ANNALES DE L'UNIVERSITÉ DE GRENOBLE

LAURENT SCHWARTZ

Généralisation de la notion de fonction, de dérivation, de transformation de Fourier et applications mathématiques et physiques

Annales de l'université de Grenoble, tome 21 (1945), p. 57-74

<http://www.numdam.org/item?id=AUG_1945__21__57_0>

GÉNÉRALISATION DE LA NOTION DE FONCTION, DE DÉRIVATION, DE TRANSFORMATION DE FOURIER ET APPLICATIONS MATHÉMATIQUES ET PHYSIQUES

par M. Laurent SCHWARTZ.

Introduction.

Depuis l'introduction du calcul symbolique, les physiciens se sont couramment servis de certaines notions ou de certaines formules dont le succès était incontestable, alors qu'elles n'était pas justifiées mathématiquement. C'est ainsi que la fonction $y(x)$ de la variable réelle x, égale à 0 pour $x \leqslant 0$, à 1 pour $x > 0$, est couramment considérée comme ayant pour dérivée la « fonction de Dirac » $y'(x) = \delta(x)$, nulle pour $x \neq 0$, égale à $+\infty$ pour $x = 0$, et telle que, de plus $\int_{-\infty}^{+\infty} \delta(x)dx = +1$. Un tel « abus de langage » est malgré tout incompatible avec la notion habituelle de fonction et de dérivation ! Et que penser alors de la considération des dérivées successives de la fonction de Dirac ! Et pourtant de telles expressions rendent de constants services en électricité et sont très adaptées à l'étude de la transformation de Laplace ou de Fourier et de la mécanique ondulatoire. Le but de cet article est de faire un très bref résumé (et sans démonstrations) d'un travail qui sera publié ultérieurement sous forme de mémoire ou de monographie et qui apportera une justification complète au langage précédent [1]. Il se

[1] J'ai exposé ces idées dans des leçons au Collège de France (Cours Peccot, janvier-avril 1946).

trouve que ce langage, ainsi réhabilité dans le domaine de l'analyse, est fécond dans beaucoup de questions diverses : électricité et physique mathématique, transformations de Fourier et Laplace, équations et inéquations aux dérivées partielles (dont la théorie des fonctions harmoniques et surharmoniques), étude théorique des variétés et formes différentielles. J'ai obtenu jusqu'à présent en particulier beaucoup de simplifications ou d'explications remarquables de formules connues, et sans vouloir faire des affirmations prématurées, je pense qu'il y a intérêt à utiliser ce formalisme dans un grand nombre de domaines de l'analyse ([1]).

§ 1. — Généralisation de la notion de fonction.

Les éléments sur lesquels il faut raisonner sont plus généraux que des fonctions. Ainsi $\delta(x)$ n'est pas une fonction, c'est une *mesure* ou *distribution de masses,* d'un type particulièrement simple : elle comporte une masse $+1$ placée à l'origine. Une distribution de masses (μ) est entièrement définie par la connaissance de la masse $\mu(a, b)$ contenue dans tout intervalle (a, b) ; c'est un nombre réel de signe quelconque ou même un nombre complexe. $\mu(a, b)$ ne peut pas être une fonction quelconque d'intervalle, elle doit vérifier une condition qui exprime que la somme des modules des masses est finie et une condition d'additivité. (μ) permet de définir une fonction d'ensemble A par

$$\mu(\mathrm{A}) = \int_{\mathrm{A}} d\mu$$

et plus généralement une fonctionnelle $\mu(\varphi)$ définie au moins pour toute fonction continue $\varphi(x)$, nulle en dehors d'un intervalle fini :

$$(1) \qquad \mu(\varphi) = \int_{-\infty}^{+\infty} \varphi(x)\, d\mu \quad \text{(intégrale de Stieltjes).}$$

Par exemple, si δ est la *distribution de Dirac* formée d'une masse $+1$ à l'origine,

$$\delta(\varphi) = \int_{-\infty}^{+\infty} \varphi(x)\, d\delta = \varphi(0).$$

La formule (1) se trouve être la mieux adaptée à la définition

([1]) L'exposé que je vais faire ici vise avant tout à la simplicité ; l'exposé ultérieur, plus détaillé, emploiera un langage plus purement mathématique d'analyse fonctionnelle, d'espaces vectoriels, etc.

même d'une distribution de masses. On voit que $\mu(\varphi)$ *est une fonctionnelle linéaire de* φ *et que, si* φ *reste nul en dehors d'un intervalle fini fixe et converge uniformément vers* o, $\mu(\varphi)$ *converge vers* o.

Réciproquement [1] toute fonctionnelle vérifiant ces propriétés est de la forme (1) et désormais nous ne parlerons plus d'une distribution de masses autrement que *comme définie par la forme linéaire continue* $\mu(\varphi)$.

En quoi une fonction $f(x)$ est-elle un cas particulier d'une mesure ? En ce qu'elle définit une mesure de « densité » $f(x)$ pour laquelle la masse contenue dans un intervalle (a, b) est $\int_a^b f(x)\,dx$; plus généralement, la fonctionnelle définie dans la formule (1) s'écrit dans ce cas :

$$\int_{-\infty}^{+\infty} f(x)\,\varphi(x)\,dx \tag{2}$$

que nous appellerons $f(\varphi)$. Dans la suite nous identifierons une fonction $f(x)$ à la mesure qu'elle définit, et nous dirons fréquemment : la mesure μ se réduit à la fonction f, ce qui voudra dire que

$$\int_{-\infty}^{+\infty} \varphi(x)\,d\mu = \int_{-\infty}^{+\infty} \varphi(x) f(x)\,dx,$$

$f(x)$ est alors n'importe quelle fonction sommable sur tout intervalle fini et elle est définie à un ensemble de mesure nulle près. Comme nous l'avons dit plus haut, la mesure δ n'est pas une fonction.

