Composition de Mathématiques – B, Filière MP (X)

1. Présentation du sujet

L'étude de la fonction zêta ζ de Riemann, objet de la célèbre conjecture du même nom, est un thème fascinant pour ses relations avec l'arithmétique des entiers, notamment la répartition des nombres premiers. Dans ce sujet, on calcule certaines des valeurs de ζ, puis on donne une interprétation arithmétique de celles-ci.

Plus précisément, la première partie calcule les valeurs $\zeta(2)$, $\zeta(4)$ et $\zeta(6)$. Les parties II à IV du sujet démontrent que la probabilité que k entiers inférieurs à n choisis au hasard soient premiers entre eux tend vers $\frac{1}{\zeta(k)}$, lorsque n tend vers l'infini.

Le sujet commence par le calcul du développement en série entière de $x\,\mathrm{cotan}(x)$ au moyen d'une équation fonctionnelle. Ainsi, on montre qu'une série donnée et $x\,\mathrm{cotan}(x)$ vérifient la même équation fonctionnelle, puis que $x\,\mathrm{cotan}(x)$ égale cette série. On conclut alors en développant en série entière chaque terme de la série obtenue ; puis en permutant les signes somme.

Ensuite, on utilise le principe de prolongement analytique pour évaluer la série entière de $x\mathrm{cotan}(x)$ le long de $i\mathbb{R}$. On en déduit « un algorithme » pour calculer les valeurs de la fonction ζ de Riemann aux entiers pairs.

Notons qu'il est seulement demandé d'appliquer ce procédé pour $\zeta(2)$, $\zeta(4)$ et $\zeta(6)$ et non pas de le formaliser comme un algorithme.

La deuxième partie du sujet concerne l'ensemble $\mathscr{M}(\mathbb{N}^*)$ des mesures de probabilité sur $\mathbb{N}^*$ et relève davantage de l'analyse fonctionnelle. On voit $\mathscr{M}(\mathbb{N}^*)$ comme une partie de l'espace vectoriel $\mathcal{B}$ des fonctions bornées sur l'ensemble des parties de $\mathbb{N}^*$ à valeurs réelles. Cet espace vectoriel est muni de la norme infinie $\|\cdot\|_\infty$. On montre alors que la convergence d'une suite μ_n de mesures de probabilité vers une mesure de probabilité μ au sens de la norme $\|\cdot\|_\infty$ est équivalente à la convergence ponctuelle sur les singletons de $\mathbb{N}^*$. Un exemple montre ensuite que cette équivalence n'est plus valide si μ n'est plus supposée de probabilité.

Le sujet introduit alors la notion de suite *tendue* d'éléments de $\mathscr{M}(\mathbb{N}^*)$. Par un procédé diagonal de Cantor, on montre alors que toute suite tendue d'éléments de $\mathscr{M}(\mathbb{N}^*)$ admet une valeur d'adhérence dans $\mathscr{M}(\mathbb{N}^*)$. Ceci représente donc un résultat de compacité.

Dans la troisième partie on démontre un résultat de densité. Plus précisément, on part d'une loi de probabilité μ_X associée à une variable aléatoire X. On construit alors une suite de lois de probabilité μ_n telle que :

1. la suite μ_n converge vers μ_X ponctuellement sur les singletons ;
2. Pour tout n, il existe un nombre fini de nombres premiers qui apparaissent dans les décompositions en facteurs premiers des éléments du support de μ_n.

Dans la quatrième et dernière partie, on utilise les résultats théoriques des deux parties précédentes pour montrer le théorème suivant : la probabilité que k entiers inférieurs à n choisis au hasard soient premiers entre eux tend vers $\frac{1}{\zeta(k)}$ pour $n \to \infty$. Enfin, on conclut le sujet en utilisant la première partie, pour donner une expression de ces limites lorsque $k = 2$, 4 ou 6.

