I.R.E.M. DE LORRAINE
Université Henri Poincaré Nancy 1

Campus Scientifique Victor Grignard - BP 239 - 54506 VANDŒUVRE-LES-NANCY cedex
Tél. 03 83 68 49 41 *(secrétariat)* - 03 83 68 49 45 *(bibliothèque)* - *Fax.* 03 83 68 43 94

Refonder l'enseignement de l'analyse

Un extrait du programme de Yannis Varouchas

Jean-Pierre Ferrier

DIDACTIQUES N° 5

© **Edité et imprimé par l'Institut de Recherche sur l'Enseignement des Mathématiques** - (Université Henri Poincaré - Nancy I - Faculté des Sciences) - B.P. 239 - VANDŒUVRE-les-NANCY Cedex.
Dépôt légal : 2ème trimestre 2003
n° de la publication : 2 - 85406 - 173 - X
Responsable de la publication : Le Directeur de l'IREM, Jean-Pierre FERRIER

Didactiques 5

REFONDER L'ENSEIGNEMENT DE L'ANALYSE
Un extrait du programme de Yannis Varouchas

rédigé et commenté par Jean-Pierre Ferrier
Irem de Lorraine

Μὴδε βάλητε τοῦς μαργαρίτας ὑμῶν ἔμπροσθεν τῶν χοίρων

Mathieu 7-6

J'ai choisi d'écrire ce petit texte dans l'espoir faire vivre cet investissement considérable que Yannis Varouchas a entrepris vis-à-vis de l'enseignement des mathématiques à l'université. Alors que beaucoup de collègues se contentent de puiser dans ce qui existe de façon parfois approximative, il a toujours été animé par la volonté de renouveler les contenus et les approches en respectant une rigueur sans faille.

Ce besoin de renouvellement présente lui-même deux aspects. Il y a bien sûr des considérations d'ordre strictement pédagogique, découlant du souci d'être accessible par des étudiants pas toujours motivés. Je qualifierai cet aspect de didactique, même s'il n'y a pas toujours eu derrière une véritable expérimentation et surtout l'analyse correspondante.

Cependant la motivation principale est plus esthétique, liée à la très heureuse utopie suivant laquelle les mathématiques doivent rester simples. Autrement dit il s'agit de réécrire Bourbaki. Le fameux traité a été conçu au départ comme un acte pédagogique et il ne faudrait pas l'oublier, même s'il n'est pas très adapté à nos étudiants d'aujourd'hui et qu'il ne l'a été qu'à une infime minorité des étudiants d'hier. Je qualifierai donc cet aspect de refondateur. Il possède une logique propre, indépendante de l'acte d'enseignement à un niveau donné.

Pour éviter des rappels incessants, j'emploierai le pronom personnel "on" pour tout ce qui relève des aspects refondateur ou didactique. J'emploierai la première personne du singulier pour mes remarques personnelles et la première personne du pluriel pour tout ce qui concerne des discussions entre nous deux, lesquelles ont été nombreuses.

Je commence par l'aspect refondateur, pour dire d'abord qu'il a été fortement inspiré par un goût, poussé très loin, pour le constructivisme. Je ne parlerai cependant pas ici des aspects les plus techniques dans la mesure où la traduction dans l'enseignement n'a pas été opérée.

L'exigence constructiviste demande par exemple de renoncer à l'axiome du choix général, et de se contenter de la récurrence à choix multiples. En d'autres termes, au lieu de supposer qu'un produit d'ensembles non vides est non vide, on suppose qu'est non vide la limite projective d'un suite de tels ensembles.

Je me suis largement inspiré de ces préceptes dans mon cours d'Analyse fonctionnelle de maîtrise, sous une forme moins radicale cependant. Par exemple j'ai d'abord donné la compacité d'un produit dénombrable d'espaces compacts métrisables, avant d'évoquer le cas général et de démontrer aussi celle d'un produit dénombrable d'espaces compacts généraux, comme limite projective d'une suite de produits finis. J'ai pris grand soin de montrer que la boule unité faible du dual d'un espace normé séparable était compacte et métrisable. Ce faisant, d'une part j'évitais d'utiliser le cas général et d'autre part d'imposer une distance a priori pour définir la topologie faible.

Il y a quelquechose d'absurde dans la considération des espaces compacts généraux en Analyse. En effet pratiquement personne ne songe à parler d'un cadre adéquat pour la convergence. Celui des suites est parfaitement inadapté. Aussi faudrait-il impérativement parler de filtres ou d'un équivalent. Cela est pour montrer que renoncer à l'axiome du choix général n'est pas une coquetterie. C'est, bien au contraire, l'assurance d'un discours cohérent.

De la même façon j'ai d'abord donné le théorème de Hahn-Banach dans le cas d'un espace normé séparable avant d'évoquer le cas général. D'ailleurs le théorème de Hahn-Banach n'est pas d'une importance aussi grande qu'on le pense parfois. Pour l'utiliser avec l'espace E il faut en effet connaître son espace dual. Or, en déterminant ce dual dans les cas utiles comme les L^p, on montre en fait que l'orthogonal d'un sous-espace fermé strict est non nul. Et c'est par cette propriété qu'on utilise vraiment Hahn-Banach.

Une autre exigence constructiviste, et c'est aussi un point sur lequel nous nous sommes rencontrés, est de ne pas se laisser tenter par la généralisation factice. Yannis Varouchas m'expliquait que, plus jeune, il aurait mis un point d'honneur à ne supposer sur les fonctions considérées que les propriétés minimales nécessaires.

Avec plus d'expérience, il se rendait compte que cela n'avait aucune importance. De mon côté, je me souviens lui avoir dit que toute fonction partout définie était nécessairement continue. Il m'a demandé si c'était un théorème ou un principe d'action. Nous sommes tombés d'accord sur ce principe qu'il n'a transgressé qu'à une occasion, me semble-t-il, pour supposer équilibrées les fonctions continues par morceaux alors qu'il aurait dû les supposer non définies aux points de continuité et les équilibrer après coup.

De façon semblable, nous nous sommes entendus pour rejeter toute convergence simple. Pour les séries de Fourier, il ne doit être question que de convergence L^2, normale ou uniforme. Un énoncé, comme celui de Riemann suivant lequel la série de Fourier d'une fonction non presque partout nulle ne peut converger simplement en tout point vers 0, est à jeter aux oubliettes. D'une part sa démonstration est difficile, non éclairante et complètement réfractaire aux méthodes modernes; d'autre part il ne sert, à proprement parler, à rien.

Ce souci de ne pas compliquer inutilement le discours va de pair avec une exigence de pureté qui ne semble plus à la mode, ni dans les cours universitaires faits de bric et de broc, ni dans le modèle de l'Agrégation qui préfère une tête bien pleine à une tête bien faite. Cela impose un exposé fidèle à la même ligne d'un bout à l'autre. S'y tenir conduit souvent à innover de façon très heureuse. J'en donnerai beaucoup d'exemples.

J'en viens à des préoccupations plus pédagogiques, mais qui témoignent aussi d'une rigueur intellectuelle sans faille. Aujourd'hui il est admis qu'un élève peut traverser avec succès toute sa scolarité de l'école maternelle à l'université sans ne jamais avoir rien compris. Alors que les programmes sont libres à l'université, beaucoup de collègues se moquent de l'appropriation éventuelle par les étudiants des connaissances qu'ils enseignent.

Les préoccupations dont je parlerai relèvent d'une tout autre vision des choses. Au départ elles peuvent emprunter quelques sentiers battus. Par exemple, pour tenter de faire comprendre à ses étudiants la notion de limite, on fera calculer explicitement, pour une suite (u_n) dont cherche à montrer qu'elle converge vers l et un $\epsilon > 0$ donné, un nombre N tel que $n \geq N$ implique $|u_n - l| \leq \epsilon$. Mais cela est sans doute naïf. On finira alors par adopter un point de vue radicalement différent où il n'est plus du tout question de ϵ, de façon visible au moins.

Ces préoccupations s'opposent également à une perversion courante qui conduit à multiplier inutilement les définitions axiomatiques. Est-il besoin de savoir tout de suite ce qu'est un espace compact en général? Il suffit dans un premier de connaître la compacité d'un segment, d'un rectangle ou d'un disque fermés.

Evidemment, chacun l'aura compris, comprendre est ici tout le contraire du "avoir entendu parler de" qui sévit aujourd'hui. Les programmes parlent de continuité, de dérivabilité au lycée mais tout y est admis, y compris les définitions. De leur côté, les classes préparatoires ont inventé une théorie de l'intégration pour laquelle les théorèmes principaux sont également admis. Pourquoi ne pas se contenter d'énoncés plus faibles, mais suffisants, dont la démonstration peut être exigée?

Pour finir, je ne devrais pas ignorer le soin méticuleux que Yannis Varouchas apportait à la typographie. Qu'il me pardonne donc d'utiliser TeX, qui est le système passe-partout aujourd'hui. Bien que détestant Microsoft, il était un adepte de Word qu'il savait utiliser d'une façon personnelle, élégante et systématique, certainement plus conforme à l'esprit de Donald Knuth que l'usage que beaucoup font de TeX aujourd'hui. Evidemment il s'agissait de Word 5 pour Macintosh avec son langage de balises pour les formules mathématiques. Rien à voir avec l'éditeur d'équations qui sévit dans les versions plus récentes. Bien sûr cela demandait beaucoup d'efforts en amont. Personne ne pouvait le suivre sur son terrain.

