

OC
ES
T

201

Irem de Besançon

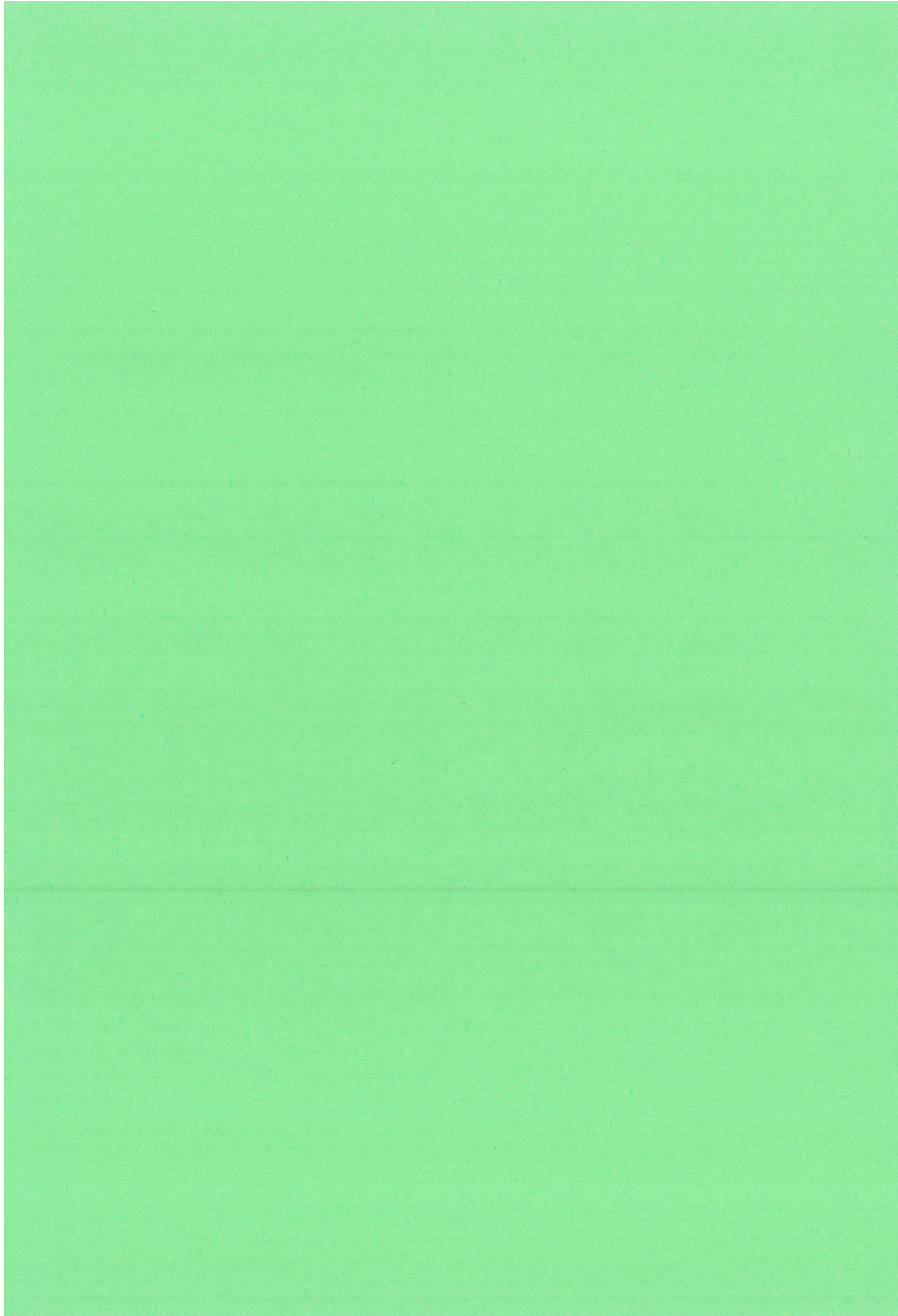
IBC97019.PDF

Le Théorème de Cayley Hamilton

IRM de LYON
BIBLIOTHÈQUE
Université Claude Bernard - LYON I
43, Bd du 11 Novembre 1918
69622 VILLEURBANNE Cedex



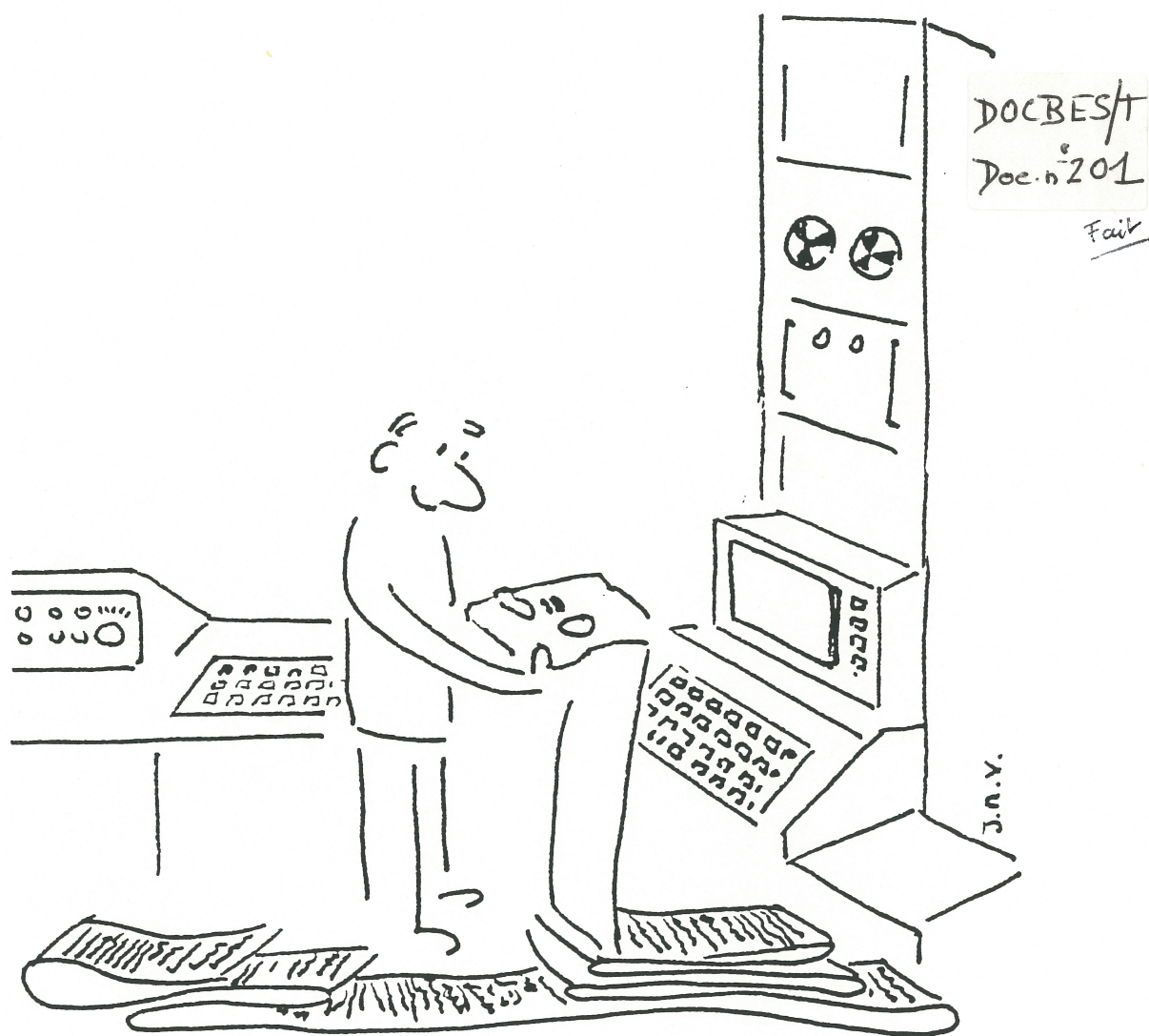
Jean MERKER
Février 1992



Le Théorème de Cayley Hamilton

IREM de LYON
BIBLIOTHEQUE

Université Claude Bernard - LYON I
43, Bd du 11 Novembre 1918
69622 VILLEURBANNE CEDEX



Jean MERKER

Février 1992

LE THEOREME DE CAYLEY HAMILTON

Bonsoir à tous !

Voici, sous forme de "fascicule-IREM", la première partie de l'U.V. Alg. II du C.T.U. de Besançon (DEUG SSM). Elle traite du théorème de Cayley-Hamilton et la réduction de Jordan des matrices (les notions de valeurs propres et vecteurs propres étant supposées connues).

Les pages numérotées T., se réfèrent au cours proprement dit prolongé par les pages E., qui proposent des compléments sous forme d'exercices commentés dans la partie numérotée C..

Nous avons donné une tournure genre "rébus-mathématique" à cette première partie qui est une digression autour du théorème de Cayley-Hamilton.

1. Voici le théorème :

Théorème :

Prenez une matrice *carrée* A . Formez la matrice $A - \lambda I$ (I , matrice unité). Calculez le déterminant de cette matrice. Ordonnez par rapport aux puissances de λ . Remplacez λ par "la variable X ". Vous obtenez ainsi un polynôme en X .

Remplacez maintenant X par A . Si vous effectuez alors la somme qui est sous vos yeux, vous trouvez 0 .

Miracle : cela marche toujours ! Ne vous évertuez pas de prendre les matrices, p. ex. à 3781 lignes et colonnes, une par une et de vérifier que cela marche : vous seriez pires que Saint-Thomas !

Comme ce théorème se trouve écrit dans presque tous les livres d'algèbre, il y a de fortes chances qu'il soit vrai (le continent australien existe, enfin, je crois !)

Nous pourrions donc nous passer, le cœur léger, d'une démonstration de ce théorème. Mais comme nous sommes des apprentis mathématiciens et que les programmes sont ce qu'ils sont (soupir), nous allons donc chercher la raison de ce théorème ou plutôt les multiples raisons qui militent en faveur de ce théorème, ou mieux, rébus : pourquoi est-ce que cela marche toujours ?

2 - Si on veut comprendre que quelque chose marche *toujours*, il faut pousser du côté du *général* : sinon c'est l'arbre qui risque de cacher la forêt.

Soit donc $A = (a_{i,j})_{i,j}$; une matrice $n \times n$. On forme le déterminant

$$\det(A - \lambda I).$$

On remplace, calculs effectués, λ par X pour obtenir un polynôme $\chi_A(X)$, *polynôme dit caractéristique* de A .

On écrit : $\det(A - X.I) = \chi_A(X)$.

Ensuite on remplace X par A et le théorème dit : $\chi_A(A) = 0$

Ces écritures mathématiques nous paraissent parfaitement anodines, mais en fait, elles sont semées d'embûches.

Par exemple, voici une démonstration fausse :

Nous avons $\chi_A(X) = \det(A - X.I)$. Remplaçant X par A donne :

$\chi_A(A) = \det(A - A.I) = \det(0) = 0$ et le théorème est démontré.

Vous chercherez pourquoi : ex 2.1.

C'est moi qui vous ai gentiment induit en erreur. Voici comment :

On peut dire que $\det(A - \lambda I)$ est une fonction de λ et on écrit $\chi(\lambda) = \det(A - \lambda I)$. Une vérification simple (ou un simple coup d'oeil) nous montre qu'en fait cette fonction χ est une fonction polynomiale. Il lui correspond donc un polynôme qu'on confond avec χ . On écrit $\chi(X) = \det(A - X.I)$: voici le coup de force décisif : on remplace λ par X , mais ici le coup de force est justifiable, par contre, en continuant sur la bonne pente et en remplaçant ensuite X par A on fait une erreur; erreur facilement reconnue par ses conséquences visibles mais qui pose alors le problème angoissant : *Pourquoi est-ce que je peux écrire $\chi(X) = \det(A - X.I)$ en remplaçant λ par X ?*

La réponse à cette question va vous étonner.

Avant d'aller plus loin, notons le coup de force qu'on a fait dans les lignes précédentes en disant qu'il y a identité entre "fonction polynôme" et "polynôme".

Jusqu'ici, la nature des nombres qui figurent *dans* les matrices qu'on manipule n'est pas intervenue.

L'énoncé du théorème de CH est "écrivable" si on suppose que l'ensemble des "nombres" qu'on utilise forme un *anneau commutatif* $\mathcal{A}$ et le théorème est *vrai* dans ce cas.

Dans la pratique des apprentis-mathématiciens $\mathcal{A}$ est soit l'ensemble des nombres réels $\mathbb{R}$, ou bien $\mathbb{C}$, le corps des complexes. Mais pour la pratique d'un ordinateur qui exécute le programme contenu dans l'énoncé du théorème, jamais on ne fait appel à l'inverse d'un nombre et on peut donc conjecturer hardiment :

Si le théorème est vrai quand les nombres employés forment un corps, il doit aussi être vrai quand cet ensemble n'est qu'un anneau commutatif (possédant un élément unité noté 1).

Ce qui change quand on manipule des matrices "à coefficients dans un anneau"? Pas grand chose sauf que pour l'inverse d'une matrice c'est encore pire : si vous écrivez "2" pour "1 + 1" vous n'avez pas le droit d'écrire sans vérifications préalables "1/2", car dans un anneau, certains nombres ont un inverse et d'autres pas.

Nous allons donc faire un progrès psychologique en étant d'accord de manipuler des matrices à coefficients dans un anneau, *ceci quel que soit cet anneau* évidemment, voir exo T2.5.

Nous voyons donc que les "coefficients" utilisés peuvent être dans des anneaux variés. Mais alors le coup de force d'écrit au milieu de la page T1 risque de ne plus être valable.

Vous savez que si les coefficients sont dans $\mathbb{R}$, ou $\mathbb{C}$ ou plus généralement dans un corps *infini*, il y a identité entre polynômes et fonctions polynômes. De façon plus technique, on a une flèche $K[X] \longrightarrow \mathcal{F}(K)$ ou $\mathcal{F}(K)$ désigne les fonctions de K dans K , définie par $P(x) \longrightarrow (la fonction \lambda \longrightarrow P(\lambda))$ fonction obtenue en substituant λ à X (et non X à λ).

C'est-à-dire tout polynôme fournit, produit une fonction polynôme, mais qu'une fonction polynôme pourrait fort bien être produite par deux polynômes *différents*. Si le corps est infini, dire qu'il y a "identité" entre polynômes, signifie que cela ne peut pas arriver.

Par contre, prenons $K = \mathbb{Z}/3\mathbb{Z}$ et $P_1 = X^3 - X$ $P_2 = 0$.. On a $P_1 \neq P_2$, mais les fonctions produites sont égales.

Considérons maintenant la matrice à coefficients dans K avec, toujours, $K = \mathbb{Z}/3\mathbb{Z}$

$$A = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

Ses valeurs propres sont $0, 1, -1$.

Formons la matrice $A - \lambda I$ avec $\lambda \in K$ et calculons $\det(A - \lambda I)$. On trouve que ce déterminant vaut $-\lambda^3 + \lambda$. Or cette quantité est nulle (pour $\lambda \in K$). Dans une quantité nulle, comment maintenant remplacer λ par X ? Angoisse. Mais non, vous me direz, il faut *laisser* l'écriture $-\lambda^3 + \lambda$, *ne pas* utiliser les simplifications dues au fait particulier qu'ici $K = \mathbb{Z}/3\mathbb{Z}$. Il faut remplacer λ par X dans $-\lambda^3 + \lambda$.

Or, "ne pas se laisser aller à des simplifications" revient exactement à ne pas raisonner avec des fonctions polynômes, mais avec des polynômes. Voici donc maintenant la réponse étonnante annoncée au milieu de la page T1.

Le polynôme $\chi_A(X)$ dont il est question dans le théorème de CH et qu'on appelle polynôme caractéristique s'obtient de la manière suivante :

On forme la matrice $A - XI$

Cette matrice a des coefficients dans l'anneau $\mathcal{A} = K[X]$.

Et on prend le déterminant de cette matrice déterminant qui est un élément de $\mathcal{A}$ c.-à-d. un polynôme en X : $\chi_A(X) = \det(A - XI)$.

En remplaçant $X \longrightarrow \lambda$ et en égalant à 0 : $\chi_A(\lambda) = 0$, on a l'équation aux valeurs propres de A .

En remplaçant X par A , on obtient une matrice à coefficients dans K et cette matrice est nulle : théorème de CH.

C'est donc X qui est prioritaire et on a les substitutions $X \longrightarrow \lambda$ et $X \longrightarrow A$ et bien d'autres encore que nous allons étudier maintenant au n° 3.

3 - Polynômes matriciels. L'exponentielle d'une matrice.

Dans ce numéro, nous allons toujours rester dans les "embûches" des notations mathématiques, étudier les substitutions " $X \longrightarrow$ quelque chose" et faire un divertissement que j'espère vous allez apprécier, car n'est-il pas beau de pouvoir écrire e^A où $e = 2,7...$ et où A est une matrice carrée quelconque ?

Mais peut-être êtes-vous fatigués, arrivés à ce point. Alors faites une pause ! Car penser est difficile, sauf quand on sait déjà, et comme on ne sait pas, si souvent...

Nous allons étudier la substitution $X \longrightarrow A$, une matrice.

Soit $P(X) = a_0 + a_1 X + \dots + a_n X^n$ un polynôme, on pose
 $P(A) = a_0 I + a_1 A + \dots + a_n A^n$.

Cela suppose que $a_i A^i$ a un sens : cela revient à demander que les a_i multiplient les coefficients de A^i .

Voici des situations qui rendent cela possible :

- 1) Les coefficients de P sont dans K , anneau ou corps les coefficients de A sont aussi dans K
- 2) Les coefficients de P sont dans $\mathbb{C}$ (resp $\mathbb{R}$), les coefficients de A sont dans $\mathbb{R}$ (resp $\mathbb{C}$)
- 3) Si $\mathcal{A}$ est un sous anneau de $\mathcal{B}$, anneau, les coefficients de P sont dans $\mathcal{A}$ (resp $\mathcal{B}$) les coefficients de A sont dans $\mathcal{B}$ (resp $\mathcal{A}$).

Considérons seulement le premier cas : on a donc défini une application

$$\sigma : K[X] \longrightarrow \mathcal{M}_{n,n}(K)$$

de l'anneau des polynômes $K[X]$ dans l'anneau des matrices $n \times n$ sur K , obtenue en substituant A à X .

$P(A)$ s'appelle un polynôme matriciel.

Nous avons les formules

$$\begin{aligned} (P_1 + P_2)(A) &= P_1(A) + P_2(A) \text{ où } P_1, P_2 \in K[X] \\ (P_1 P_2)(A) &= P_1(A) P_2(A) \end{aligned} \quad (h)$$

dont l'évidence saute aux yeux, et pourtant ... !