2. Conseils généraux pour les candidats.es

Traditionnellement, le sujet de l'épreuve de Mathématiques - B porte sur des points importants du programme de la filière MP.

La première partie du sujet met à l'honneur l'analyse : séries numériques, séries de fonctions, séries entières, fonctions continues, formules de trigonométrie... La deuxième partie touche à la topologie, à l'étude de différents modes de convergence et de leurs liens, au procédé d'extraction diagonale de Cantor. La troisième partie requiert des compétences en probabilité (transfert, lien entre probabilité et espérance, étude d'événements) et dans l'étude de limites de probabilités. La quatrième utilise des compétences en arithmétique (décomposition en produit de nombres premiers), en probabilité (calculs, limites, ...) et demande de savoir utiliser de manière opportune les résultats des parties précédentes du sujet.

Pour traiter le sujet d'une manière satisfaisante, il était nécessaire de manipuler avec soin et précision ces nombreuses notions. Le sujet a révélé des lacunes importantes chez certains candidats dont on cite quelques exemples ci-dessous :

1. La question 1a est très simple à condition de bien expliciter les étapes du raisonnement pour obtenir la convergence ponctuelle de la série.
2. Trop de candidats ont manipulé les termes généraux des séries dans les questions 1b, 1c et 2b sans apporter de justification satisfaisante à ces opérations.
3. La question 1d a été généralement mal traitée : pas de majoration sur un compact, problèmes aux bornes dans les majorations uniformes,... La question 3a a posé des difficultés similaires.
4. Les calculs de sommabilité double comme ceux de la question 5a.
5. La manipulation des séries entières aux questions 6, 7a et 7b.
6. Les doubles passages à la limite comme aux questions 12d.
7. Les propriétés qu'une probabilité doit vérifier en 13.
8. Le passage à la limite pour des suites d'événements monotones.
9. Bien qu'élémentaire, un recours au théorème de Gauss, la double inclusion de la question 17 a posé de nombreux problèmes.
10. Des calculs élémentaires de probabilités en 19, 20a et 21.

Toutes ces erreurs montrent de la part de trop nombreux candidats des bases en analyse réelle et en probabilité bien fragiles. Certaines de ces notions sont pourtant abordées dès la première année de CPGE en MPSI, et plus d'un an après, elles devraient être bien assimilées par les candidats.

Comme chaque année, il est préférable de s'attacher à traiter correctement plusieurs questions consécutives et parmi elles des questions plus difficiles, plutôt que d'essayer de survoler toutes les parties et de tenter de « grappiller » des points sur les questions les plus faciles. Le barème est établi de sorte qu'une telle stratégie est forcément vouée à l'échec. À l'opposé, passer du temps pour, par exemple, réussir à traiter correctement les questions les plus difficiles comme les 5a, 10b donnait les clés pour réussir les Parties I et II et s'assurer un nombre de points suffisant pour atteindre la barre d'admissibilité. Il est par contre tout à fait autorisé de sauter une question que l'on ne parviendrait pas à résoudre, puis d'en utiliser le résultat plus tard. Il faut alors veiller à ne pas oublier de vérifier soigneusement toutes les hypothèses requises pour appliquer ces « boîtes noires ». Enfin, le candidat a grand intérêt à lire le sujet intégralement avant de commencer à le traiter et à faire preuve de perspicacité pendant cette lecture.

Le soin apporté aux copies pose toujours problème. Il reste encore trop de candidats qui ne mettent pas en avant, dans la rédaction de leurs réponses, les arguments clés de la démonstration et qui présentent dans leur copie des calculs ou des raisonnements qui n'aboutissent pas. Nous devons donc rappeler aux candidats que l'usage d'un brouillon est indispensable afin de ne présenter sur sa copie que les étapes essentielles d'un raisonnement ou d'un calcul et de ne pas y faire figurer des arguments faux ou trop incomplets.

