Corrigé du devoir n° 1 (méthode de Cardan)

1 Substituons $X - A$ à x dans l'équation $(\star)$ $\quad ax^3 + bx^2 + cx + d = 0$ (avec $a \neq 0$); nous obtenons $a(X-A)^3 + b(X-A)^2 + c(X-A) + d = 0$, et donc, en développant,

$$aX^3 + (b - 3aA)X^2 + (c - 2bA + 3aA^2)X + d - cA + bA^2 - aA^3 = 0.$$

Choisissant $A = b/3a$, on voit que le terme en X^2 s'annule, et donc que l'équation $(\star)$ équivaut au système $X = x + A = x + b/3a$ et $(\star\star)$ $\quad X^3 = pX + q$, où p et q sont donnés par

$$p = \frac{-3aA^2 + 2bA - c}{a} = \frac{b^2 - 3ac}{3a^2}$$
$$q = \frac{aA^3 - bA^2 + cA - d}{a} = \frac{9abc - 27a^2d - 2b^3}{27a^3}$$

On en déduit (après substitution) que l'équation $8x^3 - 24x^2 + 40x - 33 = 0$ équivaut au système $X = x - 1$ et $X^3 = -2X + 9/8$.

2 Introduisons une nouvelle inconnue $u \neq 0$ telle que $X = u + p/3u$ (ce qui revient à introduire u et v, avec $u + v = X$ et $3uv = p$). L'équation $(\star\star)$ devient alors $u^3 + pu + p^2/3u + p^3/27u^3 = pu + p^2/3u + q$, qui équivaut ($u$ n'étant pas nul) à $u^6 - qu^3 + p^3/27 = 0$. Cette dernière équation se ramène à une équation du second degré : $(\star\star\star)$ $U^2 - qU + p^3/27 = 0$, par le changement d'inconnue $U = u^3$ (et on reviendra aisément à u, puisqu'on sait que tout réel admet une racine cubique unique, notée $u = \sqrt[3]{U} = U^{1/3}$); $(\star\star\star)$ admet des solutions réelles si, et seulement si, $4p^3 - 27q^2 \leq 0$.

3 Pour que $u = 0$ soit solution de l'équation précédente, il faut (et il suffit) que $p = 0$. Dans ce cas, le changement d'inconnue de **2** ne saurait évidemment convenir, mais en revanche, l'équation $(\star\star)$, $X^3 = q$, possède la solution évidente (et unique) $X = q^{1/3}$; ce cas ne présente donc pas d'inconvénient pratique.

4 Si $U_1 \neq 0$ est solution de $(\star\star\star)$ (et donc si $4p^3 - 27q^2 \leq 0$), c'est que $u_1 = U_1^{1/3}$ est solution de $u^3 + pu + p^2/3u + p^3/27u^3 = pu + p^2/3u + q$, et donc que $X_1 = u_1 + p/3u_1$ est solution de $(\star\star)$. Ainsi, à chaque solution (non nulle) de $(\star\star\star)$ correspond une solution de $(\star\star)$; comme le produit des deux solutions U_1 et U_2 de $(\star\star\star)$ vaut $p^3/27$, c'est que $u_1u_2 = p/3$; en définitive, on aura $X_1 = u_1 + p/3u_1 = p/3u_2 + u_2 = X_2$, et cettte méthode ne peut donc donner (au plus) qu'une solution de $(\star\star)$ (et on voit qu'en aucun cas, il ne peut apparaître de solutions parasites). Réciproquement, si X_0 est solution de $(\star\star)$, et qu'il existe un $u_0 \neq 0$ tel que $X_0 = u_0 + p/3u_0$, $U_0 = u_0^3$ sera solution de $(\star\star\star)$, mais cela ne se produit que si l'équation $u^2 - uX_0 + p/3$ admet des solutions (réelles), c'est-à-dire si $X_0^2 \geq 4p/3$; on verra en **6** qu'on peut ne pas être dans ce cas, et que donc la méthode peut ne pas trouver toutes les solutions, ou même n'en trouver aucune, alors qu'il en existe.

5 Appliquons la méthode précédente à l'équation $X^3 = -2X + 9/8$. On est donc amené à résoudre $u^6 - 9u^3/8 - 8/27$, puis l'équation du second degré en $U = u^3$: $(\star\star\star)$ $U^2 - 9U/8 - 8/27 = 0$, d'où la solution

$$U_1 = \frac{9}{16} + \frac{11\sqrt{105}}{144} \text{ et } U_2 = \frac{9}{16} - \frac{11\sqrt{105}}{144},$$

puis $X = \sqrt[3]{U_1} - 2/3\sqrt[3]{U_1}$, et la solution analogue pour U_2 (la remarque faite en **4** montrant que ces deux solutions sont en fait égales). Or, bien que la formule finale soit si effrayante qu'on ne l'a pas recopiée, la calculatrice montre que ce nombre vaut $0,4999999...$ (à 10^{-12} près); on pourrait donc penser que U_1, par exemple, est un cube parfait. Ce n'est pourtant pas le cas; ainsi, à ce stade du calcul, il est seulement extrêmement vraisemblable que $1/2$ soit solution de $(\star\star)$, mais il est aisé de contrôler que c'est bien le cas, puisque $(1/2)^3 = 1/8 = -2(1/2) + 9/8$. On obtient alors facilement (par identification) la factorisation $X^3 + 2X - 9/8 = (X-1/2)(X^2+X/2+9/4)$, et le trinôme $X^2+X/2+9/4$ n'ayant pas de racines réelles (puisque son discriminant vaut $-35/4$), on voit finalement que l'équation $(\star\star)$ a pour seule solution (réelle) $X = 1/2$; revenant à $x = X+1$ (comme on l'a vu en **1**), l'équation $(\star)$ $8x^3 - 24x^2 + 40x - 33 = 0$ a donc pour solution $\mathcal{S} = \{3/2\}$.

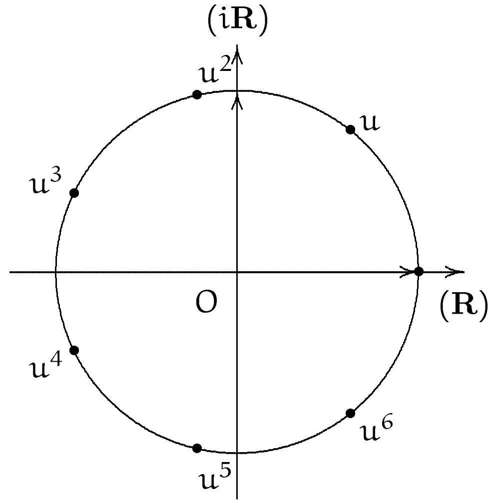
6 Déterminons les variations de $f\colon x \mapsto f(x) = x^3 - px - q$; on a $f'(x) = 3x^2 - p$. Si p est négatif, f est donc (strictement) croissante, et comme elle varie de $-\infty$ à $+\infty$, elle s'annule pour une valeur unique x_0, qui sera donc l'unique solution de $(\star\star)$. Si p est positif, f est croissante sur l'intervalle $]-\infty, -\sqrt{p/3}\,]$ et sur l'intervalle $[\sqrt{p/3}, +\infty[$; et décroissante sur $[-\sqrt{p/3}, \sqrt{p/3}\,]$. Elle a donc un maximum (relatif) en $-\sqrt{p/3}$, qui vaut $f(-\sqrt{p/3}) = (2p/3)\sqrt{p/3} - q$; et un minimum relatif (en $\sqrt{p/3}$), qui vaut $f(\sqrt{p/3}) = -(2p/3)\sqrt{p/3} - q$; ce qui correspond au «tableau de variation»

$$\begin{array}{c|ccccccc|}
x & -\infty & & -\sqrt{\frac{p}{3}} & & \sqrt{\frac{p}{3}} & & +\infty \\ \hline
f'(x) & & + & 0 & - & 0 & + & \\ \hline
 & & & \frac{2p}{3}\sqrt{\frac{p}{3}} - q & & & & +\infty \\
f(x) & & \nearrow & & \searrow & & \nearrow & \\
 & -\infty & & & & -\frac{2p}{3}\sqrt{\frac{p}{3}} - q & &
\end{array}$$

On voit sur ce tableau (en utilisant le théorème des valeurs intermédiaires) que $(\star\star)$ (et par conséquent aussi $(\star)$) aura toujours au moins une solution, et qu'il y en aura trois si et seulement si les deux valeurs qu'on vient de trouver sont de signes opposés, c'est-à-dire si leur produit $f(-\sqrt{p/3})f(\sqrt{p/3}) = q^2 - 4p^3/27$ est négatif; or ce nombre est de toute façon positif si p est négatif. On aboutit donc, en regroupant les deux cas possibles sur le signe de p, à la «condition de Cardan» : il y a trois «solutions» si et seulement si $4p^3 - 27q^2 > 0$ (et deux solutions si $4p^3 - 27q^2 = 0$). On peut remarquer que c'est précisément dans ce cas que l'équation $(\star\star\star)$ n'a pas de solutions (réelles), et donc que la méthode vue en **2** donne bien une solution (unique) quand il n'y en a qu'une, et échoue quand il y en a trois !

