SUR L'INTÉGRALE DE LEBESGUE.

PAR

FRÉDÉRIC RIESZ

à KOLOZSVÁR.

(Extrait d'une lettre à M. G. MITTAG-LEFFLER.)

L'été dernier, quand nous parlions des idées de M. LEBESGUE, vous m'avez aimablement invité à vous donner, pour les Acta mathematica, quelques détails au sujet d'une note que j'ai publiée antérieurement (Comptes rendus, le 4 mars 1912) et dans laquelle j'ai indiqué comment je m'imagine l'introduction de l'intégrale indépendamment de la théorie de la mesure. Les voilà; mais avant tout, permettez-moi de vous rappeler que ce n'est pas moi le premier qui ai tâché de refondre l'œuvre admirable de M. LEBESGUE, que c'était plutôt la lecture d'une note de M. BOREL sur ce sujet (Comptes rendus, le 12 février 1912) qui m'a donné l'occasion d'exposer mes idées sur la même question. Depuis-là, peu après, M. BOREL a développé sa méthode dans un mémoire détaillé qu'il faisait paraître dans le Journal des mathématiques (6e série, t. 8, p. 159). En étudiant ce mémoire, je vois de nouveau que le point de vue que j'adopte diffère beaucoup de celui de M. BOREL; moi, je tiens à définir l'intégrale aussitôt que possible, pendant que M. BOREL commence par introduire la mesure, seulement que, au lieu de la définir suivant M. LEBESGUE, il se sert de la définition qui est son propre et qui conduit aux ensembles moins généraux que l'on appelle, avec M. LEBESGUE, les ensembles mesurables B. Sans doute, cette définition a ses avantages quand il s'agit de certaines questions délicates, comme par exemple dans l'étude approfondie des ensembles de mesure nulle; mais on ne doit pas oublier que le grand mérite de la théorie de M. LEBESGUE consiste en ce qu'elle est compréhensive sans être sensible, c'est à dire qu'elle permet d'ajourner les questions délicates sans les abandonner. À cet égard, je dois préférer une autre idée de M. BOREL (Comptes rendus, le 14 et le 28 février 1910) dont j'ai pris connaissance seulement après la publication de ma note; elle consiste à définir l'intégrale en négligeant d'abord un nombre fini ou une infinité d'inter valles dont la somme est arbitrairement petite et puis faire tendre cette somme vers zéro. Il faut regretter que M. BOREL ne continuait pas de développer cette idée féconde, sur laquelle on pourrait baser effectivement une théorie élémentaire telle que je l'ai en vue; d'ailleurs, l'idée fut reprise récemment avec succès par M. HAHN (Monatsberichte f. Math. u. Phys., 1915, p. 3). Une autre définition où la mesure n'intervient pas, est due à M. PERRON (Sitzungsberichte, Heidelberg 1914, 14. Abh.); il s'y agit d'une généralisation très subtile, mais un peu artificielle, de l'idée classique d'envisager l'intégration comme l'inverse de la différentiation. Enfin, je me rappelle d'avoir parcouru incidemment, pendant mon dernier séjour à Stockholm, une note de M. YOUNG sur le même sujet qui est imprimée dans les Comptes rendus de 1915 ou de 1916; malheureusement, à cause de la guerre, ces Comptes rendus ne se trouvent pas actuellement à notre bibliothèque.

Dans ce qui suit, je ne parlerai que des fonctions d'une seule variable; mais j'ajoute que, pour l'intégration des fonctions à plusieurs variables, on n'aura presque rien à changer.

1.

En commençant d'exposer ma méthode, il me faut tout de suite avouer que je n'ai pas l'intention d'écarter complètement la mesure. Au contraire, le point le plus essentiel de ma méthode consiste à faire intervenir dès le commencement une classe particulière d'ensembles mesurables, savoir les ensembles de *mesure nulle*. On y entend, d'après M. LEBESGUE, les ensembles qu'on peut enfermer dans un nombre fini ou une infinité dénombrable d'intervalles de manière que la somme des longueurs soit arbitrairement petite. Il découle immédiatement de cette définition que *tout ensemble compris dans un ensemble de mesure nulle est lui-même de mesure nulle* et que *la somme d'un nombre fini ou d'une infinité dénombrable de tels ensembles, c'est à dire l'ensemble que l'on obtient en les réunissant, l'est aussi*. En particulier, *chaque ensemble fini ou dénombrable de points est de mesure nulle.*

C'est tout ce qu'il nous faudra savoir sur ces ensembles particuliers.

2.

Envisageons maintenant une infinité dénombrable d'ensembles $E_1, E_2, \ldots$, appartenant à l'intervalle ab et formé chacun d'un nombre fini d'intervalles que l'on peut supposer de ne pas empiéter les uns sur les autres. Désignons par

l_n la longueur totale de l'ensemble E_n — on pourrait aussi l'appeler mesure de E_n — c'est à dire la somme des longueurs des intervalles dont il se compose. De plus, considérons l'ensemble E^* des points communs à une infinité d'entre les ensembles E_n. Je dis que *E^* ne peut être de mesure nulle que si $l_n \to 0$.*