Analysons de plus près par exemple la formule $(P_1 P_2)(A) = P_1(A) P_2(A)$.

Introduisons la lettre " π " pour "produit de deux polynômes" et " μ " pour "multiplication de deux matrices". La lettre σ est introduite plus haut ("substitution de A à X ").

Alors la formule en question peut être schématisée par " $\sigma \circ \pi = \mu \circ \sigma$ " ce qui signifie en gros que les opérations de substitution et de multiplication commutent. Or la commutation de deux opérations n'est pas toujours vraie ! Il y en a même qui sont remarquablement fausses et donnent lieu à des identités remarquables : par exemple $(a+b)^2 = a^2 + 2ab + b^2$.

(Voir ex.4 1).

En plus, il y a des notations mathématiques qui peuvent nous pousser à faire cette erreur (de commutation, alors qu'il n'en est rien). Par exemple, $\det(A - XI)$!! Soit " S " la lettre qui désigne l'opération qui consiste à prendre le déterminant de la matrice $A - XI$. Alors, dans l'énoncé du théorème de CH on nous dit de faire " $\sigma \circ S$ " : d'abord " S " puis " σ ". Si vous faites l'inverse, ce que vous suggère insidieusement l'écriture mathématique, " $S \circ \sigma$ ", vous vous perdez dans les sables de la confusion. Tout cela pour dire que ces formules (celles du milieu de cette page) doivent être *réalisées* par chacun comme étant justes, ce qui ne vous posera pas de problème ; par contre, il est probable que vous ressentirez un certain malaise si quelqu'un de mauvaise foi vous demande de les *démontrer*, car les notations mathématiques sont parfois choisies de sorte qu'elles font éclater l'évidence de certaines relations mais qu'elles en rendent la démonstration plus obscure.

Les formules (h) expriment en langage savant que σ est un homomorphisme de l'anneau $K[X]$ dans l'anneau $\mathcal{M}_{n,n}(K)$

Mais ce langage savant est celui qui est approprié pour dégager le pourquoi de ces relations (voir éventuellement ex4.3).

La situation précédente se généralise de la façon suivante : on suppose que K est un sous anneau de l'anneau $\mathcal{A}$ et on fixe un élément $a \in \mathcal{A}$.

On a alors une application σ :

$$\sigma : K[X] \longrightarrow \mathcal{A} : P(X) \longrightarrow P(a)$$

qui est un homomorphisme d'anneaux. (l'anneau $\mathcal{M}_{n \times n}$ est remplacé par $\mathcal{A}$.)

Un polynôme matriciel est donc une somme finie $P(A) = \sum_{i=1}^n a_i A^i$

Serait-il possible de définir des sommes infinies ?

Considérons la somme infinie de nombres réels

$$\sum_{n=0}^{\infty} \frac{x^n}{n!} \quad \text{où } x \in \mathbb{R}.$$

Vous savez que cette somme existe et vaut :

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$$

Cette série converge absolument, ce qui entraîne sa convergence. Mais la convergence de cette série entraîne aussi son absolue-convergence : car

$$(i) \quad |x^n / n!| = |x|^n / n!$$

et on reconnaît le terme général de la série qui a comme somme $e^{|x|}$.

Cette remarque permet maintenant de dire que la série $\sum_{n=0}^{\infty} \frac{z^n}{n!}$ ($z \in \mathbb{C}$) est convergente ;

$$\text{On note } e^z = \sum_{n=0}^{\infty} \frac{z^n}{n!}$$

En effet, cette série est absolument convergente :

$$(ii) \quad |z^n / n!| = |z|^n / n!$$

série qui donne $e^{|z|}$. Remarquons que dans (i) et (ii) une inégalité aurait suffi !

Alors pourquoi ne pas oser écrire :

$$e^A = \sum_{n=0}^{\infty} \frac{A^n}{n!}$$

Pour cela deux conditions doivent être remplies :

1) Une série $S = \sum_{n=0}^{\infty} u_n$ n'étant rien d'autre qu'une limite d'une suite : $S = \lim_{n \rightarrow \infty} S_n$

(avec $S_n = \sum_{k=0}^n u_k$ on doit pouvoir définir dans le cas des matrices la limite d'une suite de matrices :

$$B = \lim_{n \rightarrow \infty} B_n$$

Or, dans le cas des nombres ceci signifie que $|B - B_n|$ tend vers 0.

Il faut donc dans le cas des matrices l'analogue d'une valeur absolue, notée $|B|$.

Les propriétés qu'on demande à cette valeur absolue sont :

$$\begin{aligned} \text{(i)} \quad & |B| = 0 \iff B = 0 \\ \text{(ii)} \quad & |A + B| \leq |A| + |B| \end{aligned}$$

par exemple, si $B = (b_{i,j})$ on peut prendre $|B| = \sqrt{\sum (b_{i,j})^2}$

2) La deuxième condition est qu'on puisse refaire la démonstration utilisée dans le cas de z .

Pour cela, il suffirait que la valeur absolue $|B|$ ait une dernière propriété :

$$\text{(iii)} \quad |AB| \leq |A| |B|$$

ce qui entraîne $|A^n| \leq |A|^n$

La valeur absolue $\sqrt{\sum (b_{i,j})^2}$ ne convient pas. On peut trouver une autre valeur absolue, "équivalente" à la première et qui possède en plus la propriété (iii) (voir ex 6).

Ecrivons donc tranquillement $e^A = \sum_{n=0}^{\infty} \frac{A^n}{n!}$

Pour bien montrer que l'exponentielle de matrice n'est que le prolongement naturel de la construction de l'exponentielle d'un nombre réel, remplaçons "A" par "a" :

$$e^a = \sum_{n=0}^{\infty} \frac{a^n}{n!}$$

Alors ou bien $a \in \mathbb{R}$, ou $a \in \mathbb{C} = \mathbb{R}^2$ ou $a \in \mathbb{R}^{n^2}$, car une matrice $(n \times n)$ peut être représentée par un point de $\mathbb{R}^{n^2}$.

Il faut maintenant regarder si les propriétés classiques de l'exponentielle restent encore valables :

La propriété fondamentale est :

$$e^{a+b} = e^a e^b$$

Or, si vous regardez la démonstration de cette formule dans un cours d'analyse, pour le cas $a \in \mathbb{R}$ ou $a \in \mathbb{C}$, et si vous faites bien attention, vous remarquerez qu'on utilise la propriété évidente des nombres : $ab = ba$.

Vous en tirerez comme conclusion que la formule $e^{a+b} = e^a e^b$ reste valable pour les matrices a, b si elles commutent :

$$e^{a+b} = e^a e^b \text{ si } ab = ba$$

De cette formule résulte que e^a est toujours inversible et vous trouverez facilement son inverse (voir ex 7.3).

Une autre propriété fondamentale de l'exponentielle réelle est

$$\frac{de^x}{dx} = e^x \quad \text{ou} \quad \frac{de^{ta}}{dt} = a.e^{ta}.$$

Cette formule aussi reste valable dans le cas général (voir ex 7.2) à condition de savoir lire cette formule dans le cas général : la fonction qu'on dérive est la fonction de la variable $t \in \mathbb{R}$ à valeur dans $\mathbb{R}^k$ avec $k = n^2$, car e^{ta} est une matrice $(n \times n)$ c.-à-d. un point de $\mathbb{R}^k$. Il s'agit donc de dériver une fonction vectorielle. Le second membre est le produit de la matrice a par la matrice e^{ta} . Ce produit est aussi égal à e^{ta} . (voir ex 7.3).

Cette formule montre l'intérêt de la fonction exponentielle pour résoudre une équation différentielle :

$$x' = a.x$$

Le cas élémentaire : x est une fonction inconnue de t à valeurs réelles. Solution : $x = ce^{ta}$ où $c \in \mathbb{R}$ est la condition initiale : $x(0) = c$. Mais l'équation $x' = a.x$ peut avoir un sens plus général : x est une fonction inconnue de $t \in \mathbb{R}$ à valeurs vectorielles $\in \mathbb{R}^n$ et a est une matrice $n \times n$.

Les solutions à cette équation sont donc de même $x = e^{ta}.c$.

Question : Pourquoi le c passe t-il de l'autre côté?

Dans les calculs pratiques on est donc ramené à calculer e^{ta} pour une matrice a donnée.

4- Après cet interlude, les notations étant précisées, les embûches étant évitées, commençons à cerner dans ce n° un peu plus le théorème de CH.

Il faut, avant tout, dire que l'opération essentielle qui intervient dans la construction du polynôme caractéristique χ_A est le déterminant ! Or, le déterminant dans le cas général des matrices $n \times n$, ouille, ouille !!

C'est pourquoi dans ce numéro, allons nous tourner autour du pot en essayant d'en dire le plus sur CH en parlant le moins possible du "dét".

Dans cette optique, nous avons deux voies de contournement :

1) ne connaissant pas vraiment le polynôme χ_A (à cause de notre ignorance sur dét) nous allons nous intéresser à tous les polynômes $P(X)$ qui sont tels que $P(A) = 0$.

2) On va tester CH pour des matrices A qui sont d'un type simple par rapport à la fonction dét, et on espère après ces divers tests en avoir assez fait pour que le cas général en découle, ceci toujours sans savoir grand chose sur dét !

1) Commençons donc par considérer l'ensemble

$$N = \{P \in K[X] / P(A) = 0\}.$$

Cette ensemble n'est pas vide, car le polynôme $0 \in N$ et on a les propriétés de stabilité suivante :

$$(J) \quad \begin{aligned} &\text{Si } P_1 \in N \text{ et } P_2 \in N \text{ alors } P_1 + P_2 \in N \\ &\text{Si } Q \in K[X] \text{ et si } P \in N \text{ alors } QP \in N. \end{aligned}$$

En langage savant, on dit que N est un idéal de l'anneau $K[X]$. C'est d'ailleurs le noyau de l'homomorphisme σ de la page T4.

Or, un sous ensemble qui a les propriétés de stabilité (J) dans $K[X]$ n'est pas quelconque à cause de la division euclidienne de deux polynômes (la même chose doit donc être vraie pour l'anneau $\mathbb{Z}$!).

En effet, on voit facilement que :

Dans N existe un polynôme unitaire de degré minimum (et un seul), P_0 , tel que tout polynôme P de N soit un multiple de P_0 : $N = K[X] \cdot P_0 = \{QP_0 / Q \in K[X]\}$.

En effet, si P_0 est choisi de degré minimum dans N et si P est quelconque dans N , on peut écrire $P = QP_0 + R$ où R est soit nul, soit un polynôme de degré strictement plus petit que celui de P_0 , or la deuxième alternative est exclue, car en écrivant $R = P - QP_0$ on voit que d'après (J), R est un élément de N ; or P_0 a le plus petit degré de tous les polynômes de N .

NB : il existe une légère erreur dans ce qui précède !

Le polynôme P_0 s'appelle le *polynôme minimal* de A : on le notera P_A : c'est le polynôme unitaire de plus petit degré qui annule A et, d'après ce que nous venons de démontrer :

Le polynôme minimal P_A divise χ_A , pol-car

$$\chi_A = Q_0 P_A, \text{ d'après de th de CH.}$$

Il en résulte que *toute racine de P_A est racine de χ_A , c'est-à-dire valeur propre de A* .

Si $K = \mathbb{R}$ vous savez qu'il y a des polynômes, qui peuvent être caractéristiques, sans racines dans $K = \mathbb{R}$, par contre étant décomposables entièrement dans $K = \mathbb{C}$. On dit que $\mathbb{C}$ est la "clôture algébrique" de $\mathbb{R}$.

On démontre que pour tout corps K , il existe une telle "clôture algébrique" $\bar{K}$. On peut donc dire que *toutes les racines de P_A sont des racines de χ_A* .

Nous allons voir que la réciproque est vraie : toute racine de χ_A est racine de P_A de sorte que polynôme minimal et caractéristique ont les mêmes racines avec des multiplicités différentes.

Il résulte de la relation du haut que :

$$\deg P_A \leq \deg \chi_A = n \text{ (A matrice } n \times n \text{)}.$$

Il se peut que l'inégalité soit stricte. Rappelons que toutes ces conclusions supposent la validité du théorème de CH.

Mais laissons pour le moment l'étude de la collection des polynômes Q tels que $Q(A) = 0$. Nous allons y revenir en exploitant l'idée de "vecteur propre".

2) Faisons maintenant quelques tests sur CH.

a) Essayez vous-même avec les matrices suivantes :

$$A = 0; \quad A = I, \quad A = \begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix} \quad A = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & \dots \\ \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & \lambda_n \end{pmatrix}$$

Le théorème de Cayley-Hamilton est donc facilement vérifié pour des matrices diagonales. Cela fait peu, mais plus qu'on ne croit. En effet, vous savez diagonaliser une matrice A (si elle se laisse faire !). Cela revient à chercher une matrice inversible S telle que la matrice

$$A' = S^{-1} A S$$

soit une matrice diagonale.

Deux matrices A et A' qui vérifient la relation précédente sont dites semblables.

Question : Est-ce que le théorème de CH est vrai pour une matrice A semblable à une matrice A diagonale ?

Il faudrait qu'on puisse facilement montrer que

$$\chi_{A'}(A') = 0 \text{ entraîne } \chi_A(A) = 0 .$$

Or, on a tout simplement (ex) : $\chi_{A'}(X) = \chi_A(X)$.

Et il va en résulter facilement que $\chi_A(A) = 0$ sachant que $\chi_{A'}(A') = 0$ (ex).

Malheureusement, une matrice quelconque n'est pas nécessairement diagonalisable. Par contre, on peut justifier la phrase :

En général, une matrice est diagonalisable ;

Pour simplifier, considérons les matrices 2×2 sur $\mathbb{R}$. Soit $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$

Le calcul des valeurs propres vous donnent le polynôme :

$$X^2 - (a + d)X + (ad - bc) = X^2 - \text{Tr}A \cdot X + \det A .$$

Or, vous savez que si les valeurs propres sont distinctes, alors la matrice est diagonalisable.

Mais, un polynôme $X^2 + \alpha X + \beta$ a "en général" des racines distinctes.

Plus précisément, un tel polynôme peut être représenté par le point (α, β) de $\mathbb{R}^2$.

Seuls les points (α, β) situés sur la courbe $\alpha^2 - 4\beta = 0$ (une parabole) posent problème. Mais cet ensemble, par rapport à l'ensemble de *tous* les points de $\mathbb{R}^2$, est d'une certaine façon "négligeable".

Nous allons, une autre fois, quand vous serez bien reposés, rendre cette idée intuitive de "négligeable" rigoureuse et vous aurez alors une démonstration du théorème de CH.

Si vous êtes toujours dans le coup, vous me ferez la remarque que non seulement les points (α, β) avec $\alpha^2 - 4\beta = 0$ posent problème, mais tous les points tels que $\alpha^2 - 4\beta < 0$ et alors ça en fait beaucoup.

Mais, en fait, il n'en est rien, car si $\alpha^2 - 4\beta < 0$ c.-à-d. quand il n'y a pas de valeurs propres réelles du tout, il y a néanmoins des valeurs propres complexes et elles sont distinctes ! Donc, en raisonnant maintenant comme si la matrice (à coefficients réels !) était une matrice sur $\mathbb{C}$ et diagonalisable sur $\mathbb{C}$, on en déduit que son polynôme caractéristique "sur $\mathbb{C}$ " annule la matrice. Or, que vous calculiez le polynôme caractéristique de A "sur $\mathbb{R}$ " ou "sur $\mathbb{C}$ ", le résultat est toujours le même !! Donc, ces points ne posent en fait pas problème. Retenez bien la remarque qui est faite dans les dernières lignes : elle reservira dans le test suivant.

b) Essayez maintenant CH avec la matrice

$$A = \begin{pmatrix} \lambda & 0 \\ a & \mu \end{pmatrix}$$

Essayez ensuite

$$A = \begin{pmatrix} \lambda & 0 & 0 \\ a & \mu & 0 \\ c & b & \nu \end{pmatrix}$$

(surtout : laissez votre déterminant sous forme de produit de facteurs).

Quand vous aurez fini et que vous serez de nouveau installés dans votre fauteuil, nous continuerons.

Le pas suivant consiste évidemment à tester les matrices du type :

$$A = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ n' & \lambda_2 & \dots & \dots \\ im & por & \dots & \dots \\ te & \dots & quoi & \lambda_n \end{pmatrix}$$

matrices triangulaires (inférieures). Vous serez guidés par l'énoncé de l'exercice 10.2. Mais vous avez le temps de faire ce test. Restez donc avec moi pour le moment.

Une matrice semblable à une matrice triangulaire vérifie bien sûr aussi le test de CH : il suffit de refaire (texto !) le raisonnement qu'on vous a demandé de faire au milieu de la page T.9..

Or maintenant, si vos souvenirs sont bons, il est vrai que sur $\mathbb{C}$ toute matrice est semblable à une matrice triangulaire.

Ce qui veut dire que CH est bel et bien démontré dans le cas où $K = \mathbb{C}$.

On peut rapidement régler le sort de $K = \mathbb{R}$ si vous voulez bien vous rappeler la remarque page précédente sur laquelle j'ai attiré votre attention.

Mais si vous n'êtes pas fatigués, je commence à l'être.

Alors ! A la semaine prochaine ! Bonsoir.

BONSOIR !

Le feuilleton continue. Et comme pour tout feuilleton qui se respecte, un résumé pour bien rappeler où se situent les drames !

Nous avons finalement passé beaucoup de temps à mettre en place l'énoncé du théorème de CH. Le résultat essentiel de cette mise en place est l'intuition que vous avons que ce théorème doit être vraie *quel que soit la nature des nombres employés*.

Ensuite, nous avons tourné autour du théorème en empruntant deux voies. La deuxième voie nous a donné une démonstration du théorème. Nous me direz : "Bon, ça y est ! On l'a démontré et on peut donc s'arrêter. Bonne nuit !".

D'abord, j'espère que vous n'êtes pas fatigués en commençant cette quinzaine, ensuite, vous avez oublié qu'on avait remarqué dès le départ qu'on pouvait, le coeur léger, se passer d'une démonstration !

La démonstration est secondaire, c'est toute l'activité mathématique autour d'une idée qui importe.

Continuons donc nos investigations. Nous allons le faire en développant nos deux voies. (voir p. T.7).

1) *Développons la première voie :*

L'idée annoncée était de géométriser le problème : à la matrice A correspond un endomorphisme u de $\mathbb{R}^n$ (la base canonique ayant été choisie).

