(j)}I_j+(I_pI_q-I_qI_p)\Big]u=0,$$

d'où il résulte, compte tenu de ce que le vecteur u est arbitraire, les relations suivantes entre les transformations infinitésimales :

$$I_qI_p-I_pI_q=\sum_{j=1}^{r}C_{pq}^{(j)}I_j \quad (p,q=1,2,\ldots,r). \qquad (208)$$

Nous avons déterminé les I_j et démontré les relations (208) en partant d'un groupe continu donné G et en utilisant l'équation (204). Montrons que cette équation ou, ce qui revient au même, le système (205) a une solution unique pour la condition initiale donnée (206). Supposons deux solutions possibles. Compte tenu du fait que l'équation (204) est linéaire, leur différence doit aussi vérifier l'équation et se transformer en vecteur nul, si $\alpha_s = 0$. Ainsi, il faut montrer que la solution $\mathbf{x}$ de l'équation (204), pour des conditions initiales nulles, est identiquement nulle. Pour simplifier les notations, supposons $r = 3$. Soit $\mathbf{x}\ (\alpha_1, \alpha_2, \alpha_3)$ la solution indiquée. Ecrivons l'équation (204) pour $p = 1$ et faisons dans le second membre $\alpha_2 = \alpha_3 = 0$. On obtient une équation différentielle ordinaire par rapport à la variable indépendante α_1 et pour des conditions initiales nulles. Compte tenu du théorème d'unicité (tome II, [II-3-1]), elle est identiquement nulle, c'est-à-dire que $\mathbf{x}\ (\alpha_1, 0, 0) \equiv \equiv 0$. Ecrivons l'équation (204) pour $p = 2$ et faisons $\alpha_3 = 0$ dans le second membre. Cette équation différentielle ordinaire par rapport à la variable indépendante α_2 a, comme on vient de le montrer, pour des conditions initiales nulles la solution: $\mathbf{x}\ (\alpha_1, \alpha_2, 0) = 0$ pour $\alpha_2 = 0$. Ecrivons l'équation (204) pour $p = 3$. Cette équation différentielle ordinaire pour des conditions initiales: $\mathbf{x}\ (\alpha_1, \alpha_2, \alpha_3) = 0$ pour $\alpha_3 = 0$, a la solution : $\mathbf{x} \equiv (\alpha_1, \alpha_2, \alpha_3) \equiv 0$, ce qu'on voulait démontrer.

Ainsi, l'équation (204) peut conduire seulement à une transformation finie (202) pour des transformations infinitésimales I_j et des $T_{jp}\ (\alpha_s)$ données, définies à partir de l'opération de groupe (193). Autrement dit, les transformations infinitésimales forment un groupe. Ceci est essentiel pour la suite. On démontre l'existence d'une solution unique de l'équation (204) à partir d'un théorème général des équations aux dérivées partielles qui, appliqué à l'équation (204), s'énonce comme suit: pour que l'équation (204) ait une solution, quelle que soit la condition initiale (206), il faut et il suffit que le terme entre crochets dans la formule (207) soit identiquement nul par rapport à α_s quels que soient p et q. Par la suite on n'utilisera pas ce théorème d'existence.

III-3-3. Groupe des rotations. Considérons comme exemple le groupe des rotations de l'espace autour de l'origine O. Les matrices correspondantes du troisième ordre dépendent de trois paramètres. Ces paramètres peuvent être, par exemple, les trois angles d'Euler. Introduisons alors d'autres paramètres α_1, α_2, α_3 avec lesquels on effectuera tous les calculs ultérieurs. Toute rotation peut être considérée comme une rotation autour d'un axe orienté l, passant par l'origine des coordonnées, d'un angle non supérieur à π dans le sens inverse des aiguilles d'une montre. Alors deux rotations d'angle π par rapport à des axes opposés conduisent à une même position finale. On peut ainsi représenter toute rotation par un vecteur issu de l'origine des coordonnées, dirigé suivant l'axe de rotation et dont la longueur est égale à l'angle de rotation. Les projections de ce vecteur sur les axes de coordonnées sont les paramètres $(\alpha_1, \alpha_2, \alpha_3)$.

Si nous prenons la sphère V de centre à l'origine et de rayon π, et si nous identifions les extrémités de l'un quelconque de ses diamètres, il existe entre les points $(\alpha_1, \alpha_2, \alpha_3)$ de la sphère V et les éléments du groupe des rotations une correspondance biunivoque. Ici, cette correspondance n'aura pas lieu seulement au voisinage de l'origine des coordonnées et de l'élément unité du groupe mais pour tout le groupe si l'on prend toute la sphère V. On peut exprimer toutes les matrices qui figurent dans le groupe des rotations en fonction des paramètres $\alpha_1, \alpha_2, \alpha_3$ et vérifier ainsi la continuité et l'existence de dérivées, dont nous avons parlé ci-dessus.

Nous n'établirons pas dans ce cas la formule (193) pour l'opération fondamentale de groupe mais nous déterminerons les constantes de structure en calculant directement les matrices des transformations infinitésimales.

Pour calculer I_1 supposons $\alpha_2 = \alpha_3 = 0$. On différentie la matrice de la transformation par rapport à α_1, et on fait ensuite $\alpha_1 = 0$. Mais si $\alpha_2 = \alpha_3 = 0$, on a une rotation d'angle α_1 autour de l'axe X, d'où :

$$\begin{aligned} x_1' &= x_1, \\ x_2' &= x_2 \cos\alpha_1 - x_3 \sin\alpha_1, \\ x_3' &= x_2 \sin\alpha_1 + x_3 \cos\alpha_1. \end{aligned}$$

En différentiant la matrice de cette transformation par rapport à α_1, puis en faisant $\alpha_1 = 0$, il vient :

$$I_1 = \begin{Vmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{Vmatrix}. \tag{209}$$

Et de même :

$$I_2 = \begin{Vmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{Vmatrix}; \qquad I_3 = \begin{Vmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{Vmatrix}. \tag{210}$$

Après quoi on peut calculer directement les premiers membres des relations (208) et, par là même, déterminer les constantes de structure. Ce calcul élémentaire conduit aux trois relations suivantes :

$$I_1 I_2 - I_2 I_1 = I_3; \quad I_2 I_3 - I_3 I_2 = I_1; \quad I_3 I_1 - I_1 I_3 = I_2. \tag{211}$$

Si le second membre de l'équation (202) est développé en série entière de α_s, en se limitant aux termes du premier ordre, on obtient avec cette précision :

$$\mathrm{x} \doteq \mathrm{u} + (\alpha_1 I_1 + \alpha_2 I_2 + \alpha_3 I_3)\, \mathrm{u}.$$

Ainsi, le vecteur u subit par suite de cette transformation la variation :

$$\delta \mathbf{u} \doteq \alpha_1 I_1 \mathbf{u} + \alpha_2 I_2 \mathbf{u} + \alpha_3 I_3 \mathbf{u}.$$

Chaque composante du second membre donne la variation de u pour une petite rotation autour de l'un des axes de coordonnées. Ainsi, lorsque le vecteur u tourne d'un petit angle α_1, autour de l'axe X, on obtient les variations suivantes des composantes (u_1, u_2, u_3) :

$$\delta u_1 \doteq 0 \ ; \quad \delta u_2 \doteq -u_3 \alpha_1 \ ; \quad \delta u_3 \doteq u_2 \alpha_1.$$

Ce faisant on se limite aux termes du premier ordre par rapport à α_1.

III-3-4. Transformations infinitésimales et représentations du groupe des rotations. Etudions maintenant la relation entre les transformations infinitésimales et les représentations du groupe des rotations. Supposons l'existence d'une représentation biunivoque au voisinage de la transformation identique par des matrices $F(\alpha_1, \alpha_2, \alpha_3)$ d'ordre n, et les éléments de ces matrices sont considérés comme des fonctions continues et différentiables des paramètres $\alpha_1, \alpha_2, \alpha_3$. Toute rotation D peut être obtenue sous forme du produit d'un nombre fini de rotations dans le voisinage donné et le produit des matrices correspondantes de la représentation donne des représentations de D. Mais ainsi on peut obtenir une représentation multivoque du groupe des rotations puisqu'en changeant continûment les paramètres de la rotation, on peut revenir à la rotation initiale et obtenir pour elle une nouvelle représentation. C'était le cas pour la représentation à deux déterminations du groupe complet des rotations [III-2-5].

On a la même opération de groupe pour les matrices $F(\alpha_1, \alpha_2, \alpha_3)$ et pour les rotations elles-mêmes et, par suite, les mêmes constantes de structure. Le groupe G' des matrices $F(\alpha_1, \alpha_2, \alpha_3)$ peut être formé à partir de transformations infinitésimales I_k. Ce sont des matrices d'ordre n liées par les relations (211). S'il existe des matrices I_k, on peut écrire pour le vecteur $\mathbf{x}$ de R_n les équations différentielles (204), car les $T_{jp}(\alpha_s)$ ne sont définies que par l'opération de groupe. Ces équations ne peuvent, pour des conditions initiales données (206), avoir qu'une solution ; mais cette solution peut évidemment être la transformation :

$$\mathbf{x} = F(\alpha_1, \alpha_2, \alpha_3)\, \mathbf{u},$$

qui donne une représentation du groupe des rotations au voisinage de la transformation identique.

Dans le cas considéré $r = 3$ et si dans l'équation (204) on passe aux composantes du vecteur $\mathbf{x}$, on obtient $3n$ équations pour les n composantes du vecteur

$$\mathbf{x}(x_1, x_2, \ldots, x_n).$$

Dans la suite, il sera important seulement de savoir que l'équation (204) ne peut pas avoir plus d'une solution pour les conditions initiales (206) données. On a déjà dit que ceci peut être énoncé de la façon suivante : *la représentation du groupe des rotations est entièrement déterminée par ses transformations infinitésimales I_1, I_2, I_3.*

Ainsi, tout se réduit à la détermination des transformations infinitésimales de représentations. Au lieu des matrices cherchées I_1, I_2, I_3, on introduit les nouvelles matrices :

$$A_1 = -I_2 + iI_1; \quad A_2 = I_2 + iI_1; \quad A_3 = iI_3. \tag{212}$$

Il est aisé de vérifier qu'au lieu de (211) on obtient les relations suivantes :

$$\left.\begin{aligned} A_3A_1 - A_1A_3 &= A_1, \\ A_3A_2 - A_2A_3 &= -A_2, \\ A_1A_2 - A_2A_1 &= 2A_3. \end{aligned}\right\} \tag{213}$$

Dans la représentation au moyen des matrices $F(\alpha_1, \alpha_2, \alpha_3)$, on doit avoir, en particulier, la représentation du groupe abélien des rotations autour de l'axe Z dont les éléments correspondent aux matrices $F(0, 0, \alpha_3)$. En choisissant convenablement les vecteurs de base, toutes ces matrices peuvent être simultanément mises sous forme diagonale, car les représentations irréductibles du groupe abélien sont des représentations du premier ordre. Pour les vecteurs qui jouent dans ce cas le rôle de vecteurs de base, la transformation $F(0, 0, \alpha_3)$ est de la forme [III-2-5] :

$$F(0, 0, \alpha_3)\mathbf{v} = e^{l\alpha_3}\mathbf{v}$$

soit en faisant $l \doteq -im$ et en remplaçant $\mathbf{v}$ par $\mathbf{v}_m$:

$$F(0, 0, \alpha_3)\mathbf{v}_m = e^{-im\alpha_3}\mathbf{v}_m.$$

La représentation étant choisie univoque seulement au voisinage de $\alpha_s = 0$ $(s = 1, 2, 3)$, on ne doit pas considérer m comme entier. D'après la définition de I_3 on en déduit que :

$$A_3\mathbf{v}_m \doteq iI_3\mathbf{v}_m = i\left[\frac{\partial}{\partial\alpha_3}F(0, 0, \alpha_3)\mathbf{v}_m\right] = i\left(\frac{\partial}{\partial\alpha_3}e^{-im\alpha_3}\mathbf{v}_m\right) = m\mathbf{v}_m.$$

Ainsi :

$$A_3\mathbf{v}_m = m\mathbf{v}_m, \tag{214}$$

c'est-à-dire que $\mathbf{v}_m$ est un vecteur propre de l'opérateur A_3 correspondant à la valeur propre m. S'il y a plusieurs vecteurs propres, on désigne par $\mathbf{v}_m$ l'un d'entre eux.

