Composition de Mathématiques – B, Filière MP (X)

Cette année encore, le concours d'admission de l'École Polytechnique s'est déroulé à l'issue d'une année scolaire marquée par des conditions difficiles liées à la crise sanitaire.

1. Présentation du sujet

Le sujet de l'épreuve de Mathématiques - B 2021 est découpé en quatre parties. La Partie I est indépendante des Parties II et III et la Partie IV fait la synthèse des résultats obtenus dans les parties I et III.

La première partie du problème porte sur l'étude de la loi ζ en probabilités. Rappelons qu'à partir de toute série positive convergente, de somme non nulle, on peut définir une loi de probabilité discrète. Ici, on considère la série numérique $\sum n^s$ pour $s > 1$ un nombre réel. Cette loi ζ permet de relier l'indépendance d'évènements, propriété de multiplicativité des probabilités, à la multiplicativité des fonctions arithmétiques via la décomposition en facteurs premiers d'un entier. Cette approche permet de retrouver dès la question 2 du sujet une expression classique de la fonction ζ de Riemann comme produit infini sur l'ensemble des nombres premiers. Elle permet aussi de démontrer l'indépendance des images d'une variable aléatoire suivant la loi ζ par les valuations p-adiques.

Ensuite, on introduit la fonction g définie pour tout entier $n \geq 1$ par

$$g(n) = r_1(n) - r_3(n) \qquad \text{où} \qquad r_i(n) = \mathrm{Card}\{d \in \mathbb{N} \mid d \equiv i[4] \text{ et } d|n\},\ i \in \{0, 1, 2, 3\}.$$

Cette fonction est arithmétique et, à l'aide du caractère de Dirichlet χ_4 et de la décomposition en facteurs premiers des entiers, on montre à la fin de cette Partie I que l'espérance de l'image d'une variable aléatoire suivant la loi ζ de paramètre $s > 1$ par la fonction g est égale à la fonction β de Dirichlet en s.

Mentionnons que la Partie I illustre de manière élémentaire le fait que l'approche probabiliste est déjà fortement présente dans un grand nombre de résultats d'arithmétique. Nous pensons là aux résultats qui relèvent de la théorie probabiliste des nombres après Hardy, Ramanujan, Erdös, Kăc et bien d'autres. Ces résultats portent sur des théorèmes de répartition asymptotique de nombres premiers ou d'autres familles d'entiers comme dans le Théorème des nombres premiers, celui de la progression arithmétique de Dirichlet ou encore dans la preuve de l'existence de progressions arithmétiques de longueurs arbitraires dans l'ensemble des nombres premiers par Green et Tao.

La Partie II du problème a pour objectif d'établir un développement de la fonction $x \mapsto \sin(\pi x)$ en produit de Weierstrass via des techniques élémentaires d'analyse réelle et bien entendu sans recourir à l'analyse complexe.

La Partie III démarre par l'introduction de la fonction Γ d'Euler sous forme de produit infini. Il est amusant de constater que ce produit infini fait intervenir la constante γ d'Euler. La suite de la partie nous fait démontrer l'identité classique entre les fonctions β et Γ d'Euler via la propriété de log-convexité de la fonction Γ. Enfin nous déduisons de cette expression et des résultats de la Partie II une expression sous forme d'intégrale impropre de la fonction $x \mapsto \frac{\pi}{\sin(\pi x)}$ sur l'intervalle $]0, 1[$.

Le problème se termine par une quatrième partie consacrée à une méthode de calcul des valeurs aux points entiers impairs positifs de la fonction β de Dirichlet. Cela se fait de manière itérative en montrant que ces valeurs sont, à un facteur simple près, les coefficients du développement en série entière de la fonction $x \mapsto \frac{1}{\cos(x)}$. Cette relation s'obtient à l'aide d'un développement en série de fonctions rationnelles de $x \mapsto \frac{\pi}{\sin(\pi x)}$ et en utilisant l'expression de cette fonction trouvée à la partie IV. Ces méthodes de calcul sont habituelles en combinatoire où l'on emploie les séries génératrices pour calculer des quantités difficiles à dénombrer directement.

