

ANNALES SCIENTIFIQUES DE L'É.N.S.

MAURICE FRÉCHET

La notion de différentielle dans l'analyse générale

Annales scientifiques de l'É.N.S. 3^e série, tome 42 (1925), p. 293-323

<http://www.numdam.org/item?id=ASENS_1925_3_42__293_0>

© Gauthier-Villars (Éditions scientifiques et médicales Elsevier), 1925, tous droits réservés.

L'accès aux archives de la revue « Annales scientifiques de l'É.N.S. » (<http://www.elsevier.com/locate/ansens>) implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
<http://www.numdam.org/>

LA NOTION DE DIFFÉRENTIELLE

DANS

L'ANALYSE GÉNÉRALE

PAR M. MAURICE FRÉCHET

(Université de Strasbourg)

Introduction.

Objet de l'Analyse générale. — L'Analyse fonctionnelle, créée par M. Volterra, a essentiellement pour objet l'étude des fonctions qui font correspondre à chaque ligne ou surface, à chaque fonction ordinaire, un *nombre* déterminé. Nous avons, dans différents mémoires, considéré les fonctions plus générales qui font correspondre un *nombre* déterminé à chaque élément d'un certain ensemble abstrait de nature quelconque. On peut encore ranger l'étude de ces fonctions, qu'on peut appeler fonctionnelles avec M. Hadamard, dans l'Analyse fonctionnelle.

Mais, sans pour cela s'enfoncer dans les nuages, et gardant un contact étroit avec toute la science acquise, on peut généraliser encore. Les fonctionnelles que nous venons de décrire peuvent être considérées comme définissant une correspondance univoque (non nécessairement biunivoque) entre l'ensemble abstrait considéré et un ensemble de nombres. Ou, si l'on aime mieux, elles définissent chacune une transformation d'un certain ensemble d'éléments ou points d'un espace abstrait en un certain ensemble de points d'une droite.

(¹) Le présent Mémoire est le développement d'une Note des *Comptes rendus* du 16 mars 1925, t. 180, p. 806 (même titre).

Or, il n'y a nulle raison pour se limiter au cas où l'ensemble transformé est linéaire. L'étude de transformations d'une espèce plus générale s'est imposée depuis longtemps.

Depuis longtemps, par exemple, on a étudié les transformations de points en droites, en plans, en sphères, ..., dans l'espace euclidien, de fonctions analytiques d'une variable complexe en fonctions d'une même nature, etc. Nous avons plus généralement considéré des transformations biunivoques et bicontinues (sous le nom d'homéomorphies) entre ensembles topologiques abstraits ⁽¹⁾. Il est donc naturel et utile de se proposer l'étude des transformations d'ensembles abstraits en ensembles abstraits. On pourra donner à cette étude le nom d'*Analyse générale* emprunté à M. E. H. Moore. Il suffira ensuite de particulariser les résultats obtenus en prenant en considération la nature des éléments envisagés pour obtenir des applications variées.

Notion de continuité. — La définition de la continuité d'une transformation ponctuelle $M = \Phi(m)$ d'un ensemble abstrait n'offre pas de difficulté quand on suppose que ces deux ensembles appartiennent à des espaces où la notion de voisinage a été étendue d'une façon que nous préciserons plus loin. On peut exprimer cette définition d'une façon intuitive en disant que tout point qui est infiniment voisin d'autres points est transformé en un point infiniment voisin des transformés des autres points, sans attacher nécessairement à l'idée de voisinage celle de distance numérique.

La Notion de différentielle en Analyse fonctionnelle. — Pour généraliser aux espaces abstraits la notion de continuité, il nous a fallu supposer qu'on ne donnait pas seulement dans chaque espace abstrait ses éléments ou points, considérés individuellement.

On supposait, en outre, donnée antérieurement à la notion de continuité d'une transformation, une certaine définition des points

⁽¹⁾ *Les dimensions d'un ensemble abstrait* (*Math. Ann.*, Bd 63, 1910, p. 145). Nous appelons espace topologique abstrait un espace dont chaque point est un élément de nature quelconque, mais où une définition des éléments d'accumulation des ensembles de points de cet espace a été fixée.

d'accumulation ou de l'opération de dérivation des ensembles par l'intermédiaire d'une généralisation de la notion de voisinage.

Pour arriver à introduire la notion de différentielle dans l'Analyse générale, il va être nécessaire de restreindre une fois de plus la généralité de l'espace considéré.

Avant de préciser en quoi consiste cette restriction, essayons de dégager ce qu'il y a d'essentiel dans la notion de différentielle.

Dans le but de donner un point de départ moins empirique à l'introduction de la différentielle dans l'Analyse fonctionnelle, M. Hadamard a formulé un premier principe.

D'après lui, le résultat fondamental du calcul différentiel a la forme suivante : « La différentielle d'une fonction est une fonction linéaire des différentielles des variables. » Il est ainsi conduit à considérer comme fonctionnelles auxquelles on peut étendre les méthodes du Calcul infinitésimal, toutes les fonctions $U(y)$ dont la variation est une fonctionnelle linéaire de la variation de y ⁽¹⁾. Il reste à définir ce que l'on nomme fonctionnelle linéaire : c'est une fonctionnelle distributive et continue.

Dans la définition précédente, la variation de la fonction $y(x)$ doit être entendue au sens de Lagrange, c'est-à-dire que c'est la différentielle pour $\alpha = \alpha_0$, d'une fonction $y(x, \alpha)$ dépendant d'un paramètre numérique α et égale pour $\alpha = \alpha_0$ à la fonction donnée.

On voit que le principe de M. Hadamard appliqué à l'Analyse générale pourrait s'exprimer ainsi :

Supposons que nous ayons adopté une définition de la différentielle d'une transformation d'un nombre α en un point abstrait. Si l'on fait dépendre m d'un nombre α , toute fonction abstraite de m , $M = \Phi(m)$ dépendra aussi du nombre α . On aura par exemple $m = g(\alpha)$ et $M = \Phi[g(\alpha)]$. Et il faudra que $d\Phi[g(\alpha)]$ soit une fonctionnelle linéaire de $dg(\alpha)$.

Ceci nous fournit une condition qui est bien dans l'esprit du Calcul différentiel; mais il reste encore à dire ce qui sera la différentielle d'une transformation $m = g(\alpha)$ d'un nombre variable α en un élément abstrait variable m . (Cette question ne se posait pas pour l'Analyse

(1) J. HADAMARD, *Leçons sur le Calcul des Variations*, t. I, Paris, 1910, p. 288.

fonctionnelle où M était un nombre U , m une fonction numérique ordinaire $y(x)$ d'une variable x et où, par conséquent, on savait définir $\frac{dU}{dz}$ et $\frac{dy}{dx}$.)

Le principe de M. Hadamard permettrait donc de ramener le cas de la correspondance $M = \Phi(m)$ entre deux éléments abstraits au cas où le second élément m est un nombre. Mais il resterait à élucider ce cas.

D'ailleurs, même dans le cas considéré par M. Hadamard, il nous a semblé qu'il y aurait intérêt à adjoindre au principe de M. Hadamard un autre principe destiné à mettre immédiatement en évidence l'importance de la notion de différentielle pour l'étude approchée de la fonction. C'est le suivant : la différentielle d'une fonction est une expression *simple, approchée* de l'accroissement de la fonction. Le qualificatif simple est précisé par le principe de M. Hadamard en y remplaçant toutefois la variation ou différentielle de l'argument y de la fonctionnelle U_y par l'accroissement Δy de y sans intervention du paramètre numérique α . Le qualificatif « approché » recevra une première précision si l'on dit que la différentielle d'une fonction doit être la *partie principale* de l'accroissement de cette fonction. Mais pour éviter les objections élevées contre cette condition dans le cas classique des fonctions d'une variable et visant les cas où la dérivée est nulle, nous préciserons ainsi le sens du mot partie principale : La différence entre la différentielle et l'accroissement de la fonction doit être infiniment petite par rapport à l'accroissement de la variable.

L'analyse précédente conduit alors pour les fonctionnelles numériques à la définition que nous avons donnée en 1911.

Une fonctionnelle $\Phi(m)$ admet une différentielle pour l'argument m_0 s'il existe une fonctionnelle $\varphi(\Delta m)$, linéaire par rapport à l'accroissement Δm de l'argument m_0 , qui ne diffère de l'accroissement correspondant $\Delta\Phi(m_0)$ que par une quantité infiniment petite par rapport à l'accroissement Δm_0 de l'argument.

Pour achever cette définition, il suffit de préciser : 1° qu'une fonctionnelle $\varphi(\xi)$ est linéaire quand elle est continue et, en outre, distributive, c'est-à-dire telle que l'on ait identiquement

$$\varphi(\xi_1 + \xi_2) = \varphi(\xi_1) + \varphi(\xi_2);$$

2° que le nombre $\Delta\Phi(m_0)$ sera considéré comme infiniment petit par rapport à Δm_0 s'il est infiniment petit par rapport au nombre qui mesure l'écart entre les éléments m_0 et $m_0 + \Delta m_0$.

Enfin, bien entendu, la différentielle de $\Phi(m)$ en m_0 sera précisément $\varphi(\Delta m)$.

