Variation Bornée et Dirichlet-Jordan

Les hypothèses du théorème de Dirichlet sur la convergence de la série de Fourier d'une fonction (à valeurs réelles) 2π-périodique f sont que celle-ci est continue par morceaux et ne possède sur $[0, 2\pi]$ qu'un nombre fini de maxima ou de minima (locaux, et stricts). Bref, elle est monotone-continue par morceaux. Mais cette condition n'est pas stable par combinaison linéaire : par exemple une fonction C^1 peut sûrement avoir une infinité de (stricts) maxima et minima locaux ($f(x) = x^3 \sin(\frac{1}{x})$) et cependant elle est toujours la somme d'une fonction continue croissante ($\int^x \max(f'(t), 0)\,dt$) et d'une fonction continue décroissante ($\int^x \min(f'(t), 0)\,dt$). La notion de << fonction de variation bornée >>, sur un intervalle $[a, b]$, fut introduite par Camille Jordan ; elle englobe à la fois les conditions de Dirichlet et les fonctions de classe C^1 par morceaux (que l'on associe en général dans les livres de premier cycle avec le théorème de Dirichlet). De plus c'est une notion extrêmement naturelle en théorie de l'intégration puisque comme nous le verrons elle caractérise les fonctions qui (à peu de choses près) sont de la forme $x \longmapsto \int_{[a,x]} d\mu(u)$ avec μ une *mesure complexe* sur $[a, b]$ au sens de l'annexe précédente.

Soit f une fonction (à valeurs complexes) sur $[a, b]$. À tout choix d'au moins deux points $a \leq x_0 \leq x_1 \leq \cdots \leq x_N \leq b$ on peut associer $V((x_j); f) = |f(x_0) - f(x_1)| + \cdots + |f(x_{N-1}) - f(x_N)|$. Et on définit $V([a, b]; f)$ comme le supremum de toutes ces sommes finies : c'est la << variation totale >> de f sur l'intervalle $[a, b]$ (si $b = a$, posons $V([a, a]; f) = 0$; si $b > a$ on peut toujours imposer $x_i < x_{i+1}$ pour tout i dans les subdivisions). On dit que f est de variation bornée sur $[a, b]$ si $V([a, b]; f) < \infty$. Des définitions analogues imposant $x_0 > a$ ou $x_N < b$ mènent à $V(]a, b]; f)$ ou $V([a, b[; f)$ ou $V(]a, b[; f)$.

Il est évident que pour $a \leq b \leq c$ on a $V([a, c]; f) = V([a, b]; f) + V([b, c]; f)$. Attention cependant à ne pas commettre d'erreur du style $V([a, c[; f) = V([a, b[; f) + V([b, c[; f)$. La formule correcte, pour $b < c$, est $V([a, c[; f) = V([a, b]; f) + V([b, c[; f)$.

Un exemple instructif est celui d'une fonction, non nécessairement continue, mais de classe C^1 par morceaux. Vous prouverez que la variation totale est la somme d'une part de tous les (modules des) << (demi)-sauts >> aux discontinuités de f, et d'autre part de $\int_a^b |f'(t)|dt$, la fonction continue par morceaux f' étant définie à l'exception d'un nombre fini de points (à propos, quelle différence entre $f(x)$ et $\int_a^x f'(t)dt$?).

Soit donc f de variation bornée sur $[a, b]$. Remarquez en passant que f est alors bornée. De plus je vais montrer que f admet une limite à droite $f(a^+)$ en a (un raisonnement analogue montre que f a en tout point une limite à droite et une à gauche). En effet comme $V([a,b];f) < \infty$ on a aussi $V(]a,b];f) < \infty$. Soit (x_j) avec $V((x_j);f) \geq V(]a,b];f) - \epsilon$ (et $x_0 > a$). Soit $y \in]a,x_0[$. En considérant la subdivision $y_0 = y$, $y_1 = x_0$, ..., on obtient $|f(y) - f(x_0| + V((x_j);f) \leq V(]a,b];f)$. Donc $|f(y) - f(x_0)| \leq \epsilon$. Donc $|f(y) - f(z)| \leq 2\epsilon$ pour $a < y, z \leq x_0$ et le critère de Cauchy est vérifié pour l'existence de $f(a^+)$.

Je vous laisse maintenant la formule suivante comme un exercice assez facile : $V([a,b];f) = |f(a) - f(a^+)| + V(]a,b];f)$. Plus généralement

$$V([x,y];f) = |f(x) - f(x^+)| + V(]x,y[;f) + |f(y^-) - f(y)|$$

J'aurai aussi besoin plus tard de la propriété suivante :

$$\lim_{y \to x^+} V(]x,y];f) = 0$$

La technique de preuve est semblable à celle pour établir l'existence de $f(a^+)$. Soit $\epsilon > 0$ et soit (x_j) une subdivision de $]x,b]$ avec $x_0 > x$ et $V((x_j);f) \geq V(]x,b];f) - \epsilon$. Supposons maintenant $x < y \leq x_0$, et considérons une subdivision quelconque (y_k) de l'intervalle $]x,y]$. En mettant les (y_k) ensemble avec les (x_j) et en considérant la variation associée, on obtient la majoration $V((y_k);f) \leq \epsilon$. Donc $V(]x,y];f) \leq \epsilon$. Ainsi $\lim_{y \to x} V(]x,y];f) = 0$ est établi et on en déduit aussi $\lim_{y \to x} V([x,y];f) = |f(x^+) - f(x)|$. Dans le même style on a par exemple pour $a < y \leq z \leq b$: $\lim_{x \to y^-} V([x,z];f) = |f(y) - f(y^-)| + V([y,z];f)$, etc...