2. Conseils généraux pour les candidats.es

Comme chaque année, il est préférable de traiter correctement plusieurs questions consécutives et parmi elles des questions plus difficiles, plutôt que d'essayer de survoler toutes les parties en tentant 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. A l'opposé, passer du temps pour réussir à traiter correctement des questions plus difficiles permet de s'assurer un nombre de points suffisant pour atteindre une note supérieure à 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 vérifier soigneusement toutes les hypothèses requises pour appliquer ces « boîtes noires ».

Le soin apporté à certaines 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 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.

Il est aussi demandé aux candidats une lecture plus attentive des énoncés et des questions posées. Certains ne répondent qu'à la « moitié » de la question posée sans dire un seul mot a propos du reste. Préciser que l'on admet le résultat correspondant au reste de la question facilite la lecture des correcteurs. De même, les phrases du type « ...et à partir de là il est facile de voir que... » alors que justement, le point délicat reste à traiter, n'apportent pas de points. Il est inutile de compter sur un manque de vigilance de la part du correcteur.

Enfin, la lisibilité des copies peut parfois poser un réel problème aux correcteurs. Nous rencontrons encore trop de copies remplies de râtures 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 1 398 candidats ont une moyenne de 10,11/20 avec un écart-type de 4,25.

4. Examen détaillé des questions

Ce qui suit n'est pas un corrigé de l'épreuve mais une liste de commentaires et de remarques des correcteurs, question par question.

Partie I

Question 1a. Il s'agit ici essentiellement d'utiliser la définition de la probabilité d'un événement. Taux de réussite : 80%.

Question 1b. Certains candidats vérifiaient l'indépendance de deux événements donnés seulement, ce qui n'est pas suffisant pour répondre à la question. Taux de réussite : 79%.

Question 2a. Il faut utiliser la question 1a et l'indépendance mutuelle des événements $\{p_i \nmid X\}$. Taux de réussite : 92%.

Question 2b. Il convient de passer à la limite dans la question précédente. La justification utilise le résultat sur la continuité d'une famille décroissante d'événements, ce qui a été négligé par un certain nombre d'élèves. Taux de réussite : 73%.

Question 3a. Ceci est probablement la première question qui ait posé des difficultés à beaucoup de candidats. Une des erreurs rencontrées consistait à traiter les événements $\{p_k^n|X\}$ et $\{p_k^{n+1} \nmid X\}$ comme des événements indépendants, ce qui n'est pas le cas.

Une manière simple surmonter cette difficulité est d'observer que

$$\{\nu_{p_k}(X) = n\} = A_n \backslash A_{n+1} := \{\nu_{p_k}(X) \geq n\} \backslash \{\nu_{p_k}(X) \geq n + 1\},$$

$A_{n+1} \subset A_n$, et ensuite de prendre la différence des probabilités correspondantes. On peut aussi revenir à la formule de transfert et faire des calculs explicites. Taux de réussite : 46%.

Question 3b. Cette question assez technique a donné le coup d'arrêt au travail sur la Partie I d'une majorité de candidats.

Certains ont répondu à cette question en utilisant « la formule de crible de Poincaré » qui est hors programme en filière MP. On pouvait l'appliquer donc à condition de l'avoir démontrer au préalable. Taux de réussite : moins de 3%.

Question 3c. Malgré la rédaction un peu encombrante, cette question est un corollaire simple des questions 1a et 3b. Taux de réussite : 15%.

Question 4a. Soient m et n deux entiers naturels non nuls et premiers entre eux ; il convient de remarquer que tout diviseur d de mn admet une unique écriture comme $d = d_1d_2$, où d_1 divise m et d_2 divise n.

La suite de cette question se fait par un calcul combinatoire à l'aide de congruences ; ceci a mis en difficulté un bon nombre d'élèves. Taux de réussite : 25% .

Question 4b. Un calcul direct des quantités $r_1(n), r_3(n)$ selon la congruence de n modulo 4 suffit à répondre à la question. Taux de réussite : 59%.

Question 5. Malgré sa simplicité, cette question est cruciale pour la bonne compréhension de passages à la limite dans les questions 7b.-9c.