Démontrons à présent le lemme suivant :

Lemme. *Si un vecteur* $\mathbf{v}$ *est un vecteur propre de l'opérateur* A_3 *correspondant à la valeur propre* a, *le vecteur* $A_1\mathbf{v}$, *s'il n'est pas nul, est aussi vecteur propre de* A_3 *et correspond à la valeur propre* $(a+1)$, *et de même* $A_2\mathbf{v}$ *est vecteur propre de* A_3 *et correspond à la valeur propre* $(a-1)$.

Par hypothèse du lemme $A_3\mathbf{v} = a\mathbf{v}$ et compte tenu de (213) :

$$A_3(A_1\mathbf{v}) = (A_1A_3 + A_1)\mathbf{v} = A_1(A_3\mathbf{v}) + A_1\mathbf{v} = A_1(a\mathbf{v}) + A_1\mathbf{v} = (a+1)A_1\mathbf{v},$$

et de même $A_3(A_2\mathbf{v}) = (a-1)A_2\mathbf{v}$.

Le nombre de valeurs propres différentes de A_3 n'est pas supérieur à n. Parmi ces valeurs, une ou plusieurs ont une partie réelle maximum. On désigne ces valeurs propres ou l'une d'entre elles par j, et soit $\mathbf{v}_j$ le vecteur propre correspondant (l'un d'entre eux s'il y en a plusieurs). D'après le lemme, le vecteur $A_1\mathbf{v}_j$ doit être associé à la valeur propre $(j+1)$, mais il résulte de la définition de j que A_3 n'a pas de valeur propre de ce type et, par conséquent, on doit avoir :

$$A_1\mathbf{v}_j = 0. \tag{215}$$

D'après ce qui a été démontré, les vecteurs :

$$\mathbf{v}_{j-1} = A_2\mathbf{v}_j;\quad \mathbf{v}_{j-2} = A_2\mathbf{v}_{j-1};\ \ldots, \tag{216}$$

s'ils ne sont pas nuls, correspondent aux valeurs propres $(j-1)$, $(j-2)$, ... de l'opérateur A_3. La suite des vecteurs (216) doit bien sûr aboutir à un vecteur nul, dans la mesure où le nombre de valeurs propres différentes de A_3 n'est pas supérieur à n. Démontrons maintenant que

$$A_1\mathbf{v}_k = \rho_k\mathbf{v}_{k+1} \qquad (k = j,\ j-1,\ j-2,\ \ldots), \tag{217}$$

où les ρ_k sont des entiers. D'après (215), cette formule est vraie si $k = j$, de plus, $\rho_j = 0$ et pour $\mathbf{v}_{j+1}$ on peut prendre, par exemple, le vecteur nul. Supposons maintenant la formule (217) vraie pour un k quelconque et démontrons qu'elle est vraie pour $(k-1)$. Compte tenu de (213), (216) et (217), on a :

$$A_1\mathbf{v}_{k-1} = A_1(A_2\mathbf{v}_k) = (A_2A_1 + 2A_3)\mathbf{v}_k =$$
$$= A_2(A_1\mathbf{v}_k) + 2A_3\mathbf{v}_k = A_2(\rho_k\mathbf{v}_{k+1}) + 2k\mathbf{v}_k = (\rho_k + 2k)\mathbf{v}_k.$$

Notons que si $k = j$, on n'utilise pas la formule :

$$A_2\mathbf{v}_{k+1} = \mathbf{v}_k,$$

car $\rho_k=0$ pour $k=j$. Ainsi, la formule (217) est démontrée et les nombres ρ_k se déterminent à partir des relations:

$$\rho_{k-1}=\rho_k+2k;\quad \rho_j=0\quad (k=j,\ j-1,\ \ldots).$$

En effectuant les calculs successifs, on obtient:

$$\rho_k=j(j+1)-k(k+1),$$

c'est-à-dire:

$$A_1\mathbf{v}_k=[j(j+1)-k(k+1)]\,\mathbf{v}_{k+1}\quad (k=j,\ j-1,\ \ldots). \tag{218}$$

D'après ces équations on détermine l'indice s du premier des vecteurs (216), lequel est nul, c'est-à-dire tel que $\mathbf{v}_s=0$ et $\mathbf{v}_{s+1}\neq 0$. Il résulte alors de (217) que $\rho_s=0$, c'est-à-dire:

$$j(j+1)-s(s+1)=0.$$

Cette équation du second degré en s a les racines $s=j$ et $s=-(j+1)$. La valeur $s=j$ ne convient pas, car le vecteur $\mathbf{v}_j$ n'est pas nul et n'appartient pas à la suite (216). Ainsi dans la suite (216) les vecteurs

$$\mathbf{v}_j,\ \mathbf{v}_{j-1},\ \ldots,\ \mathbf{v}_{-j+1},\ \mathbf{v}_{-j} \tag{219}$$

ne sont pas nuls et $A_2\mathbf{v}_{-j}=0$. Le nombre de ces vecteurs est égal à $(2j+1)$, d'où on voit que j est soit un entier non négatif, soit un demi-entier positif. Si $2j+1=n$, on peut prendre les vecteurs (219) comme vecteurs de base dans l'espace R_n, si $2j+1<n$, ils forment dans R_n un sous-espace L_{2j+1}. Supposons ce cas. Chaque vecteur $\mathbf{v}_k$ de la suite (219) vérifie l'équation:

$$A_3\mathbf{v}_k=k\mathbf{v}_k\quad (k=j,\ j-1,\ \ldots,\ -j+1,\ -j).$$

On a $A_2\mathbf{v}_k=\mathbf{v}_{k-1}$, $\mathbf{v}_{j-1}=0$, et la formule (218). Par là même les opérateurs A_1, A_2 et A_3 transforment le sous-espace L_{2j+1} en lui-même et les formules définissent parfaitement les opérateurs dans ce sous-espace. De plus, il résulte des formules (216) et (218) qu'aucun sous-espace L_k situé dans L_{2j+1} pour lequel $0<k<2j+1$ n'est invariant si l'on applique les opérateurs A_1, A_2, A_3. Ayant déterminé A_k, on peut construire pour le sous-espace L_{2j+1} les équations (204) que doit vérifier le vecteur

$$\mathbf{x}=F_j(\alpha_1,\alpha_2,\alpha_3)\,\mathbf{u} \tag{220}$$

de la représentation cherchée dans le sous-espace L_{2j+1}. Cette représentation ne peut laisser invariant aucun sous-espace L_k appartenant à L_{2j+1}, c'est-à-dire que la représentation est irréductible dans L_{2j+1}, car s'il en était ainsi, tout A_s, d'après la définition, devrait laisser invariant L_k, ce qui n'est pas vrai, comme on vient de le voir. Si $2j+1=n$, les raisonnements effectués s'appliquent à tout R_n. Si $2j+1<n$, de la représentation générale on isole, dans R_n, une représentation irréductible au sens indiqué d'ordre $(2j+1)$, c'est-à-dire qui ne laisse invariant aucun sous-espace L_k pour $0<k<2j+1$. Il résulte de ces raisonnements qu'il n'existe qu'une représentation irréductible d'un ordre donné (aux représentations semblables près). On a déjà construit [III-2-5] des représentations unitaires irréductibles d'un ordre quelconque. Elles épuisent par là même toutes les représentations irréductibles possibles, et les représentations indiquées, obtenues à partir des opérateurs A_s construits dans L_{2j+1}, doivent leur être semblables.

Les vecteurs (219) peuvent être multipliés par des facteurs numériques arbitraires, non nuls. Alors dans les relations (216) et (218) apparaissent également des facteurs numériques. Les facteurs indiqués peuvent être choisis de sorte qu'on ait finalement les relations:

$$\left.\begin{aligned} A_1\mathrm{v}_k &= \sqrt{j(j+1)-k(k+1)}\,\mathrm{v}_{k+1}, \\ A_2\mathrm{v}_k &= \sqrt{j(j+1)-k(k-1)}\,\mathrm{v}_{k-1}, \\ A_3\mathrm{v}_k &= k\mathrm{v}_k, \end{aligned}\right\} \tag{221}$$

et $\mathrm{v}_{j+1}=0$, $\mathrm{v}_{-j-1}=0$.

Si on choisit ainsi les facteurs, on obtient les représentations que l'on a déjà construites [III-2-5] en partant de:

$$\eta_l = \frac{x_1^{j+l}x_2^{j-l}}{\sqrt{(j+l)!\,(j-l)!}}. \tag{222}$$

Cette méthode permet d'isoler, dans n'importe quelle représentation, les parties irréductibles. Tout se ramène à la recherche des vecteurs propres v_j de l'opérateur A_3 ayant la plus grande valeur propre et à la construction de (216).