Pour le voir, considérons d'abord le cas particulier où chaque ensemble E_n contient tous les suivants. Dans ce cas, la suite des l_n étant décroissante, elle tendra vers une limite déterminée positive ou zéro, et pour voir que c'est précisément vers zéro qu'elle tend, il suffit de montrer que l'hypothèse que $l_n > \delta > 0$, pour tous les n, aboutit à une contradiction. Or, l'ensemble E^* étant de mesure nulle, il peut être enfermé dans un nombre fini ou une infinité dénombrable d'intervalles tels que leur somme soit inférieure à δ et on pourra aussi s'arranger que chaque point de E^* soit intérieur à un au moins de ces intervalles; en effet, on pourrait, s'il était nécessaire, dilater tous les intervalles sans que leur somme cesse d'être inférieure à δ. Ces intervalles ne recouvriront complètement aucun des ensembles E_n, puisque dans le cas contraire il existerait d'après un théorème bien connu de M. Borel, un nombre fini d'entre eux ayant la même propriété; mais c'est impossible, leur somme étant inférieure à δ et la longueur totale de E_n y étant supérieure. C'est à dire que, en excluant de E_n les points intérieurs aux intervalles où nous avons enfermé l'ensemble E^*, il reste encore un certain ensemble fermé $\bar{E}_n$. Or, les ensembles fermés $\bar{E}_n$ dont chacun contient les suivants, admettent au moins un point compris dans chacun d'eux; ce point étant compris dans tous les E_n, il l'est aussi dans E^*; mais, d'autre part, en définissant les ensembles $\bar{E}_n$, nous avons supprimé tous les points de E^*. Voici la contradiction annoncée.

3.

Allons maintenant au cas général! Si les l_n ne tendaient pas vers zéro, il existerait une infinité d'ensembles E_n dont les longueurs totales l_n surpasseraient une certaine quantité $\delta > 0$. Alors en désignant par E_m^n l'ensemble qu'on obtient en réunissant les ensembles $E_m, E_{m+1}, \ldots, E_n \, (m \leqq n)$ et par l_m^n la longueur totale de cet ensemble, on aurait aussi, pour m quelconque et pour n suffisamment grand $l_m^n > \delta$. C'est à dire que, en désignant par λ_m la limite de la suite croissante l_m^n (m fixe, $n \to \infty$), on aurait $\lambda_m > \delta$ pour tous les m.

Soit ε une quantité positive inférieure à δ, d'ailleurs quelconque, et choisissons successivement les indices $n_1 \leqq n_2 \leqq n_3 \ldots$ de sorte qu'on ait pour tous les m

$$l_m^{n_m} > \lambda_m - \frac{\varepsilon}{2^m}.$$

Désignons par $E^{(m)}$ la partie commune des ensembles $E_1^{n_1}, E_2^{n_2}, \ldots, E_m^{n_m}$ et par $l^{(m)}$ la longueur totale de $E^{(m)}$. Je dis qu'on a

$$l^{(m)} > \delta - \varepsilon$$

et cela pour tous les m. Cette inégalité suit immédiatement, à plus forte raison, de l'inégalité plus précise

$$l^{(m)} > \lambda_m - \varepsilon\left(1 - \frac{1}{2^m}\right)$$

que nous allons vérifier par récurrence. En effet, comme $l^{(1)} = l_1^{n_1}$, l'inégalité est vraie, par hypothèse, pour $m = 1$; supposons qu'elle le soit encore pour $m - 1$; il faut montrer qu'elle subsiste pour m. Or, l'ensemble $E^{(m)}$ est la partie commune de $E^{(m-1)}$ et de $E_m^{n_m}$ pour lesquels on a, par hypothèse,

$$l^{(m-1)} > \lambda_{m-1} - \varepsilon\left(1 - \frac{1}{2^{m-1}}\right), \quad l_m^{n_m} > \lambda_m - \frac{\varepsilon}{2^m},$$

et qui sont tous les deux compris dans l'ensemble $E_{m-1}^{n_m}$ dont la longueur totale ne dépasse pas la limite λ_{m-1}. Par conséquent on a

$$l^{(m)} > l^{(m-1)} + l_m^{n_m} - \lambda_{m-1} > \lambda_{m-1} - \varepsilon\left(1 - \frac{1}{2^{m-1}}\right) + \lambda_m - \frac{\varepsilon}{2^m} - \lambda_m = \lambda_m - \varepsilon\left(1 - \frac{1}{2^m}\right),$$

ce qu'il fallait montrer.

L'inégalité que nous venons de démontrer, indique que les longueurs totales $l^{(m)}$ des ensembles $E^{(m)}$ dont chacun contient les suivants, restent supérieures à une quantité $\delta' = \delta - \varepsilon > 0$; d'autre part, l'ensemble des points communs à tous les $E^{(m)}$ étant compris dans E^*, sera de mesure nulle, mais ce sont précisément les hypothèses correspondant au cas particulier déjà traité et qui nous ont conduit à une contradiction.

4.

Le résultat que nous venons d'obtenir, suffit complètement pour introduire l'intégrale, et c'est seulement pour éviter double emploi que je fais déjà ici une remarque qui nous sera essentielle dans la suite; il s'y agit de ce que j'ai démontré, en réalité, un peu plus que le fait annoncé. En effet, après avoir déduit de l'hypothèse que l_n ne tend pas vers zéro, l'inégalité $\lambda_m > \delta$, j'ai fait voir comment cette inégalité aboutit à une contradiction, sans qu'il nous aurait fallu retourner à l'hypothèse faite. C'est à dire que nous avons démontré, en réalité, que l'hypothèse que l'ensemble E^* soit de mesure nulle, entraîne la relation

$$\lambda_m \to 0 \qquad (m \to \infty),$$

ou bien, ce qui revient au même, la relation

$$l_m^n \to 0 \qquad (m \to \infty,\ n \to \infty);$$

cette relation comprend, pour $m = n$, la relation $l_n \to 0$. (D'ailleurs, on aurait pu aussi déduire la relation $l_m^n \to 0$ de cette dernière relation bien qu'elle ait une forme plus particulière; en effet, l'hypothèse faite sur la suite $\{E_n\}$ est également remplie pour toutes les suites $\{E_m^n\}$.)