Il s'agit d'une version du théorème de convergence dominée (= TCD) pour des fonctions $f : \mathbb{N}^* \to \mathbb{R}$. La référence au TCD « classique » pour des applications $f : \mathbb{R} \to \mathbb{R}$ ne suffisait pas pour la réponse complète à la question. Les candidats devaient ou bien déduire la version du théorème de convergence requise en passant par des fonctions « en escalier » bien choisies, ou bien en imitant la démonstration du TCD « classique ».

Enfin, un appel simple « au passage à la limite sous l'intégrale grace à la convergence uniforme » ne suffit pas non plus, et ceux qui ont choisi ce chemin, devaient détailler les étapes de leur argument.

Ce type de question fait partie « des bases » d'analyse et de probabilités, qui doivent être largement maîtrisées par les candidats au concours d'admission de l'X. Taux de réussite : 35%.

Question 6a. L'astuce consiste à remarquer que

$$\sum_{n\geq 1} r(n)n^{-s} = \left(\sum_{n\geq 1} n^{-s}\right)^2.$$

Taux de réussite : 23%.

Question 6b. Application immédiate de la question précedente. Taux de réussite : 56%.

Question 7a. Application directe de la définition de la convergence simple. Taux de réussite : 54%.

Question 7b. Cette question fait la synthèse des questions 6b, 7a, et l'application de la version de TCD obtenue à la question 5. Taux de réussite : 25%.

Question 8a. Cette question se fait par la formule de transfert à l'aide des questions 3a et 4b. Plusieurs candidats ont été bloqués par le calcul de la somme d'une série qui était pourtant relativement simple. Taux de réussite : 40%.

Question 8b. La question est similaire à la précédente. Taux de réussite : 17%.

Question 8c. Dans l'état d'esprit des questions 8a et 8b, on obtient une expression simple pour

$$E(g(p_k^{\nu_{p_k}(X)}))$$

en s'aidant de la question 4b. La question 7b justifie le passage à la limite. Taux de réussite : 17%.

Question 9a. La multiplicativité de l'application χ_4 fournit la réponse. Taux de réussite : 11%.

Question 9b. Les questions 9b et 9c présentent la synthèse de la majorité des résultats obtenus au préalable ; en particulier elles demandent un certain recul sur l'intégralité de la Partie I du sujet.

Par exemple, la question 9b requiert l'usage de questions 3c, 5, 7a et 9a. Taux de réussite : 11%.

Question 9c. La solution à cette question découle de la formule de transfert ainsi que des questions 8c et 9b. Taux de réussite : 11%.

Partie II

L'objectif de la Partie II est de démontrer la formule d'analyse classique suivante

$$\sin(\pi x) = \pi x \lim_{n\to+\infty} \prod_{k=1}^{n} \left(1 - \frac{x^2}{k^2}\right).$$

On y utilise exclusivement des techniques élémentaires de manipulation des sommes, produits, leurs minorations/majorations et convergences.

Question 10a. Cette question est un calcul classique relevant de la formule de Moivre. Il est fort dommage qu'un nombre considérable de candidats n'aient pas su le faire aboutir. Taux de réussite : 45%.

Question 10b. L'astuce consiste ici à re-écrire le polynôme obtenu à la question précédente d'une manière multiplicative. Taux de réussite : 32%.

Question 10c. La question s'obtient en partant des questions 10a et 10c par un changement de variable simple. Taux de réussite : 71%.

Question 11a. La question consiste à étudier la convergence des suites $(u_{m,n}(x))_{n>m}$ et $(v_{m,n}(x))_{n>m}$ *selon* n et pour m *fixé*.

L'étude de la convergence de produits *finis* $(u_{m,n}(x))_{n>m}$ se fait par passage aux équivalents, et la convergence de produits $(v_{m,n}(x))_{n>m}$ en découle, car

$$v_{m,n}(x) = \frac{\sin(\pi x)}{u_{m,n}(x)}.$$

Taux de réussite : 21%.

Question 11b. Dans cette question, on procède à la majoration et minoration fines du produit $v_{m,n}(x)$. Le point de départ de ce travail est l'inégalité « des cordes » pour la fonction sinus :

$$\frac{2}{\pi} \leq \sin x \leq x, \quad x \in [0, \frac{\pi}{2}].$$

Une fois l'inégalité double annoncée dans la question obtenue, on utilise le fait bien connu que la limite

$$\lim_{n\to+\infty} \prod_{k=1}^{n} a_k, \quad a_k > 0,$$

existe si et seulement si la série $\sum_{k=1}^{+\infty} \log a_k$ converge. Taux de réussite : 28%.