Il est ensuite nécessaire de définir des distributions plus générales que des distributions de masses et qui correspondent aux « couches multiples » (couches de doublets ou dipôles et couches plus compliquées) employées dans la théorie du potentiel. Qu'est-ce qu'un « doublet » de « moment » $+1$ placé à l'origine ? On le représente comme formé de deux masses $+\frac{1}{\varepsilon}$ et $-\frac{1}{\varepsilon}$, placées respectivement aux points ε et o, ε étant un infiniment petit. La fonctionnelle linéaire définie par une telle distribution vérifie

$$T_\varepsilon(\varphi) = \frac{\varphi(\varepsilon) - \varphi(0)}{\varepsilon}.$$

Elle n'a de limite pour $\varepsilon \to 0$ que si φ est dérivable et l'on a alors

$$T(\varphi) = \varphi'(0). \tag{3}$$

[1] Théorème classique de F. Riesz. Voir par exemple *Banach*, « Théorie des opérations linéaires », Monografie Matematyczne, Varsovie, 1932, pp. 59-65.

Un doublet définit donc une fonctionnelle linéaire pour des fonctions φ dérivables, de plus $T(\varphi)$ ne converge vers o que si la dérivée φ' converge vers o. Des couches multiples d'ordres plus élevés font intervenir les dérivées suivantes de φ.

Dans le cadre d'une absolue généralité, nous sommes ainsi amenés à ne considérer que des fonctions φ, nulles en dehors d'intervalles finis et indéfiniment dérivables. Par ailleurs on peut considérer des distributions dans l'espace à n dimensions, de coordonnées $x_1, x_2, \ldots\ldots x_n$. Nous donnerons alors les définitions précises suivantes :

Définition 1. — Φ sera l'ensemble des fonctions $\varphi(x_1, x_2, \ldots\ldots x_n)$ de n variables réelles, indéfiniment dérivables et nulles en dehors d'ensembles bornés. A chaque fonction φ correspond un « noyau », ensemble compact, dont le complémentaire est le plus grand ensemble ouvert sur lequel $\varphi \equiv 0$.

Définition 2. — On appellera « distribution » de l'espace à n dimensions toute fonctionnelle ou forme linéaire $T(\varphi)$ définie pour toutes les φ de Φ, et vérifiant de plus la condition de continuité suivante :

Si une suite de fonctions φ_i, ont leurs noyaux contenus dans un compact fixe et si elles convergent uniformément vers o, ainsi que chacune de leurs dérivées, alors les $T(\varphi_i)$ convergent vers o.

On appelle noyau de T le plus petit ensemble fermé « contenant toute la distribution » ; c'est-à-dire que son complémentaire est le plus grand ensemble ouvert tel que, pour toute fonction φ ayant son noyau dans cet ouvert, $T(\varphi)$ soit nul. (Ainsi le noyau de la mesure de Dirac est réduit à l'origine.) Le noyau d'une distribution, toujours fermé, n'est en général pas borné, donc pas compact ; il peut être tout l'espace.

Une distribution T peut être, comme cas particulier, une mesure μ ; cela signifie que $T(\varphi_i)$ converge vers o si les φ_i, gardant leurs noyaux contenus dans un compact fixe, convergent uniformément vers o, aucune hypothèse n'ayant besoin d'être faite sur la convergence des dérivées des φ_i. Comme cas plus particulier encore, cette mesure μ peut être une fonction f, sommable dans toute région bornée (et définie à un ensemble de mesure nulle près).

Définition 3. — Les distributions T forment elles-mêmes un espace vectoriel puisqu'elles peuvent être additionnées et multipliées par des nombres réels ou complexes ; nous l'appellerons $\mathcal{T}$, espace des distributions.

§ 2. — Dérivation des distributions.

Soit f une fonction continue et dérivable à dérivées continues. Définie comme une distribution, elle satisfait à

$$f(\varphi)=\iint\cdots\int f(x_1, x_2, \ldots x_n)\,\varphi(x_1, x_2, \ldots x_n)\,dx_1, dx_2 \ldots dx_n.$$

Sa dérivée f'_{x_i}, en tant que distribution, satisfait à

$$f'_{x_i}(\varphi)=\iint\cdots\int f'_{x_i}(x_1, \ldots x_n)\,\varphi(x_1, \ldots x_n)\,dx_1 \ldots dx_n$$