Si λ est une valeur propre de A les vecteurs propres associés forment un sous-espace E_λ de $E = \mathbb{R}^n$. Dire que u est diagonalisable revient à dire que E est somme directe des E_λ . Ceci n'est pas toujours le cas :

$$E \supset E_\lambda \oplus \dots \oplus E_{\lambda_r} \text{ seulement, en général.}$$

On peut écrire $E_\lambda = \text{Ker}(A - \lambda I)$ et $A - \lambda I$ correspond au facteur $(X - \lambda)$ du polynôme caractéristique χ .

Ce polynôme peut s'écrire :

$$\chi_A = \prod_{i=1}^r (X - \lambda_i)^{k_i} \quad (k_i, \text{ multiplicité de } \lambda_i).$$

Nous allons voir qu'on a *toujours* :

$$E = E'_{\lambda_1} \oplus \dots \oplus E'_{\lambda_r} \text{ toujours où } E'_{\lambda_i} = \text{Ker}(A - \lambda_i I)^{k_i}.$$

Ceci découle de la relation de Bezout. Au lieu de raisonner avec χ_A prenons un polynôme $P(X)$ quelconque tel que $P(A) = 0$ et supposons que $P = Q_1 Q_2$ où Q_1 et Q_2 sont premiers entre eux.

Alors $E = E_1 \oplus E_2$ où $E_i = \text{Ker } Q_i(A)$.

(Vous avez noté qu'il faudrait mettre "u" à la place de "A" dans les divers "Ker" si on veut être puriste).

Comme Q_1 et Q_2 sont premiers entre eux, on a

$$1 = RQ_1 + SQ_2 \quad \text{pour deux polynômes } R, S$$

D'où $I = R(A)Q_1(A) + S(A)Q_2(A)$
 ou bien $\text{id}_E = R(u) \circ Q_1(u) + S(u) \circ Q_2(u)$
 si vous préférez.

Soit $x \in E$ quelconque ; la ligne qui précède donne :

$$x = Q_1(u)(R(u) \cdot x) + Q_2(u)(S(u) \cdot x) .$$

(on a écrit les produits RQ_1 et SQ_2 dans l'autre sens).

$$x = Q_1(u)(y_1) + Q_2(u)(y_2)$$

en posant $y_1 = R(u) \cdot x$ et $y_2 = S(u) \cdot x$

$$x = x_2 + x_1$$

en posant $x_2 = Q_2(u)(y_2)$, $x_1 \dots$.

or $x_2 \in E_2$ car

$$Q_2(u)(x_2) = Q_2(u)(Q_1(u)(y_1)) = P(u)y_1 = 0 \quad \text{car } P(u) = 0 .$$

De même $x_1 \in E_1$ c.-à-d. $E = E_1 + E_2$.

Il reste à montrer que cette somme est directe : si $x \in E_1 \cap E_2$, montrons que $x = 0$.

Nous avons toujours la première égalité de départ :

$$x = R(u)(Q_1(u)(x)) + S(u)(Q_2(u)(x))$$

Or, cette fois-ci, x n'est plus quelconque, mais $x \in E_1$ c.-à-d. $Q_1(u) \cdot x = 0$ et $x \in E_2$ c.-à-d. $Q_2(u) \cdot x = 0$. Ce qui donne par addition $x = 0$.

Le résultat que nous venons de démontrer s'étend par récurrence :

Si $P(A) = 0$ et si $P = P_1 P_2 \dots P_s$ où les polynômes P_i sont premiers entre eux deux à deux, alors $E = E_1 \oplus \dots \oplus E_s$ où $E_i = \text{Ker } P_i(A)$.

Si le corps K est le corps des complexes, ou, si le polynôme caractéristique a toutes ses racines dans K , alors nous pouvons prendre pour P , $P = \chi_A$ et $P_i(x) = (X - \lambda_i)^{k_i}$ et nous avons démontré le résultat annoncé page T11.

Mais, dans la voie de contournement n°1, nous ne savons toujours pas si $\chi_A(A) = 0$.

Par contre, si nous remplaçons le χ_A , par P_A , le polynôme minimal de A , et si

$$P_A(X) = \prod_{i=1}^t (X - \mu_i)^{l_i} \quad \text{alors } E = E''_{\mu_1} \oplus \dots \oplus E''_{\mu_t}$$

où $E''_{\mu_i} = \text{Ker}(A - \mu_i)^{l_i}$.

Mais comme de toute façon $\chi_A(A) = 0$ (qu'on le veuille ou non), nous sommes donc confrontés à deux décompositions, l'une relative à χ_A et l'autre à P_A .

Comparons ces deux décompositions. Pour cela, rappelons-nous : P_A divise χ_A . Donc toute racine de P_A est une racine de χ_A .

On peut donc écrire :

$$\chi_A(X) = \prod_{i=1}^r (X - \lambda_i)^{k_i} \quad (\lambda_i \text{ distinctes}).$$

$$P_A(X) = \prod_{i=1}^r (X - \lambda_i)^{l_i} \quad l_i \leq k_i .$$

Remarquez que nous avons remplacé "t" (p. T11, 11ème ligne du bas) par r , le même pour les deux produits. On peut se permettre ceci si on admet $l_i \geq 0$ (l_i peut être nul), car nous

n'avons pas encore vu pourquoi χ_A et P_A ont les mêmes racines : à priori, P_A pourrait en avoir moins (de racines distinctes).

Les deux décompositions s'écrivent :

$$E = E'_{\lambda_1} \oplus \dots \oplus E'_{\lambda_r} \quad E'_{\lambda_i} = \text{Ker}(u - \lambda_i)^{k_i}.$$

$$E = E''_{\lambda_1} \oplus \dots \oplus E''_{\lambda_r} = \text{Ker}(u - \lambda_i)^{l_i}.$$

Remarquons d'abord les choses simples :

- Comme on tolère $l_i = 0$ (mais provisoirement, car on sait qu'on va démontrer que $l_i \neq 0 \ \forall i$) on peut donc avoir a priori $E''_{\lambda_i} = 0$.

- Comme $l_i \leq k_i$, on a $E''_{\lambda_i} \subset E'_{\lambda_i}$

Notons d''_i la dimension de E''_{λ_i} et par d'_i celle de E'_{λ_i} . Nous avons donc $d''_i \leq d'_i \ \forall i$.

Comme les sommes sont directes, on peut écrire :

$$\dim E = n = d'_1 + d'_2 + \dots + d'_r$$

et
$$\dim E = n = d''_1 + d''_2 + \dots + d''_r.$$

- Il en résulte nécessairement $d'_i = d''_i$. Combien vaut ce nombre ? (ex)

- Comme on tolère a priori $l_i = 0$ c.-à-d. $E''_{\lambda_i} = 0$ ou bien $d''_i = 0$, on doit donc dire que d'_i et d''_i sont mis en même temps. Or, d'_i ne peut être nul. Pourquoi ? Une fois que vous aurez répondu à cette question vous pourrez dire : Toute racine du polynôme caractéristique et *reciproquement*, toute racine de χ_A est racine de P_A (en admettant bien sûr CH).

Tout ceci nous fait tourner autour du polynôme caractéristique, mais ne nous fait pas encore comprendre pourquoi CH est vrai. C'est pourquoi, à partir de maintenant, n'utilisons plus $\chi_A(A) = 0$ comme nous l'avons fait jusqu'ici en essayant de comprendre quelque chose.

Pour que nous soyons bien d'accord, recentrons le propos : nous essayons par une voie contournée d'arriver à CH. L'idée qui nous guide est l'idée de décomposition de l'espace E liée à la notion de vecteur propre. Cette idée généralisée nous amène au résultat page T15. *Oublions pour le moment tout ce qui suit cette proposition et écrivons un autre texte à la place* (on se souviendra bien sûr des pages 11, 12 quand on aura entrevu CH). *Le texte p. T 11 bis* court maintenant de la façon suivante :

Prenons un polynôme P le polynôme minimal : $P = P_A$. Par définition de ce polynôme $P_A(A) = 0$.

Si on a $P_A(X) = P_1 \dots P_t$ où les P_i sont premiers entre eux deux à deux, alors $E = E_1 \oplus \dots \oplus E_t$ où $E_i = \text{Ker } P_i(u)$

Si le corps K est tel que P_A a toutes ses racines dans K , alors P_i est de la forme $P_i(X) = (X - \mu_i)^{l_i}$ ($l_i \neq 0$, μ_i distincts).

Dans ce cas, ($P = P_A$), nous pouvons donc quelque chose de plus, au sujet des P_i :

Les sous-espaces E_i sont stables par u : $u(E_i) \subset E_i$ (voir ex 14). On peut donc considérer la restriction de u à E_i , soit $u_i : E_i \longrightarrow E_i$.

Alors, P_i est le polynôme minimal de u_i car par définition de E_i on a $P_i(u_i) = 0$; le polynôme minimal de u_i doit donc diviser P_i . Appelons Q_i ce polynôme minimal. Un peu d'effort va vous montrer que le polynôme $Q = Q_1 Q_2 \dots Q_t$ annule u (voir ex 14). Par construction, on a " Q divise P " comme chaque Q_i divise P_i ce qui exige $P = Q$ (à un détail près : voir ex 14).

Interprétons ce renseignement à la lumière de $P_i = (X - \mu_i)^{l_i}$.

On peut écrire $u_i = (u_i - \mu_i \text{id}_{E_i}) + \mu_i \text{id}_{E_i}$.

Posons $v_i = u_i - \mu_i \text{id}_{E_i}$ c.-à-d. $u_i = v_i + \mu_i \text{id}_{E_i}$

alors $P_i(u_i) = 0$ signifie que v_i est nilpotent d'ordre l_i . (ex 14).

Le devoir n° 1 (voir à la fin du fascicule) nous montre alors que la matrice de u est constituée par des blocs diagonaux, chaque bloc étant la matrice relative à u_i .

Cette dernière matrice s'obtient en prenant $\mu_i \times$ la matrice unité et en rajoutant la matrice de v_i comme dans le devoir.

Les blocs sont donc de la forme :

$$\begin{pmatrix} \mu & 0 & 0 \\ 1 & \mu & 0 \\ 0 & 1 & \mu \end{pmatrix}$$

Une telle matrice est une "matrice de Jordan" dont nous avons déjà calculé l'exponentielle. Une matrice quelconque ne peut pas être diagonalisée, mais elle peut toujours être mise sous forme de blocs de matrices de Jordan (réduction de Jordan). Cette voie nous mène aussi à CH car il suffit de tester CH pour une telle matrice (voir ex 14 5).

La réduction de Jordan généralise donc la réduction à la forme diagonale. Quelle est l'intérêt de cette réduction ? La forme diagonale est commode dans beaucoup de calculs, en particulier pour le calcul de e^{ta} :

Si a est une matrice diagonale en blocs (les blocs n'étant pas de même dimension), on calcule e^{ta} en calculant cette expression séparément pour chaque bloc. (exo).

- Si les blocs sont du type "Jordan", on sait calculer son exponentielle.

En exercice vous verrez un de ces calculs. Cette méthode est lourde.

Des gens astucieux ont trouvé de meilleures méthodes pour calculer e^{ta} .

Nous y reviendrons.

Pourquoi calculer e^{ta} ? Pour le déplaisir de faire un exercice ? Si vous voulez.

Sinon, rappelez vous que e^{ta} nous permet de résoudre un système d'équations différentielles linéaires.

Nous sommes donc arrivés au même point que dans la voie 2 : CH est démontré en faisant un test sur une matrice de type particulier. Mais, on supposait toujours que les racines du polynôme caractéristique étaient toutes contenues dans K ce qui est le cas si $K = \mathbb{C}$.

Nous n'avons pas avancé dans la déduction du théorème CH dans la situation : $K =$ anneau quelconque. Or, une compréhension valable de CH ne peut pas éviter ce cas général, car comme nous l'avons senti, une machine à calculs (un ordinateur) est automatiquement placé dans ce cas.

Alors, sommes-nous moins forts qu'un ordinateur ?

Non, car les ordinateurs actuels sont des bêtes qui vous sortent "Beuh = 0". CH est bien de ce type là. Mais pourquoi Beuh "doit être" = 0, cela l'ordinateur ne vous le dit pas. Or, c'est uniquement cela qui nous intéresse.

Alors, attaquons le problème.

Si K est un anneau quelconque, une bonne connaissance de "dét" vous donne CH en quelques lignes. Mais, ce n'est, je crois, pas la peine de vous le rappeler, le "dét", on l'évite un max. Mais essayons néanmoins d'en dire le plus possible sur la forme générale de :

$$\chi_A(X) = \det(A - XI).$$

Pour bien comprendre ce qui suit, reprenez l'exercice n° 1 du tout début, mais faites-le pour une matrice 3×3 .

$$A = \begin{pmatrix} a & b & c \\ a' & b' & c' \\ a'' & b'' & c'' \end{pmatrix}$$

De nouveau, pour bien marquer la généralité de A , introduisez des lettres indéterminées :

$$A = \begin{pmatrix} x_{1,1} & x_{1,2} & x_{1,3} \\ \dots & \dots & \dots \\ x_{3,1} & x_{3,2} & x_{3,3} \end{pmatrix}$$

La vérification de CH se fait en deux étapes :

- on calcule $\chi_A(X)$.

Ce polynôme est de degré 3 (de degré n si A est $n \times n$). Le coefficient de X^n est ± 1 , le coefficient de $\pm X^{n-1}$ est $x_{1,1} + x_{2,2} + \dots + x_{n,n}$.

Ce nombre (élément de l'anneau K de base car les $x_{ij} \in K$) s'obtient à partir du polynôme à plusieurs variables (à n^2 variables, mais seulement n de ces variables y figurent effectivement) :

$$P_1(X_{ij}) = X_{1,1} + X_{2,2} + \dots + X_{n,n},$$

en y injectant $x_{i,i}$ à la place de $X_{i,i}$.

Les autres coefficients s'obtiennent de la même façon : il faut injecter dans des polynômes :

$P_k(X_{ij})$ les valeurs x_{ij} aux indéterminées X_{ij} .

La remarque fondamentale est : ces polynômes sont les mêmes :

- 1) quelque soit la matrice A considérée ;
- 2) quelque soit l'anneau de base K .

Ce qui dépend de A ou de K ce sont les *valeurs* de ces polynômes. Ensuite : les coefficients de ces polynômes sont des entiers.

Pour vérifier CH, on remplace maintenant X par A dans $\chi_A(X)$. (faites le pour $n = 3$).

Pour exécuter cela, on commence par calculer A^2 , puis A^3 .

Si vous faites le calcul, vous vous rendrez compte que les coefficients de A^2 s'obtiennent en injectant les x_{ij} dans certains polynômes à coefficients entiers, polynômes qui sont indépendants de A et de K .

De même pour A^3 etc.

finalement, les coefficients de la matrice $\chi_A(A)$ s'obtiennent de la même façon : substituer des $x_{i,j}$ aux $X_{i,j}$ dans des polynômes $Y_{i,j}$ à coefficients entiers, c.-à-d. à coefficients dans l'anneau $\mathbb{Z}$.

Ce que dit CH, c'est *qu'après* substitution des x_{ij} aux indéterminées $X_{i,j}$ dans les polynômes $Y_{i,j}$, on trouve 0, c.-à-d. le miracle s'accomplit et ce, quelque soit l'anneau de base K choisi (il y en a des millions !). Ce qui est plus probable, c'est que les polynômes $Y_{i,j}$ eux-mêmes sont nuls, car alors automatiquement n'importe quelle substitution donne 0.

Mais comment montrer qu'un polynôme est nul ? Il faut montrer que tous ses coefficients sont nuls. Ici, les polynômes $Y_{i,j}$ considérés sont à coefficients dans $\mathbb{Z}$.

Prenons un polynôme :

$$P(x) = a_0 + a_1 X + \dots + a_n X^n$$

à une indéterminée à coefficients dans $\mathbb{Z}$. Si ce polynôme s'annule, si on substitue à X n'importe quel entier $\in \mathbb{Z}$, alors ce polynôme est identiquement nul :

$$(P(x) = 0 \quad \forall x \in \mathbb{Z}) \Rightarrow P(X) = 0.$$

(Car $P(X)$ peut être considéré comme un polynôme sur $\mathbb{R}$ et vous savez qu'un polynôme non nul sur $\mathbb{R}$ ne peut avoir qu'un nombre fini de racines).

Ce qui est vrai pour les polynômes à une variable est encore vrai pour les polynômes à plusieurs indéterminées, donc pour les $Y_{i,j}$.

C.-à-d., il faut que nous montrions que $Y_{i,j}(x_{i,j})$ pour n'importe quelle substitution $X_{i,j} \longrightarrow x_{i,j}$ où $x_{i,j} \in \mathbb{Z}$, ce qui veut dire, il faut que nous montrions que $\chi_A(A) = 0$ quand la matrice A est à coefficients dans $\mathbb{Z}$ et nos raisonnements montrent alors que $\chi_A(A) = 0$ pour A à coefficients dans n'importe quel anneau K !!

Or, une matrice à coefficients dans $\mathbb{Z}$ est un cas particulier de matrices à coefficients dans $\mathbb{C}$! Et pour $K = \mathbb{C}$, nous avons la certitude totale que CH est vraie c.-à-d. $\chi_A(A) = 0$: calculer $\chi_A(X)$ "sur $\mathbb{Z}$ " ou "sur $\mathbb{C}$ ", c'est kif-kif.

Voilà, on est sorti d'affaire !

Le raisonnement précédent est difficile, si cela peut vous rassurer, pour un "débutant". Le matheux professionnel le fait aussi rapidement que $3 \times 13 = 39$. C'est une question d'habitude !

La démarche du raisonnement précédent est en résumé : Pour vérifier CH pour tout anneau K et il suffit de le vérifier pour l'anneau $\mathbb{Z}$. Et pour l'anneau $\mathbb{Z}$ il l'est vérifié s'il l'est pour $\mathbb{C}$.

On peut mettre un peu "d'ordre" dans cette "variation" de l'anneau de base K . Cet ordre, on le trouve en analysant les lignes 9 et 10 du bas de la page T9. Voici comment :

Soient K et K' deux anneaux (commutatifs) et soit $\sigma : K \longrightarrow K'$ un homomorphisme (on suppose toujours qu'un homomorphisme échange les éléments unités ($\sigma(1) = 1'$)). On dit alors que K et K' sont "comparables". On considère deux cas particuliers :

- $\sigma : K \longrightarrow K'$ est injectif (on dit que K "précède" K')
- $\sigma : K \longrightarrow K'$ est surjectif (K' suit K).

Le premier cas est réalisé par $K = \mathbb{Z}$, $K' = \mathbb{C}$, σ l'inclusion $\mathbb{Z} \subset \mathbb{C}$.

