MT40 Automne 2004

Corrigé de l'examen médian du 17 novembre 2004

Correction de l'exercice 1.
On pourra faire tourner le script matlab `mediancorexo1MT40_A04.m`, disponible à l'adresse habituelle `http://utbmjb.chez.tiscali.fr/`

1. Comme dans le cours, on produit une table des valeurs des différences divisées de la fonction f type suivant :

x_i	$f[x_i]$	$f[x_i, x_{i+1}]$	$f[x_i, x_{i+1}, x_{i+2}]$	$f[x_i, x_{i+1}, x_{i+2}, x_{i+3}]$	$f[x_0, x_1, x_2, x_3, x_4]$
0	1				
		4			
1	5		$-1,5$		
		1		$0,5$	
2	6		0		0
		1		$0,5$	
3	7		$1,5$		
		4			
4	11				

   On déduit de cette table, l'expression du polynôme p_4 :

$$p_4(x) = 1 + 4 \times x - 1,5 \times x(x-1) + 0,5 \times x(x-1)(x-2) + 0 \times x(x-1)(x-2)(x-3),$$

   soit encore

$$\boxed{p_4(x) = 1 + 4x - 1,5x(x-1) + 0,5x(x-1)(x-2).} \tag{1}$$

   On en déduit donc

$$\boxed{p_4(2,5) = 6,3125.} \tag{2}$$

   Remarque 1. On remarque que $f[x_0, x_1, x_2, x_3, x_4]$ est nul. Puisque

$$p_4(x) = p_3(x) + f[x_0, x_1, x_2, x_3, x_4]\, x(x-1)(x-2)(x-3),$$

   on en déduit que

$$p_4(x) = p_3(x).$$

2. Pour chaque $i \in \{0, ..., 3\}$, on sait que le polynôme p_i est de degré au plus i. De plus, la courbe représentative du polynôme p_i passe par les points de coordonnées $(x_0, f(x_0))$, ..., $(x_i, f(x_i))$. On peut donc tracer de façon qualitative les polynômes p_0, p_1, p_2 et p_3.

   Voir la figure 1 page suivante.

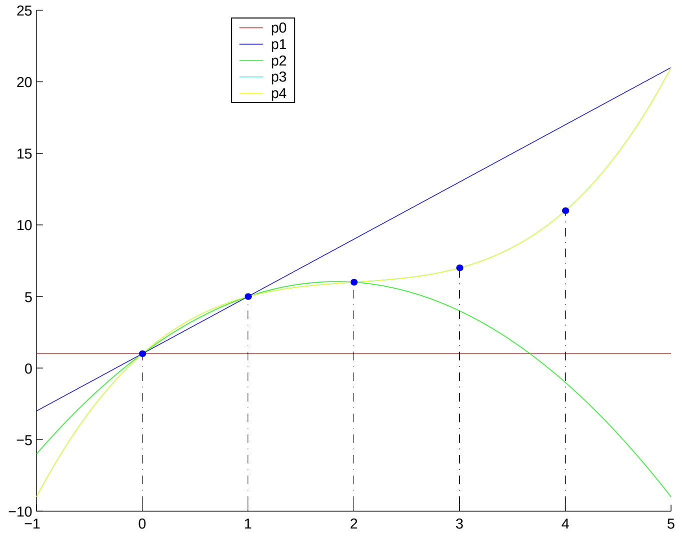

FIG. 1 – Les polynômes p_0, p_1, p_2, p_3 et p_4.

Remarque 2. Sur cette figure, on a aussi tracé le polynôme p_4 ; on constate que $p_3 = p_4$, ce qui est conforme à la remarque 1.

3. On a déjà vu que $f[x_0, x_1, x_2, x_3, x_4]$ était nulle. D'après la remarque 1, cela implique que

$$\boxed{p_4(x) = p_3(x).} \tag{3}$$

4. (a) Dire que l'erreur $e_3 = f - p_3$ est identiquement nulle revient à dire que $f = p_3$, c'est-à-dire que

$$\boxed{f \text{ est un polynôme de degré 3.}} \tag{4}$$

   (b) On ne connaît de f que ses valeurs aux points x_i, ce qui nous donne aucune information sur la nature de f et ne nous permet donc pas de savoir si (4) est vérifiée.

   (c) Si cette condition est vérifiée, alors $f = p_3$ et d'après (1) et (3) on a donc

$$\boxed{f(x) = 1 + 4x - 1,5x(x-1) + 0,5x(x-1)(x-2).} \tag{5}$$

Correction de l'exercice 2.

Seuls des éléments de correction sont fournis ici.

1. On sait que la fonction proposée f est continue sur $[1, 2]$ donc intégrable sur cet intervalle. Il vient évidemment :

$$I = \int_1^2 f(x)dx = \left[-\frac{1}{x}\right]_1^2,$$

et donc

$$\boxed{I = \frac{1}{2}.} \tag{6}$$

Remarque 3. Le fait de trouver I négative est un erreur de calcul, certes mais grave pour un ingénieur qui doit remarquer que l'intégrale d'une fonction positive sur l'intervalle proposé doit être positive. Trouver le contraire est inquiétant car signe un manque d'attention, d'interrogation des résultats trouvés.