III-3-5. Représentations du groupe de Lorentz. Considérons le groupe des transformations linéaires à déterminant égal à l'unité:

$$\begin{aligned} x_1' &= ax_1+bx_2, \\ x_2' &= cx_1+dx_2. \end{aligned} \qquad (ad-bc=1) \tag{223}$$

La matrice de transformation contient quatre coefficients complexes liés entre eux. Seules trois grandeurs complexes restent arbitraires, ce qui se réduit à six paramètres réels. Introduisons ces paramètres dans la notation de la matrice de transformation:

$$A = \left\| \begin{matrix} 1+\alpha_1+i\alpha_2 & \alpha_3+i\alpha_4 \\ \alpha_5+i\alpha_6 & d(\alpha_s) \end{matrix} \right\|, \tag{224}$$

où :

$$d(\alpha_s) = \frac{1 + (\alpha_3 + i\alpha_4)(\alpha_5 + i\alpha_6)}{1 + \alpha_1 + i\alpha_2}.$$

Nous obtenons six transformations infinitésimales I'_k faciles à construire. Par exemple, pour obtenir I'_1, il suffit d'annuler dans la matrice A tous les α_s à l'exception de α_1, de différentier la matrice par rapport à α_1 et ensuite de faire $\alpha_1 = 0$. Ainsi, on a :

$$I'_1 = \begin{Vmatrix} 1 & 0 \\ 0 & -1 \end{Vmatrix}; \quad I'_2 = \begin{Vmatrix} i & 0 \\ 0 & -i \end{Vmatrix}; \quad I'_3 = \begin{Vmatrix} 0 & 1 \\ 0 & 0 \end{Vmatrix};$$

$$I'_4 = \begin{Vmatrix} 0 & i \\ 0 & 0 \end{Vmatrix}; \quad I'_5 = \begin{Vmatrix} 0 & 0 \\ 1 & 0 \end{Vmatrix}; \quad I'_6 = \begin{Vmatrix} 0 & 0 \\ i & 0 \end{Vmatrix}.$$

Les constantes de structure $C^{(j)}_{pq}$ qui figurent dans les relations (208) doivent, par définition, être réelles et peuvent être obtenues à partir de ces relations :

$$I'_pI'_q - I'_qI'_p = \sum_{j=1}^{6} C^{(j)}_{pq} I'_j \quad (p < q;\ p, q = 1, 2, \ldots, 6).$$

A remarquer de plus qu'entre les matrices I'_j on n'a pas de relation linéaire (en dehors de la relation triviale) avec des coefficients réels, et on obtient donc les 15 égalités suivantes :

$$I'_1I'_3 - I'_3I'_1 = 2I'_3, \quad I'_1I'_4 - I'_4I'_1 = 2I'_4, \quad I'_2I'_3 - I'_3I'_2 = 2I'_4,$$
$$I'_1I'_5 - I'_5I'_1 = -2I'_5, \quad I'_1I'_6 - I'_6I'_1 = -2I'_6, \quad I'_2I'_5 - I'_5I'_2 = -2I'_6,$$
$$I'_3I'_5 - I'_5I'_3 = I'_1, \quad I'_3I'_6 - I'_6I'_3 = I'_2, \quad I'_4I'_5 - I'_5I'_4 = I'_2.$$
$$I'_2I'_4 - I'_4I'_2 = -2I'_3, \quad I'_1I'_2 - I'_2I'_1 = 0,$$
$$I'_2I'_6 - I'_6I'_2 = 2I'_5, \quad I'_3I'_4 - I'_4I'_3 = 0,$$
$$I'_4I'_6 - I'_6I'_4 = -I'_1, \quad I'_5I'_6 - I'_6I'_5 = 0,$$

Si $I_k\ (k = 1, 2, \ldots, 6)$ sont des transformations infinitésimales pour n'importe quelle représentation du groupe considéré, il existe entre elles aussi 15 relations :

$$I_pI_q - I_qI_p = \sum_{j=1}^{6} C^{(j)}_{pq} I_j,$$

avec les mêmes coefficients $C^{(j)}_{pq}$. Si nous introduisons les notations :

$$I_3 + iI_4 = 2A_1; \quad I_5 + iI_6 = 2A_2; \quad I_1 + iI_2 = 4A_3;$$
$$I_3 - iI_4 = 2B_1; \quad I_5 - iI_6 = 2B_2; \quad I_1 - iI_2 = 4B_3, \qquad (225)$$

les 15 relations s'écrivent sous la forme suivante :

$$A_pB_q - B_qA_p = 0 \qquad (p, q = 1, 2, 3), \qquad (226)$$

et de plus:

$$A_3A_1-A_1A_3=A_1,\quad A_3A_2-A_2A_3=-A_2,\quad A_1A_2-A_2A_1=2A_3, \tag{227}$$

$$B_3B_1-B_1B_3=B_1,\quad B_3B_2-B_2B_3=-B_2,\quad B_1B_2-B_2B_1=2B_3. \tag{228}$$

Remarquons que les relations (226) et (227) sont vérifiées de façon triviale pour les matrices I'_k, car alors $A'_k=0$ ($k=1, 2, 3$). Les relations (227) et (213) coïncident, et les raisonnements du paragraphe précédent restent valables. On applique les relations écrites aux transformations infinitésimales de n'importe quelle représentation linéaire du groupe (223). Si $\mathbf{v}_j$ est un vecteur propre correspondant à la plus grande valeur propre de l'opérateur A_3, on a $(2j+1)$ vecteurs propres $\mathbf{v}_k$ ($k=j, j-1, \ldots, -j+1, -j$) de l'opérateur A_3 qui se transforment au moyen des opérateurs A_1, A_2, A_3, compte tenu des formules (221), $\mathbf{v}_{j+1}=0$ et $\mathbf{v}_{-j-1}=0$. Soit $L^{(j)}$ un sous-espace formé de tous les vecteurs propres de l'opérateur A_3 correspondant à la valeur propre j. Montrons que si le vecteur $\mathbf{v}$ appartient à $L^{(j)}$, les vecteurs $B_q\mathbf{v}$ ($q=1, 2, 3$) eux aussi appartiennent à $L^{(j)}$. En effet, d'après (226):

$$A_3(B_q\mathbf{v})=B_q(A_3\mathbf{v})=B_q(j\mathbf{v})=jB_q\mathbf{v},$$

d'où il résulte que $B_q\mathbf{v}$ est un vecteur propre de A_3 correspondant à la valeur propre j (ou bien un vecteur nul), c'est-à-dire que $B_q\mathbf{v}$ appartient à $L^{(j)}$. On peut refaire pour ce sous-espace les raisonnements de [III-3-4] en remplaçant les opérateurs A_k par les opérateurs B_k. Par suite, on peut construire dans $L^{(j)}$ la série de vecteurs $\mathbf{v}_{jk'}$ ($k'=j', j'-1, \ldots, -j'+1, -j'$) qui se transforment, d'après (221), si on remplace j par j' et A_k par B_k. Tout vecteur $\mathbf{v}_{jk'}$ si on applique deux fois l'opération A_2, donne $(2j+1)$ vecteurs $\mathbf{v}_{kk'}$ ($k=j, j-1, \ldots, -j+1, -j$). Ainsi on obtient finalement $(2j+1)(2j'+1)$ vecteurs $\mathbf{v}_{kk'}$, pour lesquels on a les relations:

$$\left.\begin{aligned}
A_1\mathbf{v}_{kk'}&=\sqrt{j(j+1)-k(k+1)}\,\mathbf{v}_{k+1,k'},\\
A_2\mathbf{v}_{kk'}&=\sqrt{j(j+1)-k(k-1)}\,\mathbf{v}_{k-1,k'},\\
A_3\mathbf{v}_{kk'}&=k\mathbf{v}_{kk'},\\
B_1\mathbf{v}_{kk'}&=\sqrt{j(j+1)-k'(k'+1)}\,\mathbf{v}_{k,k'+1},\\
B_2\mathbf{v}_{kk'}&=\sqrt{j(j+1)-k'(k'-1)}\,\mathbf{v}_{k,k'-1},\\
B_3\mathbf{v}_{kk'}&=k'\mathbf{v}_{kk'}.
\end{aligned}\right\} \tag{229}$$

Ces formules définissent les opérateurs A_p et B_q dans l'espace à $(2j+1)(2j'+1)$ dimensions, et les formules (225) définissent les opérateurs I_k. L'équation (204) ne peut alors conduire qu'à une seule représentation linéaire du groupe. C'est la représentation que l'on a obtenue en [III-2-16].

Nous avons suivi dans ces derniers paragraphes l'exposé donné par Van der Waerden dans son livre « Die gruppentheoretische Methode in der Quantenmechanik ».

III-3-6. Formules auxiliaires. Reprenons les formules de [III-3-1]. Nous avons :

$$G_\beta G_\alpha = G_\gamma, \tag{230}$$

et les γ_s s'expriment en fonction des α_s et des β_s, d'après les formules (189) ou (192) qui définissent l'opération fondamentale de groupe. Formons la matrice dépendant des variables α_s et β_s, c'est-à-dire des éléments G_α et G_β du groupe. Désignons cette matrice par $S(G_\beta, G_\alpha)$ et définissons ses éléments par :

$$S_{ik}(G_\beta, G_\alpha) = \frac{\partial \gamma_i}{\partial \beta_k} \qquad (i, k = 1, 2, \ldots, r). \tag{231}$$

Nous avons déjà étudié cette matrice [III-3-1] pour $\beta_s = 0$, c'est-à-dire lorsque $G_\beta = E$, où E est l'élément unité du groupe. Etudions-en les propriétés. Il résulte de sa définition que :

$$S(G_\beta, E) = I. \tag{232}$$

Démontrons encore que :

$$S(G_\beta, G_\alpha)\cdot S(E, G_\beta) = S(E, G_\beta G_\alpha). \tag{233}$$

Supposons $G_\alpha = G_{\alpha''} G_{\alpha'}$, de sorte que :

$$G_\gamma = G_\beta G_\alpha = (G_\beta G_{\alpha''})\, G_{\alpha'} = G_\delta G_{\alpha'} \qquad (G_\delta = G_\beta G_{\alpha''}).$$

En utilisant la règle de différentiation des fonctions de fonctions, on a :

$$\frac{\partial \gamma_i}{\partial \beta_k} = \sum_{s=1}^{r} \frac{\partial \gamma_i}{\partial \delta_s}\cdot\frac{\partial \delta_s}{\partial \beta_k} = \sum_{s=1}^{r} S_{is}(G_\delta, G_{\alpha'})\, S_{sk}(G_\beta, G_{\alpha''}),$$

d'où :

$$S(G_\beta, G_{\alpha''} G_{\alpha'}) = S(G_\delta, G_{\alpha'})\, S(G_\beta, G_{\alpha''}).$$

En faisant dans cette égalité $G_\beta = E$, $G_{\alpha''} = G_\beta$ et $G_{\alpha'} = G_\alpha$, on obtient l'égalité (233). Si $G_\alpha = G_\beta^{-1}$, on obtient l'expression de la matrice inverse de la matrice $S(E, G_\beta)$:

$$S^{-1}(E, G_\beta) = S(G_\beta, G_\beta^{-1}). \tag{234}$$

La matrice $S(E, G_\beta)$, avec les notations de [III-3-1], sera $S(\beta_s)$ et la matrice inverse $T(\beta_s)$. Désignons-les par les symboles $S(G_\beta)$ et $T(G_\beta)$:

$$S(E, G_\beta) = S(G_\beta)\,; \quad S^{-1}(E, G_\beta) = T(G_\beta). \tag{235}$$