On verra bientôt un exemple du deuxième type.

Nous allons voir la vérité de :

- Si σ est injectif, si CH est vrai pour K' , alors CH est vrai pour K .
- Si σ est surjectif, si CH est vrai pour K , alors CH est vrai pour K' .

Le premier cas généralise donc tout simplement la remarque un peu plus haut (l 12). Les quelques déductions que nous allons faire vous montrent aussi l'intérêt du concept "d'homomorphisme".

Voici ce que déclenche la connaissance d'un homomorphisme $\sigma : K \longrightarrow K'$ quelconque. Il permet de définir un homomorphisme encore noté σ de $K[X]$ dans $K'[X]$ défini de la façon suivante :

$$\sigma.P(X) = \sigma \left(\sum_{i=1}^n a_i X^i \right) = \sum_{i=1}^n a'_i X^i \quad \text{avec} \quad a'_i = \sigma(a_i).$$

De même, σ définit un homomorphisme encore noté σ de $\mathcal{M}_{n \times n}(\mathbf{K})$ dans $\mathcal{M}_{n \times n}(\mathbf{K}')$. En posant :

$$\begin{aligned} \sigma(a_{i,j})_{n \times n} &= (a'_{i,j})_{n \times n} \text{ avec } a'_{ij} = \sigma(a_{i,j}) \\ \text{et } \sigma(aA) &= \sigma(a) \sigma(A) \quad a \in \mathbf{K}, A \in \mathcal{M}_{n \times n}(\mathbf{K}). \end{aligned}$$

On a alors les formules :

$$\begin{aligned} \sigma(P(A)) &= (\sigma.P)(\sigma.A) \\ \sigma.\det B &= \det(\sigma.B). \end{aligned}$$

Ne vous inquiétez pas pour ces formules, vous les comprendrez facilement en faisant les quelques exercices qui s'y réfèrent.

Montrons maintenant la vérité du premier point : (p. T16)

Soit A dans $\mathcal{M}_{n \times n}(\mathbf{K})$ et soit $\chi_A(X)$ son polynôme caractéristique.

Posons $A' = \sigma A$.

Alors $\chi_{A'}(A') = 0$, où $\chi_{A'}(X)$ est le polynôme caractéristique de A' (on suppose que sur $\mathbf{K}'$ CH est vrai).

$$\begin{aligned} \text{Or, } \chi_{A'}(X) &= \det(A' - XI') &= \det(\sigma.A - XI') \\ &= \det(\sigma(A - XI)) &= \sigma.\det(A - XI) \\ &= \sigma.\chi_A(X) \end{aligned}$$

Donc formule : $\chi_{A'}(X) = \sigma.\chi_A(X)$ ou bien $\chi_{A'} = \sigma.\chi_A$.

Calculons : $\chi_{A'}(A') = \chi_{A'}(\sigma.A) = (\sigma.\chi_A)(\sigma.A) = \sigma.\chi_A(A)$.

Nous avons donc $\sigma.\chi_A(A) = 0$. Ceci entraîne $\chi_A(A) = 0$ comme σ est injectif. Le premier point est donc vrai. En fait, nous avons démontré que si CH est vrai pour la matrice $A' = \sigma.A$, alors CH est vrai pour la matrice A . Si les formules précédentes vous posent problème, passez provisoirement outre. Ces pages sont en fait un exercice développé.

Passons à la vérité du deuxième point : les formules générales de cette page sont bien sûr toujours valables.

Cette fois-ci, $\sigma = \mathbf{K} \longrightarrow \mathbf{K}'$ est surjectif et on *part* avec une matrice A' sur $\mathbf{K}'$. Mais comme σ est surjectif, on peut écrire $A' = \sigma.A$ où A est sur $\mathbf{K}$.

comme CH est vrai sur $\mathbf{K}$, nous avons cette fois $\chi_A(A) = 0$ et il faut en déduire $\chi_{A'}(A') = 0$. Il suffit pour cela de relire la formule au milieu de cette page.

Pourquoi cette remarque en long et en large ? Pas pour son importance intrinsèque. Sur les dernières pages, nous avons en fait traité un exercice d'écriture.

Néanmoins, cette gymnastique nous permet maintenant de formaliser une remarque faite beaucoup plus haut : p. T 9.

On avait vu "qu'en général" une matrice est diagonalisable donc que CH est vrai "en général" et qu'il *suffit* de le vérifier uniquement pour les matrices diagonales.

Nous allons, dans les pages qui suivent, rendre cette intuition formelle.

Mais, jusqu'où va-t-on continuer comme cela ? Vous en avez peut-être ralus bolus de Messieurs Cayley et Hamilton ? Alors, arrêtez-vous là !

Si le coeur vous en dit, restez encore deux à trois pages avec moi. Ensuite, la première partie de votre unité sera achevée. Je recopierai à la fin pour vous, comme curiosité, une démonstration de CH que ferait un professionnel du déterminant.

Allons-y pour une dernière promenade avant la mi-temps. Comme ce qui m'importe, ce n'est pas de "démontrer" CH, mais de vous faire sentir une idée mathématique, nous allons nous restreindre aux matrices 2×2 . L'anneau $\mathbf{K}$ de base est $\mathbb{C}$. Car nous y sommes ramenés. Nous savons que pour une matrice diagonale CH est vrai et nous voulons montrer que ce test suffit pour régler le sort d'une matrice quelconque.

La forme générale d'une matrice 2×2 est $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$.

Pour bien montrer que les constantes a, b, c, d sont quelconques, on peut préférer écrire

$$A = \begin{pmatrix} x & y \\ z & t \end{pmatrix} \quad \text{où donc } x, y, z, t \in \mathbb{C}.$$

En général, pour des x, y, z, t quelconques, il n'y a pas de lien entre ces nombres.

Pour exprimer mathématiquement cette situation général, on introduit maintenant les indéterminées X, Y, Z, T et la matrice

$$\underline{A} = \begin{pmatrix} X & Y \\ Z & T \end{pmatrix}.$$

Cette matrice $\underline{A}$ n'est absolument pas la même que la matrice A ! Tout simplement parce que A est à coefficients dans $\mathbb{C}$ et que $\underline{A}$ est à coefficients dans l'anneau des polynômes à quatre indéterminées et à coefficients dans $\mathbb{C}$, c.-à-d. à coefficients dans

$$\mathbb{C}[X, Y, Z, T] = \mathcal{A}.$$

La matrice $\underline{A}$ a ses valeurs propres distinctes, elle est donc diagonalisable. Là, vous devriez me dire "hue-là" ! D'abord, les coefficients de $\underline{A}$ sont dans un anneau $\mathcal{A}$! Or, pour diagonaliser, on écrit $S^{-1} \underline{A} S$ et donc S^{-1} ; on a besoin de prendre des inverses dans $\mathcal{A}$, ce qui n'est pas toujours possible. Pour inverser une matrice, il nous faut travailler sur un corps et $\mathcal{A}$ n'en est pas un.

Ensuite, vous me direz : "J'ai calculé le polynôme caractéristique "en λ " ce qui donne :

$$\lambda^2 - (X + T)\lambda + XT - YZ = 0$$

et ensuite j'ai été très ennuyé, car je ne savais pas où trouver des λ qui vérifient une telle relation ! Dans $\mathcal{A}$, je suis quasiment sûr qu'il n'y en a pas".

Voici comment on contourne ces difficultés :

1) $\mathcal{A}$ n'est pas un corps ? $\mathbb{Z}$ non plus n'en était pas. Mais à partir de $\mathbb{Z}$, on peut bricoler $\mathbb{Q}$ le corps des "fractions". Un élément de $\mathbb{Q}$ s'écrit p/q où $p, q \in \mathbb{Z}$.

Cette écriture n'est pas unique : on a $p/q = p'/q'$ si $pq' = p'q$. Malgré cette ambiguïté, on peut définir une addition et une multiplication par :

$$p/q + p'/q' = (pq' + p'q)/qq'; \quad p/q \cdot p'/q' = pp'/qq' \quad \text{et } \mathbb{Q} \text{ devient ainsi un corps.}$$

On peut de même bricoler à partir de $\mathcal{A} = \mathbb{C}[X, Y, Z, T]$ un corps noté $\mathbb{C}(X, Y, Z, T)$, appelé "corps des fraction rationnelles" à quatre indéterminées. Le procédé est le même.

2) Le "où" de λ s'est déjà posé au sujet de $\lambda^2 + 1 = 0$. Si vous ne connaissez que $\mathbb{R}$, il n'y a pas de place pour un tel λ . C'est pourquoi on a construit $\mathbb{C}$. On démontre, mais c'est

une autre et longue histoire, qu'on peut faire cela avec n'importe quel corps. Donc les λ , il y a en quelque part et ils sont distincts car le discriminant est $\neq 0$.

Donc, CH est vrai sur $\mathcal{A} = \mathbb{C}[X, Y, Z, T]$ pour les matrices diagonalisables, et en particulier pour la matrice $\underline{A}$!

On va maintenant se ramener aux matrices $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ sur $\mathbb{C}$.

On a un homomorphisme surjectif : $\sigma : \mathbb{C}[X, Y, Z, T] \longrightarrow \mathbb{C}$ défini par $\sigma(X) = a$, $\sigma(Y) = b$, $\sigma(Z) = c$, $\sigma(T) = d$ et $\sigma(\alpha) = \alpha$ si $\alpha \in \mathbb{C}$.

(il y a quelques détails à justifier pour vous-même).

On a ainsi
$$\sigma \begin{pmatrix} X & Y \\ Z & T \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Nous savons d'après le point 2 p. T 22 que CH est vrai sur $\mathbb{C}$ sachant qu'il est vrai sur $\mathbb{C}[X, Y, Z, T]$.

Voici maintenant le texte de professionnel.

$$f(u) = \det(a_i^j - u\delta_i^j) = \begin{vmatrix} a^1_{1-u} & a^1_2 & \dots \\ a^2_1 & a^2_{2-u} & \dots \\ \dots & \dots & \dots \end{vmatrix}$$

appelé polynôme caractéristique de la matrice A .

On ne peut pas conclure immédiatement de là que $m(u)$ divise $f(u)$, mais c'est ce qui a lieu, d'après le résultat suivant (Hamilton, Cayley) :

THEOREME : Toute matrice vérifie son équation caractéristique.

Appelons b_i^j le cofacteur de $(a_i^j - u\delta_i^j)$ dans $f(u)$, de sorte que

$$\sum_j (a^k_j - u\delta^k_j) b_i^j = \sum_j (a_i^j - u\delta_i^j) b^k_j = \delta^k_i f(u),$$

Ce qui signifie que, posant $B = (b_i^j)$, on a :

$$B(A - uI) = (A - uI)B = (-1)^n u^n I + \dots = f(u)I.$$

B est un polynôme de degré $(n-1)$ en u , dont les coefficients sont des matrices, soit :

$$B = M_0 + M_1 u + \dots + (-u)^{n-1} I.$$

les égalités ci-dessus donnent

$$AM_0 + (AM_1 - M_0)u + \dots = M_0A + (M_1A - M_0)u + \dots = f(u)I,$$

dont les matrices

$$AM_0 = M_0A, \quad AM_1 - M_0 = M_1A - M_0, \quad \dots$$

valent I à un facteur près. En remplaçant maintenant u par A , il vient

$$f(A) = (A - A)B = B(A - A) = 0$$

comme on l'avait annoncé.

et je ne vous dénoncerai pas si vous n'y avez rien compris !

A la semaine prochaine

Bonsoir !



J.N.Y

1.1 - Vérifier à la main ce que dit le théorème de Cayley-Hamilton pour les matrices suivantes :

$$\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} a & 0 \\ c & d \end{pmatrix}, \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

1.2 - Prendre la matrice $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ et montrer que la "fausse démonstration", p.T1, est vraiment fausse.

2.1 - Rechercher dans un cours quelconque l'idée et la définition "Anneau d'entiers modulo n ", anneau qui est noté par $\mathbb{Z}/n\mathbb{Z}$ ou $\mathbb{Z}/(n)$ etc

Par exemple $\mathbb{Z}/8\mathbb{Z} = \{0, \underline{1}, \underline{2}, \dots, \underline{7}\}$ où r est "la classe modulo 8 du nombre entier r ". Chercher les éléments inversibles.

Même question pour $\mathbb{Z}/7\mathbb{Z}$.

.2 - Rappelez-vous le théorème de Bezout. En déduire que $\mathbb{Z}/p\mathbb{Z}$ est un corps si p est un nombre premier.

Montrer en général que m est inversible dans $\mathbb{Z}/n\mathbb{Z}$ si n et m sont premier entre eux.

.3 - Soit p un entier premier. Soit $x \in \mathbb{Z}/p\mathbb{Z}$, montrez que $x^p - x = 0$.

.4 - On prend $\mathcal{A} = \mathbb{Z}/6\mathbb{Z}$ et la matrice $A = \begin{pmatrix} \underline{1} & \underline{1} \\ \underline{1} & \underline{1} \end{pmatrix}$.

Vérifier le théorème de CH sur A .

.5 - On prend l'anneau $\mathcal{A}$ des polynômes à une indéterminée : $\mathcal{A} = K[X]$ (K est un corps) et on considère la matrice à coefficients dans $\mathcal{A}$:

$$A = \begin{pmatrix} X & 1+X \\ -X^2 & X \end{pmatrix}$$

Calculer le déterminant de A . De quelle nature est le résultat de votre calcul ?

Complétez la phrase :

Si A est une matrice dans un anneau $\mathcal{A}$ (commutatif avec élément unité), alors $\det A$ est un élément de

On complique : Au lieu de prendre $K = \mathbb{C}$ ou $K = \mathbb{R}$, ce qui serait moral et bon, on prend $K = \mathbb{Z}/2\mathbb{Z}$. (on note toujours 1 l'élément unité, c.-à-d. $\underline{1} = 1$).

Calculer A^2 , $(\det A)^2$ et $\det(A^2)$. A quelle relation générale cela vous fait penser ?

.6 - Le "on complique" précédent vient du fait qu'on considère des polynômes dont les coefficients sont dans un domaine plus général (et plus particulier) que ce dont on a l'habitude. Comme pour les matrices, on peut calculer de façon ordinaire avec les polynômes dont les coefficients sont seulement dans un anneau $\mathcal{A}$ au lieu d'être dans un corps.

Exemple : on prend $\mathcal{A} = \mathbb{Z}/4\mathbb{Z}$.

On pose $P_1 = X^2 - 3X$; $P_2 = X^2 + X + 1$

Calculer $P_1 + P_2$.

Question : remarquer que dans $\mathbb{Z}/4\mathbb{Z}$, on a $\underline{4} = \underline{0}$ et avec un peu d'habitude, on écrit $\underline{0}$ au lieu de $\underline{0}$; $\underline{1}$ au lieu de $\underline{1}$, $\underline{2}$ au lieu de $\underline{2}$ etc.... Mais on n'écrit pas pour autant :

$$X^4 = X^0 = 1.$$

Pourquoi ?

.7 - Compliquons encore : on prend $\mathcal{A} = K[Y]$ et on considère les polynômes en X à coefficients dans $\mathcal{A}$. L'anneau de ces polynômes se note $\mathcal{A}[X]$ c.-à-d. $K[Y][X]$.

On pose $P_1 = 1 + YX + Y^2 X^2$

$$P_2 = 1 + Y^2 X^2 + YX$$

Quels sont les coefficients de P_1 , de P_2 ?

Calculer $P_1 + P_2$ et $P_1 P_2$. Donner les coefficients de ces polynômes.

Donner une justification intuitive (donc fondamentale) de la ligne suivante :

$$K[Y][X] = K[X][Y] = K[X, Y] \text{ (où l'un des termes est à définir).}$$

.8 - $K = \mathbb{Z} / 3 \mathbb{Z}$. Pourquoi le polynôme (en haut de la page T3) $P_1 \neq 0$?

Montrer que la fonction polynôme produite par P_1 est la fonction partout nulle.

3 - Choisissez une matrice 2×2 , A , particulière.

Former la matrice $A - XI$. Ecrire sur une ligne les quatre coefficients de cette matrice. Essayez de comprendre la phrase suivante :

«Une matrice à coefficients dans K peut toujours être considérée comme étant à coefficients dans $K[X]$ en associant à $a \in K$, le polynôme constant correspondant, de sorte que la matrice $A - XI$ est une différence de matrices de même type».

4.1 - Introduisez deux opérateurs pour d'écrire la formule $(a + b)^2 = a^2 + b^2$.

Cette formule est-elle fausse si $a, b \in K = \mathbb{Z} / 2 \mathbb{Z}$?

.2 - Montrer que $(a + b)^p = a^p + b^p$ si $a, b \in \mathbb{Z} / p \mathbb{Z}$ où p est un entier premier.

.3 - Soit K un anneau (commutatif avec un élément unité) sous anneau d'un anneau $\mathcal{A}$ (commutatif etc). Soit $a \in \mathcal{A}$ un élément fixe de $\mathcal{A}$.

Montrer qu'il existe un homomorphisme d'anneaux et un seul

$$\sigma : K[X] \longrightarrow \mathcal{A}$$

tel que

$$\sigma(X) = a.$$

6. - Une matrice $(a_{i,j})_{n \times n}$ peut être considérée comme un point de $\mathbb{R}^{n^2}$. La notion de "valeur absolue" à laquelle nous faisons allusion est donc celle de "norme dans $\mathbb{R}^{n^2}$ ". Il faut donc regarder dans un livre d'analyse la définition de "normes équivalentes". Parmi les normes que vous trouverez dans ce livre, essayez d'en trouver une qui nous conviendrait.

7.1 - Soit $z \in \mathbb{C}$. Est-ce que e^z peut s'annuler ? Quel est l'inverse de e^z .

Chercher de même à montrer que e^a est inversible et trouver son inverse quand $a \in \mathbb{R}^{n^2}$.

.2- Montrez que si la formule $\frac{d}{dt}(e^{ta}) = a e^{ta}$ est vraie pour $t = 0$, alors elle est vraie pour

tout t .

Montrez ensuite que la fonction est dérivable en 0 en ouvrant un bouquin d'analyse qui traite du cas de e^x ($x \in \mathbb{R}$).

.3 - Montrez que les matrices e^{ta} et a commutent.

Commencez d'abord par montrer que $P(a)$ et a commutent avec $P \in \mathbb{K}[X]$. Essayez ensuite de justifier le "passage à l'infini".

.4 - La notation $a.x$ signifie qu'on considère la base canonique (e_i) de $\mathbb{R}^n$ et la matrice a définit alors une application linéaire de $\mathbb{R}^n$ dans $\mathbb{R}^n$ qu'on note toujours a et on note alors $a(x) = a.x$.