J'ajoute un complément. C'est un exposé de Calcul intégral que j'ai fait en petit comité en décembre dernier et qui complète certains points abordés ici. Yannis Varouchas y assistait. C'est la dernière fois que nous avons eu l'occasion de discuter ensemble.

CALCUL DIFFERENTIEL
Refondation

Continuité.

Soit f une fonction d'une variable réelle, définie sur un intervalle I. Conformément à notre principe d'action, on lui imposera d'être continue. Cependant on n'entend pas par là que c'est une fonction continue en tout point de I. On veut précisément dire que c'est une fonction uniformément continue sur tout segment de I.

Une fonction f est dite **continue** sur un segment $[a, b]$ s'il existe un module de continuité ω tel que

$$|f(x) - f(y)| \leq \omega(|x - y|)$$

pour tous x, y dans $[a, b]$.

Un module de continuité est une fonction $\omega \geq 0$ croissante sur la demi-droite $[0, +\infty[$ qui prend sur $]0, +\infty[$ des valeurs arbitrairement petites; cela impose $\omega(0) = 0$ entre autres.

Cela n'est pas exactement la définition proposée dans le programme, où il est demandé qu'il existe une fonction δ de l'intervalle $]0, +\infty[$ dans lui-même telle que $\omega(\delta(\epsilon)) \leq \epsilon$ pour tout $\epsilon > 0$. Certaine dernière propriété correspond très exactement à l'existence pour tout $\epsilon > 0$ d'un $\delta > 0$ convenable. C'en est une version algébrisée.

Chacun aura reconnu une définition équivalente à la définition classique, compte-tenu du théorème de Heine.

Par exemple le module de continuité $x \mapsto kx$ conduit aux fonctions lipschitziennes dans le rapport k.

On notera qu'une fonction continue sur un segment est toujours bornée. On montre sans difficulté la continuité d'une somme, d'un produit, d'une composée. Pour l'inverse de f, il faut supposer, dans un premier temps, la fonction $|f|$ minorée.

On remarque que l'extension à plusieurs variables ne présente aucune difficulté particulière. On écrirait

$$|f(x, y) - f(x', y')| \leq \omega(\max(|x - x'|, |y - y'|))$$

en deux variables par exemple.

Les propriétés des fonctions continues sur un segment $[a, b]$ s'établissent de façon plus constructive, par dichotomie, que dans la démarche classique.

C'est notamment le cas pour le théorème des valeurs intermédiaires, qui fournit donc l'occasion de développer la méthode dichotomique pour résoudre l'équation $f(x) = 0$. Contrairement à ce beaucoup de gens croient, cette méthode, telle qu'exposée habituellement, ne relève pas du calcul numérique. On y néglige en effet l'imprécision des données et les erreurs de calcul.

Une attention particulière doit être cependant apportée au théorème de la borne inférieure. En effet, pour établir la continuité de l'inverse d'une fonction continue f ne s'annulant pas, il faut savoir que la fonction $|f|$ est minorée par une constante $c > 0$.

Prenons le cas des fonctions lipschitziennes, même si ce qu'on va dire peut s'étendre au cas général. Aussi bien pour le théorème de la borne supérieure que pour celui des valeurs intermédiaires, on peut calculer une à une les décimales d'un nombre où f atteindra son maximum ou bien où f s'annulera. On rejoint ainsi la philosophie d'Henri Lebesgue, pour qui un nombre était avant tout un développement décimal (limité ou non). Cela évite d'avoir à présupposer une propriété des nombres réels, comme celle des segments emboités.

Dérivabilité.

Il s'agit ici de donner des pistes pour réécrire une partie du traité de Bourbaki, plus précisément une partie du livre IV sur les fonctions d'une variable réelle. Ce dernier est remarquablement écrit et relativement élémentaire. Ce qui suit est encore plus basique évidemment.

Bien évidemment, suivant notre principe d'action, il ne saurait être question de dérivée en dehors des fonctions continûment dérivables.

On dira que la fonction (continue) F admet la fonction (continue) f comme **dérivée** sur I s'il existe une fonction g continue sur $I \times I$ telle que

$$\begin{cases} F(x) - F(y) = (x-y)g(x,y) \\ f(x) = g(x,x) \end{cases}$$

pour tous x, y dans I. On écrira $f = F'$.

La fonction g sera appelée **fonction de pente**. C'est l'existence d'une fonction de pente continue qui définit la dérivabilité.

On notera que g est unique, et donc aussi f. C'est le principe de prolongement des égalités, qu'on établit d'habitude par passage à la limite.

Cependant une démonstration sans parler de limite est possible. Supposons que $g_1(x,y) = g_2(x,y)$ pour $x \neq y$ alors que $g_1(a,a) \neq g_2(a,a)$, soit $|g_1(a,a) - g_2(a,a)| = \epsilon > 0$. Considérons un module de continuité ω commun à g_1 et g_2. Il existe $\eta > 0$ tel que $\omega(\eta) < \epsilon/2$. Alors $|g_1(a, a+\eta) - g_2(a, a+\eta)| \leq 2\omega(\eta) < \epsilon$. C'est absurde.

Il n'est pas difficile d'établir dans ce cadre la dérivée d'une somme, d'un produit, d'une fonction composée, de la fonction $x \mapsto 1/x$, donc d'un inverse et d'un quotient.

Un exemple.

Voyons comment **calculer une dérivée**, en prenant l'exemple de la fonction $\sin x$. Ma mémoire fait ici défaut. Je ne sais plus comment il faut s'y prendre et je vais improviser dans le désordre.

Posant $h = x - y$, on a

$$\sin x - \sin y = 2 \sin \frac{h}{2} \cos \frac{x+y}{2} = (x-y).\phi(\frac{x-y}{2}) \cos \frac{x+y}{2}$$

où $\phi(h) = \sin h/h$ pour $h \neq 0$ et $\phi(0) = 1$.

Par suite la dérivée sera donnée en x par

$$\phi(0) \cos x = \cos x\ .$$

Cependant un problème reste à régler. Comment montrer que la fonction ϕ est continue. Autrement dit dit comment calculer un module de continuité pour cette fonction? C'est là que je ne sais plus faire.

Une autre idée, un peu détournée, serait celle-ci. On suppose que l'on défini π comme l'aire du disque unité.

On calcule l'aire d'un triangle rectangle en A dont un côté AB de l'angle droit est 1 et dont l'angle en B est x. On obtient

$$\tan x = \int_0^x \frac{dt}{\cos^2 t} dt$$

en multipliant par 2. Cette relation vaut pour $-\pi/2 < x < \pi/2$.

Il en résultera que la dérivée de $\tan x$ est donnée par

$$\frac{1}{\cos^2 x} = 1 + \tan^2 x \ .$$

Maintenant

$$\cos x = \frac{1}{\sqrt{1 + \tan^2 x}}$$

sur l'intervalle considéré. En dérivant il vient

$$-\frac{1}{2} \frac{1}{(1 + \tan^2 x)^{3/2}} .2 \tan x . (1 + \tan^2 x) = -\sin x \ .$$

Le reste des formules de dérivation en découle.

On peut obtenir plus directement la fonction sin en calculant l'aire comprise entre un arc d'angle au centre x et sa corde, dans le disque unité. On obtient

$$\frac{x - \sin x}{2} = \int_0^x \sin^2 \frac{t}{2} \, dt$$

et on en déduit que la dérivée de $x - \sin x$ eest $1 - \cos x$.

Cette méthode passe par une intégrale. Cela suppose que l'on a établi la dérivabilité de l'intégrale par rapport à sa borne supérieure, ce que dont on parlera à la prochaine section. Cependant le calcul d'aire sur lequel on s'est appuyé est beaucoup plus facile à établir que l'équivalence en 0 entre la longueur de l'arc et celle de la corde.

Le calcul utilise l'encadrement d'une petite aire qui faist passer de x à $x + \Delta x$. On s'appuie sur l'aire d'un secteur angulaire, qu'on ramène au disque unité par homothétie et qu'on obtient ensuite par proportionnalité.

Dérivée et intégrale.

Proposition. *Etant donnés une fonction continue f sur I et un point a de I, posons*

$$F(x) = \int_a^x f(t)dt$$

pour x dans I. Alors F est dérivable et f est la dérivée de F.

En effet on aura

$$F(x) - F(y) = \int_x^y f(t)dt = (x-y)\int_0^1 f((1-u)x+uy)du$$

en opérant le changement de variable $t = (1-u)x + uy$. En résulte la propriété de définition, où

$$g(x,y) = \int_0^1 f((1-u)x+uy)du$$

et $g(x,x) = f(x)$.

Bien sûr cela relève de la continuité sous le signe intégrale. Mais ici ce n'est pas difficile car

$$|g(x,y) - g(x',y')| \leq \omega(\max(|x-x'|, |y-y'|))$$

si ω est un module de continuité de f sur un segment donné.

Proposition. *Supposons F dérivable sur I, de dérivée f. Alors*

$$F(b) - F(a) = \int_a^b f(t)dt$$

si a, b sont dans I.