On a :

$$S(G_\beta)\, T(G_\beta) = T(G_\beta)\, S(G_\beta) = E. \tag{236}$$

La formule (233) donne :

$$S(G_\beta, G_\alpha) = S(E, G_\gamma)\, S^{-1}(E, G_\beta) = S(G_\gamma)\, S^{-1}(G_\beta), \tag{237}$$

et la relation (231) peut être mise sous la forme :

$$\frac{\partial \gamma_i}{\partial \beta_k} = \sum_{s=1}^{r} S_{is}(G_\gamma) T_{sk}(G_\beta). \tag{238}$$

En multipliant les deux membres par $T_{mi}(G_\gamma)$ et en sommant sur i, on obtient, compte tenu de (236) :

$$\sum_{i=1}^{r} T_{mi}(G_\gamma) \frac{\partial \gamma_i}{\partial \beta_k} = T_{mk}(G_\beta). \tag{239}$$

En différentiant (238) par rapport à β_l, on a :

$$\frac{\partial^2 \gamma_i}{\partial \beta_k \, \partial \beta_l} = \sum_{s,\, p=1}^{r} \frac{\partial S_{is}(G_\gamma)}{\partial \gamma_p} \frac{\partial \gamma_p}{\partial \beta_l} T_{sk}(G_\beta) + \sum_{s=1}^{r} S_{is}(G_\gamma) \frac{\partial T_{sk}(G_\beta)}{\partial \beta_l},$$

d'où, en exprimant $\dfrac{\partial \gamma_p}{\partial \beta_l}$, compte tenu de (238) :

$$\frac{\partial^2 \gamma_i}{\partial \beta_k \, \partial \beta_l} = \sum_{s,\, p,\, q=1}^{r} \frac{\partial S_{is}(G_\gamma)}{\partial \gamma_p} S_{pq}(G_\gamma) T_{ql}(G_\beta) T_{sk}(G_\beta) + \sum_{s=1}^{r} S_{is}(G_\gamma) \frac{\partial T_{sk}(G_\beta)}{\partial \beta_l}.$$

En permutant dans le second membre k et l, en utilisant le fait que le premier membre est indépendant de l'ordre des différentiations et en permutant les variables de sommation s et q, on obtient :

$$\sum_{s,\, p,\, q=1}^{r} \left[\frac{\partial S_{is}(G_\gamma)}{\partial \gamma_p} S_{pq}(G_\gamma) - \frac{\partial S_{iq}(G_\gamma)}{\partial \gamma_p} S_{ps}(G_\gamma) \right] T_{ql}(G_\beta) T_{sk}(G_\beta) =$$

$$= -\sum_{s=1}^{r} S_{is}(G_\gamma) \left[\frac{\partial T_{sk}(G_\beta)}{\partial \beta_l} - \frac{\partial T_{sl}(G_\beta)}{\partial \beta_k} \right].$$

On multiplie les deux membres par le produit $S_{lf}(G_\beta) S_{kg}(G_\beta) T_{hi}(G_\gamma)$ et on somme par rapport à i, k et l de 1 à r. En tenant compte de (236), on obtient le système équivalent d'équations :

$$\sum_{i,\, p=1}^{r} \left[\frac{\partial S_{ig}(G_\gamma)}{\partial \gamma_p} S_{pf}(G_\gamma) - \frac{\partial S_{if}(G_\gamma)}{\partial \gamma_p} S_{pg}(G_\gamma) \right] T_{hi}(G_\gamma) =$$

$$= -\sum_{k,\, l=1}^{r} S_{lf}(G_\beta) S_{kg}(G_\beta) \left[\frac{\partial T_{hk}(G_\beta)}{\partial \beta_l} - \frac{\partial T_{hl}(G_\beta)}{\partial \beta_k} \right].$$

On passe facilement de ces égalités aux précédentes en multipliant les deux membres par le produit $T_{fl_1}(G_\beta) T_{gk_1}(G_\beta) S_{i_1h}(G_\gamma)$ et en sommant par rapport à f, g et h. Dans cette dernière égalité le premier membre dépend seulement de γ_s et le second membre seulement de β_s. Ainsi, compte tenu du fait que dans la formule (230) G_α est arbitraire et que par là même β_s et γ_s sont indépendants, les deux membres doivent être égaux à une même constante, et en particulier :

$$\sum_{k,\, l=1}^{r} S_{lj}(G_\beta)\, S_{kg}(G_\beta)\left[\frac{\partial T_{hk}(G_\beta)}{\partial \beta_l}-\frac{\partial T_{hl}(G_\beta)}{\partial \beta_k}\right]=C_{jg}^{(h)}.$$

En changeant les indices, on peut écrire :

$$\sum_{s,\, t=1}^{r} S_{ti}(G_\alpha)\, S_{sk}(G_\alpha)\left[\frac{\partial T_{ps}(G_\alpha)}{\partial \alpha_t}-\frac{\partial T_{pt}(G_\alpha)}{\partial \alpha_s}\right]=-C_{ik}^{(p)}. \tag{240}$$

Si on fait dans cette identité $G_\alpha = E$, c'est-à-dire si $\alpha_s = 0$ $(s = 1, \ldots, r)$, compte tenu de ce que $S(E) = E$, on obtient :

$$-C_{ik}^{(p)}=\left[\frac{\partial T_{pk}(G_\alpha)}{\partial \alpha_i}-\frac{\partial T_{pi}(G_\alpha)}{\partial \alpha_k}\right].$$

En comparant avec la formule (199) de [III-3-1] on voit que les $C_{ik}^{(p)}$ sont les constantes de structure que l'on a déjà définies. En multipliant les deux membres de (240) par $T_{il}(G_\alpha)\, T_{km}(G_\alpha)$ et en sommant par rapport à i et k, on obtient, compte tenu de (236) :

$$\frac{\partial T_{pm}(G_\alpha)}{\partial \alpha_l}-\frac{\partial T_{pl}(G_\alpha)}{\partial \alpha_m}=-\sum_{i,\, k=1}^{r} C_{ik}^{(p)} T_{il}(G_\alpha)\, T_{km}(G_\alpha). \tag{241}$$

Revenons aux formules (207) et (208). (208) s'obtient, comme on l'a vu, de l'annulation de la quantité entre crochets dans (207) pour $\alpha_s = 0$ $(s = 1, \ldots, r)$. En utilisant (241), on montre qu'*il résulte de* (208) *que la quantité entre crochets dans la formule* (207) *est nulle quel que soit* α_s.

Le deuxième terme de cette quantité entre crochets peut être mis sous la forme :

$$\sum_{j,\, k=1}^{r} T_{jp} T_{kq} I_j I_k-\sum_{j,\, k=1}^{r} T_{jq} T_{kp} I_j I_k,$$

et ce faisant on n'écrit pas l'argument G_α dans T. En remplaçant, dans le terme que l'on retranche, j par k et k par j, on obtient :

$$\sum_{j,\, k=1}^{r} T_{jp} T_{kq}(I_j I_k - I_k I_j)=\sum_{j,\, k,\, s=1}^{r} T_{jp} T_{kp} C_{jk}^{(s)} I_s.$$

En transformant le premier terme entre crochets dans la formule (207) :

$$\sum_{j=1}^{r}\left(\frac{\partial T_{jp}}{\partial \alpha_q}-\frac{\partial T_{jq}}{\partial \alpha_p}\right) I_j,$$

compte tenu de (241), on obtient directement le même résultat, mais avec le signe inverse. A côté de la matrice $S(G_\beta, G_\alpha)$ considérons la matrice $S'(G_\beta, G_\alpha)$ dont les éléments sont définis par les formules :

$$\frac{\partial \gamma_l}{\partial \alpha_h}=S'(G_\beta, G_\alpha). \tag{242}$$

Comme ci-dessus, on peut démontrer les formules :

$$\left.\begin{aligned} &S'(E,\ G_\alpha)=I,\\ &S'(G_\beta G_\alpha,\ E)=S'(G_\beta,\ G_\alpha)\,S'(G_\alpha,\ E),\\ &S'^{-1}(G_\alpha,\ E)=S'(G_\alpha^{-1},\ G_\alpha), \end{aligned}\right\} \tag{243}$$

que l'on utilisera dans la suite.

III-3-7. **Construction d'un groupe à partir des constantes de structure.** Dans ce paragraphe, nous aborderons de façon très générale la construction de l'opération de groupe et du groupe des transformations linéaires à partir des constantes de structure $C_{ik}^{(p)}$ données, qui vérifient les relations (194_1) et (194_2). Cette construction est fondée sur un théorème de la théorie des équations aux dérivées partielles que nous énoncerons maintenant.

Soit le système d'équations différentielles aux dérivées partielles :

$$\begin{gathered} \frac{\partial z_i}{\partial x_k}=X_{ik}(x_1,\ \ldots,\ x_n;\ z_1,\ \ldots,\ z_m)\\ (i=1,\ 2,\ \ldots,\ m;\ k=1,\ 2,\ \ldots,\ n). \end{gathered} \tag{244}$$

De ce système, on tire la condition :

$$\frac{\partial^2 z_i}{\partial x_k\,\partial x_l}=\frac{\partial^2 z_i}{\partial x_l\,\partial x_k}.$$

Elle est de la forme :

$$\frac{\partial X_{ik}}{\partial x_l}+\sum_{s=1}^{m}\frac{\partial X_{ik}}{\partial z_s}\cdot\frac{\partial z_s}{\partial x_l}=\frac{\partial X_{il}}{\partial x_k}+\sum_{s=1}^{m}\frac{\partial X_{il}}{\partial z_s}\cdot\frac{\partial z_s}{\partial x_k},$$

soit en remplaçant $\frac{\partial z_s}{\partial x_l}$ et $\frac{\partial z_s}{\partial x_k}$ par le second membre du système (244) :

$$\frac{\partial X_{ik}}{\partial x_l}+\sum_{s=1}^{m}\frac{\partial X_{ik}}{\partial z_s}\cdot X_{sl}=\frac{\partial X_{il}}{\partial x_k}+\sum_{s=1}^{m}\frac{\partial X_{il}}{\partial z_s}\,X_{sk}\qquad (k\neq l). \tag{245}$$

Cette égalité est une relation entre les variables x_k, z_i.

Théorème. *Si au voisinage des valeurs $x_k=x_k^{(0)}$, $z_i=z_i^{(0)}$ (et en ces points) les fonctions X_{ik} sont continues et ont les dérivées partielles continues qui figurent dans les relations (245) et si ces relations sont identiquement vérifiées par rapport à x_k et z_i, le système (244) avec les conditions initiales :*

$$z_i\big|_{x_k=x_k^{(0)}}=z_i^{(0)}$$

a une solution et cette solution est unique.

Le fait que toutes les relations (245) sont identiquement vérifiées avec les conditions de continuité s'appelle condition d'intégrabilité complète du système (244). Décrivons maintenant le schéma de construction de l'opération de groupe et du groupe des transformations linéaires à partir des constantes de structure.