Question 11c. Ceci est une question de synthèse récapitulant la majorité de résultats obtenus dans la Partie II du sujet. Notamment, à partir de la relation

$$\sin(\pi x) = u_{m,n}(x) v_{m,n}(x),$$

on effectue d'abord le passage à la limite par rapport à n. La deuxième étape consiste à passer à la limite par rapport à m en utilisant la question 11b. Il convient ici de prendre soin de passer à la limite successivement lorsque n, puis m tendent vers l'infini.

De plus, les cas $x \in \mathbb{Z}$ devait être traité séparemment. Taux de réussite : 21%.

Partie III

Cette partie contient un dévelopement sur les propriétés de la fonction $\Gamma(x), x > 0$

$$\Gamma(x) = \frac{e^{-\gamma x}}{x} \lim_{n\to+\infty} \prod_{k=1}^{n} \frac{e^{x/k}}{1 + x/k}, \tag{1}$$

où γ est la constante d'Euler.

Malgré le fait qu'un bon nombre de candidats a essayé de traiter cette partie, peu y ont obtenu des résultats raisonnables.

Notons une démarche qui a porté un grand préjudice à un bon nombre de candidats. Bien que la fonction Γ puisse être définie par la formule intégrale

$$\Gamma(x) = \int_0^{+\infty} t^{x-1} e^{-t}\, dt, \quad x > 0, \tag{2}$$

ce n'est pas le choix fait dans l'énoncé. En effet, ici la fonction Γ est définie à partir d'un produit infini. Ainsi, toute réponse aux questions utilisant (2) sans démonstration est hors sujet.

Question 12. C'est un cas de figure classique d'étude de la convergence simple d'une série (ou d'un produit). Il convient d'utiliser la positivité des termes du produit, de passer au log et d'utiliser des équivalents standards (dépendant d'un paramètre $x, x > 0$) pour obtenir le résultat voulu. Taux de réussite : 34%.

Question 13. Cette question est un calcul direct qui rélève d'une simplification judicieuse de facteurs et de la définition de la constante d'Euler

$$\gamma = \lim_{n\to+\infty} \left(\sum_{k=1}^{n} \frac{1}{k} - \log n \right).$$

Taux de réussite : 12%.

Question 14a. La réponse à cette question est une application immédiate du théorème sur la dérivabilité d'une série de fonctions. Puisqu'il s'agit de la deuxième dérivée $(\log \Gamma)''(x), x > 0$, il fallait appliquer ledit théorème deux fois consécutivement. La vérification minutieuse des hypothèses requises par le théorème était attendue. Taux de réussite : 10%.

Question 14b. La question découle de la précédente et de l'observation que la série représentant la fonction $(\log\Gamma)''$ converge *normalement* sur tout intervalle $]a,+\infty[$, $a>0$. Taux de réussite : 25%.

Question 15a. Plusieurs candidats ont identifié cette question comme pouvant être « grappillée », or sa solution demande un raisonnement assez fin.

La 1-périodicité se déduit facilement de la question 13 et des hypothèses sur la fonction f. Pour la convexité de la fonction $(\log(f/\Gamma))$, on utilise d'une manière essentielle la 1-périodicité démontrée précédemment qui permet d'écrire

$$\begin{aligned}(\log(f/\Gamma))''(x) &= (\log f)''(x)-(\log\Gamma)''(x)=\lim_{n\to+\infty}[(\log f)''(x+n)-(\log\Gamma)''(x+n)]\\ &= \lim_{n\to+\infty}(\log f)''(x+n)\geq 0,\end{aligned}$$

d'où le résultat. Taux de réussite : 2%.

Question 15b. Dans une moindre mesure, la remarque faite à la question 15a, s'applique ici aussi. Beaucoup trop de candidats donnaient des solutions fantasques (ou des explications non-rigoureuses) à un résultat demandé plutôt transparent.

En effet, la fonction convexe et 1-périodique ne peut qu'être constante, car sa dérivée est alors croissante, 1-périodique et donc constante... Enfin, il convient de calculer la valeur $(f/\Gamma)(1)=1$ comme à la question 13 pour s'assurer que $f(x)=\Gamma(x)$. Taux de réussite : 15%.