On introduit la fonction g de pente. Supposant $a \leq b$ par exemple, on découpe le segment $[a,b]$ en segments de longueur h assez petite et on écrit

$$F(a+nh+h) - F(a+nh) = hg(a+nh, a+nh)$$

à une erreur près, que l'on peut majorer par

$$h.\omega(h)$$

où ω est un module de continuité de g. Le reste suit.

On notera qu'avec la dérivée classique la dernière propriété fait appel au théorème des accroissements finis, au moins pour caractériser les fonctions constantes. Or cet énoncé est l'un des plus difficiles de l'Analyse; il nécessite une construction rigoureuse des nombres réels.

Il faut voir que la dérivée fait perdre toute l'information sur la partie tangente à zéro dans une relation comme

$$f(t) = f(a) + (t-a)f'(a) + o(|t-a|) \ .$$

Si l'on avait pu instiller dans le o de Landau une quelconque indication sur son uniformité par rapport à a, tout aurait été plus facile. Mais l'existence d'une dérivée en chaque point ne la donne pas.

Commentaires.

Dans sa fameuse construction, Henri Lebesgue cherchait à faire rattraper à l'intégrale son retard sur la dérivée. Il ne pouvait pas se douter qu'il condamnait la notion classique de dérivée au profit de sa nouvelle intégrale.

Aujourd'hui on dérive les fonctions **au sens des distributions**. Dire que la fonction, i.e. fonction localement intégrable f, est la dérivée de la fonction F au sens des distributions sur l'intervalle ouvert I signifie très exactement que F est continue et que

$$F(x) - F(a) = \int_a^x f(t)dt$$

pour tout x dans I, où a est un point fixé de I. Ainsi la dérivée découle-t-elle de l'intégrale.

Cela vaut en particulier si f est continue ou continue par morceaux. Déjà cela nous donne sign x comme dérivée de $|x|$. Il faut dire qu'il est particulièrement irritant de ne pas pouvoir dériver une telle fonction.

Bien sûr cela reporte la difficulté sur la définition de l'intégrale. On peut voir cette dernière comme une **aire plane**.

On peut être plus précis. Si l'on découpe $[a, b]$ en intervalles égaux de longueur h, on a

$$\int_a^b f(t)dt = h[f(a) + \cdots + f(a+kh) + \cdots + f(a+(n-1)h)]$$

avec une erreur que l'on peut majorer par $(b-a)\omega(h)$ où ω est un module de continuité uniforme de h, ou encore par $|f(b) - f(a)|.h$ si f est monotone.

CALCUL DIFFERENTIEL
Aspect didactique

Comment faut-il définir un module de continuité? La définition algébrisée par la condition $\omega(\delta(x)) \leq x$ n'est pas très parlante.

Bien sûr, en demandant que la fonction ω prenne sur $]0, +\infty[$ des valeurs arbitrairements petites, on introduit un quantificateur. Malgré tout la continuité en 0 d'une fonction monotone reste beaucoup plus simple que celle d'une fonction générale. Ou bien ω prend sur $]0, +\infty[$ des valeurs arbitrairement petites, ou alors il existe $\epsilon > 0$ tel que $\omega \geq \delta$ sur cet intervalle.

De toute façon la définition proposée semble bien adaptée pour les deux premières années de l'université. On pourrait réserver la définition classique et l'équivalence avec la précédente à la troisième année universitaire, où l'on serait peut-être alors moins enclin à parler d'espaces trop généraux.

Le module de continuité $\omega(x) = kx$ conduit aux fonctions lipschitziennes dans le rapport k; il suffit pour étudier les fonctions polynômiales et trigonométriques, et établir le théorème des valeurs intermédiaires dans ce cadre. On aurait pu y penser en écrivant les programmes de l'enseignement secondaire. C'est toujours mieux que de définir comme continues les fonctions mentionnées au bulletin officiel ou que d'admettre l'existence de primitives pour les fonctions ... dérivables !

Au lycée comme à l'université, démontrer la continuité d'une somme, d'un produit, d'un inverse, d'une composée dans le cadre présenté ici sont des exercices faciles et instructifs. C'est l'occasion d'établir des majorations et, sur ce point parmi d'autres, on ne peut pas dire que l'enseignement actuel soit très performant.

De même, démontrer la dérivabilité d'une somme, d'un produit, d'un inverse, d'une composée sont des exercices également faciles et instructifs.

Prenons l'exemple de la fonction composée de

$$u = f(x)$$

et

$$y = g(u)$$

où l'on aura noté F et G les fonctions de pente pour ces fonctions.

On a

$$\Delta y = \Delta u.G(u + \Delta u, u)$$

et

$$\Delta u = \Delta x.F(x + \Delta x, x) \ .$$

Par suite
$$\Delta y = \Delta x . G(u + \Delta u, u) . F(x + \Delta x, x)$$
et la dérivée de la fonction composée est donnée par
$$g'(u) f'(x) \; .$$

Cependant, au lycée, jouer avec les fonctions lipschitziennes est peut-être trop ambitieux déjà. On peut se contenter d'y présenter une dérivée comme une pente limite, une intégrale comme une aire algébrique. Le mot limite ne serait pas défini, par plus qu'on ne parlerait de continuité, de dérivabilité et a fortiori d'intégrabilité. On se contenterait de calculer des dérivées, en levant l'indétermination, et des intégrales.

Voici, par exemple, comment introduire le mot limite pour illustrer une dérivée. On va se placer, pour indiquer le format, dans le cadre toujours considéré ici. La version plus modeste consisterait à présenter la même chose sur des exemples.

On fixe $y = a$. On a
$$\frac{F(x) - F(a)}{x - a} = f'(a) + g(x, a) - g(a, a)$$
où $|g(x, a) - g(a, a)|$ est majoré par un certain $C|x - a|$. On dira qu'on obtient $f'(a)$ à la limite quand x tend vers a.

En fait, même si le cadre considéré ne convient peut-être pas au plus grand nombre des élèves, au moins a-t-il le mérite de fournit un socle épistémologique cohérent. Il met en avant des énoncés de base démontrables : propriété des valeurs intermédiaires, dérivation par rapport à la borne supérieure d'une intégrale, formule intégrale des accroissements finis, dérivation des fonctions trigonométriques. C'est déjà infiniment supérieur à la présentation de définitions incompréhensibles donnant lieu à des résultats qui sont pratiquement tous admis.

Pour finir, le fait de faire reposer la dérivée sur l'intégrale permet aussi de tirer un parti plus sûr de l'intuition sur le seul point qui n'est pas formalisé et qui concerne la définition du nombre π. En effet les propriétés des aires planes sont suffisamment naturelles pour ne pas avoir besoin d'être explicitées dans un premier temps. Les risques d'erreurs sont beaucoup moins présents à propos des aires qu'ils ne le sont avec les limites.

CALCUL INTEGRAL
Refondation

Théorie complète.

Il s'agit de la théorie de Lebesgue. On peut encore ajouter que le but final de la théorie, tel qu'il apparaît aujourd'hui, n'est pas d'obtenir, comme beaucoup le croient, le plus grand nombre de fonctions intégrables qu'il est possible. Le but est de construire un espace de Hilbert L^2 qui soit complet. Cela figurait même parmi les questions posées par Yannis Varouchas à ses examens de Licence.

On cherche une définition des fonctions intégrables à partir d'un espace vectoriel réticulé $\mathcal{E}$ de fonctions simples dont on connait l'intégrale. On suppose que l'on déjà discuté les questions de négligeabilité. Il n'est pas facile d'en trouver une qui satisfasse les trois exigences suivantes.

(i) La définition vaut pour la mesure de Lebesgue sur $\mathbf{R}$, sur $\mathbf{R}^n$, pour une mesure de Radon ou encore pour une mesure abstraite sur une algèbre (pas encore une tribu).

(ii) Elle conduit aussi bien à la définition de l'espace L^1 qu'à celle des espaces L^p, pour $1 \leq p < +\infty$.

(iii) Elle s'étend au cas de fonctions vectorielles, à valeurs dans un espace de Banach.

C'est à ce prix là qu'on pourra prétendre imposer la réécriture du traité de Bourbaki, précisément celle du livre VI et surtout du chapitre 6 d'intégration vectorielle.

Une telle définition est possible si l'on pense à définir les espaces L^p comme des complétés. C'est la voie que j'ai suivi. Cependant on perd a priori la relation entre les L^p. La propriété (ΓB) qui va suivre permet de la récupérer.

Dans la suite on désignera par des lettres grecques les fonctions de $\mathcal{E}$, sur lesquelles l'intégrale est déjà connue: fonctions continues (ou seulement continues par morceaux) à support compact sur $\mathbf{R}$, sur $\mathbf{R}^n$ ou sur un espace localement compact, ou encore combinaisons linéaires de fonctions caractéristiques de parties de mesure finie d'une algèbre abstraite.

Définition. *La fonction f est dite intégrable si elle vérifie la condition:*

(ΓB) *il existe des suites ϕ_n et ψ_n telles que*
la fonction f est limite presque partout de la suite ϕ_n,
$|\phi_n| \leq \psi_n$,
la suite ψ_n est croissante ,
la suite $\int \psi_n$ est majorée.

L'existence d'une telle suite ψ_n sera qualifiée de condition de *domination mixte*; il est équivalent de dire que les intégrales des enveloppes supérieures finies des $|\phi_n|$ sont uniformément bornées.