Ainsi, supposons données les constantes $C_{ik}^{(p)}$, où i, k, $p=1,\ 2,\ \ldots,\ r$, et ces constantes vérifient les relations (194_1) et (194_2).

Si l'on résout le système (241) par rapport aux dérivées partielles, on peut vérifier que les relations indiquées sont les conditions d'intégrabilité complète du système (241). Ainsi, il existe une matrice unique $T(G_\alpha)$, d'éléments $T_{pq}(G_\alpha)$ $(p, q = 1, 2, \ldots, r)$, qui se transforme en matrice unité si $G_\alpha = E$, c'est-à-dire si $\alpha_s = 0$ $(s = 1, 2, \ldots, r)$, et qui vérifie le système (241). Connaissant $T(G_\alpha)$ on peut construire la matrice inverse $S(G_\alpha) = T^{-1}(G_\alpha)$. Pour construire l'opération de groupe on utilise le système (238). Les seconds membres de ces équations sont des fonctions connues de β_s et γ_s $(s = 1, 2, \ldots, r)$. On peut vérifier que le système (241) exprime les conditions d'intégrabilité complète du système (238). Par suite, il existe une solution unique du système (238) qui vérifie les conditions initiales:

$$\gamma_i\big|_{\beta_k=0} = \alpha_i.$$

La solution ainsi construite donne l'opération de groupe. Les conditions initiales expriment que l'élément G_γ défini par la formule (230) se transforme en G_α pour $\beta_s = 0$ $(s = 1, \ldots, r)$. Passons maintenant à la construction du groupe des transformations linéaires, c'est-à-dire du groupe des matrices dont l'ordre est donné, à partir des constantes de structure. Comme il a été dit ci-dessus, on a déjà la matrice $T(G_\alpha)$. On a déjà montré en [III-3-2] que les conditions d'intégrabilité complète du système (204) ou du système (205) se réduisent au fait que le terme entre crochets de l'équation (207) est identiquement nul quel que soit le choix des indices et ces dernières conditions sont vérifiées, comme on l'a montré en [III-3-6], si les matrices I_s vérifient les relations (208). Ainsi, la résolution du problème doit commencer par la construction des matrices I_s d'ordre donné qui vérifient les relations (208). C'est un problème algébrique compliqué. Connaissant les matrices I_s, on peut déjà affirmer que le système (205) a une solution unique vérifiant les conditions initiales (206). Cette solution donne précisément le groupe des matrices à constantes de structure $C_{ik}^{(p)}$ données.

On peut montrer que l'intégration du système (241) avec les conditions initiales $T(E) = I$ se ramène à l'intégration du système d'équations différentielles ordinaires linéaires à coefficients constants. Enonçons le résultat correspondant. Soit le système d'équations différentielles ordinaires linéaires à coefficients constants:

$$\frac{dw_{ik}(t)}{dt} = \delta_{ik} + \sum_{p,\,q=1}^{r} C_{pq}^{(i)}\alpha_p w_{qk}(t),$$

où $\delta_{ik} = 0$ pour $i \neq k$ et $\delta_{ii} = 1$ et soient $\alpha_1, \alpha_2, \ldots, \alpha_r$ des constantes données. Alors les fonctions $T_{ik}(\alpha_s) = w_{ik}(t)$ vérifient le système (241) et les conditions initiales $T(E) = I$.

III-3-8. Intégration sur un groupe. Nous avons démontré en [III-2-12] et [III-2-13] une série de relations qui contiennent les sommes de certaines quantités dépendant des éléments du groupe, en outre, la sommation était étendue à tous les éléments du groupe. Dans le cas d'un groupe continu, la sommation peut être remplacée par une intégration par rapport aux paramètres qui définissent les éléments du groupe. Supposons qu'un groupe continu G défini dans l'espace réel T_r à r dimensions par les paramètres $\alpha_1, \alpha_2, \ldots, \alpha_r$ correspond à un domaine fermé borné V, de sorte qu'à tout élément de G correspond un point défini de V et inversement. A l'intérieur du domaine V les fonctions $\varphi_j(\beta_1, \ldots, \beta_r; \alpha_1, \ldots, \alpha_r)$ qui définissent l'opération de groupe sont continues et différentiables un nombre suffisant de fois. De plus, ces fonctions et leurs dérivées sont considérées continues jusqu'à la frontière de V. La relation entre les paramètres $\tilde{\alpha}_s$ correspondant à l'élément G_α^{-1} et les paramètres α_s est aussi considérée comme continue. Les groupes jouissant de pareilles propriétés sont dits *compacts*. Pour définir l'intégration sur le groupe, considérons le déterminant de la matrice $S'(G_\beta, G_\alpha)$ [III-3-6] et introduisons les notations suivantes:

$$\Delta'(G_\beta, G_\alpha) = \left| \frac{\partial \gamma_i}{\partial \alpha_k} \right|_1^r . \tag{246}$$

Il résulte de (243) que:

$$\Delta'(E, G_\alpha) = 1, \tag{247$_1$}$$

$$\Delta'(G_\beta G_\alpha, E) = \Delta'(G_\beta, G_\alpha)\cdot\Delta'(G_\alpha, E). \tag{247$_2$}$$

En désignant $\Delta'(G_\beta, E)$ par $\delta'(G_\beta)$, on peut écrire:

$$\Delta'(G_\beta, G_\alpha) = \frac{\delta'(G_\beta G_\alpha)}{\delta'(G_\alpha)}, \tag{248}$$

d'où puisque $\delta'(E) = \Delta'(E, E) = 1$, on obtient:

$$\Delta'(G_\alpha^{-1}, G_\alpha) = \frac{1}{\delta'(G_\alpha)} . \tag{249}$$

Introduisons encore la notation:

$$u'(G_\alpha) = \Delta'(G_\alpha^{-1}, G_\alpha). \tag{250}$$

Compte tenu des hypothèses faites ci-dessus, $u'(G_\alpha)$ est aussi une fonction continue dans le domaine fermé V. Elle ne s'annule pas, car:

$$\frac{1}{u'(G_\alpha)} = \delta'(G_\alpha) = \Delta'(G_\alpha, E)$$

est aussi une fonction continue. De ce que $u'(E) = 1$, on peut affirmer que $u'(G_\alpha)$ et $\delta(G_\alpha)$ sont des fonctions positives. On peut en dire autant de $\Delta'(G_\beta, G_\alpha)$ d'après (248).

Soit $f(G_\alpha) = f(\alpha_1, \ldots, \alpha_r)$ une fonction continue quelconque dans le domaine fermé V.

On définit l'intégrale de cette fonction sur le groupe G d'après la formule:

$$\int_G f(G_\alpha)\, dG_\alpha = \int_V f(\alpha_1, \ldots, \alpha_r)\, u'(G_\alpha)\, d\alpha_1 \ldots d\alpha_r, \tag{251}$$

où l'intégrale du second membre est une intégrale habituelle étendue au domaine V. Démontrons que cette intégrale a la propriété d'invariance à gauche:

$$\int_G f(G_\alpha)\, dG_\alpha = \int_G f(G_\beta G_\alpha)\, dG_\alpha, \tag{252}$$

ou bien qu'on a les coordonnées:

$$\int_V f(\alpha_1, \ldots, \alpha_r)\, u'(G_\alpha)\, d\alpha_1 \ldots d\alpha_r =$$

$$= \int_V f(\gamma_1, \ldots, \gamma_r)\, u'(G_\alpha)\, d\alpha_1 \ldots d\alpha_r, \tag{253}$$

où G_β est un élément donné quelconque du groupe G. Pour le démontrer, on remplace dans l'intégrale du premier membre l'élément variable G_α par l'élément variable G_δ, en supposant $G_\alpha = G_\beta G_\delta$, en outre, le domaine de variation des paramètres $\delta_1, \ldots, \delta_r$ est V comme auparavant. Le déterminant de la transformation est :

$$\left|\frac{\partial \alpha_i}{\partial \alpha_k}\right|_1^r = \Delta'(G_\beta, G_\delta) = \frac{\delta'(G_\beta G_\delta)}{\delta'(G_\delta)} = \frac{u'(G_\delta)}{u'(G_\beta G_\delta)} = \frac{u'(G_\delta)}{u'(G_\alpha)},$$

et on obtient :

$$\int_V f(\alpha_1, \ldots, \alpha_r)\, u'(G_\alpha)\, d\alpha_1 \ldots d\alpha_r =$$

$$= \int_V f(\alpha_1, \ldots, \alpha_r)\, u'(G_\alpha) \frac{u'(G_\delta)}{u'(G_\alpha)}\, d\delta_1 \ldots d\delta_r = \int_G f(G_\beta G_\delta)\, dG_\delta.$$

Ce qui coïncide avec (253). Le fait de remplacer dans le second membre G_α par G_δ est sans importance.

On obtient de façon analogue l'intégrale invariante à droite. Introduisons le déterminant de la matrice $S(G_\beta, G_\alpha)$:

$$\Delta(G_\beta, G_\alpha) = \left|\frac{\partial \gamma_i}{\partial \beta_k}\right|_1^r. \tag{254}$$

Comme ci-dessus, on a :

$$\left.\begin{aligned} &\Delta(G_\alpha, E) = 1, \\ &\Delta(E, G_\beta G_\alpha) = \Delta(G_\beta, G_\alpha)\, \Delta(E, G_\beta), \\ &\Delta(G_\beta, G_\alpha) = \frac{\delta(G_\beta, G_\alpha)}{\delta(G_\beta)}, \end{aligned}\right\} \tag{255}$$

où $\delta(G_\alpha) = \Delta(E, G_\alpha)$. Introduisons la fonction positive :

$$u(G_\alpha) = \frac{1}{\delta(G_\alpha)} = \Delta(G_\alpha, G_\alpha^{-1}), \tag{256}$$

l'intégrale est donnée par la formule :

$$\int_V f(\alpha_1, \ldots, \alpha_r)\, u(G_\alpha)\, d\alpha_1 \ldots d\alpha_r = \int_G f(G_\alpha)\, \widetilde{dG_\alpha}. \tag{257}$$

Le signe au-dessus de la différentielle sert à distinguer cette intégrale de (251). On a alors la propriété de l'invariance à droite :

$$\int_G f(G_\alpha)\, \widetilde{dG_\alpha} = \int_G f(G_\alpha G_\beta)\, \widetilde{dG_\alpha}. \tag{258}$$