Enfin, la lisibilité des copies peut parfois poser un réel problème aux correcteurs. Nous rencontrons encore trop de copies remplies de ratures et/ou parfaitement illisibles du fait d'une graphie microscopique ou indéchiffrable. Dans les cas où malgré tous nos efforts de déchiffrage, certaines parties du texte restent incompréhensibles pour le correcteur, et dans le doute, les points ne sont pas attribués. Dans l'autre sens, il est évident qu'une copie bien présentée met le correcteur dans de bien meilleures dispositions au moment d'attribuer des points à une question.

3. Statistiques générales

Les résultats de l'épreuve sont en accord avec les directives statistiques générales du concours, ce qui garantit l'influence respective convenable des différentes épreuves.

- La moyenne des notes des 1429 candidats français est de 9,73/20 avec un écart-type de 3,90
- La moyenne des notes des 456 candidats internationaux est de 8,41/20 avec un écart-type de 3,93.

Nous aimerions exprimer nos impressions quant au niveau général des candidats. Ce qui suit ne repose que sur ce que nous avons pu constater lors de la correction de cette épreuve de mathématiques en particulier, mais nous pensons que cette analyse peut se retrouver sur l'ensemble des épreuves ainsi que dans d'autres filières et dans d'autres concours.

En effet, au regard de notre expérience de correcteur des années passées, nous avons pu constater une nette baisse du niveau moyen des copies. Ainsi, le bon usage de notions de base telles que limite, équivalent, continuité, quantificateur… suffit souvent à faire la différence. Le sujet en lui-même n'était pas non plus particulièrement difficile ou original et ne peut en aucun cas expliquer cette situation. A l'inverse, les conséquences de la crise sanitaire sur la formation des candidats pourraient être une explication conjoncturelle.

Nous souhaitons donc encourager les élèves de CPGE à renforcer leur compréhension des notions fondamentales du programme et à s'exercer à les manipuler avec précision et rigueur.

4. Examen détaillé des questions

Ce qui suit n'est pas un corrigé de l'épreuve mais une liste de commentaires et de remarques inspirés par la correction des copies.

Partie I

Question 1a. Une étude élémentaire de la convergence ponctuelle d'une série de fonctions par comparaison à une série de Riemann. Réussite : 97%.

Question 1b. On utilisait ici les propriétés de base de fonctions trigonométriques (pour la fonction f) et des manipulations simples des termes de la série (pour la fonction g). La question est réussie par la quasi-totalité des candidats.

Question 1c. Malgré la ressemblance avec des questions précédentes, cette question contient une subtilité. Plus précisément, la série

$$\tilde{g}(x) = \sum_{n=-\infty}^{+\infty} \frac{1}{x-n} \tag{1}$$

est *divergente*, tandis que la série

$$g(x) = \frac{1}{x} + \sum_{n=1}^{+\infty} \left(\frac{1}{x+n} + \frac{1}{x-n} \right) \tag{2}$$

converge absolument selon la question 1a. Le choix dans l'ordre de sommation de la série (1) est essentiel pour la transformer en une série convergente (2).

Une des manières licites d'obtenir le résultat demandé, est de passer aux sommes partielles de rang N dans la série (2), faire les calculs nécessaires, et de faire tendre N vers l'infini. Cette remarque s'applique également à la question 2b. Réussite : 35%.

Question 1d. Tandis que la continuité de la fonction f découle des propriétés élémentaires des fonctions trigonométriques standards, la continuité de la fonction D demande la continuité de la fonction g.

La démonstration de cette dernière continuité sur $\mathbb{R}\backslash\mathbb{Z}$ utilise le théorème standard sur les séries de fonctions et nécessite de montrer la convergence localement uniforme. Le chapitre sur les suites/séries de fonctions de tout cours d'analyse contient des nombreux exemples de ce type.

La manipulation des inégalités que nécessite toute preuve de convergence uniforme pose des difficultés considérables à une majorité de candidats. Certains majorent, par exemple, la borne supérieure d'une valeur absolue de fonctions par une valeur négative. Réussite : 45%.