Il est évident qu'une fonction f est à variation bornée si et seulement si Re(f) et Im(f) le sont. Je suppose maintenant que f est *à valeurs réelles*. Considérons la fonction $x \mapsto f_1(x) = V([a,x];f)$. Elle est croissante. Définissons $f_2(x) = f_1(x) - f(x)$, et examinons pour $x < y$ la quantité $f_2(y) - f_2(x) = f_1(y) - f_1(x) - (f(y) - f(x)) = V([x,y];f) - (f(y) - f(x))$. La dernière expression est certainement ≥ 0. Donc f_2 est aussi une fonction croissante. Autrement dit $f = f_1 - f_2$ est la différence de deux fonctions croissantes.

Il est par ailleurs évident que des combinaisons linéaires de fonctions de variations bornées sont de variations bornées et que les fonctions croissantes ou décroissantes sont de variations bornées. On a donc obtenu :

Théoréme : *les fonctions à valeurs complexes de variations bornées sont exactement les fonctions de la forme* $f = f_1 - f_2 + i(f_3 - f_4)$ *avec* f_1, f_2, f_3, f_4 *croissantes.*

Remarquons que toute fonction de variation bornée est Riemann intégrable puisque les fonctions monotones sont Riemann intégrables. Aussi elle est bornée et n'a au pire qu'un nombre dénombrable de discontinuités puisque c'est le cas des fonctions monotones.

Variation bornée et mesures complexes

Vous pouvez sauter cette section dans une première lecture, et continuer avec la discussion du théorème de Dirichlet-Jordan en sautant ensuite les paragraphes parlant de mesures complexes. Le théorème précédent est fortement réminiscent de la décomposition $\mu = \mu_1 - \mu_2 + i(\mu_3 - \mu_4)$ d'une mesure complexe en une combinaison de mesures positives finies. Rappelons qu'il existe un chapitre essentiel du cours d'intégration qui montre qu'à toute fonction g croissante et continue à gauche sur $[a,b[$, on peut associer une mesure positive, dite de Lebesgue-Stieltjes, telle que $\mu([a,x[) = g(x) - g(a)$ pour $a \leq x < b$, autrement dit

$$\mu([x,y[) = g(y) - g(x)$$

pour $a \leq x \leq y < b$ (on notera que $\mu(\{a\}) = g(a^+) - g(a)$, plus généralement $\mu(\{x\}) = g(x^+) - g(x)$ pour tout x). Et à toute mesure positive μ sur $[a,b[$ on peut associer les fonctions croissantes et continues à gauche $g(x) = C + \mu([a,x[)$, avec $C = g(a)$ arbitraire (la mesure sera finie si et seulement si $g(b^-) < \infty$).

Donc, revenons à notre fonction à valeurs complexes et de variation bornée sur $[a,b]$. On l'écrit $f = f_1 - f_2 + i(f_3 - f_4)$ avec f_1, f_2, f_3, f_4 croissantes, et on remplace chaque f_j par f_j^-. Chaque fonction croissante f_j n'a au pire qu'un nombre dénombrable de points de discontinuités, et donc la fonction croissante $x \to f_j^-(x) = f_j(x^-)$ (en $x = a$, on convient que $f_j^-(a) = f_j(a)$), qui est continue *à gauche*, ne peut différer de f_j qu'en un nombre dénombrable de points. Alors $f^* = f_1^- - f_2^- + i(f_3^- - f_4^-)$ est continue à gauche et ne diffère de f qu'en au pire un nombre dénombrable de points. Par construction $f^*(a) = f(a)$ et par ailleurs comme $f^*(y) = f(y)$ à un nombre dénombrable d'exceptions près et que $f^*(x) = \lim_{y \to x, y<x} f^*(y)$ on a aussi $f^*(x) = f(x^-)$ pour tout $x > a$ (rappel : on sait que $f(x^-)$ existe !), donc f^* ne dépend pas des choix faits et est égal à f^-. En considérant les mesures positives finies μ_j associées aux f_j^-, on aboutit donc à une écriture $f(x^-) = f(a) + \mu([a,x[)$ pour une certaine mesure complexe μ sur $[a,b[$. On pourra alors définir $\mu(\{b\})$ comme valant $f(b) - f(b^-)$. Réciproquement les formules $x \mapsto C + \mu([a,x[)$ et $f(b) = C + \mu([a,b])$ définissent une fonction à variation bornée sur $[a,b]$ et continue à gauche sur $[a,b[$.