7 Et c'est bien parce qu'il est clair que $(\star\star)$ a toujours des solutions que les mathématiciens italiens du $16^{\text{ème}}$ siècle s'enhardirent à introduire des valeurs «imaginaires» de U, en espérant des simplifications analogues à celles de **4**, qui leur donneraient le moyen de «deviner» les solutions réelles dont ils connaissaient *a priori* l'existence. Examinons par exemple le cas $p = 3$, $q = 0$. On a l'équation $X^3 = 3X$, de solution $\mathcal{S} = \{0; \sqrt{3}; -\sqrt{3}\}$. L'équation $(\star\star\star)$ devient $U^2+1 = 0$, et donc $U = i$ (en utilisant une notation moderne; les textes italiens n'hésitent pas à écrire $U = \sqrt{-1}$, et même $u = \sqrt[3]{\sqrt{-1}}$!). On voit donc que pour pouvoir utiliser cette méthode, il faut «inventer» une «racine cubique» à i. On constate aisément que $-i$ convient ($i^3 = -i$, donc $(-i)^3 = i$), mais on n'obtient ainsi que la solution $X = 0$; pour obtenir les autres, on peut par exemple remarquer que $u^3 - i = (u+i)(u^2 - iu - 1)$, résoudre l'équation du second degré par la méthode habituelle, obtenant $u = (\sqrt{3} \pm i)/2$, et contrôler (par calcul direct) qu'on a bien $u^3 = i$ (et donc que i possède (au moins) trois «racines cubiques», à savoir $-i$, u_1 et u_2); il n reste plus alors qu'à calculer $u + 1/u$ (ce qu'on peut faire en remarquant qu'on doit avoir $(\sqrt{3}+i)(\sqrt{3}-i) = 2$), obtenant finalement $X = \sqrt{3}$; c'est ce genre de manipulations, et la confiance progressive dans les résultats ainsi obtenus (en dépit du caractère «magique» des calculs) qui ont amené à la construction rigoureuse des complexes au siècle suivant.

Corrigé du devoir n° 2 (sommes de Gauss)



1 Posons $z = [\rho; \theta]$; on sait que $z^7 = 1$ équivaut à $\rho^7 = 1$ et $7\theta \equiv 0 \pmod{2\pi}$; c'est-à-dire que les solutions de l'équation $z^7 = 1$ sont les sept nombres $e^{2ik\pi/7}$ avec $k \in \{0; 1; 2; 3; 4; 5; 6\}$. On constate que parmi eux, seul celui dont l'argument est compris entre 0 et $\pi/2$, c'est-à-dire $u = e^{2i\pi/7} = \cos 2\pi/7 + i \sin 2\pi/7$, est de parties réelles et imaginaires strictement positives; et qu'on peut donc exprimer la solution en terme de puissances de u sous la forme $\mathcal{S} = \{1; u; u^2; u^3; u^4; u^5; u^6\}$. Les images des solutions forment un heptagone régulier, inscrit dans le cercle unité.

2 On a, pour tout n, $u^{7n} = 1$. Remarquant que $P(z) = z^7 - 1 = (z-1)(z^6 + z^5 + z^4 + z^3 + z^2 + z + 1) = (z-1)Q(z)$ (la formule des suites géométriques) et que $u^k \neq 1$ pour $1 \leq k \leq 6$, on voit que les u^k vérifient $P(u^k) = 0 = (u^k - 1)Q(u^k)$, et donc sont les racines du polyn «me $Q(z) = z^6 + z^5 + z^4 + z^3 + z^2 + z + 1$, et finalement que l'on a la factorisation

$$Q(z) = z^6+z^5+z^4+z^3+z^2+z+1 = (z-u)(z-u^2)(z-u^3)(z-u^4)(z-u^5)(z-u^6).$$

3 On a posé $a = u + u^2 + u^3$. On sait que la somme des racines $n^{\text{èmes}}$ de l'unité est nulle : $1 + u + u^2 + u^3 + u^4 + u^5 + u^6 = (u^7 - 1)/(u - 1) = 0$. Comme on voit aisément que $u^4 = \overline{u^3}$ (puisque $e^{8i\pi/7} = e^{-6i\pi/7}$), que $u^5 = \overline{u^2}$ et que $u^6 = \overline{u}$, on en déduit que

$$1+a+\overline{a} = 1+(u+u^2+u^3)+\overline{(u+u^2+u^3)} = 1+u+u^2+u^3+u^6+u^5+u^4 = 0,$$

et donc que $1 + 2\,\mathrm{Re}(a) = 0$, c'est-à-dire que $\mathrm{Re}(a) = -1/2$. Géométriquement, la somme des 7 vecteurs images des solutions étant nulle, et la figure étant symétrique par rapport à Ox, on voit que ce résultat s'interprète naturellement par projection des vecteurs sur l'axe des réels. Écrivant u, u^2 et u^3 sous forme trigonométrique, on obtient finalement $\mathrm{Re}(u + u^2 + u^3) = \cos 2\pi/7 + \cos 4\pi/7 + \cos 6\pi/7$, et donc

$$\cos 2\pi/7 + \cos 4\pi/7 + \cos 6\pi/7 = -1/2.$$

Comme on sait que $\cos 2\alpha = 2\cos^2\alpha - 1$ et que $\cos 3\alpha = 4\cos^3\alpha - 3\cos\alpha$, on peut, en posant $\alpha = 2\pi/7$ et $X = \cos 2\pi/7$, en déduire que $X + (2X^2 - 1) + (4X^3 - 3X) = -1/2$, et donc que X est racine du polyn» me $P = 4X^3 + 2X^2 - 2X - 1/2$. Une analyse soignée du raisonnement précédent montrerait d'ailleurs que les deux autres racines de P sont $\cos 4\pi/7$ et $\cos 6\pi/7$.

4 Comme on a $u^7 = 1$, on voit que $u^{7n+k} = u^k$, et donc qu'on peut écrire $s = 1 + u + u^4 + u^2 + u^2 + u^4 + u = 1 + 2(u + u^2 + u^4)$. On a par conséquent

$$s = 1 + 2(\cos 2\pi/7 + \cos 4\pi/7 + \cos 8\pi/7) + 2i(\sin 2\pi/7 + \sin 4\pi/7 + \sin 8\pi/7);$$

la calculatrice donne $\mathrm{Re}(s) = 0$ (à 10^{-9} près) et $\mathrm{Im}(s) = 2,6458751311$; s semble donc imaginaire pur.

5 $s-2a-1=2(u^4-u^3)$. Comme on sait que $u^4=1/u^3$ (puisque $u^7=1$), et que $u^4\overline{u^4}=|u^4|^2=1$, on en déduit que $u^4=\overline{u^3}$, et donc que $b=s-2a-1=2(\overline{u^3}-u^3)$ est imaginaire pur (puisque $\overline{b}=-b$). Comme on a vu que $\mathrm{Re}(a)=-1/2$, on en déduit que s est également imaginaire pur.

6 Développons $s^2=\left(1+2(u+u^2+u^4)\right)^2$. On obtient

$$s^2=1+4u+8u^2+8u^3+8u^4+8u^5+8u^6+4u^8;$$

et comme $u^8=u$, on a donc $s^2=1+8(u+u^2+u^3+u^4+u^5+u^6)$; or on a rappelé en **2** que $1+u+u^2+u^3+u^4+u^5+u^6=0$; on en déduit donc que $s^2=-7$, et compte tenu de la valeur numérique (positive) trouvée en **3**, que $s=i\sqrt{7}$.

7 La généralisation cherchée semble porter sur l'étude des «sommes de Gauss» définies par $s(n)=1+u+u^4+\cdots+u^{k^2}+\cdots+u^{(n-1)^2}=\sum\limits_{k=0}^{n-1}u^{k^2}$, où u est une racine $n^{\text{ème}}$ de l'unité «bien choisie» (on prendra ici $u=e^{2i\pi/n}$). On vient de voir que $s(7)=i\sqrt{7}$; déterminons $s(n)$ pour $n\leq 6$. On obtient

$$\begin{array}{llll}
s(2)= & 1+(-1) & & =0\\
s(3)= & 1+j+j^4 & =1+2j & =i\sqrt{3}\\
s(4)= & 1+i+i^4+i^9 & =2(1+i) & =(1+i)\sqrt{4}\\
s(5)= & 1+a+a^4+a^9+a^{16} & =1+2(a+\overline{a}) & =\sqrt{5}\\
s(6)= & 1+b+b^4+b^9+b^{16}+b^{25} & =1+2(b+b^4)+(-1) & =0
\end{array}$$

(où $a=e^{2i\pi/5}=\cos 2\pi/5+i\,\mathrm{Im}(a)=(\sqrt{5}-1)/4+i\sin 2\pi/5$, et $b=e^{i\pi/3}=-j^2=1/2+i\sqrt{3}/2$); résultats qui laissent entrevoir une certaine régularité, au moins pour n impair, et permettent de conjecturer que $s(9)=\sqrt{9}=3$ (ou peut-être $3i$).

De fait, il est aisé de vérifier numériquement que

$$s(9)=1+u+u^4+u^9+u^{16}+u^{25}+u^{36}+u^{49}+u^{64}=3+2u+2u^4+2u^7=3$$

(en prenant par exemple des valeurs numériques des cosinus), mais ce genre de vérification numérique ne prouve pas vraiment le résultat ; en fait, une preuve très simple consiste à remarquer que $u^3=j$, et que $s(9)$ est donc égal à $3+2u(1+j+j^2)$, et donc bien à 3. Il ne semble pas si facile de déterminer $s(11)$, mais la calculatrice donne (à 10^{-12} près) $s(11)\simeq i\sqrt{11}$. La valeur de $s(n)$ dans le cas général est beaucoup plus délicate à obtenir ; un programme de calcul tel que celui-ci (établi pour `Casio 7500`, pour ne pas trop utiliser les raccourcis que procurent des machines plus coûteuses) :

```
"N"?→N: 0→R: 0→S: 0→K: Lbl 1: R+cos(2πK²/N) → R:
S+sin(2πK²/N) → S: K+1→K: K<N⇒Goto 1: R◢ S◢
```

permet d'obtenir assez vite une table de valeurs de $s(n)$ pour $n\leq 100$, d'où on peut conjecturer la forme générale que l'on va donner plus loin ; il n'est pas trop difficile (pour un bon mathématicien, ayant déjà de sérieuses connaissances en théorie des nombres !) de montrer des simplifications analogues à celles déjà vues dans le calcul de s^2, mais Gauss lui-même a avoué avoir passé plusieurs années (!) sur le problème du signe de s. On obtient finalement $s(n)=0$ si $n=4k+2$, $s(4k)=(1+i)\sqrt{4k}$, $s(4k+1)=\sqrt{4k+1}$ et $s(4k+3)=i\sqrt{4k+3}$; comme on vient de le dire, donner même une idée de la démonstration dépasserait de très loin nos possibilités (à titre de curiosité, une démonstration n'utilisant que les connaissancs de Spé (MP*) constituait le sujet du problème d'algèbre (1ère épreuve, 6 heures) de l'ENS Ulm-Lyon de 2001).