Question 2a. Cette question se fait par un calcul élémentaire utilisant les formules trigonométriques standards. La réponse étant donnée dans l'énoncé, il est important de rappeler les formules utilisées à chaque étape du calcul. Réussite : 34%.

Question 2b. La bonne manière de traiter cette question était d'obtenir l'identité fonctionnelle demandé pour les sommes partielles de la série (2), et ensuite passer à la limite. Une fois de plus, les transformations par un réarrangement « arbitraire » des termes de la série (2) ne sont pas acceptables, cf. la discussion de la question 1c. Réussite : 24 %.

Question 3a. Par périodicité, il suffit de montrer la continuité en 0. Des développements limités des fonctions f et g au point $x_0 = 0$ permettaient de conclure. Une erreur commune à cette question, qui pourtant est enseignée dans toutes les classes préparatoires, consistait à ajouter des équivalents de manière illicite. Ainsi, on voit encore trop souvent des erreurs du type : « f et g sont équivalentes à $\frac{1}{x}$ en 0 donc $f - g$ tend vers 0 ».

Il convient en outre d'observer qu'il est (en principe) impossible de répondre à cette question en utilisant l'équation fonctionnelle

$$D(x/2) + D((1 + x)/2) = 2D(x), \quad x \in \mathbb{R}$$

et la continuité de la fonction $D(x)$ sur $\mathbb{R}\backslash\mathbb{Z}$ seulement. Réussite : 30%.

Question 3b. Le résultat de cette question découle facilement de la question 3a et du théorème de Weierstrass sur les bornes d'une fonction continue sur un segment. Réussite : 92%.

Question 4. En grande partie, cette question représente la synthèse des questions 1a-3b. En effet, on peut voir que

$$M = \sup_{x\in[0,1]} \tilde{D}(x) = \lim_{n\to+\infty} \tilde{D}(\alpha/2^n) = \tilde{D}(0) = 0.$$

Alors, l'imparité de la fonction $\tilde{D}$ montrée à la question 1b permettait de conclure que $\tilde{D} \equiv 0$ sur $\mathbb{R}$ tout entier. Enfin, la formule demandée est une conséquence immédiate de cette égalité.

Une erreur courante dans cette question consistait à conclure la nullité de D de l'annulation de son sup, confondant sup de D et sup de $|D|$. Réussite : 57%.

Question 5a. Dans cette question, il fallait utiliser la formule obtenue à la question précédente, observer que

$$\frac{x^2}{x^2 - n^2} = -\sum_{k=1}^{+\infty} \frac{x^{2k}}{n^{2k}}, \quad |x| < 1,$$

et permuter ensuite l'ordre de sommation dans la somme double obtenue. Naturellement, le changement d'ordre de sommation devait être méticuleusement justifié selon les résultats du cours (les théorèmes de Fubini, de sommation par paquets). Réussite : 25%.

Question 5b. Utilisation simple de la représentation des fonctions trigonométriques par l'exponentielle complexe. Réussite : 83%.

Cette question donnait un coup d'arrêt du travail sur la Partie I pour un bon nombre de candidats, qui passaient ensuite aux Parties II et III.

Question 6. Cette question nécessitait un raisonnement subtil sur les séries entières. Il s'agissait en effet d'identifier deux séries entières en utilisant l'unicité du développement tout en justifiant précisément les convergences. C'est ce que l'on appelle le principe de prolongement analytique. Réussite : 11%.

Question 7a. L'écriture de h comme fraction donnait le caractère C^∞ sur $\mathbb{R}^*$. Son écriture comme série entière donnait ce résultat au voisinage de 0. La formule de Taylor et le DSE donnaient les dérivées successives en 0. Réussite : 61%.

Question 7b. Une question plutôt subtile. La formule de Cauchy pour le produit des séries entières de la question 6 permettait de répondre. La formule de Leibniz pouvait aussi être utilisée. Réussite : 11%.