2. Etude de E_N^S

   (a) Dérivabilité de f

   - f est une fonction rationnelle définie sur $[1, 2]$ donc indéfiniment dérivable sur cet intervalle comme tout quotient de fonction polynômes non nulles sur l'intervalle considéré. Par suite on peut calculer les nombres dérivés successifs ; ce n'est pas le fait de calculer qui la prouve dérivable, c'est le fait qu'elle le soit, qui autorise le calcul ! La nuance est de taille...
   - On écrit $f(x) = x^{-2}$ d'où $f'(x) = (-2)x^{-3}$ et on continue en évitant soigneusement le théorème du quotient qui relève du masochisme ! Bref, on trouve :

   $$f^{(4)}(x) = \frac{120}{x^6}.$$

   Par suite, vu que $f^{(4)}$ est positive sur $[1, 2]$ (c'est crucial !), décroissante sur cet intervalle car sa dérivée y est négative, grâce à une rapide étude des variations de $f^{(4)}$ sur $[1, 2]$, on déduit que $\left|f^{(4)}\right|$ atteint un maximun en $x = 1$ c'est-à-dire

   $$\forall x \in [1, 2] \quad \left|f^{(4)}(x)\right| \leq 120.$$

   Remarque 4. Si $f^{(4)}$ avait été négative, le maximum de valeur absolue eût été atteint en $x = 2$. Attention aux affirmations foireuses ou incomplètes !

   (b) Le cours nous montre que l'erreur de méthode commise en intégration de Simpson est définie par :

   $$E_N^S = -h^4 \frac{(2-1)}{2880} f^{(4)}(\xi) \quad \text{avec } \xi \in [1, 2] \,.$$

   Par suite

   $$\left|E_N^S\right| = \left|h^4 \frac{1}{2880} f^{(4)}(\xi)\right| = h^4 \frac{1}{2880} \left|f^{(4)}(\xi)\right| \leq \frac{h^4}{24}.$$

   Or on sait que N et h sont liés par la relation : $Nh = 1$. Par conséquent, le nombre N est tel que :

   $$\boxed{\left|E_N^S\right| \leq \frac{1}{24N^4}.} \tag{7}$$

   Or nous voulons que $\left|E_N^S\right| < \varepsilon$. Il suffit que (dire si et seulement si est faux) le majorant trouvé soit lui-même plus petit que ε. Ainsi

   $$\frac{1}{24N^4} < \varepsilon.$$

   On résout cette inéquation en N, sachant que N est cherché entier naturel (ce qui est crucial pour les logiciels connexes !). Il vient par recours à la fonction $t \mapsto 1/t$ décroissante sur $\mathbb{R}_+^*$

   $$N^4 > \frac{1}{24\varepsilon} \quad \text{d'où} \quad N > (24\varepsilon)^{-1/4} \,.$$

Nous savons que N doit être entier, donc en général le plus petit entier naturel N vérifiant la contrainte proposée sera défini par :

$$\boxed{N=\mathrm{E}\left((24\varepsilon)^{-1/4}\right)+1.} \tag{8}$$

(c) La question présente est ici pour signifier que le raisonnement c'est fait en termes de N et non de pas h comme d'autres fois... Evidemment : $h=1/N$.

(d) Applications numériques

Pour les différentes valeurs proposées en ε, on trouve respectivement :

$$N_1=2;\quad N_2=5;\quad N_3=15.$$

3. Etude de la valeur approchée obtenue sous méthode de Simpson

(a) (i) De façon évidente on a

$$S(4)=1+\frac{1}{4}+\frac{1}{9}+\frac{1}{16}\quad \text{et}\quad \mathcal{I}(4)=1+\frac{1}{9}\,;\,\mathcal{P}(4)=\frac{1}{4}+\frac{1}{16}.$$

Les détails des calculs et vérifications sont laissés ici.

(ii) Les preuves proposées dans les copies ont été le plus souvent incorrectes, approximatives...

Considérons donc un entier naturel k quelconque. $\mathcal{P}(k)$ est obtenue en sommant tous les termes de la forme $1/(2j)^2$ pour tous les termes entiers j tels que $2\leq 2j\leq k$. Ainsi les j qui répondent sont les entiers naturels compris entre 1 et $k/2$; selon la parité de k, $k/2$ n'est pas toujours entier mais la valeur maximale de j sera toujours l'entier majoré par $k/2$ c'est-à-dire sa partie entière $\mathrm{E}(k/2)$. il vient donc :

$$\mathcal{P}(k)=\sum_{j=1}^{\mathrm{E}(k/2)}\frac{1}{(2j)^2}=\frac{1}{4}\sum_{j=1}^{\mathrm{E}(k/2)}\frac{1}{j^2}=\frac{1}{4}S\left(\mathrm{E}(k/2)\right).$$

(iii) Par suite évidemmment, comme $S(k)=\mathcal{P}(k)+\mathcal{I}(k)$, on déduit :

$$\mathcal{I}(k)=S(k)-\frac{1}{4}S\left(\mathrm{E}(k/2)\right).$$

(iv) Vu la définition de K_1 et K_2, il est clair, pour K_1 que par simple différence de termes absents

$$K_1=\sum_{i=1}^{4N}\frac{1}{i^2}-\sum_{i=1}^{2N-1}\frac{1}{i^2}=S(4N)-S(2N-1)\,;$$