Corrigé du devoir n° 4 (l'inéquation $a^x \geq x$)

1 La fonction f_a a pour dérivée $f'_a(x) = a^x \ln a - 1$, qui est toujours négative si $a \leq 1$, et qui s'annule sinon pour $a^x = 1/\ln a$, c'est-à-dire pour $x \ln a = \ln(1/\ln a)$, soit $x = x_a = (-\ln\ln a)/\ln a$; a^x étant croissante pour $a > 1$, on en déduit qu'alors $f'_a(x)$ est négative jusqu'à $x = x_a$, positive ensuite ; et donc que f_a atteint un minimum absolu en x_a, valant $f_a(x_a) = a^{x_a} - x_a = e^{x_a \ln a} - x_a = 1/\ln a - x_a$. Les limites sont obtenues en utilisant les relations de comparaison, ainsi par exemple $\lim\limits_{+\infty} f_a = +\infty$ si $a > 1$, car «l'exponentielle l'emporte» (c'est-à-dire que $x \underset{+\infty}{\ll} e^x$, ou encore que $\lim\limits_{x \to +\infty} x/e^x = 0$). On peut finalement établir les deux tableaux suivants :

$\boxed{\text{Cas } a < 1}$

x	$-\infty$		$+\infty$
$f'_a(x)$		$-$	
$f_a(x)$	$-\infty$	$\searrow$	$+\infty$

$\boxed{\text{Cas } a > 1}$

x	$-\infty$		$x_a = -\frac{\ln\ln a}{\ln a}$		$+\infty$
$f'_a(x)$		$-$	0	$+$	
$f_a(x)$	$+\infty$	$\searrow$	$f_a(x_a) = \frac{1+\ln\ln a}{\ln a}$	$\nearrow$	$+\infty$

Le cas $a = 1$ ($f_1(x) = 1 - x$) manque d'intérêt : le graphe est une droite, et f_1 s'annule pour $x = 1$.

2 D'après l'étude des limites de f_a (quand $a < 1$), on voit que f_a passe de valeurs positives à des valeurs négatives ; comme elle est continue, le théorème des valeurs intermédiaires nous dit qu'elle s'annule. Si on avait $f_a(x_1) = f_a(x_2) = 0$, on ne pourrait avoir $x_1 < x_2$, puisque f_a est strictement décroissante ; on voit donc que la solution de $f_a(x) = 0$ existe et est unique (on la notera désormais $s(a)$). De plus, $f_a(1) = a - 1 < 0$, et $f_a(a) = a^a - a$ est > 0, puisque a^x est décroissante et que donc $a < 1 \Rightarrow a^a > a^1$; on en déduit que f_a s'annule entre a et 1, et donc que $a < s(a) < 1$. Si alors $x < s(a)$, f_a étant strictement décroissante, on aura $f_a(x) > f_a(s(a)) = 0$, et donc $a^x > x$; on voit de même que pour $x > s(a)$, $a^x < x$, et en définitive, la solution de l'inéquation $a^x \geq x$ est

$$\mathcal{S} =]-\infty, s(a)]$$

3 Pour $a < b < 1$, les fonctions f_a et f_b sont décroissantes, et de plus, pour $x > 0$, on a $a^x < b^x$, et par conséquent $f_a(x) < f_b(x)$. On a donc $f_a(s(a)) = 0 < f_b(s(a))$ (car $s(a) > 0$), et donc $f_b(s(a)) > 0 = f_b(s(b))$. Comme f_b est décroissante, on en déduit que $s(a) < s(b)$ (pour obtenir ce résultat, il serait plus naturel en fait d'utiliser la fonction g^{-1} de la question **7**).

4 On a vu en **1** que le minimum de f_a (pour $a > 1$) est atteint en $x_a = -\ln\ln a/\ln a$, et vaut $(1 + \ln\ln a)/\ln a$; on en déduit que si a est tel que ce minimum est nul, l'équation $f_a(x) = 0$ possédera une solution unique (égale à x_a). Il faut donc déterminer a_0 tel que $1 + \ln\ln a_0 = 0$; on voit que cela équivaut à $a_0 = e^{1/e}$.

Une analyse des variations de f_a montre que si $a \geq a_0$, le minimum de f_a sera positif, et l'inéquation $a^x \geq x$ sera vérifiée pour tout x ($\mathcal{S} = \mathbf{R}$) ; par contre, si $1 < a < a_0$, on voit que f_a sera négative sur un intervalle $[\alpha_a, \beta_a]$ (avec $\alpha_a < x_a < \beta_a$) ; on en déduit qu'alors la solution sera

$$\mathcal{S} =]-\infty, \alpha_a] \cup [\beta_a, +\infty[$$

5 Si $1 < a < a_0$, il y a à présent deux solutions (α_a et β_a) à l'équation $a^x = x$; on voit que l'on peut donc définir deux fonctions ($s_1(a) = \alpha_a$ et $s_2(a) = \beta_a$) par des conditions telles que (par exemple) $s_1(a) < x_a$ et $f_a(s_1(a)) = 0$. Montrons que s_1 est croissante (on verrait de même que s_2 est décroissante) : on a (comme en **3**) $a < b \Rightarrow f_a(x) < f_b(x)$, et sur l'intervalle $[-\infty, x_a\ ]$, la fonction f_a est décroissante; comme $x_b > x_a$, il en est de même de f_b. En supposant que $s_1(b) \leq s_1(a) < x_a$; on aurait alors $f_b(s_1(b)) = 0 > f_a(s_1(b))$ et donc $0 > f_a(s_1(b)) > f_a(s_1(a)) = 0$, ce qui est absurde; on voit donc qu'on doit avoir $s_1(b) > s_1(a)$.

6 Par définition de a_0, le minimum (nul) de f_{a_0} est atteint en $x_{a_0} = \dfrac{-\ln\ln a_0}{\ln a_0}$, on voit que l'équation $a_0^x = x$ a pour solution unique $x_0 = (-\ln(1/e))/(1/e)) = e$; il est d'ailleurs évident en effet que $(e^{1/e})^e = a_0^e = e$!

7 Posant $g(x) = \ln x/x$, on a (pour x positif) $a^x \geq x \iff x\ln a \geq \ln x \iff \ln a \geq g(x)$ (et on sait que $a^x > x$ pour tout $x \leq 0$). On voit donc qu'en posant $b = \ln a$, l'étude proposée se ramène à celles d'équations de la forme $g(x) \leq b$; g ayant pour tableau de variation

x	0	$x_0 = e$	$+\infty$
$g'(x) = (1 - \ln x)/x^2$	$+$	0	$-$
$g(x)$	$-\infty$ ↗	$g(e)=1/e$	↘ 0

A l'aide de ce tableau, on voit par exemple que l'équation $a^x = x$ admet une solution pour $b \leq 0$, c'est-à-dire pour $a \leq 1$; ou encore une solution unique pour $b_0 = 1/e$, c'est-à-dire pour $a_0 = e^{1/e}$.

Définissons une fonction «réciproque» de g, notée g^{-1}, et allant de $]-\infty, 0]$ vers $]0, 1]$, par la «formule» $g^{-1}(y) = x \iff g(x) = y$ (le théorème des valeurs intermédiaires en garantit l'existence). On aura donc $a^x = x \iff \ln a = g(x) \iff x = g^{-1}(\ln a)$; ainsi,

$$s(a) = g^{-1}(\ln a)$$

Si on définit de même une fonction $h(x)$ (allant de $]0, e^{1/e}]$ vers $]1, e]$) par $h(y) = x \iff g(x) = y$ et $x \leq e$, on aura

$$s_1(a) = h(\ln a)$$

8 Avec les notations précédentes, $a^x = x^a \iff x\ln a = a\ln x \iff g(x) = g(a)$; comme g est injective sur $]0, 1]$, on voit que si $a \in]0, 1]$, la seule solution sera $x = a$. Si $a \in]1, e[$, une seconde solution apparait, car la fonction g prend deux fois chaque valeur; en notant par exemple k la fonction allant de $]0, e^{1/e}]$ vers $[e, +\infty[$ telle que $k(y) = x \iff g(x) = y$ et $x \geq e$, on aura comme seconde solution $x = k(g(a))$. En particulier, comme le seul entier m de l'intervalle $]1, e]$ est 2, on voit que l'équation $m^n = n^m$ ne peut avoir qu'une solution (telle que $m < n$) au plus; on constate sur le graphe de g que $n = 4$ pourrait convenir, et en effet $2^4 = 4^2 = 16$.

