Devoir des vacances d'été – corrigé

Table des matières

1 Commentaires

Problème d'algèbre linéaire

Une des intentions du problème d'algèbre linéaire est clairement mentionnée dans le sujet : vous donner un aperçu de la réduction des endomorphismes ou des matrices. Le type de réduction le plus fréquemment rencontré en mathématiques, et qui occupera la majeure partie du chapitre d'algèbre linéaire cette année, est la réduction à une matrice diagonale que nous pratiquons dans la deuxième partie du problème. Notez bien ce qu'on y fait : on ramène la matrice A à une matrice diagonale *via* une relation de similitude, parce qu'il est bien plus facile *a priori* d'extraire les racines carrées d'une matrice diagonale, simplement en extrayant les racines carrées de ses coefficients diagonaux. Mais il y a une très grande subtilité : pourquoi les racines carrées d'une matrice diagonale seraient-elles nécessairement diagonales ? Il y a alors une *deuxième réduction*, plus difficile parce qu'on l'effectue sur des matrices inconnues (les racines carrées de A, qu'on cherche justement à déterminer). Constatez la très grande différence de méthodologie entre la réduction de A et celle de M, où M est une racine carrée de A.

Dans ce problème, la réduction pratique d'une matrice A à une matrice diagonale est guidée. Vous aurez plus tard des outils pour y parvenir seuls. On retiendra néanmoins que la clé est de produire une base $(\vec{e}_1, \dots, \vec{e}_n)$ constituée de vecteurs vérifiant $f(\vec{e}_i) = \lambda_i \vec{e}_i$, où f est l'endomorphisme canoniquement associé à A et les λ_i des scalaires. C'est la préoccupation de tout algébriste qui cherche à réduire : produire une base sur laquelle l'action de f est « simple ». Et peut-on faire plus simple qu'agir sur un vecteur en le dilatant ? Ce procédé fournit une matrice diagonale après changement de base. C'est toutefois une approche qui peut échouer, et c'est pourquoi nous proposons une autre réduction des matrices nilpotentes dans la partie III. On l'appelle **réduction de Jordan**.

Cet objectif de réduction est un prétexte pour introduire la notion « d'espace **cyclique** », qui a déjà un intérêt significatif à elle seule (ce n'est pas au programme mais cela fait régulièrement l'objet de problèmes). Il s'agit d'un espace vectoriel engendré par une famille de la forme $(\vec{x}, f(\vec{x}), f^2(\vec{x}), \dots)$, où $\vec{x}$ et f sont respectivement un vecteur et un endomorphisme fixés. Bien qu'une telle famille ne soit pas aussi appréciable que celles dont nous parlions plus haut (encore que !), on note que l'action de f sur cette famille est très simple : l'image par f d'un vecteur donne le vecteur suivant de la famille. Un intérêt de ce type d'espace vectoriel est que, connaissant l'action de f sur UN SEUL vecteur $\vec{x}$, on peut en déduire (en réitérant l'image, puis par linéarité) l'action de f sur TOUT VECTEUR. Vous avez déjà vu comment la donnée d'une base peut souvent suffire à en déduire un résultat valable pour tout vecteur : c'est le même principe ici (et on le retrouvera en théorie des groupes, lorsqu'on démontrera un résultat pour tout élément d'un groupe en le démontrant uniquement avec un système de générateurs).

Un autre lieu commun des mathématiques, illustré en plusieurs endroits de ce problème, est le « **principe de conjugaison** ». À savoir : si x est un objet mathématique « d'un certain type » (cette année, en pratique, ce sera souvent : une permutation ou une application linéaire, souvent dotée d'une certaine interprétation géométrique concrète), et si g est un objet mathématique appartenant à un groupe (et pour lequel le produit par x admet un sens), alors gxg^{-1} est du « même type » que x, et les caractéristiques « géométriques » (ensemble de points fixes...) de gxg^{-1} s'obtiennent à partir de celles de x, en prenant leur image par g (si cela a un sens). Cela revêt de très nombreuses significations, selon qu'on soit en théorie des groupes, des espaces vectoriels, etc. Par exemple, lorsque nous parlerons d'ordre d'un élément d'un groupe, nous verrons que si x est d'ordre d, alors gxg^{-1} l'est également ; dans un tout autre contexte, si f est une symétrie par rapport à F et parallèlement à G, et si g est un automorphisme, alors $g \circ f \circ g^{-1}$ est une symétrie par

rapport à $g(F)$ et parallèlement à $g(G)$. Une application du principe de conjugaison est de passer d'un objet x « compliqué » à un autre objet du même type gxg^{-1} dont l'étude est plus familière. Nous l'illustrerons en des contextes variés cette année (en fait, la théorie de la réduction peut être présentée comme un de ses nombreux cas particuliers).

Dans le cadre de ce devoir : le principe de conjugaison nous permet d'étudier les puissances et les racines carrées de A en nous ramenant, *via* le principe de conjugaison, aux puissances et racines carrées d'une matrice diagonale (uniquement dans la deuxième partie), ou d'une matrice plutôt simple de la forme $J_{k_1,\dots,k_r}$ (appelée **matrice de Jordan**), simple parce que son exponentiation se fait par simple décalage « vers le haut » de ses coefficients non nuls.

Ce qu'on retiendra en bref. Comment résoudre $M^2 = A$ en se ramenant à deux matrices diagonales. Utiliser la formule du changement de base. Matrice de f dans une base de la forme $(f^i(\vec{x}))_i$. Indice de nilpotence et noyaux itérés des matrices $J_{k_1,\dots,k_r}$. Une matrice nilpotente est d'indice de nilpotence inférieur ou égal à son ordre.

Questions faciles ou classiques à retravailler

— PREMIÈRE PARTIE : toutes les questions ;

— DEUXIÈME PARTIE : toutes les questions ;

— TROISIÈME PARTIE : les questions 8 et 9 ;

— QUATRIÈME PARTIE : les questions 17, 18, 20 et 21.

Problème d'analyse

Ce problème est une adaptation inédite (à ma connaissance) du théorème de la progression arithmétique de Dirichlet, de sorte que certains cas particuliers puissent être traités avec les outils de MPSI (certes subtilement). Même si je sais que le lien entre la fonction dzêta de Riemann et la répartition des nombres premiers est intrigant pour beaucoup d'étudiants de MP, il doit encore rester mystérieux pour le moment : il faut des outils d'analyse complexe pour en saisir toute la richesse, et ce n'est pas au programme des classes préparatoires (même si nous l'effleurerons avec la théorie des séries entières).

Ainsi ce problème est majoritairement un prétexte pour vous initier à l'analyse telle qu'on la pratique quand on atteint une certaine maturité scientifique : les fonctions les plus intéressantes des mathématiques ont souvent une description sous forme d'une somme (à support infini) de fonctions. Il faut donc savoir manipuler les opérations classiques de l'analyse avec de telles sommes : dérivation, intégration, limite, pour citer les principales. Ce problème vous fait concrètement étudier les difficultés qui se présentent : 1° lorsqu'on veut intégrer une somme à support infini en intégrant d'abord les sommes partielles, puis en passant à la limite (pour obtenir : $-\ln(1-x) = \sum_{k=1}^{+\infty} \frac{x^k}{k}$), 2° lorsqu'on veut montrer la continuité en un point (calcul de : $\lim_{s\to 1^+} \mathrm{L}_\chi(s)$) en faisant apparaître toutes les quantités « petites » en jeu (notamment : le reste de la série dont la somme égale $\mathrm{L}_\chi(s)$) et des sommes à support fini (puisque dans ce cas, on sait qu'il n'y a pas de problème : une somme de fonctions continues est continue), 3° lorsqu'on veut calculer une limite, dans le cas où on s'attend à ce qu'elle soit infinie. Les élèves observateurs se doivent de remarquer que, aussi bien pour 1° que pour 2°, **la clé est la majoration du reste** ! (Plus précisément, pour 1°, il s'agit de majorer l'intégrale du reste.) Ce sera une préoccupation extrêmement fréquente cette année.

Vous aurez cette année un théorème pour obtenir des limites de sommes (à support infini) de fonctions, lorsque la limite est FINIE. Je voulais profiter de ce problème pour vous faire étudier le cas non couvert explicitement par le théorème : que faire pour calculer une limite du type : $\lim_{x\to a} \sum_{n=0}^{+\infty} \bigstar$, quand la limite est infinie ? Il s'avère que la monotonie est une hypothèse très commode pour répondre à cette question. Vous deviez remarquer (et je l'utilise abondamment dans le corrigé) que dans ce cas très particulier, tout se passe « comme on pense », que la limite soit finie ou non d'ailleurs, puisqu'il suffit d'écrire : $\lim_{x\to a} \sum_{n=0}^{+\infty} \bigstar = \sum_{n=0}^{+\infty} \lim_{x\to a} \bigstar$. Il vous sera très utile de retenir la démarche pour la reproduire durant l'année.

Ce devoir vous permet également de voir comment s'utilise le **théorème de sommation par paquets**, puisqu'il intervient souvent (et avec des choix différents de paquets), ou comment peuvent s'utiliser les « autres » conséquences du théorème des séries alternées (puisque souvent, en dehors de la conclusion portant sur la convergence de la série étudiée, ses conséquences peuvent vous paraître nébuleuses).

Ce qu'on retiendra en bref. Limite de ζ en 1 (et équivalent asymptotique si demandé). Produit eulérien de ζ et L_χ. Importance de la convergence absolue. Interversion somme-limite lorsque la limite est infinie. Majoration du reste d'une série alternée pour obtenir la continuité en un réel d'une somme de série de fonctions.

Questions faciles ou classiques à retravailler

— PREMIÈRE PARTIE : toutes les questions sauf la dernière et, éventuellement, la question 4 (on aura de meilleurs outils pour la traiter ultérieurement) ;

— DEUXIÈME PARTIE : les questions 7 et 8 ; aussi, ne pas hésiter à observer où l'emploi du théorème spécial des séries alternées permet de simplifier la résolution des questions 10 et 13.

2 Corrigé

Problème d'algèbre linéaire

Première partie : propriétés de base

1. Soient A et B deux matrices semblables de $\mathrm{M}_n(\mathbb{C})$. Il existe donc $P \in \mathrm{GL}_n(\mathbb{C})$ telle que : $A = PBP^{-1}$. Fixons une telle matrice P. On note alors que pour toute matrice $M \in \mathrm{M}_n(\mathbb{C})$, on a :

$$M \in \mathcal{R}(A) \iff M^2 = A \iff P^{-1}M^2P = B \overset{(*)}{\iff} \left(P^{-1}MP\right)^2 = B \iff P^{-1}MP \in \mathcal{R}(B),$$

l'équivalence $(*)$ se justifiant par le calcul élémentaire suivant : $\left(P^{-1}MP\right)^2 = \left(P^{-1}MP\right)\left(P^{-1}MP\right) = P^{-1}M^2P$. On montrerait de même que M appartient à $\mathcal{R}(B)$ si et seulement si $PMP^{-1} \in \mathcal{R}(A)$, de sorte que les applications suivantes :

$$\Phi : \begin{cases} \mathcal{R}(A) & \to & \mathcal{R}(B) \\ M & \mapsto & P^{-1}MP \end{cases}, \quad \Psi : \begin{cases} \mathcal{R}(B) & \to & \mathcal{R}(A) \\ M & \mapsto & PMP^{-1} \end{cases}$$

soient correctement définies. Elles sont trivialement réciproques l'une de l'autre, donc bijectives. Ainsi $\mathcal{R}(A)$ et $\mathcal{R}(B)$ sont effectivement en bijection, d'où le résultat.

Questions à se poser, réflexes à acquérir. Noter là une 1re illustration du principe de conjugaison.

2. Soient $M \in \mathcal{R}\left(0_{\mathrm{M}_n(\mathbb{C})}\right)$ et $P \in \mathrm{GL}_n(\mathbb{C})$. On a : $\left(PMP^{-1}\right)^2 = PM^2P^{-1}$, et comme $M \in \mathcal{R}\left(0_{\mathrm{M}_n(\mathbb{C})}\right)$ cela implique : $\left(PMP^{-1}\right)^2 = P \times 0_{\mathrm{M}_n(\mathbb{C})} \times P^{-1} = 0_{\mathrm{M}_n(\mathbb{C})}$. On a ainsi montré : $\forall P \in \mathrm{GL}_n(\mathbb{C})$, $PMP^{-1} \in \mathcal{R}\left(0_{\mathrm{M}_n(\mathbb{C})}\right)$. Or toute matrice semblable à M est de cette forme, d'où le résultat voulu.
Un usage soigneux de la 1re question permet également de répondre très rapidement à celle-ci.

3. Vérification facile. Revoir si besoin le cours de 1re année, où il fut démontré que $\sim$ définit une relation d'équivalence sur $\mathrm{M}_n(\mathbb{C})$. On l'a simplement restreinte à $\mathcal{R}\left(0_{\mathrm{M}_n(\mathbb{C})}\right)$.

Deuxième partie : un premier résultat théorique, avec un peu de réduction

4. Soit $M \in \mathcal{R}(A)$. Alors : $AM = M^2M = M^3 = MM^2 = MA$. Or, si l'on définit les coefficients de A et M ainsi : $A = ((\lambda_i\delta_{i,j}))_{1\leqslant i,j\leqslant n}$ (le symbole de Kronecker est là pour rappeler que les coefficients de A sont nuls hors de la diagonale), et : $M = ((m_{i,j}))_{1\leqslant i,j\leqslant n}$, alors on a :

$$AM = \begin{pmatrix} \lambda_1 m_{1,1} & \cdots & \lambda_1 m_{1,n} \\ \vdots & \ddots & \vdots \\ \lambda_n m_{n,1} & \cdots & \lambda_n m_{n,n} \end{pmatrix} = ((\lambda_i m_{i,j}))_{1\leqslant i,j\leqslant n}, \quad MA = \begin{pmatrix} \lambda_1 m_{1,1} & \cdots & \lambda_n m_{1,n} \\ \vdots & \ddots & \vdots \\ \lambda_1 m_{n,1} & \cdots & \lambda_n m_{n,n} \end{pmatrix} = ((\lambda_j m_{i,j}))_{1\leqslant i,j\leqslant n}.$$

En effet, si $(i,j) \in [\![1,n]\!]^2$, alors le coefficient d'indice (i,j) de AM est : $\sum_{k=1}^{n} \lambda_i \delta_{i,k} m_{k,j} = \lambda_i m_{i,j}$. Calcul analogue pour le coefficient d'indice (i,j) de MA.
Déduisons-en que M est une matrice diagonale. Considérons $(i,j) \in [\![1,n]\!]^2$ avec i et j distincts. L'égalité $AM = MA$ donne, en comparant les coefficients d'indice (i,j) : $\lambda_i m_{i,j} = \lambda_j m_{i,j}$. Or par hypothèse de l'énoncé on a, pour i et j distincts : $\lambda_i \neq \lambda_j$, donc l'égalité précédente implique : $m_{i,j} = 0$. Ceci vaut pour tout $(i,j) \in [\![1,n]\!]^2$ avec i et j distincts, donc M est une matrice diagonale.
On en déduit que l'égalité $M^2 = A$ équivaut, en comparant les coefficients diagonaux de M^2 et A, à : $\forall i \in [\![1,n]\!]$, $m_{i,i}^2 = \lambda_i$. Or, pour tout $i \in [\![1,n]\!]$, l'équation $x^2 = \lambda_i$ d'inconnue $x \in \mathbb{C}$ admet exactement deux solutions si λ_i est non nul (nous noterons μ_i et $-\mu_i$ ces deux solutions ci-dessous), et une unique solution si λ_i est nul (l'unique solution est 0). En résumé, on a montré que :

— *si* $\lambda_i \neq 0$ *pour tout* $i \in [\![1,n]\!]$, alors :

$$\mathcal{R}(A) = \{((\varepsilon_i \mu_i \delta_{i,j}))_{1\leqslant i,j\leqslant n} \mid (\varepsilon_i)_{1\leqslant i\leqslant n} \in \{-1,1\}^n\} = \left\{ \begin{pmatrix} \varepsilon_1\mu_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \varepsilon_n\mu_n \end{pmatrix} \mid (\varepsilon_i)_{1\leqslant i\leqslant n} \in \{-1,1\}^n \right\}$$

(le raisonnement ci-dessus ne démontre que l'inclusion directe, mais l'inclusion réciproque est triviale), donc il y a une bijection naturelle entre $\{-1,1\}^n$ et $\mathcal{R}(A)$, et on en déduit :

$$\operatorname{card}(\mathcal{R}(A)) = 2^n \ ;$$

— *s'il existe* $i \in [\![1,n]\!]$ *tel que* $\lambda_i = 0$: alors ce i est unique puisque les λ_i sont tous distincts ; quitte à remplacer A par $P_\tau A P_\tau^{-1}$, où $P_\tau \in \mathrm{M}_n(\mathbb{C})$ est la matrice associée à la transposition $(i\ n)$, on peut supposer que $i = n$ (c'est-à-dire $\lambda_n = 0$) ; puisque $\mathcal{R}(A)$ et $\mathcal{R}(P_\tau A P_\tau^{-1})$ sont en bijection d'après la première question, cela ne changera pas le cardinal cherché ; cette hypothèse étant faite, on a alors :

$$\mathcal{R}(A) = \left\{ \begin{pmatrix} \varepsilon_1\mu_1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \varepsilon_{n-1}\mu_{n-1} & 0 \\ 0 & \cdots & 0 & 0 \end{pmatrix} \mid (\varepsilon_i)_{1\leqslant i\leqslant n-1} \in \{-1,1\}^{n-1} \right\}$$

ce dont on déduit, *via* une bijection naturelle entre $\{-1,1\}^{n-1}$ et $\mathcal{R}(A)$:

$$\operatorname{card}(\mathcal{R}(A)) = 2^{n-1}.$$

En résumé :

$$\operatorname{card}(\mathcal{R}(A)) = \begin{cases} 2^n & \text{si : } \forall i \in [\![1,n]\!],\ \lambda_i \neq 0, \\ 2^{n-1} & \text{si : } \exists i \in [\![1,n]\!],\ \lambda_i = 0. \end{cases}$$

Si l'on veut éviter une disjonction de cas, il suffit de remarquer qu'on a aussi montré :

$$\operatorname{card}(\mathcal{R}(A)) = 2^{\operatorname{rg}(A)}.$$

Autre démonstration que M est diagonale, plus géométrique. La stratégie qui suit est plus proche en esprit des raisonnements que nous ferons cette année. Soient f_A et f_M les endomorphismes de $\mathbb{C}^n$ canoniquement associés à A et M respectivement. Soit $(\vec{e}_1, \ldots, \vec{e}_n)$ la base canonique de $\mathbb{C}^n$. Comme : $(f_M)^2 = f_A$, un calcul similaire à celui ci-dessus démontre que f_A et f_M commutent. On a alors, pour tout $i \in [\![1,n]\!]$:

$$\lambda_i f_M(\vec{e}_i) = f_M(\lambda_i \vec{e}_i) = f_M(f_A(\vec{e}_i)) = f_A(f_M(\vec{e}_i)),$$

donc : $\forall i \in [\![1,n]\!]$, $f_M(\vec{e}_i) \in \ker(f_A - \lambda_i \mathrm{Id}_{\mathbb{C}^n}) = \operatorname{Vect}(\vec{e}_i)$ (la détermination de ce noyau est très facile si l'on constate que $A - \lambda_i \mathrm{I}_n$ est échelonnée pour tout $i \in [\![1,n]\!]$, avec une i^{e} colonne nulle). Ainsi $f_M(\vec{e}_i)$ et $\vec{e}_i$ sont proportionnels pour tout $i \in [\![1,n]\!]$, ce qui signifie exactement que M est diagonale.

Questions à se poser, réflexes à acquérir.

— Se convaincre qu'il n'y avait rien d'évident à ce que M soit diagonale, même si A l'est : trouver des contre-exemples dans le cas où les coefficients diagonaux ne sont pas tous distincts ;

— Sait-on écrire les « bijections naturelles », se convaincre qu'elles le sont, et détecter ce qui fait que la première bijection n'en est plus une si $\lambda_i = 0$ se produit ?

— Comprend-on bien la réécriture avec le rang ? Pourquoi cette écriture peut-elle être préférable, outre le fait qu'elle dispense d'une distinction de cas ?

— Comparer les mérites des deux démonstrations. Que peut-on dire si les coefficients diagonaux ne sont plus supposés distincts ? Regarder ce que cela change aussi bien par l'argument matriciel que celui géométrique.

— Voyez que $X^2 - A \in (\mathrm{M}_n(\mathbb{C}))[X]$ peut avoir plus de racines que son degré : pourquoi n'est-ce pas absurde ?

5. D'après les questions 1 et 4, on a : $\mathrm{card}\,(\mathcal{R}(A)) = 2^{\mathrm{rg}(A)}$, parce que le rang est invariant par similitude.

