Algèbre Linéaire 2

Avant-propos.

Ce document sert de polycopié pour le cours "Algèbre Linéaire 2", basé sur le cours "Algèbre Multilinéaire" donné avant 2012, et que j'ai donné depuis 2010. La première année j'ai donné le cours en rédigeant des notes manuscrites ; la rédaction du cours sous forme informatique a eu lieu essentiellement en automne 2011, et a apporté des changements importants à son organisation. Le cours avait été donné les deux années précédentes respectivement par mes collègues Jean Souville et par Rupert Yu, et je leur suis reconnaissant de m'avoir transmis les documents qu'ils ont préparés pendant ce temps. Les deux exemples m'ont montré qu'il existe un vaste éventail de possibilités pour présenter la matière, pourtant classique, qui est l'objet de ce cours.

J'ai cherché à trouver une présentation personnelle qui tienne compte à la fois du niveau limité des connaissances acquises dans le cours d'algèbre linéaire de la première année universitaire (qui est commun aux différents parcours de sciences et mathématiques), et de la nécessité pour les étudiants de mathématiques de se familiariser avec des notions plus structurelles en algèbre (comme les sous-espaces, choix de bases, homomorphismes, quotients). La maîtrise de ces notions est réputée plus difficile que celles des méthodes calculatoires, mais j'ai essayé de limiter le degré d'abstraction utilisé. Si certaines définitions sont formulées dans un cadre général (espaces vectoriels sur un corps commutatif K, idéaux dans $K[X]$, déterminant d'une matrice carrée à coefficients dans un anneau commutatif), c'est parce que cette généralité s'avère directement utile dans le cours.

Je considère comme but principal de ce cours d'aborder les questions autour des problèmes des vecteurs propres, et plus généralement les notions de "réduction d'endomorphismes" (une expression consacrée en France ; je n'ai jamais compris dans quel sens un endomorphisme peut être réduit). Les polynômes et déterminants figurent dans ce cours principalement comme outils pour atteindre ce but, la notion de polynôme caractéristique touchant de façon évidente aux deux sujets. Malgré son rôle utilitaire, le traitement des déterminants est assez complet, et il peut être considéré comme but secondaire du cours ; cela se justifie non seulement par le titre du cours, mais aussi par le fait qu'aucun cours ultérieur ne traitera les déterminants en profondeur, pendant qu'on se servira d'eux à diverses occasions. Pour les polynômes la situation est différente, et pour cette raison leur traitement se limite aux connaissances utiles pour ce cours. Je regrette d'ailleurs que le chapitre 3 fasse une interruption si importante des considérations de l'algèbre linéaire, et pour parler des choses qu'on pourrait supposer déjà familières, mais il est nécessaire d'assurer un fondement solide sur les polynômes, pour pouvoir en faire un nombre d'applications (notamment les polynômes d'un endomorphisme et la définition de son polynôme minimal). J'ai profité de l'occasion pour renforcer dans ce chapitre d'autres connaissances de base qui à un moment ou un autre doivent être acquises, notamment concernant les structures quotient.

L'organisation du cours de 2010–2011 était basée sur l'idée de ne pas attendre l'introduction des déterminants (et donc du polynôme caractéristique) avant d'aborder la recherche des valeurs propres, le faisant dans un premier temps à l'aide du polynôme minimal. Mais cette approche me semble maintenant une erreur pédagogique : le fait que la définition complète du polynôme caractéristique demande plus de préparation que celle du polynôme minimal (et que son degré est parfois plus élevé) ne semble en rien dissuader les étudiants de l'utiliser (même avant qu'il ne soit introduit dans le cours !). Pour cette raison l'introduction du polynôme minimal est maintenant reportée au dernier chapitre du cours.

En rédigeant, le texte est devenu beaucoup plus long que mes notes manuscrites ne l'étaient, ou que je ne l'avais imaginé au départ. Je reconnais avoir un style de discours peu succinct, et être incapable d'écrire dans le style définition–théorème–preuve souvent trouvé dans les textes français. Aussi de nombreuses remarques que je jugeait importantes de faire quelque part se sont glissées dans le texte au fur et à mesure. Je souligne que le texte est fait pour expliquer, et non pas pour être appris par cœur. J'espère que le lecteur saura reconnaître les quelques énoncés importants à retenir tels quels (souvent marqués "théorème") ; en cas de doute, un résumé des objectifs du cours est donné à la fin. Sinon, divers livres sous le titre *Algèbre Linéaire* consultables à la BU peuvent servir de référence complémentaire.

septembre 2013 *Marc van Leeuwen*

Introduction.