Ces résultats s'étendent presque immédiatement à un tableau infini à double entrée $\{E_{ik}\}$; on aura seulement à préciser le sens qu'il faudra donner aux notations E^*, E_m^n, l_m^n et λ_m. Quant à l'ensemble E^*, convenons qu'il soit formé des points appartenant à une infinité d'entre les ensembles E_{ik}, en exigeant de plus que *tous les deux indices* croissent indéfiniment. C'est à dire que les points de E^* sont caractérisés par le fait que pour m quelconque et pour chaque point appartenant à E^*, il existe des ensembles E_{ik} avec $i \geqq m$, $k \geqq m$ qui le contiennent. Par E_m^n nous entendrons l'ensemble formé en réunissant tous les ensembles E_{ik} $(i = m, \ldots, n;\ k = m, \ldots, n)$; l_m^n est la longueur totale de E_m^n et λ_m désigne la limite de l_m^n quand n va à l'infini. Ces conventions faites, les raisonnements du numéro précédent s'appliquent entièrement au cas considéré et fournissent des résultats analogues à ceux que nous venons d'obtenir.

Je vais énoncer nos résultats sous la forme que j'en aurai besoin, la relation particulière $l_n \to 0$ pour le cas d'une suite et la relation plus générale $l_m^n \to 0$ pour le cas du tableau.

Lemme particulier: *Étant donnée une infinité dénombrable d'ensembles E_1, $E_2, \ldots$, compris dans le même intervalle et formé chacun d'un nombre fini d'intervalles, soient $l_1, l_2, \ldots$ leurs longueurs totales et soit E^* l'ensemble des points appartenant à une infinité d'entre eux. Si l'ensemble E^* est de mesure nulle, on a, n allant à l'infini:*

$$l_n \to 0.$$

Lemme général: *Étant donné un tableau à double entrée $\{E_{ik}\}$ d'ensembles du type considéré, soit E^* l'ensemble des points tels que, pour chaque point et pour chaque nombre m, il existe des ensembles E_{ik} avec $i \geqq m$, $k \geqq m$ qui le contiennent; de plus, soit E_m^n l'ensemble formé en réunissant les ensembles E_{ik} $(i, k = m, m+1, \ldots, n)$*

et soit l_m^n *sa longueur totale. Si l'ensemble* E^* *est de mesure nulle, on a, m et n allant à l'infini:*

$$l_m^n \to 0.$$

Bien entendu, ces deux lemmes sont compris dans des théorèmes connus appartenant à la théorie générale de la mesure; ainsi, par exemple, le premier est compris dans un théorème de M. BOREL (Comptes rendus, le 17 décembre 1903), mais il peut aussi être regardé comme cas particulier du théorème fondamental de M. LEBESGUE concernant l'intégration terme à terme des suites bornées. Pour nous, il s'agissait seulement d'établir ces lemmes indépendamment de la théorie générale. D'ailleurs, quant au premier lemme, il fut déjà établi avant LEBESGUE par ARZELÀ pour le cas particulier où il n'y a aucun point appartenant à l'ensemble E^* (Memorie Ist. Bologna 1899, p. 131).

5.

Pour définir l'intégrale, je commence par les fonctions particulières appelées *fonctions simples* et définies comme il suit: on partage l'intervalle $ab\,(a<b)$ en un nombre fini de segments et on fait correspondre à chaque segment une quantité constante; quant aux points de division, on y peut attacher des valeurs quelconques; d'ailleurs, ces dernières n'importent pas au point de vue de l'intégration. Pour ces fonctions simples, on définira l'intégrale comme d'ordinaire en ajoutant les produits de chaque segment et de la valeur correspondante.

Soit maintenant $\{\varphi_n(x)\}$ *une suite infinie de fonctions simples, bornées dans leur ensemble et tendant presque partout, c'est-à-dire sauf peut-être pour un ensemble de mesure nulle, vers une fonction limite* $f(x)$. *Je dis que les intégrales des* $\varphi_n(x)$ *tendent vers une valeur limite et que cette limite est toujours la même pour deux suites tendant vers la même fonction* $f(x)$. Pour le voir, considérons d'abord le cas particulier où $f(x)=0$ sur tout l'intervalle. Désignons par E_n l'ensemble formé par les intervalles où $|\varphi_n(x)| \geqq \varepsilon$, ε étant une quantité positive arbitrairement choisie, et soit l_n la longueur totale de cet ensemble. Les fonctions $\varphi_n(x)$ étant bornées dans leur ensemble, leur module reste inférieure à une constante G. Alors le module de l'intégrale de $\varphi_n(x)$ sera inférieure à

$$l_n G + \varepsilon (b-a),$$

et comme les hypothèses de notre lemme particulier sont remplies pour les ensembles E_n, la quantité l_n tend vers zéro. C'est-à-dire que, pour n suffisamment grand, l'intégrale de $\varphi_n(x)$ sera, en valeur absolue, inférieure à la quantité arbitrairement petite $\varepsilon(b-a)$. Donc l'intégrale tend vers zéro avec $\frac{1}{n}$.