6. Résoudre $AX = 0_{\mathrm{M}_{2,1}(\mathbb{C})}$, d'inconnue $X \in \mathrm{M}_{2,1}(\mathbb{C})$, montre que $\vec{x}_1 = (1,1)$ et $\vec{x}_2 = (1,-1)$ conviennent. **Remarque pour les cinq-demis, et plus tard pour toute la classe.** Puisque A est une matrice symétrique réelle, on sait qu'elle est diagonalisable par le théorème spectral et que ses sous-espaces propres (il y en a nécessairement deux : pourquoi ?) sont de dimension 1 et supplémentaires orthogonaux. On en déduit que : 1° il suffit de déterminer un seul vecteur propre, pour chaque valeur propre, afin d'engendrer tout le sous-espace propre associé, 2° dès qu'on a déterminé un sous-espace propre, il suffit de prendre son supplémentaire orthogonal pour avoir le second. Or on trouve immédiatement un vecteur propre associé à 3 en remarquant que la somme des coefficients de chaque ligne égale 3 : cela assure que (1,1) engendre $\ker(f - 3\mathrm{Id}_{\mathbb{C}^2})$. En considérant le vecteur $(1,-1)$, clairement orthogonal à (1,1), on obtient le sous-espace propre associé à $\ker(f - \mathrm{Id}_{\mathbb{C}^2})$. Ainsi aucune résolution de système linéaire n'est nécessaire pour trouver $\vec{x}_1$ et $\vec{x}_2$.

7. On a : $\begin{vmatrix} 1 & 1 \\ 1 & -1 \end{vmatrix} = -2 \neq 0$, donc $(\vec{x}_1, \vec{x}_2)$ est une base de $\mathbb{C}^2$. Comme : $f(\vec{x}_1) = 3\vec{x}_1$, et : $f(\vec{x}_2) = \vec{x}_2$, on a : $\mathrm{M}_{(\vec{x}_1,\vec{x}_2)}(f) = \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix}$. Posons alors : $P = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$. C'est la matrice de passage de la base canonique $\mathcal{B}_{can}$ dans la base $(\vec{x}_1, \vec{x}_2)$. En appliquant la formule du changement on obtient, du fait que $A = \mathrm{M}_{\mathcal{B}_{can}}(f)$:

$$A = P \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix} P^{-1},$$

ce qu'il fallait démontrer. Ainsi $P^{-1}AP$ est une matrice diagonale dont les coefficients diagonaux sont distincts, donc d'après la question 4 on a :

$$\mathcal{R}(P^{-1}AP) = \left\{ \begin{pmatrix} \varepsilon_1\sqrt{3} & 0 \\ 0 & \varepsilon_2 \end{pmatrix} \mid (\varepsilon_1, \varepsilon_2) \in \{-1,1\}^2 \right\}.$$

D'après les deux bijections de la question 1, on conclut :

$$\begin{aligned} \mathcal{R}(A) &= \left\{ P \begin{pmatrix} \varepsilon_1\sqrt{3} & 0 \\ 0 & \varepsilon_2 \end{pmatrix} P^{-1} \mid (\varepsilon_1, \varepsilon_2) \in \{-1,1\}^2 \right\} \\ &= \left\{ \frac{1}{2}\begin{pmatrix} \sqrt{3}+1 & \sqrt{3}-1 \\ \sqrt{3}-1 & \sqrt{3}+1 \end{pmatrix}, \frac{1}{2}\begin{pmatrix} \sqrt{3}-1 & \sqrt{3}+1 \\ \sqrt{3}+1 & \sqrt{3}-1 \end{pmatrix}, -\frac{1}{2}\begin{pmatrix} \sqrt{3}+1 & \sqrt{3}-1 \\ \sqrt{3}-1 & \sqrt{3}+1 \end{pmatrix}, -\frac{1}{2}\begin{pmatrix} \sqrt{3}-1 & \sqrt{3}+1 \\ \sqrt{3}+1 & \sqrt{3}-1 \end{pmatrix} \right\}. \end{aligned}$$

Remarque pour les cinq-demis, et plus tard pour toute la classe. Il suffit de noter que $\vec{x}_1$ et $\vec{x}_2$ sont des vecteurs propres associés à des valeurs propres différentes, pour en déduire que la famille $(\vec{x}_1, \vec{x}_2)$ est libre. Par un argument de cardinalité, c'est une base de $\mathbb{C}^2$.

Questions à se poser, réflexes à acquérir.

— En peu de mots, comment résumer cette 2^e partie ? Quelle est la stratégie globale ? En quoi illustre-t-elle l'intérêt de la réduction des matrices, dont j'ai parlé en juin et plus longuement dans *Présentation des chapitres de MP* ?

— Autre illustration du principe de conjugaison : comparer les solutions de $AX = \lambda X$, pour $\lambda = 1$ puis $\lambda = 3$, à celles de $\begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix} X = \lambda X$ pour ces mêmes valeurs de λ. Qu'observe-t-on ? Que proposer comme généralisation ?

Troisième partie : réduction des matrices nilpotentes

★ 8. L'ensemble $\left\{i \in \mathbb{N} \mid f^i(\vec{x}) = \vec{0}\right\}$ est une partie de $\mathbb{N}$, non vide puisqu'elle contient l'indice de nilpotence de f (qui est aussi celui de A), donc elle admet un plus petit élément. Ainsi $h_{\vec{x}}$ existe bien.
Montrons que la famille $\mathcal{B}_{\vec{x}}$ est une base de $\mathcal{F}_{\vec{x}}$. Commençons par montrer qu'elle est génératrice. Soit $\vec{y} \in \mathcal{F}_{\vec{x}}$. Par définition de ce sous-espace vectoriel, il existe $d \in \mathbb{N} \setminus \{0\}$ et $(a_k)_{0\leqslant k\leqslant d-1} \in \mathbb{C}^d$ tels que :

$$\vec{y} = \sum_{k=0}^{d-1} a_k f^k(\vec{x}).$$

Si $d \leqslant h_{\vec{x}}$, alors on a clairement : $\vec{y} \in \mathrm{Vect}(\mathcal{B}_{\vec{x}})$. Si $d > h_{\vec{x}}$, alors il suffit de noter que l'on a :

$$\forall k \in \mathbb{N},\ \left(k \geqslant h_{\vec{x}} \Longrightarrow f^k(\vec{x}) = \vec{0}\right). \qquad (*)$$

On peut le démontrer aisément par récurrence sur k : pour $k = h_{\vec{x}}$ c'est vrai par définition de $h_{\vec{x}}$, et pour l'hérédité il suffit d'écrire : $f^{k+1}(\vec{x}) = f(f^k(\vec{x}))$, et de constater ainsi que la proposition au rang k implique la proposition au rang $k+1$ (du fait que $f(\vec{0}) = \vec{0}$ par linéarité de f). Ainsi, si $d \geqslant h_{\vec{x}}$, on écrit :

$$\vec{y} = \sum_{k=0}^{d-1} a_k f^k(\vec{x}) = \sum_{k=0}^{h_{\vec{x}}-1} a_k f^k(\vec{x}) + \sum_{k=h_{\vec{x}}}^{d-1} a_k f^k(\vec{x}) \overset{(*)}{=} \sum_{k=0}^{h_{\vec{x}}-1} a_k f^k(\vec{x}) \in \mathrm{Vect}(\mathcal{B}_{\vec{x}}).$$

On a bien montré : $\forall \vec{y} \in \mathcal{F}_{\vec{x}}$, $\vec{y} \in \mathrm{Vect}(\mathcal{B}_{\vec{x}})$ (s'il vous choque d'avoir f^k au lieu de $f^{h_{\vec{x}}-1-k}$ dans le terme général de la somme, il suffit d'utiliser le changement d'indice $k \mapsto h_{\vec{x}} - 1 - k$). Donc : $\mathcal{F}_{\vec{x}} \subseteq \mathrm{Vect}(\mathcal{B}_{\vec{x}})$, ce qui achève de démontrer que $\mathcal{B}_{\vec{x}}$ est une famille génératrice de $\mathcal{F}_{\vec{x}}$.
Montrons à présent que c'est une famille libre de $\mathcal{F}_{\vec{x}}$. Soit $(a_k)_{0\leqslant k\leqslant h_{\vec{x}}-1} \in \mathbb{C}^{h_{\vec{x}}}$. On suppose : $\sum_{k=0}^{h_{\vec{x}}-1} a_k f^{h_{\vec{x}}-1-k}(\vec{x}) = \vec{0}$. Montrons : $\forall k \in [\![0, h_{\vec{x}} - 1]\!]$, $a_k = 0$. Une façon de procéder est *via* une récurrence soignée (soyez prudents dans ce cas-là, puisque la proposition voulue ne doit pas être démontrée pour tout $k \in \mathbb{N}$). Il revient au même de raisonner par l'absurde (ce qui ne doit pas vous étonner si vous vous souvenez comment fut démontré le principe de récurrence ; je choisis de procéder ainsi pour m'affranchir de la subtilité dont je parle ci-dessus). Supposons donc au contraire qu'il existe $k \in [\![0, h_{\vec{x}} - 1]\!]$ tel que : $a_k \neq 0$, et soit k_0 l'entier maximal pour cette propriété. Cette maximalité assure que $a_k = 0$ pour tout $k \in [\![k_0 + 1, h_{\vec{x}} - 1]\!]$. La relation de dépendance linéaire ci-dessus se réécrit donc :

$$\sum_{k=0}^{k_0} a_k f^{h_{\vec{x}}-1-k}(\vec{x}) = \vec{0}.$$

Prenons l'image par f^{k_0} de ce vecteur. La linéarité de cette application implique :

$$\sum_{k=0}^{k_0} a_k f^{h_{\vec{x}}+(k_0-1-k)}(\vec{x}) = \vec{0}.$$

Pour tout $k \in [\![0, k_0 - 1]\!]$, on a : $h_{\vec{x}} + (k_0 - 1 - k) \geqslant h_{\vec{x}}$, donc d'après $(*)$ on a l'identité : $\forall k \in [\![0, k_0 - 1]\!]$, $f^{h_{\vec{x}}+(k_0-1-k)}(\vec{x}) = \vec{0}$. L'égalité ci-dessus se simplifie donc, pour donner : $a_{k_0} f^{h_{\vec{x}}-1}(\vec{x}) = \vec{0}$. Or : $f^{h_{\vec{x}}-1}(\vec{x}) \neq \vec{0}$, donc finalement : $a_{k_0} = 0$, ce qui contredit la définition de k_0. Par l'absurde, on a montré : $\forall k \in [\![0, h_{\vec{x}} - 1]\!]$, $a_k = 0$. Ainsi la seule relation de dépendance linéaire est la relation triviale, donc la famille $\mathcal{B}_{\vec{x}}$ est libre.
Étant une famille libre et génératrice de $\mathcal{F}_{\vec{x}}$, c'est une base de cet espace vectoriel : d'où le résultat.

Questions à se poser, réflexes à acquérir.

— Bien vérifier l'indexation après le changement d'indice $k \mapsto h_{\vec{x}} - 1 - k$;

— Si ma démonstration de l'indépendance linéaire vous paraît obscure : la reprendre, éventuellement l'écrire avec peu de termes (ou avec des points de suspension informels) pour comprendre ce que je fais ; le choix de prendre l'image par f^{k_0} (pourquoi cette puissance ?) doit être compris ; en quoi cette démonstration illustre le principe selon lequel on a souvent besoin d'autant d'équations que d'inconnues ?

— Revoir la démonstration du principe de récurrence, et son lien avec un raisonnement par l'absurde ; pourquoi m'intéressé-je au plus *grand* k tel que $a_k \neq 0$, et non au plus petit ?

★ 9. L'hypothèse $h_{\vec{x}} = n$ assure que $\mathcal{B}_{\vec{x}}$ est une famille libre de $\mathbb{C}^n$ (d'après la question précédente) de cardinal maximal, donc c'est une base de $\mathbb{C}^n$. On a : $f\left(f^{h_{\vec{x}}-1}(\vec{x})\right) = f^{h_{\vec{x}}}(\vec{x}) = \vec{0}$ par définition de $h_{\vec{x}}$, et : $\forall k \in [\![1, h_{\vec{x}} - 1]\!]$, $f\left(f^{h_{\vec{x}}-1-k}(\vec{x})\right) = f^{h_{\vec{x}}-k}(\vec{x}) = f^{h_{\vec{x}}-1-(k-1)}(\vec{x})$. On en déduit que la matrice représentative de f dans la base $\mathcal{B}_{\vec{x}}$ est :

$$\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & & & 0 & 1 \\ 0 & 0 & \cdots & 0 & 0 \end{pmatrix} = J_n.$$

Soit P la matrice de passage de la base canonique dans la base $\mathcal{B}_{\vec{x}}$. D'après la formule du changement de base, on a : $A = PJ_nP^{-1}$, ce qu'il fallait démontrer.

Questions à se poser, réflexes à acquérir.
- Ce qu'on a démontré admet-il une réciproque ? c'est-à-dire, si la matrice de f dans une base convenable est J_n, cette base est-elle de la forme $\mathcal{B}_{\vec{x}}$ pour un certain $\vec{x}$?
- Question très ouverte : quel peut être l'intérêt, dans ce problème ou ailleurs, d'avoir une base de $\mathbb{C}^n$ de la forme $\mathcal{B}_{\vec{x}}$, pour étudier un endomorphisme f, non nécessairement nilpotent ? (C'est une question qui vous poursuivra peut-être durant toute votre carrière mathématique : nul besoin d'y répondre immédiatement.)

10. Le fait que $\mathcal{F}_{\vec{x}} \cap \ker(f)$ soit un sous-espace vectoriel de $\mathbb{C}^n$ ne pose pas de problème, puisque c'est une intersection de sous-espaces vectoriels. Pour en donner une base et la dimension, nous allons l'expliciter. Soit $\vec{y} \in \mathcal{F}_{\vec{x}} \cap \ker(f)$. Il existe $(a_k)_{0\leqslant k\leqslant h_{\vec{x}}-1} \in \mathbb{C}^{h_{\vec{x}}}$ tel que : $\vec{y} = \sum_{k=0}^{h_{\vec{x}}-1} a_k f^{h_{\vec{x}}-1-k}(\vec{x})$. De plus, on a : $f(\vec{y}) = \vec{0}$. En combinant ces deux égalités, on obtient par linéarité de f : $\sum_{k=0}^{h_{\vec{x}}-1} a_k f^{h_{\vec{x}}-k}(\vec{x}) = \vec{0}$. Le premier terme de la somme est nul, étant donné que : $f^{h_{\vec{x}}}(\vec{x}) = \vec{0}$. Par conséquent, quitte à faire un changement d'indice, l'égalité précédente équivaut à : $\sum_{k=0}^{h_{\vec{x}}-2} a_{k+1} f^{h_{\vec{x}}-k-1}(\vec{x}) = \vec{0}$. Or la famille $\mathcal{B}_{\vec{x}} = \left(f^{h_{\vec{x}}-1-k}(\vec{x})\right)_{0\leqslant k\leqslant h_{\vec{x}}-1}$ est libre d'après la question 8, donc la famille extraite $\left(f^{h_{\vec{x}}-1-k}(\vec{x})\right)_{0\leqslant k\leqslant h_{\vec{x}}-2}$ l'est également. On en déduit : $\forall k \in [\![0, h_{\vec{x}} - 2]\!]$, $a_{k+1} = 0$, et donc : $\vec{y} = \sum_{k=0}^{h_{\vec{x}}-1} a_k f^{h_{\vec{x}}-1-k}(\vec{x}) = a_0 f^{h_{\vec{x}}-1}(\vec{x}) \subseteq \text{Vect}\left(f^{h_{\vec{x}}-1}(\vec{x})\right)$. Ce raisonnement montre que l'on a : $\mathcal{F}_{\vec{x}} \cap \ker(f) \subseteq \text{Vect}\left(f^{h_{\vec{x}}-1}(\vec{x})\right)$. Pour l'inclusion réciproque, on note que l'inclusion $\text{Vect}\left(f^{h_{\vec{x}}-1}(\vec{x})\right) \subseteq \mathcal{F}_{\vec{x}}$ découle de la définition de $\mathcal{F}_{\vec{x}}$, tandis que l'inclusion $\text{Vect}\left(f^{h_{\vec{x}}-1}(\vec{x})\right) \subseteq \ker(f)$ est une conséquence du calcul suivant, déjà apparu tantôt : $f\left(f^{h_{\vec{x}}-1}(\vec{x})\right) = f^{h_{\vec{x}}}(\vec{x}) = \vec{0}$. Ainsi : $\text{Vect}\left(f^{h_{\vec{x}}-1}(\vec{x})\right) \subseteq \mathcal{F}_{\vec{x}} \cap \ker(f)$, donc :

$$\text{Vect}\left(f^{h_{\vec{x}}-1}(\vec{x})\right) = \mathcal{F}_{\vec{x}} \cap \ker(f).$$

Ainsi la famille $\left(f^{h_{\vec{x}}-1}(\vec{x})\right)$ engendre $\mathcal{F}_{\vec{x}} \cap \ker(f)$, et c'est une famille libre en tant que famille extraite de la famille libre $\mathcal{B}_{\vec{x}}$ (comme c'est une famille constituée d'un seul vecteur, un autre argument serait simplement de constater que ce vecteur est non nul). On en déduit que $\left(f^{h_{\vec{x}}-1}(\vec{x})\right)$ est une base de $\mathcal{F}_{\vec{x}} \cap \ker(f)$, qui est donc un espace vectoriel de dimension 1.

11. **Coquille de l'énoncé.** On suppose dans les questions 11 à 13 qu'il existe $m \in \mathbb{N}\setminus\{0\}$ et $(\vec{z}_1, \ldots, \vec{z}_m) \in (\text{im}(f))^m$ tels que $\mathcal{B}_{\vec{z}_1} \cup \cdots \cup \mathcal{B}_{\vec{z}_m}$ soit une base de $\text{im}(f)$.
Par définition de F, la famille $\mathcal{B} = \mathcal{B}_{\vec{y}_1} \cup \cdots \cup \mathcal{B}_{\vec{y}_m}$ en est une famille génératrice. Il suffit donc de démontrer que c'est une famille libre. Soient $a_{1,0}, \ldots, a_{1,h_{\vec{y}_1-1}}, \ldots, a_{m,0}, \ldots, a_{m,h_{\vec{y}_m-1}}$ des nombres

complexes. On suppose :

$$\sum_{i=1}^{m}\sum_{j=0}^{h_{\vec{y}_i}-1} a_{i,j} f^{h_{\vec{y}_i}-1-j}(\vec{y}_i) = \vec{0}. \qquad (*)$$

Montrons que tous les scalaires sont nuls. Pour cela, prenons l'image par f de ce vecteur. Par linéarité, on a : $\sum_{i=1}^{m}\sum_{j=0}^{h_{\vec{y}_i}-1} a_{i,j} f^{h_{\vec{y}_i}-j}(\vec{y}_i) = \vec{0}$.

Remarquons que pour tout $i \in [\![1,m]\!]$ et pour $j=0$, on a : $f^{h_{\vec{y}_i}-j}(\vec{y}_i) = f^{h_{\vec{y}_i}}(\vec{y}_i) = \vec{0}$. D'autre part, du fait que $f(\vec{y}_i)$ soit égal à $\vec{z}_i$ pour tout $i \in [\![1,m]\!]$, on obtient : $\sum_{i=1}^{m}\sum_{j=1}^{h_{\vec{y}_i}-1} a_{i,j} f^{h_{\vec{y}_i}-1-j}(\vec{z}_i) = \vec{0}$. Or l'égalité suivante : $\forall i \in [\![1,m]\!]$, $\forall k \in \mathbb{N}\setminus\{0\}$, $f^k(\vec{y}_i) = f^{k-1}(\vec{z}_i)$, facile à démontrer, implique que l'on a : $\forall i \in [\![1,m]\!]$, $h_{\vec{y}_i} = h_{\vec{z}_i} + 1$. On en déduit que la relation ci-dessus s'écrit finalement, après le changement d'indice $j \mapsto j-1$:

$$\sum_{i=1}^{m}\sum_{j=0}^{h_{\vec{z}_i}-1} a_{i,j+1} f^{h_{\vec{z}_i}-1-j}(\vec{z}_i) = \vec{0}.$$

Or la famille $\mathcal{B}_{\vec{z}_1}\cup\cdots\cup\mathcal{B}_{\vec{z}_m}$ est libre, puisque c'est une base de $\mathrm{im}(f)$ par hypothèse. On en déduit que pour tous $i \in [\![1,m]\!]$ et $j \in [\![0, h_{\vec{z}_i}-1]\!] = [\![0, h_{\vec{y}_i}-2]\!]$, on a : $a_{i,j+1} = 0$. Ainsi la relation de dépendance $(*)$ est devenue :

$$\sum_{i=1}^{m} a_{i,0} f^{h_{\vec{y}_i}-1}(\vec{y}_i) = \vec{0}.$$

Or, comme on l'a noté plus haut, on a : $\forall i \in [\![1,m]\!]$, $h_{\vec{y}_i} - 1 = h_{\vec{z}_i} \geqslant 1$, donc en particulier : $\forall i \in [\![1,m]\!]$, $f^{h_{\vec{y}_i}-1}(\vec{y}_i) = f^{h_{\vec{z}_i}}(\vec{y}_i) = f^{h_{\vec{z}_i}-1}(\vec{z}_i)$. Autrement dit : $\sum_{i=1}^{m} a_{i,0} f^{h_{\vec{z}_i}-1}(\vec{z}_i) = \vec{0}$. On invoque encore une fois la liberté de la famille $\mathcal{B}_{\vec{z}_1}\cup\cdots\cup\mathcal{B}_{\vec{z}_m}$ pour conclure : $\forall i \in [\![1,m]\!]$, $a_{i,0} = 0$.