K_2 se déduit des égalités obtenues juste avant, en ayant remarqué que la somme K_2 est constituée de la somme des inverses des carrés de nombres pairs compris entre 1 et $4N$ en supprimant ceux compris entre 1 et $2N-1$; par suite

$$K_2=\sum_{\substack{i=1\\ i\text{ pair}}}^{4N}\frac{1}{i^2}-\sum_{\substack{i=1\\ i\text{ pair}}}^{2N-1}\frac{1}{i^2}=\left(S(4N)-\frac{1}{4}S\left(\mathrm{E}(4N/2)\right)\right)-\left(S(2N-1)-\frac{1}{4}S\left(\mathrm{E}((2N-1)/2)\right)\right).$$

Or $(2N-1)/2=N-1/2$ donc $\mathrm{E}((2N-1)/2)=N-1$ d'où l'égalité proposée.

ε	N, I_N^S	$\lvert I_N^S - 1/2\rvert$
10^{-2}	$5,00417607 \times 10^{-1}$	$4,1 \times 10^{-4}$
10^{-4}	$5,00012433 \times 10^{-1}$	$1,2 \times 10^{-5}$
10^{-6}	$5,00000111 \times 10^{-1}$	$1,1 \times 10^{-7}$

TAB. 1 – Quelques valeurs N, de l'intégrale approchée I_N^S et de l'erreur $\left|\left|I_N^S - 1/2\right|\right|$.

(b) Utilisation pour le calcul de I_N^S

Ici l'idée est de prendre un pas de h entre 1 et 2. Sous les notations du cours, les points X_i définis par $X_i = ih/2$ correspondent aux points x_i du support lorsque i est pair, aux points milieux m_i lorsque i est impair, pour tous les i compris entre $2N$ et $4N$.

(i) Le premier point à vérifier est trivial ; il est destiné à s'approprier les notations.

(ii) L'analyse de la formule d'intégration de Simpson donnée en cours, montre que les points X_{2N} et X_{4N} interviennent une seule fois, les autres d'indices pairs deux fois, ceux d'indices impairs quatre fois. Pour uniformiser le traitement on double l'intervention des extrêmes quitte à les retirer une fois chacun. Quant aux points d'indices impairs, ils sont comptés deux fois comme les autres dans la première somme ; pour atteindre quatre, on doit les compter tout seuls deux fois encore à part. L'écriture proposée est la traduction de ces faits. On interprète alors en termes de K_1 et K_2, ce qui donne la dernière écriture en K_1 et K_2.

(iii) On remplace alors par les expressions de K_1 et K_2 trouvées ci-dessus, sachant que $Nh = 1$. On obtient alors

$$\boxed{I_N^S = \frac{1}{6N}\left[8N^2\left\{2S\left(4N\right) - 2S\left(2N-1\right) - \frac{1}{4}\left[S\left(2N\right) - S\left(N-1\right)\right]\right\} - \frac{5}{4}\right].} \tag{9}$$

Cette expression peut aussi s'écrire

$$\boxed{I_N^S = \frac{N}{3}\big(8S(4N) - S(2N) - 8S(2N-1) + S(N-1)\big) - \frac{5}{24N}.} \tag{10}$$

L'évènement essentiel, à ce stade, est alors que la somme évaluée pour fournir une valeur approchée de I_N^S est obtenue grâce à quatre appels de la fonction S pour les arguments $4N$, $2N-1$, $2N$ et $N-1$.

(c) Les valeurs proposées dans la table correspondent précisément à celles qui sont nécessaires pour les valeurs $N_1 = 2$, $N_2 = 5$, $N_3 = 15$ trouvées dans la question 2 ci-dessus ; les valeurs de la fonction S en $4N_i$, $2N_i - 1$, $2N_i$ et $N_i - 1$ pour i élément de $\{1, 2, 3\}$.

Numériquement, en utilisant les valeurs de S données dans l'énoncé on obtient les valeurs de N, I_N^S et de $|I_N^S - 1/2|$ pour différentes valeurs de ε. On pourra utiliser la fonction `integrale_simpson_invcarre`. Voir tableau 1.

On vérifie facilement que ces résultats sont cohérents avec la précision promise, en étudiant leur écart à $1/2$.

(d) Prolongement

On connaît trois valeurs de la fonction $N \mapsto g(N) = I_N^S$, grâce à l'étude antérieure, en les points $\{N_1, N_2, N_3\}$. On pourrait avoir l'idée de déterminer le polynôme d'interpolation p_2 de g sur le support $\{N_1, N_2, N_3\}$. Les résultats seraient médiocres en raison de la concavité insatisfaisante de p_2 comparée à la situation réelle de la focntion g. La zone de validité de l'interpolation proposée serait trop réduite comme une étude graphique menée avec les outils matlab proposés le montrerait.

Compléments

On peut aussi éviter le calcul de la somme $S(N)$ en la remplaçant par une approximation (voir (35), en annexe B, justifiée en annexe A).