Le cas général se ramène immédiatement au cas particulier que nous venons de considérer, et celà en remarquant que, dans le cas général, ce sont les suites $\{\varphi_m(x)-\varphi_n(x)\}$ qui tendent presque partout vers zéro; m et n y peuvent aller à l'infini de toute sorte possible, indépendamment l'un de l'autre. On en conclut, en appliquant le résultat particulier établi, que la différence des intégrales de φ_m et de φ_n tend aussi vers zéro et il s'en suit, d'après le critère général de CAUCHY, que l'intégrale de φ_n tend vers une limite déterminée. Quand il s'agit de deux suites distinctes $\{\varphi_n\}$ et $\{\varphi'_n\}$ tendant presque partout vers la même fonction $f(x)$, on n'aura qu'à appliquer le même raisonnement à la suite $\{\varphi_n-\varphi'_n\}$; on en conclut que les intégrales des φ_n et des φ'_n tendent vers la même limite.

Après le résultat que nous venons d'établir, la définition de l'intégrale est immédiate. *Appelons fonction sommable toute fonction $f(x)$ du type considéré, c'est-à-dire qui est la limite — presque partout — d'une suite bornée de fonctions simples $\varphi_n(x)$. Comme valeur d'intégrale, nous y attachons la limite de l'intégrale des $\varphi_n(x)$; d'après ce que nous venons de voir, cette limite existe et elle est parfaitement déterminée par la fonction $f(x)$, indépendamment du choix particulier de la suite approchante.*

6.

On pourrait être tenté d'appliquer le même procédé une seconde fois, en espérant de définir ainsi l'intégrale pour une classe de fonctions plus large. Je dis que le procédé ne fournira rien de nouveau. D'une façon précise, je dis que *la fonction limite d'une suite bornée de fonctions sommables qui converge presque partout, est elle-même une fonction sommable et que, de plus, la suite peut être intégrée terme à terme.*

Démontrons d'abord le lemme suivant, qui n'est qu'un cas particulier d'un théorème important de M. EGOROFF (Comptes rendus, le 30 janvier 1911): *Si une suite $\{\varphi_n(x)\}$ de fonctions simples converge presque partout dans l'intervalle ab, on peut enlever de cet intervalle un système (fini ou dénombrable) d'intervalles tels que leur somme soit aussi petite que l'on veut, et que d'autre part, la convergence devienne uniforme sur l'ensemble qui reste.*

Pour le voir, soient ε une quantité positive arbitrairement petite et p un nombre entier positif, d'ailleurs quelconque, et désignons par E_{ik} l'ensemble où $|\varphi_i(x)-\varphi_k(x)|\geqq\frac{1}{p}\cdot$ Les φ tendant presque partout vers une limite déterminée, l'ensemble E^* correspondant au tableau $\{E_{ik}\}$ sera de mesure nulle; car les points où les φ convergent n'y appartiennent pas. Il s'en suit, d'après notre lemme général, que $l_m^n \to 0$; donc il y a un premier indice m_p tel qu'on ait $l_m^n < \frac{\varepsilon}{2^p}$ pour $m = m_p$ et pour tous les $n \geqq m_p$. Considérons le système d'intervalles que l'on définira en prenant d'abord les intervalles dont se compose l'ensemble $E_m^m = E_{mm}$, puis les intervalles formant l'ensemble $E_m^{m+1} - E_m^m$, puis ceux formant l'ensemble $E_m^{m+2} - E_m^{m+1}$ et ainsi de suite. En prenant un nombre fini quelconque de ces intervalles, ils seront compris, pour n suffisamment grand, dans l'ensemble E_m^n et comme ils n'empiètent pas les uns sur les autres, leur somme ne surpassera pas la longueur totale l_m^n de cet ensemble. C'est-à-dire que, pour $m = m_p$, la somme de tous les intervalles choisis sera inférieure à $\frac{\varepsilon}{2^p}$. D'autre part, ces intervalles renferment tous les ensembles E_{ik} tels que $i \geqq m_p$, $k \geqq m_p$. Faisons maintenant parcourir à p tous les entiers positifs, en gardant la même quantité donnée ε; la somme des intervalles correspondants sera inférieure à $\frac{\varepsilon}{2} + \frac{\varepsilon}{4} + \cdots = \varepsilon$. Excluons tous ces intervalles; sur l'ensemble qui reste, la convergence des φ_n vers leur limite sera uniforme. En effet, on y a $|\varphi_i(x) - \varphi_k(x)| < \frac{1}{p}$ pour tous les $i, k \geqq m_p$.