Donc, toute fonction f de variation bornée sur $[a,b]$ définit de manière unique une mesure complexe sur $[a,b]$ via les formules $\mu([a,x[) = f(x^-) - f(a)$, $\mu(\{b\}) = f(b) - f(b^-)$ et f coïncide avec $f(a) + \mu([a,x[)$ en tout point où elle est continue à gauche. L'on a $\mu([a,b]) = f(b) - f(a)$. On peut aussi si l'on préfère, utiliser des formules du genre $C + \mu([a,x])$ (continue à droite) ou $C + \mu([x,b])$ (continue à gauche) ou $f(x^+) = f(b) + \mu(]x,b])$ (continue à droite). Ou encore on conviendra que $f(x) = f(a) + \mu([a,x[) + \frac{1}{2}\mu(\{x\})$ pour $a < x < b$ ce qui impose $f(x) = \frac{1}{2}(f(x^-) + f(x^+))$ en tout point de $]a,b[$, ce qui est utile en théorie des séries de Fourier.

Un dernier mot (assez long, désolé). Soit f de variation bornée sur $[a,b]$ et normalisée de manière à être en chaque x soit continue à droite, soit continue à gauche, soit à vérifier la formule $f(x) = \frac{1}{2}(f(x^-) + f(x^+))$ en tout point intérieur (ce que l'on veut en fait c'est que f(x) soit sur le segment dans le plan complexe $[f(x^-), f(x^+)]$ pour tout x de sorte que $|f(x^+) - f(x^-)| = |f(x)-f(x^-)|+|f(x^+)-f(x)|$). Soit μ la mesure de Lebesgue-Stieltjes associée, qui vérifie donc $\mu(\{a\}) = f(a^+) - f(a)$, $\mu(\{b\}) = f(b) - f(b^-)$, $\mu(]x, y[) = f(y^-) - f(x^+)$.

Théorème : $V([a, b]; f) = |\mu|([a, b])$ *est la variation totale de* μ. (*)

Tout d'abord si $a = x_0 < x_1 < \cdots < x_N = b$, alors $V((x_j)) \leq |f(a) - f(a^+)| + |f(a^+) - f(x_1^-)| + |f(x_1^-) - f(x_1)| + |f(x_1) - f(x_1^+)| + \ldots$. Or, par nos conventions sur f en chaque discontinuité, $|f(x_1^-) - f(x_1)| + |f(x_1) - f(x_1^+)| = |f(x_1^-) - f(x_1^+)|$, etc..., donc $V((x_j)) \leq |\mu(\{a\})| + |\mu(]a, x_1[)| + |\mu(\{x_1\})| + |\mu(]x_1, x_2[)| + \ldots$ ce qui est toujours inférieur ou égal à $|\mu|([a, b])$ par définition de ce dernier. Donc en tout cas $V([a, b]; f) \leq |\mu|([a, b])$. Le problème est de montrer l'inégalité dans l'autre sens, sachant que $|\mu|([a, b])$ est défini comme étant le supremum des sommes $\sum |\mu(A_n)|$ pour toutes les partitions dénombrables $[a, b] = \bigcup_n A_n$.

Considérons (ne me demandez pas pourquoi, je suis mon instinct, guidé par le premier mouvement de la sérénade pour cordes de Piotr Tchaikowski) la fonction croissante $g(x) = V([a, x]; f)$. Il est associé à g une mesure positive finie ν par les formules $\nu([a, x[) = g(x^-) - g(a) = g(x^-) = V([a, x[; f)$, $\nu(\{b\}) = g(b) - g(b-) = |f(b) - f(b^-)|$, $\nu([a, b]) = g(b) - g(a) = V([a, b]; f)$.

Comparons $\mu(]x, y[)$ et $\nu(]x, y[)$. Tout d'abord on a $\nu(]x, y[) = g(y^-) - g(x^+)$ et cela en fait coïncide avec $V(]x, y[; f)$ (en effet, on sait que $g(x^+) = g(x) + \lim_{z \to x^+} V([x, z]; f) = g(x) + |f(x^+) - f(x)|$ et $g(y^-) = V([a, y[; f) = g(y) - |f(y) - f(y^-)|$, donc $g(y^-) - g(x^+) = g(y) - g(x) - |f(y) - f(y^-)| - |f(x^+) - f(x)| = V([x, y]; f) - |f(y) - f(y^-)| - |f(x^+) - f(x)| = V(]x, y[; f)$). Par ailleurs $\mu(]x, y[) = f(y^-) - f(x^+)$ donc $|\mu(]x, y[)| \leq |f(y^-) - f(x^+)| \leq V(]x, y[; f)$. Ainsi