La définition est alors complète dans l'Analyse fonctionnelle — où $\Phi(m)$ est un nombre — lorsque l'argument m est une fonction ordinaire. Il suffit de choisir pour l'écart de deux fonctions $m(x)$ et $m_0(x)$ une définition en rapport avec la nature des fonctions envisagées.

Par exemple, lorsque m varie dans le champ des fonctions continues sur un intervalle fixe J , on prendra, pour l'écart de m et m_0 , le maximum de $|m(x) - m_0(x)|$ sur J . Lorsque m est une fonction arbitraire de carré sommable sur J , on prendra pour écart

$$\sqrt{\int_J [m(x) - m_0(x)]^2 dx} \dots$$

La définition de la différentielle dans l'Analyse générale. — Nous allons voir qu'on peut étendre sans modification la définition donnée plus haut aux fonctionnelles qui sont, aussi bien que leur argument, de nature quelconque, c'est-à-dire aux transformations d'éléments abstraits en éléments abstraits. La seule difficulté consiste à préciser le sens à attribuer aux différentes parties de cette définition. C'est en le faisant que nous serons amenés à restreindre la catégorie des ensembles abstraits transformés et à transformer, sans pourtant avoir à définir la nature de leurs éléments.

Si la définition est bien choisie, elle ne s'appliquera qu'à des transformations continues; nous prévoyons donc qu'on devra avoir défini d'avance dans chaque espace l'opération de dérivation des ensembles. Mais cela ne sera pas suffisant.

Nous remarquons qu'en regardant la différentielle d'une transformation comme une transformation simple, approchée, au voisinage d'un point, de la transformation donnée, nous sommes sûrs de rester dans la voie tracée par les créateurs de la notion de différentielle d'une fonction et nous pouvons attendre de cette différentielle généralisée

les mêmes services qu'a rendus la différentielle classique dans un domaine plus restreint. C'est seulement dans l'interprétation de la définition que nous risquons de faire fausse route.

Analyse de la définition de la différentielle. — Reprenons donc en détail la définition de la page 296 en donnant à la fonctionnelle $M = \Phi(m)$ la signification d'une transformation univoque de points m (formant un ensemble e) d'un espace abstrait en points M (formant un ensemble E) d'un espace abstrait, distinct ou non du premier.

On peut représenter cette définition sous la forme symbolique

$$(I) \quad \Delta\Phi(m_0) - d\Phi(m_0) = o(\Delta m_0)$$

où nous employons la notation $o(\dots)$ empruntée à la théorie des fonctions pour indiquer que $o(\Delta m_0)$ est infiniment petit par rapport à Δm_0 ; et où $d\Phi(m_0)$ est une fonctionnelle linéaire de Δm_0 . On peut aussi représenter symboliquement la définition des fonctionnelles linéaires $\varphi(\xi)$ par la double condition

$$(II) \quad \varphi(\xi_1 + \xi_2) = \varphi(\xi_1) + \varphi(\xi_2).$$

(III) Si ξ est voisin d'un ensemble d'éléments η , $\varphi(\xi)$ est voisin de l'ensemble d'éléments $\varphi(\eta)$.

Il reste à interpréter ces trois conditions.

Tout d'abord, m_0 et m étant deux positions de m , qu'entendra-t-on par accroissement Δm_0 de m_0 à m , lorsque m n'est pas un nombre?

On pourra s'inspirer de la géométrie vectorielle usuelle où cet accroissement ne serait autre que le *vecteur géométrique* $m_0 m$.

Alors Δm_0 est un vecteur abstrait du premier espace, $\Delta M_0 = \Delta\Phi(m_0)$ est un vecteur du second espace, $d\Phi(m_0) = \varphi(\Delta m_0)$ est un vecteur du second espace transformé linéaire d'un vecteur du premier. Et pour interpréter complètement la signification du premier membre de (I), il faudra supposer qu'on ait antérieurement adopté une définition de la différence entre deux vecteurs du second espace, à savoir ici $\Delta\Phi(m_0) - \varphi(\Delta m_0)$.

D'autre part, la condition (II) suppose qu'on a donné un sens à la

somme de deux vecteurs ξ_1, ξ_2 du premier espace et à la somme de deux vecteurs $\varphi(\xi_1), \varphi(\xi_2)$ du second espace.

On voit alors qu'un moyen très général — sinon le plus général — de donner les significations qui viennent d'être précisées consiste à *supposer que les deux espaces considérés sont des espaces abstraits affines* au sens que j'ai indiqué ailleurs et que nous rappellerons plus loin.

Mais nous n'avons pas encore tenu compte de la condition (III). Celle-ci n'acquerrait de sens dans les espaces affines tels qu'ils sont habituellement définis que si l'on ne faisait tendre un élément ou point abstrait vers un autre que suivant un chemin rectiligne. Si l'on veut restituer à la notion de continuité toute sa portée, nous serons donc ici amenés à admettre que les deux espaces abstraits où varient m et M sont non seulement affines mais topologiquement affines au sens que nous avons précisé ailleurs. Cela se traduit par une nouvelle restriction qui sera rappelée plus loin.

Enfin, reste à interpréter la signification de $o(\Delta m_0)$ dans l'égalité (I). C'est un vecteur A du second espace qui dépend d'un vecteur $\lambda = \Delta m_0$ du premier espace. Et le vecteur abstrait A doit être infiniment petit par rapport au vecteur λ . Dans l'espace topologiquement affine le plus général, on peut attacher à chaque vecteur un *nombre* ≥ 0 qu'on peut appeler longueur ou norme de ce vecteur. On pourrait donc être tenté de dire que le vecteur abstrait A est infiniment petit par rapport au vecteur abstrait λ si la norme ou longueur de A est infiniment petite par rapport à celle de λ . Mais on sait que, dans un espace affine, les valeurs absolues des longueurs n'ont pas de signification propre, que seuls importent les rapports de deux longueurs et encore, en les supposant prises sur la même droite abstraite ou sur deux droites parallèles. Il faut donc donner une définition plus stricte du symbole $o(\Delta m_0)$. Un moyen très général consistera à supposer que, dans chacun des deux espaces topologiques considérés, les points d'accumulation peuvent être définis au moyen d'un écart numérique. Nous entendons par là qu'à tout couple de points A, B , on peut faire correspondre un nombre $(A, B) \geq 0$ — nul seulement si A et B ne sont pas distincts — de sorte que A soit point d'accumulation d'un ensemble de points B , si la borne inférieure des écarts (A, B) positifs est nulle et réciproquement.

Alors la signification du symbole $o(\Delta m_0)$ sera évidemment la suivante : le vecteur $\overline{M_0 M_1} = o(\overline{m_0 m})$ devra être tel que le nombre qui mesure l'écart de M_0 et de M_1 soit infiniment petit par rapport au nombre qui mesure l'écart de m_0 et de m quand ce dernier tend vers zéro.

Ainsi, nous voyons qu'on pourra définir la différentiabilité d'une transformation $M = F(m)$ de points m d'un espace abstrait en points M d'un espace abstrait, toutes les fois que ces deux espaces sont des espaces topologiquement affines où la définition des points d'accumulation peut être donnée par l'intermédiaire d'un écart numérique.

Toutefois, on est obligé pour ces espaces d'établir une distinction, au premier abord subtile, entre la longueur et l'écart. Cette distinction est dans la nature des choses comme on pourra le voir par les exemples de tels espaces ($E\omega$, \mathfrak{N} , ...) que nous avons donnés ailleurs ⁽¹⁾. Mais au moment d'introduire la notion si importante de différentielle d'une transformation d'espaces abstraits en espaces abstraits, cette distinction risque d'égarer l'esprit sur des détails non essentiels. Nous nous bornerons donc dans la suite au cas où les espaces abstraits considérés sont des « espaces (ω) vectoriels », espaces un peu moins généraux que « les espaces \mathcal{E} topologiquement affines » considérés ci-dessus et un peu plus généraux que les espaces définis par M. Banach ⁽²⁾ et que nous appellerons « espaces (ω) vectoriels complets ».

Résumé. — Nous allons dans la suite rappeler la définition précise des espaces considérés plus haut successivement et formuler la définition correspondante de la différentielle d'une transformation d'ensembles abstraits. Nous étudierons ensuite ses propriétés. Dans un second Mémoire, nous définirons les transformations d'ordre entier et les différentielles d'ordres supérieurs. Dans toute cette étude, nous utiliserons, \rightarrow sans avoir beaucoup à y changer pour en étendre considérablement la portée — un grand nombre des raisonnements présentés dans nos quatre mémoires : *Une définition*

⁽¹⁾ *Les espaces topologiquement affines* (*Acta mathematica*, 1925).

⁽²⁾ *Fundamenta Mathematicae*, t. III, 1922, p. 133.

fonctionnelle des polynomes (Nouv. Ann. de Math., t. 9, 1909); *Les fonctionnelles continues* (Ann. Éc. Norm. sup., t. 26, 1909, p. 193); *Sur la notion de différentielle dans le Calcul fonctionnel* (C. R. Congrès Soc. sav., 1912, p. 45); *Les fonctions d'une infinité de variables* (C. R. Congrès Soc. sav. 1909, p. 5).