Voilà une conséquence immédiate du lemme établi. Soit $\{\varphi_n(x)\}$ une suite bornée de fonctions simples, convergeant presque partout vers la fonction $f(x)$, et soit ε une quantité positive arbitrairement donnée. Sans restreindre la généralité, on peut admettre que les φ_n soient comprises entre les bornes supérieure et inférieure de $f(x)$; s'il n'en était pas ainsi, on remplacerait la valeur de $\varphi_n(x)$, partout où elle surpasse la borne supérieure de $f(x)$, par cette borne elle-même et on ferait la modification analogue pour la borne inférieure. Cela étant, enlevons de l'intervalle ab un système d'intervalles dont la somme $< \varepsilon$ et tels que la suite converge uniformément sur l'ensemble qui reste. Alors on aura sur cet ensemble et pour n suffisamment grand $|f(x) - \varphi_n(x)| < \varepsilon$. D'autre part, comme il est permis d'intégrer terme à terme, la différence des intégrales de f et de φ_n sera aussi, pour n suffisamment grand, aussi petite qu'on voudra, par exemple $< \varepsilon$. C'est-à-dire *qu'on peut, pour chaque fonction sommable $f(x)$, choisir une fonction simple $\varphi(x)$ de sorte qu'elle soit comprise entre les deux bornes de $f(x)$, que de plus la différence des intégrales de $f(x)$ et de $\varphi(x)$ soit inférieure à ε et enfin que la différence elle-même des deux fonctions soit inférieure à ε, sauf peut-être pour un ensemble de points que l'on peut enfermer dans des intervalles dont la somme est inférieure à ε.*

Pour démontrer maintenant le théorème annoncé, soit $\{f_n(x)\}$ une suite bornée de fonctions sommables, tendant presque partout vers la fonction $g(x)$.

Faisons correspondre à chaque fonction $f_n(x)$ une fonction simple $\varphi_n(x)$ d'après la loi que nous venons d'établir, en y posant $\varepsilon = \frac{1}{2^n}$. Je dis qu'on a, presque partout, $f_n(x) - \varphi_n(x) \to 0$. En effet on peut, pour chaque indice n, enlever de l'intervalle ab un système d'intervalles dont la somme $< \frac{1}{2^n}$, de sorte que, sur l'ensemble qui reste, on ait $|f_n(x) - \varphi_n(x)| < \frac{1}{2^n}$. Enlevons les intervalles correspondant à $n = k+1, k+2, \ldots$; leur somme sera inférieure à $\frac{1}{2^{k+1}} + \frac{1}{2^{k+2}} + \cdots = \frac{1}{2^k}$. Sur l'ensemble qui reste on aura pour tous les n, à partir de $n = k+1$, $|f_n(x) - \varphi_n(x)| < \frac{1}{2^n}$, par conséquent, $f_n(x) - \varphi_n(x)$ y tend vers zéro. C'est-à-dire que les points où cette différence ne tend pas vers zéro, peuvent être enfermés dans des intervalles dont la somme est inférieure à $\frac{1}{2^k}$ et comme k est arbitraire cette somme peut être rendue arbitrairement petite. Par conséquent, ces points d'exception, s'il en existe, forment un ensemble de mesure nulle. Il s'en suit que les φ_n tendent presque partout vers la même limite que les f_n, savoir vers $g(x)$. De plus, les $\varphi_n(x)$ restant par hypothèse entre les mêmes bornes que les $f_n(x)$, ils sont bornés dans leur ensemble. Donc $g(x)$ est la limite, presque partout, d'une suite bornée de fonctions simples; elle est par conséquent sommable. Enfin, son intégrale est la limite des intégrales des $\varphi_n(x)$, ou bien, ces intégrales étant égales, à $\frac{1}{2^n}$ près, à celles des $f_n(x)$, ces dernières intégrales tendent aussi vers l'intégrale de $g(x)$. Le théorème annoncé est donc démontré.

7.

Jusqu'ici il s'agissait toujours des fonctions sommables bornées; pour arriver aux fonctions sommables *non bornées*, on pourrait, au lieu d'exiger de la suite $\{\varphi_n(x)\}$ d'être bornée, faire quelque autre hypothèse plus générale; en tout cas, il faudrait choisir cette hypothèse de sorte qu'on puisse en conclure la convergence des intégrales. Voilà une telle hypothèse: supposons que l'intégrale de $\varphi_n(x)$, prise sur un nombre fini d'intervalles dont la longueur totale est suffisamment petite, soit elle-même aussi petite que l'on voudrait et cela indépendamment de n. Sous cette hypothèse, les raisonnements faits pour les suites bornées s'appliquent presque sans modification et conduisent précisément à toutes les fonctions sommables au sens de LEBESGUE. Cependant, du moins à première vue, l'hypothèse proposée pourrait paraître un peu artificielle, et il me semble plus naturel d'aller une voie qu'on a suivi déjà bien des fois quand il s'agissait d'étendre l'intégrale à des fonctions non bornées. Tout d'abord, introduisons les *fonctions mesurables;* nous appelons ainsi chaque fonction $f(x)$, bornée ou non, qui est la limite — presque partout — d'une suite, bornée ou non, de fonctions simples. *Lorsque la fonction mesurable $f(x)$ est bornée, elle est aussi sommable;* en effet, si la suite correspondante $\{\varphi_n(x)\}$ n'était pas bornée, on pourrait la modifier, comme nous venons de voir, en remplaçant la valeur de $\varphi_n(x)$, partout où elle ne serait pas compris entre les deux bornes de $f(x)$, par la borne plus proche. Donc chaque fonction mesurable et bornée est sommable; inversement, à plus forte raison, chaque fonction bornée et sommable est mesurable. Lorsque la fonction mesurable $f(x)$ n'est pas bornée, désignons par $f(x; c, d)$, où $c \leqq d$, la fonction égale à $f(x)$ partout où $c \leqq f(x) \leqq d$, égale à c partout où $f(x) \leqq c$ et à d partout où $f(x) \geqq d$. Soit $\{\varphi_n(x)\}$ une suite de fonctions simples tendant presque partout vers $f(x)$; alors les fonctions correspondantes $\varphi_n(x; c, d)$ formeront une suite bornée de fonctions simples qui tend presque partout vers $f(x; c, d)$. Par conséquent, la fonction bornée $f(x; c, d)$ sera sommable. Cela étant, faisons tendre c vers $-\infty$ et d vers $+\infty$ et cela de toutes les manières possibles; si alors l'intégrale de $f(x; c, d)$ tend toujours vers la même limite finie déterminée, nous attacherons cette limite comme valeur d'intégrale à la fonction $f(x)$ et nous convenons de dire que la fonction non bornée $f(x)$ est sommable. On voit aisément qu'une condition nécessaire et suffisante pour qu'il en soit ainsi, consiste en ce que l'intégrale de $f(x; c, d)$ soit une fonction bornée des paramêtres c et d.