Cette intégrale est étendue à tout l'espace, mais φ est nul en dehors d'une région bornée ; il n'y a aucune difficulté à intégrer par parties :

$$f'_{x_i}(\varphi)=-\iint\cdots\int f(x_1 \ldots x_n)\,\varphi'_{x_i}(x_1, \ldots x_n)\,dx_1 \ldots dx_n$$

ou

$$f'_{x_i}(\varphi)=f(-\varphi'_{x_i}). \tag{4}$$

Telle est l'égalité fondamentale vérifiée par une fonction continue et dérivable et sa dérivée f'_{x_i}, au sens usuel du mot. Cette égalité permet de généraliser la notion de dérivée et *de définir la dérivée* T'_{x_i} *d'une distribution quelconque* T comme une nouvelle distribution, entièrement connue comme fonctionnelle :

$$T'_{x_i}(\varphi)=T(-\varphi'_{x_i}). \tag{5}$$

Ainsi toute distribution est indéfiniment dérivable et on peut intervertir l'ordre des dérivations ; en employant la notation de Leibniz :

$$\frac{\partial^{p_1+p_2+\cdots+p_n}T}{\partial x_1^{p_1}\partial x_2^{p_2}\ldots\partial x_n^{p_n}}(\varphi)=T\left((-1)^{p_1+\cdots+p_n}\frac{\partial^{p_1+\cdots+p_n}\varphi}{\partial x_1^{p_1}\ldots\partial x_n^{p_n}}\right). \tag{6}$$

Il en résulte, en particulier, que toute fonction continue (ou même sommable sur toute région bornée) est indéfiniment dérivable ; en général sa dérivée n'est ni une fonction ni même une mesure, c'est une distribution. Cependant, si f est dérivable au sens usuel du mot et si $\frac{\partial f}{\partial x_i}$ est continue, sa distribution dérivée s'identifie bien avec sa fonction dérivée usuelle $\frac{\partial f}{\partial x_i}$, d'après (4).

Donnons quelques exemples, dans le cas d'une seule variable.

1° Calculons la dérivée de $y(x)$, fonction égale à o pour $x \leqslant 0$, à 1 pour $x > 0$.

$$(7) \qquad y'(\varphi) = -y(\varphi') = -\int_0^{+\infty} \varphi'(x)\,dx = \varphi(0) = \delta(\varphi).$$

y' est donc bien égale à la distribution de Dirac.

Nous voyons toute la richesse de cette dérivation. Non seulement elle est toujours possible, mais toutes les singularités de la distribution se retrouvent dans sa dérivée. Au sens usuel du mot, y' est nul pour $x > 0$ et pour $x < 0$, nulle à droite et à gauche pour $x = 0$, de sorte que rien dans la dérivée usuelle ne permet de retrouver la discontinuité $+1$ de $y(x)$ à l'origine ; ici cette discontinuité se retrouve dans la masse $+1$ placée à l'origine.

La dérivée suivante y'' est la dérivée δ' de la mesure de Dirac :

$$\delta'(\varphi) = -\delta(\varphi') = -\varphi'(0).$$

C'est donc en vertu de (3) un doublet de moment -1 placé à l'origine. Plus généralement, la dérivée $\delta^{(n)}$ est définie par :

$$(8) \qquad \delta^{(n)}(\varphi) = (-1)^n \varphi^{(n)}(0).$$

2° Soit $f(x)$ une fonction indéfiniment dérivable par morceaux. Cela signifie qu'on peut partager l'axe réel en intervalles partiels dans chacun desquels f est indéfiniment dérivable, f et toutes ses dérivées ayant donc des discontinuités de première espèce aux extrémités de ces intervalles. La distribution dérivée première est la somme de la fonction dérivée usuelle $f'(x)$ et d'un système de masses ponctuelles aux extrémités des intervalles, chacune étant égale à la discontinuité de f. La distribution dérivée seconde est la somme de la fonction dérivée seconde usuelle $f''(x)$, d'une distribution de doublets dont les moments sont les opposés des discontinuités de f, et d'une distribution de masses ponctuelles égales aux discontinuités de f', etc.

3° Voici un exemple de nature différente. La dérivée de la fonction $f(x)$, nulle pour $x \leqslant 0$, égale à $1/\sqrt{x}$ pour $x > 0$, n'est certainement pas la fonction dérivée usuelle, nulle pour $x \leqslant 0$, égale à $-\frac{1}{2} x^{-3/2}$ pour $x > 0$, car cette fonction n'étant pas sommable au voisinage de $x = 0$, ne définit pas une distribution.

Le calcul direct donne :

$$f'(\varphi) = -f(\varphi') = -\int_0^{+\infty} \frac{\varphi'(x)}{\sqrt{x}}dx = -\lim_{\varepsilon \to 0} \int_\varepsilon^{+\infty} \frac{\varphi'(x)}{\sqrt{x}}dx$$

$$= \lim_{\varepsilon \to 0} \left[-\int_\varepsilon^{+\infty} \frac{1}{2} x^{-3/2} \varphi(x) dx + \frac{\varphi(0)}{\sqrt{\varepsilon}} \right].$$

On obtient donc ce que M. Hadamard [1] a appelé « partie finie » d'une intégrale divergente et qu'il a noté par le symbole :

$$f'(\varphi) = \left| \overline{\int_0^{+\infty}} \left(-\frac{1}{2} x^{-3/2} \right) \varphi(x) dx. \right. \tag{9}$$

De sorte que nous pouvons considérer que f' est une pseudo-fonction que nous représenterons par 0 pour $x \leqslant 0$, et par

$$\left| \overline{-\frac{1}{2} x^{-3/2}} \right. \quad \text{pour} \quad x > 0.$$

On démontrerait de même, en utilisant toujours le symbole $\overline{|}$ de M. Hadamard, que la dérivée d'ordre k de la fonction nulle pour $x \leqslant 0$, égale à x^α pour $x > 0$, $(\alpha > -1)$ est nulle pour $x \leqslant 0$, et représentée par la pseudofonction :

$$\left| \overline{\alpha(\alpha - 1) \dots (\alpha - k + 1) x^{\alpha - k}} \right. \quad \text{pour} \quad x > 0.$$

La dérivée de la fonction $\log|x|$ est la pseudo-fonction

$$\left| \overline{1/x} \right. = v.\ p.\ 1/x,$$

$v.\ p.$ désignant la valeur principale de Cauchy

$$v.\ p.\ \frac{1}{x}(\varphi) = v.\ p. \int_{-\infty}^{+\infty} \frac{\varphi(x)}{x} dx = \lim_{\varepsilon \to 0} \left(\int_{-\infty}^{-\varepsilon} + \int_\varepsilon^{+\infty} \right). \tag{10}$$

Les dérivées des fonctions usuelles donnent des types d'intégrales généralisées d'une richesse très grande comme peuvent le montrer ces quelques exemples.