Corrigé du devoir n° 8 (la règle de l'Hospital)

1 Si f et g sont dérivables en 0, c'est par définition que $\lim\limits_{x\to 0}(f(x)-f(0))/(x-0)=\lim\limits_{x\to 0} f(x)/x=f'(0)$ et que $\lim\limits_{x\to 0} g(x)/x=g'(0)$; on a donc $\lim\limits_{x\to 0}(f(x)/g(x))=\lim\limits_{x\to 0}(f(x)/x)\times(x/g(x)=f'(0)/g'(0)$. Posant $X=x+a$, on en déduit que si $f(a)=g(a)=0$, que f et g sont dérivables en a et que $g'(a)\neq 0$, on aura $\lim\limits_{X\to a} f(X)/g(X)=f'(a)/g'(a)$. Ces calculs supposent la nullité de f et de g (et d'ailleurs, le cas général aboutirait à $\lim\limits_{X\to a}(f(X)-f(a))/(g(X)-g(a))=f'(a)/g'(a)$), mais ce n'est pas un problème en pratique, car si $g(a)\neq 0$, la limite de f/g est $f(a)/g(a)$, et si $g(a)=0$ et $f(a)\neq 0$, cette limite est infinie, et ne demande que l'étude du signe de g.

2 La «règle» formulée par l'énoncé (et dont on verra la version correcte en **5**) correspond à une généralisation de la règle précédente si f et g sont de classe $\mathcal{C}^1$ (en a), car alors si $f'(a)/g'(a)$ existe, il est égal à la limite donnée; mais on peut très bien avoir des applications de cette règle au cas où $g'(a)=0$; ou encore où $f'(a)$ et $g'(a)$ n'existent pas, comme on le verra plus loin.

3 La formule des accroissements finis ne permet pas d'obtenir directement la règle, car le c dont elle prouve l'existence dépend de la fonction étudiée : tout ce qu'on sait, en fait, c'est qu'il existe c_1 et c_2 (compris entre 0 et x) tels que $f(x)/g(x)=f'(c_1)/g'(c_2)$, et cela ne suffit pas pour conclure (car on ne sait pas passer à la limite dans ce cas).

4 Soit h une fonction dérivable (sur $]a,b[$) telle que $h(a)=h(b)=0$; appliquant le lemme de Rolle à la fonction $k(x)=h(x)-h(a)$, on voit que $k'(x)=h'(x)$ s'annule sur $]a,b[$. Si h est de la forme $Af+Bg$, avec $A\neq 0$, on voit qu'on aura donc (en un point c de $]a,b[$) $Af'(c)+Bg'(c)=0$, et donc (si $g'(c)\neq 0$) $f'(c)/g'(c)=-B/A$ (et si $g'(c)=0$, $f'(c)=0$ également). Prenant $A=g(b)-g(a)$, et $B=f(a)-f(b)$, on vérifie aisément que $h(a)=Af(a)+Bg(a)=g(b)f(a)-g(a)f(b)=Af(b)+Bg(b)=h(b)$; le résultat précédent s'applique et on en déduit qu'il existe un c (avec $a<c<b$) tel que

$$(\star)\qquad \frac{f'(c)}{g'(c)}=\frac{f(b)-f(a)}{g(b)-g(a)}$$

ou tel que $f'(c)=g'(c)=0$.

5 Supposons alors que la limite $\lim\limits_{x\to a} f'(x)/g'(x)$ existe (et vaille L, fini ou infini); la définition de Cauchy nous dit que les inégalités correspondantes (du type $|F(x)-L|<\varepsilon$) sont valables pour tout x assez proche de a, et donc pour tout c compris entre a et un de ces x. Comme la règle de l'Hospital n'a de sens que si la limite cherchée en a un, on voit que $g(x)-g(a)$ doit être non nul dans un voisinage de a, et donc que le calcul précédent s'applique (on éliminerait de même le cas $f'(c)=g'(c)=0$ en remarquant qu'alors la limite de $f'(x)/g'(x)$ ne serait pas définie). Utilisant alors la formule $(\star)$, on en déduit que $(f(x)-f(a))/(g(x)-g(a))$ vérifie la même inégalité, et donc que $\lim\limits_{x\to a}(f(x)-f(a))/(g(x)-g(a))=L$. En prolongeant par continuité f et g (en supposant, bien sûr, que $\lim\limits_{x\to a} f(x)=\lim\limits_{x\to a} g(x)=0$, sinon la question n'a pas d'intérêt, comme on l'a dit en **1**), on aboutit à la formulation exacte suivante :

Règle de l'Hospital *Si* $\lim_{x\to a} f(x) = \lim_{x\to a} g(x) = 0$, *et si* f *et* g *sont drivables au voisinage de* a, *et si* $\lim_{x\to a} f'(x)/g'(x)$ *existe, alors* $\lim_{x\to a} f(x)/g(x) = \lim_{x\to a} f'(x)/g'(x)$

Cette règle n'est pas tout à fait équivalente à celle énoncée en **2** ; et en particulier, la première limite peut exister et pas la seconde : prenant $f(x) = x^2 \sin(1/x)$ et $g(x) = x$ (et $f(0) = 0$), on sait que $\lim_{x\to 0} f(x)/x = 0$, alors que $f'(x)/1$ n'est pas continue en 0.

6 $\lim_{x\to a} f'(x)/g'(x)$ peut être à son tour une forme indéterminée (du type «0/0»), et on peut alors envisager de lui réappliquer la règle. Par récurrence, on voit que celle-ci devient : si f et g sont k fois dérivables au voisinage de a, et si $\lim_{x\to a} f^{(k)}(x)/g^{(k)}(x) = L$ (avec L fini ou infini) et si (c'est cette condition qui est le plus souvent oubliée !) $\lim_{x\to a} f(x) = \lim_{x\to a} g(x) = \lim_{x\to a} f'(x) = \lim_{x\to a} g'(x) = \cdots = \lim_{x\to a} f^{(k-1)}(x) = \lim_{x\to a} g^{(k-1)}(x) = 0$, alors $\lim_{x\to a} f(x)/g(x) = L$. De même, on sait, posant $X = 1/x$, que $\lim_{x\to+\infty} f(x)/g(x) = \lim_{X\to 0^+} f(1/X)/g(1/X)$, or, posant $h(X) = f(1/X)$ et $k(X) = g(1/X)$, on a $h'(X) = (-1/X^2)f'(1/X)$ et $k'(X) = (-1/X^2)g'(1/X)$; appliquant la règle de l'Hospital (généralisée) à h et k, on voit finalement que la règle précédente reste valable même si $a = +\infty$.

7 En appliquant la règle de l'Hospital précédente (en $+\infty$), on obtient (après avoir contrôlé que la limite est bien de la forme «0/0»)

$$\lim_{x\to+\infty} x(\operatorname{Arc\,tg} x - \pi/2) = \lim_{x\to+\infty} (1/(1+x^2))/(-1/x^2) = -1.$$

De même,

$$\lim_{x\to 0} \frac{1-\cos x}{\operatorname{tg} x} = \lim_{x\to 0} \frac{\sin x}{1+\operatorname{tg}^2 x} = 0.$$

La dernière limite cherchée est de la forme «$\infty - \infty$», mais on peut écrire

$$\begin{aligned}\lim_{x\to 0} \frac{1}{\sin^2 x} - \frac{1}{x^2} &= \lim_{x\to 0} \frac{x^2 - \sin^2 x}{x^2 \sin^2 x} \\ &= \lim_{x\to 0} \frac{x - \sin x}{x \sin^2 x} \times \lim_{x\to 0} \frac{x + \sin x}{\sin x} = 2 \lim_{x\to 0} \frac{x - \sin x}{x^2 \sin x};\end{aligned}$$

en appliquant deux fois la règle de l'Hospital à cette dernière limite (comme on l'a vu en **6**) après avoir vérifié qu'on est bien dans un cas où la règle généralisée s'applique (c'est-à-dire qu'on obtient des formes indéterminées du type «0/0»), on trouve

$$\begin{aligned}\lim_{x\to 0} \frac{1}{\sin^2 x} - \frac{1}{x^2} &= 2 \lim_{x\to 0} \frac{\sin x}{(2-x^2)\sin x + 4x\cos x} \\ &= 2 \lim_{x\to 0} \frac{1}{2 - x^2 + 4(x/\sin x)\cos x} = \frac{1}{3}.\end{aligned}$$

8 Montrons le résultat demandé par récurrence : on sait que $\lim_{x\to 0} \frac{e^x - 1}{x} = 1$; supposons que $\lim_{x\to 0} \frac{e^x - \sum_{p=0}^{k} \frac{x^p}{p!}}{x^{k+1}} = \frac{1}{(k+1)!}$ soit vrai pour tout $k \leq n$; montrons que cette formule est alors encore vraie pour $k = n+1$. La dérivée de $e^x - \sum_{p=0}^{k} \frac{x^p}{p!}$ étant $e^x - \sum_{p=0}^{k-1} \frac{x^p}{p!}$ (comme on le vérifie aisément), les limites successives cherchées par la règle de l'Hospital (généralisée) sont toutes de la forme «0/0» (d'après l'hypothèse de récurrence) jusqu'à celle correspondant aux dérivées d'ordre $n+1$: $\lim_{x\to 0} \frac{e^x - 1}{(n+2)!x} = \frac{1}{(n+2)!}$; par récurrence, la formule étant vraie pour $n = 0$, elle est toujours vraie. Pour la limite correspondante à $\ln(1+x)$, il faut remarquer que $\ln(1+x)' = 1/(1+x) = 1 - x + x^2 - \cdots + (-1)^{n-1}x^{n-1} + (-1)^n x^n/(1+x)$ (comme on le voit en appliquant la formule des suites géométriques); posant $P_n(x) = x - x^2/2 + x^3/3 + \cdots + (-1)^{n-1}x^n/n$, on constate (la règle de l'Hospital s'applique, puisque le numérateur tend vers 0) que

$$\lim_{x\to 0} \frac{\ln(x+1) - P_n(x)}{x^{n+1}} = \lim_{x\to 0} \frac{1/(x+1) - (1 - x + x^2 - \cdots)}{(n+1)x^n},$$

et donc, d'après la remarque précédente, que

$$\lim_{x\to 0} \frac{\ln(x+1) - P_n(x)}{x^{n+1}} = \frac{(-1)^n}{n+1}.$$

Corrigé du devoir n°5 (fonction Beta)

1 La fonction à intégrer étant continue sur $]0,1[$, il suffit de vérifier qu'elle est définie en 0 et 1 (ou prolongeable par continuité) ; on doit donc avoir $u-1 \geq 0$ et $v-1 \geq 0$, c'est-à-dire $(u,v) \in [1,+\infty[\times [1,+\infty[$. On a

$$B(v,u) = \int_0^1 t^{v-1}(1-t)^{u-1}\,dt = \int_1^0 (1-T)^{v-1}T^{u-1}(-dT)$$

(à l'aide du changement de variable $T = -t$) et donc (d'après Chasles) $B(v,u) = B(u,v)$. Calculons $B(u,v+1)$: on a

$$\begin{aligned} B(u,v+1) &= \int_0^1 t^{u-1}(1-t)^v\,dt = \int_0^1 t^{u-1}(1-t)^{v-1}(1-t)\,dt \\ &= \int_0^1 t^{u-1}(1-t)^{v-1}\,dt - \int_0^1 t^u(1-t)^{v-1}\,dt = B(u,v) - B(u+1,v) \end{aligned}$$

et donc $B(u,v) = B(u,v+1) + B(u+1,v)$.