Posons en particulier $c = 0$ ou $d = 0$; il s'en suit que *la partie positive et la partie négative d'une fonction sommable sont aussi sommables et que, par conséquent, le module $|f(x)|$ l'est aussi; inversement, lorsque pour une fonction mesurable $f(x)$, le module $|f(x)|$ est sommable, la fonction elle-même est aussi sommable.*

Il sera bien de fois utile de voir si une fonction est mesurable, avant de s'occuper de la question si elle est sommable. On a évidemment les règles suivantes qui découlent immédiatement de la définition: la somme, la différence, le produit et dans le cas où le diviseur ne s'annule pas, le quotient de deux fonctions mesurables le sont également. Mais il y a plus: *la limite d'une suite de fonctions mesurables qui converge presque partout, sera elle-même une fonction mesurable.* Ce fait ressort des raisonnements du numéro précédent qui se simplifient même, puisque maintenant n'interviennent plus les hypothèses dont on avait besoin pour intégrer terme à terme.

8.

Les propriétés principales des fonctions sommables découlent, grâce aux définitions données, presque immédiatement, par passage à la limite, des propriétés analogues des fonctions simples; permettez que je n'entre pas dans les détails. Je me contente d'indiquer que, entre autres, la somme et la différence de deux fonctions sommables le seront aussi et que leurs intégrales s'obtiennent par addition et par soustraction des intégrales relatives aux deux fonctions, que de plus le produit de deux fonctions sommables dont l'une est au moins bornée, est aussi sommable, enfin que $f(x)$ étant sommable dans deux des intervalles ab, bc, ac $(a < b < c)$, elle le sera aussi dans le troisième et que l'intégrale prise sur ac est la somme des deux autres. Cela suffit pour pouvoir dire quelques mots sur *la mesure des ensembles.* Vous savez certainement comment on peut définir la mesure après avoir introduit l'intégrale: soit E un ensemble et soit $f(x)$ la fonction égale à 1 pour les points x appartenant à cet ensemble et égale à 0 ailleurs. Si la fonction $f(x)$ est mesurable et par conséquent sommable dans un intervalle quelconque contenant E, elle l'est évidemment dans chaque intervalle et son intégrale est la même sur tous les intervalles qui contiennent E. Alors nous dirons que l'ensemble E est mesurable et nous y attacherons comme mesure $m(E)$ la valeur de l'intégrale considérée. À chaque ensemble mesurable correspond de cette façon une certaine fonction sommable caractéristique ne prenant que les valeurs 0 et 1, et en appliquant à ces fonctions les résultats trouvés, on aura les théorèmes principaux sur la mesure. Ainsi par exemple, désignons par $f(x)$, $f_n(x)$ les fonctions caractéristiques qui correspondent aux ensembles E, E_n et pour fixer les idées, supposons que tous ces ensembles soient compris dans l'intervalle ab. À l'ensemble complémentaire $C(E)$ de E par rapport à l'intervalle ab correspond une fonction égale à $1-f(x)$ sur cet intervalle et égale à 0 ailleurs; donc $C(E)$ *est mesurable et on a* $m(E)+m(C(E))=b-a$. À la partie commune des ensembles $E_1, \ldots, E_n$ correspond le produit $f_1(x)\ldots f_n(x)$; à la partie commune d'une infinité dénombrable d'ensembles $E_1, E_2, \ldots$ correspond le produit infini $f_1(x)f_2(x)\ldots$; or, produit et limite de fonctions mesurables étant elles-mêmes mesurables, *les ensembles considérés sont mesurables.* Dans le cas particulier où chaque ensemble E_n contient les suivants, la valeur du produit infini n'est autre que la limite de $f_n(x)$, alors son intégrale sera la limite de celle de $f_n(x)$, c'est à dire qu'elle sera la limite de $m(E_n)$; donc *dans ce cas, la mesure de la partie commune est fournie par la limite de* $m(E_n)$. D'autre part, réunissons un nombre fini ou une infinité dénombrable d'ensembles $E_1, E_2, \ldots$; quant à l'ensemble qui vient, il y correspond la fonction

$1-(1-f_1(x))(1-f_2(x))\ldots$, et cette expression se réduit à $f_1(x)+f_2(x)+\cdots$ dans le cas particulier où le produit de deux fonctions $f_n(x)$ à indices différents s'annule, c'est-à-dire lorsque les ensembles sont sans points communs; par conséquent, *l'ensemble-somme est mesurable et de plus, dans le cas particulier où il n'existe pas des points communs, on a* $m(E_1+E_2+\cdots)=m(E_1)+m(E_2)+\cdots$.