Démontrons à présent que le fait de remplacer G_α par G_α^{-1} dans la fonction à intégrer transforme l'invariance à gauche en une invariance à droite et inversement. Différentions par rapport à α_s l'égalité $G_\lambda = G_\alpha G_\beta$, écrite en fonction des paramètres, et dans toutes les formules considérons $G_\beta = G_\alpha^{-1}$:

$$\frac{\partial \lambda_i}{\partial \alpha_k} + \sum_{s=1}^{r} \frac{\partial \lambda_i}{\partial \beta_s} \cdot \frac{\partial \beta_s}{\partial \alpha_k} = 0, \quad \text{d'où} \quad \left|\frac{\partial \lambda_i}{\partial \alpha_k}\right|_1^r = (-1)^r \left|\frac{\partial \lambda_i}{\partial \beta_k}\right|_1^r \left|\frac{\partial \beta_i}{\partial \alpha_k}\right|_1^r,$$

et, en tenant compte de (245) et (246) :

$$\left|\frac{\partial\beta_i}{\partial\alpha_k}\right|_1^r=(-1)^r\frac{\Delta(G_\alpha,G_\alpha^{-1})}{\Delta'(G_\alpha,G_\alpha^{-1})}=(-1)^r\frac{u(G_\alpha)}{u'(G_\alpha^{-1})}. \tag{259}$$

On peut donner une autre représentation de ce déterminant. En effet, il résulte de l'égalité

$$\left|\frac{\partial\beta_i}{\partial\alpha_k}\right|_1^r\left|\frac{\partial\alpha_i}{\partial\beta_k}\right|_1^r=1$$

que

$$\left|\frac{\partial\alpha_i}{\partial\beta_k}\right|_1^r=(-1)^r\frac{u'(G_\alpha^{-1})}{u(G_\alpha)}, \tag{260}$$

soit, en changeant la place de G_α et G_α^{-1} :

$$\left|\frac{\partial\beta_i}{\partial\alpha_k}\right|=(-1)^r\frac{u'(G_\alpha)}{u(G_\alpha^{-1})}. \tag{261}$$

Considérons maintenant les intégrales. En effectuant un changement des variables d'intégration, on obtient en utilisant (260) :

$$\int_V f(\alpha_1,\ldots,\alpha_r)\,u'(G_\alpha)\,d\alpha_1\ldots d\alpha_r=$$

$$=\int_V f(\tilde\beta_1,\ldots,\tilde\beta_r)\,u'(G_\alpha)\left\|\frac{\partial\alpha_i}{\partial\beta_k}\right|_1^r\right|d\beta_1\ldots d\beta_r=$$

$$=\int_V f(\tilde\beta_1,\ldots,\tilde\beta_r)\,u'(G_\alpha)\frac{u(G_\beta)}{u'(G_\alpha)}\,d\beta_1\ldots d\beta_r.$$

En simplifiant par $u'(G_\alpha)$ et en remplaçant l'élément variable G_β par l'élément variable G_α, on obtient :

$$\int_V f(\tilde\alpha_1,\ldots,\tilde\alpha_r)\,u(G_\alpha)\,d\alpha_1\ldots d\alpha_r=$$

$$=\int_V f(\alpha_1,\ldots,\alpha_r)\,u'(G_\alpha)\,d\alpha_1\ldots d\alpha_r \tag{262}$$

De même, en tenant compte de (261), on obtient :

$$\int_V f(\tilde\alpha_1,\ldots,\tilde\alpha_r)\,u'(G_\alpha)\,d\alpha_1\ldots d\alpha_r=$$

$$=\int_V f(\alpha_1,\ldots,\alpha_r)\,u(G_\alpha)\,d\alpha_1\ldots d\alpha_r. \tag{263}$$

Jusqu'à présent on n'a pas utilisé le fait qu'il s'agit d'un groupe compact. Le domaine V peut être infini. Mais alors il faut supposer la fonction $f(\alpha_1,\ldots,\alpha_r)$ telle que toutes les intégrales écrites aient un sens. Partant du fait qu'on a un groupe compact démontrons que $u(G_\alpha) = u'(G_\alpha)$. Pour cela considérons le déterminant :

$$D(G_\beta, G_\alpha) = \left| \frac{\partial \mu_i}{\partial \beta_k} \right|_1^r, \tag{264}$$

où $G_\mu = G_\alpha^{-1} G_\beta G_\alpha$ et démontrons que :

$$D(G_\beta, G_{\alpha''} G_{\alpha'}) = D(G_{\alpha''}^{-1} G_\beta G_{\alpha''}, G_{\alpha'}) D(G_\beta, G_{\alpha''}). \tag{265}$$

On peut écrire :

$$G_\mu = (G_{\alpha''} G_{\alpha'})^{-1} G_\beta (G_{\alpha''} G_{\alpha'}) = G_{\alpha'}^{-1} G_\nu G_{\alpha'},$$

où $G_\nu = G_{\alpha''}^{-1} G_\beta G_{\alpha''}$, c'est pourquoi :

$$\left| \frac{\partial \mu_i}{\partial \beta_k} \right|_1^r = \left| \frac{\partial \mu_i}{\partial \nu_k} \right|_1^r \cdot \left| \frac{\partial \nu_i}{\partial \beta_k} \right|_1^r = D(G_\nu, G_{\alpha'}) D(G_\beta, G_{\alpha''}),$$

d'où (265). En faisant $G_\beta = E$, on obtient :

$$D(E, G_{\alpha''} G_{\alpha'}) = D(E, G_{\alpha'}) D(E, G_{\alpha''}). \tag{266}$$

Introduisons la fonction numérique de l'élément :

$$\eta(G_\alpha) = D(E, G_\alpha). \tag{267}$$

D'après (266), on peut écrire :

$$\eta(G_{\alpha''} G_{\alpha'}) = \eta(G_{\alpha''}) \eta(G_{\alpha'}), \tag{268}$$

c'est-à-dire que le produit des éléments signifie le produit des valeurs correspondantes de la fonction $\eta(G_\alpha)$. On a :

$$\eta(E) = 1 \quad \text{et} \quad \eta(G_\alpha) \eta(G_\alpha^{-1}) = 1, \tag{269}$$

et la fonction $\eta(G_\alpha)$ est continue et positive dans le domaine fermé V.

Le groupe étant compact, démontrons que $\eta(G_\alpha) = 1$ pour tout élément G_α. Supposons que pour un élément G_α on ait $\eta(G_\alpha) \neq 1$. Si, par exemple, $\eta(G_\alpha) < 1$, d'après (269) : $\eta(G_\alpha^{-1}) > 0$ et $\eta(G_\alpha) > 1$, ce qu'on peut toujours considérer. On a :

$$\eta(G_\alpha^n) = [\eta(G_\alpha)]^n \longrightarrow \infty \text{ si } n \longrightarrow \infty.$$

Ce qui est en contradiction avec le fait que la fonction $\eta(G_\alpha)$, continue dans le domaine fermé V, doit être bornée. Établissons la relation entre $u(G_\alpha)$ et $u'(G_\alpha)$. Soit :

$$G_\gamma = G_\beta G_\alpha = G_\alpha^{-1}(G_\alpha G_\beta) G_\alpha = G_\alpha^{-1} G_\rho G_\alpha \qquad (G_\rho = G_\alpha G_\beta).$$

Nous avons :

$$\left| \frac{\partial \gamma_i}{\partial \beta_k} \right|_1^r = \Delta(G_\beta, G_\alpha).$$

Mais, par ailleurs :

$$\left| \frac{\partial \gamma_i}{\partial \beta_k} \right|_1^r = \left| \frac{\partial \gamma_i}{\partial \rho_k} \right|_1^r \left| \frac{\partial \rho_i}{\partial \beta_k} \right|_1^r = D(G_\rho, G_\alpha) \Delta'(G_\alpha, G_\beta),$$

c'est-à-dire :

$$\Delta(G_\beta, G_\alpha) = D(G_\alpha G_\beta, G_\alpha) \cdot \Delta'(G_\alpha, G_\beta).$$

En supposant $G_\beta = G_\alpha^{-1}$, on obtient :

$$\Delta(G_\alpha^{-1}, G_\alpha) = D(E, G_\alpha) \cdot \Delta'(G_\alpha, G_\alpha^{-1}),$$

c'est-à-dire :

$$u(G_\alpha^{-1}) = \eta(G_\alpha)\, u'(G_\alpha^{-1}), \text{ soit } u(G_\alpha^{-1}) = u'(G_\alpha^{-1})$$

pour tout G_α car $\eta(G_\alpha) = 1$. Ainsi, pour un groupe compact l'intégrale invariante à gauche (251) coïncide avec l'intégrale invariante à droite (257). De plus, il résulte de (262) ou de (263) que cette intégrale coïncide aussi avec l'intégrale :

$$\int_V f(\tilde{\alpha}_1, \ldots, \tilde{\alpha}_r)\, u'(G_\alpha)\, d\alpha_1 \ldots d\alpha_r.$$

Dans le cas de groupes non compacts l'intégrale invariante à gauche peut être différente de l'intégrale invariante à droite. Comme exemple considérons le groupe des transformations linéaires de la forme :

$$z' = e^{\alpha_1} z + \alpha_2,$$

où α_1 et α_2 varient de $(-\infty)$ à $(+\infty)$. Dans ce cas $r = 2$ et V est le plan tout entier. La composition de deux transformations donne :

$$z' = e^{\alpha_1} z + \alpha_2; \quad z'' = e^{\beta_1} z' + \beta_2; \quad \gamma_1 = \varphi_1(\beta_1, \beta_2; \alpha_1, \alpha_2) = \beta_1 + \alpha_1;$$

c'est-à-dire :

$$z'' = e^{\beta_1 + \alpha_1} z + (e^{\beta_1}\alpha_2 + \beta_2); \quad \gamma_2 = \varphi_2(\beta_1, \beta_2; \alpha_1, \alpha_2) = e^{\alpha_1}\beta_2 + \alpha_2.$$

A l'élément unité correspondent les paramètres $\alpha_1 = \alpha_2 = 0$. L'élément G_α^{-1} a pour paramètres $\tilde{\alpha}_1 = -\alpha_1$, $\tilde{\alpha}_2 = -\alpha_2 e^{-\alpha_1}$. Calculons les déterminants fonctionnels :

$$\Delta'(G_\beta, G_\alpha) = \begin{vmatrix} 1, & 0 \\ 0, & e^{\alpha_1} \end{vmatrix} = e^{\alpha_1}; \quad \delta'(G_\alpha) = e^{\alpha_1}; \quad u'(G_\alpha) = e^{-\alpha_1};$$

$$\Delta(G_\beta, G_\alpha) = \begin{vmatrix} 1, & e^{\alpha_1}\beta_2 \\ 0, & 1 \end{vmatrix} = 1; \quad \delta(G_\alpha) = u(G_\alpha) = 1.$$

L'intégrale invariante à gauche est de la forme :

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(\alpha_1, \alpha_2)\, e^{-\alpha_1}\, d\alpha_1\, d\alpha_2$$

et l'intégrale invariante à droite :

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(\alpha_1, \alpha_2)\, d\alpha_1\, d\alpha_2.$$

Remarquons que pour démontrer l'égalité des deux intégrales, c'est-à-dire $u(G_\alpha) = u'(G_\alpha)$, on peut considérer une autre condition que celle d'un groupe compact.