Montrer que $x' = a.x$ est une écriture condensée pour un système de n équations différentielles du premier ordre.

.5 - Considérer une équation différentielle ordinaire du second ordre :

$$x'' = ax' + bx \quad a, b \in \mathbb{R} \quad x \in \mathbb{R}.$$

Elargir le sens de cette écriture au cas où $x \in \mathbb{R}^n$.

Montrer qu'une telle équation est équivalente à une équation du premier ordre " $y' = \alpha y$ " en considérant comme variables "position" et "vitesse".

.6 - Calculez e^a où $a = \begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix}$.

Généralisez au cas où a est diagonale $n \times n$.

.7 - Calculer e^a où a est la matrice carrée $n \times n$ avec λ sur la diagonale principale, 1 sur la diagonale en dessous, 0 partout ailleurs.

$$a = \begin{pmatrix} \lambda & 0 & \dots & 0 \\ 1 & \lambda & \dots & \dots \\ \dots & \dots & \dots & 0 \\ 0 & \dots & 1 & \lambda \end{pmatrix}$$

Commencez par regarder les cas $n = 2$ et $n = 3$. Puis passez au cas général. Que pensez-vous de considérer la matrice $a - \lambda I$?

Une telle matrice s'appelle matrice de Jordan.

.8 - Pouvez-vous donner une justification mathématique de l'intérêt de la théorie des ensembles ?

8.1 - Pouvez-vous dénicher l'erreur dans le raisonnement de la page T8 ? Si vous analysez pourquoi dans la phrase (parfaitement correcte) "R est soit le polynôme nul, soit un polynôme de degré strictement plus petit...." on distingue deux cas ; vous pouvez être amenés à la voir.

.2 - Une relation du type $P(A) = 0$ a une forte odeur de polynôme. Pourtant, en considérant l'espace vectoriel $M_{n \times n}$ des matrices carrées une telle relation peut être considérée comme une relation *linéaire*. Cette remarque vous servira pour rendre le raisonnement de la page T10 parfaitement correct.

.3 - Vérifier CH pour une matrice diagonale.

.4 - Montrer que les polynômes caractéristiques de deux matrices semblables sont égaux. En déduire que CH pour l'une entraîne CH pour l'autre.

10.1 - Tester CH pour une matrice 3×3 triangulaire.

.2 - Rappelez-vous comment on calcule le déterminant d'une matrice triangulaire $n \times n$.

.3 - Si A est la matrice de la p. T10, posez $A_i = A - \lambda_i I$.

Calculez $A_1 A_2$: quelle forme a cette matrice ?

Calculez $A_1 A_2 A_3$: forme ?

En déduire que $A_1 A_2 \dots A_n = 0$.

11.1 - Expliquez la phrase "seulement en général" à côté de la formule $E = E_{\lambda_1} + \dots + E_{\lambda_r}$.

.2 - Au lieu d'injecter A dans $P(X)$, on peut injecter u : rendez-vous compte du sens de $P(u)$. Par exemple, prenez $P = X^2 + X + 1$. Que signifie $P(u).x$ (ou $P(u)(x)$) si $x \in E$?

Réalisez les formules :

$$(PQ)(u)(x) = (PQ)(u).x = (P(u) \circ Q(u)).x = P(u)(Q(u)(x)) = P(u).(Q(u).x).$$

12.1 - Si vous en sentez le besoin, exécutez la récurrence annoncée en milieu de page.

.2 - Si $P = Q_1 \cdot Q_2$ comme p. T 11, montrez qu'on a aussi $E = E'_1 + E'_2$ où $E'_i = \text{Im } Q_i(u)$. Comparez E_i et E'_j .

13.1 - Montrer que par définition de " λ est valeur propre de u " les entiers d'_i sont tous non nuls.

.2 - Nous avons une troisième relation : $n = k_1 + k_2 + \dots + k_r$

Dites pourquoi.

Montrez que $d'_i = k_i$ pour tout i . Ceci vous paraîtra plus difficile. Aidez-vous par un livre. Ne suffirait-il pas de démontrer $d'_i \leq k_i$?

14.1 - Avec les notation de la page T 11, montrez que E_i est stable par u .

.2 - On suppose que $E = F_1 + F_2$ où chaque F_i est stable par un endomorphisme $u : E \longrightarrow E$. On appelle u_i la restriction de u à F_i .

On suppose que le polynôme R_1 annule u_1 et que R_2 annule u_2 .

Montrez que $R = R_1 R_2$ annule u .

Il vous faudra montrer que $v = R(u)$ est l'endomorphisme nul, c.-à-d. $v(x) = 0 \ \forall x \in E$.

.3 - Remarquez que si $P(X)$ est "un" polynôme minimal; $2P(X)$ en est un autre ($K = \mathbb{R}$). Comment singulariser "le" polynôme minimal de sorte que l'égalité $Q = P$ soit exacte ?

.4 - On a $v_i^{h_i} = 0$ si h_i est le plus petit exposant k tel que $v_i^k = 0$. Démontrez donc que h_i est effectivement l'ordre de v_i .

.5 - Faites le test CH qu'on vous demande de faire. Vous l'avez peut-être déjà fait quelque part ?

.6 - Soit A une matrice formée de deux blocs B_1, B_2 diagonaux :

$$A = \begin{pmatrix} B_1 & 0 \\ 0 & B_2 \end{pmatrix} \quad B_1 \text{ de type } r \times r, B_2 \text{ de type } s \times s, \text{ avec } s + r = n$$

Comment calculer e^A connaissant e^{B_i} ?

.7 - On considère la matrice ($K = \mathbb{R}$)

$$A = \begin{pmatrix} 1 & 0 & -1 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

- Calculez $\chi_A(X)$ qu'on décomposera en produits de facteurs premiers.
- Déterminer $P_A(X)$, le polynôme minimal
- Ecrire la décomposition de $E = \mathbb{R}^4$ de la page T11 relative au polynôme P_A .
- Calculer e^{tA} .

16.1 - On considère $\sigma : \mathbb{Z} \longrightarrow \mathbb{Z} / 8\mathbb{Z}$ définie par $\sigma(n) = \tilde{n}$ = Classe modulo 8 de n .
Ainsi $\sigma(24) = 0$.

On prend $P_1(X) = 1 + 2X + 3X^2$
 $P_2(X) = X + 6X^2$.

Vérifier les formules $\sigma(P_1 + P_2) = \dots$ $\sigma(P_1 P_2) = \dots$ en utilisant la notation " n " et ensuite faites le calcul en maintenant la notation $\sigma(n)$.

On prend $A = \begin{pmatrix} 1 & 2 \\ -3 & 4 \end{pmatrix}$ $B = \begin{pmatrix} 2 & 1 \\ 0 & 3 \end{pmatrix}$

Mêmes questions.

Vérifiez la formule $\sigma \cdot P_1(A) = (\sigma \cdot P_1)(\sigma A)$ en prenant les notations n et $\sigma(n)$.

Vérifier $\sigma(\det B) = \det(\sigma B)$.

.2 - On prend $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. On suppose $\sigma \cdot A = 0$ avec σ injectif.

Montrez que $A = 0$.

.3 - On prend σ comme dans 1 -, et on prend $A' = \begin{pmatrix} 3 & 1 \\ -1 & 0 \end{pmatrix}$. Quelle matrice A prendre pour que $\sigma \cdot A = A'$.

17.1 - Nous avons vu que pour définir un homomorphisme σ de $K[X]$ dans K (par exemple) tel que $\sigma(\alpha) = \alpha$ quand $\alpha \in K$, il suffit de fixer la valeur de X , soit $\sigma(X) = a$; il en résulte alors la formule $\sigma(P(X)) = P(a)$. La valeur a peut être arbitraire (mais fixée pour chaque σ). Montrez que σ est alors surjectif.

Pour définir $\sigma : K[X, Y, Z, T] \longrightarrow K$, on peut choisir arbitrairement les nombres $\sigma(X)$, $\sigma(Y)$, $\sigma(Z)$, $\sigma(T)$, car il n'y a aucune relation entre les indéterminées X, Y, Z, T , sauf les relations de commutation $XY = YX$ etc qui exigent

$$\sigma(YX) = \sigma(Y) \cdot \sigma(X) = \sigma(X) \cdot \sigma(Y) = \sigma(XY)$$

qui sont bien sûr vérifiées comme on considère systématiquement des anneaux de coefficients k qui sont commutatifs.

Il en résulte $\sigma(P(X, Y, Z, T)) = P(a, b, c, d)$ (si $\sigma(X) = a$, $\sigma(Y) = b$...)

.2 - Montrer qu'une matrice A est nilpotente, si son polynôme caractéristique est de la forme X^n . (Aider-vous d'une triangularisation si $K = \mathbb{R}$).

Avez-vous une idée pour K quelconque ?

.3 - Soit E , un espace de dimension n sur le corps K , $u = E \longrightarrow E$ un endomorphisme, λ une valeur propre de multiplicité $k (\lambda \in K)$.

Montrez que $E = E_1 + E'_1$ où $E_1 = \text{Ker}(u - \lambda.1)^k$ et où E'_1 est à définir à l'aide de la page T14 et de l'exercice E 14.2.

Remarquez que E_1 est stable par u . Quel est le polynôme caractéristique de la restriction de u à E_1 ?

.4 - Montrer qu'une matrice est diagonalisable si son polynôme minimal n'a que des racines simples.

.5 - Calcul de e^{tA}

Voici une bonne méthode (il y a encore mieux, bien sûr).

On part d'une équation diff. linéaire : (E) $x^{(n)} - a_{n-1}x^{(n-1)}, \dots, a_1x' - a_0x = 0$ où x est une fonction inconnue de t et où $a_i \in \mathbb{R}$.

Ce qu'on sait (voir un livre d'analyse) :

- L'ensemble des solutions de (E) forme un espace vectoriel E .
- Cet espace est de dimension n .
- Si on donne arbitrairement les n nombres $c_0, c_1, \dots, c_{n-1}$, alors il existe une solution et une seule $x(t)$ telle que $x(0) = c_0, x^{(1)}(0) = c_1, \dots, x^{(n-1)}(0) = c_{n-1}$.
- D'après le point précédent, on peut trouver n solutions φ_i telles que $\varphi_i^{(j)}(0) = \delta_i^j$. Ces n solutions forment une base de E .
- On sait trouver une base de solutions de la forme $e^{\alpha t} t^k$ où α est un zéro du "polynôme caractéristique de (E)" :

$$R(X) = X^n - a_{n-1}X^{n-1} - \dots - a_1X - a_0$$

et où $k < \text{multiplicité de } \alpha$. (Vous avez sûrement vu le cas $n = 2$).

Soit $(f_i)_{1 \leq i \leq n}$ une telle base

- Si f est une solution quelconque de (E), on peut décomposer f sur la base (φ_i) , on a

$$f = f(0)\varphi_0 + f'(0)\varphi_1 + \dots + f^{(n-1)}(0)\varphi_{n-1}.$$

Vérification !

Soit maintenant une matrice A $m \times m$ dont on veut calculer l'exponentielle e^{tA} . Voici la marche opératoire.

- Déterminer un polynôme de bas degré qui annule A (le polynôme minimal est celui qui convient le mieux).

Appelez $R(X)$ ce polynôme supposé écrit sous la forme :

$$R(X) = X^n - a_{n-1}X^{n-1} - \dots - a_1X - a_0.$$

- A ce polynôme associez l'équa diff (E) dont R est justement le polynôme caractéristique.
- Déterminer une base de solutions $f_1, \dots, f_n$ de (E) de la forme $e^{\alpha t} t^k$.

A partir de cette base (f_i) , déterminez la base (φ_i) en résolvant un système linéaire grâce à la relation du milieu de la page écrite pour chaque f_i .

- Vous avez ainsi la formule :

$$e^{tA} = \varphi_0(t)I + \varphi_1(t)A + \dots + \varphi_{n-1}(t)A^{n-1}$$

Joli, non !!

Cette formule vous sera expliquée dans le corrigé.

Pour le moment, appliquez la recette au cas de la matrice de l'exercice 17.7.

COMMENTAIRES SUR LES EXERCICES PROPOSES :

1.1 - Il suffit bien sûr de traiter le cas général $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ et puis injecter si on veut $a \mapsto$,
 $b \mapsto$, $c \mapsto$, $d \mapsto$.

Pour bien montrer que dans le cas général les nombres a, b, c, d sont arbitraires, changeons de notation et prenons la matrice $A = \begin{pmatrix} x_1 & x_2 \\ x_3 & x_4 \end{pmatrix}$

Le polynôme caractéristique s'écrit alors

$$\chi_A(X) = X^2 - (x_1 + x_4)X + x_1x_4 - x_2x_3 .$$

La matrice $\chi_A(A)$ s'écrit a priori

$$\chi_A(A) = \begin{pmatrix} y_1 & y_2 \\ y_3 & y_4 \end{pmatrix} .$$

Par identification, on trouve :

$$y_1 = x_1^2 + x_2x_3 - (x_1 + x_4)x_1 + x_1x_4 - x_2x_3$$

$$y_2 = x_1x_2 + x_2x_4 - (x_1 + x_4)x_2 + 0$$

$$y_3 = x_1x_3 + x_3x_4 - (x_1 + x_4)x_3 + 0$$

$$y_4 = x_2x_3 + x_4^2 - (x_1 + x_4)x_4 + x_1x_4 - x_2x_3$$

$$\text{c'est-à-dire } y_1 = y_2 = y_3 = y_4 = 0 .$$

On remarque que le calcul ne fait pas intervenir la nature des nombres a, b, c, d ou x_1, x_2, x_3, x_4 et en plus la simplification dans le calcul $y_i = 0$ ne fait intervenir que des suppressions mais aucune relation particulière qui serait due à la nature particulière des nombres avec lesquels on travaille.

1.2 - Prenons donc la matrice $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$.

L'opération consiste à remplacer λ par A dans la matrice $A - \lambda I$ et de dire qu'on obtient ainsi la matrice $A - A = 0$. Si nous faisons effectivement cette opérations dans $A' = A - \lambda I$, nous obtenons la matrice :

$$\begin{pmatrix} 1-A & 0 \\ 0 & -A \end{pmatrix} .$$

La nature de cette matrice n'est pas évidente. Il y a une façon naturelle (une seule) de rendre un sens à l'écriture précédente : Apparemment, il y a mélange de natures des coefficients : les symboles 0 et 1 sont dans K , le symbole A est dans $M_{n \times n}(K)$, anneau des matrices sur K .

Mais que signifie alors $1 - A$ par exemple ?

On fait alors la convention suivante : si $a \in K$, on associe à la matrice diagonale $\begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix} = \underline{a}$

(et donc à 0 la matrice $\underline{0} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$).

Le sens de $\begin{pmatrix} 1-A & 0 \\ 0 & -A \end{pmatrix}$ est donc : il s'agit d'une matrice 2×2 à coefficients dans l'anneau

$$M_{n \times n}(K) \text{ égale à } \begin{pmatrix} 1-A & 0 \\ 0 & -A \end{pmatrix} = \begin{pmatrix} \alpha & \beta \\ \delta & \varphi \end{pmatrix}$$

$$\text{avec } \alpha = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \quad \beta = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \quad \delta = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \quad \varphi = \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix}$$

Or les coefficients α et φ sont non nuls : la matrice " $A - A$ " n'est pas nulle !

2.1 - Éléments inversibles de $\mathbb{Z}/8\mathbb{Z}$:

$$\begin{array}{ll} \underline{1} & \underline{1.1} = \underline{1} \\ \underline{2.4} = 0 & \underline{2} \text{ ne peut être inversible} \\ \underline{3} & \underline{3.6} = \underline{9} = \underline{1} \\ \underline{4.2} = 0 & \underline{4} \text{ ne peut être inversible} \\ \underline{5} & \underline{5.5} = \underline{25} = \underline{1} \\ \underline{6.4} = 0 & \\ \underline{7} & \underline{7.7} = \underline{49} = \underline{1} \end{array}$$

Tous les éléments de $\mathbb{Z}/7\mathbb{Z}$ sauf $\underline{0}$ sont inversibles.

2.2 - Soit p un nombre premier et $0 < q < p$. Alors p et q sont premiers entre eux. On a donc une relation de Bezout :

$$1 = ap + bq$$

qui donne, en passant aux classes modulo p :

$\underline{1} = \underline{a.p} + \underline{b.q}$ c.-à-d. comme $\underline{p} = 0$, $\underline{1} = \underline{bq}$, ce qui montre que $\underline{q}$ est inversible et que $\mathbb{Z}/p\mathbb{Z}$ est un corps.

Si p n'est pas premier, on peut écrire :

$$p = rs \text{ avec } 1 < r < p \quad 1 < s < p \quad \text{d'où}$$

$\underline{p} = \underline{rs}$ c.-à-d. $\underline{r.s} = 0$ et $\underline{r}$ (ou $\underline{s}$) ne peuvent être inversibles et $\mathbb{Z}/p\mathbb{Z}$ ne peut être un corps.

Le raisonnement précédent se généralise facilement :

Si m et n sont premiers entre eux on a une relation de Bezout $1 = am + nb$ d'où $\underline{1} = \underline{a.m} + \underline{0}$ ce qui montre que $\underline{m}$ est inversible. Si $m < n$ et si m, n ne sont pas premiers entre eux, soit r facteur commun : $m = rm'$ et $n = rn'$. Il en résulte $\underline{mn'} = \underline{rm'n'}$ et $\underline{mn'} = \underline{n.m'} = 0$ et $\underline{m}$ ne peut être inversible.

2.3 - si $x = 0$ la relation est vraie.

si $x \neq 0$ on a $x^p - x = 0$ entraîne $x^{p-1} - 1 = 0$ ("simplifier" par x , c.-à-d. multiplier à droite et à gauche par x^{-1} qui existe d'après 2.

Or, $\mathbb{Z}/p\mathbb{Z} - \{0\}$ est un groupe d'ordre $p - 1$ (c'est le groupe multiplicatif du corps $\mathbb{Z}/p\mathbb{Z}$).

Or, il y a un résultat qui dit : Si G est un groupe fini k , pour tout $x \in G$, on a $x^k = 1$.

Dans le cas présent, $k = p - 1$.

.5 - On a simplement :

$$\det A = X^2 + X^2(1 + X) = X^3 + 2X^2.$$

Le résultat est un polynôme, c.-à-d. $\det A$ est de même nature que les coefficients de la matrice A : $\det A \in \mathcal{A}$.