2 En intégrant par parties, on a

$$B(u+1,v) = \left[\frac{t^u(1-t)^v}{-v}\right]_0^1 - \int_0^1 \frac{ut^{u-1}(1-t)^v}{-v}\,dt = \frac{u}{v}B(u,v+1)$$

et comme on a déjà obtenu une relation entre cette intégrale et $B(u,v)$, on a $B(u+1,v) = (u/v)(B(u,v) - B(u+1,v))$, donc

$$B(u+1,v) = \frac{u}{u+v}B(u,v)$$

En particulier, si u est entier, on va pouvoir exprimer $B(u,v)$ en fonction de $B(1,v)$, puisqu'on aura $B(u,v) = \dfrac{u-1}{u+v-1} \times \dfrac{u-2}{u+v-2} \times \cdots \times \dfrac{1}{v+1}B(1,v)$, et comme $B(1,v) = 1/v$, on a donc

$$B(n,p) = \frac{(n-1)!}{(p+n-1)(p+n-2)\cdots(p+1)p}$$

Si p est aussi entier, on en déduit la «formule d'Euler»

$$B(n,p) = \frac{(n-1)!(p-1)!}{(n+p-1)!}$$

Calculons directement $B(n,p)$ dans le cas n et p entiers : on a (d'après la formule du binôme)

$$B(n,p) = \int_0^1 \sum_{k=0}^{p-1} (-1)^k C_{p-1}^k t^{n+k-1}\,dt = \sum_{k=0}^{p-1} \frac{(-1)^k C_{p-1}^k}{n+k}$$

et par conséquent on obtient

$$\sum_{k=0}^{p-1} \frac{(-1)^k C_{p-1}^k}{n+k} = \frac{(n-1)!(p-1)!}{(n+p-1)!}$$

3 $B(3/2,3/2) = \displaystyle\int_0^1 \sqrt{t(1-t)}\,dt$; on peut calculer cette intégrale par le changement de variable $T = t - 1/2$ (forme canonique) puis en posant $X = \operatorname{Arc\,sin} 2T$; mais il est plus rapide de remarquer que le graphe de $f(t) = \sqrt{t - t^2}$ est un demi-cercle, et donc que $B(3/2, 3/2) = \pi/8$

Appliquant (illégalement) la «formule d'Euler», on devrait avoir $B(3/2, 3/2) = \dfrac{(1/2)!^2}{2}$, et donc $(1/2)! = \sqrt{\pi}/2$; c'est ce que confirment les calculettes (la TI-92, ou la HP-48S) de manière approchée, et Maple (exactement). Euler a généralisé la fonction factorielle en définissant, par une autre intégrale, ce qu'il appelle la fonction Gamma : il pose

$$\Gamma(x) = \lim_{a\to\infty} \int_0^a t^{x-1} e^{-t}\, dt$$

et il démontre que $\Gamma(n) = (n-1)!$ et que $B(u, v) = \dfrac{\Gamma(u)\Gamma(v)}{\Gamma(u+v)}$; on a montré par la suite que cette généralisation est (en un certain sens) la meilleure possible, d'où le choix de $\Gamma(x+1)$ comme prolongement de $x!$.

4 Posons $t = \sin^2 x$ (et donc $dt = 2t^{1/2}(1-t)^{1/2}dx$), on obtient (par changement de variable)

$$\int_0^{\pi/2} \sin^p x \cos^q x\, dx = \frac{1}{2}\int_0^1 t^{(p-1)/2}(1-t)^{(q-1)/2} dt = \frac{B(\frac{p+1}{2}, \frac{q+1}{2})}{2}$$

l'intégrale est évidemment calculable pour p et q positifs, ce qui semble contredit par le domaine de $B(\frac{p+1}{2}, \frac{q+1}{2})$ (on devrait en effet avoir p et q supérieurs à 1) Mais en réalité, le changement de variable dans cette direction (correspondant donc à $x = \varphi(t) = \operatorname{Arc}\sin\sqrt{t}$) n'a pas été contrôlé (il n'est pas dérivable en 0 et 1) et il n'est même pas évident que le résultat soit correct quand les fonctions sont définies. Cependant, le changement de variable réciproque $x = \varphi^{-1}(t) = \phi(t) = \sin^2 t$ est légitime dans le calcul de la fonction Beta (c'est-à-dire qu'on a bien $\int_0^1 x^p(1-x)^q\, dx = \int_0^{\pi/2} \phi(t)^p(1-\phi(t))^q \phi'(t)\, dt$), ce qui justifie le résultat précédent pour p et q supérieurs à 1. Pour comprendre ce qui se passe pour $p < 1$, il faut en fait passer à la limite, comme on va le faire dans les questions suivantes.

Lorsque p et q sont entiers, ces intégrales se calculent facilement par linéarisation. De plus, si $q = 0$, on retrouve les intégrales de Wallis, et la récurrence classique entre ces intégrales correspond à la formule déjà obtenue en **2**.

5 Si la formule précédente était correcte, on aurait

$$B(1/2, 1/2) = 2\int_0^{\pi/2} \sin^0 x \cos^0 x\, dx = \pi.$$

Montrons que cette valeur est la limite de $\displaystyle\int_x^{1-x} \frac{dt}{\sqrt{t(1-t)}}$ quand x tend vers 0; en effet, sur l'intervalle $[x, 1-x]$ le changement de variable $t = \sin^2 T$ est légal (bijectif, dérivable et de fonction réciproque dérivable), et on obtient

$$\int_x^{1-x} \frac{dt}{\sqrt{t(1-t)}} = \int_{\operatorname{Arc}\sin\sqrt{x}}^{\operatorname{Arc}\sin\sqrt{1-x}} 2dT = 2(\operatorname{Arc}\sin\sqrt{1-x} - \operatorname{Arc}\sin\sqrt{x})$$

dont la limite en 0 est bien π.

6 Soit $F(t)$ une primitive de $t^{u-1}t^{v-1}$ (sur $]0, 1[$), on a donc par définition $B^*(u, v) = \lim_{x\to 0} F(1-x) - F(x)$. Si $B(u, v)$ est définie, c'est que F existe sur $[0, 1]$; F étant une primitive est continue et par conséquent $B^*(u,v) = F(1) - F(0) = B(u,v)$; B^* est donc un prolongement de B. On vient de voir qu'elle est définie en $(1/2, 1/2)$; plus généralement, la fonction $F(1-x)-F(x)$ ayant pour dérivée $-f(1-x)-f(x) < 0$, elle est décroissante et aura une limite en 0 si elle est majorée (borne supérieure) (et sinon cette limite sera infinie). Or on a par exemple $t^p(1-t)^q \leq t^p$ si $q \geq 0$, ce qui montre que $F(1-x) - F(x) \leq (1-x)^{p+1} - x^{p+1}/(p+1)$ (pour $p \neq -1$), et par conséquent que $B^*(p,q)$ est définie pour $p > 0$ et q positif. Inversement, si p est ≤ 0, un encadrement analogue montre que la limite donnant B^* est de même nature que celle de $\lim\limits_{x\to 0} x^p$ qui est infinie. On voit qu'il est donc impossible de prolonger B de cette manière à p (ou q) négatif.

Montrons que B^* vérifie les relations de **1** et **2**. C'est évident pour la relation $B^*(p,q+1) + B^*(p+1,q) = B^*(p,q)$ (le calcul ne dépendait pas des bornes) et la relation de «symétrie» $B^*(p,q) = B^*(q,p)$ se déduit de ce que les bornes x et $1-x$ s'échangent dans le changement de variable $T = -t$. La relation de **2** découle d'un passage à la limite : en effet, il subsiste (après intégration par parties) le terme $\left[\frac{t^p(1-t)^q}{-q}\right]_x^{1-x}$, dont la limite est nulle si p et q sont positifs. Finalement, on obtient bien la relation

$$B^*(u+1,v) = \frac{u}{u+v} B^*(u,v)$$

Comme le terme de gauche est défini pour $u \geq -1$, on peut en déduire un prolongement (si $u \neq 0$) et par exemple on aura $B^*(-1/2, 1/2) = 0$. Mais il semble peu probable que cela corresponde encore à une intégrale, puisqu'on a vu que même en se limitant à $]0,1[$, B^* devrait devenir infinie.