En conclusion, nous avons bien démontré que pour tous entiers $i \in [\![1,m]\!]$ et $j \in [\![0, h_{\vec{y}_i}-1]\!]$, on a : $a_{i,j} = 0$. L'unique relation de dépendance linéaire entre vecteurs de $\mathcal{B}$ est donc la triviale, ce qui prouve que c'est une famille libre, et plus précisément une base de F. D'où le résultat.

Une économie possible dans la rédaction. On reprend les notations ci-dessus, et on pose : $\forall i \in [\![1,m]\!]$, $\vec{x}_i = \sum_{j=0}^{h_{\vec{y}_i}-1} a_{i,j} f^{h_{\vec{y}_i}-1-j}(\vec{y}_i)$, de sorte que la relation de dépendance linéaire supposée vraie équivaille à : $\sum_{i=1}^{m}\vec{x}_i = \vec{0}$. Prendre l'image par f, et utiliser le fait que $\mathcal{B}_{\vec{z}_1}\cup\cdots\mathcal{B}_{\vec{z}_m}$ soit libre, implique (un peu comme dans le raisonnement ci-dessus) : $\forall i \in [\![1,m]\!]$, $f(\vec{x}_i) = \vec{0}$. Ainsi : $\forall i \in [\![1,m]\!]$, $\vec{x}_i \in \mathcal{F}_{\vec{y}_i} \cap \ker(f)$. D'après la question précédente (c'est là qu'elle intervient), on sait que $\vec{x}_i$ est proportionnel à $f^{h_{\vec{y}_i}-1}(\vec{y}_i)$. On retrouve ce qui est apparu vers la fin du raisonnement ci-dessus.

Questions à se poser, réflexes à acquérir.

— Vérifier ce que j'ai laissé au lecteur en exercice, notamment l'égalité : $h_{\vec{y}_i} = h_{\vec{z}_i} + 1$;

— Pourquoi ne me suis-je pas fatigué à quantifier ainsi : « soit $(a_{i,j})_{\substack{0\leqslant i\leqslant m\\ 0\leqslant j\leqslant h_{\vec{y}_i}-1}} \in \bigstar$... » ? Quelle est la difficulté rédactionnelle ? Comment pourrait-on la résoudre ? (Ne cherchez pas trop loin.)

12. Montrons d'abord que F et G sont en somme directe, en montrant qu'ils sont d'intersection triviale. On sait que $F\cap\ker(f)$ et G sont supplémentaires dans $\ker(f)$. On a donc : $(F\cap\ker(f))\cap G = \{\vec{0}\}$. Or : $G \subseteq \ker(f)$, donc : $F\cap G = F\cap(G\cap\ker(f)) = (F\cap\ker(f))\cap G = \{\vec{0}\}$. Cela démontre que F et G sont en somme directe. Montrons à présent : $\mathbb{C}^n \subseteq F+G$, l'inclusion réciproque étant triviale. Pour cela, démontrons d'abord l'égalité $f(F) = \mathrm{im}(f)$ suggérée dans l'énoncé. L'inclusion $f(F)\subseteq\mathrm{im}(f)$ est vraie car $F\subseteq\mathbb{C}^n$; démontrons que l'on a : $\mathrm{im}(f)\subseteq f(F)$. Soit $\vec{z}\in\mathrm{im}(f)$. Comme $\mathcal{B}_{\vec{z}_1}\cup\cdots\cup\mathcal{B}_{\vec{z}_m}$ est une base de $\mathrm{im}(f)$, en particulier cette famille engendre $\mathrm{im}(f)$, donc il existe des

nombres complexes $a_{1,0},\dots,a_{1,h_{\vec{z}_1-1}},\dots,a_{m,0},\dots,a_{m,h_{\vec{z}_{m-1}}}$, tels que : $\vec{z}=\sum_{i=1}^{m}\sum_{j=0}^{h_{\vec{z}_i}-1}a_{i,j}f^{h_{\vec{z}_i}-1-j}(\vec{z}_i)=\sum_{i=1}^{m}\sum_{j=0}^{h_{\vec{y}_i}-2}a_{i,j}f^{h_{\vec{y}_i}-1-j}(\vec{y}_i)$ (on a repris l'égalité : $\forall i\in[\![1,m]\!]$, $h_{\vec{y}_i}=h_{\vec{z}_i}+1$, vue à la question précédente).
Alors, par linéarité de f :

$$\vec{z}=f\left(\sum_{i=1}^{m}\sum_{j=0}^{h_{\vec{y}_i}-2}a_{i,j}f^{h_{\vec{y}_i}-2-j}(\vec{y}_i)\right),$$

et le vecteur $\sum_{i=1}^{m}\sum_{j=0}^{h_{\vec{y}_i}-2}a_{i,j}f^{h_{\vec{y}_i}-2-j}(\vec{y}_i)$ appartient à F puisqu'il est une combinaison linéaire de vecteurs de la famille $\mathcal{B}_{\vec{y}_1}\cup\cdots\cup\mathcal{B}_{\vec{y}_m}$, dont on sait qu'elle engendre F par définition de ce sous-espace vectoriel. Donc : $\vec{z}\in f(F)$. Ceci vaut pour tout $\vec{z}\in\mathrm{im}(f)$, d'où l'inclusion voulue. On a bien démontré : $f(F)=\mathrm{im}(f)$.
Voyons en quoi cela va nous aider à démontrer l'inclusion : $\mathbb{C}^n\subseteq F+G$. Soit $\vec{x}\in\mathbb{C}^n$. Alors : $f(\vec{x})\in\mathrm{im}(f)=f(F)$, donc il existe $\vec{y}\in F$ tel que : $f(\vec{x})=f(\vec{y})$. Par linéarité de f, on en déduit : $f(\vec{x}-\vec{y})=\vec{0}$, donc il existe $\vec{z}\in\ker(f)$ tel que : $\vec{x}-\vec{y}=\vec{z}$. Or : $\ker(f)=(F\cap\ker(f))\oplus G$, donc il existe $\vec{a}\in F\cap\ker(f)$ et $\vec{b}\in G$ tels que : $\vec{z}=\vec{a}+\vec{b}$. En compilant tout cela, on obtient :

$$\vec{x}=\underbrace{\vec{y}+\vec{a}}_{\in F}+\underbrace{\vec{b}}_{\in G}$$

d'où le résultat : $\mathbb{C}^n\subseteq F+G$. L'inclusion réciproque étant triviale, on a montré :

$$\mathbb{C}^n=F+G=F\oplus G.$$

Autre démonstration. Après avoir démontré que F et G sont en somme directe, on peut aussi conclure avec un argument dimensionnel, en montrant : $\dim(F)+\dim(G)=\dim(\mathbb{C}^n)$. Pour cela, on utilise d'abord le théorème du rang avec la restriction de f à F, pour avoir l'égalité : $\dim(F)=\dim(\ker(f)\cap F)+\dim(f(F))$. Comme $F\cap\ker(f)$ et G sont supplémentaires dans $\ker(f)$, la première dimension égale $\dim(G)-\ker(f)$; ayant démontré que $f(F)=\mathrm{im}(f)$, comme ci-dessus, on en déduit : $\dim(F)=\dim(\ker(f))-\dim(G)+\dim(\mathrm{im}(f))=\dim(\mathbb{C}^n)-\dim(G)$ (toujours grâce au théorème du rang, mais cette fois-ci appliqué à f), ce qui donne l'égalité dimensionnelle voulue.

Questions à se poser, réflexes à acquérir.
- Soit $\vec{y}\in\mathbb{C}^n$ fixé ; quel résultat structurel peut-on donner concernant $f^{-1}(\{f(\vec{y})\})=\{\vec{x}\in\mathbb{C}^n\mid f(\vec{x})=f(\vec{y})\}$, qui aurait pu être invoqué ici ? En quoi est-ce conséquence d'un résultat général de 1re année ?
- Se convaincre de l'égalité $\dim(F)=\dim(\ker(f)\cap F)+\dim(f(F))$ (surtout pour le noyau).

13. Comme F et G sont supplémentaires dans $\mathbb{C}^n$ d'après la question précédente, la concaténation d'une base de F et d'une base de G donne une base de $\mathbb{C}^n$. On connaît déjà une base de F par la question 11, à savoir la famille $\mathcal{B}_{\vec{y}_1}\cup\cdots\cup\mathcal{B}_{\vec{y}_m}$. Posons donc : $\forall i\in[\![1,m]\!]$, $\vec{x}_i=\vec{y}_i$, et soit $(\vec{x}_{m+1},\dots,\vec{x}_\ell)$ une base de G (notons que l'on a $\ell-m=\dim(G)$ par définition de ℓ, donc l'indexation est cohérente). La famille $\mathcal{B}_{\vec{x}_1}\cup\cdots\cup\mathcal{B}_{\vec{x}_m}\cup(\vec{x}_{m+1},\dots,\vec{x}_\ell)$ est une base de $\mathbb{C}^n$ d'après ce qui vient d'être dit : vérifions qu'elle est effectivement de la forme voulue par l'énoncé. Pour cela, on rappelle que G est inclus dans $\ker(f)$, de sorte que : $\forall i\in[\![m+1,\ell]\!]$, $f(\vec{x}_i)=\vec{0}$, et on en déduit que les vecteurs $\vec{x}_{m+1},\dots,\vec{x}_\ell$ sont tous de hauteur 1. Par conséquent : $\forall i\in[\![m+1,\ell]\!]$, $\mathcal{B}_{\vec{x}_i}=(\vec{x}_i)$, et on a :

$$\mathcal{B}_{\vec{x}_1}\cup\cdots\cup\mathcal{B}_{\vec{x}_m}\cup(\vec{x}_{m+1},\dots,\vec{x}_\ell)=\mathcal{B}_{\vec{x}_1}\cup\cdots\cup\mathcal{B}_{\vec{x}_m}\cup\mathcal{B}_{\vec{x}_{m+1}}\cup\cdots\cup\mathcal{B}_{\vec{x}_\ell}.$$

C'est une base de $\mathbb{C}^n$, ce qui démontre le résultat voulu.

14. La différence entre cette question et la précédente, est que cette fois on ne suppose plus l'existence de $\vec{z}_1,\dots,\vec{z}_m$ tels que $\mathcal{B}_{\vec{z}_1}\cup\cdots\cup\mathcal{B}_{\vec{z}_m}$ soit une base de $\mathrm{im}(f)$. Notre récurrence permettra de s'y ramener. Pour tout $n\in\mathbb{N}$, soit P_n la proposition :

« Pour tout $\mathbb{C}$-espace vectoriel E de dimension au plus n, et pour tout endomorphisme nilpotent f de E, alors soit $\dim(E) = 0$, soit il existe $\ell \in \mathbb{N} \setminus \{0\}$ et $(\vec{x}_1, \dots, \vec{x}_\ell) \in (\mathbb{C}^n)^\ell$ tels que $\mathcal{B}_{\vec{x}_1} \cup \dots \cup \mathcal{B}_{\vec{x}_\ell}$ soit une base de E. »

(Les définitions de $\mathcal{B}_{\vec{x}}$ et de la hauteur d'un vecteur restent les mêmes, même si l'espace vectoriel considéré n'est plus $\mathbb{C}^n$.)

Nous allons démontrer que P_n est vraie pour tout $n \in \mathbb{N}$, par récurrence sur n.

L'initialisation est immédiate. Passons à l'hérédité. Soit $n \in \mathbb{N}$. Supposons P_n et démontrons P_{n+1}. Soient E un $\mathbb{C}$-espace vectoriel de dimension au plus $n+1$ et f un endomorphisme nilpotent de E. Si E est de dimension nulle, alors il n'y a rien à raconter. Supposons donc : $\dim(E) \geqslant 1$. L'idée est de construire, en partant de f, un endomorphisme défini sur un sous-espace vectoriel strict de E, afin d'utiliser P_n, et les questions précédentes incitent plus précisément à se ramener à $\mathrm{im}(f)$. Encore faut-il bien avoir : $\dim(\mathrm{im}(f)) \leqslant n$. C'est ce que nous justifions à présent, en remarquant qu'un endomorphisme nilpotent ne peut pas être inversible. Un argument parmi d'autres est par l'absurde : supposons que f est inversible. Si p est l'indice de nilpotence de f, alors composer l'égalité $f^p = 0_{\mathrm{L}(E)}$ par f^{-1} donne : $f^{p-1} = 0_{\mathrm{L}(E)}$, ce qui contredit la minimalité de l'indice de nilpotence : d'où la contradiction. Ainsi f n'est pas inversible, et comme c'est un endomorphisme d'un espace vectoriel de dimension finie cela implique en particulier que f n'est pas surjectif. D'où : $\mathrm{im}(f) \neq E$, puis : $\dim(\mathrm{im}(f)) \leqslant n$. Considérons alors l'endomorphisme de $\mathrm{im}(f)$ défini par :

$$g : \begin{cases} \mathrm{im}(f) & \to & \mathrm{im}(f) \\ \vec{x} & \mapsto & f(\vec{x}) \end{cases}.$$

Le fait que ce soit un endomorphisme est clair, et il est aussi nilpotent parce que f l'est. On peut donc appliquer l'hypothèse de récurrence P_n avec l'espace vectoriel $\mathrm{im}(f)$ muni de l'endomorphisme nilpotent g, et on en déduit que soit $\dim(\mathrm{im}(f)) = 0$ (cas qu'on traite en second lieu), soit il existe $m \in \mathbb{N} \setminus \{0\}$ et $(\vec{z}_1, \dots, \vec{z}_m) \in (\mathrm{im}(f))^m$ tels que $\mathcal{B}_{\vec{z}_1} \cup \dots \cup \mathcal{B}_{\vec{z}_m}$ soit une base de $\mathrm{im}(f)$. Dans ce dernier cas, la proposition P_{n+1} découle alors de la question précédente. Si $\mathrm{im}(f) = \{\vec{0}\}$ alors f est l'endomorphisme nul, et tout vecteur non nul est de hauteur 1. On en déduit, en raisonnant comme dans la question précédente, que *toute* base de E est de la forme $\mathcal{B}_{\vec{x}_1} \cup \dots \cup \mathcal{B}_{\vec{x}_\ell}$. Ce qui achève la récurrence (notons que pour tout $\vec{x} \in \mathrm{im}(f)$ non nul, $\mathcal{B}_{\vec{x}}$ décrit la même famille, que ce soit en appliquant la définition de $\mathcal{B}_{\vec{x}}$ avec g ou f).

Ainsi P_n est vraie pour tout $n \in \mathbb{N} \setminus \{0\}$. En prenant $E = \mathbb{C}^n$ et f l'endomorphisme de l'énoncé, on a bien montré qu'il existe $\ell \in \mathbb{N} \setminus \{0\}$ et $(\vec{x}_1, \dots, \vec{x}_\ell) \in (\mathbb{C}^n)^\ell$ tels que $\mathcal{B}_{\vec{x}_1} \cup \dots \cup \mathcal{B}_{\vec{x}_\ell}$ soit une base de $\mathbb{C}^n$.

Questions à se poser, réflexes à acquérir.

- Pourquoi l'initialisation est immédiate ? Comprendre pourquoi j'ai voulu initialiser à $n = 0$ plutôt qu'à $n = 1$ (bien que n soit supposé non nul dans l'énoncé), et se demander comment on s'en serait sorti malgré tout dans le second cas.
- Se convaincre qu'absolument tout ce qui est quantifié dans P_n est essentiel : qu'est-ce qui aurait été embêtant, si je n'avais pas écrit « pour tout $\mathbb{C}$-espace vectoriel E » dans P_n, puisque après tout, je ne travaille que dans $\mathbb{C}^n$ dans ce problème ? Question analogue avec la quantification de l'endomorphisme f. Qu'est-ce qui aurait posé problème si je n'avais pas inclus « soit $\dim(E) = 0$ » à la formulation ? Qu'aurait-on pu proposer d'autre à la place ?
- Se convaincre que la définition de $\mathcal{B}_{\vec{x}}$ est la même avec g ou f.
- Trouver d'autres démonstrations qu'un endomorphisme nilpotent n'est pas inversible.

15. D'après la question précédente, il existe $\ell \in \mathbb{N} \setminus \{0\}$ et $(\vec{x}_1, \dots, \vec{x}_\ell) \in (\mathbb{C}^n)^\ell$ tels que $\mathcal{B}_{\vec{x}_1} \cup \dots \cup \mathcal{B}_{\vec{x}_\ell}$ soit une base de $\mathbb{C}^n$. Appelons $\mathcal{B}$ une telle base, et reprenons les notations ci-avant. En reprenant le raisonnement de la question 9, on observe que la matrice représentative de f dans la base $\mathcal{B}$ est : $J_{k_1,\dots,k_\ell}$, où l'on a posé : $\forall i \in [\![1, \ell]\!]$, $k_i = h_{\vec{x}_i}$. La formule du changement de base appliquée à f, entre la base canonique et la base $\mathcal{B}$, assure alors que A est semblable à une matrice de la forme $J_{k_1,\dots,k_\ell}$: d'où le résultat.

☢ 16. Pour répondre à cette question, il vaut mieux comprendre dans un premier temps ce qu'est le noyau de J_k pour tout $k \in \mathbb{N} \setminus \{0\}$, ainsi que les noyaux de ses puissances. Soit $k \in \mathbb{N} \setminus \{0\}$, et soit f_k

l'endomorphisme de $\mathbb{C}^k$ canoniquement associé à J_k. Notons $(\vec{e}_1, \dots, \vec{e}_k)$ la base canonique de $\mathbb{C}^k$. Il est plus facile d'étudier les puissances de J_k en passant par f_k. On a en effet une description très simple de $f(\vec{e}_i)$ pour tout $i \in [\![1, k]\!]$:

$$\forall i \in [\![1,k]\!], \quad f(\vec{e}_i) = \begin{cases} \vec{e}_{i-1} & \text{si} \quad i \geqslant 2, \\ \vec{0} & \text{si} \quad i = 1. \end{cases}$$

Une récurrence très basique permet alors de démontrer que pour tout $j \in \mathbb{N}$, on a :

$$\forall i \in [\![1,k]\!], \quad f^j(\vec{e}_i) = \begin{cases} \vec{e}_{i-j} & \text{si} \quad i \geqslant j+1, \\ \vec{0} & \text{si} \quad i \leqslant j. \end{cases}$$

Pour $j \geqslant k$, on a donc : $f^j = 0_{\mathrm{L}(\mathbb{C}^k)}$. Matriciellement, cela donne :

$$(J_k)^0 = \mathrm{I}_k, \quad \text{et :} \quad \forall j \in \mathbb{N} \setminus \{0\},\ (J_k)^j = \begin{cases} \begin{pmatrix} 0_{k-j+1,j-1} & J_{k-j+1} \\ 0_{j-1} & 0_{j-1,k-j+1} \end{pmatrix} & \text{si} \quad j < k, \\ 0_{\mathrm{M}_k(\mathbb{C})} & \text{si} \quad j \geqslant k. \end{cases}$$

De cela on déduit aisément que pour tout $j \in \mathbb{N}$ et tout $k \in \mathbb{N} \setminus \{0\}$:

$$\mathrm{rg}\left[(J_k)^j\right] = \begin{cases} k-j & \text{si } j < k, \\ 0 & \text{si } j \geqslant k, \end{cases}, \quad \dim\left[\ker\left((J_k)^j\right)\right] = \begin{cases} j & \text{si } j < k, \\ k & \text{si } j \geqslant k. \end{cases}$$

En effet, si $j < k$, alors $(J_k)^j$ a j colonnes nulles et les $k-j$ autres sont linéairement indépendantes car échelonnées. On peut éviter une distinction de cas en écrivant :

$$\forall k \in \mathbb{N} \setminus \{0\},\ \forall j \in \mathbb{N}, \quad \dim\left[\ker\left((J_k)^j\right)\right] = \min(j,k). \tag{$*$}$$

On considère à présent $\ell \in \mathbb{N} \setminus \{0\}$ et $(k_1, \dots, k_\ell) \in (\mathbb{N} \setminus \{0\})^\ell$. On déduit des calculs ci-dessus la dimension du noyau de $J_{k_1,\dots,k_\ell} = \mathrm{diag}(J_{k_1}, \dots, J_{k_\ell})$, et du noyau de ses puissances, *via* l'identité suivante :

$$\forall j \in \mathbb{N}, \quad \dim\left[\ker\left((J_{k_1,\dots,k_\ell})^j\right)\right] = \sum_{i=1}^{\ell} \dim\left[\ker\left((J_{k_i})^j\right)\right]. \tag{$\dagger$}$$

Pour se convaincre de la justesse de cette égalité, on peut par exemple raisonner sur le rang de $(J_{k_1,\dots,k_\ell})^j$: ses colonnes non nulles sont des vecteurs distincts de la base canonique que l'on a permutés, donc elles sont linéairement indépendantes. Ainsi, pour déterminer le rang (et, par le théorème du rang, obtenir la dimension du noyau), il suffit de compter le nombre de colonnes non nulles. Pour cela, on compte le nombre de colonnes non nulles parmi les k_1 premières (ce qui revient à étudier $(J_{k_1})^j$), puis parmi les k_2 suivantes, etc., jusqu'aux k_ℓ dernières (qui correspondent à l'étude de $(J_{k_\ell})^j$). Cela donne la relation : $\forall j \in \mathbb{N}$, $\mathrm{rg}\left[(J_{k_1,\dots,k_\ell})^j\right] = \sum_{i=1}^{\ell} \mathrm{rg}\left[(J_{k_i})^j\right]$, et par le théorème du rang on obtient l'identité ci-dessus (ne pas oublier que les matrices J_{k_i} sont de taille k_i, et que $J_{k_1,\dots,k_r}$ est de taille $\sum_{i=1}^{r} k_i = n$).