Rappelons que l'on a montré que l'on a

$$\boxed{I = I_N^S + E_N^S,} \tag{11}$$

où I_N^S est donnée par (10), où I est donné par (6) et E_N^S est majorée d'après (7). Nous allons remplacer I_N^S par une approximation en $0(1/N^4)$. On peut réécrire (35) sous la forme

$$S(N) = d + Q\left(\frac{1}{N+1}\right) + R_N, \tag{12a}$$

où

$$d = \frac{\pi^2}{6}, \tag{12b}$$

$$Q(x) = -x - \frac{x^2}{2} - \frac{x^3}{6}, \tag{12c}$$

$$\forall N \geq 2, \quad |R_N| \leq \frac{1}{3N^5}. \tag{12d}$$

Après calculs, on a donc, selon (10),

$$I_N^S = \frac{N}{3}\left(8Q\left(\frac{1}{4N+1}\right) - Q\left(\frac{1}{2N+1}\right) - 8Q\left(\frac{1}{2N}\right) + Q\left(\frac{1}{N}\right)\right) - \frac{5}{24N} + \widetilde{R}_N, \tag{13}$$

où

$$\left|\widetilde{R}_N\right| \leq \frac{5}{1152N^4} + \frac{1}{36N^4\left(1-\frac{1}{2N}\right)^5} + \frac{1}{9N^4\left(1-\frac{1}{N}\right)^5}. \tag{14}$$

Ici, on se permet[1] une approximation, valable pour N grand : on remplace $\frac{1}{36N^4\left(1-\frac{1}{2N}\right)^5}$ et $\frac{1}{9N^4\left(1-\frac{1}{N}\right)^5}$ respectivement par $\frac{1}{36N^4}$ et $\frac{1}{9N^4}$. Moyennant cet abus, (14) est équivalent à

$$\left|\widetilde{R}_N\right| \leq \frac{165}{1152N^4}. \tag{15}$$

D'après (7), (11) et (13), on a donc

$$I = \frac{N}{3}\left(8Q\left(\frac{1}{4N+1}\right) - Q\left(\frac{1}{2N+1}\right) - 8Q\left(\frac{1}{2N}\right) + Q\left(\frac{1}{N}\right)\right) - \frac{5}{24N} + \widehat{R}_N,$$

où

$$\left|\widehat{R}_N\right| \leq \frac{71}{384N^4}.$$

[1] Pour une fois ... En toute rigueur, il faudrait utiliser un développement asymptotique de $\left(1-\frac{1}{2N}\right)^5$ en $1/N$.

Le polynôme Q étant défini par (12c), on a donc une approximation $I(N)$ de I :

$$\boxed{I(N)=\frac{N}{3}\left(8Q\left(\frac{1}{4N+1}\right)-Q\left(\frac{1}{2N+1}\right)-8Q\left(\frac{1}{2N}\right)+Q\left(\frac{1}{N}\right)\right)-\frac{5}{24N},} \tag{16}$$

qui vérifie

$$\boxed{\forall N\geq 2,\quad |I-I(N)|\leq \frac{71}{384N^4}.} \tag{17}$$

Pour avoir une erreur inférieur à $\varepsilon>0$, on choisit donc N tel que

$$\boxed{\varepsilon>0,\quad N=\mathrm{E}\left(\left(\frac{384}{71}\varepsilon\right)^{-1/4}\right)+1.} \tag{18}$$

Sur le plan théorique, on sait que la limite de $I(N)$ est égale à $I=1/2$. Cela est confirmé par le script `theorie_simpson_approche` qui montre, en formel, que

$$I(N)=\frac{1}{72}\,\frac{18432\,N^7+41472\,N^6+38016\,N^5+18144\,N^4+4752\,N^3+620\,N^2+18\,N-3}{(4\,N+1)^3\,(2\,N+1)^3\,N}.$$

et que

$$\lim_{N\to\infty} I(N)=\frac{1}{2},$$

ce qui est cohérent avec (6). De plus, ce même script permet de montrer que

$$I(N)=\frac{1}{2}-\frac{7}{9216}\frac{1}{N^5}+\frac{5}{4096}\frac{1}{N^6}-\frac{31}{24576}\frac{1}{N^7}+\frac{35}{32768}\frac{1}{N^8}-\frac{635}{786432}\frac{1}{N^9}+\frac{595}{1048576}\frac{1}{N^{10}}+O\left(\frac{1}{N^{11}}\right),$$

ce qui implique donc, pour N tendant vers l'infini,

$$I(N)=\frac{1}{2}-\frac{7}{9216}\frac{1}{N^5}+O\left(\frac{1}{N^5}\right). \tag{19}$$

Or, d'après (17), on a pour N tendant vers l'infini,

$$I(N)=\frac{1}{2}+O\left(\frac{1}{N^4}\right).$$

On obtient donc un résultat plus précis que celui prévu par la théorie : dans le développement asymptotique de $I(N)$ à l'infini, le terme en $1/N^4$ ne devrait pas être nul. Nous n'avons pas d'explications à ce phénomène !