On précise maintenant $K = \mathbb{Z} / 2\mathbb{Z}$. Le calcul fait reste valable, sauf que certaines simplifications vont apparaître dans les coefficients des éléments (polynômes) de A , dues à la nature particulière de K . On a dans le cas précis ici $2 = 0$, c.-à-d. $\det A = X^3$.

$$A^2 = \begin{pmatrix} X^2 - X^2 - X^3 & X + X^2 + X + X^2 \\ -X^3 - X^3 & -X^2 - X^3 + X^2 \end{pmatrix} = \begin{pmatrix} X^3 & 0 \\ 0 & X^3 \end{pmatrix}$$

car dans K , $1 = -1$ et $-2 = 2 = 0$, d'où $\det A^2 = X^6 = (\det A)^2$.

La relation générale à laquelle est faite allusion est : $\det(AB) = \det A \det B$

Remarquez : $\det(AB) = \det(BA)$.

$$\begin{aligned} \text{.6 - On a} \quad P_1 + P_2 &= X^2 - 3X + X^2 + X + 1 \\ &= 2X^2 - 2X + 1 \\ &= 2X^2 + 2X + 1 \end{aligned}$$

car $-2 = 2$ dans $\mathbb{Z} / 4\mathbb{Z}$.

$$\begin{aligned} \text{et} \quad P_1 P_2 &= X^4 + X^3 + X^2 - 3X^3 - 3X^2 - 3X \\ &= X^4 + 2X^3 + 2X^2 + X \end{aligned}$$

car $-3 = 1$ dans $\mathbb{Z} / 4\mathbb{Z}$.

L'exposant de X^k n'est pas de même nature que les coefficients d'un polynôme. L'entier k désigne une case, une place, dans l'écriture d'un polynôme (k indique le degré de la case). On peut aussi remarquer que X^4 est le produit de X par lui-même répété 4 fois. On ne peut pas répéter une opération 2,3 fois ou π fois, ou 6 fois où 6 est une classe d'équivalence. Mais, il se pourrait bien qu'une opération répété 4 fois donne un résultat particulier, ainsi :

$$2^4 = 2.2.2.2$$

Mais, si le 2 qu'on utilise est celui du $2 = \underline{2}$ de $\mathbb{Z} / 4\mathbb{Z}$, on a bien évidemment $2^4 = 0$. Ceci ne peut pas arriver avec $2 \longrightarrow X$, car par construction même des polynômes et de l'indéterminée X , il n'y a aucune relation $= 0$ où figure X sauf celle où les coefficients figurant devant un X^k sont tous nuls.

.7 - P_1 et P_2 sont des polynômes en X :

$$P_1 = a_0 + a_1X + a_2X^2$$

$$P_2 = b_0 + b_1X + b_2X^2$$

On a donc $a_0 = b_0$; $a_1 = b_1 = Y$; $a_2 = b_2 = Y^2$ et donc $P_1 = P_2$.

Ici, $a_i \in \mathcal{A} = K[Y]$.

$$P_1 + P_2 = 2P_1 = 2 + 2YX + 2Y^2X^2$$

$$\begin{aligned} P_1 \cdot P_2 &= P_1^2 = 1 + Y^2 X^2 + Y^4 X^4 + 2YX + 2Y^2 X^2 + 2Y^3 X^3 \\ &= 1 + 2YX + 3Y^2 X^2 + 2Y^3 X^3 + Y^4 X^4 \end{aligned}$$

coefficients = $1, 2Y, 3Y^2, 2Y^3, Y^4$.

Les écritures précédentes font apparaître des polynômes en X et Y , polynômes à 2 indéterminées. On notera $K[X, Y]$ l'ensemble de ces polynômes. Un élément de $K[X, Y]$ est une somme formelle finie

$$\sum_{i,j} a_{i,j} X^i Y^j$$

où on a mis arbitrairement X à gauche de Y .

En postulant les relations

$$X^i Y^j = Y^j X^i,$$

on peut donc mettre aussi bien les X à droite de Y .

L'égalité $K[X, Y] = K[Y][X]$ s'obtient en ordonnant la somme précédente "par rapport aux puissances de X " et l'autre relation : $K[X, Y] = K[X][Y]$ s'obtient en ordonnant par rapport à l'indéterminée Y .

.8 - Le polynôme P_1 du haut de la page T3 est non nul car le coefficient de X^3 (par exemple) est non nul car égal à $1 \neq 0$ donc $P_1 \neq 0$.

Par contre, si on substitue x à X , où $x \in \mathbb{Z}/3\mathbb{Z}$, le (petit) théorème de Fermat nous dit $x^3 - x = 0$. La fonction polynôme produite par P_1 est donc la fonction nulle.

3 - Prenons $A = \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix}$, la matrice $A - XI = \begin{pmatrix} 1-X & 2 \\ 1 & 1-X \end{pmatrix}$

On voit donc naturellement apparaître une matrice à coefficients dans $K[X]$ alors qu'au départ, A était à coefficients dans K et que XI est une matrice à coefficients dans $K[X]$. Cela vient de l'écriture d'un polynôme $a_0 + a_1 X + a_2 X^2 + \dots$ où sans le dire, on assimile a_0 au polynôme constant $a_0 X^0$.

On peut donc "lire" la matrice A comme étant une matrice à coefficients dans $K[X]$.

4.1 - Les opérateurs sont " σ " pour "addition" et " φ " pour "élévation au carré".

La formule s'exprime donc par " $\varphi \circ \sigma = \sigma \circ \varphi$ ". Cette commutation est vraie si $K = \mathbb{Z}/2\mathbb{Z}$ car alors $2ab = 0 \cdot ab = 0$.

.2 - La formule $(a+b)^2 = a^2 + b^2$ se généralise : $(a+b)^p = a^p + b^p$ dans $\mathbb{Z}/p\mathbb{Z}$ (p étant premier).

En effet pour tout x dans $\mathbb{Z}/p\mathbb{Z}$: $x^p = x$

En fait la relation $(a+b)^p = a^p + b^p$ est vraie dans tout anneau $\mathcal{A}$ où

$$1 + \dots + 1 \text{ (p fois)} = 0$$

par exemple dans $(\mathbb{Z}/p\mathbb{Z})[X]$. Vous pouvez le démontrer en utilisant la formule du binôme.

.3 - L'unicité se démontre ainsi :

Soit σ un tel homomorphisme : nous avons donc nécessairement

$$\sigma(X^k) = (\sigma(X))^k = a^k$$

car σ est un homomorphisme multiplicatif.

Ensuite, $\sigma(\alpha X^k) = \sigma(\alpha X^0 \cdot X^k) = \sigma(\alpha X^0) \sigma(X^k)$.

Si on rajoute donc l'hypothèse que σ est égale à l'identité sur les polynômes constants on a $\sigma(\alpha X^k) = \alpha a^k$. (On dit aussi que σ doit être un homomorphisme d'algèbre, une "algèbre" étant un anneau + espace vectoriel ; exemples typiques $K[X]$ et $M_{n \times n}(K)$).

Finalement, si $P(X) = \alpha_0 + \alpha_1 X + \dots + \alpha_n X^n$ on a par additivité (ou par "linéarité") que $\sigma(\alpha_0 + \alpha_1 X + \dots + \alpha_n X^n) = \alpha_0 + \alpha_1 a + \dots + \alpha_n a^n$.

Pour montrer l'existence de σ on pose donc maintenant a priori :

$$\sigma P(X) = \sigma \left(\sum \alpha_i X^i \right) = \sum \alpha_i a^i$$

Il en résulte $\sigma(\alpha) = \alpha$ $\sigma(X) = a$ $\sigma(X^k) = a^k$.

$$\begin{aligned} \text{Soit} \quad P_1(X) &= \sum \alpha_i X^i \\ P_2(X) &= \sum \beta_i X^i \quad (\text{somme finie}) \end{aligned}$$

$$\text{On a} \quad (P_1 + P_2)(X) = \sum (\alpha_i + \beta_i) X^i.$$

$$\begin{aligned} \text{D'où} \quad \sigma(P_1 + P_2) &= \sum (\alpha_i + \beta_i) a^i \\ &= \sum \alpha_i a^i + \sum \beta_i a^i \\ &= \sigma(P_1) + \sigma(P_2) \end{aligned}$$

Pour le produit : $(P_1 P_2)(X) = \sum \gamma_i X^i$ (finie) avec

$$\gamma_i = \sum \alpha_r \beta_s, \quad \text{somme étendue à tous les } r \text{ et } s \text{ tels que } r + s = i.$$

$$\text{Donc, } \sigma(P_1 P_2) = \sum \gamma_i a^i.$$

Or, si on effectue le produit dans A :

$$(\alpha_0 + \alpha_1 a + \dots + \alpha_n a^n) (\beta_0 + \beta_1 a + \dots + \beta_n a^n)$$

en regroupant par rapport aux puissances de a , on trouve le même coefficient α_i devant a^i ce qui termine la démonstration.

Le lecteur devrait rester avec un léger malaise après cette démonstration, car elle ne nous fait guère avancer par rapport à l'intuition première : "les calculs qu'on fait avec les a^i sont formellement homologues à ceux qu'on fait avec les X^i ".

Pour décortiquer davantage le phénomène évident à nos yeux, il faut alors fouiller davantage la nature de $K[X]$.

6 - Vous trouverez dans $M_{n \times n}(K) \simeq K^{n^2}$ la norme suivante :

$$\|a\| = \|a\| = n \sup |a_{i,j}| \quad \text{si } a = (a_{i,j})$$

On a alors, si $b = (b_{j,k})$ et $ab = (c_{i,k})$

$$\|ab\| = n \sup |c_{i,k}| = n \sup \left(\left| \sum_j a_{i,j} b_{j,k} \right| \right)$$

$$\leq n \sup \left(\sum_j |a_{i,j}| |b_{j,k}| \right) \leq n^2 \sup |a_{i,j}| \sup |b_{j,k}|$$

$$\text{c.-à-d.} \quad \|ab\| \leq \|a\| \|b\|.$$

La norme de $\mathbb{R}^k$: $\|x\| = \sup |x_i|$ est une norme classique de $\mathbb{R}^k$ équivalente à la norme euclidienne $\sqrt{\sum x_i^2}$.

Pour avoir la compatibilité avec le produit, il suffit de multiplier cette norme par n quand $k = n^2$.

7.1 - On a $e^z e^{-z} = e^{z-z} = e^0 = 1$.

L'inverse de e^z existe donc et vaut e^{-z} . Le nombre e^z ne peut s'annuler comme étant inversible (ou bien vous auriez $0 \cdot e^{-z} = 1$).

Cette démonstration s'applique évidemment aux matrices car a et $b = -a$ commutent :

$$e^a \cdot e^{-a} = e^{a-a} = e^0 = I.$$

La conclusion est la même.

.2 - Calculons la dérivée en t_0 : il faut regarder le rapport $\frac{e^{ta} - e^{t_0 a}}{t - t_0} = \Delta(t, t_0)$

Or,
$$\Delta(t, t_0) = e^{t_0 a} \cdot \frac{e^{(t-t_0)a} - 1}{t - t_0}$$

Car $t_0 a$ et ta sont 2 matrices qui commutent (1 pour I).

Le rapport $\Delta(t, t_0)$ a donc une limite si et seulement si le rapport $\Delta(h, 0)$ en a une (poser $h = t - t_0$).

Et on a
$$\frac{d}{dt} e^{ta} \Big|_{t=t_0} = e^{t_0 a} \cdot \frac{d(e^{ta})}{dt} \Big|_{t=0}$$

Reste à calculer
$$\lim_{h \rightarrow 0} \frac{e^{ha} - 1}{h}$$

Les notations ne sont pas différentes du cas $h, a, 1 \in \mathbb{R}$, de façon intentionnelle, car la détermination de cette limite dans le cas de $\mathbb{R}$ se transpose intégralement au cas des matrices :

$$e^{ha} = 1 + ha + \frac{h^2 a^2}{2!} + \dots$$

$$\frac{e^{ha} - 1}{h} = a + h \frac{a^2}{2!} + \dots \text{ et donc limite } \frac{e^{ha} - 1}{h} = a, \text{ ce qui donne :}$$

$$\frac{d}{dt} e^{ta} \Big|_{t=t_0} = e^{t_0 a} \cdot a$$

.3 - $P(a)$ et a commutent car a commute avec chaque terme de $P(a)$, donc avec la somme de ces termes. Or, e^{ta} est une limite d'une suite $S_m(t, a) = \sum \frac{t^n a^n}{n!}$ par définition de l'exponentielle (t et a sont fixés ici)

La matrice a commute avec chaque S_m , S_m étant un polynôme en a :

$$a S_m = S_m a$$

Il faut donc passer à la limite dans cette égalité ce qui est justifié si les applications $x \mapsto ax$ et $x \mapsto xa$ de $\mathbb{R}^k$ dans $\mathbb{R}^k$ sont continues ($k = n^2$). Ceci est vrai pour $k = 1$.

Si vous revoyez la démonstration élémentaire de cette propriété dans le cas de $\mathbb{R}$, vous saurez la faire immédiatement dans le cas de $\mathbb{R}^k$ grâce à la propriété de la "valeur absolue" :

$$|xy| \leq |x| |y|$$

dans le cas des matrices.

.4 - Si $x = (x_1, x_2, \dots, x_n)$, "x" fonction inconnue de t est donc en fait une suite de n fonctions inconnues x_i de la variable t .

la dérivée de la fonction vectorielle x s'obtient en dérivant chaque composante x_i :

$$x' = (x_1', \dots, x_n')$$

Si $a = (a_{i,j})$ alors $x' = a.x$ (où x' et x sont écrites sous forme "colonne") donne :

$$x_1' = a_{1,1}x_1 + a_{1,2}x_2 + \dots + a_{1,n}x_n$$

$$x_i' = a_{i,1}x_1 + \dots + a_{i,n}x_n$$

$$x_n' = a_{n,1}x_1 + \dots + a_{n,n}x_n$$

.5 - Au cas où $x \in \mathbb{R}^n$, l'équation prend un sens si a et b sont des matrices $n \times n$.

L'astuce consiste à prendre "deux" fonctions inconnues : x_1 et x_2 où

$$x_1 = x$$

$$x_2 = x'$$

L'équation $x'' = ax' + bx$ s'écrit alors :

$$x_2' = ax_2 + bx_1$$

et finalement

$$\begin{cases} x_1' = x_2 \\ x_2' = bx_1 + ax_2 \end{cases}$$

Si on pose $y = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ ce système devient

$$y' = \begin{pmatrix} x_1' \\ x_2' \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ b & a \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

c.-à-d. $y' = \alpha.y$ où α est une matrice "en bloc".

En fait, y consiste en $2n$ fonctions inconnues.

.6 - On a $a^2 = \begin{pmatrix} \lambda^2 & 0 \\ 0 & \mu^2 \end{pmatrix}$ et en général

$$\begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix} + \begin{pmatrix} \lambda' & 0 \\ 0 & \mu' \end{pmatrix} = \begin{pmatrix} \lambda + \lambda' & 0 \\ 0 & \mu + \mu' \end{pmatrix}$$

D'où de façon évidente

$$\sum \frac{a^n}{n!} = \begin{pmatrix} \sum \lambda^n / n! & 0 \\ 0 & \sum \mu^n / n! \end{pmatrix} \quad \text{c.-à-d.} \quad e^a = \begin{pmatrix} e^\lambda & 0 \\ 0 & e^\mu \end{pmatrix}.$$

Les principes de calcul précédents ne sont pas tributaires de la dimension 2. Donc si

$$a = \begin{pmatrix} \lambda_1 & \dots & 0 \\ \dots & \dots & \dots \\ 0 & \dots & \lambda_n \end{pmatrix} \quad \text{alors} \quad e^a = \begin{pmatrix} \dots & \dots & 0 \\ \dots & e^{\lambda_i} & \dots \\ 0 & \dots & \dots \end{pmatrix}$$

.7 - On a
$$a - \lambda I = b = \begin{pmatrix} 0 & 0 & \dots & 0 \\ 1 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & \dots & 1 & 0 \end{pmatrix}$$

Une matrice du type b , c.-à-d. triangulaire avec que des zéros sur la diagonale principale, est une matrice nilpotente, c.-à-d. qu'il existe un k tel que $b^k = 0$.

L'application de la règle de calcul du produit de deux matrices montre en effet tout de suite que le calcul de b^2 fait apparaître un 0 en $i = 2, j = 1$. Le cran au-dessus (b^3) fait apparaître deux 0 supplémentaires en $i = 3; j = 1$ et $j = 2$.

De proche en proche, on fait ainsi apparaître des 0 dans le triangle inférieur et au cran n on a : $b^n = 0$.

Ceci est une démonstration.

Une autre, plus élégante, consiste à introduire une application linéaire, dont la matrice par rapport à une base (e_i) est une matrice triangulaire c (inf) avec que des 0 sur la diagonale principale : le produit c^2 correspond alors à $v \circ v = v^2$. Montrer que $c^n = 0$ revient à montrer que $v^n = 0$ c.-à-d. que $v^n(e_i) = 0$ pour tout i de 1 à n .

Appelons E_0 , le sous espace engendré par $e_1, \dots, e_n$, E_1 le sous-espace engendré par $e_2, \dots, e_n$, E_2 celui engendré par $e_3, \dots, e_n$ etc... jusqu'à E_{n-1} celui engendré par le vecteur e_n , et $E_n = 0$. On a ainsi une suite décroissante (appelée un "drapeau") :

$$\mathbb{R}^n = E_0 \supset E_1 \supset E_2 \supset \dots \supset E_{n-1} \supset e_0 = (0)$$

Or, vue la forme de la matrice, on a de façon évidente :

$$v(E_0) \subset E_1$$

$$v(E_1) \subset E_2$$

.....

$$v(E_{n-1}) = 0 = E_n$$

et ensuite

$$v(v(E_0) \subset v(E_1) \subset E_2$$

c.-à-d.

$$v^2(E_0) \subset E_2$$

et ainsi de suite :

$$v^3(E_0) \subset E_3$$

et finalement

$$v^n(E_0) \subset E_n = (0)$$

ce qui signifie

$$v^n(e_i) = 0 \quad \forall i.$$

Cette démonstration est plus "communicable" à quelqu'un d'autre que soi. La première est aussi convaincante mais demande une vérification personnelle.

Mais revenons à notre problème : si on sait calculer $e^{a-\lambda I}$, on en déduira e^a car

$$e^a = e^{a-\lambda I} \cdot e^{\lambda I}$$

Nous sommes donc ramenés à calculer e^b avec $b = \begin{pmatrix} 0 & \dots & \dots & \dots & 0 \\ 1 & 0 & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & 1 & 0 \end{pmatrix}$

Comme $b^n = 0$ on peut écrire $e^b = 1 + b + \frac{b^2}{2!} + \dots + \frac{b^{n-1}}{(n-1)!}$.