$$x < y \Longrightarrow \quad |\mu(]x, y[)| \leq \nu(]x, y[)$$

De plus on a $\nu([a, x[) = g(x^-) - g(a) = V([a, x[; f)$ et $\mu([a, x[) = f(x^-) - f(a)$ donc $|\mu([a, x[)| \leq |\nu([a, x[)|$. Et finalement $\nu(]y, b]) = g(b) - g(y^+) = g(b) - g(y) - |f(y^+) - f(y)| = V([y, b]; f) - |f(y^+) - f(y)| = V(]y, b]; f)$ et $\mu(]y, b]) = f(b) - f(y^+)$ donc aussi $|\mu(]y, b])| \leq \nu(]y, b])$. Tout ouvert U pour la topologie induite dans l'intervalle [a, b] est une union disjointe dénombrable d'intervalles du type [a, x[,]x, y[, et]y, b]. Donc on a $|\mu(U)| \leq \nu(U)$ pour tout ouvert. Cela implique donc $|\mu(A)| \leq \nu(A)$ pour (maintenant c'est un rêve d'amour de Franz Liszt, interprété par François René Duchable. Ça change des choses qui commencent par Blo et terminent par cage) toute intersection dénombrable (que l'on peut prendre décroissante) d'ouverts. Bon, je dois réfléchir. Considérons une fonction continue à valeurs complexes $\phi(x)$ quelconque ; alors soit $\eta > 0$ et décomposons le plan complexe (merci Mozart et Anne-Sophie Mutter) en les translatés

(*) Attention : il aurait été plus prudent d'écrire $V(]a, b[; f) = |\mu|(]a, b[)$ (c'est équivalent car $|\mu|(\{a\}) = |\mu(\{a\})| = |f(a) - f(a^+)|$, etc...) car alors plus généralement $V(]c, d[; f) = |\mu|(]c, d[)$ pour tout $a \leq c \leq d \leq b$; Mais comme $|\mu|(\{c\}) = |f(c^+) - f(c^-)|$, tandis que $V([c, d]; f)$ ne tient compte que de $|f(c^+) - f(c)|$ on peut avoir $V([c, d]; f) < |\mu|([c, d])$.

du << carré >> $\{0 < \mathrm{Re}(z), \mathrm{Im}(z) < \eta\} \cup [0,\eta[\cup[0,i\eta[$. Les images réciproques de ces carrés sont chacune une intersection dénombrable d'ouverts car ϕ est continue. Donc en examinant $\int_{[a,b]} \phi(x)d\mu(x)$ avec un η-découpage des valeurs de ϕ, en faisant tendre η vers 0, j'obtiens (Toréador, toréador... l'amour, l'amour d'abord) $\left|\int_{[a,b]} \phi(x)d\mu(x)\right| \leq \int_{[a,b]} |\phi(x)|d\nu(x)$ pour toute fonction continue sur $[a,b]$, à valeurs complexes.

Petite pause Stade 2. Reprise. En utilisant que les fonctions continues sont denses dans l'espace $L^1([a,b], |\mu| + \nu)$,(*) nous pouvons alors affirmer que $\left|\int_{[a,b]} \phi(x)d\mu(x)\right| \leq \int_{[a,b]} |\phi(x)|d\nu(x)$ pour toute fonction dans cet espace par exemple pour toute fonction mesurable bornée. Je rappelle qu'il existe une fonction $h(x)$ de module 1 telle que l'on puisse écrire $d\mu(x) = h(x)d|\mu|(x)$, on utilise $\phi(x) = \overline{h(x)}$ et on obtient

$$|\mu|([a,b]) = \int_{[a,b]} d|\mu|(x) \leq \int_{[a,b]} d\nu(x) = V([a,b];f)$$

Le Théorème est ainsi démontré.

Autrement dit ce qui a été prouvé ici c'est que lorsque μ est une mesure complexe sur $[a,b]$ sa variation totale $\|\mu\| = |\mu|([a,b])$ au sens de l'annexe précédente est aussi le supremum de toutes les sommes finies $|\mu([a,x_1[)| + |\mu([x_1,x_2[)| + \cdots + |\mu([x_{N-1},b[)| + |\mu(\{b\})|$, pour toutes les subdivisions $a = x_0 < x_1 < \cdots < x_N = b$. Une troisième caractérisation de $\|\mu\|$ est à portée de main si vous reprenez notre technique de démonstration c'est que $\|\mu\|$ est la plus petite constante C avec $|\int_{[a,b]} \phi(x)\,d\mu(x)| \leq C \cdot \|\phi\|_\infty$ pour toutes les fonctions continues $\phi : [a,b] \dashrightarrow \mathbb{C}$. Cependant je n'en parle pas plus car cela m'entraînerait sur la pente glissante vers une discussion du fondamental << Théorème de représentation de Riesz >> (à ne pas confondre avec le théorème de Fréchet-Riesz) qui caractérise les mesures complexes comme étant les éléments du dual de l'espace de Banach des fonctions continues muni de la norme sup (de la convergence uniforme). Je dois laisser un peu de travail à mes collègues sévissant en M1 et M2, et de plus j'ai déjà réécrit au fil de ces annexes une trop grande partie du Rudin << Analyse réelle et complexe >>, si vous vous sentez intéressé(e) allez le regarder (il vous faudra surmonter votre dégoût de l'horrible typographie de l'édition française actuelle, qui induit même des erreurs dans certaines preuves ou énoncés d'exercices !).