Revenons à présent au contexte de l'énoncé. Soient $(\ell, m) \in (\mathbb{N} \setminus \{0\})^2$ et $(k_1, \dots, k_\ell, k'_1, \dots, k'_m) \in (\mathbb{N} \setminus \{0\})^{\ell+m}$ tels que : $k_1 \leqslant \cdots \leqslant k_\ell$, $k'_1 \leqslant \cdots \leqslant k'_m$, et : $n = \sum_{i=1}^{\ell} k_i = \sum_{i=1}^{m} k'_i$. Soit $A \in \mathrm{M}_n(\mathbb{C})$ une matrice semblable à $J_{k_1,\dots,k_\ell}$ et $J_{k'_1,\dots,k'_m}$. Commençons par justifier que l'on a : $\ell = m$. Pour cela, notons d'abord que par transitivité de la relation de similitude, les matrices $J_{k_1,\dots,k_\ell}$ et $J_{k'_1,\dots,k'_m}$ sont semblables. Elles ont donc même rang, et par le théorème du rang leurs noyaux sont de même dimension. Or, d'après l'identité ci-dessus avec $j = 1$:

$$\dim\left(\ker\left(J_{k_1,\dots,k_\ell}\right)\right) \overset{(\dagger)}{=} \sum_{i=1}^{\ell} \dim\left(\ker\left(J_{k_i}\right)\right) \overset{(*)}{=} \sum_{i=1}^{\ell} \min(k_i,1) = \sum_{i=1}^{\ell} 1 = \ell,$$

et de même : $\dim\left(\ker\left(J_{k'_1,\dots,k'_\ell}\right)\right) = m$. Par égalité des dimensions, on a donc bien : $\ell = m$.
Passons à la démonstration des égalités : $\forall i \in [\![1,\ell]\!]$, $k_i = k'_i$. On suit l'indication de l'énoncé. Pour tout $j \in \mathbb{N}\setminus\{0\}$ on a, toujours d'après $(*)$ et $(\dagger)$:

$$\begin{aligned}\dim\left[\ker\left((J_{k_1,\dots,k_\ell})^{j+1}\right)\right] - \dim\left[\ker\left((J_{k_1,\dots,k_\ell})^{j}\right)\right] &\overset{(\dagger)}{=} \sum_{i=1}^{\ell}\left[\dim\left[\ker\left((J_{k_i})^{j+1}\right)\right] - \dim\left[\ker\left((J_{k_i})^{j}\right)\right]\right]\\ &\overset{(*)}{=} \sum_{i=1}^{\ell}\left(\min(j+1,k_i) - \min(j,k_i)\right).\end{aligned}$$

Examinons de plus près cette somme. Son i^{e} terme, pour tous $j \in \mathbb{N}\setminus\{0\}$ et $i \in [\![1,\ell]\!]$, est égal à 1 si et seulement si $k_i \geqslant j+1$, et il est égal à 0 dans tous les autres cas. On a donc, plus simplement :

$$\forall j \in \mathbb{N}\setminus\{0\},\quad \dim\left[\ker\left((J_{k_1,\dots,k_\ell})^{j+1}\right)\right] - \dim\left[\ker\left((J_{k_1,\dots,k_\ell})^{j}\right)\right] = \operatorname{card}\left(\{i \in [\![1,\ell]\!] \mid k_i \geqslant j+1\}\right).$$

Cette égalité vaut aussi pour $j = 0$, comme on le vérifie immédiatement. L'égalité entre les dimensions des noyaux de $(J_{k_1,\dots,k_\ell})^j$ et $(J_{k'_1,\dots,k'_\ell})^j$, déjà justifiée plus haut, implique donc :

$$\forall j \in \mathbb{N},\quad \operatorname{card}\left(\{i \in [\![1,\ell]\!] \mid k_i \geqslant j+1\}\right) = \operatorname{card}\left(\{i \in [\![1,\ell]\!] \mid k'_i \geqslant j+1\}\right).$$

Comme :

$$\forall j \in \mathbb{N},\ \operatorname{card}\left(\{i \in [\![1,\ell]\!] \mid k_i = j+1\}\right) = \operatorname{card}\left(\{i \in [\![1,\ell]\!] \mid k_i \geqslant j+1\}\right) - \operatorname{card}\left(\{i \in [\![1,\ell]\!] \mid k_i \geqslant j+2\}\right),$$

l'égalité ci-dessus implique :

$$\forall j \in \mathbb{N},\quad \operatorname{card}\left(\{i \in [\![1,\ell]\!] \mid k_i = j+1\}\right) = \operatorname{card}\left(\{i \in [\![1,\ell]\!] \mid k'_i = j+1\}\right).$$

Très concrètement, cette égalité signifie que pour tout $j \in \mathbb{N}$, l'entier $j+1$ apparaît autant de fois dans la suite $(k_1,\dots,k_\ell)$ que dans la suite $(k'_1,\dots,k'_\ell)$. Si l'on note, pour tout $j \in \mathbb{N}$, m_j le nombre d'apparitions de $j+1$ dans la suite $(k_1,\dots,k_\ell)$ (qui est le même que le nombre d'apparitions de $j+1$ dans la suite $(k'_1,\dots,k'_\ell)$), alors il existe une unique suite de ℓ éléments, *croissante*, prenant exactement m_j fois la valeur j pour tout $j \in [\![1,n]\!]$ (sachant qu'il est inutile de considérer le cas $j > n$, puisque tous les k_i et k'_i doivent être inférieurs ou égaux à n, du fait que l'on ait : $n = \sum_{i=1}^{\ell} k_i = \sum_{i=1}^{\ell} k'_i$). Or les suites $(k_1,\dots,k_\ell)$ et $(k'_1,\dots,k'_\ell)$ en sont également. L'unicité assure que l'on a : $(k_1,\dots,k_\ell) = (k'_1,\dots,k'_\ell)$, d'où le résultat.

Remarque. Grâce à ce qu'on a démontré ci-dessus, on peut démontrer que l'on a plus précisément, pour tout $j \in \mathbb{N}\setminus\{0\}$:

$$\begin{aligned}\operatorname{card}\left(\{i \in [\![1,\ell]\!] \mid k_i = j+1\}\right) = 2\dim\left[\ker\left((J_{k_1,\dots,k_\ell})^{j+1}\right)\right] - \dim\left[\ker\left((J_{k_1,\dots,k_\ell})^{j+2}\right)\right]\\ - \dim\left[\ker\left((J_{k_1,\dots,k_\ell})^{j}\right)\right]\end{aligned}$$

Remarque. Pour toute permutation σ de l'ensemble $\{k_1,\dots,k_r\}$, les matrices $J_{k_1,\dots,k_r}$ et $J_{\sigma(k_1),\dots,\sigma(k_r)}$ sont semblables, comme on le démontrerait en passant de la base canonique à une base obtenue par permutation convenable des vecteurs (le faire soigneusement, pour se convaincre qu'on en est capable). C'est pourquoi il est nécessaire d'ordonner les k_i pour obtenir un résultat d'unicité.

Remarque. En passant, on déduit du calcul matriciel de cette question que J_k est nilpotente d'indice de nilpotence k. Ce sera utile pour la question 18 notamment.

Questions à se poser, réflexes à acquérir.

- Se demander pourquoi j'ai préféré passer par f^j pour obtenir A^j. Quel est le gain ? En quelles circonstances y penser ? Voir, aussi, si on parvient à calculer A^j directement.
- Se convaincre par d'autres arguments matriciels de l'identité vérifiée par le rang de $J^j_{k_1,\dots,k_\ell}$ (de même avec le noyau). Idéalement, trouver une démonstration utilisant l'endomorphisme canoniquement associé (ce sera sans doute plus facile après avoir découvert les sommes directes de $k \geqslant 3$ sous-espaces vectoriels, ou bien : se contenter de démontrer l'identité avec J_{k_1,k_2}).
- Se convaincre, à la lecture de cette question, que l'intérêt des matrices $J_{k_1,\dots,k_\ell}$ est que TOUT ce qui est intéressant pour une matrice se voit à l'œil nu sur une telle matrice.

Quatrième partie : racines carrées des matrices nilpotentes

17. Soit $A \in \mathrm{M}_n(\mathbb{C})$ une matrice nilpotente, dont on note p l'indice de nilpotence. Soit $M \in \mathcal{R}(A)$. On a : $M^{2p} = (M^2)^p = A^p = 0_{\mathrm{M}_n(\mathbb{C})}$, ce qui démontre que M est nilpotente.

18. Soit $P \in \mathrm{GL}_n(\mathbb{C})$ une matrice telle que : $A = PJ_{k_1,\ldots,k_\ell}P^{-1}$. Une telle matrice existe puisqu'on suppose A semblable à $J_{k_1,\ldots,k_\ell}$. On montre alors, par une récurrence facile, que l'on a :

$$\forall m \in \mathbb{N}, \quad A^m = P\left(J_{k_1,\ldots,k_\ell}\right)^m P^{-1} = P\begin{pmatrix} J_{k_1}^m & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & J_{k_\ell}^m \end{pmatrix} P^{-1}.$$

Or le calcul matriciel effectué au début de la question 16 montre que pour tout $k \in \mathbb{N}\setminus\{0\}$, on a : $(J_k)^{k-1} \neq 0_{\mathrm{M}_k(\mathbb{C})}$, et : $(J_k)^k = 0_{\mathrm{M}_k(\mathbb{C})}$ (ce dont on déduit que pour tout entier $m \geqslant k$, on a encore $(J_k)^m = 0_{\mathrm{M}_k(\mathbb{C})}$, tandis que $(J_k)^m$ est non nulle pour tout entier naturel $m < k$). Autrement dit : pour tout $k \in \mathbb{N}\setminus\{0\}$, l'indice de nilpotence de J_k est k. L'égalité ci-dessus implique donc que si l'on pose : $p = \max(k_1,\ldots,k_\ell)$, alors pour tout $i \in [\![1,\ell]\!]$ on a : $(J_{k_i})^p = 0_{\mathrm{M}_{k_i}(\mathbb{C})}$ (car $p \geqslant k_i$), et donc : $A^p = 0_{\mathrm{M}_n(\mathbb{C})}$. En revanche A^{p-1} n'est pas la matrice nulle : si l'on note $i_0 \in [\![1,\ell]\!]$ l'indice tel que : $p = k_{i_0}$, alors on a : $(J_{k_{i_0}})^{p-1} = (J_{k_{i_0}})^{k_{i_0}-1} \neq 0_{\mathrm{M}_{k_{i_0}}(\mathbb{C})}$, donc A^{p-1} contient des coefficients non nuls.

On a donc justifié que l'indice de nilpotence de A est : $p = \max(k_1,\ldots,k_\ell)$. Ceci démontre en particulier que p est inférieur ou égal à n : en effet, en comparant la taille de A avec celle de $J_{k_1,\ldots,k_\ell}$, on voit que l'on a : $\sum_{i=1}^{\ell} k_i = n$, donc : $\forall i \in [\![1,\ell]\!]$, $k_i \leqslant n$ (on utilise que les k_i sont positifs). Leur maximum p vérifie donc aussi cette inégalité. On en déduit : $n - p \geqslant 0$, puis : $A^n = A^{n-p} \times A^p = 0_{\mathrm{M}_n(\mathbb{C})}$.

Remarque. On redémontrera ce résultat cette année grâce au théorème de Cayley-Hamilton.

> **Questions à se poser, réflexes à acquérir.** Faire la « récurrence facile » omise.

19. Soit $M \in \mathcal{R}\left(0_{\mathrm{M}_n(\mathbb{C})}\right)$. On a alors : $M^2 = 0_{\mathrm{M}_n(\mathbb{C})}$, donc M est nilpotente et d'indice de nilpotence inférieur ou égal à 2. Le fait que M soit nilpotente assure, par la question 15, qu'il existe $\ell \in \mathbb{N}\setminus\{0\}$ et $(k_1,\ldots,k_\ell) \in (\mathbb{N}\setminus\{0\})^\ell$ tels que M soit semblable à $J_{k_1,\ldots,k_\ell}$, et le fait que l'indice de nilpotence de M soit inférieur ou égal à 2 assure, par la question précédente, que l'on a : $\forall i \in [\![1,\ell]\!]$, $k_i \leqslant 2$. Ainsi toute matrice $M \in \mathcal{R}\left(0_{\mathrm{M}_n(\mathbb{C})}\right)$ est dans la classe d'équivalence d'une matrice la forme $J_{k_1,\ldots,k_\ell}$ avec : $(k_1,\ldots,k_\ell) \in \{1,2\}^\ell$. Comme il y a un nombre fini de ℓ-uplets $(k_1,\ldots,k_\ell) \in \{1,2\}^\ell$, il y a un nombre fini de matrices $J_{k_1,\ldots,k_\ell}$ auxquelles une matrice M de $\mathcal{R}\left(0_{\mathrm{M}_n(\mathbb{C})}\right)$ peut être semblable, et donc il y a un nombre fini de classes d'équivalence pour la relation de similitude.

Remarque. On peut faire un peu mieux, en notant que la relation $\sum_{i=1}^{\ell} k_i = n$ impose qu'au plus $\left\lfloor \frac{n}{2} \right\rfloor$ entiers parmi les k_i sont égaux à 2. Avec un peu plus de travail, cela permettrait de démontrer qu'il y a exactement $\left\lfloor \frac{n}{2} \right\rfloor + 1$ classes d'équivalence (exercice).

> **Questions à se poser, réflexes à acquérir.** Comment formaliser ce qui fut démontré, en écrivant une inclusion ou une application injective (ou surjective selon la façon de poser l'application) convenable ?

20. **Coquille de l'énoncé.** L'inégalité $2p - 2n \geqslant n$ doit être remplacée par : $2p - 2 \geqslant n$.
Supposons : $2p - 2 \geqslant n$, et raisonnons par l'absurde en supposant que $\mathcal{R}(A)$ est non vide. Soit $M \in \mathcal{R}(A)$. Comme A est nilpotente, M est nilpotente aussi d'après la question 17. D'après la question 18, on a : $M^n = 0_{\mathrm{M}_n(\mathbb{C})}$, et donc aussi, du fait que $2p-2$ soit supérieur ou égal à n :

$$A^{p-1} = M^{2p-2} = M^{(2p-2)-n} \times M^n = 0_{\mathrm{M}_n(\mathbb{C})}.$$

Or on a aussi : $A^{p-1} \neq 0_{\mathrm{M}_n(\mathbb{C})}$, puisque A est supposée d'indice de nilpotence p. C'est absurde. Ce raisonnement montre que l'on a nécessairement $\mathcal{R}(A) = \emptyset$ si $2p - 2 \geqslant n$.

Exemple. La matrice $J_3 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$ n'admet pas de racine carrée, puisque son indice de nilpotence vaut 3, et vérifie : $2 \cdot 3 - 2 \geqslant 3$. Plus généralement, si $k \geqslant 2$ alors J_k n'admet pas de racine carrée.

Questions à se poser, réflexes à acquérir. Peut-on améliorer cette inégalité ? Par exemple, peut-on trouver des matrices nilpotentes A dont l'indice de nilpotence est $\lfloor \frac{n}{2} \rfloor$ (c'est le plus grand entier à *ne pas* vérifier l'inégalité $2p - 2 \geqslant n$), et qui admettent des racines carrées ?

21. Soit f l'endomorphisme de $\mathbb{C}^{2n}$ canoniquement associé à J_{2n}. Notons $\mathcal{B}_1 = (\vec{e}_1, \dots, \vec{e}_{2n})$ la base canonique de $\mathbb{C}^{2n}$. On rappelle que d'après le calcul du début de la question 16, on a :

$$\forall i \in [\![1,2n]\!], \quad f^2(\vec{e}_i) = \begin{cases} \vec{e}_{i-2} & \text{si} \quad i \geqslant 3, \\ \vec{0} & \text{si} \quad i \in \{1,2\}. \end{cases}$$

En particulier, on en déduit que les deux premières colonnes de J_{2n}^2 sont nulles. Pour obtenir la matrice $\begin{pmatrix} J_n & 0_{\mathrm{M}_n(\mathbb{C})} \\ 0_{\mathrm{M}_n(\mathbb{C})} & J_n \end{pmatrix}$, une première idée est qu'il suffit de permuter les colonnes nulles de sorte qu'elles soient en première et $(n+1)^{\mathrm{e}}$ positions, ce qui revient à écrire la matrice de f dans une nouvelle base où $\vec{e}_2$ est en $(n+1)^{\mathrm{e}}$ position. Mais cela ne suffit pas à faire apparaître J_n parmi les blocs diagonaux : en effet, dans J_{2n}^2, les indices des coefficients non nuls sont de la forme $(i, i-2)$ (décalage de deux unités entre l'indice de la ligne et l'indice de la colonne), tandis que dans J_n ces indices sont de la forme $(i, i-1)$. Toutes ces considérations nous amènent à considérer la base :

$$\mathcal{B}_2 = (\vec{e}_1, \vec{e}_3, \dots, \vec{e}_{2n-1}, \vec{e}_2, \vec{e}_4, \dots, \vec{e}_{2n}) = (\vec{e}_{2k-1})_{1 \leqslant k \leqslant n} \cup (\vec{e}_{2k})_{1 \leqslant k \leqslant n}\,.$$

C'est bien entendu une base de $\mathbb{C}^{2n}$, puisqu'elle est obtenue par simple permutation des vecteurs de la base canonique. En réexploitant ce qui fut rappelé ci-dessus, on a :

$$\forall k \in [\![1,n]\!], \quad f^2(\vec{e}_{2k-1}) = \begin{cases} \vec{e}_{2(k-1)-1} & \text{si} \quad k \geqslant 2, \\ \vec{0} & \text{si} \quad k = 1, \end{cases}, \quad f^2(\vec{e}_{2k}) = \begin{cases} \vec{e}_{2(k-1)} & \text{si} \quad k \geqslant 2, \\ \vec{0} & \text{si} \quad k = 1, \end{cases}$$

ce qui permet de démontrer que la matrice représentative de f^2 dans la base $\mathcal{B}_2$ est : $\begin{pmatrix} J_n & 0_{\mathrm{M}_n(\mathbb{C})} \\ 0_{\mathrm{M}_n(\mathbb{C})} & J_n \end{pmatrix}$.
Comme la matrice de f^2 dans la base canonique $\mathcal{B}_1$ est J_{2n}^2, la formule du changement de base implique que J_{2n}^2 est semblable à $\begin{pmatrix} J_n & 0_{\mathrm{M}_n(\mathbb{C})} \\ 0_{\mathrm{M}_n(\mathbb{C})} & J_n \end{pmatrix}$, ce qu'il fallait démontrer.

Pour démontrer que la matrice J_{2n+1}^2 est semblable à $\begin{pmatrix} J_n & 0_{\mathrm{M}_{n,n+1}(\mathbb{C})} \\ 0_{\mathrm{M}_{n+1,n}(\mathbb{C})} & J_{n+1} \end{pmatrix}$, on procède de même : il faut simplement appliquer la formule du changement de base à l'endomorphisme de $\mathbb{C}^{2n+1}$ canoniquement associé à J_{2n+1}, entre la base canonique $(\vec{e}_1, \dots, \vec{e}_{2n+1})$ et la base $(\vec{e}_{2k})_{1 \leqslant k \leqslant n} \cup (\vec{e}_{2k-1})_{1 \leqslant k \leqslant n+1}$.

Questions à se poser, réflexes à acquérir. Varier les changements de base de cette forme, pour un peu mieux comprendre le type de matrice qu'on peut obtenir par simple permutation des vecteurs de la base. On peut notamment se demander à quelles conditions deux matrices, dont l'une est obtenue à partir de l'autre par permutations de lignes et colonnes, sont semblables (on sait qu'elles sont équivalentes). Si vous avez besoin d'un fil directeur : sauriez-vous démontrer que J_n est semblable à $J_n^\top$? Et, par extension, que toute matrice nilpotente est semblable à sa transposée ?