Corrigé du devoir n°6 (puissances de matrices)

1 Les définitions générales se traduisent dans l'anneau $(\mathcal{M}_n(\mathbf{R}),+,\times)$ par : M est unipotente si il existe un entier $k>0$ tel que $M^k = \underbrace{M\times M\times\cdots\times M}_{k \text{ fois}} = I_n$ (la matrice unité de $\mathcal{M}_n(\mathbf{R})$ définie par $I_n = (\delta_{ij})_{1\leq i,j\leq n}$; où δ_{ij} est le symbole de Kronecker) ; M est nilpotente si il existe un entier $k>0$ tel que $M^k = O_n$ (la matrice nulle, c'est-à-dire celle dont tous les termes sont nuls).

2 Si x est unipotent, on a $x^k = x\star(x^{k-1}) = (x^{k-1})\star x = 1_A$ (avec $k\geq 1$) ; la définition des éléments inversibles (a est inversible $\iff \exists b\in A,\ a\star b = b\star a = 1_A$) montre donc que x est inversible et que $x^{-1} = x^{k-1}$. Supposons que x soit nilpotent et inversible, on aura $(x^{-1})^k\star x^k = (x^{-1}\star x)^k = 1_A$ (puisque x et x^{-1} commutent) et comme $x^k = 0_A$, on aurait donc $1_A = 0_A$ ce qui est absurde (sauf si l'anneau n'a qu'un élément !) ; aucun élément nilpotent n'est donc inversible.

3 Calculons par récurrence $(y\star x\star y^{-1})^n$: supposons (hypothèse de récurrence) que l'on ait $(y\star x\star y^{-1})^k = y\star x^k\star y^{-1}$, on aura alors

$$\begin{aligned}(y\star x\star y^{-1})^{k+1} &= (y\star x\star y^{-1})^k\star(y\star x\star y^{-1})\\ &= y\star x^k\star y^{-1}\star y\star x\star y^{-1} = y\star x^{k+1}\star y^{-1}\end{aligned}.$$

La formule étant évidemment vraie pour $k=1$, elle est vraie par récurrence pour tout n. Supposons alors que x soit unipotent, et que $x^k = 1_A$; on aura donc $(y\star x\star y^{-1})^k = y\star 1_A\star y^{-1} = 1_A$, et $y\star x\star y^{-1}$ sera donc également unipotent ; de même, si $x^k = 0_A$, on aura $(y\star x\star y^{-1})^k = y\star 0_A\star y^{-1} = 0_A$.

4 Soit $A = \begin{pmatrix} a & b\\ 0 & d\end{pmatrix}$ une matrice triangulaire supérieure 2×2, on sait (et on le vérifie aisément par récurrence) que $A^n = \begin{pmatrix} a^n & b_n\\ 0 & d^n\end{pmatrix}$ (b_n peut aussi être obtenu par récurrence, mais la formule compliquée correspondante n'est pas nécessaire pour la suite du problème). Si alors A est nilpotente, cela veut dire qu'il existe un k pour lequel $a^k = d^k = b_k = 0$; on a donc $a = d = 0$, et comme réciproquement on a alors $A^2 = O$, on voit qu'on a donc trouvé toutes les matrices nilpotentes de la forme cherchée (le cas général passerait par la théorie de la diagonalisation), que l'on se place dans $\mathbf{R}$ ou dans $\mathbf{C}$. Si on veut $A^k = I$, on voit qu'on doit avoir $a^k = d^k = 1$, et donc (dans $\mathbf{R}$) $a = \pm 1$ et $d = \pm 1$. Prenant au besoin $-A$ (et remarquant que $A^k = I \Rightarrow (-A)^{2k} = I$), on voit qu'on peut supposer $a = 1$. Si alors $d = 1$, on a $A = I + B$, avec $B^2 = O$, et d'après la formule du binôme $(I+B)^n = I + nB \neq I$ si $B\neq O$. I et $-I$ sont donc les seules matrices unipotentes pour lesquelles $a = d$. Si par contre on prend $a = 1$ et $d = -1$, on vérifie aisément que $A^2 = I$ pour toute valeur de b ; les matrices $\pm\begin{pmatrix} 1 & b\\ 0 & -1\end{pmatrix}$ sont donc toutes unipotentes. Ces calculs ne sont plus valables dans $\mathbf{C}$, puisqu'on pourrait alors avoir $a = e^{2i\pi/n}$ par exemple ; il est alors plus facile d'utiliser la théorie de la diagonalisation, et le résultat final est que les matrices unipotentes (de la forme demandée) sont du type $e^{2i\pi/n}I$, d'une part, et du type $\begin{pmatrix} e^{2i\pi/p} & b\\ 0 & e^{2i\pi/q}\end{pmatrix}$, d'autre part.

5 Si $X^2 = I_n$, on a $(X + aI_n)^2 = 2aX + (a^2+1)I_n$; ainsi, si $Y = X + I_n$, on aura $Y^2 = 2Y$. Réciproquement, si $Y^2 = 2Y$, posant $X = Y - I$, on aura $(X+I)^2 = X^2 + 2X + I = 2X + 2I$, donc $X^2 = I$.

6 La formule demandée peut s'obtenir par récurrence, mais il est plus simple de remarquer que A commute avec I, et qu'on peut donc appliquer la formule de Newton; comme $A^k = O$ pour tout $k \geq 3$, elle devient ici $(I+A)^n = I + nA + \dfrac{n(n-1)}{2}A^2$.

7 A et I commutant, on peut appliquer l'identité des suites géométriques : $(I + A + A^2 + \cdots + A^n)(I - A) = I - A^{n+1}$. Prenant $n = k-1$, où k est le plus petit entier tel que $A^k = O$, on voit que $(I + A + \cdots + A^{k-1})(I - A) = I$, ce qui montre que $I - A$ est inversible. La matrice $B = -A$ étant également nilpotente, on aura de même $I - B = I + A$ inversible (et $(I+A)^{-1} = I - A + A^2 - \cdots + (-1)^{k-1}A^{k-1}$). Un calcul analogue aboutissant, si A est unipotente, à $(I + A + \cdots + A^{k-1})(I - A) = O$, on pourrait penser que $I - A$ n'est pas inversible, mais en fait, cela montre seulement qu'alors $(I + A + \cdots + A^{k-1})$ est nulle, ce qui peut se produire : prendre par exemple $A = -I$!

8 Supposons que A et B commutent; on sait qu'on peut alors appliquer la formule du binôme $(A+B)^n = \sum_{k=0}^{n} C_n^k A^k B^{n-k}$. Si on a $A^p = O$ et $B^q = O$, prenons alors $n = p+q$, et décomposons la sommation précédente en $\sum_{k=0}^{p} C_n^k A^k B^{q+(p-k)} + \sum_{k=p+1}^{n} C_n^k A^k B^{n-k}$. Il est clair que toutes les puissances de B sont nulles dans la première somme, et que toutes les puissances de A sont nulles dans la seconde; on aura donc $(A+B)^{p+q} = O$, et $A+B$ sera bien nilpotente. Ce résultat n'est plus valable si A et B ne commutent pas : ainsi, prenant $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ et $B = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$, on a $A^2 = B^2 = O$, mais $C = A + B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ vérifie $C^2 = I$, et donc C^n, qui vaut C ou I (suivant la parité de n), n'est jamais nulle.

9 On peut en fait sans inconvénient poser $\exp(A) = \sum_{k=0}^{n} \dfrac{A^k}{k!}$, en prenant n supérieur au premier k pour lequel $A^k = 0$. Comme on a supposé que A et B commutent (ce qui est de toute façon nécessaire pour que $\exp(A+B)$ ait un sens, comme on l'a vu plus haut), on peut d'une part appliquer la formule du binôme aux $(A+B)^k$, d'autre part effectuer le produit $\exp(A)\exp(B)$ (par distributivité) sans prendre de précautions (d'ailleurs inutiles dans ce cas). On aboutit d'une part à

$$\exp(A)\exp(B) = \left(\sum_{j=0}^{n} \frac{A^j}{j!}\right)\left(\sum_{k=0}^{n} \frac{B^k}{k!}\right) = \sum_{0 \leq j,k \leq n} \frac{A^j B^k}{j!k!}$$

(en prenant n assez grand comme plus haut) et d'autre part à

$$\exp(A+B) = \sum_{k=0}^{n} \frac{(A+B)^k}{k!} = \sum_{k=0}^{n} \frac{\sum_{j=0}^{k} C_k^j A^j B^{k-j}}{k!}$$

Cherchons alors dans la seconde expression les termes correspondant à un monôme $A^p B^q$ donné : il est clair que le seul terme convenable se trouve dans le développement de $(A+B)^{p+q}$, et vaut $C_{p+q}^p A^p B^q/(p+q)!$; mais comme $C_{p+q}^p/(p+q)! = 1/p!q!$, on en déduit que ce terme est identique à celui correspondant à la première expression, et donc que $\exp(A+B) = \exp(A)\exp(B)$. O étant nilpotente, on a

$\exp(O) = I$; la formule précedente montre donc que $\exp(A)\exp(-A) = \exp(O) = I$, ce qui prouve que $\exp(A)$ est inversible, et que $(\exp(A))^{-1} = \exp(-A)$.

10 Les résultats des questions **5**, **6** et **7** sont valables dans un anneau quelconque (au prix d'un changement de notation : ainsi, par exemple, si $a^k = 0_A$, $1_A - a$ est inversible, et $(1_A - a)^{-1} = 1 + a + a^2 + \cdots + a^{k-1}$) ; mais on ne peut (dans le cas général) définir l'exponentielle d'un élément nilpotent, car, même en convenant que les entiers sont «plongés» dans l'anneau, et en identifiant 3 avec $3_A = 1_A + 1_A + 1_A$, la notation $x/k!$ ne signifie rien (par exemple, dans $\mathbf{Z}$, $n/5!$ n'est défini que si n est un multiple de 120).