(*) damenède ! je n'ai prouvé cela que pour dx. Bon, voyons voir, oui, je rappelle que l'idée (communément appelée << technique Burnolienne >> dans la littérature ; euh...non, pas encore, mais cela va venir), dans la démonstration du cours, c'est de montrer que les parties mesurables A pour lesquelles $\mathbb{1}_A$ est approchable (dans le cours par des fonctions en escalier, ici par des fonctions continues) forment une tribu, ensemble vide ok, complémentaire...ok, union croissante......ok si union finie, union de deux.........les << identités Burnoliennes >> font le job ici aussi ; et bien sûr cette tribu contient les intervalles ouverts, voyons, oui, ok on approxime par en dessous, convergence dominée, c'est bon. Sous-jacent à cela sans le dire il y a la propriété de << régularité >> des mesures de Lebesgue-Stieltjes, $\nu(A)$ est l'infimum des $\nu(U)$ pour U ouvert contenant A et le supremum des $\nu(K)$ pour K compact inclus dans A. Si vous voulez approfondir, direction le Rudin ou tout livre sérieux sur la théorie de l'intégration.

Dirichlet-Jordan

Théorème : *soit* f *une fonction* 2π*-périodique et de variation bornée sur* $[0,2\pi]$. *Alors en tout* x *on a*

$$\lim S_N(f)(x) = \frac{f(x^-)+f(x^+)}{2}$$

Par le théorème de localisation si on suppose seulement f de variation bornée sur un sous-intervalle $[a,b]$ alors la conclusion vaudra pour tous les $x \in]a,b[$. Je vais faire une première preuve, qui utilise la mesure complexe μ de Lebesgue-Stieltjes associée à f et le théorème de la convergence dominée (on peut modifier la preuve pour n'utiliser que la convergence uniforme, mais à partir du moment où l'on parle de mesures, cela n'a pas grand sens de vouloir éviter l'emploi du théorème de la convergence dominée ou celui de la convergence monotone !). Dans cette preuve je vais utiliser la formule

$$\sum_{k=1}^{\infty} \frac{\sin(kx)}{k} = \frac{\pi - x}{2} \qquad (0 < x < 2\pi)$$

et surtout la propriété

$$\exists C < \infty \qquad \forall x \in \mathbb{R} \quad \forall N \in \mathbb{N} \quad \left|\sum_{k=1}^{N} \frac{\sin(kx)}{k}\right| \leq C$$

L'existence d'un tel $C < \infty$ a été traitée en exercice (voir en particulier l'exercice ??, pour une parmi de multiples méthodes possible).

Soit $L = \frac{f(x^-)+f(x^+)}{2}$ et $g(t) = \frac{f(x+t)+f(x-t)}{2} - L$ sur $]0,\pi]$, de sorte que $g(0^+) = 0$. On pose par convention $g(0) = 0$. Il s'agit de montrer que $\lim \int_0^\pi g(t)D_N(t)dt = 0$, sachant que g est de variation bornée et que $g(0^+) = 0$. Pour ceux qui ont peur des mesures complexes, sautez ce qui suit. Une variante qui utilise plutôt la deuxième formule de la moyenne est donnée ensuite ; on y utilise toujours l'existence d'une majoration uniforme pour les sommes $\sum_{k=1}^N \frac{\sin(kx)}{k}$.

L'intégrale ne change pas si je modifie g en le nombre dénombrable de points où elle est discontinue et donc je peux lui imposer d'être continue à gauche. J'écris $g(t) = \int_{[0,t[} d\mu(u)$, et je prends note que $\mu(\{0\}) = g(0^+) = 0$, donc je peux aussi écrire $g(t) = \int_{]0,t[} d\mu(u)$. J'applique Fubini (Fubini vaut pour μ_1, μ_2, μ_3, μ_4 de l'écriture $\mu = \mu_1 - \mu_2 + i(\mu_3 - \mu_4)$), soigneusement, car ici on a une mesure qui peut donner une masse aux singletons :

$$\begin{aligned}\int_{0<t<\pi} g(t)D_N(t)dt &= \iint_{0<u<t<\pi} D_N(t)\,dt\,d\mu(u)\\ &= \int_{0<u<\pi}\left(\int_u^\pi D_N(t)\,dt\right)\,d\mu(u)\end{aligned}$$

Je calcule $\int_u^\pi D_N(t)\,dt = \int_u^\pi (1 + 2\cos t + \cdots + 2\cos Nt)\,dt = \pi - u - 2\sum_{k=1}^N \frac{\sin(ku)}{k}$. Je sais que pour chaque $u \in]0,\pi]$ cela tend vers 0 lorsque $N \to \infty$ et surtout que toutes ces expressions sont bornées indépendamment de u et de N, donc il y a convergence dominée vers 0 sur $]0,\pi]$ (ou alors vous travaillez sur $[0,\pi]$ et il y a alors convergence presque partout puisque $|\mu|(\{0\}) = 0$). Le théorème de la convergence dominée (valable pour μ car valable pour μ_1, ..., μ_4) donne alors la conclusion recherchée.