On renvoie à la fonction `integrale_simpson_invcarre` qui calcule l'intégrale approchée de la fonction $1/x^2$ sur $[1,2]$ par la méthode de Simpson, à la précision ε, avec plusieurs variantes :

- La variante 1 correspond à l'utilisation de la fonction `int_fcn`, programmée en TP 3.A. L'entier N est défini par (8).
- La variante 2 correspond à l'utilisation de la formule (10), où S est calculé par le biais de la fonction gamma de deuxième espèce, notée `Psi`, sous matlab. Voir la fonction `sommeinvcarre`. L'entier N est défini par (8).

ε	N	$I_{\text{approch}}(\varepsilon)$	$\lvert I_{\text{approch}}(\varepsilon) - 1/2\rvert$
10^{-2}	2	$5,004176114890401 \times 10^{-1}$	$4,1761 \times 10^{-4}$
10^{-4}	5	$5,000124698607190 \times 10^{-1}$	$1,2469 \times 10^{-5}$
10^{-6}	15	$5,000001588211296 \times 10^{-1}$	$1,5882 \times 10^{-7}$
10^{-10}	143	$5,000000000193049 \times 10^{-1}$	$1,9304 \times 10^{-11}$
10^{-20}	45181	$4,999999999999996 \times 10^{-1}$	$4,4408 \times 10^{-16}$

TAB. 2 – Quelques valeurs N, de l'intégrale approchée $I_{\text{approch}}(\varepsilon)$ et de l'erreur $\lvert I_{\text{approch}}(\varepsilon) - 1/2\rvert$ pour les variantes 1 à 3.

ε	N	$I_{\text{approch}}(\varepsilon)$	$\lvert I_{\text{approch}}(\varepsilon) - 1/2\rvert$
10^{-2}	3	$4,999981016313585 \times 10^{-1}$	$1,8983 \times 10^{-6}$
10^{-4}	7	$4,999999638188485 \times 10^{-1}$	$3,6181 \times 10^{-8}$
10^{-6}	21	$4,999999998275828 \times 10^{-1}$	$1,7241 \times 10^{-10}$
10^{-10}	208	$4,999999999999981 \times 10^{-1}$	$1,9428 \times 10^{-15}$
10^{-20}	65575	$4,999999999999998 \times 10^{-1}$	$1,6653 \times 10^{-16}$
10^{-30}	20736325	$5,000000000000000 \times 10^{-1}$	0

TAB. 3 – Quelques valeurs N, de l'intégrale approchée $I_{\text{approch}}(\varepsilon)$ et de l'erreur $\lvert I_{\text{approch}}(\varepsilon) - 1/2\rvert$ pour la variante 4.

- La variante 3 correspond à l'utilisation de la formule (10), réécrite ainsi : on remarque que

$$S(4N) - S(2N-1) = \sum_{k=2N}^{4N} \frac{1}{k^2},$$

et que

$$S(2N) - S(N-1) = \sum_{k=N}^{2N} \frac{1}{k^2}.$$

D'après (10), on a donc

$$I_N^S = \frac{N}{3}\left(8\sum_{k=2N}^{4N}\frac{1}{k^2} - \sum_{k=N}^{2N}\frac{1}{k^2}\right) - \frac{5}{24N}.$$

Chacune des sommes est calculée directement sous matlab. L'entier N est défini par (8).
- La variante 4 correspond à l'utilisation de la formule (16). L'entier N est défini par (18).

On constate que les variantes 1 à 3 fournissent à peu près les mêmes valeurs numériques. On a indiqué dans le tableau 2 les valeurs de N, de l'intégrale approchée $I_{\text{approch}}(\varepsilon)$ et de l'erreur $\lvert I_{\text{approch}}(\varepsilon) - 1/2\rvert$ pour quelques valeurs de ε pour les variantes 1 à 3.
On a indiqué dans le tableau 3 les valeurs de N, de l'intégrale approchée $I_{\text{approch}}(\varepsilon)$ et de l'erreur $\lvert I_{\text{approch}}(\varepsilon) - 1/2\rvert$ pour quelques valeurs de ε pour la variante 4.
Enfin, la fonction `test_toc` permet de comparer chacune des quatre variantes de façon temporelle.

Annexe A. Développement asymptotique du reste de la fonction Dzêta de Riemann

Cette annexe est inspirée de la correction de l'exercice 11 page 467, chapitre IX de [AF88]. On considère la fonction Dzêta de Riemann définie par

$$\forall \alpha \in]1,+\infty[, \quad \zeta(\alpha)=\sum_{n=1}^{\infty}\frac{1}{n^{\alpha}}. \tag{20}$$

Cette fonction est définie pour tout $\alpha \in]1,+\infty[$, puisque selon la règle de Riemann, la série de terme général $(1/n^{\alpha})$ est absolument convergente. On considère le reste de la fonction Dzêta, défini par

$$\forall N \in \mathbb{N}^*, \quad \forall \alpha \in]1,+\infty[, \quad F(N,\alpha)=\sum_{n=N}^{\infty}\frac{1}{n^{\alpha}}. \tag{21}$$

Nous proposons une méthode de calcul du développement asymptotique de ce reste. Seuls les résultats essentiels sont donnés. On se réferrera à l'exercice 11 page 467, chapitre IX de [AF88].
On pose, pour tout $k \in \mathbb{N}^*$ et pour tout $\alpha \in]1,+\infty[$,

$$A_k(\alpha)=(-1)^{k+1}\frac{\alpha(\alpha+1)(\alpha+2)...(\alpha+k-1)}{k!}. \tag{22}$$