Entre les ensembles et les fonctions mesurables, il y a encore une relation très intime qui joue un rôle important dans la théorie de M. LEBESGUE; elle consiste en ce que, *pour* $f(x)$ *mesurable et pour* c *quelconque, les ensembles où* $f(x)>c$, $f(x)\geqq c$, $f(x)<c$, $f(x)\leqq c$, $f(x)=c$, *sont mesurables et que, inversement, lorsqu'il en est ainsi, la fonction* $f(x)$ *est mesurable.* Qu'il nous suffise de prouver la première partie de cette assertion, sans insister sur la réciproque. Alors on pourra se borner évidemment, sans restreindre la généralité, au cas où c est positif et à l'ensemble pour lequel $f(x)\geqq c$. Or, la fonction caractéristique de cet ensemble est la limite de la suite des fonctions $\frac{1}{c^n}f_n(x;0,c)$ $(n\to\infty)$; ces fonctions étant mesurables, il en sera de même quant à leur limite.

9.

On définira d'une manière tout analogue l'intégrale d'une fonction $f(x)$ sur un ensemble E, n'importe que $f(x)$ soit défini seulement dans E ou bien sur un intervalle ou dans un ensemble quelconque contenant E. À ce but, envisageons la fonction $f(x)$ égale à $g(x)$ sur E et s'annulant ailleurs; si elle est sommable dans un des intervalles contenant E, elle le sera aussi dans tous ces intervalles et son intégrale ne dépend pas du choix particulier de l'intervalle. Lorsqu'il en est ainsi, nous dirons que $f(x)$ est sommable dans E et nous y attacherons comme intégrale l'intégrale de $f_1(x)$ prise sur l'un quelconque des intervalles considérés. Cette définition semble de ne faire aucune hypothèse à l'égard de l'ensemble E; observons qu'on ne restreindra pas la généralité en supposant E mesurable; en effet, même si l'ensemble E lui-même n'était pas mesurable, la partie où $f(x)\neq 0$ le sera certainement puisqu'elle se confonda vec l'ensemble où la fonction sommable $f_1(x)\neq 0$.

Lorsque $f(x)$ *est sommable dans* E, *elle l'est aussi dans chaque sous-ensemble mesurable.* En effet, désignons par $g(x)$ la fonction caractéristique correspondant au sous-ensemble; alors le passage de l'ensemble entier au sous-ensemble revient à multiplier la fonction sommable $f_1(x)$ par la fonction sommable et bornée $g(x)$, ce qui ne touche pas la sommabilité. En particulier, lorsque $f(x)$ est sommable dans l'intervalle ab, il l'est aussi dans chaque sous-ensemble mesurable.

Au point de vue que nous venons d'adopter, les considérations du numéro précédent peuvent être regardées comme correspondant à la fonction $f(x) = 1$, en effet, la mesure d'un ensemble n'est autre que l'intégrale de cette constante prise sur l'ensemble. On se démande, si les résultats acquis au numéro précédent subsistent encore, lorsqu'on y remplace la mesure des ensembles qui interviennent, par l'intégrale d'une fonction sommable quelconque par rapport à ces ensembles. La réponse sera affirmative, comme on voit immédiatement pour les fonctions bornées, pour lesquelles on n'aura presque rien à changer dans les raisonnements; pour les fonctions non bornées, tout revient à observer que lorsqu'on remplace $f(x)$ par la fonction bornée $f(x; c, d)$ et qu'on choisit convenablement les quantités c et d, l'erreur commise relative aux intégrales considérées sera, pour toutes ces intégrales, aussi petite que l'on voudra; bien entendu, on suppose que la fonction soit sommable dans un ensemble E comprenant tous les ensembles considérés. En effet, supposons d'abord $f(x)$ positif; dans ce cas, l'erreur commise en remplaçant $f(x)$ par $f(x; 0, d)$ ne surpassera pas, pour aucun des ensembles, l'erreur relative à l'ensemble E. Or, le cas général se ramène au cas considéré en décomposant $f(x)$ en ses deux parties positive et négative. Des résultats que l'on obtient de cette manière, énonçons le suivant: *Lorsque $E_1, E_2, \ldots$ sont des ensembles mesurables sans points communs et que de plus, $f(x)$ est sommable sur l'ensemble $E_1 + E_2 + \cdots$, l'intégrale par rapport à cet ensemble-somme vient en ajoutant les intégrales relatives aux ensembles E_n.* Dans le même ordre d'idées, on arrivera aussi au théorème suivant qui est un des plus utiles dans la théorie de M. LEBESGUE: *Lorsque l'ensemble E varie de sorte que $m(E)$ tende vers zéro, il en est de même pour l'intégrale sur E de toute fonction $f(x)$ qui est sommable dans un ensemble contenant E.*

10.