Corrigé du devoir n° 7 (Endomorphismes dans un espace de matrices)

1 Calculons $\Phi(A)(M+\lambda N) = A(M+\lambda N)-(M+\lambda N)$. On obtient $\Phi(A)(M+\lambda N) = AM - MA + \lambda(AN - NA) = \Phi(a)(M) + \lambda\Phi(A)(N)$, ce qui montre que $\Phi(A)$ est une application linéaire ; comme $\Phi(A)(M) \in E$, c'est donc un endomorphisme de E (ce qu'on note $\Phi(A) \in \mathcal{L}(E)$).

2 Soit alors l'application Φ qui à A (appartenant à E) fait correspondre l'endomorphisme $\Phi(A)$. Calculons $\Phi(A+\lambda B)$: c'est l'endomorphisme qui à toute matrice M fait correspondre la matrice $(A+\lambda B)M-M(A+\lambda B) = (AM-MA)+\lambda(BM-MB) = \Phi(A)(M)+\lambda\Phi(B)(M)$, ce qui, d'après la définition des opérations sur les endomorphismes, signifie que $\Phi(A + \lambda B) = \Phi(A) + \lambda\Phi(B)$, et donc que Φ est linéaire ; on a donc $\Phi \in \mathcal{L}\big(E, \mathcal{L}(E)\big)$.

3 On a (pour tout M) $\Phi(aI)(M) = aIM - M(aI) = O$, ce qui montre que l'endomorphisme $\Phi(aI)$ est l'endomorphisme nul, c'est-à-dire le vecteur nul de $\mathcal{L}(E)$; par définition, cela veut dire que $aI \in \operatorname{Ker}\Phi$.

4 Réciproquement, soit $A \in \operatorname{Ker}\Phi$, on a donc (pour tout M de E) $AM - MA = O$; prenant en particulier $M = M_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, et posant $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, on obtient $\begin{pmatrix} 0 & a \\ 0 & c \end{pmatrix} = \begin{pmatrix} c & d \\ 0 & 0 \end{pmatrix}$, ce qui montre que $a = d$ et $c = 0$; de même, prenant $M = M_3 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$, on obtient $b = 0$.

5 A est donc une matrice d'homothétie $A = \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix} = aI$; le résultat de la question précédente et celui-ci nous permettent donc d'affirmer que $\operatorname{Ker}\Phi = \{aI\}_{a\in\mathbf{R}} = G$. D'après la formule du rang, on a $\operatorname{rg}(\Phi) = \dim(E) - \dim(\operatorname{Ker}\Phi) = 4 - 1 = 3$, l'image $\operatorname{Im}\Phi$ est donc un sous-espace vectoriel de $\mathcal{L}(E)$ de dimension 3 (rappelons que $\dim(\mathcal{L}(E)) = 16$).

6 Soit alors D l'ensemble des endomorphismes u de E tels que (pour tout M et N de E) on ait $u(MN) = u(M)N + Mu(N)$. Si u et v appartiennent à D, on aura (pour tout M et N) $(u + \lambda v)(MN) = u(MN) + \lambda v(MN) = u(M)N + Mu(N) + \lambda(v(M)N + Mv(N)) = (u + \lambda v)(M)N + M(u + \lambda v)(N)$, ce qui montre que $u + \lambda v$ appartiendra aussi à D. L'endomorphisme nul appartenant évidemment à D, D est donc un sous-espace vectoriel de $\mathcal{L}(E)$. Montrons que $\Phi(A)$ vérifie la propriété caractéristique de D : on a en effet pour tout M et N

$$\begin{aligned}\Phi(A)(MN) = AMN - MNA &= AMN - MAN + MAN - MNA \\ &= (AM - MA)N + M(AN - NA) \\ &= \Phi(A)(M)N + M\Phi(A)(N)\end{aligned}$$

Ainsi, tout endomorphisme de l'image de Φ appartient à D, et donc $\operatorname{Im}\Phi \subset D$.

7 Pour tout M et N de E) on ait $u(MN) = u(M)N + Mu(N)$. Prenant $M = N = I$, on obtient en particulier $u(I) = u(I)I + Iu(I) = 2u(I)$, et donc $u(I) = O$; de même, comme $M_2.M_2 = O$, on aura $u(O) = O = M_2u(M_2) + u(M_2)M_2$; posant $u(M_2) = \begin{pmatrix} x & y \\ z & t \end{pmatrix}$, on en déduit que $\begin{pmatrix} z & t \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & x \\ 0 & z \end{pmatrix} = O$, d'où on tire $z = 0$ et $t = -x$.

8 On obtient de même $u(M_3) = \begin{pmatrix} -t & 0 \\ z & t \end{pmatrix}$; la relation $y + z = 0$ découle du calcul de $u(M_1^2) = u(M_1) = M_1 u(M_1) + u(M_1)M_1$, ce qui aboutit à $u(M_1) = \begin{pmatrix} 0 & y \\ z & 0 \end{pmatrix}$ puis de ce que $u(M_2.M_3) = u(M_1)$.

9 Soit alors $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ avec les relations données, et donc $A = \begin{pmatrix} y+d & -t \\ -x & d \end{pmatrix}$, avec d quelconque et $y = -z$. On vérifie aisément que $AM_2 - M_2A = \begin{pmatrix} x & y \\ 0 & -x \end{pmatrix}$ et que $AM_3 - M_3A = \begin{pmatrix} -t & 0 \\ -y & t \end{pmatrix}$, et donc que $\Phi(A)(M_2) = u(M_2)$ et que $\Phi(A)(M_3) = u(M_3)$; on sait d'autre part que $M_2M_3 = M_1$, et donc que $u(M_1) = u(M_2)M_3 + u(M_3)M_2 = \begin{pmatrix} 0 & t \\ x & 0 \end{pmatrix} = AM_1 - M_1A$; de même $u(M_4) = -u(M_1) = \Phi(A)(M_4)$. Ainsi, u et $\Phi(A)$ prennent les mêmes valeurs sur les quatre matrices de la base canonique; cela montre que $u = \Phi(A)$ (car si f et g sont deux applications linéaires d'un espace de dimension fini E vers F, et si pour tous les e_i d'une base de E, on a $f(e_i) = g(e_i)$, alors $f = g$) et donc que toute matrice u de D est dans l'image de Φ (puisque pour u fixé, on peut évidemment déterminer un A vérifiant ces conditions); d'après les résultats précédents, on a donc $\operatorname{Im}\Phi = D$.

10 Les calculs précédents donnent directement les valeurs de $u(M_i) = AM_i - M_iA$ (sachant qu'ici $x = 1$, $y = 3$, $z = -3$ et $t = -2$), on en déduit la matrice de u :

$$U = \begin{pmatrix} 0 & 1 & 2 & 0 \\ -2 & 3 & 0 & 2 \\ -1 & 0 & -3 & 1 \\ 0 & -1 & -2 & 0 \end{pmatrix};$$

on remarquera qu'il était possible d'obtenir ce résultat sans avoir fait aucune des questions précédentes ! On sait que $\operatorname{Ker}\Phi = \{\lambda I\}$, mais ce n'est pas la question ici ! Remarquant que le rang de U est 2 (on a par exemple $C_4 = -C_1$ et $C_3 = 2C_2 + 3C_1$), on en déduit que $\dim \operatorname{Ker} u = 2$ (formule du rang); comme I et A sont évidemment dans le noyau, et que la famille (I, A) est libre, cette famille est donc une base du plan vectoriel $\operatorname{Ker} u$. Plus aisément, les colonnes de U représentant les images des M_i, on voit que $(u(M_4) = 2M_2 + M_3, u(M_2) = M_1 + 3M_2 - M_4)$ est une base de $\operatorname{Im} u$.

11 Cette question revient en fait à déterminer les valeurs propres de A ! On pouvait la résoudre sans connaître la théorie du chapitre 19, mais il est plus simple de l'utiliser : le polynôme caractéristique $\det(A - \lambda I)$ vaut ici $P_A(\lambda) = (4-\lambda)(1-\lambda) + 2 = \lambda^2 - 5\lambda + 6 = (\lambda - 2)(\lambda - 3)$; les deux valeurs propres sont donc $\alpha = 2$ et $\beta = 3$, et on a bien $|\alpha - \beta| = 1$

12 Dire que $u - \mathrm{id}_E$ et $u + \mathrm{id}_E$ ne sont pas bijectives revient à montrer que 1 et -1 sont valeurs propres de u; c'est-à-dire qu'il existe deux matrices non nulles P_1 et P_{-1} telles que $u(P_1) = AP_1 - P_1A = P_1$ et que $u(P_{-1}) = AP_{-1} - P_{-1}A = -P_{-1}$. Ces matrices ne semblant pas «évidentes», on va effectuer le calcul dans la base (M_i) : on doit résoudre des systèmes d'équations tels que

$$\begin{cases} y + 2z & = x \\ -2x + 3y + 2t & = y \\ -x - 3z + t & = z \\ -y - 2z & = t \end{cases},$$

avec $P_1 = xM_1 + yM_2 + zM_3 + tM_4$. On obtient $x = 2a$, $y = 4a$, $z = -a$ et $t = -2a$, donc $P_1(a) = a(2M_1 + 4M_2 - M_3 - 2M_4)$ (avec $a \neq 0$); de même, $P_{-1}(b) = b(M_1+M_2-M_3-M_4)$ est vecteur propre (avec $b \neq 0$). Les vecteurs (non nuls) I et A du noyau de u étant propres (pour la valeur propre 0), et formant une base avec $P_1(1)$ et $P_{-1}(1)$ (comme on le vérifie aisément), on a donc diagonalisé U, c'est-à-dire qu'on a $P^{-1}UP = \Delta$, où P est la matrice (inversible) de passage (la matrice des «vecteurs» $P_1(1)$, $P_{-1}(1)$, I et A dans la base M_i) et Δ la matrice des valeurs propres (puisque $u(V) = \lambda V$) :

$$P = \begin{pmatrix} 2 & 1 & 1 & 4 \\ 4 & 1 & 0 & 2 \\ -1 & -1 & 0 & -1 \\ -2 & -1 & 1 & 1 \end{pmatrix}; \qquad \Delta = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Corrigé du devoir n° 8 (vissages)

L'espace est rapporté au repère $(O, \overrightarrow{\imath}, \overrightarrow{\jmath}, \overrightarrow{k})$, et on identifie $\mathbf{u} = X\overrightarrow{\imath} + Y\overrightarrow{\jmath} + Z\overrightarrow{k}$ au vecteur-colonne $\begin{pmatrix} X \\ Y \\ Z \end{pmatrix}$.