Question 16. C'est probablement la question la plus difficile et la plus technique du sujet ; très peu de candidats ont essayé de l'aborder, et seulement 2% des candidats ont eu 40% des points à cette question.

Après avoir vérifié la convergence de l'intégrale en question pour $x>0$ et $a>0$, il faut considérer la fonction

$$f(x)=\frac{\Gamma(x+a)}{\Gamma(a)}\int_0^{+\infty}\frac{t^x}{(1+t)^{x+a+1}}\,dt.$$

Ensuite, on doit vérifier (en passant, entre outre, par l'inégalité de Cauchy-Schwarz) que la fonction f satisfait les hypothèses énoncées après la question 14, et, enfin, identifier les fonctions $f(x)$ et $\Gamma(x)$ à l'aide des questions 15a et 15b.

Question 17. La question consiste à démontrer que

$$\int_0^{+\infty}\frac{t^{x-1}}{1+t}\,dt=\Gamma(x)\Gamma(1-x)=\frac{\pi}{\sin(\pi x)} \tag{3}$$

en utilisant les calculs dans l'esprit de la question 13. En posant $a=1-x$ dans la formule de la question 16, on identifie la partie de gauche de l'égalité donnée ci-dessus, avec l'intégrale de la question 17.

La question n'a été réussie par aucun candidat.

Partie IV

Par manque de temps, cette partie a été abordée par très peu d'élèves.

Question 18a. Pour répondre à cette question, on doit utiliser les résultats de la question précédente, à savoir l'égalité de la fonction $\pi/\sin(\pi x)$ à l'intégrale de la relation (3). Il convient de re-écrire l'intégrale comme

$$\int_0^{+\infty}\frac{t^{x-1}}{1+t}\,dt=\int_0^{1}\frac{t^{x-1}}{1+t}\,dt+\int_1^{+\infty}\frac{t^{x-1}}{1+t}\,dt.$$

En justifiant soigneusement la convergence de sommes et d'intégrales obtenues, on développe ensuite le terme $1/(1+t)$ de la première intégrale en série entière et on effectue l'intégration « terme par terme ». Pour la deuxième intégrale, on procéde au changement de variable $u = 1/t$ et on continue le calcul d'une manière analogue à la première intégrale. Taux de réussite : 1%.

Question 18b. C'est une bonne idée de s'appuyer sur la question 18a et d'effectuer le changement de variable $x \mapsto x + \frac{1}{2}$. L'étape suivante consiste à re-écrire la série obtenue comme une série double en développant les expressions du type

$$\frac{(-1)^n}{2n+1+2x}, \quad \frac{(-1)^n}{2n+1-2x},$$

en série entière par rapport à $(2x/(2n+1))$. Il convient de justifier les convergences de séries, la positivité de rayons de convergence des séries entières, le changement d'ordre de sommation, etc.

1% des candidats ont eu 30% des points attribués à cette question.

Question 18c. Cette question traite des propriétés de base de la série entière donnée. Après un changement de variable $x \mapsto \pi x$, le résultat découle de l'identification des coefficients de la série avec les dérivées correspondantes de sa somme. Taux de réussite : 3%.

Question 19a. Dans l'état d'esprit de la question précédente, il faut exploiter le lien entre les dérivées $v^{(2k)}(0)$ de la fonction v de la question 18c, et les dérivées de la fonction

$$t \mapsto \frac{e^{it}}{\cos(t)}, \quad t \in]-\frac{1}{2}, \frac{1}{2}[,$$

obtenus à l'aide de formule de Leibniz.

Les valeurs E_0, E_2 et E_4 se calculent par une récurrence simple à partir de la relation de la première partie de la question.

2% de candidats ont comptabilisé 20% de la note sur cette question.

Question 19b. A l'aide de la question 9c, on se rend compte que l'espérance $E(g(x))$, où X est une variable aléatoire de loi ζ et de paramètre $s = 2k+1$, est exactement la quantité à gauche de l'égalité obtenue à la question 18c. La réponse à la question est alors donnée par la question 19a. Taux de réussite : 1%.