La définition d'une fonction de L^p se fait de même, en remplaçant $\int \psi_n$ par $\int \psi_n^p$. Le passage aux fonctions vectorielles se fait en remplaçant $|\phi_n| \leq \psi_n$ par $\|\phi_n\| \leq \psi_n$.

On note au passage que la fonction réelle f est intégrable (ou dans L^p) si et seulement f^+ et f^- le sont; en particulier $|f|$ l'est avec f. Les fonctions intégrables constituent un espace vectoriel réticulé $\mathcal{I}$. Si f est positive, alors f est dans L^p si et seulement si f^p est intégrable.

Exemples.

1) On désigne par $\mathcal{E}^*$ (resp. $\mathcal{E}_*$) l'ensemble des fonctions réelles qui sont limites croissantes (resp. décroissantes) $f = \lim \phi_n$ de fonctions de $\mathcal{E}$ telles que la suite $\int \phi_n$ soit majorée (resp. minorée).

Les fonctions de $\mathcal{E}^*$ et $\mathcal{E}_*$ sont presque partout finies; elles sont intégrables.

La vérification est facile. Soit par exemple $f = \lim \phi_n$ dans $\mathcal{E}^*$. On se ramène au cas où f, puis les ϕ_n sont positives. Supposons les $\int \phi_n$ majorées par M. Donnons-nous $\epsilon > 0$. On définit A_n comme la partie sur laquelle

$$\phi_n \geq M/\epsilon \,.$$

Il est clair que la mesure de A_n est majorée par ϵ. D'où la propriété.

2) On suppose que

$$f = \sum_n \chi_n$$

simplement presque partout, la série $\int |\chi_n|$ étant convergente. Alors la fonction f est intégrable.

En effet la suite ϕ_n des sommes partielles, définie par

$$\phi_n = \sum_{k \leq n} \chi_k$$

et la suite ψ_n définie par

$$\psi_n = \sum_{k \leq n} |\chi_k|$$

vérifient la propriété (ΓB).

Noter que la seule convergence de la série $\int |\chi_n|$ implique la convergence simple absolue presque partout de la série χ_n; en effet $\sum |\chi_n|$ est dans $\mathcal{E}^*$.

Pour définir l'intégrale, on s'appuiera sur le résultat suivant.

Propriété fondamentale. *Sous l'hypothèse* (ΓB), *la suite double*

$$\int |\phi_m - \phi_n|$$

tend vers 0 *quand* $m, n \to \infty$.

Avant d'établir cette propriété, voyons ce que l'on peut en tirer. D'abord la suite $\int \phi_n$ vérifie le critère de Cauchy, puisque

$$\left| \int \phi_m - \int \phi_n \right| \leq \int |\phi_m - \phi_n| \, .$$

Elle admet donc une limite. De plus cette limite est indépendante du choix de ϕ_n. Etant données des suites ϕ_n et ϕ'_n, pour lesquelles on peut prendre la même suite ψ_n, on peut fabriquer une nouvelle suite qui les mélange. Il en résulte que $\int \phi_n$ et $\int \phi'_n$ ont la même limite.

Définition. *L'intégrale de* f *est la limite commune à toutes les suites* $\int \phi_n$.

On notera que la propriété (ΓB) implique alors que $\int |f - \phi_n| \to 0$.

Retour sur les exemples.

1) Pour une fonction $f = \lim \phi_n$ de $\mathcal{E}^*$ ou $\mathcal{E}_*$, on a

$$\int f = \lim \int \phi_n$$

sous les hypothèses indiquées.

2) Pour une fonction $f = \sum \chi_n$, avec les hypothèses indiquées, on a

$$\int f = \int \sum \chi_n \, .$$

On notera que l'on a aussi

$$\int \left| \sum \phi_n \right| \leq \sum \int |\phi_n|$$

sous les mêmes hypothèses.

Voici une autre conséquence. Si f vérifie (ΓB), quitte à remplacer les suites ϕ_n et ψ_n par des suites extraites, on peut supposer

$$\int |\phi_{n+1} - \phi_n| \leq 2^{-n} \, .$$

Posant $\chi_0 = \phi_0$ et $\chi_n = \phi_{n+1} - \phi_n$, cela montre que l'on peut écrire

$$f = \sum_n \chi_n$$

simplement presque partout, où la série $\int |\chi_n|$ est convergente. Autrement dit *l'exemple 2) est le cas général.*

Au passage on a montré que l'on pouvait remplacer, dans la propriété (ΓB), la condition de domination mixte par

$\int |\phi_m - \phi_n| \to 0$ *quand* m, $n \to \infty$.

Voici une application immédiate.

Théorème de complétion. *Soit* f_p *une suite de fonctions intégrables telle que*

$$\int |f_p - f_q| \to 0$$

quand p, $q \to \infty$. *Alors il existe une fonction intégrable* f *telle que*

$$\int |f_p - f| \to 0$$

quand $p \to \infty$.

Il suffit d'établir la propriété pour une suite extraite. On peut se ramener au cas où $\int |f_{p+1} - f_p| \leq 2^{-p}$. Pour p fixé, soit ϕ_p telle que $\int |f_p - \phi_p| \leq 2^{-p}$. On a évidemment

$$\int |\phi_{p+1} - \phi_p| \leq 2^{-p+2} \; .$$

de sorte que la suite ϕ_p converge simplement presque partout vers une fonction intégrable f et que $\int |\phi_p - f| \to 0$.

On en vient à **démontrer la propriété fondamentale**. On raisonne par l'absurde. Si la conclusion n'était pas vraie on pourrait trouver $\epsilon > 0$ et des suites m_p et n_p tendant vers l'infini telles que

$$\int |\phi_{m_p} - \phi_{n_p}| \geq \epsilon \; .$$

Posant

$$\xi_p = |\phi_{m_p} - \phi_{n_p}| \quad , \quad \eta_p = \psi_{\max(m_0,\ldots,m_p,n_0,\ldots,n_p)} \; ,$$

on aurait $\int \xi_p \geq \epsilon$, alors que la suite ξ_p converge vers 0 simplement presque partout, que

$$0 \leq \xi_p \leq \eta_p \; ,$$

que la suite η_p est croissante et la suite $\int \eta_p$ est majorée.

On va alors utiliser le lemme qui suit, pour lequel on suppose seulement que $\mathcal{E}$ est un espace vectoriel réticulé.

Lemme clé. *Etant donnés une suite ϕ_n qui vérifie une condition de domination mixte et $\epsilon > 0$, si les ϕ_n sont ≥ 0 et convergent simplement vers 0 en dehors d'une partie N, on peut trouver une suite décroissante θ_n avec les mêmes propriétés et telle que*

$$\int \theta_n \geq \int \phi_n - \epsilon \, .$$

On pose

$$\psi_{n,p} = \max(\phi_n, \ldots, \phi_{n+p}) \, .$$

Pour chaque n, on choisit p_n tel que $\psi_n = \psi_{n,p_n}$ vérifie

$$\int \psi_n \geq \sup_p \int \psi_{n,p} - \epsilon 2^{-n-2} \, .$$

C'est possible puisque la suite croissante $(\int \psi_{n,p})_p$ est majorée. La suite ψ_n converge vers 0 comme ϕ_n. Comme $(\psi_{n+1} - \psi_n)^+ \leq \psi_{n,p} - \psi_n$ pour p assez grand, on a

$$\int (\psi_{n+1} - \psi_n)^+ \leq \epsilon 2^{-n-1} \, .$$

On considère maintenant

$$\theta_n = \min(\psi_0, \ldots, \psi_n) \, .$$

La suite θ_n est décroissante et converge vers 0 comme ψ_n. De plus

$$\theta_n + (\psi_n - \psi_{n-1})^+ + \cdots + (\psi_1 - \psi_0)^+ \geq \psi_n \geq \phi_n$$

d'où $\int \theta_n + \epsilon \geq \int \phi_n$.

En appliquant le lemme clé, quitte à changer ϵ, on se ramène au cas où la suite ϕ_n est décroissante en gardant $\int \phi_n \geq \epsilon$.

On se débarrasse ensuite du presque partout. Pour cela on construit une suite croissante $\gamma_n \geq 0$ telle que $\sup \gamma_n \geq \phi_0$ sur la partie négligeable N à exclure et $\int \gamma_n \leq \epsilon/2$. On remplace alors ϕ_n par $(\phi_n - \gamma_n)^+$.

Tout résulte alors de la propriété suivante, qui est vérifiée par l'intégrale dans $\mathcal{E}$, et dont la démonstration dépend du choix de cet espace :

si la suite ϕ_n décroissante converge simplementvers 0, alors $\int \phi_n \to 0$.

Voici encore des applications de tout ce qu'on vient d'établir.

Propriété d'encadrement. *Soient f une fonction réelle intégrable et $\epsilon > 0$. Il existe alors des fonctions g de $\mathcal{E}_*$ et h $\mathcal{E}^*$ telles que $g \leq f \leq h$ presque partout et que*

$$\int h - \int g \leq \epsilon .$$

On écrit en effet f sous la forme $\sum_n \chi_n$ et on considère N tel que

$$\sum_{n>N} \int |\chi_n| \leq \epsilon.$$

On pose

$$g = \chi_0 + \cdots + \chi_N - \sum_{n>N} \chi_n^- \quad \text{et} \quad h = \chi_0 + \cdots + \chi_N + \sum_{n>N} \chi_n^+ .$$

Les propriétés demandées se vérifient aussitôt.