CHAPITRE I.

ESPACES ABSTRAITS ET TRANSFORMATIONS PARTIELLES ABSTRAITES.

Définition d'un champ de vecteurs abstraits. — De nombreuses définitions axiomatiques des champs de vecteurs ont été données qui ne diffèrent que par quelques variantes. Nous nous arrêterons à la suivante :

Nous appellerons *champ de vecteurs* un système composé d'un ensemble σ d'éléments de nature quelconque (et que nous appellerons *vecteurs*) et de trois opérations portant sur ces éléments et sur des nombres, opérations représentées par les symboles $+$, \cdot , $\| \dots \|$ et soumises aux conditions suivantes :

Soient ξ, η, ζ trois vecteurs quelconques du champ σ , a et b des nombres quelconques ⁽¹⁾.

1° $\xi + \eta$ est un élément bien déterminé du champ σ ; 2° $\xi + \eta = \eta + \xi$;
 3° $(\xi + \eta) + \zeta = \xi + (\eta + \zeta)$; 4° $\xi + \eta = \xi + \zeta$ entraîne $\eta = \zeta$;
 5° il existe dans le champ σ un vecteur qu'on peut désigner par O tel que $\xi + O = \xi$; 6° $a \cdot \xi$ est un vecteur déterminé du champ σ ⁽¹⁾;
 7° $a \neq 0$ et $a \cdot \xi = a \cdot \eta$ entraînent $\xi = \eta$; 8° $\xi \neq O$ et $a \cdot \xi = b \cdot \xi$ entraînent $a = b$; 9° $a \cdot (\xi + \eta) = a \cdot \xi + a \cdot \eta$; 10° $(a + b) \cdot \xi = a \cdot \xi + b \cdot \xi$;
 11° $1 \cdot \xi = \xi$; 12° $ab \cdot \xi = a \cdot (b \cdot \xi)$; 13° $\| \xi \|$ est un nombre ≥ 0 appelé longueur de ξ ; 14° $\| \xi \| = 0$ équivaut à $\xi = O$; 15° $\| a \cdot \xi \| = |a| \| \xi \|$.

Définition d'un espace abstrait affine. — Nous appellerons *espace abstrait affine* un système composé d'un ensemble d'éléments de nature quelconque appelés *points abstraits* et d'un champ σ de vecteurs

⁽¹⁾ Pour éviter des complications qui sont secondaires relativement à l'idée essentielle de ce Mémoire, nous ne définirons $a \cdot \xi$ que pour a réel.

associé à cet ensemble de façon à réaliser les conditions suivantes.

Appelons *couple ordonné de points*, deux points abstraits pris dans un certain ordre.

I. A tout couple ordonné A, B de points abstraits correspond un vecteur déterminé ξ du champ σ et l'on exprime cette correspondance par la notation $\overline{AB} = \xi$.

II. Étant donnés un point abstrait A quelconque et un vecteur ξ quelconque du champ σ , il existe un point abstrait B et un seul tel que $\overline{AB} = \xi$.

III. Quel que soit le point abstrait A, $\overline{AA} = O$.

IV. Quels que soient les points abstraits A, B, C, on a

$$\overline{AB} = \overline{AC} + \overline{CB}.$$

Par analogie avec l'espace euclidien, on peut aussi définir dans un espace abstrait affine des droites abstraites, des translations abstraites et des homothéties.

On appellera *droite abstraite* passant par deux points abstraits distincts A et B, le lieu des points abstraits M_ρ obtenus en faisant varier le nombre ρ dans la relation

$$\overline{AM_\rho} = \rho \cdot \overline{AB}.$$

On appellera *translation abstraite* déterminée par un vecteur abstrait ξ , la transformation ponctuelle qui fait correspondre à un point abstrait M, un point abstrait M' par la relation

$$\overline{MM'} = \xi.$$

On appellera *homothétie abstraite* de centre A et de rapport k , la transformation ponctuelle qui fait correspondre à un point abstrait M un point abstrait M'' par la relation

$$\overline{AM''} = k \cdot \overline{AM} \text{ (}^1\text{)}.$$

(¹) On peut généraliser de la même façon un certain nombre de définitions d'éléments géométriques euclidiens et de leurs propriétés. Voir *Sur une définition géométrique des espaces abstraits affines* (*Ann. Soc. Math. polonaise*, 1925).

Intervention des considérations de continuité. — Nous avons fait remarquer ailleurs⁽¹⁾ qu'un ensemble à peu près quelconque, E , peut être considéré comme ensemble des points abstraits d'un espace affine. (Si, par exemple, E a la puissance du continu, il suffit de lui faire correspondre l'ensemble des points d'une droite euclidienne et de lui associer un champ de vecteurs de la même façon qu'aux points de la droite.) La notion d'espace affine n'offre donc un intérêt que lorsque les opérations $+$, \cdot , $\| \dots \|$ effectuées sur les vecteurs de l'espace considéré sont en relation avec la nature des points de l'espace affine. En particulier, la nature de ces points détermine sinon logiquement, du moins pratiquement, la façon de traduire dans l'espace considéré la notion de voisinage des points de l'ensemble.

Par conséquent, il n'y aura intérêt à adopter une façon de considérer un ensemble de points comme espace affine que si cette façon est en relation régulière avec la définition adoptée dans cet ensemble pour le voisinage de ses points.

Un cas simple est celui où le voisinage plus ou moins étroit de deux points de l'espace considéré peut être apprécié numériquement par un nombre et même plus précisément quand l'espace considéré est un espace (\mathfrak{O}) .

Nous disons qu'un espace où une certaine définition de la convergence d'une suite de points a été adoptée *est un espace* (\mathfrak{O}) lorsque :

A tout couple de points A, B de l'espace correspond un certain nombre $(A, B) = (B, A) \geq 0$ qu'on appelle *distance* de A et de B .

$(A, B) = 0$ équivaut à dire que A et B ne sont pas distincts.

$\lim_{n \rightarrow \infty} (A, A_n) = 0$ équivaut à dire que A_n tend vers A quand n tend vers l'infini.

Quels que soient les points A, B, C , on a

$$(A, B) \leq (A, C) + (C, B).$$

Ceci étant, un cas particulièrement simple, quoique très général, où le champ vectoriel associé à un espace (\mathfrak{O}) est en relation directe et naturelle avec les notions de continuité sur cet espace est le suivant.

⁽¹⁾ *Les espaces topologiquement affines* (*Acta mathematica*, 1925).

Nous appellerons *espace* (\mathfrak{O}) *vectoriel*, un espace (\mathfrak{O}) auquel on peut associer un champ de vecteurs, de sorte que la longueur de tout vecteur AB soit égale à la distance de ses extrémités A, B . Autrement dit, nous supposons que $\| \overline{AB} \| = (A, B)$.

[Parmi les espaces (\mathfrak{O}) vectoriels figurent ceux qui ont été étudiés par M. S. Banach dans le Tome III de *Fundamenta Mathematicæ*, et qu'on peut appeler espaces (\mathfrak{O}) vectoriels *complets*.]

Il n'est pas indispensable de confondre, comme dans les espaces (\mathfrak{O}) vectoriels, les notions de longueur et de distance. On peut même plus généralement considérer des espaces abstraits affines où le voisinage ne peut s'exprimer numériquement.

Nous étudierons ces espaces plus généraux dans un Mémoire ultérieur, parce qu'on peut y définir et y étudier sans trop de complications les transformations ponctuelles abstraites d'ordre entier.

Mais nous désirons mettre bien en lumière la simplicité de l'extension que nous voulons donner à la notion de différentielle et qui est l'objet essentiel de ce Mémoire. Nous écarterons donc les difficultés accessoires en nous bornant, pour définir la différentielle, au cas des transformations ponctuelles d'espaces (\mathfrak{O}) vectoriels en espaces de même nature.

On trouvera des exemples de tels espaces dans le Mémoire de M. Banach cité ci-dessus et dans notre dernier Mémoire du *Bull. Calcutta Math. Soc.*, 1925, sur « Les espaces vectoriels abstraits ».

Transformation vectorielle linéaire. — Considérons une transformation $V = \Psi(v)$ qui fait correspondre à chaque vecteur v d'un champ σ de vecteurs abstraits défini comme plus haut, un vecteur V d'un champ analogue de vecteurs abstraits, les deux champs étant distincts ou non. Nous dirons que cette transformation vectorielle est *linéaire* si elle est d'abord *distributive*, c'est-à-dire telle que, pour tout couple v_1, v_2 de vecteurs du premier champ

$$\Psi(v_1 + v_2) = \Psi(v_1) + \Psi(v_2).$$

Elle doit être ensuite *continue*, c'est-à-dire telle que $\| \Psi(v_2) - \Psi(v_1) \|$ tende vers zéro quand $\| v_2 - v_1 \|$ tend vers zéro, ou encore telle que $\| \Psi(v) \|$ tende vers zéro quand $\| v \|$ tend vers zéro.

On conclut de la distributivité qu'en prenant pour v_2 le vecteur O

du premier champ, on a

$$\Psi(v_1) = \Psi(v_0) + \Psi(0);$$

donc le transformé $\Psi(0)$ du vecteur 0 du premier champ est le vecteur 0 du second champ.