La matrice b étant non seulement une matrice triangulaire avec des 0 sur la diagonale principale mais une matrice bien plus particulière, une matrice de Jordan, un calcul direct ou un raisonnement géométrique utilisant v , montre que pour obtenir b^2 il faut faire descendre la diagonale des "1" de un cran, et ainsi de suite pour b^3 , jusqu'à $b^n = 0$.

.8 - La justification qui est visée ici est la suivante : très souvent on énonce des problèmes qui n'ont pas de solutions calculables, ni même de façon approchée : les équations qui régissent les mouvements des planètes sont des équations différentielles pour lesquels on démontre qu'il existe une solution et une seule les conditions initiales étant précisées.

Cela signifie que dans le modèle mathématique, la connaissance des positions et des vitesses des planètes à un instant, détermine complètement l'évolution des planètes dans le futur (et dans le passé). Or pratiquement, on ne peut calculer que de façon approchée ces solutions mathématiques, le calcul demandant l'injection des conditions initiales. Or ces conditions sont entachées d'une erreur de mesure physique. La conjonction de ces deux approximations entraîne que, si on veut connaître la position d'une planète longtemps plus tard ("long" pouvant être relativement court) on ne peut strictement plus rien prédire (la position de la planète pourrait être n'importe où).

Etant dans l'impossibilité de dire quoique ce soit au sujet d'une solution, on considère la collection de toutes les solutions (obtenues en faisant varier les conditions initiales) et on essaie de dire quelque chose au sujet de l'ensemble de ces solutions.

8.1 - Si parmi les polynômes de N il y en a qui ont un degré, c.-à-d. qui sont non nuls, alors on peut en choisir un de degré minimum.

Mais si jamais $N = (0)$?

Il nous reste donc, pour être parfaitement correct, à démontrer que $N \neq (0)$.

.2 - $P(A) = \sum a_i A^i$ est une relation linéaire entre les vecteurs $v_i = A^i$ de $M_{n \times n}$.

Considérons les $n^2 + 1$ vecteurs $v_0, v_1, \dots, v_{n^2}$. Comme l'espace vectoriel $M_{n \times n}$ est de dimension n^2 , ce système est lié ; il existe donc des scalaires $\lambda_0, \lambda_1, \dots, \lambda_{n^2}$ non tous nuls tels que :

$$\lambda_0 v_0 + \lambda_1 v_1 + \dots + \lambda_{n^2} v_{n^2} = 0$$

c.-à-d.

$$Q(A) = 0 \quad \text{avec} \quad Q(x) = \sum_{i=0}^{n^2} \lambda_i x^i \quad \text{et} \quad Q \neq 0.$$

.3 - Soit $A = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_n \end{pmatrix}$.

Alors $\chi_A(X) = (\lambda_1 - X)(\lambda_2 - X) \dots (\lambda_n - X)$

et $\chi_A(A) = (\lambda_1 I - A) (\lambda_2 I - A) \dots (\lambda_n I - A) .$

La matrice $\lambda_i I - A$ est une matrice diagonale avec un 0 à la place (i,i).
Comme pour faire le produit de matrices diagonaux on multiplie tout simplement en chaque place (i,i) , le produit des matrices $(\lambda_1 I - A) \dots (\lambda_n I - A)$ est nul. En effet en chaque place, il y a au moins une des matrices dont le terme correspondant est nul.

.4 - Si $A' = S^{-1}AS$, on a $A' - \lambda I = S^{-1}(A - \lambda I)S$

car $S^{-1}(A - \lambda I)S = (S^{-1}A - \lambda S^{-1}I)S = S^{-1}AS - \lambda S^{-1}IS = S^{-1}AS - \lambda I = A' - \lambda I$

et $\chi_{A'}(X) = \det(A' - XI) = \det(S^{-1}.A - XI.S) = \det S^{-1}.\det.(A - XI).\det S = \det(A - XI)$

car $\det S^{-1}.\det S = 1$.

Donc $\chi_{A'}(X) = \chi_A(X) .$

Si on sait $\chi_A(A) = 0$, on a donc $\chi_{A'}(A) = 0 .$

Or, $\chi_{A'}(A') = \chi_{A'}(S^{-1}AS) = S^{-1}\chi_{A'}(A)S$ car $S^{-1}(\sum a_i A^i)S = \sum a_i S^{-1}A^i S$

Il en résulte donc $\chi_{A'}(A') = 0 .$

10.1 - Soit $A = \begin{pmatrix} \lambda & 0 & 0 \\ a & \mu & 0 \\ b & c & \nu \end{pmatrix} .$

On sait que

$$\det A = \lambda \mu \nu .$$

Donc, de même

$$\det.(A - XI) = (\lambda - X) (\mu - X) (\nu - X) .$$

En calculant directement

$$(\lambda - A) ,$$

puis

$$(\lambda - A) (\mu - A) ,$$

et finalement

$$(\lambda - A) (\mu - A) (\nu - A) ,$$

on trouve 0 .

.2 - Cette façon de procéder se généralise facilement au cas général.

Seulement, cette méthode est convaincante si chacun le fait soi-même à la main.

Cherchons donc une méthode plus communicable :

Il faut reprendre l'idée géométrique de l'exercice 7.8 :

Ecrivons $\chi_A(A) = (\lambda_n I - A) \dots (\lambda_1 I - A)$ si $A = \begin{pmatrix} \lambda_1 & \dots & 0 \\ \dots & \dots & \dots \\ 0 & \dots & \lambda_n \end{pmatrix} .$

Introduisons le drapeau (E_i) comme plus haut et posons v_i égal à l'endomorphisme linéaire dont la matrice par rapport à la base (e_i) est $\lambda_i I - A$.

Cette fois nous avons toujours :

$$v_1(E_0) \subset E_1$$

mais seulement

$$v_1(E_1) \subset E_1 .$$

En général, nous avons : $v_i(E_{i-1}) \subset E_i$.

De $v_1(E_0) \subset E_1$, nous trouvons donc $v_2 \circ v_1(E_0) \subset v_2(E_1) \subset E_2$

En continuant, cela donne :

$$v_3 \circ v_2 \circ v_1(E_0) \subset v_3(E_2) \subset E_3$$

et finalement

$$v_n \circ v_{n-1} \circ \dots \circ v_1(E_0) \subset E_n = (0) .$$

Ce qui veut dire que l'endomorphisme $v_n \circ v_{n-1} \circ \dots \circ v_1$ est nul.

Or la matrice de cet endomorphisme est justement $\chi_A(A)$.

11.1 - Le "seulement en général" fait appel à la page T9.

.2 - Si $P(X) = 1 + X + X^2$, alors $P(u) = \text{id} + u + u^2$ où $u = E \longrightarrow E$. Alors $u^2 = u \circ u$ et on sait additionner deux endomorphismes. Le sens de $P(u)$ est donc bien clair en général. On sait que $\text{End } E$ est une algèbre et on a un homomorphisme d'algèbres.

$$\sigma: K[X] \longrightarrow \text{End } E ; P(X) \longmapsto P(u) .$$

On sait que σ est complètement défini par $\sigma(X)$ qui vaut u .

$$P(u).x = (\text{id} + u + u^2).x = \text{id}(x) + u(x) + u^2(x) .$$

Les formules $(PQ)(u) = P(u) Q(u)$ etc. sont l'analogue de $(PQ)(A) = P(A)Q(A)$. Elles sont trop évidentes pour être facilement démontrées.

12.1 - La récurrence s'amorce ainsi :

On écrit $P = Q_1 Q_2$ avec $Q_1 = P_1 \dots P_{n-1}$, $Q_2 = P_n$.

Alors, par hypothèse Q_1 et Q_2 sont premiers entre eux.

On a donc une décomposition $E = E'_1 + E_n$ où

$$E_n = \text{Ker } Q_2(u) = \text{Ker } P_n(u) \text{ et } E'_1 = \text{Ker } Q_1(u) .$$

E'_1 est stable par u et Q_1 annule la restriction de u à E'_1 .

On applique alors l'hypothèse de récurrence à E'_1 , u/E'_1 , Q_1 où $Q_1 = P_1 \dots P_{n-1}$ etc. (cqfd)

.2 - Il suffit de remarquer que :

$$E_1 = \text{Ker } Q_1(u) = \text{Im } Q_2(u)$$

car si comme à la page T12 $1 = R Q_1 + S Q_2$, on a

$$x = R(u)(Q_1(u).x) + Q_2(u)(S(u).x)$$

si donc $x \in \text{Ker } Q_1(u)$, on a

$$x = Q_2(u).y \text{ avec } y = S(u).x$$

c.-à-d. $\text{Ker } Q_1(u) \subset \text{Im } Q_2(u)$.

L'inclusion inverse est évidente car un élément de $\text{Im } Q_2(u)$ est de la forme $Q_2(u).y$ et donc $Q_1(u)(Q_2(u).y) = Q_1 Q_2(u).y$ qui est nul car $Q_1 Q_2$ annule u .

13.1 - " λ est valeur propre de u " veut dire qu'il existe un vecteur x non nul tel que $u.x = \lambda x$. L'espace propre E_λ est donc au moins de dimension 1 car il est non réduit à (0) .

.2 - Il suffit bien sûr de montrer que $d'_i \leq k_i$ pour tout i (même raisonnement qu'à la page T13)

L'espace E'_i est stable par u . Appelons u_i la restriction de u à E'_i . Donc u_i est un endomorphisme de E'_i . Posons $v_i = u_i - \lambda_i \text{Id}_{E'_i}$. Alors v_i est aussi un endomorphisme de E'_i .

Par définition de E'_i nous avons $v_i^{k_i} = 0$ c.-à-d. v_i est un endomorphisme nilpotent. Mais l'entier k_i n'est pas forcément l'ordre de v_i : on peut voir $v_i^{r_i} = 0$ avec $r_i \leq k_i$, r_i étant l'ordre de v_i .

En regardant le corrigé du devoir D1, on sait alors que le système

$$x, v(x), \dots, v^{r_i-1}(x) \neq 0$$

est une base de E'_i ; on a donc bien $d'_i = \dim E'_i \leq k_i$.

14.1 - Soit $x_1 \in E_1$. Il faut montrer que $u.x_1 \in E_1$ c.-à-d. que $Q_1(u)(u(x_1)) = 0$.

Or $Q_1(u).u = u.Q_1(u)$

Donc $Q_1(u)(u(x_1)) = u(Q_1(u).x) = 0$

car $Q_1(u).x = 0$ puisque $x \in E_1$.

.2 - Il faut montrer que $R(u)$ est l'endomorphisme nul c.-à-d. que $R(u).x = 0 \quad \forall x \in E$.

Pour cela, il suffit de démontrer que $R(u).x = 0 \quad \forall x \in E_1$ et $\forall x \in E_2$.

Or, dans ces deux cas c'est évident par hypothèse (il faut éventuellement écrire $R_2 R_1$ à la place de $R_1 R_2$)

.3 - On appelle "le" polynôme minimal celui dont le coefficient de plus haut degré vaut 1. Deux polynômes minimaux sont obligatoirement de même degré ; ils ne peuvent différer que par un facteur constant.

.4 - Sachant que $P_i = (X - \mu_i)^{l_i}$ est le polynôme minimal de u_i , si on avait $v_i^s = 0$ avec $s < l_i$, le polynôme $(X - \mu_i)^s$ annulerait u_i et il serait de degré strictement plus petit que celui du polynôme minimal ce qui n'est pas possible.

.5 - Ce test a bien sûr déjà été fait : cas d'une matrice triangulaire.

.6 - On voit facilement (par la règle de calcul du produit de deux matrices) :

$$A^2 = \begin{pmatrix} B_1^2 & 0 \\ 0 & B_2^2 \end{pmatrix} \text{ etc}$$

Il en résulte alors facilement que :

$$e^A = \begin{pmatrix} e^{B_1} & 0 \\ 0 & e^{B_2} \end{pmatrix}.$$

Une autre raison du calcul de A^2 est la suivante : une matrice telle que A correspond à une décomposition d'un espace E en somme directe $E = E_1 + E_2$ où $\dim E_1 = r$ et $\dim E_2 = s$ et à un endomorphisme u qui laisse E_1 et E_2 stables. La matrice A^2 correspond à u^2 qui laisse aussi E_i stable et u^2 est déterminée par u_i^2 ou u_i est la restriction de u à E_i ($i = 1, 2$).

.7 - On trouve $\chi_A(X) = -X(1 - X)^3$

On a $(A - I)^3 = \begin{pmatrix} 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{pmatrix}$

D'où $\text{Ker } (A - I)^3 = \{(x_1, x_2, x_3, x_4) / x_3 = x_4\}$

et $\text{Ker } A = \{(x_1, x_2, x_3, x_4) / x_2 = x_3 = x_1 + x_4 = 0\}$

On a $\mathbb{R}^4 = E_1 \oplus E_2$ avec $E_1 = \text{Ker } (A - I)^3$, $E_2 = \text{Ker } A$.

Si $x = (x_1, x_2, x_3, x_4) \in \mathbb{R}^4$ on peut le décomposer $x = u + v$ ou $u \in E_1$ et $v \in E_2$.

En procédant par coefficients indéterminés :

$$x = (\alpha, \beta, \varphi, \varphi) + (\delta, 0, 0, -\delta)$$

on trouve $u = (x_1 - x_3 + x_4; x_2, x_3, x_3)$

$$v = (x_3 - x_4, 0, 0, -x_3 + x_4)$$

On a $e^{tA}.x = e^{tA}.u + e^{tA}.v$

et $e^{tA}.u = e^t \cdot e^{t(A-I)}.u = e^t [I + t(A-I) + \frac{t^2}{2}(A-I)^2] (u)$

Le calcul montre que $(A - I)^2.u = 0$. Il reste

$$e^{tA}.u = e^t [I + t(A - I)].u$$

$e^{tA}.v = v$ même méthode.

c.-à-d.
$$e^{tA}.x = \begin{pmatrix} 1 & 0 & -t & t \\ 0 & 1 & t & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & t & 1-t \end{pmatrix} \begin{pmatrix} x_1 - x_2 + x_4 \\ x_2 \\ x_3 \\ x_3 \end{pmatrix} + \begin{pmatrix} x_3 - x_4 \\ 0 \\ 0 \\ x_3 + x_4 \end{pmatrix}$$

$$= (e^t x_1 + (1 - e^t)x_3 - (1 - e^t)x_4, e^t x_2 + te^t x_3, e^t x_3, -(1 - e^t)x_3 + x_4)$$

D'où finalement :

$$e^{tA} = \begin{pmatrix} e^t & 0 & 1-e^t & -(1-e^t) \\ 0 & e^t & te^t & 0 \\ 0 & 0 & e^t & 0 \\ 0 & 0 & -(1-e^t) & 1 \end{pmatrix}$$

16.1 - On a $P_1 + P_2 = 1 + 3X + 9X^2$

$$P_1 P_2 = X + 8X^2 + 15X^3 + 18X^4$$

1) D'où $\sigma(P_1 + P_2) = 1 + 3X + 1X^2$

$$\sigma(P_1 P_2) = 1x + 7X^3 + 2X^4$$

$$\sigma(P_1) = 1 + 2X + 3X^2$$

$$\sigma(P_2) = 1X + 6X^2$$

Le calcul $\sigma(P_1) + \sigma(P_2)$ et $\sigma(P_1) \sigma(P_2)$ redonnent les deux premières lignes.

2) En maintenant la rotation σ , les calculs s'écrivent :

$$\sigma(P_1 + P_2) = \sigma(1) + \sigma(3) X + \sigma(9) X^2$$

$$\sigma(P_1 P_2) = \sigma(1) X + \sigma(8) X^2 + \sigma(15) X^3 + \sigma(18) X^4$$

$$\sigma(P_1) = \sigma(1) + \sigma(2) X + \sigma(3) X^2$$

$$\sigma(P_2) = \sigma(1) X + \sigma(6) X^2$$

et $\sigma(P_1) + \sigma(P_2) = \sigma(1) + (\sigma(1) + \sigma(2)) X + (\sigma(3) + \sigma(6)) X^2$

Or, $\sigma(1) + \sigma(2) = \sigma(1+2) = \sigma(3) = 3$

$$\sigma(3) + \sigma(6) = \sigma(9) = \sigma(1) = 1.$$

$$\sigma(P_1) \cdot \sigma(P_2) = \sigma(1) \cdot \sigma(1) X + ((\sigma(1)\sigma(2) + \sigma(6)\sigma(1)) X^2 + (\sigma(2)\sigma(6) + \sigma(3)\sigma(1)) X^3 + \sigma(3)\sigma(6) X^4$$

$$\text{Or } \sigma(1) \cdot \sigma(1) = \sigma(1.1) = \sigma(1) = \underline{1}$$

$$\sigma(1)\sigma(2) + \sigma(6)\sigma(1) = \sigma(1.2) + \sigma(6.1) = \sigma(2) + \sigma(6) = \sigma(8) = \underline{0}$$

$$\sigma(2)\sigma(6) + \sigma(3)\sigma(1) = \sigma(2.6) + \sigma(3.1) = \sigma(12) + \sigma(3) = \sigma(15) = \sigma(7) = \underline{7}$$

$$\sigma(3)\sigma(6) = \sigma(3.6) = \sigma(18) = \underline{2}$$

On retrouve bien sûr la même chose. Le deuxième calcul montre clairement que c'est le fait que σ est un homomorphisme d'anneaux qui rend les deux formules justes.

$$.2 - \sigma.A = \begin{pmatrix} \sigma.a & \sigma.b \\ \sigma.c & \sigma.d \end{pmatrix} \text{ par définition.}$$

$$\sigma.A = 0 \iff \sigma.a = \sigma.b = \sigma.c = \sigma.d = 0 \text{ par définition.}$$

Or, $\sigma.a = 0 \Rightarrow a = 0$ par hypothèse (σ injectif).

De même pour b, c, d .

$$.3 - \text{Bien sûr } A = \begin{pmatrix} 3 & 1 \\ -1 & 0 \end{pmatrix}$$

17.1 - σ est surjectif car il y a autant d'éléments dans K qu'il y a de polynômes constants ;
or $\sigma(\text{polynôme constant } c) = c$.

La formule $\sigma(P(X, Y, Z, T)) = P(a, b, c, d)$ est de nouveau d'une évidence troublante.

.2 - Si A est nilpotente on a $A^k = 0$. Le polynôme X^k annule donc A . Le polynôme minimal divise donc X^k ; il est donc de la même forme X^r . Les seules racines du polynôme minimal sont donc $\lambda = 0$. Le polynôme caractéristique a les mêmes racines. Il a donc comme seule racine $\lambda = 0$ c.-à-d. $\chi_A(X) = X^n$. Réciproquement, si $\chi_A(X) = X^n$, le théorème de CH dit $A^n = 0$ et A est donc nilpotente.