§ 3. — Intégration des distributions.

Dans la théorie usuelle, une fonction *continue* à dérivées nulles est une constante. Mais la fonction $y(x)$, nulle pour $x \leqslant 0$, égale à

[1] *Hadamard*, « Le problème de Cauchy et les équations aux dérivées partielles du type hyperbolique ». Paris, Hermann, 1932, p. 163-206.

1 pour $x > 0$, n'est pas une constante bien que sa dérivée soit nulle. Dans l'espace de distributions, $y'(x)$ n'est pas nulle, puisque c'est la mesure δ de Dirac. On peut démontrer ce qui suit :

THÉORÈME. — *Une distribution dont les dérivées du 1^er ordre sont nulles est une fonction constante* (1).

L'intégration des distributions d'une variable x est toujours possible, la primitive d'une distribution est définie à une constante près. La théorie des dérivations et intégrations d'ordre non entier devient ici particulièrement simple et élégante. Dans le cas de plusieurs variables, on peut traiter divers problèmes ; par exemple (en nous bornant au cas de 2 variables x, y) trouver une distribution S vérifiant, T étant donnée, $\dfrac{\partial S}{\partial x} = T$.

Le problème a toujours des solutions, et la différence entre deux solutions, vérifiant $\dfrac{\partial S}{\partial x} = 0$, est « une distribution qui ne dépend que de y », c'est-à-dire invariante par toute translation parallèle à $x'x$.

On peut ensuite chercher une distribution S vérifiant, A et B étant deux distributions données

$$\frac{\partial S}{\partial x} = A, \qquad \frac{\partial S}{\partial y} = B.$$

S'il y a une solution, elle est définie à une fonction constante près. Mais il y a une condition de compatibilité.

$$\frac{\partial A}{\partial y} = \frac{\partial B}{\partial x}. \tag{11}$$

Ce résultat est remarquable. Si A et B sont des fonctions continues, (11) est la condition usuelle pour qu'elles soient les dérivées partielles d'une fonction continue S (x, y) ; mais cette condition, habituellement, n'a de sens que si A et B sont dérivables, ce qui n'a aucune raison d'être dans le problème posé ; ici, $\dfrac{\partial A}{\partial y}$ et $\dfrac{\partial B}{\partial x}$ ont toujours un sens en tant que distributions, et (11) est la condition pour que l'on puisse trouver S. Même en restant dans le cadre de A, B, S, fonctions continues, la condition (11) rend obligatoire le langage des distributions.

(1) Une distribution T est nulle si c'est la distribution identique à la fonction 0, ou si $T(\varphi) = 0$ quelle que soit φ.

La résolution des équations aux dérivées partielles telles que

$$\frac{\partial^{p_1+\cdots+p_n}S}{\partial x_1^{p_1}\cdots\partial x_n^{p_n}}=0 \qquad \text{ou} \qquad =T,$$

n'offre aucune difficulté. Ce qui est intéressant, c'est que l'on peut définir ainsi des solutions discontinues d'équations aux dérivées partielles. Ainsi la solution générale de

$$(12) \qquad \frac{\partial^2 S}{\partial x \partial y}=0,$$

est

$$S=A(x)+B(y),$$

où A est une distribution indépendante de y, B une distribution indépendante de x. A et B peuvent être en particulier des fonctions sommables discontinues ou des fonctions continues non dérivables au sens usuel, S n'en est pas moins solution de (12).

Nous sommes ainsi amenés à l'étude des équations différentielles et aux dérivées partielles. Je me contenterai d'énoncer le résultat suivant : dans le cas d'une équation différentielle à une variable, il n'y a pas d'autres solutions que les solutions usuelles, fonctions continues et dérivables.

Dans le cas d'équations aux dérivées partielles, c'est tout différent ; une équation telle que (12), hyperbolique, a des solutions nouvelles, localement très irrégulières ; de même l'équation :

$$(13) \qquad \frac{\partial^2 S}{\partial x^2}-\frac{\partial^2 S}{\partial y^2}=0,$$

a pour solution générale :

$$S=A(x+y)+B(x-y),$$

où A et B sont des distributions, donc en particulier peuvent être des fonctions sommables discontinues ou continues non dérivables au sens usuel. Une telle circonstance se produira toujours pour des équations du 1^er^ ordre à coefficients réels.

Au contraire, une équation elliptique telle que celle de Laplace

$$(14) \qquad \Delta S=\frac{\partial^2 S}{\partial x^2}+\frac{\partial^2 S}{\partial y^2}=0,$$

ou celle des fonctions polyharmoniques

$$(15) \qquad \Delta^n S=0.$$

n'a pas d'autres solutions que les solutions usuelles, fonctions analytiques (à un changement près sur un ensemble de mesure nulle).