22. D'après la question 1, pour déterminer si $\mathcal{R}(A)$ est vide ou non, il suffit de déterminer ce qu'il en est pour $\mathcal{R}(J_{3,4,4,4})$. Les calculs qui suivent consistent à construire une racine carrée de $J_{3,4,4,4}$.
Posons : $M = \begin{pmatrix} J_7 & 0_{\mathrm{M}_{7,8}(\mathbb{C})} \\ 0_{\mathrm{M}_{8,7}(\mathbb{C})} & J_8 \end{pmatrix}$. Alors : $M^2 = \begin{pmatrix} J_7^2 & 0_{\mathrm{M}_{7,8}(\mathbb{C})} \\ 0_{\mathrm{M}_{8,7}(\mathbb{C})} & J_8^2 \end{pmatrix}$, et d'après la question précédente il existe $(P,Q) \in \mathrm{GL}_7(\mathbb{C}) \times \mathrm{GL}_8(\mathbb{C})$ tel que : $J_7^2 = P \begin{pmatrix} J_3 & 0_{\mathrm{M}_{3,4}(\mathbb{C})} \\ 0_{\mathrm{M}_{4,3}(\mathbb{C})} & J_4 \end{pmatrix} P^{-1}$, et :

$J_8^2 = Q\begin{pmatrix} J_4 & 0_{\mathrm{M}_4(\mathbb{C})} \\ 0_{\mathrm{M}_4(\mathbb{C})} & J_4 \end{pmatrix}Q^{-1}$. Posons alors : $R = \begin{pmatrix} P & 0_{\mathrm{M}_{7,8}(\mathbb{C})} \\ 0_{\mathrm{M}_{8,7}(\mathbb{C})} & Q \end{pmatrix}$. Un calcul direct montre que R est inversible et d'inverse : $R^{-1} = \begin{pmatrix} P^{-1} & 0_{\mathrm{M}_{7,8}(\mathbb{C})} \\ 0_{\mathrm{M}_{8,7}(\mathbb{C})} & Q^{-1} \end{pmatrix}$. On a alors :

$$\left(R^{-1}MR\right)^2 = R^{-1}M^2R = \begin{pmatrix} P^{-1}J_7^2P & 0_{\mathrm{M}_{7,8}(\mathbb{C})} \\ 0_{\mathrm{M}_{8,7}(\mathbb{C})} & Q^{-1}J_8^2Q \end{pmatrix} = \mathrm{diag}(J_3, J_4, J_4, J_4) = J_{3,4,4,4}.$$

Ainsi : $R^{-1}MR \in \mathcal{R}(J_{3,4,4,4})$, donc : $\mathcal{R}(J_{3,4,4,4}) \neq \emptyset$, et par la question 1 on a également : $\mathcal{R}(A) \neq \emptyset$.
Supposons à présent que $A \in \mathrm{M}_{17}(\mathbb{C})$ est une matrice semblable à $J_{3,4,4,6}$. Là encore, nous allons plutôt étudier $\mathcal{R}(J_{3,4,4,6})$ et montrer que c'est un ensemble vide. Pour cela, raisonnons par l'absurde et supposons qu'il est non vide. Soit $M \in \mathcal{R}(J_{3,4,4,6})$. Comme $J_{3,4,4,6}$ est nilpotente d'indice de nilpotence 6 (d'après la question 18), la matrice M est aussi nilpotente par la question 17, et donc semblable à une matrice de la forme $J_{k_1,\dots,k_\ell}$ avec $\ell \setminus \{0\}$ et $(k_1, \dots, k_\ell) \in (\mathbb{N} \setminus \{0\})^\ell$ d'après la question 15. Quitte à conjuguer par une matrice de permutation convenable, on peut supposer que les k_i sont rangés dans l'ordre croissant : il en fut question dans la remarque de la question 16. On en déduit que $M^2 = J_{3,4,4,6}$ est semblable à $(J_{k_1,\dots,k_\ell})^2$ (il suffit pour cela d'élever au carré une relation de similitude : on a déjà effectué ce raisonnement, implicitement, à plusieurs reprises dans ce sujet). Or : $(J_{k_1,\dots,k_\ell})^2 = \mathrm{diag}((J_{k_1})^2, \dots, (J_{k_\ell})^2)$, et par la question précédente cette matrice est semblable à :

$$\mathrm{diag}\left(J_{k_1'}, J_{k_1''}, \dots, J_{k_\ell'}, J_{k_\ell''}\right),$$

où les k_i' et k_i'' sont définis ainsi :

$$\forall i \in [\![1, \ell]\!], \quad (k_i', k_i'') = \begin{cases} (k_i, k_i) & \text{si } k_i \text{ est pair,} \\ (k_i, k_i + 1) & \text{si } k_i \text{ est impair.} \end{cases}$$

On remarque que ci-dessus, je semble ignorer le cas où l'un des k_i serait égal à 1 (cas qui n'est pas couvert par la question précédente). Mais $J_1^2 = J_1$, et la question 16 assure alors que $J_{3,4,4,6}$ ne peut être semblable à une matrice $J_{m_1,\dots,m_\ell}$ dont l'un des indices m_i serait égal à 1. C'est pourquoi on sait que tous les k_i sont supérieurs ou égaux à 2, ce qui permet d'utiliser la question précédente.
Par transitivité de la relation de similitude, la matrice $J_{3,4,4,6}$ est semblable à $\mathrm{diag}\left(J_{k_1'}, J_{k_1''}, \dots, J_{k_\ell'}, J_{k_\ell''}\right)$. Par la question 16, une première conséquence est qu'il doit y avoir le même nombre de blocs, c'est-à-dire : $4 = 2\ell$, ou encore : $\ell = 2$. Toujours par la question 16, pour que $J_{3,4,4,6}$ soit semblable à $\mathrm{diag}\left(J_{k_1'}, J_{k_1''}, J_{k_2'}, J_{k_2''}\right)$, on doit avoir $3 = k_1'$ et $6 = k_2''$ (ce sont nécessairement les éléments les plus petits et grands respectivement ; l'ordre entre k_1'' et k_2' n'est pas si clair en revanche, et nécessiterait une analyse approfondie, mais c'est inutile pour conclure). Mais alors, on a $k_2' \in \{5,6\}$ d'après la définition des k_i' et k_i'' : dans tous les cas, on ne peut donc pas avoir la suite 3, 4, 4, 6. C'est absurde. On a montré, par l'absurde : $\mathcal{R}(J_{3,4,4,6}) = \emptyset$, et donc : $\mathcal{R}(A) = \emptyset$.
Supposons enfin que $A \in \mathrm{M}_9(\mathbb{C})$ est une matrice semblable à $J_{1,1,1,3,3}$. En utilisant le fait que $J_1^2 = (0) = J_1$, ainsi que la question précédente, on montre que $\mathrm{diag}(J_1, J_1, J_1, J_6)^2$ est semblable à $\mathrm{diag}(J_1, J_1, J_1, J_3, J_3)$, et donc à A. En raisonnant comme dans le premier cas ci-dessus, on en déduit : $\mathcal{R}(A) \neq \emptyset$.

Questions à se poser, réflexes à acquérir. Pourquoi n'est-il pas si clair que $k_1'' \leqslant k_2'$? Pourquoi, en revanche, la hiérarchie est-elle claire aux extrémités ?

23. Il suffit de déterminer une condition nécessaire et suffisante pour que $J_{k_1,\dots,k_\ell}$ admette des racines carrées. Rappelons encore une fois que quitte à conjuguer cette matrice par une matrice de permutation convenable, on peut supposer : $k_1 \leqslant \dots \leqslant k_\ell$. C'est ce qu'on suppose jusqu'à la fin de cette question.
Au vu des exemples de la question précédente, voici ce qu'on peut conjecturer : cette matrice admet des racines carrées si et seulement si la suite $(k_1, \dots, k_\ell)$ est constituée, dans cet ordre, d'une suite de 1, puis de couples de la forme (p, p) ou $(p, p+1)$, où p est un entier naturel non nul.

Démontrons cette conjecture. Commençons par l'implication réciproque. Supposons que la suite $(k_1,\dots,k_\ell)$ est de la forme : $(1,\dots,1,k'_1,k''_1,\dots,k'_r,k''_r)$, où r est un entier naturel éventuellement nul (auquel cas la suite n'est constituée que de 1), $s\in\mathbb{N}$ est le nombre de 1, et que pour tout $i\in[\![1,r]\!]$ il existe $p\in\mathbb{N}\setminus\{0\}$ tel que le couple (k'_i,k''_i) soit de la forme (p,p) ou $(p,p+1)$. Posons : $J=\operatorname{diag}(J_1,\dots,J_1,J_{m_1},\dots,J_{m_r})$, où l'on commence par une suite de s matrices J_1 et où, pour tout $i\in[\![1,r]\!]$, on a $m_i=2k'_i$ si $k''_i=k'_i$, et $m_i=2k'_i+1$ si $k''_i=k'_i+1$. On a évidemment : $\mathcal{R}(J^2)\neq\emptyset$ (une racine carrée de J^2 étant J), or la matrice :

$$J^2=\operatorname{diag}(J_1^2,\dots,J_1^2,(J_{m_1})^2,\dots,(J_{m_r})^2)=\operatorname{diag}(J_1,\dots,J_1,(J_{m_1})^2,\dots,(J_{m_r})^2)$$

est, par définition des m_i, semblable à la matrice $J_{1,\dots,1,k'_1,k''_1,\dots,k'_r,k''_r}=J_{k_1,\dots,k_\ell}$, donc par la question 1 on a aussi : $\mathcal{R}(J_{k_1,\dots,k_\ell})\neq\emptyset$, ce qu'on voulait démontrer.

Passons à l'implication directe. Nous allons la démontrer en raisonnant par contraposée. Supposons : $\mathcal{R}(J_{k_1,\dots,k_\ell})\neq\emptyset$, et soit M une racine carrée de $J_{k_1,\dots,k_\ell}$. Par la question 17, M est nilpotente parce que $J_{k_1,\dots,k_\ell}$ l'est, donc par la question 15 il existe $r\in\mathbb{N}\setminus\{0\}$ et $(m_1,\dots,m_r)\in(\mathbb{N}\setminus\{0\})^r$ tels que M soit semblable à $J_{m_1,\dots,m_r}$, et il ne coûte rien de les ordonner dans un ordre croissant comme on l'a dit plus haut. La matrice $M^2=J_{k_1,\dots,k_\ell}$ est alors semblable à $(J_{m_1,\dots,m_r})^2=\operatorname{diag}((J_{m_1})^2,\dots,(J_{m_r})^2)$. Si des m_i sont égaux à 1, alors ce sont les premiers de la liste (puisqu'on a ordonné les m_i en sens croissant). Soit s le nombre de m_i égaux à 1. Comme : $J_1^2=(0)=J_1$, la matrice $J_{k_1,\dots,k_\ell}$ est semblable à $\operatorname{diag}(J_1,\dots,J_1,(J_{m_{s+1}})^2,\dots,(J_{m_r})^2)$ où, pour tout $i\in[\![s+1,r]\!]$, on a : $m_i\geqslant 2$. Selon la parité de m_i, d'après la question 21, pour tout $i\in[\![s+1,r]\!]$ il existe $p\in\mathbb{N}\setminus\{0\}$ tel que la matrice $(J_{m_i})^2$ soit semblable à $\operatorname{diag}(J_p,J_p)$ ou $\operatorname{diag}(J_p,J_{p+1})$. Pour éviter les distinctions de cas, on note $\operatorname{diag}(J_{k'_i},J_{k''_i})$ la matrice semblable ainsi obtenue. Ainsi la matrice $J_{k_1,\dots,k_\ell}$ est semblable à $\operatorname{diag}\left(J_1,\dots,J_1,J_{k'_{s+1}},J_{k''_{s+1}},\dots,J_{k'_r},J_{k''_r}\right)$.

On aimerait à ce stade conclure en invoquant le résultat d'unicité de la question 16, mais pour cela encore faut-il que la suite $(1,\dots,1,k'_{s+1},k''_{s+1},\dots,k'_r,k''_r)$ soit croissante : ce n'est pas nécessairement le cas, mais on peut s'y ramener. En effet, remarquons que cette suite n'est pas croissante si et seulement s'il existe $i\in[\![s+1,r-1]\!]$ tel que m_i soit impair et $m_i=m_{i+1}$. Fixons un tel i, et considérons alors la suite maximale d'indices i, $i+1$, etc., $i+j$ tels que : $m_i=m_{i+1}=\dots=m_{i+j}$. Il existe $p\in\mathbb{N}\setminus\{0\}$ tel que $\operatorname{diag}(J_{m_i},\dots,J_{m_{i+j}})$ soit semblable à $\operatorname{diag}(J_p,J_{p+1},\dots,J_p,J_{p+1})$ (en l'occurrence, même si ce n'est pas utile pour la suite : $p=\frac{m_i-1}{2}$), où la matrice J_p apparaît j fois, de même pour J_{p+1}. En conjuguant cette matrice par une matrice de permutation convenable, on sait démontrer qu'elle est semblable à $\operatorname{diag}(J_p,\dots,J_p,J_{p+1},\dots,J_{p+1})$, qui fournit cette fois-ci une suite croissante (et, c'est important pour la suite : peu importe que j soit pair ou non, la suite des indices ne fait intervenir consécutivement que des indices de la forme (p,p), $(p,p+1)$ ou $(p+1,p+1)$).

Ce dernier paragraphe permet de conclure : soit σ une permutation des indices $k'_{s+1},k''_{s+1},\dots,k'_r,k''_r$ construite en suivant l'idée ci-dessus (laissant fixe les suites d'indices formant une suite croissante, et permutant comme indiqué les indices qui ne sont pas ordonnés en sens croissant), de sorte que $(1,\dots,1,\sigma(k'_{s+1}),\sigma(k''_{s+1}),\dots,\sigma(k'_r),\sigma(k''_r))$ soit croissante et constituée d'une suite de 1 puis de couples de la forme (p,p) ou $(p,p+1)$ avec $p\in\mathbb{N}\setminus\{0\}$. Alors $J_{k_1,\dots,k_\ell}$ est semblable à $\operatorname{diag}\left(J_1,\dots,J_1,J_{\sigma(k'_{s+1})},J_{\sigma(k''_{s+1})},\dots,J_{\sigma(k'_r)},J_{\sigma(k''_r)}\right)$, donc par la question 16 on a : $(k_1,\dots,k_\ell)=(1,\dots,1,\sigma(k'_{s+1}),\sigma(k''_{s+1}),\dots,\sigma(k'_r),\sigma(k''_r))$, où cette seconde suite est bien de la forme voulue : d'où l'implication directe, obtenue par contraposée. Notre conjecture est démontrée.

Questions à se poser, réflexes à acquérir.

- Pour l'implication réciproque, j'ai choisi de rédiger autrement que dans les exemples de la question précédente : reprendre la question précédente en imitant cette rédaction ; retenir le type de raisonnement ou de rédaction qu'on préfère.
- Pourquoi ai-je raisonné par contraposée ici, et par l'absurde dans l'exemple de la question précédente ?
- Où a servi *l'unicité* de la suite ordonnée $k_1\leqslant\dots\leqslant k_\ell$, dans tout ce problème ? Comment y penser ?
- Peut-on retrouver le résultat de la question 20 grâce à cette condition nécessaire et suffisante ?

Problème d'analyse

Première partie : produit eulérien et infinité de nombres premiers

★ 1. Soit $s > 1$. Nous proposons deux moyens de démontrer que $\zeta(s)$ tend vers l'infini quand $s \to 1$.

Première méthode : comparaison série-intégrale. L'application $t \mapsto \dfrac{1}{t^s}$ est décroissante sur $]0, +\infty[$, donc sur $[1, +\infty[$. Par la méthode de comparaison série-intégrale (que vous devez savoir justifier du début à la fin pour vous l'approprier correctement), on en déduit :

$$\forall N \in \mathbb{N} \setminus \{0\}, \quad \sum_{n=1}^{N} \frac{1}{n^s} \geqslant \int_1^{N+1} \frac{\mathrm{d}t}{t^s} = \frac{1}{s-1}\left(1 - \frac{1}{(N+1)^{s-1}}\right).$$

Quand $N \to +\infty$, on obtient :

$$\zeta(s) = \sum_{n=1}^{+\infty} \frac{1}{n^s} \geqslant \frac{1}{s-1}.$$

Or : $\lim\limits_{s \to 1^+} \dfrac{1}{s-1} = +\infty$. Donc : $\lim\limits_{s \to 1^+} \zeta(s) = +\infty$.

Deuxième méthode : minoration par les sommes partielles. Comme ζ est une fonction décroissante sur $]1, +\infty[$, en tant que somme de fonctions décroissantes (si l'on peine à s'en convaincre, du fait que la somme soit à support infini : écrire que pour tout $(s,t) \in]1, +\infty[^2$ tel que $s \leqslant t$, et pour tout $n \in \mathbb{N} \setminus \{0\}$, on a : $\dfrac{1}{n^t} \leqslant \dfrac{1}{n^s}$, par décroissance de l'application $s \mapsto \dfrac{1}{n^s}$ sur $]1, +\infty[$; sommer cette inégalité de $n = 1$ à $+\infty$, et conclure), on sait qu'elle admet une limite en 1, soit infinie soit finie. Montrons qu'elle est infinie. Puisqu'on somme des termes positifs, on a :

$$\forall N \in \mathbb{N} \setminus \{0\}, \quad \zeta(s) = \sum_{n=1}^{+\infty} \frac{1}{n^s} \geqslant \sum_{n=1}^{N} \frac{1}{n^s}.$$

Quand $s \to 1$, on en déduit : $\forall N \in \mathbb{N} \setminus \{0\}$, $\lim\limits_{s \to 1^+} \zeta(s) \geqslant \sum\limits_{n=1}^{N} \dfrac{1}{n}$, puis, quand $N \to +\infty$:

$$\lim_{s \to 1^+} \zeta(s) \geqslant \sum_{n=1}^{+\infty} \frac{1}{n}.$$

(La somme minorée existe *a priori* puisqu'elle est à termes positifs.)

Or la série harmonique $\sum\limits_{n \geqslant 1} \dfrac{1}{n}$ diverge et est à termes positifs : sa somme est donc égale à $+\infty$, et on en déduit : $\lim\limits_{s \to 1^+} \zeta(s) = +\infty$.

Remarque. Une comparaison série-intégrale plus fine permet d'obtenir l'équivalent : $\zeta(s) \underset{s \to 1}{\sim} \dfrac{1}{s-1}$.

Remarque. Notre deuxième méthode s'adapte à une situation plus générale. Soit I un intervalle de la forme $[a, +\infty[$ ou $]a, +\infty[$, avec $a \in \{-\infty\} \cup \mathbb{R}$. Pour tout $n \in \mathbb{N}$, soit $f_n : I \to \mathbb{R}$ une application décroissante et positive sur I. Alors :

$$\lim_{s \to a} \sum_{n=0}^{+\infty} f_n(s) \geqslant \sum_{n=0}^{+\infty} \lim_{s \to a} f_n(s),$$

toutes les limites et sommes étant bien définies grâce aux hypothèses sur les fonctions f_n. Mieux : en utilisant le fait que, par décroissance, on ait : $\forall s > a$, $\lim\limits_a f_n \geqslant f_n(s)$, et en sommant, on obtient : $\forall s > a$, $\sum\limits_{n=0}^{+\infty} \lim\limits_a f_n \geqslant \sum\limits_{n=0}^{+\infty} f_n(s)$. En prenant la limite dans cette inégalité quand $s \to a$, on a l'inégalité

ci-dessus mais en sens contraire. Par antisymétrie de la relation d'ordre, on a montré que, si $f_n : I \to \mathbb{R}$ est décroissante et positive sur I pour tout $n \in \mathbb{N}$, alors :

$$\lim_{s\to a} \sum_{n=0}^{+\infty} f_n(s) = \sum_{n=0}^{+\infty} \lim_{s\to a} f_n(s). \qquad (\dagger)$$

On a justifié un cas particulier du théorème de convergence monotone, dont l'énoncé général dépasse très largement le cadre du programme. Voyez comment on a raisonné : on a obtenu une inégalité en comparant d'abord les sommes partielles à la somme de la série, *puis* en passant à la limite. L'autre inégalité fut obtenue en comparant d'abord les fonctions à leurs limites, *puis* en sommant.
Nous réutiliserons cette identité (†) à plusieurs reprises dans ce problème.

Questions à se poser, réflexes à acquérir.

- N'était-il pas possible d'écrire d'emblée : $\lim_{s\to 1} \zeta(s) = \sum_{n=1}^{+\infty} \frac{1}{n}$, comme on l'aurait fait pour une somme finie ?
- Comment *penser* à une méthode de comparaison série-intégrale ? Qu'est-ce qui assurait ici que l'approche allait être fructueuse ?
- Se convaincre que nos manipulations avec des inégalités (passages à la limite notamment), clairement vraies dans $\mathbb{R}$, restent vraies si un des éléments est égal à $+\infty$.
- Chaque méthode proposée a des mérites différents. La seconde a l'air plus élémentaire, notamment elle est utilisable plus souvent (elle ne nécessite pas de savoir calculer une intégrale), donc il semble qu'elle soit à privilégier. Pourtant la comparaison série-intégrale donne plus d'informations : lesquelles ?
- Durant l'année, vous aurez des théorèmes pour calculer des limites finies de sommes de séries de fonctions. Ériger au rang de RÉFLEXES ce qu'on a fait là pour des limites infinies.
- Est-ce que la décroissance est indispensable pour la démonstration de (†) ? Si oui, où ? Et la positivité ? Si vous pensez que non : chercher des contre-exemples.

2. Soit $T \geqslant 2$. Notons $p_1, \dots, p_r$ les nombres premiers inférieurs ou égaux à T. On a :

$$\prod_{p\leqslant T} \sum_{k=0}^{+\infty} \frac{1}{p_i^{ks}} = \prod_{i=1}^{r} \sum_{k_i=0}^{+\infty} \frac{1}{p_i^{k_i s}} = \sum_{(k_1,\dots,k_r)\in\mathbb{N}^r} \frac{1}{\left(p_1^{k_1}\cdots p_r^{k_r}\right)^s} = \sum_{n=1}^{+\infty} \frac{\operatorname{card}(I_n)}{n^s} = \sum_{n\in\mathcal{N}(T)} \frac{1}{n^s},$$

où l'on a posé pour tout $n \in \mathbb{N}\setminus\{0\}$:

$$I_n = \left\{(k_1,\dots,k_r) \in \mathbb{N}^r \mid n = p_1^{k_1}\cdots p_r^{k_r}\right\}.$$

La deuxième égalité est vraie par produit fini de séries à termes positifs (et par ailleurs convergentes), tandis que la troisième égalité est vraie d'après le théorème de sommation par paquets où, pour tout $n \in \mathbb{N}\setminus\{0\}$, le n^{e} paquet est I_n. L'unicité de la décomposition de tout entier naturel non nul en facteurs premiers est ce qui assure que ce paquet est soit de cardinal 0 si $n \notin \mathcal{N}(T)$, soit de cardinal 1 dans le cas contraire. D'où le résultat.