On conclut aussi de la distributivité et par un raisonnement bien connu que l'on a

$$(1) \quad \Psi(r \cdot v) = r \cdot \Psi(v),$$

quel que soit le nombre rationnel r . Alors grâce à la continuité de Ψ , cette égalité reste vraie à la limite pour les valeurs irrationnelles de r .

L'identité (1) nous permet de prouver maintenant une propriété utile des transformations vectorielles linéaires : il existe un nombre K indépendant du vecteur v tel que

$$\|\Psi(v)\| \leq K \|v\|.$$

En effet, dans le cas contraire, il existerait une suite de vecteurs distincts v_1, v_2, \dots telle que

$$\|\Psi(v_n)\| > n \|v_n\|.$$

En posant $w_n = \frac{1}{\|v_n\|\sqrt{n}} v_n$, on aurait une suite de vecteurs w_1, w_2, \dots , telle que $\|w_n\| = \frac{1}{\sqrt{n}}$ tende vers zéro et que pourtant

$$\|\Psi(w_n)\| > \sqrt{n},$$

qui ne tend pas vers zéro.

Transformations ponctuelles abstraites. — Soient $M = F(m)$ une transformation d'un point m d'un espace (\mathfrak{O}) vectoriel en un point M d'un espace (\mathfrak{O}) vectoriel, m_0 une position particulière de m et $M_0 = F(m_0)$. On peut ramener la transformation ponctuelle $M = F(m)$ à une transformation vectorielle en l'écrivant sous la forme

$$\overline{M_0 M} = \overline{F(m_0)F(m)} = \Theta(\overline{m_0 m}),$$

ou même, en partant de points respectifs m_1, M_2 qui ne se correspondent pas, sous la forme

$$\overline{M_2 M} = A + \Theta_1(\overline{m_1 m}).$$

On peut appeler *accroissement de la variable* de m_0 à m le vecteur $\overline{m_0 m}$ et accroissement correspondant de la fonction le vecteur $\overline{M_0 M}$. On peut désigner ces accroissements par $\Delta m_0 = \overline{m_0 m}$, $\Delta \overline{M_0} = \overline{M_0 M} = \Delta F(m_0)$.

En employant la notation $m_1 = m + \overline{m m_1} = m + \Delta_1 m$, on a

$$\Delta_1 F(m) = \overline{F(m) F(m + \Delta_1 m)}.$$

En envisageant $\Delta_1 F(m)$ comme fonction de m , on peut de même considérer un accroissement $\Delta_2 [\Delta_1 F(m)]$ de cette quantité, correspondant à un second accroissement $\Delta_2 m$ de m . D'où

$$\begin{aligned} \Delta_2 \Delta_1 F(m) &= \overline{F(m + \Delta_2 m) F(m + \Delta_2 m + \Delta_1 m)} - \overline{F(m) F(m + \Delta_1 m)} \\ &= \overline{F(m + \Delta_1 m) F(m + \Delta_1 m + \Delta_2 m)} - \overline{F(m) F(m + \Delta_2 m)} \\ &= \Delta_1 \Delta_2 F(m). \end{aligned}$$

On définira de même en général l'*accroissement d'ordre n* , Δ_n , Δ_{n-1} , ..., $\Delta_1 F(m)$ correspondant aux accroissements successifs $\Delta_1 m$, $\Delta_2 m$, ..., $\Delta_n m$ et indépendant de l'ordre des indices.

Transformations ponctuelles du premier ordre. — Nous appellerons en général transformation ponctuelle d'ordre n , toute transformation $M = F(m)$ continue et telle que l'accroissement d'ordre $n + 1$ de la fonction $F(m)$ $\Delta_{n+1} \Delta_n \dots \Delta_1 F(m)$ soit identiquement nul, c'est-à-dire égal au vecteur O du second champ quand les accroissements $\Delta_1 m$, ..., $\Delta_{n+1} m$ de la variable varient de façon indépendante.

En particulier, $M = F(m)$ sera une transformation du premier ordre si $\Delta_2 \Delta_1 F(m) \equiv O$, c'est-à-dire si la transformation, toujours supposée continue, est telle que

$$\overline{F(m + \xi_1) F(m + \xi_1 + \xi_2)} = \overline{F(m) F(m + \xi_1)},$$

quels que soient les vecteurs ξ_1 et ξ_2 . En posant pour m fixe

$$\Psi(\xi) = \overline{F(m) F(m + \xi)},$$

on aura

$$(2) \quad \Psi(\xi_1 + \xi_2) = \Psi(\xi_1) + \Psi(\xi_2).$$

D'autre part, en disant que $F(m)$ est continue, nous entendons que $\|\overline{F(m) F(m_1)}\|$ tend vers zéro avec $\|\overline{m m_1}\|$. De sorte que $\Psi(\xi)$ sera

non seulement distributive en vertu de (2), mais aussi continue. Par suite, $\Psi(\xi)$ est une transformation vectorielle linéaire. En employant pour M une notation similaire à $m_1 = m + \overline{mm_1}$, nous aurons donc pour toute transformation $M = F(m)$ du premier ordre

$$F(m + \xi) = F(m) + \Psi(\xi),$$

où $V = \Psi(v)$ est une transformation vectorielle linéaire. La réciproque est évidente.

En employant la terminologie géométrique que nous avons utilisée ailleurs (1), on observera que toute transformation du premier ordre transforme toute droite en une droite, tout plan en un plan, deux droites parallèles en deux droites parallèles, etc.

Composition de deux transformations du premier ordre. — Considérons une transformation du premier ordre du point α dans le point m

$$m = m'_0 + \Psi_1(\overline{\alpha'_0 \alpha'_1}).$$

Et cherchons quelle est la transformation, résultant de la succession de cette transformation et de la transformation du premier ordre

$$N = N_0 + \Psi(\overline{m_0 m}).$$

On a

$$\Psi(\overline{m_0 m}) = \Psi(\overline{m_0 m'_0}) + \Psi[\Psi_1(\overline{\alpha'_0 \alpha'_1})],$$

d'où

$$N = N'_0 + \Psi[\Psi_1(\overline{\alpha'_0 \alpha'_1})].$$

Quels que soient les vecteurs ξ_1, ξ_2 du champ des α

$$\Psi[\Psi_1(\xi_1 + \xi_2)] = \Psi[\Psi_1(\xi_1) + \Psi_1(\xi_2)] = \Psi[\Psi_1(\xi_1)] + \Psi[\Psi_1(\xi_2)],$$

donc $\Psi[\Psi_1(\xi)]$ est distributive. Elle est d'autre part continue. Finalement, la composition de deux transformations ponctuelles du premier ordre fournit une transformation du premier ordre.

(1) *Ann. Soc. Math. polonaise*, 1925, loc. cit.

CHAPITRE II.

PROPRIÉTÉS DES TRANSFORMATIONS ABSTRAITES DIFFÉRENTIABLES.

Définition de la différentielle d'une transformation ponctuelle abstraite. — Conformément aux explications que nous avons données dans l'Introduction, nous dirons qu'une transformation ponctuelle $M = F(m)$ d'un point m d'un espace (\mathbb{O}) vectoriel en un point M d'un espace (\mathbb{O}) vectoriel, distinct ou non du précédent, est différentiable au point m_0 , s'il existe une transformation vectorielle linéaire $\Psi(\Delta m_0)$ de l'accroissement Δm_0 de la variable, qui ne diffère de l'accroissement correspondant $\Delta F(m_0)$ de la fonction que par un vecteur $o(\Delta m_0)$ infiniment petit par rapport à Δm_0 .

On a ainsi

$$(3) \quad \Delta F(m_0) = \Psi(\Delta m_0) + o(\Delta m_0),$$

la notation $o(\Delta m_0)$ signifiant que le rapport ε de la longueur du vecteur $o(\Delta m_0)$ à la longueur du vecteur Δm_0 tend vers zéro avec cette dernière longueur. On peut exprimer cette propriété en remarquant que

$$o(\Delta m_0) = \|o(\Delta m_0)\| \cdot U = \varepsilon \|\Delta m_0\| \cdot U,$$

U étant un vecteur du second espace dont la longueur est égale à l'unité.

Remarquons que pour appliquer la définition précédente, il n'est pas nécessaire de supposer que la transformation $M = F(m)$ soit définie pour toute position de m . Mais nous supposons qu'elle est au moins définie au voisinage de m_0 , c'est-à-dire que m_0 soit intérieur à l'ensemble e de point m sur lequel $F(m)$ est définie. Dans ce cas, en effet, on va pouvoir prouver facilement que si $F(m)$ est différentiable en m_0 , il n'y a aucune transformation vectorielle linéaire Θ , autre que Ψ , qui puisse jouer le rôle de Ψ .