Soit G' un sous-groupe formé d'éléments de G de la forme :

$$G_\alpha G_\beta G_\alpha^{-1} G_\beta^{-1} \tag{270}$$

ou qui s'obtiennent à partir d'éléments de cette forme par multiplication entre eux, G_α et G_β étant des éléments quelconques de G.

Il est facile de voir que si un élément G_γ est de la forme (270), l'élément inverse G_γ^{-1} est aussi de la forme (270).

De même, l'élément $G_\delta^{-1} G_\gamma G_\delta$, quel que soit G_δ appartenant à G, est de la forme (270). Autrement dit, le sous-groupe G' est engendré par les éléments (270). De ce qui a été dit il résulte que G' est un sous-groupe distingué de G. Le sous-groupe G' se réduit à l'élément unité dans le cas et seulement dans le cas où tous les éléments (270) sont des éléments unités, c'est-à-dire dans le cas où G est un groupe abélien. Le sous-groupe G' peut coïncider avec G. En particulier, ce sera le cas si G est un groupe simple, non abélien. Le sous-groupe G' est appelé d'ordinaire *groupe des commutateurs* (ou groupe dérivé) de G.

Des définitions (268) et (269) il résulte que $\eta\,(G_\alpha G_\beta G_\alpha^{-1} G_\beta^{-1}) = 1$, que $\eta\,(G_\gamma) = 1$ pour tous les G_γ appartenant à G' et que $\eta\,(G_\alpha)$ a une valeur identique pour tous les éléments d'un même ensemble par rapport au groupe G', c'est-à-dire que la fonction $\eta\,(G_\alpha)$ a une valeur définie pour tout élément du groupe quotient de G par G'. Si G' coïncide avec G, $\eta\,(G_\alpha) = 1$ pour tout G_α de G. Il en est de même si le groupe quotient est compact, mais comme

$$\eta\,(G_\alpha) = 1,$$

on a :

$$u\,(G_\alpha) = u'\,(G_\alpha).$$

III-3-9. Propriété d'orthogonalité. Exemples. Dans le cas des groupes finis la propriété d'invariance à gauche et à droite de l'intégrale est analogue à la suivante : si on prend un élément variable G_s et un élément fixe G_t, le produit $G_s G_t$ ou $G_t G_s$ ne prend qu'une seule fois toutes les valeurs du groupe. Cette propriété a été utilisée pour démontrer que toute représentation d'un groupe est équivalente à une représentation unitaire et aussi pour démontrer des propriétés d'orthogonalité. En utilisant l'intégrale invariante on peut démontrer des propositions analogues pour les groupes compacts. Si $A\,(G_\alpha)$ sont des matrices unitaires donnant une représentation linéaire irréductible du groupe compact G et $B\,(G_\alpha)$ des matrices unitaires donnant une représentation irréductible non équivalente, en désignant par deux indices les éléments des matrices, on peut établir les formules suivantes exprimant l'orthogonalité des représentations unitaires irréductibles non équivalentes :

$$\int_V \{A\,(G_\alpha)\}_{ij} \overline{\{B\,(G_\alpha)\}_{kl}}\, u\,(G_\alpha)\, d\alpha_1 \ldots d\alpha_r = 0. \tag{271}$$

Pour une représentation irréductible on a :

$$\int_V \{A\,(G_\alpha)\}_{ij} \overline{\{A\,(G_\alpha)\}_{kl}}\, u\,(G_\alpha)\, d\alpha_1 \ldots d\alpha_r = \frac{\delta_{ik}\delta_{jl}}{p} \int_V u\,(G_\alpha)\, d\alpha_1 \ldots d\alpha_r, \tag{272}$$

où p est l'ordre des matrices. De même, pour les caractères :

$$X\,(G_\alpha) = \sum_{i=1}^{p} \{A\,(G_\alpha)\}_{ii}\,; \quad X'\,(G_\alpha) = \sum_{i=1}^{q} \{B\,(G_\alpha)\}_{ii},$$

où p et q sont les ordres des matrices $A(G_\alpha)$ et $B(G_\alpha)$, on a les propriétés suivantes:

$$\int_V X(G_\alpha)\overline{X'(G_\alpha)}\,u(G_\alpha)\,d\alpha_1\ldots d\alpha_r=0, \tag{273}$$

$$\int_V X(G_\alpha)\overline{X(G_\alpha)}\,u(G_\alpha)\,d\alpha_1\ldots d\alpha_r=\int_V u(G_\alpha)\,d\alpha_1\ldots d\alpha_r. \tag{274}$$

1. Passons maintenant à l'examen d'e x e m p l e s. Soit G le groupe abélien des rotations du plan autour de l'origine. Dans ce cas $r=1$ et le seul paramètre α donne la valeur de l'angle de rotation. On considère que α appartient à l'intervalle ouvert $(0, 2\pi)$ et qu'on a identifié les extrémités de cet intervalle. Des rotations successives d'angles α et β donnent une rotation d'angle $\beta+\alpha$ de laquelle il faut au besoin retrancher 2π, pour que la somme appartienne à l'intervalle ouvert indiqué. Dans le cas considéré les déterminants fonctionnels $\Delta(G_\beta, G_\alpha)$ et $\Delta'(G_\beta, G_\alpha)$ se réduisent à la dérivée de $\beta+\alpha$ par rapport à β ou α, égale à l'unité, de sorte que $u(G_\alpha)=u'(G_\alpha)=1$. On sait que le groupe G a des représentations unitaires irréductibles du premier ordre $e^{im\alpha}$ ($m=0, \pm1, \pm2, \ldots$), et (257) et (258) donnent les formules bien connues:

$$\int_0^{2\pi} e^{im_1\alpha}\overline{e^{im_2\alpha}}\,d\alpha=\int_0^{2\pi} e^{i(m_1-m_2)\alpha}\,d\alpha=\begin{cases}0 & \text{si } m_1\neq m_2,\\ 2\pi & \text{si } m_1=m_2.\end{cases} \tag{275}$$

Remarquons qu'en vertu de la nécessité de ce que la somme $\beta+\alpha$ appartienne à l'intervalle $(0, 2\pi)$ on a des singularités pour la continuité et pour la définition des dérivées de cette somme dans le cas où α et β sont tels que la somme est égale à 2π.

2. Considérons le groupe des rotations de l'espace à trois dimensions et utilisons des paramètres quelque peu différents de ceux dont il a été question [III-3-3]. Supposons que l'on fasse tourner l'espace d'angle ω autour d'un axe qui forme les angles α, β et γ avec les axes de coordonnées X, Y et Z.

Introduisons les quatre paramètres:

$$a_0=\cos\frac{1}{2}\omega;\quad a_1=\cos\alpha\sin\frac{1}{2}\omega;\quad a_2=\cos\beta\sin\frac{1}{2}\omega;\quad a_3=\cos\gamma\sin\frac{1}{2}\omega, \tag{276}$$

qui sont liés par la relation:

$$a_0^2+a_1^2+a_2^2+a_3^2=1. \tag{277}$$

A la transformation unitaire correspondent les valeurs $a_0=1$, $a_1=a_2=a_3=0$. Nous pouvons prendre a_1, a_2, a_3 comme paramètres. Alors a_0 est considéré comme fonction de ces paramètres.

Si on effectue des rotations successives définies par les paramètres (a_0, a_1, a_2, a_3) et (b_0, b_1, b_2, b_3), les paramètres de la rotation résultante (c_0, c_1, c_2, c_3) sont définis, comme il est facile de le voir, par les formules:

$$\begin{aligned} c_0&=a_0b_0-a_1b_1-a_2b_2-a_3b_3,\\ c_1&=a_0b_1+a_1b_0+a_2b_3-a_3b_2,\\ c_2&=a_0b_2-a_1b_3+a_2b_0+a_3b_1,\\ c_3&=a_0b_3+a_1b_2-a_2b_1+a_3b_0.\end{aligned} \tag{278}$$

En considérant que a_0 est une fonction de a_1, a_2, a_3, compte tenu de (277), on obtient :

$$a_0 \frac{\partial a_0}{\partial a_j} + a_j = 0, \quad (j = 1, 2, 3),$$

d'où $\frac{\partial a_0}{\partial a_j} = 0$ pour E. En utilisant ces résultats, on peut former facilement le déterminant fonctionnel, lorsque $b_0 = 1$, $b_1 = b_2 = b_3 = 0$:

$$\frac{D(c_1, c_2, c_3)}{D(b_1, b_2, b_3)} = \begin{vmatrix} a_0 & -a_3 & a_2 \\ a_3 & a_0 & -a_1 \\ -a_2 & a_1 & a_0 \end{vmatrix} = a_0(a_0^2 + a_1^2 + a_2^2 + a_3^2) = \\ = a_0 = \sqrt{1 - a_1^2 - a_2^2 - a_3^2}.$$

L'intégrale invariante est de la forme :

$$\int_V f(a_1, a_2, a_3) \frac{1}{\sqrt{1 - a_1^2 - a_2^2 - a_3^2}} da_1\, da_2\, da_3. \tag{279}$$

Le domaine V est la sphère de centre à l'origine des coordonnées et de rayon unité. Notons que les formules (278) s'obtiennent directement à partir de la règle de multiplication des quaternions :

$$c_0 + c_1 i + c_2 j + c_3 k = (a_0 + a_1 i + a_2 j + a_3 k)(b_0 + b_1 i + b_2 j + b_3 k),$$

et les unités i, j et k obéissent à la loi de multiplication suivante :

$$i^2 = j^2 = k^2 = -1\,; \quad ij = -ji = k\,; \quad jk = -kj = i\,; \quad ki = -ik = j.$$

Il est facile d'établir la relation entre les paramètres (a_0, a_1, a_2, a_3) et les angles d'Euler α, β, γ. On a les formules correspondantes :

$$a_0 = \cos\frac{1}{2}\beta \cos\frac{1}{2}(\alpha + \gamma)\,; \qquad a_2 = \sin\frac{1}{2}\beta \sin\frac{1}{2}(\gamma - \alpha)\,;$$

$$a_1 = \sin\frac{1}{2}\beta \cos\frac{1}{2}(\gamma - \alpha)\,; \qquad a_3 = \cos\frac{1}{2}\beta \sin\frac{1}{2}(\alpha + \gamma).$$

L'intégrale invariante s'écrit à l'aide des paramètres (α, β, γ) sous la forme

$$\int_V f(\alpha, \beta, \gamma) \sin\beta \sin^2\frac{1}{2}(\alpha - \gamma)\, d\alpha\, d\beta\, d\gamma, \tag{280}$$

et $0 \leqslant \alpha < 2\pi$, $0 \leqslant \beta < \pi$, $0 \leqslant \gamma < 2\pi$. On notera que dans l'intégrale (279) la fonction

$$\frac{1}{a_0} = \frac{1}{\sqrt{1 - a_1^2 - a_2^2 - a_3^2}}$$

devient infinie si $\omega = \pi$. Ceci est lié au fait que dans les formules (276) pour a_1, a_2, a_3 au lieu de ω on a $\sin\frac{1}{2}\omega$. Notons, en liaison avec ceci, que les propriétés dont il a été question [III-3-8] pour la définition d'un groupe compact doivent être vérifiées seulement pour un certain choix des paramètres. Si on change de paramètres, ces propriétés peuvent disparaître. De plus, le groupe des rotations de l'espace à trois dimensions a des singularités dans la continuité et dans la détermination des dérivées dont il a déjà été question à la fin du premier exemple concernant le groupe des rotations du plan autour de l'origine.