Remarque. Si A est négligeable, on peut trouver une fonction g positive de $\mathcal{E}^*$ infinie sur A. Il en résulte que la propriété d'encadrement peut être exigée partout.

Convergence simple absolue des séries. *Soit (f_n) une suite de fonctions intégrables. On suppose que la série $\int |f_n|$ converge. Alors la série f_n converge simplement absolument presque partout.*

En particulier si une suite f_n de fonctions intégrables vérifie

$$\int |f_n| \to 0$$

on peut en extraire une suite qui converge simplement presque partout vers 0.

On montre d'abord la première assertion. On supposera $f_n \geq 0$. Grâce à la propriété d'encadrement, on peut trouver h_n dans $\mathcal{E}*$ telle que $f_n \leq h_n$ presque partout et que

$$\int h_n - \int f_n \leq 2^{-n} .$$

Chaque h_n s'écrit comme limite croissante d'une suite croissante $(\phi_{n,p})_p$. On peut se ramener au cas où la suite double $\phi_{n,p}$ est aussi croissante en n, en prenant le maximum sur les indices précédents. On pose alors

$$g_p = \sum_{n \leq p} \phi_{n,p} \quad \text{et} \quad g = \lim g_p .$$

On a

$$\sum |f_n| \leq \sum |h_n| \leq g$$

presque partout et

$$\int g \leq \sum \int |g_p| \leq \sum \int |f_n| + 2$$

montre que g est presque partout finie.

On termine par les énoncés fondamentaux.

Théorème de convergence décroissante vers 0. *Si la suite f_n de fonctions intégrables ≥ 0, vérifiant une condition de domination mixte, converge vers 0 simplement presque partout, alors $\int f_n \to 0$.*

On peut d'abord trouver g_n dans $\mathcal{E}_*$ telle que $0 \leq g_n \leq f_n$ presque partout et que

$$\int f_n - \int g_n \leq 2^{-n} .$$

Il suffit bien sûr de montrer que $\int g_n \to 0$. En appliquant le lemme clé avec des ϕ_n dans $\mathcal{E}_*$, on se ramène à faire la démonstration dans le cas où la suite g_n est décroissante.

Chaque g_n est la limite d'une suite décroissante $(\phi_{n,p})_p$ et on se ramène facilement au cas où la suite double $\phi_{n,p}$ est aussi décroissante en n. Dans ce cas la suite $\phi_{n,n}$ converge vers 0 presque partout. Il en résulte que $\int \phi_{n,n} \to 0$, puis que $\int g_n \to 0$.

On en déduit aussitôt le théorème suivant où l'on retrouve (ΓB) dans une situation plus générale.

Théorème de convergence mixte. *Soient f_n, g_n des fonctions intégrables. On suppose les propriétés qui suivent pour une constante $M \geq 0$.*

(i) *$f_n \to f$ simplement presque partout.*
(ii) *$|f_n| \leq g_n$ pour tout n.*
(iii) *$g_n \leq g_{n+1}$ pour tout n.*
(iv) *$\int g_n \leq M$ pour tout n.*
Alors f est intégrable et

$$\int f = \lim \int f_n .$$

En effet, en reprenant l'argumentation utilisé pour la propriété fondamentale, on montre aussitôt que

$$\int |f_m - f_n| \to 0$$

à l'aide du théorème de convergence vers 0. Le théorème de complétion montre alors l'existence d'une fonction intégrable g telle que

$$\int |f_n - g| \to 0 \; .$$

Une suite extraite de la suite f_n converge alors vers g presque partout. Ainsi $f = g$ presque partout. Enfin $|\int f_n - \int f| \leq \int |f_n - f|$ permet de conclure.

Compléments.

On peut facilement définir l'intégrale supérieure $\int^* f$ d'une fonction réelle f comme $+\infty$ si f est majorée par aucune fonction de $\mathcal{E}^*$ et comme borne inférieure des $\int g$ où g est dans $\mathcal{E}^*$ et majore f sinon. On définit symétriquement l'intégrale inférieure $\int_* f$. Il apparaît facilement que

la fonction f est intégrable si et seulement si elle vérifie $\int^* f = \int_* f$ et si cette valeur commune est finie.

On définit les fonctions mesurables à la manière des fonctions intégrables, mais en omettant la condition de domination mixte. On montre sans difficulté que

si f est mesurable et si $|f| \leq g$ où g est intégrable, alors f est intégrable.

Remarques.

Le lemme clé combine deux fonctions. D'une part il sert à remplacer une suite presque décroissante par une suite exactement décroissante comme on le fait dans le théorème de Beppo-Levi. D'autre part il permet de ramener la convergence simple à des convergences monotones comme on le fait dans le théorème de convergence dominée; dans ce cas cependant, il évite une double limite en passant par une limite presque décroissante.

Dans ce lemme, l'hypothèse suivant laquelle les ϕ_n sont ≥ 0 et convergent vers 0 en dehors de N peut être omise; la conclusion est alors l'existence d'une suite décroissante θ_n telle que

$$\lim \theta_n \leq \limsup \phi_n \quad \text{et} \quad \int \theta_n \geq \int \phi_n - \epsilon \; .$$

En effet on a successivement $\limsup \psi_n \leq \limsup \phi_n$ et $\lim \theta_n \leq \limsup \psi_n$. On a préféré donner l'énoncé dans le cas particulier qui est utile, pour éviter la limite supérieure.

Toujours dans ce lemme, on notera que l'utilisation d'une écriture du type $(f-g)^+$ est purement cosmétique. Il n'est pas nécessaire de prendre des différences. En effet

$$(f-g)^+ = f - \min(f,g)$$

et

$$\int (f-g)^+ \leq a \quad \text{signifie} \quad \int f \leq \int \min(f,g) + a,$$

l'écriture utilisée étant surtout plus facile à interpréter. Ainsi peut-on énoncer le lemme pour un cône réticulé $\mathcal{C}$.

Cela peut en avoir un intérêt si l'on cherche à définir l'intégrale en commençant par l'étendre à $\mathcal{E}^*$ et en procédant ensuite par encadrement. Dans ce cas on n'introduira pas à l'avance les parties négligeables.

Ces considérations permettraient aussi d'établir le théorème de prolongement d'une mesure de probabilités définie sur une algèbre. Dans ce cas il n'y a pas de condition de domination mixte. Le lemme clé permet de remplacer une suite A_n par une suite B_n décroissante telle que

$$\lim B_n \leq \limsup A_n \quad et \quad P(B_n) \geq P(A_n) - \epsilon \ .$$

Théories incomplètes.
En fait la propriété (ΓB) n'a pas été introduite pour l'intégrale de Lebesgue. Elle l'a été dans le programme refondateur pour démontrer le théorème d'interversion d'une intégrale et d'une série dans le cas des intégrales absolument convergentes, comme celles des fonctions continues, ou continues par morceaux ou encore localement Riemann-intégrables sur un intervalle I.

C'est elle qui remplace la propriété de domination dans ce cadre. En effet, si pour l'intégrale de Lebesgue, il est possible de décomposer en montrant d'abord l'intégrabilité de la limite ψ des ψ_n, cela ne l'est plus dans le cas considéré, puisqu'on ne sait rien a priori de la régularité de ψ.

Théorème de convergence mixte. *Soient f, f_n, g_n des fonctions admettant des intégrales absolument convergentes. On suppose les propriétés qui suivent pour une constante $M \geq 0$.*
(i) $f_n \to f$ *simplement.*
(ii) $|f_n| \leq g_n$ *pour tout* n.
(iii) $g_n \leq g_{n+1}$ *pour tout* n.
(iv) $\int g_n \leq M$ *pour tout* n.
Alors

$$\int f = \lim \int f_n \ .$$

La démonstration repose sur la méthode indiquée pour le lemme. Les choses sont un peu plus simples. D'abord il n'y a plus d'ensemble négligeable. Ensuite on travaille sur $|f_n - f|$ pour se ramener au théorème de convergence vers 0.

Pour ce dernier, on utilise directement le lemme clé en montrant que si une suite θ_n de fonctions positives admettant des intégrales absolument convergentes décroît simplement vers 0, alors $\int \theta_n$ tend vers 0.

A cette fin on se donne $\epsilon > 0$ et on considère un segment S tel que l'intégrale de θ_0 sur le complémentaire soit majorée par $\epsilon/2$. On est ainsi ramené au cas d'un segment.

Noter que l'exposé originel reposait sur un théorème de convergence uniformément bornée d'Arzelà sur un segment $[a, b]$. Ce même énoncé se déduisait d'un lemme clé de convergence uniformément bornée, monotone décroissante vers 0. Ce lemme s'appliquait non pas à des fonctions supposées Riemann-intégrables, mais à l'intégrale inférieure de fonctions quelconques.

La présentation qui en a été donnée ici s'en inspire évidemment. Cependant, pour des raisons de commodité, on n'a pas fait apparaître l'intégrale inférieure.

Sur le plan technique, la différence principale avec l'intégrale de Lebesgue réside dans le fait que l'on aura utilisé une seule fois le lemme clé au lieu de deux.