Question 7c. Question calculatoire basée sur la question 7b. Il s'agissait de retrouver les valeurs numériques :

$$\zeta(2) = \frac{\pi^2}{6}, \qquad \zeta(4) = \frac{\pi^4}{90}, \qquad \zeta(6) = \frac{\pi^6}{945}.$$

De nombreux candidats ont fait le choix délibérés de ne répondre à cette question que partiellement, pour gagner du temps. Réussite : 33%.

Partie II

Pour un ensemble dénombrable $E = (x_i)_{i\in\mathbb{N}^*}$ indexé par $\mathbb{N}^*$, on définit l'espace $\mathcal{B}(\mathcal{P}(E), \mathbb{R})$ des fonctions bornées sur les parties de E à valeurs réelles. L'ensemble $\mathcal{M}(E)$ des mesures de probabilité sur E est vu comme une partie de $\mathcal{B}(\mathcal{P}(E), \mathbb{R})$. Dans cette partie, on considère et compare diverses notions de convergence pour des suites de fonctions dans $\mathcal{B}(\mathcal{P}(E), \mathbb{R})$ ou $\mathcal{M}(E)$.

Question 8a. Question élémentaire pour se familiariser avec les notations de l'énnoncé. Réussite : 56%.

Question 8b. Les vérifications étant triviales, il s'agit d'énumérer les propriétés exigées d'une norme. Une question de cours en somme. Réussite : 86%.

Question 9. La convergence en norme dans $\mathcal{B}(\mathcal{P}(E), \mathbb{R})$ implique la convergence simple des fonctions sur E associées. Preuve directe avec des epsilons. Réussite : 74%.

Question 10a. Réponse attendue en deux temps. D'abord, existence de F_ϵ en identifiant $\mu(E)$ a une série convergente. Puis, existence de N_ϵ par convergence d'une somme finie de suites convergentes. Ce dernier point de vue économique n'a que très peu été utilisé par les candidats qui ont majoritairement essentiellement refait la preuve du résultat sur les sommes finies de suites convergentes. Réussite : 59 %.

Question 10b. L'inégalité triangulaire et le passage au complémentaire pour une mesure de probabilité étaient les ingrédients à utiliser ici. Réussite : 43 %.

Question 10c. Ceci est la synthèse des questions 9-10b. Réussite : 61 %.

Question 11. Cette question teste la faculté du candidat à appliquer les résultats théoriques des questions 9-10c à un cas concret. On devait montrer par l'absurde que la suite (δ_k) ne converge pas dans $\mathcal{B}(\mathcal{P}(E), \mathbb{R})$. La converge ponctuelle montrait que la seule limite possible était 0. Mais celle-ce ne convenait pas car δ_k est de norme 1. Réussite : 35 %.

Question 12a. Une récurrence et le théorème de Bolzano-Weierstrass étaient les ingrédients de toute bonne réponse. Réussite : 46 %.

Question 12b. Application directe du résultat du cours sur les suites extraites d'une suite convergente. Réussite : 38 %.

Question 12c. Question facile si l'on ne commettait par l'erreur de croire que ψ est une fonction composée. Réussite : 38 %.

Question 12d. Dans le terme de gauche de la deuxième inégalité il y a deux infinis. Dans ce cas, il est important de raisonner par étapes. On part de

$$\sum_{i=1}^{N} \mu_{\psi(k)}(x_i) \leq 1.$$

Dans un premier temps, on fait tendre k vers l'infini, N étant fixé. Puis, dans un second temps, on fait tendre N vers l'infini. Réussite : 15 %.

Question 12e. Le fait que la mesure limite μ_∞ est une mesure de probabilité, vient immédiatement du fait que la suite (μ_n) est tendue. Ensuite, on utilise la question 10c et on observe que dans la présente situation, la convergence simple des mesures implique la convergence en norme de $\mathcal{B}(\mathcal{P}(E), \mathbb{R})$. Réussite : 13 %.