§ 4. — Structures algébriques dans l'espace des distributions.

On peut donc dans l'espace $\mathcal{T}$ des distributions, définir diverses propriétés et structures intéressantes.

1° On peut définir une structure d'ordre en définissant une distribution $\geqslant 0$:

$$T \geqslant 0 \quad \text{si, quelle que soit} \quad \varphi \geqslant 0, \quad T(\varphi) \geqslant 0.$$

Théorème. — *Toute distribution* $\geqslant 0$ *est une mesure (ou distribution de masses)* $\geqslant 0$.

Une application intéressante concerne les fonctions surharmoniques ; on appelle ainsi une fonction sommable semi-continue inférieurement dont la valeur en chaque point majore la moyenne sur toute sphère centrée en ce point.

Si c'est une fonction continue et 2 fois dérivable, le fait pour T d'être surharmonique est équivalent à l'inéquation :

$$\Delta T \leqslant 0. \tag{16}$$

La définition donnée est plus générale, puisque valable pour des fonctions non dérivables. Mais dans l'espace des distributions, tout est dérivable ; on montre que les distributions solutions de (16) sont justement les fonctions « presque surharmoniques » c'est-à-dire égales presque partout aux fonctions surharmoniques, ce qui réhabilite le laplacien ! Remarquons que d'après (16) et le théorème ci-dessus, ΔT est une mesure, qui joue un rôle important dans la décomposition de Riesz, que nous verrons plus loin. Cette décomposition de Riesz sera d'ailleurs généralisable à toutes les inéquations aux dérivées partielles analogues à (16).

2° On peut multiplier une distribution par une fonction indéfiniment dérivable : αT est une nouvelle distribution définie par

$$\alpha T(\varphi) = T(\alpha\varphi). \tag{17}$$

Le produit se dérive suivant la règle habituelle

$$(\alpha T)' = \alpha' T + \alpha T'.$$

Exemple. — Dans le cas d'une variable, δ étant la distribution de Dirac : $x\delta' = -\delta$ car

$$(x\delta')(\varphi) = \delta'(x\varphi) = -(\varphi + x\varphi')_0 = -\varphi(0) = -\delta(\varphi).$$

3° On peut définir le produit tensoriel de 2 distributions.

Dans le cas de fonctions d'une variable, le produit tensoriel des deux fonctions f et g est la fonction de deux variables $f(x)g(y)$. Le produit tensoriel des deux mesures μ et ν est la mesure produit à deux variables $\mu(x)\nu(y)$ par rapport à laquelle l'intégrale d'une fonction $\varphi(x, y)$ est :

$$(\mu_x \times \nu_y)\varphi(x, y) = \iint \varphi(x, y)\, d\mu(x)\, d\nu(y).$$

De même le produit tensoriel d'une distribution $S(x_1, x_2, \ldots x_m)$ à m variables par une distribution $T(y_1, y_2, \ldots y_n)$ à n variables est une distribution à $m+n$ variables $x_1 x_2 \ldots x_m y_1 y_2 \ldots y_n$ qu'on représente par $S \times T$. Pour la connaître, il faut connaître $(S \times T)(\varphi)$ pour toute fonction $\varphi(x_1, x_2, x_m, y_1 \ldots y_n)$; on montre que $S \times T$ est entièrement déterminée par le fait que si

$$\varphi = A(x_1, x_2 \ldots x_m) B(y_1, y_2 \ldots y_n),$$
$$(S \times T)(\varphi) = S(A)T(B).$$

On montre, d'autre part, que le procédé de calcul de Fubini, qui permet de remplacer une intégration multiple par plusieurs intégrations simples successives, est valable. Ainsi :

$$T_{y_1 y_2 \ldots y_n}\varphi(x_1, x_2 \ldots x_m, y_1 \ldots y_n)$$

est une fonction de $x_1 \ldots x_m$ indéfiniment dérivable, à noyau compact, et l'on a

$$(S \times T)(\varphi) = S_{x_1, x_2 \ldots x_m}(T_{y_1, \ldots y_n}(\varphi)) = T_{y_1, y_2 \ldots y_n}(S_{x_1, \ldots x_m}(\varphi)).$$

4° Mais le produit le plus important dans toute la théorie est le produit de composition (« Faltung »). Pour deux fonctions d'une variable $f(x)$, $g(x)$, on le définit d'ordinaire par la relation :

$$h(x) = f(x) * g(x) = \int_{-\infty}^{+\infty} f(x - t) g(t)\, dt, \tag{18}$$
$$= \int_{-\infty}^{+\infty} g(x - t) f(t)\, dt.$$

Ce produit de composition est associatif et commutatif:

$$f * g = g * f,$$
$$(f * g) * h = f * (g * h) \qquad \text{qu'on écrit} \qquad f * g * h. \tag{19}$$

La composition joue un rôle essentiel dans un nombre croissant de domaines de l'analyse ; elle n'est pas spéciale à la droite ou à l'espace à n dimensions, elle est liée à la structure de groupe ([1]). Dans (18), h n'existe que si f et g décroissent assez rapidement à l'infini ; nous supposerons toujours que dans un produit de composition, tous les facteurs, sauf un au plus. sont à noyaux compacts. Dans ce cas, ce produit a un sens si les facteurs sont des distributions quelconques. Pour le définir on remarque que si f, g, h, sont des fonctions continues :

$$(20)\quad h(\varphi)=\int_{-\infty}^{+\infty} h(x)\varphi(x)\,dx=\iint f(u)g(v)\varphi(u+v)\,du\,dv,$$

de sorte que $h(\varphi)$ est la valeur de la fonctionnelle linéaire produit tensoriel $f(u)\times g(v)$ pour la fonction de deux variables $\varphi(u+v)$.