CALCUL INTEGRAL
Aspect didactique

Une stratégie possible.

On voit se dessiner une progression possible pour le cours d'intégration. Au lycée d'abord, l'intégrale sera présentée comme une aire plane, en s'appuyant sur l'idée intuitive qu'on s'en fait, sans aucune définition abstraite. On fera comme on faisait autrefois pour les cas d'égalité en géométrie.

En première année de classe préparatoire ou d'université, on introduira les sommes de Riemann canoniques pour établir l'intégrale d'une fonction continue sur un segment. La notion d'intégrabilité au sens de Riemann est à éviter : on privilégie la définition en extension à la définition axiomatique. Parler par ailleurs de diverses notions d'intégrabilité ne fait qu'embrouiller. On le constate encore à l'Agrégation.

En deuxième année de classe préparatoire ou d'université, on pourra se contenter de l'intégrale absolument convergente d'une fonction continue sur un intervalle. Là encore on ne parlera pas d'intégrabilité. Cela conduit à donner deux noms différents à la même chose et à ranger deux choses différentes sous un même nom.

Etablir le théorème de convergence mixte sous l'hypothèse (ΓB) n'offre pas de grande difficulté dans le cas des intégrales absolument convergentes de fonctions continues. La seule démonstration à faire est celle du lemme clé. Or cette dernière ne prend qu'une demi-page. On devrait pouvoir la présenter à ce niveau.

Ensuite, en troisième année d'université, on passera à la théorie de Lebesgue et aux vraies fonctions intégrables.

On commencera par discuter la négligeabilité, qui est le fait vraiment nouveau et gagne sans doute à être travaillé pour lui-même, avant tout le reste.

Puis on s'appuiera sur la propriété (ΓB) pour définir les fonctions intégrables. On dispose déjà du lemme clé qui est le seul énoncé dont la démonstration soit délicate. Autrement dit tout le travail est déjà fait.

Enfin, dans le cours de probabilités, on reprendra le lemme clé pour établir le théorème de prolongement.

Ma propre vision consiste à partir de la complétion de L^1 pour aboutir à la propriété (ΓB) et non l'inverse. Elle me paraît intellectuellement plus satisfaisante parce qu'elle part de ce qu'on veut vraiment et dégage la négligeabilité de façon naturelle. Malheureusement elle est condamnée devant de vrais étudiants car elle fait une place trop belle aux constructions abstraites. Aussi la présentation donnée ici a-t-elle toute sa place.

Dans la stratégie proposée, la condition (ΓB) apparaît trois fois. Une fois pour le théorème de convergence mixte des intégrales impropres, une fois pour définir les fonctions intégrables, une fois pour les théorème mixte de l'intégrale de Lebesgue. On peut espérer que la répétition facilitera l'acquisition de ce concept. D'autre part la formulation est très proche de l'application, puisque la convergence monotone et la convergence dominée en dérivent très directement. Enfin on dispose, immédiatement et sans surprise, d'une définition des fonctions intégrales. On peut espérer que cela facilitera aussi l'apprentissage.

Un autre choix.

La piste indiquée précédemment n'a pas servi pour refonder l'enseignement du Calcul intégral en troisième année. C'est une autre qui a été prise, à propos de laquelle nous avons beaucoup discuté entre nous pour nous rejoindre là encore sur l'essentiel. Cependant un certain nombre des principes retenus restent compatibles avec la stratégie citée.

L'une des raisons de cet autre choix est qu'on ne pouvait pas se fonder sur une progression cohérente du type de celle envisagée précédemment. Il fallait un contenu indépendant de tout ce qui aurait pu être enseigné avant.

Il s'agit d'abord de fournir très vite les énoncés fondamentaux qui vont permettre aux étudiants de se familiariser avec leur maniement. Il faut en effet au moins trois semaines pour que la condition de domination, pour les suites ou les intégrales à paramètres, soit maîtrisée. Il en faut presque autant pour les intégrales multiples.

Aussi le premier principe est-il le suivant. Ces théorèmes figurent tous dans un premier chapitre, et leur démonstration est renvoyée à un chapitre ultérieur. Cet artifice ne se serait pas imposé de façon si nette dans le cadre de la stratégie précédente.

Un autre impératif est de familiariser les étudiants avec la notion de mesure, qu'ils retrouveront en théorie des probabilités. Cependant la mesurabilité au sens de Lebesgue est impossible à mettre en défaut sans recourir à des procédés interdits. Aussi fait-on, dans un premier temps, comme si tout était mesurable.

Le second principe consiste à indiquer dans le premier chapitre une liste de mesures qu'on rencontrera.

L'exemple le plus important est celui de la mesure sur la droite, ou celui de la loi uniforme sur $[0, 1]$ si l'on préfère. Il n'est pas raisonnable d'en admettre brutalement l'existence ni réaliste de l'établir péniblement en début de cours. Aussi renvoie-t-on à ce sujet à un chapitre ultérieur. Là on peut discuter sur la méthode. On peut construire une mesure régulière sur un espace localement compact à partir de sa donnée sur les parties ouvertes. On peut aussi faire découler la mesure des parties de la droite de l'intégrale des fonctions obtenue directement. Cependant, dans les deux cas, c'est difficile.

On voit les avantages de ces choix en matière de souplesse vis-à-vis d'une population étudiante très hétérogène. Les étudiants qui ne visent pas la recherche ou l'agrégation pourront se contenter du premier chapitre, la suite du cours mettant l'accent sur les relations avec ce qui a été présenté dans les années précédentes.

Complément : complétion L^1

décembre 2002

Fonctions en escalier.

Dans l'algèbre des réunions finies d'intervalles [disjoints], on définit la mesure comme la longueur totale. On etablit la σ-additivité, *pour des parties de longueur finie*, sous la forme:

si une suite décroissante A_n est d'intersection vide, alors $\text{mes}(A_n) \to 0$.

Soit $\epsilon > 0$. On remplace A_n par sa fermeture B_n, ce qui revient à ajouter les extrémités manquantes. Alors l'intersection B_n est compacte, de la forme $D = \{x_0, \ldots, x_n, \ldots\}$. On la recouvre par les intervalles ouverts $]x_n - \epsilon 2^{-n}, x_n + \epsilon 2^{-n}[$, donc aussi par un nombre fini d'entre eux. De plus cette réunion finie contient B_n pour n assez grand et sa mesure est au plus 4ϵ.

Dans l'espace vectoriel $\mathcal{E}$ des sommes finies $\sum \lambda_I 1_I$, où chaque I est un intervalle fini, pour lequel $\int f = \sum \lambda_I \text{mes}(I)$, on en déduit aussitôt la propriété de convergence monotone:

si une suite f_n de $\mathcal{E}$ décroît vers 0, *alors* $\int f_n \to 0$.

Pour $\epsilon > 0$ donné, il apparaît que la mesure de l'ensemble des points x tels que où $f_n(x) \geq \epsilon$ tend vers zéro.

L'espace L^1 en une variable.

Pour une fonction réelle f, on note f^+ la partie positive, qui vaut $\max(0, f(x))$ en x, et f^- la partie négative, qui est $(-f)^+$ et vaut $\max(0, -f(x)) = -\min(0, f(x))$ en x. On a $f = f^+ - f^-$ et $|f| = f^+ + f^-$.

Nous partons de l'espace vectoriel ordonné $\mathcal{E}$ des fonctions en escalier et de l'intégrale définie sur cet espace comme forme linéaire positive.

Cet espace est muni de la norme

$$||f||_1 = \int |f(t)|dt$$

et l'intégrale est continue, de norme ≤ 1.

On définit L^1 comme le complété de $\mathcal{E}$ et on y prolonge l'intégrale. Tout le problème est d'interpréter les éléments de L^1 comme des fonctions sur la droite. Par construction un élément f de L^1 s'écrit sous la forme

$$\sum_n f_n$$

où f_n est une série absolument convergente de $\mathcal{E}$.

Il est facile de voir qu'on peut se ramener au cas où

$$f_n = \lambda_n 1_{I_n}$$

les I_n étant des intervalles et les λ_n des nombres réels tels que

$$\sum |\lambda_n| \mathrm{mes}(I_n) < +\infty \ .$$

Cependant, cela ne vaut que pour L^1; de plus, contrairement au cas de $\mathcal{E}$, on ne peut pas supposer les I_n disjoints. Enfin L'intégrale est donnée par

$$\sum \int f_n \ .$$

En un point x on cherche à définir $f(x) = \sum_n f_n(x)$, par exemple par

$$\sum_n f_n^+(x) - \sum_n f_n^-(x)$$

ce qui est possible si chacune des deux sommes est finie, ou si $\sum |f_n|$ l'est.

Disons provisoirement qu'un ensemble A de la droite est *exceptionnel* s'il existe une série (f_n) de fonctions ≥ 0 de $\mathcal{E}$ telle que $\sum \|f_n\|_1 < +\infty$ et $\sum f_n(x) = +\infty$ sur A. Alors on pourra définir $f(x)$ en dehors d'un tel ensemble; c'est une lapalissade.

On verra que modifier les valeurs d'une fonction sur un tel ensemble est sans effet sur l'intégrale; on préfère donc parler d'ensemble **négligeable** ou de mesure nulle. Une propriété sera dite vraie **presque partout** (en abrégé pp) si elle est vraie en dehors d'un ensemble négligeable.