Questions à se poser, réflexes à acquérir.

- À chaque sommation par paquets (ici et ailleurs), comprendre ce qui motive le choix des paquets, afin de faire le bon choix par vous-mêmes au moment venu.
- On a utilisé l'unicité de la décomposition en facteurs premiers dans cette question ; est-ce que l'existence est utilisée aussi ? Se poser la question éventuellement plus tard.

★ 3. Remarquons qu'en utilisant l'identité : $\forall x \in]-1,1[,\ \frac{1}{1-x} = \sum_{k=0}^{+\infty} x^k$, avec $x = \frac{1}{p} \in \left[-\frac{1}{2}, \frac{1}{2}\right] \subseteq]-1,1[$, l'égalité de la question précédente peut s'écrire aussi : $\forall T \geqslant 2,\ \prod_{p\leqslant T} \frac{1}{1-\frac{1}{p^s}} = \sum_{n\in\mathcal{N}(T)} \frac{1}{n^s}$. Pour tout réel $T \geqslant 2$, on a donc :

$$0 \leqslant \left|\zeta(s) - \prod_{p\leqslant T} \frac{1}{1-\frac{1}{p^s}}\right| = \sum_{n\in\mathbb{N}\setminus\{0\}} \frac{1}{n^s} - \sum_{n\in\mathcal{N}(T)} \frac{1}{n^s} = \sum_{n\notin\mathcal{N}(T)} \frac{1}{n^s} \leqslant \sum_{n>T} \frac{1}{n^s}.$$

Pour la dernière inégalité, on utilise le fait qu'un entier naturel non nul n n'appartenant pas à $\mathcal{N}(T)$ admette nécessairement un diviseur premier strictement supérieur à T, donc n est lui-même strictement supérieur à T. Autrement dit : $(\mathbb{N}\setminus\{0\})\setminus\mathcal{N}(T) \subseteq (\mathbb{N}\setminus\{0\})\cap]T,+\infty[$. Comme nous sommons des termes positifs, la somme indexée par $(\mathbb{N}\setminus\{0\})\setminus\mathcal{N}(T)$ est plus petite que celle indexée par $(\mathbb{N}\setminus\{0\})\cap]T,+\infty[$. Nous avons là le reste d'une série de Riemann convergente (car $s>1$), donc il tend vers 0 quand $T\to+\infty$. Donc, d'après le théorème des gendarmes :

$$\lim_{T\to+\infty}\prod_{p\leqslant T}\frac{1}{1-\frac{1}{p^s}}=\zeta(s),$$

d'où le résultat.

Questions à se poser, réflexes à acquérir. Le fait que T soit pris réel est-il gênant, pour reconnaître le reste d'une série convergente ? Pourquoi la limite est bien nulle malgré cela ? Question duale : était-il vraiment nécessaire, dans tout ce qu'on nous fait démontrer, de prendre T réel au lieu d'entier naturel ?

4. Soient $n\in\mathbb{N}$ et $x\in]-1,1[$. On sait que l'on a : $\dfrac{1-x^{n+1}}{1-x}=\sum_{k=0}^{n}x^k$. En intégrant cette relation de 0 à x on obtient :

$$\sum_{k=0}^{n}\frac{x^{k+1}}{k+1}=\sum_{k=0}^{n}\int_0^x t^k\mathrm{d}t=\int_0^x\sum_{k=0}^{n}t^k\mathrm{d}t=\int_0^x\frac{1-t^{n+1}}{1-t}\mathrm{d}t=-\ln(1-x)-\int_0^x\frac{t^{n+1}}{1-t}\mathrm{d}t. \qquad (\ddagger)$$

Or, si $x\geqslant 0$, alors on a :

$$0\leqslant\int_0^x\frac{t^{n+1}}{1-t}\mathrm{d}t\leqslant x^{n+1}\int_0^x\frac{\mathrm{d}t}{1-t}=-x^{n+1}\ln(1-x),$$

et si $x\leqslant 0$, alors on majore ainsi à la place :

$$0\leqslant\left|\int_0^x\frac{t^{n+1}}{1-t}\mathrm{d}t\right|=\left|\int_x^0\frac{t^{n+1}}{1-t}\mathrm{d}t\right|\leqslant\int_x^0\frac{|t|^{n+1}}{1-t}\mathrm{d}t\leqslant(-x)^{n+1}\ln(1-x).$$

Dans tous les cas, on a l'encadrement :

$$0\leqslant\left|\int_0^x\frac{t^{n+1}}{1-t}\mathrm{d}t\right|\leqslant|x|^{n+1}\cdot\left|\ln(1-x)\right|,$$

et le majorant tend vers 0 quand $n\to+\infty$, car : $|x|<1$. Par le théorème des gendarmes : $\forall x\in]-1,1[$, $\lim_{n\to+\infty}\int_0^x\frac{t^{n+1}}{1-t}\mathrm{d}t=0$. Passer à la limite dans l'égalité $(\ddagger)$ prouve donc, d'une part, que la série $\sum_{k\geqslant 0}\frac{x^{k+1}}{k+1}$ converge pour tout $x\in]-1,1[$, et d'autre part que l'on a :

$$\forall x\in]-1,1[,\quad -\ln(1-x)=\sum_{k=0}^{+\infty}\frac{x^{k+1}}{k+1}=\sum_{k=1}^{+\infty}\frac{x^k}{k}.$$

Déduisons-en l'égalité vérifiée par $\ln\circ\zeta$. Soit $s>1$ un réel. On a montré, dans la question précédente : $\zeta(s)=\lim_{T\to+\infty}\prod_{p\leqslant T}\frac{1}{1-\frac{1}{p^s}}$. Chaque quantité dans cette égalité est positive strictement, ce qui permet de considérer leur logarithme. Par continuité du logarithme, on a :

$$\lim_{T\to+\infty}\ln\left(\prod_{p\leqslant T}\frac{1}{1-\frac{1}{p^s}}\right)=\ln\left(\lim_{T\to+\infty}\prod_{p\leqslant T}\frac{1}{1-\frac{1}{p^s}}\right)=\ln(\zeta(s)),$$

mais on a aussi, le logarithme étant un morphisme de $(\mathbb{R}_+^*,\times)$ dans $(\mathbb{R},+)$:

$$\forall T\geqslant 2,\quad\ln\left(\prod_{p\leqslant T}\frac{1}{1-\frac{1}{p^s}}\right)=\sum_{p\leqslant T}-\ln\left(1-\frac{1}{p^s}\right).$$

Comme : $\forall p \in \mathbb{P},\ \dfrac{1}{p^s} \in]-1,1[$, on peut utiliser l'identité ci-dessus et on a :

$$\ln\left(\prod_{p\leqslant T}\frac{1}{1-\frac{1}{p^s}}\right) = \sum_{p\leqslant T}\sum_{k=1}^{+\infty}\frac{(1/p^s)^k}{k} = \sum_{p\leqslant T}\sum_{k=1}^{+\infty}\frac{1}{kp^{ks}}.$$

En prenant la limite quand $T \to +\infty$, on obtient finalement :

$$\ln(\zeta(s)) = \lim_{T\to+\infty}\ln\left(\prod_{p\leqslant T}\frac{1}{1-\frac{1}{p^s}}\right) = \lim_{T\to+\infty}\sum_{p\leqslant T}\sum_{k=1}^{+\infty}\frac{1}{kp^{ks}} = \sum_{p}\sum_{k=1}^{+\infty}\frac{1}{kp^{ks}},$$

d'où le résultat.

Questions à se poser, réflexes à acquérir.

- Comment pouvait-on *penser* à démontrer ainsi l'identité vérifiée par $-\ln(1-x)$, *sans indication de l'énoncé* ? C'est TRÈS important car c'est le type de réflexe que l'on attendra de vous cette année. Conjecturer (voire démontrer) d'autres identités analogues (pour varier les situations, on n'utilisera pas que la somme géométrique).
- Pourquoi ne pouvait-on pas directement dire : « $\forall t \in [0,x]$, $\lim_{n\to+\infty}\frac{t^{n+1}}{1-t} = 0$, donc : $\lim_{n\to+\infty}\int_0^x \frac{t^{n+1}}{1-t}\mathrm{d}t = 0$ » ?
- Pourquoi mentionné-je la continuité du logarithme ? Que vient-elle faire là-dedans ?
- Vérifier la stricte positivité que j'ai omise, si elle n'est pas évidente pour vous.
- Il s'avère que l'identité vérifiée par $-\ln(1-x)$ reste valable pour $x = -1$. Est-ce que notre proposition de démonstration permet de le démontrer ? Si non, réfléchir à une alternative.

5. Comme la famille $\left(\dfrac{1}{kp^k}\right)_{(k,p)\in(\mathbb{N}\setminus\{0,1\})\times\mathbb{P}}$ est à termes positifs, on peut utiliser le théorème de Fubini et on a :

$$\sum_{(k,p)\in(\mathbb{N}\setminus\{0,1\})\times\mathbb{P}}\frac{1}{kp^k} = \sum_{p}\sum_{k=2}^{+\infty}\frac{1}{kp^k} \leqslant \sum_{p}\sum_{k=2}^{+\infty}\frac{1}{p^k} = \sum_{p}\frac{1}{p^2}\frac{1}{1-\frac{1}{p}} \leqslant \frac{1}{1-\frac{1}{2}}\sum_{p}\frac{1}{p^2} \leqslant 2\sum_{n=1}^{+\infty}\frac{1}{n^2} < +\infty,$$

donc la famille $\left(\dfrac{1}{kp^k}\right)_{(k,p)\in(\mathbb{N}\setminus\{0,1\})\times\mathbb{P}}$ est sommable.

Questions à se poser, réflexes à acquérir.

- Comment savoir dans quel ordre sommer ? Pourquoi ai-je calculé $\sum_p\sum_k \frac{1}{kp^k}$ plutôt que $\sum_k\sum_p \frac{1}{kp^k}$?
- Pourquoi utiliser le théorème de comparaison serait probablement moins fructueux que notre démonstration ? Dans quels cas analogues la sommabilité se démontrerait relativement facilement par une majoration directe de la somme, mais pas par une majoration du terme général de la famille ?

6. Soit $s > 1$ un réel. On reprend l'égalité de la question 4, en isolant le terme correspondant à $k = 1$. On obtient alors :

$$\sum_{p}\frac{1}{p^s} = \ln(\zeta(s)) - \sum_{p}\sum_{k=2}^{+\infty}\frac{1}{kp^{ks}}.$$

Déduisons-en la limite demandée. Notons d'abord qu'en invoquant le même argument de décroissance que dans la question 1, le membre de gauche de cette égalité admet une limite (finie ou infinie) quand $s \to 1$, et de même pour $\sum_p\sum_{k=2}^{+\infty}\dfrac{1}{kp^{ks}}$. Justifions que cette dernière quantité a une limite finie quand $s \to 1$. Par décroissance du terme général quand s varie, on a pour tout $s \geqslant 1$:

$$\sum_{p}\sum_{k=2}^{+\infty}\frac{1}{kp^{ks}} \leqslant \sum_{p}\sum_{k=2}^{+\infty}\frac{1}{kp^k} \overset{(q.5)}{<} +\infty.$$

Quand $s \to 1$, on obtient alors une majoration de $\lim\limits_{s\to 1} \sum\limits_p \sum\limits_{k=2}^{+\infty} \dfrac{1}{kp^{ks}}$ par un réel, ce qui prouve sa finitude. De plus, toujours d'après la question 1, et par composition de limites, on a : $\lim\limits_{s\to 1} \ln(\zeta(s)) = +\infty$. La différence d'une limite infinie et d'une limite finie est infinie ; on en déduit : $\lim\limits_{s\to 1} \sum\limits_p \dfrac{1}{p^s} = +\infty$, ce qu'il fallait démontrer.

Cette limite ne peut être infinie que s'il y a un nombre infini de nombres premiers, ce qui démontre le résultat demandé en préambule. Néanmoins l'énoncé nous en demande davantage avec la non sommabilité de la famille $\left(\dfrac{1}{p}\right)_{p\in\mathbb{P}}$. Elle découle de l'identité (†) justifiée en remarque dans la question 1 : elle implique ici :

$$\sum_p \frac{1}{p} = \lim_{s\to 1} \sum_p \frac{1}{p^s} = +\infty,$$

d'où : $\sum\limits_p \dfrac{1}{p} = +\infty$, ce qu'il fallait démontrer. Cette démonstration est due à Euler.

Questions à se poser, réflexes à acquérir.

- Bien vérifier que ce raisonnement n'utilise nulle part le fait *a priori* que $\mathbb{P}$ soit infini (sinon le raisonnement serait circulaire). Peut-on se dispenser de vérifier que $\mathbb{P}$ est non vide dans ce raisonnement ?
- Retenir la *grande idée* de cette démonstration, puisqu'elle est au cœur de la plupart des méthodes analytiques utilisées en arithmétique : quelle est la clé de voûte ?
- Peut-on utiliser autrement le produit eulérien $\zeta(s) = \prod\limits_p \frac{1}{1-\frac{1}{p^s}}$ pour obtenir le résultat arithmétique voulu, mais en nous dispensant d'un argument compliqué d'interversion d'une limite et d'une somme ? Regarder une autre valeur de s intéressante.
- Si vous connaissez, de culture scientifique, un équivalent simple du n^{e} nombre premier p_n quand $n \to +\infty$: l'utiliser pour redémontrer la non sommabilité.

Deuxième partie : infinité de nombres premiers congrus à ± 1 modulo 4

7. Soient m et n deux entiers. Remarquons que la définition de χ implique que c'est une fonction nulle en tous les entiers pairs. Par conséquent, si l'un des deux entiers est pair, disons m par exemple, alors mn est pair également, donc d'après l'observation ci-avant on a : $\chi(m) = 0$, et : $\chi(mn) = 0$. L'égalité $\chi(mn) = \chi(m)\chi(n)$ est donc bien vérifiée dans ce cas-là.

   Supposons à présent m et n impairs. Ils sont congrus modulo 4 soit à 1, soit à -1. Il existe donc $(m', n') \in \mathbb{Z}^2$ et $(\varepsilon_1, \varepsilon_2) \in \{-1,1\}^2$ tels que : $m = 4m' + \varepsilon_1$, et : $n = 4n' + \varepsilon_2$. Or la définition de χ implique : $\forall k \in \mathbb{Z}$, $\forall \varepsilon \in \{-1,1\}$, $\chi(4k+\varepsilon) = \varepsilon$. Donc : $\chi(m)\chi(n) = \varepsilon_1\varepsilon_2$. De plus :

$$mn = (4m'+\varepsilon_1)(4n'+\varepsilon_2) = 4\underbrace{(4m'n' + \varepsilon_1 n' + \varepsilon_2 m')}_{\in\mathbb{Z}} + \underbrace{\varepsilon_1\varepsilon_2}_{\in\{-1,1\}},$$

   donc : $\chi(mn) = \varepsilon_1\varepsilon_2$. On a donc bien : $\chi(mn) = \chi(m)\chi(n)$, et l'égalité est vraie dans tous les cas. D'où le résultat.

8. Soit $s > 1$ un réel. On a : $\forall n \in \mathbb{N}\setminus\{0\}$, $\left|\dfrac{\chi(n)}{n^s}\right| \leqslant \dfrac{1}{n^s}$, et la série de Riemann $\sum\limits_{n\geqslant 1} \dfrac{1}{n^s}$ converge parce qu'elle est d'exposant $s > 1$. Par le théorème de comparaison des séries à termes positifs, la série $\sum\limits_{n\geqslant 1} \dfrac{\chi(n)}{n^s}$ converge absolument pour tout $s > 1$.

Questions à se poser, réflexes à acquérir. On démontre la convergence absolue pour tout $s > 1$ dans cette question, mais ce résultat semble supplanté par les questions suivantes, où l'on démontre la convergence pour tout $s > 0$. Pour autant j'affirme que cette question n'était pas inutile : pourquoi ? Où peut bien servir la convergence absolue, là où la convergence ne suffit pas ? Avoir cette interrogation en tête dans la suite du corrigé.

9. Soit $N \in \mathbb{N}$. Comme χ s'annule en tous les entiers pairs, on a : $S_N(s) = \sum_{\substack{n=1 \\ n \text{ impair}}}^{2N+1} \frac{\chi(n)}{n^s}$. Or l'application $\phi : n \mapsto 2n+1$ induit une bijection de $[\![0, N]\!]$ dans l'ensemble des entiers impairs compris entre 1 et $2N+1$. Ce changement d'indice (♠) implique donc :

$$S_N(s) = \sum_{\substack{n=1 \\ n \text{ impair}}}^{2N+1} \frac{\chi(n)}{n^s} = \sum_{n \in \phi([\![0,N]\!])} \frac{\chi(n)}{n^s} \overset{(\spadesuit)}{=} \sum_{n \in [\![0,N]\!]} \frac{\chi(\phi(n))}{(\phi(n))^s} = \sum_{n=1}^{N} \frac{\chi(2n+1)}{(2n+1)^s}.$$

Pour avoir l'identité de l'énoncé, il reste à démontrer : $\forall n \in \mathbb{N}$, $\chi(2n+1) = (-1)^n$. On va raisonner selon la parité de n. Soit $n \in \mathbb{N}$. Si n est pair, c'est-à-dire s'il existe $k \in \mathbb{N}$ tel que : $n = 2k$, alors la définition de χ implique immédiatement : $\chi(2n+1) = \chi(4k+1) = 1 = (-1)^n$. Si n est impair, c'est-à-dire s'il existe $k \in \mathbb{N}$ tel que : $n = 2k+1$, alors par 4-périodicité on a : $\chi(2n+1) = \chi(4k+3) = \chi(3) = \chi(-1) = -1 = (-1)^n$. Ainsi on a bien l'égalité voulue pour tout $n \in \mathbb{N}$, et on conclut qu'on a bien : $S_N(s) = \sum_{n=0}^{N} \frac{(-1)^n}{(2n+1)^s}$.

Montrons à présent que les suites $(S_{2N}(s))_{N\in\mathbb{N}}$ et $(S_{2N+1}(s))_{N\in\mathbb{N}}$ sont adjacentes. Nous allons démontrer que la première est strictement décroissante, la seconde strictement croissante (la stricte monotonie n'est pas utile pour démontrer qu'elles sont adjacentes : on en aura besoin pour la question suivante), et que la différence des deux converge vers 0. Pour tout $N \in \mathbb{N}$, on a :

$$S_{2(N+1)}(s) - S_{2N}(s) = \frac{1}{(2(2N+2)+1)^s} - \frac{1}{(2(2N+1)+1)^s} < 0,$$

parce que l'application $u \mapsto \frac{1}{u^s}$ est décroissante sur $]0, +\infty[$ si $s > 0$. Ainsi $(S_{2N}(s))_{N\in\mathbb{N}}$ est strictement décroissante. Un raisonnement en tous points analogues permet de démontrer la stricte croissance de $(S_{2N+1}(s))_{N\in\mathbb{N}}$. Enfin :

$$\forall N \in \mathbb{N}, \quad S_{2N+1}(s) - S_{2N}(s) = -\frac{1}{(2(2N+1)+1)^s},$$

et comme $s > 0$ on en déduit : $\lim_{N\to+\infty} (S_{2N+1}(s) - S_{2N}(s)) = 0$. Ceci achève de démontrer que les suites $(S_{2N}(s))_{N\in\mathbb{N}}$ et $(S_{2N+1}(s))_{N\in\mathbb{N}}$ sont adjacentes.

Remarque. Dans cette question, la suivante et la question 13, je vous fais implicitement redémontrer une partie du théorème des séries alternées. L'invoquer directement réduit la longueur des raisonnements. La raison pour laquelle je vous le fais redémontrer est pour vous préparer à une préoccupation fréquente en 2e année de classes préparatoires, dans ce qu'on appelle les *problèmes d'interversion* : le contrôle de la taille du reste d'une série. Montrer qu'il est « petit », en un sens que nous formulerons précisément lorsque nous définirons la convergence uniforme d'une série de fonctions, est souvent crucial pour résoudre ces problèmes. C'est ce que nous ferons dans les questions 13 à 15.

> **Questions à se poser, réflexes à acquérir.**
> — Bien comprendre la formule de changement d'indice que j'ai utilisée. Notamment : pourquoi l'indexation par $n \in \phi([\![0, N]\!])$ n'est pas dans le membre de l'égalité où figure $\frac{\chi(\phi(n))}{(\phi(n))^s}$. Éventuellement comparer avec la formule du changement de variable dans les intégrales, même si ce parallèle a ses limites. Si elle n'est pas toujours parlante pour vous : poser $n' = 2n+1$, etc., comme vous aviez peut-être l'occasion de faire en 1re année.
> — Au vu de la définition de χ, il aurait peut-être pu paraître plus naturel de scinder la somme $S_N(s)$ en quatre, selon la classe de congruence de n modulo 4. Le faire, et se convaincre que cela permet d'aboutir au résultat, puis se demander pourquoi j'ai malgré tout privilégié ma démonstration.