Plus généralement, si $\mathrm{S}(x_1, x_2 \ldots x_n)$, $\mathrm{T}(x_1, x_2 \ldots x_n)$ sont deux distributions sur le même espace à n dimensions (l'une au moins étant à noyau compact, pour éviter toute difficulté à l'infini) nous appelerons produit de composition $\mathrm{S} * \mathrm{T}$ une nouvelle distribution dans le même espace à n dimensions, définie comme une fonctionnelle linéaire, prenant pour la fonction $\varphi(x_1, x_2 \ldots x_n)$, la valeur que prend le produit tensoriel $\mathrm{S}_{u_1, u_2, \ldots u_n} \times \mathrm{T}_{v_1, v_2, \ldots v_n}$ pour la fonction de $2n$ variables $\varphi(u_1+v_1, u_2+v_2, \ldots u_n+v_n)$. Nous écrirons :

$$\begin{aligned}(\mathrm{S}*\mathrm{T})_{x_1, x_2 \ldots x_n}&(\varphi(x_1, x_2, \ldots x_n))\\ &=(\mathrm{S}_{u_1, u_2, \ldots u_n}\times \mathrm{T}_{v_1, v_2, \ldots v_n})\left[\varphi(u_1+v_1, \ldots u_n+v_n)\right],\end{aligned}$$

que l'on peut remplacer, d'après Fubini, par :

$$\begin{aligned}&\mathrm{S}_{u_1, u_2, \ldots u_n}\Big[\mathrm{T}_{v_1, v_2, \ldots v_n}\big\{\varphi(u_1+v_1, \ldots u_n+v_n)\big\}\Big]\\ &=\mathrm{T}_{v_1, v_2, \ldots v_n}\Big[\mathrm{S}_{u_1, u_2, \ldots u_n}\big\{\varphi(u_1+v_1, \ldots u_n+v_n)\big\}\Big].\end{aligned}$$

Il est impossible de donner ici les nombreuses applications du produit de composition ainsi étendu. Son importance vient en grande partie de ce que la dérivation est une opération de composition. Si δ est dans l'espace à n dimensions la masse $+1$ à l'origine, on démontre aisément que $\delta * \mathrm{T} = \mathrm{T}$.

([1]) Pour l'étude générale du produit de composition, voir A. Weil : « L'intégration dans les groupes topologiques et ses applications ». Act. sc. et ind. n° 869, Paris, 1940, p. 46-60.

Car

$$(\delta * \mathrm{T})_{x_1, x_2, \dots x_n}[\varphi(x_1, x_2, \dots x_n)]$$
$$= \mathrm{T}_{v_1, v_2, \dots v_n}[\delta_{u_1, \dots u_n}\varphi(u_1+v_1, \dots u_n+v_n)]$$
$$= \mathrm{T}_{v_1, v_2, v_n}[\varphi(0+v_1, \dots 0+v_n)] = \mathrm{T}(\varphi).$$

Autrement dit δ est l'opérateur unité de la composition.

On peut voir ensuite que

$$\frac{\partial \delta}{\partial x_i} * \mathrm{T} = \frac{\partial \mathrm{T}}{\partial x_i}. \tag{21}$$

En effet,

$$\left(\frac{\partial \delta}{\partial x_i} * \mathrm{T}\right)_{x_1, \dots x_n}[\varphi(x_1, \dots x_n)]$$
$$= \mathrm{T}_{v_1, \dots v_n}\left[\left(\frac{\partial \delta}{\partial u_i}\right)_{u_1, \dots u_n}\varphi(u_1+v_1, \dots u_n+v_n)\right]$$
$$= \mathrm{T}_{v_1, \dots v_n}\left[-\left\{\frac{\partial \varphi}{\partial u_i}(u_1+v_1, \dots u_n+v_n)\right\}_{u_1=\dots=u_n=0}\right]$$
$$= \mathrm{T}_{v_1, \dots v_n}\left[-\frac{\partial \varphi}{\partial v_i}(v_1, \dots v_n)\right]$$
$$= \left(\frac{\partial \mathrm{T}}{\partial x_i}\right)_{x_1, \dots x_n}\left\{\varphi(x_1, \dots x_n)\right\}. \qquad \text{C. q. f. d.}$$

Aussi est-il commode de remplacer l'écriture $\frac{\partial}{\partial x_i}\delta$ par $\frac{\partial}{\partial x_i}$ tout court ; dans le produit de composition c'est purement et simplement l'opérateur de dérivation et $\frac{\partial \mathrm{T}}{\partial x_i} = \frac{\partial}{\partial x_i} * \mathrm{T}$. On en déduit des formules telles que la suivante :

$$\frac{\partial^2}{\partial x \partial y} * (\mathrm{S} * \mathrm{T}) = \left(\frac{\partial^2}{\partial x \partial y} * \mathrm{S}\right) * \mathrm{T} = \left(\frac{\partial}{\partial x} * \mathrm{S}\right) * \left(\frac{\partial}{\partial y} * \mathrm{T}\right).$$

Choisissons seulement une application particulièrement apte à mettre en évidence les qualités du formalisme : la formule de Poisson pour les potentiels et la décomposition de Riesz pour les fonctions surharmoniques.