On vérifie qu'on peut imposer $\sum \|f_n\|_1 \leq \epsilon$ où $\epsilon > 0$ est donné dans la définition d'une partie négligeable. On montre alors facilement qu'une réunion dénombrable de parties négligeables l'est aussi.

Noter qu'un intervalle I non nul ne pourra pas être négligeable: en effet $\int 1_I > 0$.

En fait la clé de la construction réside dans le

Lemme. Soit $f = \sum_n f_n$ dans L^1. Alors $f(x) \geq 0$ pp en x si et seulement $f \geq 0$ dans $\hat{E}$, i.e. si $\|F_n^-\|_1 \to 0$, où F_n est la somme partielle d'ordre n de f_n.

Noter que cela renvoie à une représentation $f = \sum_n f_n$ particulière.

1. Montrons que la condition est *suffisante*. Pour chaque q entier ≥ 0, on choisit n_q tel que

$$\|F_{n_q}^-\|_1 \leq 2^{-q} \quad \text{et} \quad \sum_{p>n_q} \|f_p\|_1 \leq 2^{-q} \ .$$

Considérons

$$g_q = F_{n_q}^- + \sum_{p>n_q} f_p^- \ .$$

En un point x tel que $f(x)$ soit défini, on a

$$-f(x) = -F_{n_q} - \sum_{p>n_q} f_p \leq g_q(x) \ .$$

Si de plus $f(x) < 0$, alors $\sum_q g_q(x) = +\infty$. Or $\sum_q g_q$ est la somme d'une série double absolument convergente. Ainsi $f(x) \geq 0$ pp.

2. Montrons que la condition est *nécessaire.* On fixe $\epsilon > 0$ et considère une série g_p de fonctions ≥ 0 de $\mathcal{E}$ telle que $\sum \|g_p\|_1 \leq \epsilon$ et que $\sum g_p(x) = +\infty$ en un point x où $f(x) < 0$ ou bien où $f(x)$ n'est pas défini.

Là où la valeur $f(x)$ est définie et positive, on a

$$F_n^-(x) \leq F_n^-(x) + \sum_{p>n} f_p^-(x) \leq F_n^+(x) + \sum_{p>n} f_p^+(x)$$

et on peut retirer $F_n^+(x)$ du second membre car $F_n^-(x) = 0$ si $F_n^+(x) > 0$. Alors

$$F_n^-(x) \leq \sum_{p>n} f_p^+(x) + \sum_p g_p(x)$$

partout.

Si n fixé, et $q \geq n$, il en résulte que

$$F_n^- = \sum_{n<p\leq q} f_p^+ + \sum_{p\leq q} g_p + h_q$$

où h_q est une suite décroissante de fonctions de $\mathcal{E}$ dont la limite est ≤ 0. Alors h_q^+ décroît vers 0, donc $\|h_q^+\|_1 \to 0$. Finalement

$$\|F_n^-\|_1 \leq \sum_{p>n} \|f_p\|_1 + \sum_p \|g_p\|_1 \leq \sum_{p>n} \|f_p\|_1 + \epsilon$$

et la propriété à démontrer suit.

De la condition suffisante il résulte que la classe associée à f est bien définie: si $\sum_n f_n = 0$ dans L^1, alors $\lim \|F_n\| = 0$ et $\sum_n f_n = 0$ pp. De la condition nécessaire il résulte que la classe caractérise f.

La positivité de l'intégrale est claire pour l'ordre dans L^1. On notera que, pour f dans L^1, on peut définir f^+, f^- et $|f|$: à partir de $f = \sum_n f_n$, on considèra $g_n = F_n^+ - F_{n-1}^+$ par exemple. Alors $\|f\|_1 = \int |f|$ s'obtient en passant à la limite.

Les fonctions (pp définies) représentant les éléments de L^1 sont appelées fonctions intégrables.

Remarque. Soient f une fonction intégrable et $\epsilon > 0$. On peut écrire f sous la forme

$$f = \sum f_n$$

où $\sum \|f_n\|_1 \leq \|f\|_1 + \epsilon$, et où la série représente f (qui est donc définie) là où elle est absolument convergente.

L'inégalité sur les normes provient directement de la définition de la norme dans le complété. Supposons là vérifiée avec $\epsilon/2$ au lieu de ϵ.

Si la série converge absolument sans représenter f sur une partie négligeable N, on considère une série g_n qui diverge sur N et telle que $\sum \|g_n\|_2 \leq \epsilon/2$. Maintenant si l'on intercale g_n et $-g_n$ et que l'on remplace par $f_0, g_0, -g_0, f_1, g_1, -g_1 \cdots$ la série f_n, on obtient le résultat cherché.

Le théorème d'intégration des séries.

Théorème. Soit u_n une série de fonctions intégrables telle que la série $\sum \int |u_n|$ converge; alors

(i) la série $u_n(x)$ converge pour presque tout x,

(ii) sa somme pp définie est intégrable $\int \sum u_n = \sum \int u_n$.

Pour le démontrer écrivons

$$f_n = \sum_p f_{n,p}$$

de façon à représenter f_n là où la série de droite converge absolument, et à vérifier

$$\sum_p \|f_{n,p}\|_1 \leq \|f_n\|_1 + 2^{-p} .$$

La série double $f_{n,p}$ est absolument convergente dans L^1 et donc absolument convergente en dehors d'une partie négligeable N. La série f_n converge alors absolument en dehors de N.

Le théorème de Fubini.

On définit l'intégrale de Lebesgue dans $\mathbf{R}^n$ de la même façon que dans $\mathbf{R}$ en partant de l'intégrale des fonctions de $\mathcal{E}(\mathbf{R}^n)$, produit tensoriel d'exemplaires de $\mathcal{E}(\mathbf{R})$.

Théorème. Soit f une fonction **intégrable** sur $\mathbf{R}^2$; alors

(i) pour presque tout y la fonction qui à x associe $f(x,y)$ est intégrable,

(ii) la fonction pp définie qui à y associe $\int f(x,y)dx$ est intégrable

(iii) et
$\int\int f(x,y)dxdy = \int\big(\int f(x,y)dx\big)dy$ qui est encore $\int\big(\int f(x,y)dy\big)dx$.

Pour le démontrer on se donne une fonction

$$f = \sum \phi_n$$

dans $L^1(\mathbf{R}^2)$, sous la forme d'une série de $\mathcal{E}(\mathbf{R}^2)$ absolument convergente. On suppose $f(x,y)$ défini et égal à la somme en tout point (x,y) tel que $\sum|\phi_n(x,y)| < +\infty$.

On a

$$\iint f = \sum \iint \phi_n = \sum \int \psi_n(y)dy \quad \text{où} \quad \psi_n(y) = \int \phi_n(x,y)dx$$

par définition de $\iint f$ d'abord, et par application du théorème dans $\mathcal{E}(\mathbf{R}^2)$ ensuite. En même temps

$$+\infty > \sum \iint |\phi_n| = \sum \int \theta_n(y)dy \quad \text{où} \quad \theta_n(y) = \int |\phi_n(x,y)|dx \ .$$

Alors, pp en y, la série $\theta_n(y)$ converge et la série $\psi_n(y)$ converge absolument. De plus

$$\iint f = \int \sum \psi_n(y)dy = \int \Big(\int \sum \phi_n(x,y)dx\Big)dy = \int\Big(\int f(x,y)dx\Big)dy$$

la seconde application du théorème sur les séries étant valable pp en y.

Le théorème de dérivation sous l'intégrale.

Théorème. Soient I est un intervalle et f une application de $I \times \mathbf{R}$ dans $\mathbf{R}$; si

(i) pour un point a dans I la fonction qui à t associe $f(a,t)$ est intégrable,

(ii) pour presque tout t la fonction qui à x associe $f(x,t)$ possède une dérivée partielle faible $\partial f/\partial x$ en x,

(iii) Cette dérivée $\partial f/\partial x$ est intégrable en (x,t) sur $S \times \mathbf{R}$ pour tout segment S de I,

alors la fonction F définie par $F(x) = \displaystyle\int f(x,t)dt$ possède une dérivée faible sur I et

$$F'(x) = \int \frac{\partial f}{\partial x}(x,t)\, dt\ .$$

On rappelle que la fonction localement intégrable h est une dérivée faible de g si

$$g(x) - g(a) = \int_a^x h(t)dt$$

pour un a et pour tout x.

Une parenthèse.

Les théorèmes de convergence découlent aussitôt du théorème sur les séries. On en déduit une caractérisation de l'intégrabilité utilisant la notion de mesurabilité.

Par définition la fonction réelle f est dite mesurable si la fonction qui à x associe $1_{[-N,N]}(x)\min(1, N/|f(x)|)f(x)$ est intégrable pour tout nombre entier N.

Il résulte du théorème de convergence dominée que si f est mesurable et $|f| \leq g$ où g est intégrable, alors f est intégrable. Il en résulte aussi la stabilité de la mesurabilité par limite simple pp.

Espaces L^p.

On peut définir les espaces L^p par complétion de la même manière que L^1. Seule l'intégrale n'a plus sa place. Cependant il faut relier les L^p entre eux.