10. Soit $s > 0$. Le théorème des suites adjacentes, appliqué aux suites $(S_{2N}(s))_{N\in\mathbb{N}}$ et $(S_{2N+1}(s))_{N\in\mathbb{N}}$, assure qu'elles convergent vers une limite commune, et donc la suite $(S_N(s))_{N\in\mathbb{N}}$ converge également vers

cette même limite. Étant donné que : $\forall N \in \mathbb{N}$, $S_N(s) = \sum_{n=0}^{N} \frac{(-1)^n}{(2n+1)^s}$, cela équivaut à la convergence de la série $\sum_{n\geqslant 0} \frac{(-1)^n}{(2n+1)^s}$.

Pour en déduire que la série $\sum_{n\geqslant 1} \frac{\chi(n)}{n^s}$ converge pour tout $s > 0$, il faut un argument supplémentaire. En effet, si l'on pose : $\forall N \in \mathbb{N}\setminus\{0\}$, $T_N(s) = \sum_{n=1}^{N} \frac{\chi(n)}{n^s}$, alors la convergence de la série $\sum_{n\geqslant 1} \frac{\chi(n)}{n^s}$ équivaut par définition à la convergence de la suite $(T_N(s))_{N\geqslant 1}$, mais ce qui précède ne démontre que la convergence que la suite extraite $(T_{2N+1}(s))_{N\geqslant 1} = (S_N(s))_{N\geqslant 1}$. Ce n'est pas très gênant, puisque l'égalité : $\forall N \in \mathbb{N}\setminus\{0\}$, $T_{2N}(s) = T_{2N+1}(s) - \frac{\chi(2N+1)}{(2N+1)^s}$ montre que, pour avoir la convergence de la suite extraite $(T_{2N}(s))_{N\geqslant 1}$, il suffit de montrer : $\lim_{N\to+\infty} \frac{\chi(2N+1)}{(2N+1)^s} = 0$. C'est immédiat grâce au fait que χ soit bornée et à l'inégalité $s > 0$. Puisque les suites extraites des indices pairs et des indices impairs convergent, il en est de même de la suite $(T_N(s))_{N\geqslant 1}$, et donc la série $\sum_{n\geqslant 1} \frac{\chi(n)}{n^s}$ converge : ce qu'on voulait démontrer.

On a montré davantage : puisque $(T_N(s))_{N\geqslant 1}$ converge vers $\sum_{n=1}^{+\infty} \frac{\chi(n)}{n^s} = \mathrm{L}_\chi(s)$, il en est de même de sa suite extraite $(T_{2N+1}(s))_{N\geqslant 1} = (S_N(s))_{N\geqslant 1}$. Or ce qui précède montre que cette dernière suite converge aussi vers $\sum_{n=0}^{+\infty} \frac{(-1)^n}{(2n+1)^s}$. Par unicité de la limite :

$$\mathrm{L}_\chi(s) = \sum_{n=0}^{+\infty} \frac{(-1)^n}{(2n+1)^s}.$$

Démontrons enfin : $\mathrm{L}_\chi(s) > 0$. Pour cela, on rappelle que d'après le théorème des suites adjacentes, la limite commune de $(S_{2N}(s))_{N\geqslant 0}$ et de $(S_{2N+1}(s))_{N\geqslant 0}$ (qui est $\mathrm{L}_\chi(s)$ comme on vient de le voir) est encadrée par ces deux suites. Mieux : comme ces deux suites sont strictement monotones, l'encadrement en question fait intervenir des inégalités strictes (nous laissons le lecteur se convaincre de ce point relativement simple à démontrer). On a donc, en particulier :

$$\mathrm{L}_\chi(s) > S_1(s) = 1 - \frac{1}{3^s} > 0,$$

d'où le résultat.

Questions à se poser, réflexes à acquérir. Se convaincre que le paragraphe laborieux où j'introduis $(T_N)_{N\geqslant 1}$ est malgré tout nécessaire. Notamment : se convaincre que la convergence de $\sum_{n\geqslant 0} u_n$ peut être fausse bien qu'il y ait convergence pour les indices impairs (néanmoins il ne manque pas grand'chose pour que cela devienne vrai : proposer un énoncé très simple du type : « si $\left(\sum_{n=0}^{2N+1} u_n\right)_{N\in\mathbb{N}}$ converge et si... alors $\sum_{n\geqslant 0} u_n$ converge »). Ou encore : se convaincre que même si l'on a deux expressions alternatives des sommes partielles, cela ne veut pas dire qu'elles correspondent à deux séries de même nature, et par extension les sommes ne sont pas nécessairement égales lorsqu'on passe à la limite.

11. Le raisonnement aboutissant à l'identité de cette question ressemble en de nombreux points à celui des questions 2 et 3. Nous l'imitons ici en ne détaillant que les arguments qui diffèrent significativement. L'ensemble $\mathcal{N}(T)$ qui apparaîtra ci-dessous est celui défini dans l'énoncé avant la question 2, et les ensembles I_n sont à nouveau définis par : $\forall n \in \mathbb{N}\setminus\{0\}$, $I_n = \{(m_1,\dots,m_r) \in \mathbb{N}^r \mid n = p_1^{m_1}\cdots p_r^{m_r}\}$.
Soient $s > 1$ et $T \geqslant 2$ deux réels. Notons d'abord que la question 7 permet d'obtenir le résultat suivant par récurrence : pour tout entier naturel non nul n de la forme : $n = p_1^{m_1}\times\cdots\times p_r^{m_r}$, avec $r \in \mathbb{N}\setminus\{0\}$, $(p_1,\dots,p_r) \in \mathbb{N}^r$ et $(m_1,\dots,m_r) \in \mathbb{N}^r$, on a : $\chi(n) = \chi(p_1)^{m_1}\cdots\chi(p_r)^{m_r}$.

Or : $\forall p \in \mathbb{P}$, $\dfrac{1}{1-\frac{\chi(p)}{p^s}} = \displaystyle\sum_{m=0}^{+\infty} \frac{\chi(p)^m}{p^{ms}}$, donc quand on fait le produit indexé par les nombres premiers $p_1, \dots, p_r$ inférieurs ou égaux à $T \geqslant 2$, on obtient :

$$\prod_{p\leqslant T} \frac{1}{1-\frac{\chi(p)}{p^s}} = \prod_{i=1}^{r} \sum_{m_i=0}^{+\infty} \frac{\chi(p_i)^{m_i}}{p_i^{m_i s}} = \sum_{(m_1,\dots,m_r)\in\mathbb{N}^r} \frac{\chi(p_1)^{m_1}\cdots\chi(p_r)^{m_r}}{(p_1^{m_1}\cdots p_r^{m_r})^s} = \sum_{n=1}^{+\infty} \frac{\operatorname{card}(I_n)\,\chi(n)}{n^s} = \sum_{n\in\mathcal{N}(T)} \frac{\chi(n)}{n^s}.$$

Cette fois-ci nous ne faisons plus un produit de sommes de réels positifs (à cause de χ), donc un argument supplémentaire est nécessaire pour le développer ainsi, puis pour utiliser le théorème de sommation par paquets. Ces égalités sont ici valables parce que pour tout $i \in [\![1, r]\!]$, la famille $\left(\dfrac{\chi(p_i)^m}{p_i^{ms}}\right)_{m\in\mathbb{N}}$ est sommable : en effet la série géométrique $\displaystyle\sum_{m\geqslant 0} x^m$ converge absolument pour tout $x \in]-1,1[$, donc en particulier pour $x = \dfrac{\chi(p_i)}{p_i^s}$, puisqu'on rappelle que χ est bornée par 1 en valeur absolue et que tout nombre premier est supérieur ou égal à 2. La famille obtenue par produit est également sommable, ce qui permet d'utiliser le théorème de sommation par paquets.
Pour conclure, on a :

$$\left|\mathrm{L}_\chi(s) - \prod_{p\leqslant T} \frac{1}{1-\frac{\chi(p)}{p^s}}\right| = \left|\sum_{n\notin\mathcal{N}(T)} \frac{\chi(n)}{n^s}\right| \leqslant \sum_{n\notin\mathcal{N}(T)} \left|\frac{\chi(n)}{n^s}\right| \leqslant \sum_{n>T} \left|\frac{\chi(n)}{n^s}\right|.$$

Nous avons là le reste d'une série convergente d'après la question 8, donc il tend vers 0 quand $T \to +\infty$. Donc, d'après le théorème des gendarmes : $\displaystyle\lim_{T\to+\infty} \prod_{p\leqslant T} \frac{1}{1-\frac{\chi(p)}{p^s}} = \mathrm{L}_\chi(s)$, d'où le résultat.

Questions à se poser, réflexes à acquérir.

— Les résultats sur les familles sommables concernent les produits de *deux* sommes : pourquoi est-ce toujours vrai pour un produit de r sommes ?

— Est-ce que cette identité n'aurait pas permis de démontrer, sans passer par des suites adjacentes, que $\mathrm{L}_\chi(s)$ est strictement positif ? En effet, il n'est pas difficile d'observer que le produit ne fait intervenir que des réels strictement positifs.

12. Soit $s > 1$ un réel. On s'est assuré dans la question 10 que $\mathrm{L}_\chi(s)$ est strictement positif, ce qui permet de considérer son logarithme. Grâce à l'identité de la question précédente, et en imitant le raisonnement de la question 4, on a :

$$\ln(\mathrm{L}_\chi(s)) = \sum_p \sum_{k=1}^{+\infty} \frac{(\chi(p)/p^s)^k}{k} = \sum_p \sum_{k=1}^{+\infty} \frac{(\chi(p))^k}{kp^{ks}}.$$

On a aussi : $\ln(\zeta(s)) = \displaystyle\sum_p \sum_{k=1}^{+\infty} \frac{1}{kp^{ks}}$. En sommant ces deux égalités, on obtient :

$$\ln(\zeta(s)) + \ln(\mathrm{L}_\chi(s)) = \sum_p \sum_{k=1}^{+\infty} \frac{1+(\chi(p))^k}{kp^{ks}} = \sum_p \frac{1+\chi(p)}{p^s} + \sum_p \sum_{k=2}^{+\infty} \frac{1+(\chi(p))^k}{kp^{ks}}.$$

Comme : $\forall n \in \mathbb{Z}$, $|\chi(n)| \leqslant 1$, la première somme du membre de droite est à termes positifs. On peut donc utiliser le théorème de sommation par paquets pour scinder cette somme en trois sommes, dont la première est indexée par les nombres premiers congrus à 1 modulo 4, la seconde indexée par ceux congrus à -1 modulo 4, et enfin la dernière uniquement indexée par $p = 2$. En faisant cela, on parcourt bien tous les nombres premiers : aucun nombre premier n'est congru à 0 modulo 4, sinon il serait multiple de 4, et seul $p = 2$ est congru à 2 modulo 4, puisque la relation $p = 4k + 2 = 2(2k + 1)$, avec $k \in \mathbb{N}$, contredit la primalité de p sauf pour $k = 0$.

Comme : $\chi(2) = 0$, et :

$$\forall p \in \mathbb{P}, \quad \chi(p) = \begin{cases} 1 & \text{si} \quad p \equiv 1 \bmod 4, \\ -1 & \text{si} \quad p \equiv -1 \bmod 4, \end{cases}$$

la sommation par paquets nous donne :

$$\sum_p \frac{1+\chi(p)}{p^s} = \sum_{p\equiv 1[4]} \frac{1+\chi(p)}{p^s} + \sum_{p\equiv -1[4]} \frac{1+\chi(p)}{p^s} + \frac{1}{2^s} = 2\sum_{p\equiv 1[4]} \frac{1}{p^s} + \frac{1}{2^s}.$$

On a donc bien montré :

$$\ln(\zeta(s)) + \ln\left(\mathrm{L}_\chi(s)\right) = 2\sum_{p\equiv 1[4]} \frac{1}{p^s} + \frac{1}{2^s} + \sum_p \sum_{k=2}^{+\infty} \frac{1+(\chi(p))^k}{kp^{ks}}. \tag{$*$}$$

Questions à se poser, réflexes à acquérir. Plus généralement, quelles sont les classes de congruence possibles pour un nombre premier modulo b, quand b est un entier naturel non nul ?

13. Soit $s > 1$. On a vu dans la question 9 que les suites $(S_{2N}(s))_{N\in\mathbb{N}}$ et $(S_{2N+1}(s))_{N\in\mathbb{N}}$ sont adjacentes (la première étant décroissante et la seconde croissante), et on en a déduit dans la question 10 qu'elles convergent et sont de limite $\mathrm{L}_\chi(s)$. Toujours d'après le théorème des suites adjacentes, on a :

$$\forall N \in \mathbb{N}, \quad S_{2N+1}(s) \underset{(\clubsuit)}{\leqslant} \mathrm{L}_\chi(s) \underset{(\heartsuit)}{\leqslant} S_{2N}(s).$$

Déduisons-en l'inégalité demandée. Soit $N \in \mathbb{N}$. Faisons une distinction de cas. Supposons N pair. Il existe $k \in \mathbb{N}$ tel que : $N = 2k$. Alors, d'après l'encadrement ci-dessus, on a :

$$|\mathrm{L}_\chi(s) - S_N(s)| \overset{(\heartsuit)}{=} S_{2k}(s) - \mathrm{L}_\chi(s) \overset{(\clubsuit)}{\leqslant} S_{2k}(s) - S_{2k+1}(s) = -\frac{(-1)^{2k+1}}{(2(2k+1)+1)^s} = \frac{1}{(2N+3)^s}.$$

Traitons à présent le cas impair : s'il existe $k \in \mathbb{N}$ tel que : $N = 2k+1$, alors, toujours d'après l'encadrement ci-dessus :

$$|\mathrm{L}_\chi(s) - S_N(s)| \overset{(\clubsuit)}{=} \mathrm{L}_\chi(s) - S_{2k+1}(s) \overset{(\heartsuit)}{\leqslant} S_{2k+2}(s) - S_{2k+1}(s) = \frac{(-1)^{2k+2}}{(2(2k+2)+1)^s} = \frac{1}{(2N+3)^s}.$$

Dans tous les cas, on a montré : $|\mathrm{L}_\chi(s) - S_N(s)| \leqslant \dfrac{1}{(2N+3)^s}$, d'où le résultat en utilisant le fait l'inégalité : $(2N+3)^s \geqslant 2N+3$, valable pour tout $s > 1$ parce que : $2N+3 \geqslant 3 \geqslant 1$.

Remarque. On a implicitement redémontré que le reste d'une série vérifiant les hypothèses du théorème des séries alternées est absolument majoré par son premier terme.

14. **Coquille de l'énoncé.** On doit montrer : $|\mathrm{L}_\chi(s) - \mathrm{L}_\chi(1)| \leqslant \dfrac{2}{2N+3} + \left|\displaystyle\sum_{n=1}^{2N+1} \chi(n)\left(\frac{1}{n^s} - \frac{1}{n}\right)\right|$, pour tous $N \in \mathbb{N}$ et $s > 1$.

Soit $s > 1$. Pour tout $N \in \mathbb{N}$, on a d'après l'inégalité triangulaire et la question précédente :

$$\begin{aligned} |\mathrm{L}_\chi(s) - \mathrm{L}_\chi(1)| &\leqslant |\mathrm{L}_\chi(s) - S_N(s)| + |S_N(s) - S_N(1)| + |S_N(1) - \mathrm{L}_\chi(1)| \\ &\leqslant \frac{1}{2N+3} + \left|\sum_{n=1}^{2N+1} \frac{\chi(n)}{n^s} - \sum_{n=1}^{2N+1} \frac{\chi(n)}{n}\right| + \frac{1}{2N+3} \\ &= \frac{2}{2N+3} + \left|\sum_{n=1}^{2N+1} \chi(n)\left(\frac{1}{n^s} - \frac{1}{n}\right)\right|, \end{aligned}$$

d'où le résultat.

Questions à se poser, réflexes à acquérir. Comment aurait-on pu penser à ce découpage en trois termes *via* l'inégalité triangulaire ? Idéalement, comprendre comment on pouvait y penser *sans* la question précédente (ce sera sans doute plus facile après avoir découvert la convergence uniforme des séries de fonctions).

15. Nous allons démontrer que $|\mathrm{L}_\chi(s) - \mathrm{L}_\chi(1)|$ devient arbitrairement petit pour s au voisinage de 1 par valeurs supérieures. On reprend l'inégalité de la question précédente. Soit $\varepsilon > 0$. Comme : $\lim\limits_{N\to+\infty} \dfrac{2}{2N+3} = 0$, il existe un rang $N_0 \in \mathbb{N}$ tel que, pour tout entier $N \geqslant N_0$, on ait : $\dfrac{2}{2N+3} \leqslant \dfrac{\varepsilon}{2}$.
Fixons un tel rang N_0, et prenons désormais : $N = N_0$.
Pour tout $n \in [\![1,2N_0+1]\!]$, l'application $s \mapsto \dfrac{1}{n^s}$ est continue sur $\mathbb{R}$, et donc en particulier en 1, donc l'application $s \mapsto \sum\limits_{n=1}^{2N_0+1} \dfrac{\chi(n)}{n^s}$ est continue en 1 en tant que combinaison linéaire de fonctions continues. On en déduit qu'il existe $\eta > 0$ tel que, pour tout $s > 1$ vérifiant : $|s-1| < \eta$, on ait : $\left|\sum\limits_{n=1}^{2N_0+1} \chi(n)\left(\dfrac{1}{n^s} - \dfrac{1}{n}\right)\right| \leqslant \dfrac{\varepsilon}{2}$.
Les majorations ci-dessus permettent de démontrer :

$$\forall \varepsilon > 0,\ \exists \eta > 0,\ \forall s \in]1,+\infty[,\quad (|s-1| < \eta \Longrightarrow |\mathrm{L}_\chi(s) - \mathrm{L}_\chi(1)| \leqslant \varepsilon)\,,$$

c'est-à-dire : $\lim\limits_{s\to1^+} \mathrm{L}_\chi(s) = \mathrm{L}_\chi(1)$. Or la question 10 montre que $\mathrm{L}_\chi(1)$ est un réel strictement positif. Comme la fonction logarithme est continue sur $\mathbb{R}_+^*$, on en déduit : $\lim\limits_{s\to1^+} \ln(\mathrm{L}_\chi(s)) = \ln(\mathrm{L}_\chi(1)) \neq \pm\infty$. D'où le résultat.
Remarque. Plus précisément, on sait démontrer que l'on a : $\mathrm{L}_\chi(1) = \dfrac{\pi}{4}$. Une des questions de ce problème peut vous donner l'inspiration nécessaire pour savoir le démontrer : le faire !

Questions à se poser, réflexes à acquérir. Pourquoi poser $N = N_0$? Qu'est-ce qui pose problème si on ne le fait pas, en se contentant de prendre $N \geqslant N_0$ quelconque ? Pourquoi ai-je pensé à revenir à la définition de la limite ?

16. Comme dans la question 6, l'idée est de démontrer que la limite quand $s \to 1$ de $\sum\limits_{p\equiv1[4]} \dfrac{1}{p^s}$ est infinie grâce à l'identité $(*)$. Les questions précédentes permettent d'obtenir la limite quand $s \to 1$ de toutes les quantités sauf $\sum\limits_p \sum\limits_{k=2}^{+\infty} \dfrac{1+(\chi(p))^k}{kp^{ks}}$. Pour l'avoir, on note que $s \mapsto \dfrac{1+(\chi(p))^k}{kp^{ks}}$ est positive et décroissante sur $[1,+\infty[$ parce que $s \mapsto \dfrac{1}{p^{ks}}$ l'est (argument déjà utilisé maintes fois) et $1+(\chi(p))^k$ est positif ou nul pour tout $p \in \mathbb{P}$ et tout $k \in \mathbb{N}\setminus\{0\}$. Donc, d'après l'identité $(\dagger)$ justifiée dans la remarque de la question 1, on a :

$$\lim_{s\to1} \sum_p \sum_{k=2}^{+\infty} \frac{1+(\chi(p))^k}{kp^{ks}} = \sum_p \sum_{k=2}^{+\infty} \lim_{s\to1} \frac{1+(\chi(p))^k}{kp^{ks}} = \sum_p \sum_{k=2}^{+\infty} \frac{1+(\chi(p))^k}{kp^k} \leqslant \sum_p \sum_{k=2}^{+\infty} \frac{2}{kp^k} < +\infty,$$

la finitude de la dernière somme ayant été démontrée dans la question 5. Donc finalement, dans le membre de droite de l'égalité suivante :

$$\sum_{p\equiv1[4]} \frac{1}{p^s} = \frac{\ln(\zeta(s))}{2} + \frac{\ln(\mathrm{L}_\chi(s))}{2} - \frac{1}{2^{s+1}} - \frac{1}{2}\sum_p \sum_{k=2}^{+\infty} \frac{1+(\chi(p))^k}{kp^{ks}},$$

toutes les quantités ont une limite finie quand $s \to 1$, sauf $\dfrac{\ln(\zeta(s))}{2}$ qui tend vers $+\infty$. On en déduit :

$$\lim_{s\to1} \sum_{p\equiv1[4]} \frac{1}{p^s} = +\infty,$$

ce qui n'est possible que si la somme est à support infini (puisque le terme général, lui, admet une limite finie en 1) : ceci démontre qu'il existe une infinité de nombres premiers p tels que $p \equiv 1 \bmod 4$ (mieux : que la somme $\sum_{p\equiv 1[4]} \frac{1}{p}$ est infinie, toujours grâce à (†)).