En effet, $\Psi(\Delta m_0) - \Theta(\Delta m_0)$ devrait être infiniment petit par rapport à Δm_0 . Or, en posant $v = \frac{1}{\|\Delta m_0\|} \cdot \Delta m_0$, on a

$$\frac{\|\Psi(\Delta m_0) - \Theta(\Delta m_0)\|}{\|\Delta m_0\|} = \frac{\|\Psi(v) - \Theta(v)\|}{\|v\|}.$$

Ce dernier rapport, qui est indépendant du nombre $\|\Delta m_0\|$, devrait tendre vers zéro quand ce nombre tend vers zéro; il est donc nul et, par suite, on a bien

$$\Psi(\Delta m_0) \equiv \Theta(\Delta m_0).$$

La transformation vectorielle linéaire $\Psi(\Delta m_0)$ est donc seule de son espèce; c'est une transformation simple approchée de la transformation du vecteur Δm_0 dans le vecteur ΔM_0 ; nous l'appellerons *la différentielle de F(m) en m₀* et nous la représenterons par $dF(m_0)$. Finalement, on a

$$(7) \quad \boxed{\Delta F(m_0) = dF(m_0) + \varepsilon \|\Delta m_0\| \cdot U}$$

U étant un vecteur (du second espace) dont la longueur est égale à l'unité et ε un nombre qui tend vers zéro quand la longueur de Δm_0 tend vers zéro.

On voit aussi qu'on a

$$F(m) = [F(m_0) + \Psi(\overline{m_0 m})] + \varepsilon \|\overline{m_0 m}\| U$$

de sorte que la transformation du premier ordre

$$(4) \quad N = F(m_0) + \Psi(\overline{m_0 m})$$

peut être considérée comme une transformation simple approchée de la transformation $M = F(m)$ au voisinage de m_0 . On peut l'appeler la transformation du premier ordre *tangente en m₀* à la transformation $M = F(m)$.

Cette façon de parler, utilisée en élasticité, rendrait service même dans l'*Analyse classique* où elle rendrait certains énoncés plus intuitifs. Par exemple, soit $M = F(m)$ une correspondance entre points de l'espace euclidien. Elle se traduit par trois relations numériques entre les coordonnées

$$(5) \quad X = F_1(x, y, z), \quad Y = F_2(x, y, z); \quad Z = F_3(x, y, z).$$

La transformation considérée sera différentiable au point x_0, y_0, z_0 si les fonctions F_1, F_2, F_3 le sont, et la transformation du premier ordre

tangente en ce point est définie par

$$(6) \quad dX = \frac{\partial F_1}{\partial x_0} dx + \frac{\partial F_1}{\partial y_0} dy + \frac{\partial F_1}{\partial z_0} dz; \quad dY = \dots, \quad dZ = \dots$$

On dit généralement que si le déterminant jacobien $\frac{D(F_1, F_2, F_3)}{D(x, y, z)}$ est différent de zéro au point x_0, y_0, z_0 , la transformation (5) est biunivoque au voisinage de ce point. Il serait aussi rigoureux et plus intuitif de dire que si la transformation tangente à la transformation (5) au point x_0, y_0, z_0 , est biunivoque, la transformation (5) elle-même est biunivoque au voisinage de ce point.

Composition des transformations différentiables. — Il résulte évidemment de la définition de la différentielle que si la transformation $M = F(m)$ est différentiable en m_0 , elle est aussi continue en m_0 .

Considérons maintenant la transformation $M = G(\alpha)$ qui résulte de la succession (de la composition) de deux transformations différentiables, $m = g(\alpha)$ effectuée sur un espace ω vectoriel et la transformation $M = F(m)$ envisagée précédemment. On a la relation (3) et une relation de même forme

$$\overline{m_0 m_1} = \Psi_1(\overline{\alpha_0 \alpha_1}) + o_1(\overline{\alpha_0 \alpha_1});$$

$$\overline{M_0 M_1} = \Psi[\Psi_1(\overline{\alpha_0 \alpha_1})] + o_2(\overline{\alpha_0 \alpha_1});$$

d'où :

$$o_2(\overline{\alpha_0 \alpha_1}) = \Psi[o_1(\overline{\alpha_0 \alpha_1})] + o(\overline{m_0 m_1}).$$

Il est bien manifeste que $\Psi[\Psi_1(\overline{\alpha_0 \alpha_1})]$ est une transformation linéaire comme Ψ et Ψ_1 (voir p. 307), il reste à voir comment se comporte $o_2(\overline{\alpha_0 \alpha_1})$ quand α_1 tend vers α_0 . Or, d'après l'alinéa précédent et la définition de o , il existe deux nombres fixes h et k et un nombre ε tendant vers zéro avec $\|\overline{\alpha_0 \alpha_1}\|$ tels que

$$\|\Psi(\Theta)\| \leq h \|\Theta\|, \quad \|\Psi_1(\overline{\alpha_0 \alpha_1})\| \leq k \|\overline{\alpha_0 \alpha_1}\|, \quad \|o_1(\overline{\alpha_0 \alpha_1})\| = \varepsilon \|\overline{\alpha_0 \alpha_1}\|,$$

d'où

$$\|\overline{m_0 m_1}\| \leq (k + \varepsilon) \|\overline{\alpha_0 \alpha_1}\|.$$

Alors

$$\|o_2(\overline{\alpha_0 \alpha_1})\| \leq \left\{ h \frac{\|\Psi_1(\overline{\alpha_0 \alpha_1})\|}{\|\overline{\alpha_0 \alpha_1}\|} + \frac{\|o(\overline{m_0 m_1})\|}{\|\overline{m_0 m_1}\|} (k + \varepsilon) \right\} \|\overline{\alpha_0 \alpha_1}\|,$$

d'où l'on déduit que $\alpha_2(\overline{\alpha_0 \alpha_1})$ est infiniment petit par rapport à $\|\overline{\alpha_0 \alpha_1}\|$.

Ainsi la composition de deux transformations différentiables successives donne une transformation différentiable. De plus, la différentielle de $F[g(\alpha)]$ est $\Psi(\Psi_1(\overline{\alpha_0 \alpha_1}))$; donc la différentielle de la fonction de fonction $F[g(\alpha)]$ s'obtient par la composition des deux transformations différentielles successives de $F(m)$ et de $g(\alpha)$. On peut aussi dire que la transformation du premier ordre tangente en α_0 à la transformation ponctuelle composée $F[g(\alpha)]$, s'obtient par la composition des transformations du premier ordre, tangentes respectivement à $g(\alpha)$ en α_0 et à $F(m)$ en $m_0 = g(\alpha_0)$.

Remarque. — Dans le cas particulier de la transformation identique $m = m$, la différentielle de m est évidemment son accroissement. Car Δm est bien une transformation linéaire Δm et elle ne diffère pas de l'accroissement de cette fonction particulière. On a donc

$$dm = \Delta m$$

lorsque m est la variable indépendante et, par conséquent, on a aussi

$$dM = dF(m) = \Psi(dm),$$

toujours quand m est la variable indépendante.

Or on a vu que si m est une fonction de α , qui est différentiable comme $F(m)$, on a

$$d[F(g(\alpha))] = \Psi[\Psi_1(\Delta\alpha)] \quad \text{et} \quad dm = \Psi_1(\Delta\alpha),$$

d'où

$$dM = dF[g(\alpha)] = \Psi(dm).$$

Ainsi, lorsque $\Psi(\Delta m)$ est la différentielle d'une fonction abstraite M de m , la relation

$$dM = \Psi(dm)$$

a lieu, non seulement quand m est une variable indépendante, mais même quand c'est une fonction différentiable d'une autre variable abstraite α .

Variable numérique. — Considérons le cas particulier où α est un nombre réel. Il peut être regardé comme appartenant à un espace (∞)

vectorel, celui formé de tous les nombres réels, avec une interprétation évidente des vecteurs.

Alors si $m = g(\alpha)$ est une transformation de cet espace en un ensemble appartenant à un espace abstrait (ω) vectoriel quelconque α et si la transformation $m = g(\alpha)$ est différentiable, on aura

$$\Delta m = \Psi_1(\Delta\alpha) + \omega \Delta\alpha \cdot U,$$

ω étant un nombre qui tend vers zéro avec $\Delta\alpha$ et U étant un vecteur unitaire. Et puisque Ψ_1 est linéaire et $\Delta\alpha$ est un nombre

$$\Psi_1(\Delta\alpha) = \Delta\alpha \cdot \xi(\alpha),$$

où $\xi(\alpha)$ est un certain vecteur de l'espace de m . D'où, finalement,

$$\Delta m = \Delta\alpha \cdot \xi(\alpha) + \omega \Delta\alpha \cdot U$$

et par suite

$$(8) \quad \frac{1}{\Delta\alpha} \cdot \Delta m = \xi(\alpha) + \omega \cdot U.$$

On voit que la différence des vecteurs $\frac{1}{\Delta\alpha} \cdot \Delta m$ et $\xi(\alpha)$ est un vecteur dont la longueur $|\omega|$ tend vers zéro avec $\Delta\alpha$, ce qu'on peut exprimer sous la forme

$$\lim_{\Delta\alpha \rightarrow 0} \frac{1}{\Delta\alpha} \cdot \Delta m = \xi(\alpha).$$