Apparemment, on n'a pas fait appel au fait que $K = \mathbb{R}$ ou non. En fait, le raisonnement marche pour un corps quelconque si on admet que chaque corps admet un "corps complexe" qui l'englobe de sorte que tout polynôme a toutes ses racines dans le corps (voir T).

Si K est un anneau ?

La réciproque marche pareille car elle dépend uniquement de CH qui est valable pour un anneau quelconque.

Supposons A nilpotente. On a l'égalité entre matrices.

$$A^p - \lambda^p I = (A - \lambda I)(A^{p-1} + \lambda A^{p-2} + \dots + \lambda^{p-1} I) \text{ (Il suffit de développer).}$$

Si $A^p = 0$, le premier nombre se réduit à $\lambda^p I$. En prenant le déterminant à droite et à gauche, on trouve une égalité de la forme $X^r = X_A(X) \cdot Q(X)$ (avec $r = np$ si A est $n \times n$).

Le polynôme $\chi_A(X)$ divise donc X^r .

On est tenté de conclure que $\chi_A(X)$ est égal à X^n , mais cette conclusion peut être fausse pour un anneau absolument quelconque.

Par exemple, si $A = \mathbb{Z}/4\mathbb{Z}$, on a :

$$X^2 = (X - \underline{2})(X - \underline{2})$$

Il y a bien sûr beaucoup d'anneaux pour lesquels ceci n'arrive pas (p. ex les anneaux intègres, les anneaux factoriels).

.3 - On a tout simplement :

$$E = E_1 \oplus E'_1 \text{ avec donc}$$

$$E_1 = \text{Ker}(u - \lambda)^k \text{ et } E'_1 = \text{Im}(u - \lambda)^k$$

car $\chi_A(X) = (X - \lambda)^k \cdot Q_2 \quad (\text{avec } Q_1 = (X - \lambda)^k) \text{ (ex 12.2)}$

Le polynôme caractéristique de la restriction de u à E_1 est $(X - \lambda)^k$.

D'abord, ce polynôme ne peut être que de cette forme (refaire le raisonnement de la page C20, en passant par le polynôme minimal), ensuite, l'exposant est nécessairement k car k est la dimension de E_1 (voir 13.2).

Une autre démonstration consiste à décrire la matrice de u relativement au sous-espace E_1 (voir aussi corrigé du devoir 1).

.4 - On a $E = E''_1 \oplus \dots \oplus E''_r$ où $E''_i = \ker(u - \lambda_i)^{l_i}$ avec les notations de la page T13.

Si les racines du polynôme minimal* sont simples, on a $E''_i = \ker(u - \lambda_i)$ c.-à-d. E''_i est l'espace propre relatif à λ_i et E est somme de ses espaces propres ce qui veut dire que u est diagonalisable.

La réciproque : supposons u diagonalisable.

On a $E = E_{\lambda_1} + \dots + E_{\lambda_r}$ où E_{λ_i} sont les espaces propres :

$$E_{\lambda_i} = \ker(u - \lambda_i).$$

Appelons u_i la restriction de u à E_{λ_i} .

Alors le polynôme minimal de u_i est égal à $X - \lambda_i$.

Le produit $\prod_{i=1}^r (X - \lambda_i)$ annule donc u (voir ex. 14.2) et le polynôme minimal contient

le facteur $X - \lambda_i$ au moins une fois. Le polynôme minimal est donc égal à

$$\prod_{i=1}^r (X - \lambda_i) ;$$

ses racines sont simples.

Retour à 14.5 :

Les calculs qu'on a faits nous montre que si $P(X) = X(X - 1)^2$, alors $P(A) = 0$. En fait, P est le polynôme minimal de A , car une vérification montre que $A(A - I)$ n'est pas nul.

Ce polynôme est de degré trois. L'espace des solutions de l'équation différentielle associée est donc de dimension trois. Le polynôme caractéristique de cette équation différentiel à comme racines :

$$\lambda = 0, \lambda = 1 \text{ double}$$

On a donc une base de solutions :

$$f_0(t) = e^{0 \cdot t} = 1, f_1(t) = e^t, f_2(t) = te^t$$

Ecrivons maintenant la relation p E 6 :

$$f_0 : 1 = 1 \cdot \varphi_0 + 0 \cdot \varphi_1 + 0 \cdot \varphi_2$$

$$f_1 : e^t = 1.\varphi_0 + 1.\varphi_1 + 1.\varphi_2$$

$$f_2 : te^t = 0.\varphi_0 + 1.\varphi_1 + 2.\varphi_2$$

Réolvons ce système en $\varphi_0, \varphi_1, \varphi_2$

$$\varphi_0 = 1$$

$$\varphi_1(t) = -2 + 2e^t - te^t$$

$$\varphi_2(t) = 1 - e^t + te^t$$

d'où

$$e^{tA} = I + (-2 + 2e^t - te^t)A + (1 - e^t + te^t)A^2$$

En calculant A^2 et en effectuant, on doit en principe retrouver le même résultat. C'est bien plus rapide !

Remarque : vous auriez pu partir avec le polynôme $X(X-1)^3$. Le résultat est encore le même, mais les calculs sont un peu plus longs, car le degré est 4, donc 4 inconnues $\varphi_0, \varphi_1, \varphi_2, \varphi_3$!

Quelques explications au sujet de cette méthode :

1) La formule $f(t) = f(0)\varphi_0(t) + f'(0)\varphi_1(t) + \dots$ se vérifie facilement si on se souvient de la définition de φ_i : on a $\varphi_i^{(j)}(0) = \delta_i^j$. On écrit a priori $f(t) = \lambda_0\varphi_0(t) + \lambda_1\varphi_1(t) +$

D'abord on fait $t=0$ ce qui donne $f(0) = \lambda_0.1$.

Ensuite on dérive la relation et on fait $t=0$ après dérivation ce qui donne : $f'(0) = \lambda_1.1$.

On dérive une deuxième fois, etc...

2) La formule e^{tA} se démontre de la façon suivante :

On pose $F(t) = \varphi_0(t) + \varphi(t)A + \dots + \varphi_{n-1}(t)A^{n-1}$ et on dérive la fonction $t \longrightarrow F(t)$.

On démontre alors la formule :

$$\frac{dF(t)}{dt} = AF(t) = F(t)A.$$

Or, la fonction $G(t) = e^{tA}$ vérifie cette même équation différentielle :

$$\frac{dX}{dt} = A.X = XA$$

avec la même condition initiale :

$$G(0) = F(0) = I$$

La théorie des équations différentielles nous apprend qu'une telle équation n'a qu'une seule solution : donc $G(t) = F(t) \forall t$ ce qu'on voulait démontrer.

DEVOIR N° 1

Première partie : n'est pas à rédiger.

Soit E un espace vectoriel sur K , $(e_1 \dots e_n)$ une base.

1. On appelle dual de E et on note E^* l'espace vectoriel $\mathcal{L}(E, K)$. Si x' est dans E^* , x' s'appelle forme linéaire ; on note $x'(x) = \langle x', x \rangle$ si $x \in E$.

Montrer les formules :

$$\begin{aligned}\langle x' + y', x \rangle &= \langle x', x \rangle + \langle y', x \rangle \\ \langle \lambda x', x \rangle &= \lambda \langle x', x \rangle \\ \langle x', x+y \rangle &= \langle x', x \rangle + \langle x', y \rangle \\ \langle x', \lambda x \rangle &= \lambda \langle x', x \rangle\end{aligned}$$

Si $x = \sum_{i=1}^n \lambda_i e_i$ écrire $\langle x', x \rangle = \dots$

Chercher une base de E^* . Montrer $\dim E^* = \dim E$.

Démontrer $(\langle x', x \rangle = 0 \quad \forall x) \Rightarrow (x' = 0)$ et $(\langle x', x \rangle = 0 \quad \forall x') \Rightarrow (x = 0)$

Dans la notation $\langle \square, 0 \rangle$ on injecte dans le carré $\square$ l'opérateur et dans le rond 0 l'opéré.

Montrer que la lecture arabe de $\langle x', x \rangle$ ré-écrite à l'occidentale, $\langle x, x' \rangle$ permet de définir un isomorphisme de E sur le dual de E^* , le bidual de E , noté E^{**} .

Deuxième partie : Soit E de dimension n sur K

Soit $\sigma \in \text{End } E$, σ nilpotent d'ordre k :

$\sigma^k = 0$ et k est le + petit entier qui ...

- Montrer qu'il existe $a \in E$ et $a' \in E^*$ t.q.

$\sigma^{k-i}(a) \neq 0$ et $\langle a', \sigma^{k-i}(a) \rangle \neq 0$

- On considère E_1 le sous-espace engendré par les vecteurs $a, \sigma(a), \dots, \sigma^{k-i}(a)$ et E'_1 le sous-espace engendré par $a', \sigma'(a'), \dots, \sigma'^{k-i}(a')$ où $\sigma' : E^* \longrightarrow E^*$ est définie par :

$$\langle \sigma'(x'), x \rangle = \langle x', \sigma(x) \rangle$$

(σ' s'appelle la transposée de σ , notée ${}^t\sigma$).

• Montrer que $\dim E_1 = \dim E'_1 = k$.

[Si $\alpha, \beta \in \text{End } E$, on regardera ${}^t(\alpha + \beta)$ et ${}^t(\alpha_0\beta)$]

- Si F est un sous-espace de E on note F^0 l'ensemble des $x' \in E^*$ tels que $\langle x', x \rangle = 0 \quad \forall x \in F$.

(De même si F' est un espace de E^* on note :

$$F'^0 = \{x \in E / \langle x', x \rangle = 0 \quad \forall x' \in F'\}$$

• Montrer que F^0 et F'^0 sont des sous-espaces.

• Montrer que si σ laisse F stable ($\sigma(F) \subset F$) alors ${}^t\sigma$ laisse F^0 stable.

- Montrer que E est somme directe de E_1 et de $(E'_1)^0$.

- En déduire qu'on peut casser E en somme directe de morceaux E_i stables par σ à E_i soit de la forme

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

COMMENTAIRE :

- Comme $\sigma^{k-1} \neq 0$, il existe au moins un $a \in E$ t.q. $\sigma^{k-1}(a) \neq 0$.

Ensuite, le couplage $\langle, \rangle$ étant non dégénérée $\langle x, x \rangle = 0 \forall x'$ entraîne $x = 0$.

Si $x = \sigma^{k-1}(a)$, on doit donc trouver un $a' \in E^*$ t.q. $\langle a', \sigma^{k-1}(a) \rangle \neq 0$

- a) Si $\lambda_0 a + \lambda_1 \sigma(a) + \dots + \lambda_{k-1} \sigma^{k-1}(a) = 0$

on en déduit :

$$\lambda_0 \sigma(a) + \lambda_1 \sigma^2(a) + \dots + \lambda_{k-2} \sigma^{k-1}(a) = 0 \text{ car } \sigma^k = 0.$$

En réappliquant σ on arrive au bout à $\lambda_0 \sigma^{k-1}(a) = 0$ ce qui entraîne $\lambda_0 = 0$.

En appliquant maintenant σ^{k-2} on trouve $\lambda_1 = 0$ etc. Le système est libre et forme donc une base.

b) Si $\mu_0 a' + \mu_1 {}^t\sigma.a' + \dots + \mu_{k-1} {}^t\sigma^{k-1}.a' = 0$ on raisonne de la même façon : pour cela il faut : $\sigma'^k.a' \neq 0$ et $\sigma'^k = 0$

or $\langle a', \sigma'^{k-1}.a \rangle = \langle \sigma'^{k-1}a', a \rangle \neq 0$ car $({}^t\sigma)^n = {}^t\sigma^n$

il en résulte $\sigma'^{k-1}.a' \neq 0$

La relation précédente donne aussi ${}^t\sigma^k = 0$.

En général, on a les relations

$${}^t(\alpha + \beta) = {}^t\alpha + {}^t\beta \text{ et } {}^t(\alpha \circ \beta) = {}^t\beta \circ {}^t\alpha$$

Regardons la deuxième :

$$\begin{aligned} \langle {}^t(\alpha \circ \beta).x', x \rangle &= \langle x', (\alpha \circ \beta)(x) \rangle \\ &= \langle x', \alpha(\beta(x)) \rangle \\ &= \langle {}^t\alpha(x').\beta(x) \rangle \\ &= \langle {}^t\beta({}^t\alpha(x')), x \rangle \\ &= \langle {}^t\beta \circ {}^t\alpha.x', x \rangle \forall x \end{aligned}$$

ce qui donne la relation.

Par itération on a $({}^t\sigma)^n = {}^t(\sigma^n)$.

Si x' et y' sont dans F^0 on a $\langle x', x \rangle = 0 \forall x \in F$ et $\langle y', x \rangle = 0 \forall x \in E$. D'où $\langle \lambda x' + \mu y', x \rangle = \lambda \langle x', x \rangle + \mu \langle y', x \rangle = 0 \forall x \in E : F^0$ est un sous-espace.

De même pour F^0 .

- Soit F t.q $\sigma(F) \subset F$. Soit $x' \in F^0$. Nous devons montrer que $\sigma.x' \in F^0$ c.-à-d.

$$\langle \sigma.x', x \rangle = 0 \quad \forall x \in F$$

or $\langle \sigma.x', x \rangle = \langle x', \sigma(x) \rangle = 0$ car $x \in F$ et donc $\sigma(x)$ aussi et $x' \in F^0$.

- Nous devons montrer que :

$$E = E_1 + (E'_1)^0 \text{ et } E_1 \cap (E'_1)^0 = 0.$$

. Montrons que $E_1 \cap (E'_1)^0 = 0$.

Soit $x = \lambda_0 a + \lambda_1 \sigma.a + \dots + \lambda_{k-1} \sigma^{k-1}.a$ un élément de E_1 qui est dans l'orthogonal de E'_1 c.-à-d.

$$\langle \sigma^i a', x \rangle = 0 \quad \forall i \text{ de } 0 \text{ à } k-1$$

Pour $i = k-1$

$$\begin{aligned} \langle \sigma^{k-1}, x \rangle &= \lambda_0 \langle a', \sigma^{k-1}.a \rangle + \lambda_1 \langle a', \sigma^k.a \rangle + \dots \\ &= \lambda_0 c + 0 = 0 \text{ avec } c \neq 0 \end{aligned}$$

D'où $\lambda_0 = 0$

Pour $i = k-2$ on trouve maintenant $\lambda_1 = 0$ etc : tous les λ_i sont nuls.

. On a $\dim F^0 = \dim E - \dim F$.

Soit $e_1 \dots e_r$ une base de F qu'on complète en une base de E , $e_{r+1}, \dots, e_n$.

Soit $e'_1, \dots, e'_n$ la base duale de (e_i) :

$$\langle e'_i, e_j \rangle = \delta_{ij}$$

On a donc $e'_{r+1}, \dots, e'_n \in F^0$ car $x' \in F^0 \iff \langle x', e_i \rangle = 0 \quad \forall i \leq r$

On a donc $\dim F^0 \geq n - r$.

Dans le cas général, on a égalité. Mais ici, le renseignement suffit car

$$\begin{aligned} \dim (E_1 + (E'_1)^0) &= \dim E_1 + \dim (E'_1)^0 \\ &\geq r + n - r = n \end{aligned}$$

On a donc nécessairement égalité car $E_1 + (E'_1)^0$ est un sous-espace de E .

Tout est donc démontré

- Le sous-espace $(E'_1)^0$ étant stable par σ on peut recommencer le raisonnement précédent avec la restriction de σ à $(E'_1)^0$ qui est toujours nilpotente, l'exposant "k" minimum pouvant être plus petit.

On peut donc écrire $(E'_1)^0 = E_2 \oplus (E'_2)^0$

où E_2 est engendré par $b, \sigma.b, \dots, \sigma^{l-1}.b$ et ainsi de suite :

$$\begin{aligned} E &= E_1 \oplus R_1 & R_1 &= (E'_1)^0 \\ E &= (E_1 \oplus E_2) \oplus R_2 & R_2 &= (E'_2)^0 \\ E &= (E_1 \oplus E_2 \oplus E_3) \oplus R_3 \end{aligned}$$

Ce processus s'arrête, car les E_i sont par construction $\neq 0$. Donc, à cause de la dimension finie de E , il y a un moment où nécessairement $R = 0$.

Tous les sous-espaces sont stables par σ .

Sur E_1 , σ a comme matrice

$$\sigma|_{E_1} = \begin{pmatrix} 0 & 0 & \dots & 0 \\ 1 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & \dots & 1 & 0 \end{pmatrix}$$

La matrice de σ par rapport à une base obtenue en choisissant dans chaque sous-espace E_i une base du type précédent.

Nous avons ainsi réalisé une "décomposition de Jordan".

Pour terminer, montrons que $\dim F^0 = \dim E - \dim F$, où F est quelconque. Avec les mêmes notations, soit

$$x' = \lambda_1 e'_1 + \dots + \lambda_r e'_r + \lambda_{r+1} e'_{r+1} + \dots + \lambda_n e'_n \text{ un élément de } F^0.$$

$$x' \in F^0 \iff \langle x', e_i \rangle = 0 \quad \forall i \leq r.$$

On en déduit que $x' = \lambda_{r+1} e'_{r+1} + \dots + \lambda_n e'_n$ c.-à-d. F^0 est engendré par $e'_{r+1} \dots e'_n$ et $\dim F^0 = n - r$.

IREM de LYON

BIBLIOTHEQUE

Université Claude Bernard - LYON I

43, Bd du 11 Novembre 1918

69622 VILLEURBANNE Cedex

I.R.E.M - FACULTE DES SCIENCES -
LA BOULOIE - ROUTE DE GRAY
25030 BESANÇON CEDEX
81 66 61 91 81 66 61 92

TITRE	LE THEOREME DE CAYLEY HAMILTON
AUTEUR	Jean Merker
PUBLIC CONCERNE	Personnes interessées par l'algèbre
NIVEAU	Bac
MOTS CLES	Matrices, Polynômes caractéristique et minimal, Matrices de Jordan
RESUME	Le prétexte de la présente brochure est un cours de télé-enseignement de premier cycle. Y était au programme le théorème de Cayley-Hamilton, qui se trouve expliqué ici en dégageant le plus possible les raisons générales cachées derrière le théorème, ceci de la façon la plus agréable possible.
DATE	février 92
FORMAT	A4
NOMBRE DE PAGES	50

IREM de Besançon.

Dépôt légal 76/92

1^{er} trimestre 92