Je terminerai ces lignes en faisant une remarque presque évidente relative à *l'intégrale au sens de* RIEMANN. Vous avez certainement observé que la définition de l'intégrale dont je viens de me servir, est quelque sorte de généralisation de celle de RIEMANN. En effet, l'expression

$$\sum_{k=0}^{m-1} f(\xi_k)(x_{k+1} - x_k)$$

qui intervient chez RIEMANN, peut être envisagée comme donnant l'intégrale de la fonction *simple* $\varphi_n(x)$ dont les valeurs constantes sur les intervalles (x_k, x_{k+1})

sont choisies égales à l'une des valeurs prises par $f(x)$ sur le même intervalle. Quant aux points x_k, on y peut supposer, pour plus de simplicité, $\varphi_n(x_k) = f(x_k)$. Lorsqu'on passe à la limite, en faisant tendre vers zéro le plus grand des intervalles, ces fonctions simples $\varphi_n(x)$ tendent certainement vers $f(x)$ partout où $f(x)$ est continu. D'autre part, pour les fonctions intégrables au sens de RIEMANN, les points de discontinuité forment un ensemble de mesure nulle; donc il vient que les fonctions $\varphi(x)$ considérées tendent presque partout vers $f(x)$. Mais il y a plus; pour tout point de continuité x_0, la convergence des $\varphi_n(x)$ vers $f(x)$ est *uniforme aux environs de* x_0. Cela veut dire, avec M. PRINGSHEIM, que, étant donnés le point x_0 et la quantité positive ε arbitrairement petite, il y correspond un certain voisinage de x_0 et un indice ν de sorte que les inégalités $|f(x) - \varphi_n(x)| < \varepsilon$ soient vérifiées pour les $n \geqq \nu$ et pour tous les points x appartenant au voisinage; ou ce qui revient au même, la relation $x_n \to x_0$ entraîne toujours $\varphi_n(x_n) \to f(x_0)$.

Par conséquent, *toute fonction intégrable au sens de* RIEMANN *peut être approchée par une suite bornée de fonctions simples convergeant presque partout et cela de sorte, que la convergence soit uniforme presque partout, c'est-à-dire aux environs de chaque point x, sauf peut-être pour un ensemble de mesure nulle.* (Il s'en suit que la convergence est uniforme sur chaque ensemble fermé qui reste après avoir exclu les points d'exception par un système d'intervalles; mais il n'est pas permis d'en conclure que la convergence doit être uniforme sur tout l'ensemble complémentaire des points d'exception; cela revient à ce que, en général, l'ensemble complémentaire ne sera pas fermé.)

La réciproque est vraie: *Lorsque une fonction $f(x)$ est, presque partout, égale à la limite d'une suite bornée de fonctions simples $\varphi_n(x)$ et si, de plus, la convergence est uniforme presque partout, la fonction $f(x)$ est intégrable au sens de* RIEMANN. En effet, de l'hypothèse faite il vient que la fonction $f(x)$ est bornée et que ses points de discontinuité forment un ensemble de mesure nulle, puisque $f(x)$ est sûrement continu en chaque point x_0 où la convergence est uniforme et les fonctions $\varphi_n(x)$ sont continues. C'est-à-dire que tous les points de discontinuité de $f(x)$ se trouvent parmi les points 1) où la suite ne converge pas (ensemble de mesure nulle), 2) où la convergence n'est pas uniforme (ensemble de mesure nulle), 3) les points de discontinuité des $\varphi_n(x)$ (ensemble dénombrable). Or, les propriétés caractérisant les fonctions intégrables d'après RIEMANN sont précisément d'être bornées et d'être continues presque partout.

Je profite de l'occasion pour faire quelques indications concernant l'origine de cette notion de convergence uniforme aux environs d'un point, dont vous venez de voir une application qui en prouve l'utilité. Je me l'ai formée, sans en chercher l'origine, en étudiant l'approximation d'une fonction par des polynomes (Jahresbericht d. deutschen Math.-V., 1908, p. 199); presque en même temps M. YOUNG, dans un travail où il analysait cette notion dont il s'est servi déjà auparavant, déclara de ne pas être sûr où elle puisse se trouver pour la première fois (Proceedings of the London Math. Soc., ser. 2, vol. 6, p. 29). Le rapporteur de ces travaux pour les Fortschritte der Mathematik, M. FABER, nous reproche (1908, p. 468, 472) de ne pas savoir que la notion se trouve déjà chez WEIERSTRASS (Œuvres 2, p. 203); seulement, au lieu d'examiner soigneusement le texte original de WEIERSTRASS, il se contente d'invoquer le témoignage de M. PRINGSHEIM (Encyklop. d. math. Wiss., II[1], p. 33, 34; édition française, p. 68). Il est vrai que M. PRINGSHEIM qui d'ailleurs s'est servi de cette notion déjà en 1894 (Math. Ann., 44, p. 64, 65, 80), en fait noblement cadeau à WEIERSTRASS, mais en regardant le texte cité de WEIERSTRASS, vous verrez qu'il s'y agit seulement de la convergence uniforme au sens ordinaire; la seule différence consiste à ne pas exiger cette uniformité à la fois pour le domaine entier que l'on envisage, mais exclusivement pour un domaine suffisamment petit entourant le point x_0. C'est-à-dire que la définition de WEIERSTRASS exige seulement qu'il existe un certain voisinage de x_0, déterminé une fois pour toutes, de sorte que, à chaque quantité positive ε, on puisse faire correspondre un nombre ν tel que les inégalités $|f(x) - f_n(x)| < \varepsilon$ soient vérifiées pour tous les $n \geqq \nu$ et pour tous les points x appartenant au voisinage considéré. Au contraire, dans la définition que nous employons, comme le fait observer aussi l'édition française de l'Encyclopédie, le choix du voisinage dépend en général de la quantité ε, de sorte qu'il pourra se réduire, pour $\varepsilon \to 0$, au point x_0.

Il s'en suit que, jusqu'à d'autres indications, c'est M. PRINGSHEIM lui-même qu'on devra honorer comme l'auteur de cette notion utile.

Kolozsvár, le 19 février 1917.