On peut dire que $\xi(\alpha)$ est la *dérivée géométrique* par rapport au paramètre α du vecteur abstrait qui joint le point abstrait m à un point fixe ω . Ou mieux, que c'est le vecteur abstrait qui représente la vitesse abstraite du point abstrait m se déplaçant sur une trajectoire abstraite et arrivant à sa position au temps α . *Dire que la transformation $m = g(\alpha)$ du nombre α en un point abstrait m — ou la représentation paramétrique d'une courbe abstraite — est différentiable, c'est dire que si α est le temps, le mouvement a lieu avec une vitesse vectoriellement déterminée $\xi(\alpha)$, définie comme la limite du vecteur $\frac{1}{\Delta\alpha} \cdot \Delta m$ lorsque Δm tend vers zéro.*

Remarquons aussi que l'on a, d'après (8),

$$\left| \|\xi(\alpha)\| - |\omega| \right| \leq \frac{\|\Delta m\|}{|\Delta\alpha|} \leq \|\xi(\alpha)\| + |\omega|,$$

ce qui donne, lorsque Δz tend zéro, l'égalité *numérique*

$$\|\xi(\alpha)\| = \lim_{\Delta\alpha \rightarrow 0} \frac{\|\Delta m\|}{|\Delta\alpha|}.$$

En particulier, supposons que la fonction $m = g(\alpha)$ soit une constante (abstraite), c'est-à-dire que m reste fixe lorsque α varie dans l'intervalle J où l'on suppose $g(\alpha)$ défini. Alors $\|\Delta m\|$ est nul et par suite la *dérivée* $\xi(\alpha)$ de la fonction $m = g(\alpha)$ est constamment le vecteur O .

Réciproquement, si la dérivée $\xi(\alpha)$ est le vecteur O pour toutes les valeurs de α dans l'intervalle J , la fonction $g(\alpha)$ est une constante (abstraite). En effet, la longueur

$$\|g(\alpha_0)g(\alpha)\|$$

— où α_0 est un point fixe de l'intervalle J et α un point variable —, est une fonction numérique $P(\alpha)$ et l'on a

$$\left| \frac{P(\alpha + \Delta\alpha) - P(\alpha)}{\Delta\alpha} \right| \leq \left\| \frac{g(\alpha)g(\alpha + \Delta\alpha)}{\Delta\alpha} \right\|.$$

Or le second membre tend vers $\|\xi(\alpha)\|$ qui est nul par hypothèse. Donc $P(\alpha)$ est une fonction numérique dont la dérivée est nulle dans l'intervalle J . Elle est donc constante. Ainsi la distance de $g(\alpha_0)$ à chacune des positions de $g(\alpha)$ est la même; en prenant $\alpha = \alpha_0$, on voit que cette distance constante est nulle. Finalement, $g(\alpha)$ est un point abstrait fixe quand α varie sur J .

Nouvelle notation. — Revenons au cas d'une variable abstraite m . Les notations $\Psi(dm)$ et $dF(m)$ pour la différentielle ne mettent chacune en évidence qu'une des données qui déterminent la différentielle. On aura donc souvent avantage à désigner la différentielle de $F(m)$ correspondant à un accroissement dm de m par la notation

$$\partial_{dm}F(m).$$

Variation. — Si nous revenons à la composition des transformations $M = F(m)$; $m = g(\alpha)$ dans le cas où α est un nombre, nous voyons que si F et g sont différentiables, la transformation composée

$M = F[g(\alpha)]$ est différentiable et que si les différentielles sont $dM = \Psi(\Delta m)$ et $dm = \Psi_1(\Delta \alpha) = \Delta \alpha \cdot \xi(\alpha)$ la différentielle composée sera

$$\Psi[\Delta \alpha \cdot \xi(\alpha)].$$

En appelant $\Delta \alpha \cdot \xi(\alpha)$, la « variation » de m , en la désignant par δm , et en posant $\Psi(\Delta m) = \partial_{\Delta m} F(m)$, on voit que *la variation de M sera*

(10)

$$\partial M = \partial_{\delta m} F(m).$$

Ce sera, en particulier, une transformation linéaire de la variation de m , résultat classique dans le Calcul des Variations.

Digression. — Il est intéressant de voir, sur un exemple, le sens de cette proposition. Prenons pour espace où varie m l'espace (\ni) des fonctions continues sur un intervalle fixe (x_0, x_1) , et pour valeur de la distance de deux « points » de cet espace, le maximum de la valeur absolue de leur différence. Prenons pour $F(m)$, en supposant que m soit une fonction numérique de la variable numérique x entre a et a ,

$$F(m) = \int_{x_0}^{x_1} m(x) dx.$$

Il est manifeste que $F(m)$ est une transformation linéaire et par suite que

$$\partial_{\Delta m} F = F(\Delta m) = \int_{x_0}^{x_1} \Delta m(x) dx.$$

Si donc $m(x)$ est remplacé par une fonction $m(x, \alpha)$, on aura

$$\delta \int_{x_0}^{x_1} m(x, \alpha) dx = \partial_{\delta m} F = \int_{x_0}^{x_1} \delta m dx$$

avec

$$\delta m = m'_\alpha(x, \alpha) \delta \alpha.$$

Mais la démonstration suppose que m est différentiable par rapport à α , c'est-à-dire que

$$\Delta m = m(x, \alpha + \Delta \alpha) - m(x, \alpha) = \Delta \alpha \cdot \xi(x, \alpha) + o \Delta \alpha \cdot U,$$

ω tendant vers zéro avec α et U étant un vecteur unitaire. Cela veut dire que

$$\left\| \frac{m(x, \alpha + \Delta\alpha) - m(x, \alpha)}{\Delta\alpha} - \xi(x, \alpha) \right\|$$

tend vers zéro avec $\Delta\alpha$. Or

$$\left| \frac{m(x, \alpha + \Delta\alpha) - m(x, \alpha)}{\Delta\alpha} - \xi(x, \alpha) \right| \leq \left\| \frac{m(x, \alpha + \Delta\alpha) - m(x, \alpha)}{\Delta\alpha} - \xi(x, \alpha) \right\|.$$

Donc $m(x, \alpha)$ doit d'abord être dérivable en α pour chaque valeur de x . Mais, de plus, le rapport $\frac{\Delta m}{\Delta\alpha} - \xi(x, \alpha)$ doit tendre uniformément vers sa limite zéro. On devrait ajouter que $\xi(x, \alpha)$ doit appartenir à l'espace (\mathfrak{E}) : cela est une conséquence de l'hypothèse précédente. Ainsi nous voyons que la formule connue

$$\delta \int_{x_0}^{x_1} m(x, \alpha) dx = \int_{x_0}^{x_1} \delta m(x, \alpha) dx$$

est, ici, établie en supposant non seulement que $m(x, \alpha)$ est continue en x et dérivable en α , mais encore que la dérivée m'_α est la limite *uniforme* de $\frac{\Delta m}{\Delta\alpha}$ quand $\Delta\alpha$ tend vers zéro. On remarquera que, par le détour de la théorie des ensembles abstraits, nous sommes amenés à exprimer sous une forme aussi simple et *plus générale* que celle qu'on donne d'habitude, la condition de dérivabilité d'une intégrale dépendant d'un paramètre. On suppose en effet ordinairement que $m(x, \alpha)$ est dérivable par rapport à α , non seulement pour la valeur considérée de α , mais pour les valeurs voisines et que $m_\alpha(x, \alpha)$ est continue par rapport à l'ensemble de x et de α . Ces conditions entraînent la condition que $\frac{\Delta m}{\Delta\alpha}$ tende uniformément entre x_0 et x_1 vers sa limite $\xi(x, \alpha) = m'_\alpha(x, \alpha)$ pour la valeur α considérée, mais la réciproque n'est pas vraie.

Fonctions abstraites constantes. — Supposons que la transformation

$$M = F(m)$$

soit définie quand m reste à une distance $\leq R$ d'un point fixe m_0 ,

m et M appartenant respectivement à deux espaces \odot vectoriels a et A distincts ou non.

Nous sommes maintenant en mesure de généraliser une propriété fondamentale des différentielles numériques : la condition nécessaire et suffisante pour que la fonction abstraite $M = F(m)$, supposée différentiable sur tous les points de la sphère abstraite Σ , $\|\overline{m_0 m}\| \leq R$, soit une constante abstraite, est que la différentielle de $F(m)$ soit identiquement nulle sur cette sphère abstraite. Nous entendons par là que le vecteur $\partial_\xi F(m)$ est le vecteur O attaché à l'espace A lorsque ξ est un vecteur quelconque attaché à l'espace a et que m est un point abstrait quelconque de la sphère Σ .

La condition est évidemment nécessaire. Montrons qu'elle est suffisante. En effet, soient m_1 un point de Σ et m un point du segment abstrait $m_0 m_1$. On aura

$$\overline{m_0 m} = \alpha \cdot \overline{m_0 m_1} \quad \text{avec} \quad 0 \leq \alpha \leq 1$$

et, par suite, pour une variation $\Delta\alpha$,

$$\Delta m = \Delta\alpha \cdot \overline{m_0 m_1}.$$

Alors ici $\delta m = \delta\alpha \cdot m_0 m_1$ et la formule (10) devient

$$\delta M = \partial\alpha \partial_{\overline{m_0 m_1}} F(m).$$

Or m appartient à la sphère Σ et l'on a $\partial_\xi F(m) = 0$ pour tout vecteur ξ . Donc $M = F(m)$ est une fonction abstraite de α dont la dérivée par rapport à α est nulle identiquement quand α varie de 0 à 1. D'après le lemme établi, page 313, il en résulte que M est fixe quand m se déplace sur le segment abstrait $m_0 m_1$. Comme m_1 est un point quelconque de Σ , il est finalement établi que M est fixe quand m est arbitraire dans la sphère Σ .