Notons encore que si les intégrales invariantes coïncident pour le groupe des rotations de l'espace, ceci peut être directement lié au fait que c'est un groupe simple, non abélien.

3. Il est facile de vérifier au moyen d'un calcul direct que les intégrales invariantes à gauche et à droite sur le groupe de Lorentz coïncident, ce groupe, comme on l'a vu, étant homomorphe au groupe des transformations linéaires, à déterminant égal à 1:

$$\left.\begin{aligned} x_1' &= a_0x_1 + a_1x_2, \\ x_2' &= a_2x_1 + a_3x_2 \end{aligned}\right\} \quad (a_0a_3 - a_1a_2 = 1). \tag{281}$$

A l'élément unité correspondent les valeurs $a_0 = a_3 = 1$, $a_1 = a_2 = 0$. On peut considérer que a_0 est une fonction de a_1, a_2, a_3 et prendre pour paramètres les parties réelles et imaginaires de a_1, a_2 et $a_3 - 1$. L'opération de groupe se réduit à la multiplication des matrices du deuxième ordre et on a:

$$c_0 = b_0a_0 + b_1a_2, \quad c_1 = b_0a_1 + b_1a_3, \quad c_2 = b_2a_0 + b_3a_2, \quad c_3 = b_2a_1 + b_3a_3. \tag{282}$$

Si l'on pose $a_k = \alpha_k' + i\alpha_k''$ $(k = 0, 1, 2, 3)$, les paramètres du groupe seront α_1', α_2', α_3', α_1'', α_2'', α_3''. En supposant de plus que $b_k = \beta_k' + i\beta_k''$ et $c_k = \gamma_k' + i\gamma_k''$, il faut, pour calculer l'intégrale invariante, calculer les déterminants fonctionnels:

$$\frac{D(\gamma_1', \gamma_2', \gamma_3', \gamma_1'', \gamma_2'', \gamma_3'')}{D(\beta_1', \beta_2', \beta_3', \beta_1'', \beta_2'', \beta_3'')} \text{ si } \beta_1' = \beta_2' = \beta_1'' = \beta_2'' = \beta_3'' = 0; \quad \beta_3' = 1,$$

u bien:

$$\frac{D(\gamma_1', \gamma_2', \gamma_3', \gamma_1'', \gamma_2'', \gamma_3'')}{D(\alpha_1', \alpha_2', \alpha_3', \alpha_1'', \alpha_2'', \alpha_3'')} \text{ si } \alpha_1' = \alpha_2' = \alpha_1'' = \alpha_2'' = \alpha_3'' = 0; \quad \alpha_3' = 1,$$

de plus, le fait que $\alpha_3' = 1$ et non nul n'est pas important pour la transformation identique du groupe. Dans les deux cas on obtient la même intégrale invariante:

$$\int_V f(\alpha_1', \alpha_2', \alpha_3', \alpha_1'', \alpha_2'', \alpha_3'') \frac{1}{\alpha_3'^2 + \alpha_3''^2} d\alpha_1' \, d\alpha_2' \, d\alpha_3', \, d\alpha_1'' \, d\alpha_2'' \, d\alpha_3''. \tag{283}$$

Le domaine V est l'espace à six dimensions tout entier. La coïncidence des intégrales invariantes est liée au fait que le sous-groupe G' du groupe (281), formé à partir des éléments générateurs $G_\alpha G_\beta G_\alpha^{-1} G_\beta^{-1}$ et dont il a été question en [III-3-8], coïncide avec le groupe lui-même. Il est facile de montrer, en effet, que G' ne se réduit pas à la transformation identique ou au sous-groupe distingué engendré par les éléments E et $(-E)$. Le calcul effectif de la densité dans l'intégrale invariante (283) peut être effectué simplement à partir du lemme qui utilise la notion de fonction analytique de plusieurs variables complexes (cf. Chapitre IV de la deuxième partie de ce tome).

L e m m e. *Soient $w_s = u_s + iv_s$ $(s = 1, 2, \ldots, k)$ des fonctions analytiques des variables complexes $z_s = x_s + iy_s$ $(s = 1, 2, \ldots, k)$. Alors le déterminant fonctionnel des fonctions $(u_1, v_1, \ldots, u_k, v_k)$ par rapport aux variables $(x_1, y_1, \ldots, x_k, y_k)$ est égal au carré du module du déterminant fonctionnel des fonctions $(w_1, \ldots, w_k)$ par rapport aux variables $(z_1, \ldots, z_k)$.*

On a (cf. Chapitres I et IV de la 2ème partie de ce tome):

$$\frac{\partial u_l}{\partial x_k} = \frac{\partial v_l}{\partial y_k}; \quad \frac{\partial v_l}{\partial x_k} = -\frac{\partial u_l}{\partial y_k},$$

et l'on peut alors écrire :

$$\frac{D(u_1, v_1, \ldots, u_k, v_k)}{D(x_1, y_1, \ldots, x_k, y_k)} = \begin{vmatrix} a_{11} & b_{11} & a_{12} & b_{12} & \ldots & a_{1k} & b_{1k} \\ -b_{11} & a_{11} & -b_{12} & a_{12} & \ldots & -b_{1k} & a_{1k} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{k1} & b_{k1} & a_{k2} & b_{k2} & \ldots & a_{kk} & b_{kk} \\ -b_{k1} & a_{k1} & -b_{k2} & a_{k2} & \ldots & -b_{kk} & a_{kk} \end{vmatrix},$$

où :

$$a_{lk} = \frac{\partial u_l}{\partial x_k}; \quad b_{lk} = \frac{\partial v_l}{\partial x_k}.$$

En ajoutant à chaque colonne d'indice impair la colonne d'indice pair, multipliée par i, on obtient le déterminant :

$$\begin{vmatrix} c_{11} & b_{11} & c_{12} & b_{12} & \ldots & c_{1k} & b_{1k} \\ ic_{11} & a_{11} & ic_{12} & a_{12} & \ldots & ic_{1k} & a_{1k} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ c_{k1} & b_{k1} & c_{k2} & b_{k2} & \ldots & c_{kk} & b_{kk} \\ ic_{k1} & a_{k1} & ic_{k2} & a_{k2} & \ldots & ic_{kk} & a_{kk} \end{vmatrix} \quad (c_{lk} = a_{lk} + ib_{lk}).$$

En retranchant ensuite de chaque ligne d'indice pair la ligne précédente d'indice impair multipliée par i, on obtient :

$$\begin{vmatrix} c_{11} & b_{11} & c_{12} & b_{12} & \ldots & c_{1k} & b_{1k} \\ 0 & \overline{c_{11}} & 0 & \overline{c_{12}} & \ldots & 0 & \overline{c_{1k}} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ c_{k1} & b_{k1} & c_{k2} & b_{k2} & \ldots & c_{kk} & b_{kk} \\ 0 & \overline{c_{k1}} & 0 & \overline{c_{k2}} & \ldots & 0 & \overline{c_{kk}} \end{vmatrix}.$$

En ramenant à gauche toutes les colonnes d'indices impairs et en haut toutes les lignes d'indices impairs on a :

$$\begin{vmatrix} c_{11} & c_{12} & \ldots & c_{1k} & b_{11} & b_{12} & \ldots & b_{1k} \\ c_{21} & c_{22} & \ldots & c_{2k} & b_{21} & b_{22} & \ldots & b_{2k} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ c_{k1} & c_{k2} & \ldots & c_{kk} & b_{k1} & b_{k2} & \ldots & b_{kk} \\ 0 & 0 & \ldots & 0 & \overline{c_{11}} & \overline{c_{12}} & \ldots & \overline{c_{1k}} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \ldots & 0 & \overline{c_{k1}} & \overline{c_{k2}} & \ldots & \overline{c_{kk}} \end{vmatrix}.$$

d'où il résulte que :

$$\frac{D(u_1, v_1, \ldots, u_k, v_k)}{D(x_1, y_1, \ldots, x_k, y_k)} = \begin{vmatrix} c_{11} & \ldots & c_{1k} \\ \cdot & \cdot & \cdot \\ c_{k1} & \ldots & c_{kk} \end{vmatrix} \begin{vmatrix} \overline{c_{11}} & \ldots & \overline{c_{1k}} \\ \cdot & \cdot & \cdot \\ \overline{c_{k1}} & \ldots & \overline{c_{kk}} \end{vmatrix} = \left| \frac{D(w_1, \ldots, w_k)}{D(z_1, \ldots, z_k)} \right|^2.$$

On passe alors au calcul de la fonction $u(G_\alpha)$ dans l'intégrale invariante. Pour cela, compte tenu du lemme, il faut calculer le déterminant fonctionnel :

$$\frac{D(c_1, c_2, c_3)}{D(b_1, b_2, b_3)} \quad \text{si} \quad b_0 = b_3 = 1\ ; \quad b_1 = b_2 = 0 \tag{284}$$

ou

$$\frac{D(c_1, c_2, c_3)}{D(a_1, a_2, a_3)} \quad \text{si} \quad a_0 = a_3 = 1\ ; \quad a_1 = a_2 = 0. \tag{285}$$

De la relation $a_0a_3 - a_1a_2 = 0$, il résulte :

$$-a_2 + a_3\frac{\partial a_0}{\partial a_1} = 0\ ; \quad -a_1 + a_3\frac{\partial a_0}{\partial a_2} = 0\ ; \quad a_0 + a_3\frac{\partial a_0}{\partial a_3} = 0.$$

De plus on a :

$$\begin{array}{lll}
\dfrac{\partial c_1}{\partial a_1} = b_0\ ; & \dfrac{\partial c_1}{\partial a_2} = 0\ ; & \dfrac{\partial c_1}{\partial a_3} = b_1\ ; \\[2ex]
\dfrac{\partial c_2}{\partial a_1} = b_2\dfrac{\partial a_0}{\partial a_1}\ ; & \dfrac{\partial c_2}{\partial a_2} = b_2\dfrac{\partial a_0}{\partial a_2} + b_3\ ; & \dfrac{\partial c_2}{\partial a_3} = b_2\dfrac{\partial a_0}{\partial a_3}\ ; \\[2ex]
\dfrac{\partial c_3}{\partial a_1} = b_2\ ; & \dfrac{\partial c_3}{\partial a_2} = 0\ ; & \dfrac{\partial c_3}{\partial a_3} = b_3,
\end{array}$$

d'où (283). En partant de la formule (284), on obtient le même résultat.

Imprimé en Union Soviétique