L. Schwartz

La « fonction » δ et les noyaux

Salam (A.) & Wigner (E.P.), éd., *Aspects of Quantum Theory*,
Cambridge University Press, 1972, p. 179-182.

Extrait des *Œuvres de Laurent Schwartz*
publiées par la Société mathématique de France, 2011.

Reprinted from *Aspects of Quantum Theory*, edited by A. Salam and E. P. Wigner
(Cambridge University Press, 1972).

11

La 'fonction' δ et les noyaux

L. Schwartz

C'est en 1935 que j'entendis parler pour la première fois de la fonction δ; j'étais étudiant, et un camarade venait d'entendre une conférence de physique théorique, et m'en a parlé en ces termes: 'Ces gens-là introduisent une soi-disant fonction δ, nulle partout sauf à l'origine, égale à $+\infty$ à l'origine, et telle que $\int\delta(x)dx = +1$. Avec des méthodes de ce genre, aucune collaboration n'est possible.' Nous y avons un peu réfléchi ensemble, et avons abandonné; je n'y ai plus repensé jusqu'en 1945. A ce moment, c'est dans un but tout-à-fait différent que j'ai défini les distributions. J'étais tourmenté par les 'solutions généralisées' d'équations aux dérivées partielles. Si nous considérons l'équation des cordes vibrantes,

$$\frac{\partial^2 u}{\partial x^2} - \frac{1}{v^2}\frac{\partial^2 u}{\partial t^2} = 0,$$

on écrit sa solution générale sous la forme $u(x, t) = f(x+vt) + g(x-vt)$, où f et g sont deux fois dérivables; si elles ne sont que continues ou une fois dérivables, on a bien cependant l'impression que la fonction u définie ci-dessus est quand même solution, dans un certain sens, de l'équation des cordes vibrantes. D'où l'idée de solution généralisée. J'en avais eu besoin pour un problème, résolu dans un cas particulier, par Deny et Choquet. La définition des solutions généralisées est simple: u est solution généralisée, si toutes ses 'régularisées', $u * \phi$ (où ϕ est une fonction suffisamment différentiable pour que $u * \phi$ ait les dérivées voulues, et à support compact), est solution de l'équation. Cette définition avait d'ailleurs déjà été introduite antérieurement par Bochner. Quelque chose restait insatisfaisant; on pouvait ainsi dire que u était solution généralisée de l'équation des cordes vibrantes, mais ni $\partial^2 u/\partial x^2$ ni $(1/v^2)(\partial^2 u/\partial t^2)$ n'avaient de sens séparément. C'est de là que sont sorties les distributions. C'est seulement après que je me suis aperçu qu'elles donnaient la solution des difficultés rencontrées dans la fonction de Dirac; celle-ci devenait la distribution de Dirac. J'ai alors regardé un certain nombre de travaux de physique théorique, et me suis aperçu avec effroi de l'énorme 'percée' qu'avaient faite les physiciens dans la manipulation des distributions, sans que les mathématiciens leur en 'donnent le droit'. La physique théorique était pleine de distributions fort complexes, notamment les 'fonctions singulières' de la physique quantique, distributions invariantes par le groupe de Lorentz; tout un chapitre de la physique s'était développé avec le plus grand succès, sans que les fondements en soient assurés. Certaines de ces fonctions singulières donnèrent encore, par la suite, bien du mal aux mathématiciens (et aux physiciens, qui mirent du temps à accepter l'impossibilité de la multiplication des distributions). Toute cette expérience a été très instructive, tant pour les physiciens que pour les mathématiciens. Cependant les leçons n'en ont pas suffisamment été tirées, et les contacts entre les deux sciences sont encore restés trop rares.