On déduit de la proposition précédente que si $M = F(m)$ est défini sur un ensemble ouvert d'un seul tenant, e , et si F est différentiable sur cet ensemble, la condition nécessaire et suffisante pour que cette fonction soit constante (c'est-à-dire que M soit fixe quand m décrit l'ensemble e) est que la différentielle de F , $\partial_\xi F(m)$, soit identiquement nulle. C'est-à-dire que $\partial_\xi F(m)$ soit le vecteur O attaché à A , quand m décrit e et quand ξ est un vecteur quelconque attaché à a .

Dans l'énoncé précédent, quand nous disons que l'ensemble ouvert e est d'un seul tenant, nous entendons par là que, pour tout couple de points m_2, m_3 de e , il existe au moins une suite d'un nombre fini de sphères appartenant à e dont la première contient m_1 , dont la dernière contient e_2 et ayant chacune au moins un point en commun avec la précédente.

Fonctions de plusieurs variables abstraites. — Supposons qu'un point M d'un espace (\mathfrak{O}) vectoriel dépende de plusieurs points m, n, \dots en nombre fini et appartenant chacun à un espace (\mathfrak{O}) vectoriel; on pourra écrire $M = F(m, n, \dots)$. Il s'agit de définir la différentielle totale dF de cette fonction de plusieurs variables. Pour cela, il nous suffira de montrer que l'ensemble des points m, n, \dots peut être considéré comme définissant un point μ d'un espace (\mathfrak{O}) vectoriel en relation simple avec ceux où se déplacent m, n, \dots . Il suffira alors d'écrire $F(m, n, \dots) = \Phi(\mu)$ et de prendre pour différentielle de F celle de Φ . Nous préciserons plus loin cette indication.

Définissons d'abord l'espace où se meut le point μ , comme un espace (\mathfrak{O}) vectoriel. Deux couples ordonnés de points μ_1, μ_2 et μ'_1, μ'_2 détermineront deux segments dirigés équipollents si les segments dirigés correspondants $m_1 m_2$ et $m'_1 m'_2$, $n_1 n_2$ et $n'_1 n'_2$, ... sont à la fois respectivement équipollents. L'ensemble des segments dirigés $\overline{\mu_1 \mu_2}$ équipollents à un segment dirigé $\overline{\mu_1 \mu_2}$ détermine un vecteur $\overline{\mu_1 \mu_2}$. La somme de deux vecteurs v_1, v_2 correspondant aux vecteurs ξ_1, ξ_2 ; μ_1, μ_2 , ... appartenant aux espaces respectifs décrits par m, n, \dots est le vecteur correspondant aux vecteurs $\xi_1 + \xi_2, \eta_1 + \eta_2, \dots$. Et si x est un nombre quelconque, $x.v_1$ est le vecteur qui correspond à $x.\xi_1, x.\eta_1, \dots$. Enfin, on peut donner plusieurs définitions de la distance de deux points μ_1, μ_2 telles que m_1, m_2, \dots tendent simultanément vers m_1, n_1, \dots quand cette distance tend vers zéro et réciproquement. Par exemple, on peut prendre

$$\|\overline{\mu_1 \mu_2}\| = \|\overline{m_1 m_2}\| + \|\overline{n_1 n_2}\| + \dots$$

On voit alors facilement que

$$\begin{aligned} \|x.\overline{\mu_1 \mu_2}\| &= |x| \|\overline{\mu_1 \mu_2}\|, \\ \|\overline{\mu_1 \mu_2}\| &\leq \|\overline{\mu_1 \mu_3}\| + \|\overline{\mu_3 \mu_2}\|, \end{aligned}$$

et que la condition nécessaire et suffisante pour que $\|\overline{\mu_1, \mu_2}\| = 0$ est que μ_1 et μ_2 coïncident.

Par analogie, une transformation vectorielle

$$V = \Psi(\xi, \eta, \dots)$$

sera linéaire par rapport à l'ensemble des ξ, η, \dots si elle est distributive et continue par rapport à cet ensemble. C'est-à-dire qu'on aura

$$(11) \quad \Psi(\xi_1 + \xi_2, \eta_1 + \eta_2, \dots) = \Psi(\xi_1, \eta_1, \dots) + \Psi(\xi_2, \eta_2, \dots)$$

et que

$$\|\Psi(\xi_1, \eta_1, \dots) - \Psi(\xi_2, \eta_2, \dots)\|$$

tend vers zéro quand

$$\|\xi_2 - \xi_1\|, \|\eta_2 - \eta_1\|, \dots$$

tendent simultanément vers zéro.

Mais on tire de l'identité (11)

$$\Psi(\xi, \eta, \zeta, \dots) - \Psi(\xi, 0, 0, \dots) = \Psi(0, \eta, \zeta, \dots),$$

$$\Psi(0, \eta, \zeta, \dots) - \Psi(0, \eta, 0, \dots) = \Psi(0, 0, \zeta, \dots),$$

$$\dots\dots\dots;$$

d'où

$$\Psi(\xi, \eta, \zeta, \dots) = \Psi(\xi, 0, \dots) + \Psi(0, \eta, 0, \dots) + \dots$$

Ainsi, si $\Psi(\xi, \eta, \zeta, \dots)$ est linéaire par rapport à l'ensemble des ξ, η, \dots , c'est d'abord une somme de transformations $\Theta(\xi) + \chi(\eta), \dots$ ne dépendant chacune que d'une variable et l'on voit facilement que chacune d'elles doit être linéaire. La réciproque est vraie.

Différentielle totale. — Définissons maintenant la différentielle d'une fonction de plusieurs variables abstraites

$$M = F(m, n, \dots).$$

Nous dirons que cette fonction est différentiable au point (m_0, n_0, \dots) , si, considérée comme une fonction $\Phi(\mu)$ de l'ensemble ordonné μ des variables m, n, \dots , elle est différentiable au point $\mu_0 (m_0, n_0, \dots)$. Alors, il existe une transformation vectorielle linéaire de l'ensemble des accroissements $\Delta m, \Delta n, \dots$, qui ne diffère de l'accroissement correspondant $\Delta\Phi$ que par une quantité infiniment petite par rapport à $\|\Delta m\| + \|\Delta n\|, \dots$. Cette transformation linéaire est la somme de

transformations linéaires séparées $\Theta(\Delta m), \chi(\Delta n), \dots$. De sorte qu'on doit avoir finalement

$$(12) \quad \Delta M = \Delta F(m_0, n_0, \dots) = \Theta(\Delta m) + \chi(\Delta n) + \dots + \varepsilon(\|\Delta m\| + \|\Delta n\| + \dots) \cdot U$$

où U est un vecteur unitaire et où ε est un nombre qui tend vers zéro avec $\|\Delta m\| + \|\Delta n\| + \dots$. Et la différentielle totale de $F(m, n, \dots)$ sera la somme des transformations linéaires $\Theta(\Delta m), \chi(\Delta n), \dots$. Remarquons qu'en faisant varier seulement, n , par exemple, l'égalité (2) devient

$$\Delta M = \chi(\Delta n) + \varepsilon \|\Delta n\| \cdot U.$$

Par suite $\chi(\Delta n)$ est « la différentielle partielle » de $F(m, n, \dots)$ considérée comme fonction de n . On pourra la représenter par la notation $\partial_{\Delta n}^{(n)} F(m, n, \dots)$. De sorte que finalement lorsque m, n, \dots sont variables

$$(13) \quad dF(m, n, \dots) = \partial_{\Delta m}^{(m)} F(m, n, \dots) + \partial_{\Delta n}^{(n)} F(m, n, \dots) + \dots$$

Ainsi, une fonction $F(m, n, \dots)$ différentiable par rapport à l'ensemble des variables m, n, \dots est différentiable par rapport à chacune d'elles, et sa différentielle totale est la somme des différentielles partielles par rapport à chacune de ces variables. Il y a lieu d'ajouter que la réciproque n'est pas vraie, même dans le cas classique où M, m, n, \dots sont des nombres.