Pour tout $q \in \mathbb{N}^*$, on considère la matrice triangulaire supérieur inversible $M_q(\alpha)$ de $\mathcal{M}_q(\mathbb{R})$ définie par

$$M_q(\alpha)=\begin{pmatrix} A_1(\alpha) & A_2(\alpha) & A_3(\alpha) & \dots & A_{q-1}(\alpha) & A_q(\alpha) \\ 0 & A_1(\alpha+1) & A_2(\alpha+1) & \dots & A_{q-2}(\alpha+1) & A_{q-1}(\alpha+1) \\ 0 & 0 & A_1(\alpha+2) & \dots & A_{q-3}(\alpha+2) & A_{q-2}(\alpha+2) \\ 0 & & \ddots & \ddots & & \vdots \\ 0 & & \dots & & & A_1(\alpha+q-1) \end{pmatrix}. \tag{23}$$

Notons $(m_i(\alpha))_{1\leq i\leq q}$ les coefficients de la première ligne de $M_q^{-1}(\alpha)$ et posons

$$m_{q+1}(\alpha)=\sum_{k=1}^{q}\frac{|m_k(\alpha)|}{(q-k+2)!}\left(\prod_{i=k}^{q}\alpha+i-1\right). \tag{24}$$

On a alors le développement asymptotique suivant :

$$\forall \alpha>0, \quad \forall N\geq 2, \quad \forall q \in \mathbb{N}^*, \quad \sum_{n=N}^{\infty}\frac{1}{n^{\alpha+1}}=\sum_{k=0}^{q-1}\frac{m_{k+1}(\alpha)}{N^{\alpha+k}}+\frac{\widetilde{m}_{q+1}(\alpha,N)}{(N-1)^{q+k}}, \tag{25}$$

où, pour tout $N\geq 2$,

$$|\widetilde{m}_{q+1}(\alpha,N)|\leq m_{q+1}(\alpha). \tag{26}$$

Remarque 5. La formule (25) peut s'obtenir de façon explicite. On considère les nombres de Bernoulli $(b_n)_{n\in\mathbb{N}}$ définis par

$$\forall t \in \mathbb{C}, \quad \frac{t}{e^t-1}=\sum_{n=0}^{\infty}\frac{t^n}{n!}b_n.$$

On pourra, pour plus d'informations, consulter l'exercice 26 de la page 146 de [AF89]. Les nombres de Bernoulli se calculent par récurrence grâce à

$$\begin{cases} b_0 = 1, \\ \forall n \in \mathbb{N}, \quad b_n = -\dfrac{1}{n+1} \displaystyle\sum_{k=0}^{n-1} C_{n+1}^k b_k, \end{cases}$$

En particulier, on a, par exemple,

$$\begin{aligned} b_0 &= 1, \\ b_1 &= -\frac{1}{2}, \\ b_2 &= \frac{1}{6}, \\ b_3 &= 0, \\ b_4 &= -\frac{1}{30}. \end{aligned}$$

On peut aussi montrer que

$$\forall n \in \mathbb{N}^*, \quad b_{2n+1} = 0. \tag{27}$$

En utilisant la formule sommatoire d'Euler Mac Laurin (voir les pages 210 à 219 de [Val55]), on peut expliciter les nombres $m_k(\alpha)$ de la somme de la formule (25) (sauf le reste) sous la forme

$$\begin{aligned} & m_1(\alpha) = \frac{1}{\alpha}, \\ & m_2(\alpha) = \frac{1}{2}, \\ & \forall p \geq 2, \quad m_{2p}(\alpha) = 0, \\ & \forall p \geq 1, \quad m_{2p+1}(\alpha) = \frac{(1+\alpha)(2+\alpha)(3+\alpha)...(2p-1+\alpha)}{(2p)!} b_{2p}. \end{aligned}$$

Nous donnons en annexe B trois applications de la formule (25).

Annexe B. Applications du développement asymptotique du reste de la fonction Dzêta

B.1. Accélération de convergence pour le calcul de ζ

On peut montrer, de façon classique, en comparant $1/n^s$ à une intégrale, que l'on a, pour tout $s > 1$ et pour tout $N \geq 2$,

$$\left| \zeta(s) - \sum_{n=1}^{N-1} \frac{1}{n^s} \right| \leq \frac{1}{(s-1)(N-1)^{s-1}}. \tag{28}$$

On peut montrer de la même façon que

$$\zeta(s) - \sum_{n=1}^{N-1} \frac{1}{n^s} \geq \frac{1}{(s-1)N^{s-1}}. \tag{29}$$