Dans l'espace à 3 dimensions, la fonction $1/r = 1/\sqrt{x^2+y^2+z^2}$ est surharmonique ; son Δ est donc $\leqslant 0$ (16), c'est une mesure négative. Comme $\Delta(1/r) = 0$ en dehors de l'origine, cette mesure est concentrée à l'origine ; c'est donc une masse $-k$ placée à l'origine. On démontre que $k = +4\pi$

$$\Delta * (1/r) = -4\pi\delta. \tag{22}$$

Dans cette formule, Δ, opérateur laplacien, veut dire la distribution $\Delta\delta$.

Dans les opérations de composition $-\frac{\Delta}{4\pi}$ et $1/r$ sont donc inverses l'un de l'autre.

Soit alors T une distribution quelconque, à noyau compact. Le potentiel électrique créé par cette distribution serait, si T était une fonction $\rho(x, y, z)$:

$$U_\rho(x, y, z) = \iiint \frac{\rho(\xi, \eta, \zeta)}{\sqrt{(x-\xi)^2+(y-\eta)^2+(z-\zeta)^2}}\, d\xi\, d\eta\, d\zeta,$$

ou

$$U_\rho = \frac{1}{r} * \rho.$$

Nous maintiendrons cette formule pour une distribution quelconque T :

$$U_T = \frac{1}{r} * T. \tag{23}$$

On a alors immédiatement :

$$\begin{aligned} \Delta * U_T &= \Delta * \left(\frac{1}{r} * T\right) = \left(\Delta * \frac{1}{r}\right) * T, \\ &= -4\pi\delta * T = -4\pi T. \end{aligned} \tag{24}$$

C'est la formule classique de Poisson pour les potentiels. Mais elle est vraie ici, non seulement pour le potentiel d'une couche simple continue, mais pour le potentiel d'une couche quelconque !

Considérons maintenant la distribution correspondant à une fonction S surharmonique (donc une fonction sommable).

$\Delta S \leqslant 0$ est une mesure $\leqslant 0$, que nous appellerons $-4\pi\mu$, μ étant une mesure $\geqslant 0$.

$$\Delta S = -4\pi\mu.$$

Mais si U_μ est le potentiel de μ, on a, d'après (24)

$$\Delta U_\mu = -4\pi\mu.$$

Alors $\Delta(S - U_\mu) = 0$; d'après (14), $S - U_\mu$ vaut presque partout une fonction harmonique H au sens usuel du mot et on a même partout, puisque deux fonctions surharmoniques égales presque partout sont égales partout :

$$S = U_\mu + H, \qquad \mu \geqslant 0. \tag{25}$$

C'est la décomposition de Riesz : toute fonction surharmonique est somme du potentiel d'une mesure $\geqslant 0$ et d'une fonction harmonique. La démonstration est devenue une simple écriture de quelques lignes algébriques.

§ 5. — Structure topologique dans l'espace des distributions.

Il est intéressant d'introduire une topologie dans l'espace des distributions, afin de pouvoir dire que les distributions T_i convergent vers une distribution limite T, ou ce qui revient au même, que les $T_i - T$ convergent vers o.

La topologie la plus intéressante est telle que :

Des distributions T_i convergent vers o si, quelle que soit φ, les $T_i(\varphi)$ convergent vers o, et cela uniformément par rapport à tout ensemble de fonctions φ à noyaux contenus dans un compact fixe, et bornées dans leur ensemble ainsi que chacune de leurs dérivées.

Cet énoncé n'est compliqué qu'en apparence, l'examen de la convergence est en général facile. Ainsi on voit immédiatement que la dérivée d'une distribution peut être définie comme la dérivée d'une fonction usuelle :

$$(26) \qquad T'(x) = \lim_{h \to 0} \frac{T(x+h) - T(x)}{h},$$

et que le développement limité de Taylor est valable pour toutes les distributions :

$$(27) \qquad T(x+h) = T(x) + hT'(x) + \cdots + \frac{h^n}{n!} T^{(n)}(x) + h^n \varepsilon(x),$$

ε étant une distribution qui tend vers o avec h.

On peut alors donner deux théorèmes qui sont, dans la pratique, les critères les plus commodes de convergence :

THÉORÈME 1. — *Si des fonctions continues f_i convergent vers une fonction continue f, uniformément sur tout compact, les distributions f_i convergent vers la distribution f.*

THÉORÈME 2. — *La dérivation est une opération linéaire continue. Autrement dit si des distributions T_i convergent vers T, les DT_i convergent vers DT, D étant un symbole de dérivation quelconque.*

Ce deuxième théorème permet de se débarrasser de toutes les difficultés inhérentes habituellement à la dérivation. On peut dans une suite, une série, une intégrale convergente, dériver sans prendre de précautions. Il est pourtant bien connu que des fonctions dérivables f_i peuvent converger uniformément vers o sans que leurs dérivées f'_i aient de limite; mais dans l'espace topologique des distributions, les f'_i convergent vers o. On déduit de là immédiatement que les solutions d'une équation aux dérivées partielles forment un ensemble fermé : une limite de solutions est une solution.

On peut démontrer les propriétés suivantes :

Dans toute région bornée, une distribution est une dérivée d'ordre fini d'une fonction continue. Au fond, nous avons introduit ces distributions pour pouvoir dériver toutes les fonctions continues ; mais nous avons introduit le minimum d'éléments nouveaux puisqu'il n'y a que les dérivées d'ordre fini des fonctions continues.