Ce n'est pas seulement pour la fonction de Dirac elle-même que Dirac s'était lancé en avant, ni même pour toutes les fonctions singulières; il avait l'idée des distributions comme noyaux. Considérons une fonction de deux variables, $(x, y) \to K(x, y)$; elle définit une transformation fonctionnelle, faisant correspondre à la fonction f une fonction g par $g(x) = \int K(x, y) f(y)\, dy$. La théorie des équations intégrales avait donné un grand développement à ces noyaux. Mais toute transformation ne peut se représenter ainsi; si l'on veut écrire l'identité, le noyau devrait être précisément $\delta(x-y)$, à cause de $f(x) = \int \delta(x-y) f(y) dy$. Dirac écrivit alors que, si l'on accepte les fonctions singulières comme noyaux, toutes les transformations fonctionnelles usuelles, même les opérateurs differentiels, peuvent se représenter par des noyaux. Il se trouve en effet que, si $K_{x,y}$ est une distribution à deux variables, et si ϕ est une fonction C^∞ à support compact, on peut lui faire correspondre une distribution T, notée $K.\phi$, qu'on peut symboliquement écrire $T_x = \int K_{x,y} \phi(y)\, dy$, et que K définit ainsi une application linéaire continue de $\mathscr{D}_y$ dans $\mathscr{D}'_x$; en outre le théorème des noyaux dit précisément que toute application linéaire continue de $\mathscr{D}_y$ dans $\mathscr{D}'_x$ peut s'écrire, d'une manière unique, à partir d'un noyau distribution. Ce théorème, que j'ai démontré en 1950, est directement inspiré de la lecture de Dirac.

Cette idée de noyau, où la distribution de Dirac joue un rôle fondamental, comme noyau de l'identité, réintervient constamment partout. Je l'ai retrouvée récemment dans l'étude des désintégrations de mesures. Soit (Ω, λ) un espace probabilisé, et soit f une fonction $\geqslant 0$ borélienne sur Ω. Soit $\mathscr{T}$ une sous-tribu de la tribu λ-mesurable. On définit alors une espérance conditionnelle de f par rapport à $\mathscr{T}$, comme suit: c'est une fonction $f^{\mathscr{T}} \geqslant 0$, appartenant à la tribu $\mathscr{T}$, et telle que, pour tout ensemble $A \in \mathscr{T}$, les intégrales $\int f^{\mathscr{T}}(\omega)d\lambda(\omega)$ et $\int_A f(\omega)d\lambda(\omega)$ soient égales. Cette espérance conditionnelle existe, et est unique à un ensemble λ-négligeable près.

On aimerait en fait avoir mieux: les probabilités conditionnelles. Le système des probabilités conditionnelles s'appelle aussi la désintégration de la mesure λ par rapport à la sous-tribu; c'est une famille de probabilités sur Ω, $(\lambda_\omega{}^{\mathscr{T}})_{\omega\in\Omega}$, indexée par Ω lui-même (donc, pour tout $\omega \in \Omega$, $\lambda_\omega{}^{\mathscr{T}}$ est une probabilité sur Ω), telle que $\omega \rightarrow \lambda_\omega{}^{\mathscr{T}}$ appartienne à la tribu $\mathscr{T}$, et que, pour tout $A \in \mathscr{T}$, l'intégrale de mesures $\int_A \lambda_\omega{}^{\mathscr{T}} d\lambda(\omega)$ soit le produit χ_A de λ par la fonction caractéristique de A. L'existence et l'unicité de la désintégration ont été demontrées par Jirina. Si on possède une désintégration, elle donne des espérances conditionnelles pour toutes les fonctions f boréliennes $\geqslant 0$, d'un seul coup: on peut prendre $f^{\mathscr{T}}(\omega) = \lambda_\omega{}^{\mathscr{T}}(f) = \int_\Omega f(\omega')d\lambda_\omega{}^{\mathscr{T}}(\omega')$. Elle joue le rôle de noyau. Poussons plus loin. Soit F une fonction sur Ω, borélienne $\geqslant 0$, non plus à valeurs numériques, mais à valeurs mesures: $(Y, \mathscr{Y})$ est un ensemble Y muni d'une tribu $\mathscr{Y}$, et, pour tout ω, $F(\omega)$ est une mesure $\geqslant 0$ sur $(Y, \mathscr{Y})$. On pourra alors chercher une espérance conditionnelle de F par rapport à la sous-tribu $\mathscr{T}$; c'est une nouvelle fonction à valeurs mesures sur $(Y, \mathscr{Y})$, $F^{\mathscr{T}}$, appartenant à la tribut $\mathscr{T}$, et telle que pour tout $A \in \mathscr{T}$, les intégrales de mesures $\int_A F^{\mathscr{T}}(\omega)d\lambda(\omega)$ et $\int_A F(\omega)d\lambda(\omega)$ soient égales. Moyennant des conditions de compacité sur $(Y, \mathscr{Y})$, on peut prouver l'existence de l'espérance conditionnelle, et son unicité à un ensemble λ-négligeable près. La liaison entre espérances conditionnelles et désintégrations devient alors bilatérale. La désintégration $\omega \rightarrow \lambda_\omega{}^{\mathscr{T}}$ n'est autre que l'espérance conditionnelle, relative à la sous-tribu $\mathscr{T}$, de la fonction à valeurs mesures $\omega \rightarrow \delta_{(\omega)}$ (où, pour $\omega \in \Omega$, $\delta_{(\omega)}$ est la mesure formée de la masse $+1$ au point ω; le $\omega \rightarrow \delta_{(\omega)}$ est encore, si Ω est $\mathbf{R}$, le fameux $\delta(x-y)$); son existence résulte donc de l'existence des espérances conditionnelles des fonctions à valeurs mesures. Et inversement, si F est une fonction à valeurs mesures, son espérance conditionnelle relative à $\mathscr{T}$ peut s'obtenir à partir de la désintégration, toujours par la formule $F^{\mathscr{T}}(\omega) = \int_\Omega F(\omega')d\lambda_\omega{}^{\mathscr{T}}(\omega')$.