Ce cours est une continuation du cours de l'algèbre linéaire de la première année. On développera les techniques de l'algèbre linéaire avec le but notamment d'étudier les *endomorphismes* d'un espace vectoriel (toujours supposé de dimension finie), c'est-à-dire les applications linéaires de l'espace vers lui-même, qui dans une base de l'espace peuvent être exprimées par une matrice *carrée*. Dans cette étude les *valeurs propres* d'un endomorphisme, et les vecteurs propres associés, joueront un rôle important. Pour la recherche effective de ces valeurs propres, on aura besoin de certains *polynômes*, notamment du polynôme caractéristique d'un endomorphisme, dont la définition fait intervenir l'opération du *déterminant*. Un chapitre du cours sera alors dédié à une courte introduction aux polynômes du point de vue algébrique (par opposition à l'étude des fonctions polynomiales), et un autre au développement général de la notion du déterminant (une notion abordée en première année, mais dont le traitement était probablement limité aux aspects calculatoires et aux matrices de petite taille). Le dernier chapitre du cours fera une synthèse des notions abordés dans le cadre de l'étude des endomorphismes.

Chapitre 1. Rappels de l'algèbre linéaire.

Dans ce premier chapitre on fera un rappel des notions de l'algèbre linéaire qui ont été introduites (ou devraient l'avoir été) en première année. Mais si en première année il y avait un accent sur le calcul vectoriel et matriciel (dont la maîtrise est bien sûr une condition nécessaire pour la compréhension de l'algèbre linéaire), on se rendra vite compte que cela ne suffit pas pour une compréhension théorique, et que pour progresser on aura besoin d'une approche plus conceptuelle. Loin d'être redondant pour ceux qui ont réussi le cours de première année, ce chapitre servira à revoir les notions de ce point de vue, dont l'expérience montre qu'il est en général considéré comme plus difficile à appréhender.

1.1. *Espaces vectoriels, sous-espaces, combinaisons linéaires, applications linéaires.*

Dans l'algèbre linéaire on commence par fixer une fois pour toute un ensemble K de *scalaires*, qui sont des valeurs numériques qui peuvent être additionnées, soustraites, multipliées et divisées entre elles (à l'exception de la division par 0 qui n'est pas définie), et par lesquelles les vecteurs pourront être multipliés. La raison de cette abstraction est que différents ensembles (plus précisément ; différents corps commutatifs) de scalaires peuvent être utilisés sans que cela change la description de l'algèbre linéaire ; on peut penser notamment au corps des nombres rationnels $\mathbf{Q}$, à celui des nombres réels $\mathbf{R}$, et au corps des nombres complexes $\mathbf{C}$.

Un espace vectoriel E sur K (ou un K-espace vectoriel pour faire court) est un ensemble de valeurs mathématiques appelées vecteurs, qui peuvent être additionnés et soustraits entre eux, ainsi que multipliés par des scalaires dans K. La nature précise des vecteurs n'est pas spécifiée : ils peuvent être des fonctions, ou des suites, des polynômes, des solutions de systèmes d'équations, ou d'autres objets encore. Ce qui rend possible l'étude de toute cette diversité de possibilités au même temps, est que l'algèbre linéaire s'intéresse uniquement aux opérations indiquées (addition, multiplication scalaire), et aux relations que peuvent être exprimés en termes de ces opérations. Il faudra bien sûr supposer que quelques propriétés de base (appelées des axiomes) soient vérifiés, pour pouvoir raisonner en toute généralité. Ces propriétés sont notamment que les opérations sont définies pour tous les arguments du bon type (deux vecteurs peuvent toujours être additionnés, sans exception, et le résultat sera un vecteur), et que certaines égalités sont toujours vérifiées (comme la loi distributive $\lambda(v+w) = \lambda v + \lambda w$ pour $\lambda \in K$ et $v, w \in E$) ; la liste complète de ces axiomes est longue mais bien connue, et on ne la redonnera pas ici.