Dans l'espace des distributions, les fonctions continues et indéfiniment dérivables sont denses. Cela permet de généraliser simplement aux distributions certaines des opérations qu'on sait faire sur des fonctions continues ou dérivables ; on montre que ces opérations sont uniformément continues lorsque les fonctions dérivables sont considérées dans l'espace topologique des distributions, et on prolonge à toutes les distributions par continuité. On peut montrer simplement comment construire une suite de fonctions indéfiniment dérivables ayant pour limite une distribution donnée T. Si ρ est indéfiniment dérivable et à noyau compact, on montre que $T * \rho$ est une fonction indéfiniment dérivable. Il suffit alors de construire une suite de fonctions $\rho_i \geqslant 0$ dont les noyaux convergent vers l'origine, et telles que $\iint \cdots \int \rho(x_1, \ldots x_n)\, dx_1, \ldots dx_n$ tende vers 1 ; dans l'espace des distributions, les ρ_i tendent vers δ, donc les $T * \rho_i$ vers $T * \delta = T$. T apparaît bien comme la limite de ses « régularisées » indéfiniment dérivables $T * \rho_i$.

Donnons quelques exemples de la convergence dans le domaine de la série et de l'intégrale de Fourier.

Théorème. — *Quel que soit le nombre réel* α, *les fonctions* $t^\alpha e^{itx}$ *convergent vers* o *pour* $t \to \pm \infty$.

Cet exemple est d'autant plus curieux que pour $\alpha > 0$, les modules, $|t|^\alpha$, de ces fonctions convergent vers $+\infty$.

La démonstration est immédiate ; si n est un entier $> \alpha$ les

$$\frac{1}{t^n} t^\alpha e^{itx}$$

convergent uniformément vers o ; et, à un facteur près, les fonctions étudiées en sont les dérivées n-ièmes.

De même, et pour la même raison, la série de Fourier, $\Sigma a_n e^{inx}$ est convergente, dès que $|a_n| = O(|n|^\alpha)$, α réel > 0 quelconque. De là on déduit simplement ceci : toute distribution $T(x)$ périodique et de période 2π, a des coefficients de Fourier qui se calculent suivant la méthode habituelle (si $e_n(x)$ est la fonction égale à e^{inx} dans une période et nulle ailleurs, on aura par exemple $a_n e^{inx} = \frac{1}{2\pi} T * e_n$) et la série de Fourier formée à partir de ces coefficients converge vers $T(x)$ dans l'espace des distributions. Il n'y a dès lors plus lieu de conserver la distinction pénible entre séries de Fourier et séries trigonométriques qui ne sont pas des séries de Fourier, quoique convergentes par un procédé de sommation vers une fonction plus ou moins bien définie ; une série trigonométrique est *toujours* et d'une seule manière la série de Fourier d'une distribution vers laquelle elle converge dans l'espace des distributions (¹).

Résultats analogues avec l'intégrale de Fourier. Pour me borner à un exemple simple :

si pour $x \to \pm\infty \; |f(x)| = O(|x|^\alpha), \qquad \alpha > 0,$

l'intégrale de Fourier est convergente :

$$\int_{-\infty}^{+\infty} f(x) e^{itx} dx.$$

Cela signifie que $\int_{-A}^{+B} e^{itx} f(x)\, dx$, qui est une fonction de t, converge vers une limite dans l'espace des distributions de la variable t, quand A et $B \to +\infty$. Il existe une formule de réciprocité, mais elle est plus compliquée que dans le cas où $f(x)$ est sommable.

Ces considérations nous amènent donc à l'étude générale des transformations de Fourier et de Laplace. Il n'est pas possible de traiter ici la question. Si l'on veut se borner aux cas les plus utiles,

(¹) Du moins sous la seule réserve vue plus haut $|a_n| = O(n^\alpha)$. On peut lever cette restriction, comme je l'indique plus loin, mais il faut alors introduire de nouveaux types de distributions. Signalons que M. Bochner a déjà défini une généralisatiion de la transformation de Fourier qui est au fond celle-ci, mais le formalisme est beaucoup moins maniable. (Bochner : « Vorlesungen über Fouriersche Integrale ». Leipzig, 1932, p. 110-169.)

ceux de la transformation d'une fonction $f(x)$ à croissance lente à l'infini ($|f(x)| = O(|x|^{\alpha})$), on peut rester dans le cadre indiqué dans cet article. Mais il est également possible d'élargir beaucoup le champ d'application des transformations de Fourier et Laplace, et de définir les transformées de toutes les distributions, quels que soient leur irrégularité et leur comportement à l'infini ; on est alors obligé d'introduire une nouvelle famille de distributions d'un maniement nettement plus compliqué et moins intuitif. Une fonction telle que e^x a alors pour transformée une masse $+1$ au point d'abscisse imaginaire $-i$ (1) ; la transformée de e^{x^2} est également intéressante. L'étude de cette transformation de Fourier-Laplace générale paraît surtout utile dans des questions d'équations aux dérivées partielles et d'équations intégrales ; je publierai prochainement un mémoire sur les fonctions moyenne-périodiques où ces diverses notions sont avantageuses.

(1) Dans le cadre des distributions introduites jusqu'ici, les points réels seuls interviennent. Les distributions sur une droite complexe ou dans un espace complexe, nécessaires pour les transformations de Fourier-Laplace, ne satisfont plus du tout aux propriétés énoncées dans cet article.