L'inégalité (28) implique que si l'on veut calculer $\zeta(s)$ à la précision $\varepsilon > 0$, on peut l'approcher par la somme partielle $\sum_{n=1}^{N-1} \frac{1}{n^s}$ à condition que N soit assez grand pour que

$$N \geq \left(\frac{1}{(s-1)\varepsilon}\right)^{1/(s-1)} + 1. \tag{30}$$

D'autre part, on peut réécrire (25) et (26) sous la forme

$$\forall s > 1, \quad \forall N \geq 2, \quad \forall q \in \mathbb{N}^*, \quad \left|\zeta(s) - \left(\sum_{n=1}^{N-1} \frac{1}{n^s} + \sum_{k=0}^{q-1} \frac{m_{k+1}(s-1)}{N^{k+s-1}}\right)\right| \leq \frac{m_{q+1}(s-1)}{(N-1)^{q+s-1}}. \tag{31}$$

Cela implique que si l'on veut calculer $\zeta(s)$ à la précision $\varepsilon > 0$, on peut l'approcher par $\sum_{n=1}^{N-1} \frac{1}{n^s} + \sum_{k=0}^{q-1} \frac{m_{k+1}(s-1)}{N^{k+s-1}}$ à condition que N soit assez grand pour que

$$N \geq \left(\frac{m_{q+1}(s-1)}{\varepsilon}\right)^{1/(q+s-1)} + 1. \tag{32}$$

Si s est «grand», l'estimation (30) fournit une valeur de N «pas trop grande» et permet de déterminer une approximation de $\zeta(s)$ à la précision $\varepsilon > 0$ en calculant la somme partielle $\sum_{n=1}^{N-1} \frac{1}{n^s}$. Cependant, l'approximation de $\zeta(s)$ par $\sum_{n=1}^{N-1} \frac{1}{n^s} + \sum_{k=0}^{q-1} \frac{m_{k+1}(s-1)}{N^{k+s-1}}$ est plus efficace puisque le nombre de termes pris en compte dans la somme partielle, défini par (32), est plus petit que celui défini par (30) ; cela est d'autant plus vrai que q est «grand» et que s est «petit». Autrement dit, d'après (28) et (29), le reste de la série de terme général $1/n^s$ tend lentement vers zéro ; au contraire, d'après (31), l'erreur tend vers zéro beaucoup plus rapidement. On parle donc d'accélération de convergence.
La fonction matlab `DL_restedzeta.m` permet de déterminer les coefficients $(m_i(s-1))_{1 \leq i \leq q+1}$. La fonction `approxim_dzeta.m` permet de calculer $\zeta(s)$ à la précision $\varepsilon > 0$ en utilisant (31) et (32).
On a testé cette fonction sur $\zeta(2)$, $\zeta(4)$ et $\zeta(6)$, explicitement connus et égaux à $\pi^2/6$, $\pi^4/90$ et $\pi^6/945$. On a indiqué, pour quelques valeurs de ε et pour $q = 5$, les valeurs correspondantes de N et de $|\zeta(s) - \zeta_{\text{approch}}(s)|$, fournit par `approxim_dzeta.m`, dans les tableaux 4, 5 et 6. On constate que l'erreur observée est bien inférieure à l'erreur prévue tant que ε n'est pas inférieur à la précision machine, de l'ordre de 10^{-16}.
On a aussi testé cette fonction sur $\zeta(1{,}01)$ (cette fois-ci inconnu explicitement). On a indiqué, pour quelques valeurs de ε et de q, les valeurs correspondantes de N et de $\zeta(1{,}01)$, fournit par `approxim_dzeta.m`, dans le tableau 7 page suivante. On constate sur ce tableau qu'à la précision 10^{-20}, on a $\zeta(1{,}01) \approx 100{,}5779433384968$. Selon les valeurs de $q \in \{5, 10, 20\}$, il a fallu prendre 5007, 63 ou 8 termes de la somme partielle, ce qui est raisonnable. Si on avait utilisé l'approximation sans accélération de convergence (28), il aurait fallu prendre, selon (30) pour $\varepsilon = 10^{-20}$,

$$N \geq \left(\frac{1}{(s-1)\varepsilon}\right)^{1/(s-1)} + 1 = \left(\frac{1}{0{,}01 \times 10^{-20}}\right)^{1/(0{,}01)} + 1 = \left(10^{22}\right)^{100} + 1 = 10^{2200} + 1,$$

ce qui aurait été largement impossible !

B.2. Calcul de sommes partielles

Supposons que l'on connaisse maintenant de façon explicite $\zeta(s)$. Les formules(25) et (26) permettent alors de calculer la somme partielle $\sum_{n=1}^{N} 1/n^s$ à la précision ε ; en effet, on peut réécrire (25) et (26)

ε	N	$\left\|\pi^2/6 - \zeta_{\text{approch}}(2)\right\|$
10^{-2}	2	$1,4240 \times 10^{-4}$
10^{-4}	4	$1,3411 \times 10^{-6}$
10^{-10}	38	$2,0805 \times 10^{-13}$
10^{-20}	1747	$2,6645 \times 10^{-15}$
10^{-30}	81044	$2,4202 \times 10^{-14}$

TAB. 4 – Quelques valeurs de N et de $\left|\pi^2/6 - \zeta_{\text{approch}}(2)\right|$ pour $q = 5$.

ε	N	$\left\|\pi^4/90 - \zeta_{\text{approch}}(4)\right\|$
10^{-2}	3	$9,1809 \times 10^{-6}$
10^{-4}	4	$7,4805 \times 10^{-7}$
10^{-8}	13	$2,0681 \times 10^{-11}$
10^{-20}	384	$6,6613 \times 10^{-16}$

TAB. 5 – Quelques valeurs de N et de $\left|\pi^4/90 - \zeta_{\text{approch}}(4)\right|$ pour $q = 5$.