À propos de la formule $\pi - u = 2\sum_{k=1}^\infty \frac{\sin(ku)}{k}$, je signale en passant qu'elle découle immédiatement du fait que $\lim \int_u^\pi D_N(t)\,dt = 0$ pour $0 < u \leq \pi$ par application du Lemme de Riemann-Lebesgue, compte tenu de la formule $D_N(t) = \dfrac{\sin((N+\frac{1}{2})t)}{\sin(\frac{t}{2})}$. Sans même d'ailleurs à avoir à invoquer le Lemme de Riemann-Lebesgue il suffit élémentairement d'intégrer par parties pour faire apparaître un $\dfrac{1}{N+\frac{1}{2}}$!

Deuxièmes formules de la moyenne

La raison pour le pluriel apparaîtra bientôt. Je signale tout de suite que toutes ces << deuxièmes >> formules s'obtiennent très élémentairement par des intégrations par parties, lorsqu'elles sont possibles. Le substitut à l'intégration par parties, c'est de se ramener à une intégrale double, avec des mesures complexes, et d'utiliser le théorème de Fubini, comme dans la preuve que j'ai faite de Dirichlet-Jordan. Mais bon, je vais faire sans ici. Soit donc g de variation bornée sur $[a,b]$ et soit f une fonction Lebesgue intégrable quelconque. Comme g est bornée le produit $f(x)g(x)$ est intégrable. On veut majorer intelligemment $I = \int_a^b f(x)g(x)\,dx$. Prenons une subdivision $a = x_0 < x_1 < \cdots < x_N = b$ quelconque. Et considérons l'approximation :

$$S = \sum_{j=1}^N \int_{x_{j-1}}^{x_j} f(x)g(x_j)\,dx$$

Déjà, que pouvons-nous dire que $I - S$? On a :

$$|I - S| \leq \sum_{j=1}^N \left(\int_{x_{j-1}}^{x_j} |f(x)|\,dx\right) V([x_{j-1}, x_j]; g) \leq \left(\max_j \int_{x_{j-1}}^{x_j} |f(x)|\,dx\right) V([a,b]; g)$$

Je rappelle que (pourquoi ?) la fonction $x \mapsto \int_a^x |f(t)|\,dt$ est continue, donc uniformément continue sur $[a,b]$. Donc on aura $S \to I$ pour tous les choix de subdivisions assurant que le pas $\max_j |x_j - x_{j-1}|$ tende vers zéro pour $N \to \infty$. Par exemple on peut prendre $x_j = a + j\frac{b-a}{N}$.

Posons maintenant $u_j = g(x_j)$, $w_j = \int_{x_j}^b f(x)\,dx$ et écrivons $\int_{x_{j-1}}^{x_j} f(x)g(x_j)\,dx = (w_{j-1} - w_j)u_j$. Ainsi (sommation d'Abel ; remarquez $w_N = 0$) :

$$S = \sum_{j=1}^N (w_{j-1} - w_j)u_j = w_0u_1 + w_1(u_2 - u_1) + \cdots + w_{N-1}(u_N - u_{N-1})$$

$$S - g(a) \int_a^b f(x)\,dx = w_0(u_1 - u_0) + w_1(u_2 - u_1) + \cdots + w_{N-1}(u_N - u_{N-1})$$

$$\left| S - g(a) \int_a^b f(x)\,dx \right| \leq \Big(\max_{0\leq j\leq N} |w_j| \Big) \big(|u_1 - u_0| + |u_2 - u_1| + \cdots + |u_N - u_{N-1}| \big)$$

$$\leq \left(\sup_{[a,b]} \left| \int_\xi^b f(x)\,dx \right| \right) V([a,b];g)$$

Nous avons donc identifié un disque fermé du plan complexe dans lequel tombe toutes les expressions S. Comme l'intégrale I est un point limite de telles S, on a donc prouvé une première version de la deuxième formule de la moyenne :

$$\int_a^b f(x)g(x)\,dx = g(a) \int_a^b f(x)\,dx + R \qquad |R| \leq \left(\sup_{[a,b]} \left| \int_\xi^b f(x)\,dx \right| \right) V([a,b];g)$$

En particulier :

$$\left| \int_a^b f(x)g(x)\,dx \right| \leq \Big(|g(a)| + V([a,b];g) \Big) \sup_{[a,b]} \left| \int_\xi^b f(x)\,dx \right|$$

Je vous conseille en supposant g de classe C^1 d'intégrer par parties l'intégrale de gauche et vous retrouverez immédiatement ce résultat. Par ailleurs dans le terme de gauche on peut remplacer g(a) par $g(a^+)$ et aussi g(b) par $g(b^-)$ et on a ainsi une autre majoration :

$$\left| \int_a^b f(x)g(x)\,dx \right| \leq \Big(|g(a^+)| + V(]a,b[;g) \Big) \sup_{[a,b]} \left| \int_\xi^b f(x)\,dx \right|$$

On peut échanger dans la formule de la moyenne les rôles de a et de b :

$$\left| \int_a^b f(x)g(x)\,dx \right| \leq \Big(|g(b^-)| + V(]a,b[;g) \Big) \sup_{[a,b]} \left| \int_a^\xi f(x)\,dx \right|$$