Culture scientifique. Soit $A \subseteq \mathbb{P}$. On définit la *densité analytique* de A, lorsqu'elle existe, comme étant la limite suivante :

$$\lim_{s\to 1} \frac{\sum\limits_{p\in A} \frac{1}{p^s}}{\sum\limits_{p} \frac{1}{p^s}}.$$

C'est une définition moins intuitive que la *densité naturelle* (qui est définie par la limite suivante, quand elle existe : $\lim\limits_{n\to+\infty} \frac{\operatorname{card}(A\cap[\![1,n]\!])}{\operatorname{card}(\mathbb{P}\cap[\![1,n]\!])}$), mais elle n'est pas sans rapport avec cette dernière (si A admet une densité naturelle, alors elle admet aussi une densité analytique, et elles sont égales ; la réciproque est fausse) et elle est plus facilement calculable que la densité naturelle. Nous vous laissons vérifier, grâce aux deux questions précédentes, que la densité analytique de l'ensemble des nombres premiers congrus à 1 modulo 4, ou à -1 modulo 4, est dans chaque cas égale à $\frac{1}{2}$ (il est possible de démontrer qu'il en est de même pour la densité naturelle). Cela implique en particulier, *grosso modo*, que les nombres premiers sont équitablement répartis dans chacune de ces deux classes de congruence. Plus généralement, l'ensemble des nombres premiers congrus à a modulo b, avec a et b premiers entre eux, admet une densité naturelle égale à $\frac{1}{\varphi(b)}$ où φ est l'indicatrice d'Euler.

17. On reprend le raisonnement de la question 12, mais en soustrayant les deux logarithmes au lieu de les additionner. Nous vous laissons compléter les détails grâce à la résolution de la question 12. On a :

$$\ln(\zeta(s)) - \ln(\mathrm{L}_\chi(s)) = \sum_p \frac{1-\chi(p)}{p^s} + \sum_p \sum_{k=2}^{+\infty} \frac{1-(\chi(p))^k}{kp^{ks}} = 2\sum_{p\equiv 3[4]} \frac{1}{p^s} + \frac{1}{2^s} + \sum_p \sum_{k=2}^{+\infty} \frac{1-(\chi(p))^k}{kp^{ks}}$$

et, comme dans la question précédente, cette identité permet d'écrire $\sum\limits_{p\equiv 3[4]} \frac{1}{p^s}$ comme combinaison linéaire de $\ln(\zeta(s))$ (qui admet une limite infinie quand $s \to 1$) et de quantités ayant une limite finie quand $s \to 1$. Donc : $\lim\limits_{s\to 1} \sum\limits_{p\equiv 3[4]} \frac{1}{p^s} = +\infty$, ce qui n'est possible que si la somme est à support infini, c'est-à-dire s'il existe une infinité de nombres premiers p tels que $p \equiv 3 \bmod 4$: d'où le résultat.

> **Questions à se poser, réflexes à acquérir.** Avec les deux dernières questions, comprendre ce qui a motivé la définition de χ, et se convaincre qu'on ne pouvait pas pour autant « faire n'importe quoi », par exemple démontrer qu'il existe une infinité de nombres premiers p tels que $p \equiv 2 \bmod 4$ en posant $\chi(2) = 1$, $\chi(1) = \chi(-1) = -1$, et $\chi(0) = 0$. Où coincerait notre raisonnement ?

☢ 18. Nous allons traiter ensemble les deux cas demandés . Soit $b \in \{3,6\}$. Remarquons que pour ces deux valeurs potentielles de b, pour tout $x \in [\![1,b]\!]$, on a :

$$\operatorname{pgcd}(x,b) = 1 \iff x \in \{1, b-1\}.$$

La vérification est immédiate par recensement exhaustif des entiers entre 1 et b, et en calculant les pgcd. Soit $\chi : \mathbb{Z} \to \mathbb{R}$ l'application définie par : $\chi(1) = 1$, $\chi(b-1) = -1$, et : $\forall a \in \{0\} \cup [\![2, b-2]\!]$, $\chi(a) = 0$ (rappelons que $[\![2,b-2]\!]$ est, d'après l'équivalence ci-dessus, l'ensemble des entiers entre 1 et b qui ne sont pas premiers avec b), puis étendue à tout entier relatif par b-périodicité. Cette définition implique en particulier que l'on a : $\forall k \in \mathbb{Z}$, $\forall \varepsilon \in \{-1,1\}$, $\chi(kb+\varepsilon) = \varepsilon$, et que χ s'annule en tous les entiers relatifs, y compris hors de $[\![2,b-2]\!]$, qui ne sont pas premiers avec b. Pour cette dernière affirmation, on utilise non seulement le fait que χ soit b-périodique, mais aussi que : $\forall x \in \mathbb{Z}$, $\forall k \in \mathbb{Z}$, $\operatorname{pgcd}(x,b) = \operatorname{pgcd}(x-kb, b)$.

Montrons que χ vérifie l'identité : $\forall (m,n) \in \mathbb{Z}^2$, $\chi(mn) = \chi(m)\chi(n)$: cette propriété fut en effet nécessaire pour obtenir le produit eulérien de la question 11. Soit $(m,n) \in \mathbb{Z}^2$. Si m ou n n'est pas

premier avec b, disons m par exemple, alors mn n'est pas premier avec b non plus (en effet, un diviseur commun à m et b divise en particulier mn et b). On a donc : $\chi(mn) = 0$, et : $\chi(m) = 0$, donc : $\chi(mn) = \chi(m)\chi(n)$. Si m et n sont premiers avec b, alors d'après l'hypothèse effectuée ils sont congrus soit à 1, soit à -1 modulo b. Soient, donc, $(m', n') \in \mathbb{Z}^2$ et $(\varepsilon_1, \varepsilon_2) \in \{-1,1\}^2$ tels que : $m = bm' + \varepsilon_1$, et : $n = bn' + \varepsilon_2$. Comme on l'a évoqué plus haut, on a : $\chi(m)\chi(n) = \varepsilon_1\varepsilon_2$, et de plus :

$$mn = (bm' + \varepsilon_1)(bn' + \varepsilon_2) = b\underbrace{(bm'n' + \varepsilon_1 n' + \varepsilon_2 m')}_{\in\mathbb{Z}} + \underbrace{\varepsilon_1\varepsilon_2}_{\in\{-1,1\}},$$

donc : $\chi(mn) = \varepsilon_1\varepsilon_2$. On a donc bien : $\chi(mn) = \chi(m)\chi(n)$, et l'égalité est vraie dans tous les cas. D'où le résultat.

Justifions à présent pourquoi tous les résultats démontrés dans cette partie restent vrais ou adaptables avec cette nouvelle fonction χ :

— question 8 : la série $\displaystyle\sum_{n\geqslant 1} \frac{\chi(n)}{n^s}$ converge toujours absolument pour tout $s > 1$, puisqu'on n'utilisait que le fait que χ soit bornée par 1 en valeur absolue (et cela reste vrai) ;

— questions 9 et 10 : nous ne pouvons plus faire apparaître des suites adjacentes, et donc toutes les conséquences de cette question sont soit à adapter, soit fausses ; néanmoins un raisonnement sur la congruence de n modulo b permet de démontrer que pour tout $s > 0$ la série $\displaystyle\sum_{n\geqslant 1} \frac{\chi(n)}{n^s}$ converge, et qu'on a :

$$\forall s > 0, \quad \mathrm{L}_\chi(s) = 1 + \sum_{n=1}^{+\infty} \left(\frac{1}{(bn+1)^s} - \frac{1}{(bn-1)^s} \right),$$

la convergence se démontrant grâce à la relation : $\displaystyle\frac{1}{(bn+1)^s} - \frac{1}{(bn-1)^s} = \underset{n\to+\infty}{O}\left(\frac{1}{n^{s+1}}\right)$ (factoriser les dénominateurs par $(bn)^s$ et faire un développement asymptotique), et à une comparaison à une série de Riemann convergente ; le signe de $\mathrm{L}_\chi(s)$ est plus délicat à obtenir : nous allons démontrer quelque chose de moins fort : 1° on a : $\mathrm{L}_\chi(1) > 0$, 2° l'application $\mathrm{L}_{\chi|[1,+\infty[}$ est continue en 1 (et donc strictement positive au voisinage de 1), ce qui sera suffisant pour notre affaire : après tout, le résultat que l'on souhaite démontrer ne dépendra que d'un calcul de limite en 1 par valeurs supérieures ; le point 2° sera démontré plus bas, et pour le point 1° nous utilisons une comparaison série-intégrale* : l'application $t \mapsto \displaystyle\frac{1}{bt+1} - \frac{1}{bt-1}$ est croissante sur $\left]\frac{1}{b}, +\infty\right[$, sa dérivée étant égale à $t \mapsto \displaystyle\frac{(2b)^2 t}{((bt)^2-1)^2}$, ce qui permet d'écrire :

$$\begin{aligned}\mathrm{L}_\chi(1) &\geqslant 1 + \left(\frac{1}{b+1} - \frac{1}{b-1}\right) + \lim_{N\to+\infty} \int_1^N \left(\frac{1}{bt+1} - \frac{1}{bt-1}\right) \mathrm{d}t \\ &= 1 + \left(\frac{1}{b+1} - \frac{1}{b-1}\right) + \frac{\ln(b-1) - \ln(b+1)}{b},\end{aligned}$$

et pour $b = 3$ cela donne la minoration : $\mathrm{L}_\chi(1) \geqslant \dfrac{3}{4} - \dfrac{\ln(2)}{3}$; ce minorant est strictement positif si et seulement si : $\ln(2) < \dfrac{9}{4}$, ce qui est trivialement vrai puisque $\ln(2) < \ln(e) = 1$; un raisonnement analogue permet de s'assurer que $\mathrm{L}_\chi(1) > 0$ pour $b = 6$ (on peut par exemple minorer trivialement $\ln(5) - \ln(7) = \ln\left(1 - \dfrac{2}{7}\right)$ par $-\dfrac{2}{7} - \dfrac{2^2}{7^2}$ grâce à la formule de Taylor avec reste intégral) ;

— question 11 : la formule reste valable pour notre nouvelle définition de χ, étant donné qu'on a seulement eu besoin de l'identité $\chi(mn) = \chi(m)\chi(n)$ valable pour tous entiers m et n, et de la majoration $|\chi| \leqslant 1$: ces deux résultats restent vrais ici ;

$*$. On sait démontrer, grâce aux techniques de MP, que $\mathrm{L}_\chi(1)$ vaut $\frac{\pi}{3\sqrt{3}}$ si $b = 3$ et $\frac{\pi}{2\sqrt{3}}$ si $b = 6$. La valeur en 1 de cette fonction a un sens arithmétique extrêmement profond qu'il ne m'est pas possible d'aborder en classes préparatoires.

— question 12 : l'identité n'est plus démontrée pour tout $s > 0$, mais pour tout s supérieur à 1 et dans son voisinage : nous avons en effet dit plus haut que nous ne parviendrons pas à avoir la stricte positivité de L_χ sur un domaine plus grand (d'ailleurs, cette positivité reste à démontrer : voir plus bas) ; cette fois-ci, on utilise le théorème de sommation par paquets pour scinder les sommes selon la classe de n modulo b, le reste du raisonnement étant en tout point identique puisque $\chi(n) = 0$ pour tout n hors de la classe de 1 ou -1 modulo b, et la quantité $1 + \chi(p)$ est nulle pour $p \equiv -1 \bmod b$; on a alors, pour tout s dans un voisinage convenable de 1 :

$$\ln(\zeta(s)) + \ln\left(\mathrm{L}_\chi(s)\right) = 2 \sum_{p \equiv 1[b]} \frac{1}{p^s} + \frac{1}{3^s} + \sum_p \sum_{k=2}^{+\infty} \frac{1 + (\chi(p))^k}{k p^{ks}},$$

du moins si $b = 3$; pour $b = 6$, nous devons encore ajouter $\dfrac{1}{2^s}$ au membre de droite ; pour ne retenir que la congruence $p \equiv -1 \bmod b$, on fait une soustraction au lieu d'une addition ;

— questions 13 à 15 : c'est le plus difficile à généraliser, et cela mérite de plus amples détails ; soit $s \geqslant 1$; on commence par majorer $|\mathrm{L}_\chi(s) - \mathrm{L}_\chi(1)|$ *via* l'inégalité triangulaire :

$$\begin{aligned}\forall N \in \mathbb{N} \setminus \{0\}, \quad |\mathrm{L}_\chi(s) - \mathrm{L}_\chi(1)| &\leqslant \left|\mathrm{L}_\chi(s) - \sum_{n=1}^{N} \frac{\chi(n)}{n^s}\right| + \left|\sum_{n=1}^{N} \frac{\chi(n)}{n^s} - \sum_{n=1}^{N} \frac{\chi(n)}{n}\right| + \left|\sum_{n=1}^{N} \frac{\chi(n)}{n} - \mathrm{L}_\chi(1)\right| \\ &= \left|\sum_{n=N+1}^{+\infty} \frac{\chi(n)}{n^s}\right| + \left|\sum_{n=1}^{N} \frac{\chi(n)}{n^s} - \sum_{n=1}^{N} \frac{\chi(n)}{n}\right| + \left|\sum_{n=N+1}^{+\infty} \frac{\chi(n)}{n}\right| ;\end{aligned}$$

et pour majorer les deux restes, nous allons utiliser ce qu'on appelle une transformation d'Abel (le nom est hors programme, mais la technique est largement répandue) ; pour tout $N \in \mathbb{N}$, posons : $V_N(s) = \sum_{n=0}^{N} \chi(n)$; il est évident que la suite $(V_N(s))_{N \in \mathbb{N}}$ est bornée par 1 ; pour tout $N \in \mathbb{N}$ et tout $N' \in \mathbb{N}$ supérieur à $N + 1$, on a :

$$\begin{aligned}\sum_{n=N+1}^{N'} \frac{\chi(n)}{n^s} = \sum_{n=N+1}^{N'} \frac{V_n(s) - V_{n-1}(s)}{n^s} &= \sum_{n=N+1}^{N'} \frac{V_n(s)}{n^s} - \sum_{n=N+1}^{N'} \frac{V_{n-1}(s)}{n^s} \\ &= \sum_{n=N+1}^{N'} \frac{V_n(s)}{n^s} - \sum_{n=N}^{N'-1} \frac{V_n(s)}{(n+1)^s} \\ &= \frac{V_{N'}(s)}{(N')^s} - \frac{V_N(s)}{(N+1)^s} + \sum_{n=N+1}^{N'-1} V_n(s) \left(\frac{1}{n^s} - \frac{1}{(n+1)^s}\right) ;\end{aligned}$$

la série $\sum_{n \geqslant 1} V_n(s) \left(\dfrac{1}{n^s} - \dfrac{1}{(n+1)^s}\right)$ est absolument convergente, puisque son terme général est dominé par $\dfrac{1}{n^s} - \dfrac{1}{(n+1)^s}$ (terme qui est positif car $s \geqslant 1 > 0$), et la série $\sum_{n \geqslant 1} \left(\dfrac{1}{n^s} - \dfrac{1}{(n+1)^s}\right)$ est convergente par le lien suite-série (étant donné que : $\lim_{n \to +\infty} \dfrac{1}{n^s} = 0$) ; de plus, la limite quand $N' \to +\infty$ de la quantité $\dfrac{V_{N'}(s)}{(N')^s}$ est 0 parce que $(V_N(s))_{N \in \mathbb{N}}$ est bornée ; par conséquent, quand on prend la limite quand $N' \to +\infty$ dans l'égalité précédente, on obtient pour tout $N \in \mathbb{N}$:

$$\sum_{n=N+1}^{+\infty} \frac{\chi(n)}{n^s} = -\frac{V_N(s)}{(N+1)^s} + \sum_{n=N+1}^{+\infty} V_n(s) \left(\frac{1}{n^s} - \frac{1}{(n+1)^s}\right),$$

et donc, toujours pour tout $N \in \mathbb{N}$ (et en rappelant que $s \geqslant 1$) :

$$\begin{aligned}\left|\sum_{n=N+1}^{+\infty} \frac{\chi(n)}{n^s}\right| &\leqslant \left|\frac{V_N(s)}{(N+1)^s}\right| + \sum_{n=N+1}^{+\infty} |V_n(s)| \cdot \left|\frac{1}{n^s} - \frac{1}{(n+1)^s}\right| \\ &\leqslant \frac{1}{N+1} + \sum_{n=N+1}^{+\infty} \left(\frac{1}{n^s} - \frac{1}{(n+1)^s}\right) \\ &= \frac{1}{N+1} + \frac{1}{(N+1)^s} \\ &\leqslant \frac{2}{N+1},\end{aligned}$$

et donc, en reprenant la majoration de $|\mathrm{L}_\chi(s) - \mathrm{L}_\chi(1)|$ plus haut, pour tout $N \in \mathbb{N}$ on a :

$$|\mathrm{L}_\chi(s) - \mathrm{L}_\chi(1)| \leqslant \frac{2}{N+1} + \left|\sum_{n=1}^{N} \chi(n)\left(\frac{1}{n^s} - \frac{1}{n}\right)\right| ;$$

le plus dur a été fait, le reste se reproduit à l'identique et permet de démontrer que : $\lim\limits_{s\to 1^+} \mathrm{L}_\chi(s) = \mathrm{L}_\chi(1)$; comme $\mathrm{L}_\chi(1) > 0$, ainsi qu'on l'a brièvement justifié, ce calcul de limite démontre qu'il existe $\eta > 0$ tel que pour tout $s \in [1,1+\eta]$, on ait : $\mathrm{L}_\chi(s) > 0$, ce qui justifie ce qu'on a annoncé plus haut sur la stricte positivité de cette fonction ;

— question 17 : la conclusion reste exactement la même.

Ainsi ce raisonnement permet de démontrer qu'il existe une infinité de nombres premiers congrus à 1 ou -1 modulo 3 ou 6. Nous laissons le lecteur en exercice comprendre ce qui coince, par exemple, si l'on veut démontrer qu'il existe une infinité de nombres premiers congrus à ± 1 ou ± 2 modulo 5. Tout n'est pas perdu cependant : la remarque culturelle à la fin du sujet donne les grandes lignes pour adapter la démonstration, et le lecteur (très) motivé s'y attellera ; cette année ou dans sa future vie de mathématicien.

Remarque. Pour certaines congruences de ce problème, la démonstration d'Euclide s'adapte. Par exemple, essayez de montrer qu'il existe une infinité de nombres premiers congrus à -1 modulo 4 en raisonnant par l'absurde, et en raisonnant sur les diviseurs premiers de l'entier $N = 4\prod\limits_{i=1}^{r} p_i - 1$ (où $p_1, \ldots, p_r$ sont des nombres premiers congrus à -1 modulo 4).

Questions à se poser, réflexes à acquérir.

— Vérifier ce que je laisse à faire.

— Pour la minoration de $\mathrm{L}_\chi(1)$, pourquoi faire une comparaison série-intégrale est une bonne idée ? Pourquoi, dans le cas $b = 6$, je parle d'utiliser la formule de Taylor avec reste intégral ? Pouvait-on anticiper que ce serait pertinent ? À éventuellement comparer avec mon discours dans *L'Art de la majoration* à ce sujet.

— Calcul de $\ln(\zeta(s)) + \ln(\mathrm{L}_\chi(s))$: pourquoi faut-il ajouter $\frac{1}{2^s}$ dans le cas $b = 6$?

— Pourquoi le fait de s'être ramené à la somme $\sum\limits_{n=N+1}^{+\infty} V_n\left(\frac{1}{n^s} - \frac{1}{(n+1)^s}\right)$ rend plus simple la majoration de $\sum\limits_{n=N+1}^{+\infty} \frac{\chi(n)}{n^s}$, alors que c'est *a priori* une expression nettement plus barbare ? Quels sont les mérites de cette nouvelle somme ? Plusieurs éléments de réponses sont possibles, et il faudra se poser la question À CHAQUE FOIS QUE CE RAISONNEMENT APPARAÎTRA CETTE ANNÉE. C'est nécessaire pour assimiler cette puissante mais subtile technique.

— Si l'on avait plutôt posé : $V_N = \sum\limits_{n=1}^{N} \frac{1}{n^s}$, qu'aurait donné la transformation que nous avons opérée ? Aurait-elle permis d'aboutir ? Cette question est éventuellement essentielle à se poser si vous n'avez pas su répondre à la précédente.

— Si vous avez lu mon document *Méthodes* sur les séries, à la section *Sommes indexées par une classe de congruence*, demandez-vous pourquoi on n'a pas fait apparaître des sommes indexées par les nombres premiers congrus à a mod b (où b est ce que vous voulez) en vous inspirant de la « formule d'orthogonalité » dont je parle dans ce document. Où cela aurait-il coincé, de sorte qu'on soit obligé d'introduire χ ?

— Concernant la remarque finale : quelles congruences marchent ? Pourquoi celles-ci et pas les autres ?