ε	N	$\left\|\pi^6/945 - \zeta_{\text{approch}}(6)\right\|$
10^{-2}	3	$4,2310 \times 10^{-6}$
10^{-5}	5	$1,8180 \times 10^{-8}$
10^{-20}	143	$4,4408 \times 10^{-16}$

TAB. 6 – Quelques valeurs de N et de $\left|\pi^6/945 - \zeta_{\text{approch}}(6)\right|$ pour $q = 5$.

q	ε	N	$\zeta_{\text{approch}}(1,01)$
5	10^{-2}	2	$100,5778922186824$
5	10^{-10}	51	$100,5779433384966$
5	10^{-12}	127	$100,5779433384968$
5	10^{-20}	5007	$100,5779433384968$
10	10^{-20}	63	$100,5779433384968$
20	10^{-20}	8	$100,5779433384968$

TAB. 7 – Quelques valeurs de N et de $\zeta_{\text{approch}}(1,01)$.

sous la forme

$$\forall s > 1, \quad \forall N \geq 2, \quad \forall q \in \mathbb{N}^*, \quad \left|\sum_{n=1}^{N} \frac{1}{n^s} - \left(\zeta(s) + \sum_{k=0}^{q-1} \frac{m_{k+1}(s-1)}{(N+1)^{k+s-1}}\right)\right| \leq \frac{m_{q+1}(s-1)}{N^{q+s-1}}. \qquad (33)$$

Cela implique que si l'on veut calculer $\sum_{n=1}^{N}\frac{1}{n^s}$ à la précision $\varepsilon>0$, on peut l'approcher par $\zeta(s)+\sum_{k=0}^{q-1}\frac{m_{k+1}(s-1)}{N^{k+s-1}}$ à condition que N soit assez grand pour que

$$N \geq \left(\frac{m_{q+1}(s-1)}{\varepsilon}\right)^{1/(q+s-1)}. \tag{34}$$

Dans ce cas, si $\zeta(s)$ est connu, la somme $\sum_{k=0}^{q-1}\frac{m_{k+1}(s-1)}{N^{k+s-1}}$ à q termes est beaucoup plus rapide à calculer que la somme partielle $\sum_{n=1}^{N}\frac{1}{n^s}$ si N est «grand».
Par exemple, pour $s=2$ et $q=4$, la fonction `DL_restedzeta.m` fournit

$$\boxed{\forall N\geq 2,\quad \left|\sum_{n=1}^{N}\frac{1}{n^2}-\left(\frac{\pi^2}{6}-\frac{1}{(N+1)}-\frac{1}{2(N+1)^2}-\frac{1}{6(N+1)^3}\right)\right|\leq\frac{1}{3N^5}.} \tag{35}$$

On peut remarquer que le terme en $1/(N+1)^4$ est nul.
Une autre façon d'obtenir le développement asymptotique (sans la majoration du reste toutefois) est d'utiliser le petit script `DL_resteinvcarre` qui utilise le calcul formel.

B.3. **Étude de ζ en 1 et en $+\infty$**

Si on explicite les formules (25) et (26) pour $N=2$ et $q=2$, on obtient

$$\forall\alpha>0,\quad \left|F(2,\alpha+1)-\frac{1}{\alpha}-\frac{1}{2}\right|\leq\frac{(\alpha+1)(\alpha+2)}{12}.$$

Puisque $F(2,\alpha+1)=\zeta(\alpha+1)-1$, on a donc

$$\forall s>1,\quad \left|\frac{1}{s-1}\zeta(s)-\frac{1}{s-1}-\frac{3}{2}\right|\leq\frac{s(s+1)}{12}. \tag{36}$$

En particulier, pour $s\to 1$, on obtient au voisinage de 1,

$$\boxed{\zeta(s)\sim\frac{1}{s-1}.} \tag{37}$$

De la même façon, si on explicite les formules (25) et (26) pour $N=3$ et $q=1$, on obtient

$$\forall\alpha>0,\quad \left|\zeta(\alpha+1)-1-\frac{1}{2^\alpha}-\frac{1}{\alpha 3^\alpha}\right|\leq\frac{1}{2^{\alpha+2}}.$$

On en déduit donc

$$\forall s>1,\quad |\zeta(s)-1|\leq\frac{1}{2^{s+1}}+\frac{1}{2^{s-1}}+\frac{1}{s3^{s-1}}. \tag{38}$$

En particulier, pour $s\to\infty$, on obtient

$$\boxed{\lim_{s\to+\infty}\zeta(s)=1.} \tag{39}$$

Références

[AF88] J.M. Arnaudiès et H. Fraysse. *Cours de mathématiques, tome 2 : Analyse.* Dunod Université, 1988.

[AF89] J.M. Arnaudiès et H. Fraysse. *Cours de mathématiques, tome 3. Compléments d'analyse.* Dunod, 1989.

[Val55] G. Valiron. *Cours d'analyse mathématique, Théorie des fonctions.* Masson, 1955.