1 ν est une transformation affine; $\overrightarrow{\nu}$ (l'application linéaire associée) a pour matrice (dans la base orthonormée directe associée au repère) $M = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$; on vérifie aisément que ${}^tMM = I_3$ (ν est donc une isométrie) et que $\det M = 1$ (ν est donc un déplacement). Si $A \left| \begin{matrix} x \\ y \\ z \end{matrix} \right.$ est un point fixe de ν, c'est donc que $x = y + a$, $y = z + b$ et que $z = x + c$; si $a + b + c \neq 0$, ce système n'a pas de solution; si $a + b + c = 0$, l'ensemble des points fixes est l'intersection des deux plans $x = y + a$ et $y = z + b$, soit la droite $M_t \left| \begin{matrix} a+b+t \\ b+t \\ t \end{matrix} \right.$ passant par le point $M_0 \left| \begin{matrix} a+b \\ b \\ 0 \end{matrix} \right.$ et de vecteur directeur $\mathbf{u} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$.

2 On sait, d'après le cours, que si $a + b + c = 0$, ν est une rotation (on vient d'en déterminer l'axe); le calcul d'angle peut se faire par la méthode qui sera rappelée en **6**, mais on peut aussi remarquer que $\nu^3 = \mathrm{Id}_{\mathcal{E}}$, ce qui montre que l'angle vaut $\pm 2\pi/3$; prenant le vecteur $\mathbf{w} = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$ (qui est orthogonal à $\mathbf{u}$), on a $\mathbf{w}' = \overrightarrow{\nu}(\mathbf{w}) = \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}$, et donc $\mathbf{w} \wedge \mathbf{w}' = \begin{pmatrix} -1 \\ -1 \\ -1 \end{pmatrix} = -\mathbf{u}$ ce qui montre que l'angle vaut $-2\pi/3$.

3 Prenons pour r une rotation correspondant à a' et b' quelconques, et $c' = -a - b$, et pour t la translation de vecteur $k\mathbf{u}$. Les formules de la transformation $t \circ r$ sont donc $x' = y + a' + k$, $y' = z + b' + k$ et $z' = x + k - a' - b'$. Par identification, on aura donc $\nu = t \circ r$ si $a = a' + k$, $b = b' + k$ et $c = k - a' - b'$; et ce système possède la solution $k = (a + b + c)/3$, $a' = (-2a + b + c)/3$, $b' = (a - 2b + c)/3$ (et $c' = (a + b - 2c)/3$). La rotation r est donc définie (d'après les calculs faits en **1**) par son axe $\Delta = (M_0, \mathbf{u})$, avec $M_0 \left| \begin{matrix} \frac{-a-b+2c}{3} \\ \frac{a-2b+c}{3} \\ 0 \end{matrix} \right.$, et son angle $\theta = -2\pi/3$.

4 Si f est une translation, on peut prendre pour Δ n'importe quel axe de direction parallèle à celle de la translation, et pour angle $\theta = 0$. Si f est une rotation, on peut prendre pour t la translation de vecteur nul (qui appartient à $\overrightarrow{\Delta}$). Dans le cas général, la transformation $g = t \circ f$ est le composé de deux déplacements; c'est donc un déplacement; comme $g(A) = t_{\overrightarrow{AB}}(f(A)) = t_{\overrightarrow{AB}}(B) = A$, g est un déplacement admettant un point fixe; c'est donc (d'après le cours) une rotation d'axe Δ passant par A.

5 Soit $\mathcal{P}$ un plan orthogonal à Δ; on a donc $f(\mathcal{P}) = (t^{-1} \circ g)(\mathcal{P})$; or $g(\mathcal{P}) = \mathcal{P}$, et l'image d'un plan $\mathcal{P}$ par une translation est un plan $\mathcal{Q}$ parallèle (ou confondu) ave $\mathcal{P}$. On peut facilement trouver un vecteur $\mathbf{w}$ orthogonal à $\mathcal{P}$ et tel que $\mathcal{P} = t_{\mathbf{w}}(\mathcal{Q})$ (poser $\{A\} = \mathcal{P} \cap \Delta$, $\{B\} = \mathcal{Q} \cap \Delta$, et $\mathbf{w} = \overrightarrow{BA}$), et on aura alors $t_{\mathbf{w}} \circ f(\mathcal{P}) = t_{\mathbf{w}}(\mathcal{Q}) = \mathcal{P}$; $\mathcal{P}$ étant (globalement) invariant par l'isométrie $t_{\mathbf{w}} \circ f$, on voit que $t_{\mathbf{w}} \circ f\big|_{\mathcal{P}}$ est une isométrie de $\mathcal{P}$; orientant $\mathcal{P}$ à l'aide de Δ, on voit que c'est un déplacement, donc une rotation de $\mathcal{P}$ de centre I (ce ne peut pas être une translation, car f en serait une aussi). Ainsi, la transformation, restreinte au plan $\mathcal{P}$, est un vissage d'axe parallèle à Δ, passant par I, d'angle α, de vecteur de translation $\mathbf{w}$. Il ne reste plus qu'à montrer que ce résultat est valable dans tout plan. L'application vectorielle $\overrightarrow{f}$ associée à f étant une rotation vectorielle d'axe $\overrightarrow{\Delta}$ et d'angle α, soit I' le point de $\mathcal{P}'$ centre de la rotation $t_{\mathbf{w}'} \circ f\big|_{\mathcal{P}'}$; on aura donc $\overrightarrow{f}(\overrightarrow{II'}) = \mathbf{w}' - \mathbf{w} + \overrightarrow{II'}$, ce qui montre que $\overrightarrow{II'} \in \overrightarrow{\Delta}$, et donc que (II') est parallèle à Δ (donc orthogonal à$\mathcal{P}$), et le même raisonnement prouve que $\mathbf{w}' = \mathbf{w}$. Finalement, tout déplacement de l'espace est donc un vissage (éventuellement dégénéré).

6 Si M est sur l'axe du vissage, on a vu que $\overrightarrow{Mf(M)} = \mathbf{w}$ est colinéaire à $\overrightarrow{\Delta}$; réciproquement, si M n'est pas sur l'axe, on a $\overrightarrow{Mf(M)} = \mathbf{w} + \overrightarrow{Mr(M)}$, et comme $\overrightarrow{Mr(M)}$ est orthogonal à Δ, il en résulte qu'alors $\overrightarrow{Mf(M)}$ n'est pas colinéaire à $\overrightarrow{\Delta}$. Étudions alors la transformation f. C'est une transformation affine, de matrice associée $F = \dfrac{1}{3}\begin{pmatrix} -2 & -1 & 2 \\ 2 & -2 & 1 \\ 1 & 2 & 2 \end{pmatrix}$; et on vérifie aisément que F est une matrice orthogonale (${}^tFF = I_3$) de déterminant 1, donc que f est un déplacement. Déterminons la direction de l'axe du vissage : il n'est pas nécessaire de diagonaliser F, puisqu'on sait que 1 est valeur propre; il suffit de chercher les $\mathbf{v}$ tels que $\overrightarrow{f}(\mathbf{v}) = \mathbf{v}$, donc de résoudre $\begin{cases} -2x - y + 2z = 3x \\ 2x - 2y + z = 3y \\ x + 2y + 2z = 3z \end{cases}$, de solution $\begin{pmatrix} a \\ a \\ 3a \end{pmatrix}$. Cherchons à présent l'angle de la rotation, en prenant pour vecteur directeur de l'axe le vecteur $\mathbf{u} = \begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix}$; comme $\mathbf{v} = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$ est orthogonal à $\mathbf{u}$, il sufit de déterminer l'angle (orienté) $(\widehat{\mathbf{v}, \overrightarrow{f}(\mathbf{v})})$, or $\mathbf{v}' = \overrightarrow{f}(\mathbf{v}) = \begin{pmatrix} -1/3 \\ 4/3 \\ -1/3 \end{pmatrix}$, et comme $\mathbf{v} \wedge \mathbf{v}' = \begin{pmatrix} 1/3 \\ 1/3 \\ 1 \end{pmatrix} = (||\mathbf{v}||^2 \sin\alpha)\mathbf{u}/||\mathbf{u}||$, on obtient finalement $\alpha = \arcsin(\sqrt{11}/6)$. La méthode précédente amène à chercher les points M tels que $\overrightarrow{Mf(M)}$ soit colinéaire à $\mathbf{u}$; posant les équations correspondantes, on aboutit au système $\begin{cases} -5x - y + 2z + 3 = k \\ 2x - 5y + z + 3 = k \\ x + 2y - z + 9 = 3k \end{cases}$, de solution la droite passant par O ($x = y = z = 0$ et $k = 3$) et de vecteur directeur $\mathbf{u}$, comme on pouvait s'y attendre; cette droite est donc l'axe du vissage. Comme $f(O) = O'$ tel que $\overrightarrow{OO'} = \mathbf{u}$, on voit que $\mathbf{u}$ est le vecteur de translation du vissage, ce qui, en revanche, est une simple coïncidence.