Lorsque l'on prend pour F successivement chacune des variables indépendantes m, n, \dots , on voit que

$$dm = \Delta m, \quad dn = \Delta n, \dots,$$

et par suite dans le cas général

$$(14) \quad dF(m, n, \dots) = \partial_{dm}^{(m)} F(m, n, \dots) + \partial_{dn}^{(n)} F(m, n, \dots) + \dots$$

Changement de variables. — Cette formule est démontrée en supposant indépendants les points abstraits m, n, \dots . Nous allons maintenant démontrer qu'elle subsiste lorsque m, n, \dots sont à leur tour des fonctions différentiables de nouvelles variables α, β, \dots , pourvu qu'on remplace dm, dn, \dots dans la formule (14) par leurs expressions en fonction de $d\alpha, d\beta, \dots$. On a par hypothèse la formule (12), et de

même

$$\begin{aligned}\Delta m &= dm + \varepsilon_1(\|\Delta\alpha\| + \|\Delta\beta\| + \dots) \cdot U_1, \\ \Delta n &= dn + \varepsilon_2(\|\Delta\alpha\| + \|\Delta\beta\| + \dots) \cdot U_2, \\ &\dots\dots\dots\end{aligned}$$

d'où en substituant dans (12)

$$(15) \quad \Delta M = \Theta(dm) + \chi(dn) + \dots + \omega$$

avec

$$\omega = \Theta(\varepsilon_1 r \cdot U_1) + \chi(\varepsilon_2 r \cdot U_2) + \dots + \varepsilon(\|\Delta m\| + \|\Delta n\| + \dots) \cdot U,$$

en posant

$$r = \|\Delta\alpha\| + \|\Delta\beta\| + \dots$$

Or on a

$$\|\Delta m\| \leq \|dm\| + |\varepsilon_1| r.$$

Mais dm est une somme de transformations linéaires

$$dm = \varphi(\Delta\alpha) + \rho(\Delta\beta) + \dots,$$

et l'on a vu que chacun des rapports

$$\frac{\|\varphi(\Delta\alpha)\|}{\|\Delta\alpha\|}, \quad \frac{\|\rho(\Delta\beta)\|}{\|\Delta\beta\|}, \quad \dots$$

a une borne supérieure finie; en appelant k la plus grande d'entre elles, on aura

$$\|dm\| \leq \|\varphi(\Delta\alpha)\| + \|\rho(\Delta\beta)\| + \dots \leq kr.$$

D'où

$$\|\Delta m\| \leq r(k + |\varepsilon_1|),$$

et par suite, il existe un nombre fini $K > 0$, tel que

$$(16) \quad \|\Delta m\| + \|\Delta n\| + \dots \leq r(K + |\varepsilon_1| + |\varepsilon_2| + \dots).$$

De même

$$\|\Theta(\varepsilon_1 r \cdot U_1) + \chi(\varepsilon_2 r \cdot U_2) + \dots\| \leq r \{ |\varepsilon_1| \|\Theta(U_1)\| + |\varepsilon_2| \|\chi(U_2)\| + \dots \}.$$

En appelant H un nombre convenablement choisi, nous savons qu'on aura

$$\|\Theta(U_1)\| \leq H, \quad \|\chi(U_2)\| \leq H, \quad \dots,$$

d'où

$$\|\Theta(\varepsilon_1 r \cdot U_1) + \chi(\varepsilon_2 r \cdot U_2) + \dots\| \leq r H (|\varepsilon_1| + |\varepsilon_2| + \dots),$$

et finalement

$$\|\omega\| \leq r \{ (|\varepsilon| + H)(|\varepsilon_1| + |\varepsilon_2| + \dots) + K|\varepsilon| \}.$$

Or si r tend vers zéro, $\varepsilon_1, \varepsilon_2, \dots$ tendent vers zéro et comme d'après (16), $\|\Delta m\| + \|\Delta n\| + \dots$ tend aussi vers zéro, il en résulte que ε tend aussi vers zéro, donc le facteur en accolade tend vers zéro. Par suite

$$\omega = r\eta V,$$

où η tend vers zéro avec r et où V est un vecteur unitaire. L'égalité (15) prouve donc que

$$\Theta(dm) + \chi(dn) + \dots,$$

quand on y remplace dm, dn, \dots en fonction de $d\alpha, d\beta, \dots$, représente la différentielle de $M = F(m, n, \dots)$ considérée comme fonction de α, β, \dots . C'est ce que nous avons annoncé.

Différentielles d'ordre supérieur. — Une étude complète des propriétés des différentielles d'ordre supérieur nous entraînerait trop loin. Nous nous contenterons ici d'indiquer la définition de la différentielle seconde.

Définition de la différentielle seconde. — Soit encore $M = F(m, n, \dots)$ une fonction abstraite de plusieurs variables abstraites, les points M, m, n, \dots étant chacun variable dans un espace (\mathfrak{O}) vectoriel particulier ou commun.

On dira tout naturellement que $F(m, n, \dots)$ a une différentielle seconde si : 1° F a une différentielle première; 2° si, quels que soient les accroissements $\Delta m, \Delta n, \dots$ qui servent à former cette différentielle première, celle-ci a elle-même une différentielle par rapport à de nouveaux accroissements $\Delta' m, \Delta' n, \dots$ indépendants des précédents. Cette nouvelle différentielle sera la différentielle seconde et l'on voit qu'elle dépendra de m, n, \dots et de deux systèmes d'accroissements indépendants, $\Delta m, \Delta n, \dots; \Delta' m, \Delta' n, \dots$.

Écrivons la différentielle première de F sous la forme

$$dF(m, n, \dots) = \partial_{\Delta m}^{(m)} F(m, n, \dots) + \partial_{\Delta n}^{(n)} F(m, n, \dots) + \dots$$

Par hypothèse, dF a, quand $\Delta m, \Delta n, \dots$ sont fixes, une différentielle première qu'on peut appeler $d'dF$ par rapport à des accroissements $\Delta'm, \Delta'n, \dots$ de m, n, \dots . En particulier, il en sera ainsi quand on prend tous les premiers accroissements nuls, sauf l'un d'eux. Donc si F a une différentielle seconde, chacune de ses différentielles partielles du premier ordre

$$\partial_{\Delta m}^{(m)} F(m, n, \dots); \quad \partial_{\Delta n}^{(n)} F(m, n, \dots); \quad \dots$$

est différentiable. La réciproque est vraie, car on aurait

$$\Delta' \partial_{\Delta m}^{(m)} F(m, n, \dots) = d' \partial_{\Delta n}^{(n)} F(m, n, \dots) + \varepsilon_1 R.V_1.$$

(et de même pour les autres différentielles partielles) en posant

$$R = \|\Delta' m\| + \|\Delta' n\| + \dots,$$

ε_1 tendant vers zéro avec R , et V_1 étant un vecteur unitaire. D'où, en ajoutant

$$\Delta' dF(m, n, \dots) = d' \partial_{\Delta m}^{(m)} F(m, n, \dots) + d' \partial_{\Delta n}^{(n)} F(m, n, \dots) + \dots + \varepsilon R.V,$$

V étant un vecteur unitaire et ε tendant vers zéro avec R puisque

$$\varepsilon R.V = \varepsilon_1 R.V_1 + \varepsilon_2 R.V_2 + \dots,$$

avec

$$|\varepsilon| \leq |\varepsilon_1| + |\varepsilon_2| + \dots$$

Donc $d'dF$ existe et l'on a

$$d'dF = d' \partial_{\Delta m}^{(m)} F(m, n, \dots) + d' \partial_{\Delta n}^{(n)} F(m, n, \dots) + \dots$$

Or si l'on pose pour un instant (puisque pour calculer les d' , on laisse les Δ fixes)

$$\varphi(m, n, \dots) = \partial_{\Delta m}^{(m)} F(m, n, \dots),$$

on aura

$$d' \varphi(m, n, \dots) = \partial_{\Delta m}^{(m)} \varphi(m, n, \dots) + \partial_{\Delta n}^{(n)} \varphi(m, n, \dots) + \dots$$

D'où

$$\begin{aligned} d'dF = & \partial_{\Delta'm}^{(m)} \partial_{\Delta m}^{(m)} F(m, n, \dots) + \partial_{\Delta'n}^{(n)} \partial_{\Delta m}^{(m)} F(m, n, \dots) + \dots \\ & + \partial_{\Delta'm}^{(m)} \partial_{\Delta n}^{(n)} F(m, n, \dots) + \partial_{\Delta'n}^{(n)} \partial_{\Delta n}^{(n)} F(m, n, \dots) + \dots, \end{aligned}$$

qu'on pourra aussi représenter par la notation abrégée

$$(15) \quad \boxed{d'dF = \partial_{\Delta'm, \Delta m}^{(m, m)} F(m, n, \dots) + \partial_{\Delta'n, \Delta m}^{(n, m)} F(m, n, \dots) + \dots} \\ + \partial_{\Delta'm, \Delta n}^{(m, n)} F(m, n, \dots) + \partial_{\Delta'n, \Delta n}^{(n, n)} F(m, n, \dots) + \dots$$

Finalement, on voit que $d'dF$ est une somme de termes dont chacun est linéaire par rapport à un des accroissements $\Delta m, \Delta n, \dots$ et aussi linéaire par rapport à un des accroissements $\Delta' m, \Delta' n, \dots$. Chacun de ces termes est une différentielle partielle du second ordre de F ⁽¹⁾.

Si l'on supposait que tous les premiers accroissements $\Delta m, \Delta n, \dots$ fussent nuls, sauf un seul, et que tous les seconds accroissements $\Delta' m, \Delta' n, \dots$ sont nuls sauf un seul, la différentielle totale du second ordre de F se réduirait successivement à chacun de ces termes.

(¹) On écrit d'habitude la différentielle seconde d'une fonction de deux variables numériques en égalant les deux systèmes d'accroissements. Mais cela ne devrait pas en faire oublier la véritable origine.