Partie III

Dans cette partie, on fait le lien avec les résultats précédents et quelques notions de la théorie des nombres. Seulement un quart des candidats se sont aventurés au-delà de cette frontière.

Question 13. Il s'agit essentiellement de connaître la définition d'un mesure de probabilité, notamment la σ-additivité. Réussite : 27 %.

Question 14. La façon la plus simple de faire c'est de montrer que la fonction

$$|\mathbf{1}_{X \in A} - \mathbf{1}_{X \in B}|$$

est majorée par $\mathbf{1}_{X \neq Y}$ et de passer à l'intégrale probabiliste. Réussite : 40 %.

Question 15a. Évident. Réussite 40 %.

Question 15b. Application directe de la définition de la variable aléatoire L. Réussite : 32 %.

Question 15c. On utilise les questions 14 et 15a pour majorer la norme de la différence des mesures. La continuité de probabilité d'une suite décroissante d'événements permet alors de montrer que le majorant tend vers 0. Réussite : 21 %.

Question 16. Une question de synthèse qui utilise non seulement tous les résultats de la Partie III mais aussi la question 9. Réussite : 12 %.

Partie IV

Comme mentionné ci-dessus, très peu d'élèves ont tâché de travailler cette partie d'une manière « sérieuse ». On y traite des variables aléatoires provenant de la théorie des nombres du point de vue de la Partie III.

Question 17a. La vérification de l'égalité ensembliste est élémentaire quoi qu'assez méticuleuse. En particulier, la double inclusion ("$A \subset B$" et "$B \subset A$") était requise. Les arguments d'arithmétiques devaient aussi apparaître clairement. Il convient d'observer aussi que la réunion dans la partie droite de l'égalité donnée est disjointe. Réussite : 11 %.

Question 17b. Une récurrence à l'aide de la question 17a. Réussite : 5 %.

Question 17c. La réponse découle de la question précédente, de la continuité de la mesure de probabilité pour une suite décroissante d'événements, et de l'égalité

$$\mu(\{x\}) = \lim_{n\to+\infty} \mu(\mathbb{N}^* r \backslash (\bigcup_{i=1}^{n} \mathbb{N}^* r p_i)).$$

Réussite : 4 %.

Question 18. La suite $(\mu_{X_n})_n$ étant tendue, par la question 12e, elle a une valeur d'adhérence dans $\mathcal{B}(\mathcal{P}(\mathbb{N}^*), \mathbb{R})$. Par ailleurs, on montre que la valeur d'adhérence en question est unique. Alors, un raisonnement par l'absurde permet de montrer que la suite converge. Réussite : 2 %.

Question 19. Il est immédiat que

$$P(r|X_n^{(i)}) = \frac{1}{n}\lfloor \frac{n}{r} \rfloor.$$

L'indépendance des variables aléatoires $(X_n^{(i)})_{i=1,\ldots,s}$ permet d'obtenir une formule similaire pour la probabilité $P(r|Z^{(s)_n})$, et le passage à la limite $n \to +\infty$ donne la conclusion souhaitée. Réussite : 5 %.

Question 20a. La formule pour $P(k|Z)$ voulue découle directement de la définition de μ_s. Réussite : 7 %.

Question 20b. Cette question fait la synthèse des questions précédentes et de la question 18. Notamment, on observe que

$$\lim_{n\to+\infty} P(r|Z^{(s)_n}) = P(r|Z),$$

la suite $(\mu_{Z_n^{(s)}})_n$ est tendue, et la convergence souhaitée découle de la question 18. Réussite : 2 %.

Question 21. Cette question était l'objectif du problème. Étant donné l'égalité

$$P_n(s) = P(Z_n^{(s)} = 1),$$

on obtient la réponse à l'aide des questions 20b et 7c. Réussite : 2 %.