Mais passons maintenant à la variante la plus générale : on modifie la preuve plus haut en écrivant certes $\int_{x_{j-1}}^{x_j} f(x)\,dx = w_{j-1} - w_j$ mais avec $w_j = -F(x_j)$ pour $F(x) = C + \int_a^x f(t)\,dt$, $C \in \mathbb{C}$ quelconque. Le réarrangement télescopique donne :

$$S = w_0 g(x_1) + w_1(g(x_2) - g(x_1)) + \cdots + w_{N-1}(g(b) - g(x_{N-1})) - w_N g(b)$$

$$S + F(a)g(a) - F(b)g(b) = w_0(g(x_1) - g(a)) + w_1(g(x_2) - g(x_1)) + \cdots + w_{N-1}(g(b) - g(x_{N-1}))$$

et le terme de droite est majoré par $(\sup_{[a,b]} |F(\xi)|) V([a,b];g)$. On conclut :

$$\left| \int_a^b f(x)g(x)\,dx - \Big(F(b)g(b) - F(a)g(a) \Big) \right| \leq \Big(\sup_{[a,b]} \big|F(\xi)\big| \Big) V([a,b];g)$$

L'analogie avec la formule d'intégration par parties fait que cette inégalité est facile à mémoriser ! On peut remplacer g(a), g(b), V([a,b];g) par $g(a^+)$, $g(b^-)$ et V(]a,b[;g). Avec $f(x) = e^{ix}$ par exemple le choix de $F(x) = -ie^{ix}$ peut être plus avantageux du point de vue de $\sup_\xi |F(\xi)|$ que $\int_a^x e^{it}\,dt$ ou $-\int_x^b e^{it}\,dt$.

Pour g monotone et f à valeurs réelles, il y a une version un peu plus précise et très classique. Supposons g croissante (non constante). Revenons à :

$$S - g(a)\int_a^b f(x)\,dx = w_0(g(x_1) - g(a)) + w_1(g(x_2) - g(x_1)) + \cdots + w_{N-1}(g(b) - g(x_{N-1}))\ ,$$

le terme de droite, divisé par $g(b) - g(a)$ est un barycentre des valeurs $w_j = \int_{x_j}^b f(x)\,dx$ prises par la fonction continue $\xi \dashrightarrow F(\xi) = \int_\xi^b f(x)\,dx$. Soit J l'intervalle (fermé borné) $F([a,b])$. On a donc $S \in g(a)\int_a^b f(x)\,dx + (g(b) - g(a))J$. C'est donc aussi le cas de l'intégrale I qui est une limite de tels S :

$$\exists\xi \in [a,b] \qquad \int_a^b f(x)g(x)\,dx = g(a)\int_a^b f(x)\,dx + (g(b) - g(a))\int_\xi^b f(x)\,dx$$

On peut remplacer $g(a)$ par $g(a^+)$ et $g(b)$ par $g(b^-)$.

$$\exists\xi \in [a,b] \qquad \int_a^b f(x)g(x)\,dx = g(a^+)\int_a^b f(x)\,dx + (g(b^-) - g(a^+))\int_\xi^b f(x)\,dx$$

En fait l'on peut toujours prendre $\xi \in\ ]a,b[$ (sautez ce paragraphe en première lecture). Si g est constante sur $]a,b[$ alors en fait $g(b^-) - g(a^+) = 0$ et la formule marche pour tous les ξ. Je suppose donc g non constante sur $]a,b[$. Supposons par exemple que $\xi = b$ marche. Il s'agit alors de montrer qu'il existe un $y \in\ ]a,b[$ avec $\int_y^b f(x)\,dx = 0$. Sinon, le signe est constant, disons strictement positif. Prenons $a < c < d < b$ avec $g(c) < g(d)$. Sur chaque sous-intervalle on applique la forme actuelle de la formule, ce qui donne $\int_a^c f(x)g(x)\,dx = g(a^+)\int_a^c f(x)\,dx + (g(c) - g(a^+))\int_{\xi_1}^c f(x)\,dx = g(a^+)\int_a^b f(x)\,dx + (g(c) - g(a^+))\int_{\xi_1}^b f(x)\,dx - g(c)\int_c^b f(x)\,dx$, puis $\int_c^d f(x)g(x)\,dx = g(c)\int_c^d f(x)\,dx + (g(d) - g(c))\int_{\xi_2}^d f(x)\,dx = g(c)\int_c^b f(x)\,dx + (g(d) - g(c))\int_{\xi_2}^b f(x)\,dx - g(d)\int_d^b f(x)\,dx$, enfin $\int_d^b f(x)g(x)\,dx = g(d)\int_d^b f(x)\,dx + (g(b^-) - g(d))\int_{\xi_3}^b f(x)\,dx$. En sommant on obtient : $\int_a^b f(x)g(x)\,dx = g(a^+)\int_a^b f(x)\,dx + (g(c) - g(a^+))\int_{\xi_1}^b f(x)\,dx + (g(d) - g(c))\int_{\xi_2}^b f(x)\,dx + (g(b^-) - g(d))\int_{\xi_3}^b f(x)\,dx$. Comme $\xi_2 \in [c,d] \subset\ ]a,b[$ le troisième terme est strictement positif, tandis que le deuxième et le quatrième sont positifs ou nuls. On a donc $\int_a^b f(x)g(x)\,dx > g(a^+)\int_a^b f(x)\,dx$ ce qui contredit le fait que $\xi = b$ marchait dans la formule. On montrerait d'une manière exactement semblable que si $\xi = a$ marche, c'est-à-dire si $\int_a^b f(x)g(x)\,dx = g(b^-)\int_a^b f(x)\,dx$ alors il existe $y \in]a,b[$ avec $\int_a^y f(x)\,dx = 0$ et donc aussi $\int_a^b f(x)g(x)\,dx = g(b^-)\int_a^b f(x)\,dx + (g(a^+) - g(b^-))\int_a^y f(x)\,dx = g(a^+)\int_a^b f(x)\,dx + (g(b^-) - g(a^+))\int_y^b f(x)\,dx$.