Employons les notations des distributions ou de la physique. Au lieu de $d\lambda(\omega)$, écrivons $\lambda(\omega)d\omega$. Les formules deviennent les suivantes. La fonction à valeurs mesures $\omega \rightarrow \delta_{(\omega)}$ devient $\delta(\omega, \omega')\, d\omega'$, avec $\int_\Omega \delta(\omega, \omega')\phi(\omega')d\omega' = \phi(\omega)$. La fonction F à valeurs mesures sera

$F(\omega, y)dy$, son espérance conditionnelle $F^{\mathscr{T}}$ sera $F^{\mathscr{T}}(\omega, y)dy$, avec la relation intégrale

$$\int_A F^{\mathscr{T}}(\omega, y)\ \lambda(\omega)d\omega = \int_A F(\omega, y)\ \lambda(\omega)d\omega.$$

La désintégration $\omega \to \lambda_\omega^{\mathscr{T}}$ se notera $\lambda^{\mathscr{T}}(\omega, \omega')d\omega'$, ou $\delta^{\mathscr{T}}(\omega, \omega')d\omega'$, avec l'égalité

$$\int_A \lambda^{\mathscr{T}}(\omega, \omega')\lambda(\omega)d\omega = \int_A \delta(\omega, \omega')\lambda(\omega)d\omega = \chi_A(\omega')\lambda(\omega').$$

Et $F^{\mathscr{T}}$ sera donnée à partir de F et de la désintégration par

$$F^{\mathscr{T}}(\omega, y) = \int_\Omega F(\omega', y)\lambda^{\mathscr{T}}(\omega, \omega')d\omega'.$$

On trouve bien alors, pour $A \in \mathscr{T}$:

$$\begin{aligned}\int_A \delta^{\mathscr{T}}(\omega, y)\lambda(\omega)d\omega &= \int_A \lambda(\omega)d\omega \int_\Omega F(\omega', y)\delta\mathrm{N}(\omega, \omega')d\omega' \\ &= \int_\Omega d\omega' F(\omega', y) \int_A \delta^{\mathscr{T}}(\omega, \omega')\lambda(\omega)d\omega \\ &= \int_\Omega F(\omega', y)\chi_A(\omega')\lambda(\omega')d\omega' = \int_A F(\omega', y)\lambda(\omega')d\omega'.\end{aligned}$$

On peut étudier ensuite les désintégrations régulières par rapport à une famille de tribus $(\mathscr{T}^t)_{t\in\mathrm{R}}$, dépendant du temps t, et obtenir des résultats, avec la même liaison bilatérale, pouvant s'appliquer fructueusement aux martingales et aux processus de Markov. Les noyaux pour définir des transformations, un noyau de Dirac pour l'identité, c'est encore la même idée, dans un contexte pourtant très différent.

Je m'excuse d'apporter ici une contribution purement suggestive à l'œuvre de Dirac, qui par ailleurs dépasse considérablement la découverte de la fonction δ !