Par exemple la négligeabilité est indépendante de p. En effet celle de A équivaut à la propriété suivante:

pour tout $\epsilon > 0$, A est incluse dans la réunion d'une suite croissante $B_n$de Atelle que $\mathrm{mes}(B_n) \leq \epsilon$.

Pour établir cette propriété, pour A est négligeable, on considère une série (f_n) de fonctions ≥ 0 de $\mathcal{E}$ telle que $\sum \|f_n\|_p \leq \epsilon$ et que $\sum f_n(x) = +\infty$ sur A. Alors la suite B_n définie par $\sum_{m\leq n} f_m(x) \geq 1$ convient.

La réciproque est facile; on donne à ϵ la suite de valeurs 2^{-k} et on somme.

En particulier, pour $1 < p < +\infty$, l'application bilinéaire $(f, g) \to fg$ de $\mathcal{E} \times \mathcal{E}$ dans $\mathcal{E}$ se prolonge en une application bilinéaire

$$L^p \times L^q \to L^1$$

de norme ≤ 1, toujours définie comme le produit en chaque point.

Maintenant la mesurabilité est aussi indépendante de p. Pour tronquer $f = \sum f_n$ de façon à maintenir $|f(x)| \leq N.1_{[-N,N]}$, on peut tronquer les sommes partielles F_n de façon que $|F_n(x)| \leq N.1_{[-N,N]}$. Alors les f_n vérifient $|f_n(x)| \leq 2N.1_{[-N,N]}$. Dans ce cas la convergence absolue de la série ne dépend pas de p, et l'intégrabilité non plus.

Notons surtout que les éléments f de L^p sont caractérisés par la propriété (ΓB) suivante:
il existe des suites (F_n), (G_n) de E telles que F_n converge pp vers f et que

$$|F_n| \leq G_n$$

la suite G_n étant croissante.
On en déduit que si f est mesurable, alors f est dans L^p si et seulement si $|f|$ est dans L^1. De plus $\|f\|_p = \left(\int |f|^p\right)^{1/p}$.

Fonctions vectorielles.

Tout ce qui a été fait peut être étendu au cas de fonctions à valeurs dans un espace de Banach. Le fait de disposer d'emblée de la complétion des espaces L^p est un avantage.

Le théorème de la convergence dominée s'obtient dans le cas scalaire en notant que $g_n = \sup_{n \leq p} f_p$ décroît vers f et qu'à n fixé, $h_n = \sup_{n \leq p \leq m} f_p$ croît vers g_n; on l'étend au cas d'une suite double (ou d'un filtre à base dénombrable).

On obtient l'énoncé vectoriel à partir du cas scalaire, en considérant la suite double $\|f_n - f_m\|$ et en montrant que (f_n) est de Cauchy dans L^1; sa limite ne peut être que f puisqu'on peut en extraire une suite qui converge pp.

La notion de mesurabilité en découle.

Appendice

Séries doubles.

Soit une série double $(u_{n,p})$ dans l'espace normé E telle que $\sum_{n,p} \|u_{n,p}\| < +\infty$. Alors

$$\sum_n \sum_p u_{n,p} = \sum_r \sum_{n+p=r} u_{n,p}$$

dès que la série interne de gauche et la série de droite convergent.

En effet, étant donné $\epsilon > 0$, on peut trouver K tel que

$$\Big\| \sum_{n \leq N} \sum_{p \leq P} u_{n,p} - \sum_{r \leq R} \sum_{n+p=r} u_{n,p} \Big\| \leq \epsilon$$

dès que N, P, $R \geq K$. Il suffit alors de passer à la limite sur r et sur p. On en déduit la limite sur n et l'égalité.

Complétion d'un espace normé.

Il est bien connu qu'on peut définir, pour un espace normé, la complétion en considérant des séries absolument convergentes plutôt que des suites de Cauchy. On va donc considérer des séries pour définir le complété.

Si E est un espace normé, on considère l'espace $L^1(\mathbf{N}, E)$ des séries absolument convergentes de E, muni de la norme

$$\|(x_n)\| = \sum \|x_n\| \, .$$

On considère aussi le sous-espace vectoriel fermé F de $L^1(\mathbf{N}, E)$ engendré par les séries $0, \ldots 0, x, 0, \ldots, 0, -x, 0, \ldots$; c'est l'adhérence du sous-espace des séries à support fini de somme nulle.

On considère enfin $\hat{E} = L^1(\mathbf{N}, E)/F$.

On note que E s'envoie naturellement dans $L^1(\mathbf{N}, E)$, puis dans $\hat{E}$, par l'application linéaire de norme ≤ 1 qui à x associe la série $x, 0, 0, \ldots$.

Pour E complet, on note aussi que l'application de $L^1(\mathbf{N}, E)$ dans E qui à une série associe sa somme est linéaire et de norme ≤ 1.

Désormais E est un espace normé quelconque.

a. L'image de E dans $\hat{E}$ est dense. En effet un élément du quotient est représenté par une série; il suffit de tronquer cette série, puis de tout concentrer tout sur le premier terme, ce qui revient à ajouter un élément de F.

b. Il en résulte que si G est un espace normé complet et $u : E \to G$ linéaire continue, il existe au plus une factorisation

$$E \to \hat{E} \to G$$

linéaire continue; c'est le principe de prolongement des identités.

c. L'existence d'une telle factorisation existe, et le fait qu'elle soit de norme au plus égale à celle de u, s'obtient en considérant la suite

$$E \to L^1(\mathbf{N}, E) \to L^1(\mathbf{N}, G) \to G\ .$$

La flèche centrale est celle qui à la série (x_n) associe la série $(u(x_n))$. Il suffit alors de passer au quotient par F.

d. Maintenant toute série absolument convergente de $\hat{E}$ est convergente, de sorte que E est complet.

On note d'abord une série absolument convergente de E converge dans $\hat{E}$ vers la classe qu'elle définit.

Si l'on considère une série absolument convergente de $\hat{E}$, on commence par représenter son terme général par un élément x_p de $L^1(\mathbf{N}, E)$ de façon à conserver la convergence absolue; on perdrait 2^{-p} à l'indice p par exemple.

Chaque x_p est une série $(x_{p,n})_n$ et la série double $(x_{p,n})_{p,n}$ est absolument convergente.

Alors la série simple de terme général

$$\sum_{p+n=r} x_{p,n}$$

définit un élément de $L^1(\mathbf{N}, E)$ qui sera, modulo F, la somme cherchée.

e. De plus, si x est dans E et si y est dans F, alors $\|x + y\| \geq \|x\|$; on commence par le cas où y est à support fini où c'est une conséquence de l'inégalité triangulaire et on passe à la limite.

On en déduit que E s'identifie à un sous-espace normé de $\hat{E}$. En particulier ce plongement est injectif.

f. On suppose donnés un espace normé complet G et une application linéaire continue $E \to G$ qui identifie E à un sous-espace dense de G; alors elle se prolonge en un isomorphisme isométrique de $\hat{E}$ sur G.

En effet, du fait de **b**, l'application se prolonge en une application linéaire continue $\hat{E} \to G$ de norme ≤ 1. Cette application conserve la norme sur E. Par densité elle la conserve sur $\hat{E}$. Ainsi l'image de $\hat{E}$ est un sous-espace complet dense de G; c'est G.

g. Soient F un espace normé et E un sous-espace dense de F. On suppose donnés un espace normé complet G et une application linéaire continue $u : E \to G$; alors elle se prolonge de façon unique à F, sans augmentation de la norme.

Pour l'existence on peut toujours se ramener au cas où F est complet, quitte à le compléter. Comme **f** permet d'identifier F à $\hat{E}$, aussitôt **b** donne le prolongement. L'unicité peut se montrer par densité.

On peut compléter cet énoncer en considérant une structure plus riche que celle d'espace normé. Les espaces que nous considérons sont aussi ordonnés, c'est à dire dotés d'un cône positif. Alors $l^1(\mathbf{N}, E)$ sera naturellement doté d'un tel cône. Quant au complété $\hat{E}$, son cône positif s'obtient en prenant l'adhérence du cône positif de E. On notera que ce n'est pas en général l'image de ce cône.

On pourrait aussi considérer une structure d'algèbre normée. Une telle structure se transmet à $l^1(\mathbf{N}, E)$ par la multiplication de Cauchy.

La même construction, à quelques détails près, est celle qui permet de construire le corps $\mathbf{R}$ en complétant le corps $\mathbf{Q}$. Aussi a-t-elle une importance culturelle évidente.

TITRE : **Refonder l'enseignement de l'Analyse, extrait du programme de Yannis Varouchas**
AUTEURS : **Jean-Pierre Ferrier**
PUBLIC VISE : **enseignants de lycée et d'université**
RESUME : **le document présente et commente diverses innovations introduites par Yannis Varouchas dans son enseignement d'Analyse à l'université; on y envisage des retombées possibles concernant le lycée.** *Thèmes abordés* : **calcul différentiel, calcul intégral.**
NOTE: **peut inspirer la réforme du LMD.**
MOTS CLE : **accroissements finis, complétion, (fonction) continue, convergence dominée, convergence monotone, dérivée, domination mixte, fonction composée, fonctions trigonométriques, (fonction) intégrable, intégrale, intégrales convergentes, intégration des séries, module de continuité, (ensemble) négligeable, presque partout, valeurs intermédiaires.**