Nous avons donc obtenu la formule célèbre (f à *valeurs réelles*, g monotone) :

$$\exists\xi \in\ ]a,b[\qquad \int_a^b f(x)g(x)\,dx = g(a^+)\int_a^\xi f(x)\,dx + g(b^-)\int_\xi^b f(x)\,dx$$

On avait g croissante non constante ; g constante est évident, et pour g décroissante on remplace g par -g. Pour g à variation bornée, ou f à valeurs complexes, on doit se contenter des inégalités établies plus haut.

Encore une fois, lorsque g est C^1, intégrez par parties, et vous aurez tout ce que peut donner la deuxième formule de la moyenne. Et en toute généralité la puissance optimale est obtenue en exprimant g avec une mesure complexe, et en utilisant Fubini, car là on a un résultat exact, pas une inégalité, ou une pseudo-égalité avec un ξ que l'on ne connaît pas. Une erreur très fréquente est d'utiliser la formule avec un ξ alors que f est à valeurs complexes, pas à valeurs réelles. L'inégalité que l'on écrit après est en général valable, mais l'étape intermédiaire était fausse.

D i r i c h l e t - J o r d a n (b i s)

Voici une formulation plus élémentaire de la preuve du théorème de Dirichlet-Jordan. On veut prouver $\lim \int_0^\pi g(t)D_N(t)dt = 0$, sachant que g est de variation bornée et que $g(0^+) = 0$. On fait la découpe usuelle pour un $\delta \in]0,\pi]$:

$$\int_0^\pi g(t)D_N(t)\,dt = \int_0^\delta g(t)D_N(t)\,dt + \int_\delta^\pi g(t)\frac{\sin((N+\frac{1}{2})t)}{\sin(\frac{t}{2})}\,dt$$

Le deuxième terme à droite tend vers 0 par le Lemme de Riemann-Lebesgue (comme $g(t)/\sin(\frac{t}{2})$ est Riemann intégrable sur $[\delta,\pi]$ il suffit d'invoquer quelque chose d'<< élémentaire >> comme l'inégalité de Bessel ou l'approximation par une fonction en escalier). On a donc

$$\forall\delta \in]0,\pi] \quad \limsup_{N\to\infty}\left|\int_0^\pi g(t)D_N(t)\,dt\right| \le \limsup_{N\to\infty}\left|\int_0^\delta g(t)D_N(t)\,dt\right|$$

On utilise la deuxième formule de la moyenne :

$$\left|\int_0^\delta g(t)D_N(t)\,dt\right| \le \Big(|g(0^+)| + V(]0,\delta[;g)\Big)\sup_{]0,\delta]}\left|\int_\xi^\delta D_N(t)\,dt\right|$$

Le supremum est majoré par une constante absolue C, en écrivant $\int_\xi^\delta D_N(t)\,dt = \int_\xi^\pi D_N(t)\,dt - \int_\delta^\pi D_N(t)\,dt$ et en rappelant que $\int_u^\pi D_N(t)\,dt = \pi - u - 2\sum_{k=1}^N \frac{\sin(ku)}{k}$ est borné indépendamment de u et de N. De plus $g(0^+) = 0$. Ainsi :

$$\left|\int_0^\delta g(t)D_N(t)\,dt\right| \le V(]0,\delta[;g)\cdot C$$

et a fortiori $\limsup_{N\to\infty}\left|\int_0^\delta g(t)D_N(t)\,dt\right| \le V(]0,\delta[;g)\cdot C$. Donc

$$\forall\delta \in]0,\pi] \quad \limsup_{N\to\infty}\left|\int_0^\pi g(t)D_N(t)\,dt\right| \le V(]0,\delta[;g)\cdot C$$

Il suffit alors de rappeler que $\lim_{\delta\to 0} V(]0,\delta];g) = 0$ pour conclure la preuve.

Dans ce contexte, je laisse en exercice $c_n(g) = \int_0^{2\pi} g(t)e^{-int}\frac{dt}{2\pi} = \mathcal{O}(\frac{1}{|n|})$ pour g à variation bornée sur $[0,2\pi]$. Plus précisément $|c_n(g)| \le \frac{V([0,2\pi];g)}{2\pi|n|}$ (attention : on impose $g(2\pi) = g(0)$... pourquoi ?). La deuxième formule de la moyenne (version générale) ou Fubini permettent de le prouver.