Si on a une collection de vecteurs $v_1, \ldots, v_l \in E$ et des scalaires $\lambda_1, \ldots, \lambda_l \in K$, on peut former la *combinaison linéaire* $\lambda_1 v_1 + \cdots + \lambda_l v_l$, qui est un vecteur de E.* Un *sous-espace vectoriel* de E est un sous-ensemble S contenant le vecteur nul $\vec{0}$, et qui est fermé pour l'addition (la somme de deux vecteurs

* Une subtilité est que parfois, en parlant de "combinaison linéaire", on désigne l'expression $\lambda_1 v_1 + \cdots + \lambda_l v_l$ elle-même plutôt que le vecteur qu'elle désigne ; notamment une "combinaison linéaire non triviale" est une telle expression avec au moins un des λ_i non nul, ce qui ne veut pas dire qu'elle désigne un vecteur non nul.

de S est toujours dans S) et pour la multiplication scalaire (un multiple scalaire d'un vecteur de S est toujours dans S). Il en découle que S est aussi fermé pour les combinaisons linéaires (si $v_1, \dots, v_l \in S$ alors on a aussi $\lambda_1 v_1 + \cdots + \lambda_l v_l \in S$, quels que soient les scalaires $\lambda_1, \dots, \lambda_l \in K$), et les sous-espaces vectoriels sont les seuls sous-ensembles non vides de E qui sont fermés pour les combinaisons linéaires.

La notion de sous-espace vectoriel est fondamentale, d'une part parce qu'ils sont eux-mêmes des espaces vectoriels (comme le suggère leur nom), mais aussi parce qu'ils interviennent par exemple dans la description des solutions d'un système d'équations linéaires (ou encore de l'image et le noyau d'une application linéaire). Certaines descriptions en algèbre peuvent être plus efficaces en termes d'ensembles, comme ici les sous-espaces vectoriels, qu'en termes d'éléments (vecteurs) individuels; toutefois, il est à noter qu'un même ensemble peut avoir plusieurs descriptions, dont parfois aucune n'est privilégiée. Un sous-espace vectoriel admet deux types de description principaux: une description par des *générateurs*, qui sont des vecteurs particuliers du sous-espace tels que tous ses vecteurs en sont des combinaisons linéaires, et une description par des *équations* (linéaires homogènes), le sous-espace étant l'ensemble de vecteurs qui vérifient l'ensemble de ces équations. Une description par générateurs peut être écrite $S = \mathrm{Vect}(v_1, \dots, v_l)$ ou $v_1, \dots, v_l$ sont les générateurs; c'est par définition l'ensemble de toutes les combinaisons linéaires $\lambda_1 v_1 + \cdots + \lambda_l v_l$. On vérifie facilement qu'un tel ensemble $\mathrm{Vect}(v_1, \dots, v_l)$ est toujours un sous-espace vectoriel, et il contient chacun des générateurs $v_1, \dots, v_l$. En fait c'est le plus petit tel sous-espace : un sous-espace vectoriel qui contient $v_1, \dots, v_l$ contiendra forcément $\mathrm{Vect}(v_1, \dots, v_l)$ tout entier.

Une autre notion fondamentale est celle d'une application linéaire entre des K-espaces vectoriels. (C'est sans doute pour traiter de façon uniforme un grand nombre d'opérations très disparates comme la différentiation et intégration de fonctions, évaluation de polynômes, rotations et projections dans l'espace, mais qui ont le caractère *linéaire* en commun, que l'algèbre linéaire comme discipline abstraite a été développée.) Si E, E' sont deux K-espaces vectoriels, une application $f : E \to E'$ est dite (K-)linéaire si elle vérifie

$$\begin{aligned} f(x+y) &= f(x) + f(y) \\ f(\lambda x) &= \lambda f(x) \end{aligned} \tag{1}$$

pour tout $x, y \in E$ et tout $\lambda \in K$. Une façon alternative de caractériser les application linéaires est qu'elles sont compatibles avec les combinaisons linéaires :

$$f(\lambda_1 v_1 + \cdots + \lambda_l v_l) = \lambda_1 f(v_1) + \cdots + \lambda_l f(v_l) \tag{2}$$

pour tout $v_1, \dots, v_l \in E$ et tout $\lambda_1, \dots, \lambda_l \in K$. Un exemple typique d'une application linéaire est l'opération de la combinaison linéaire elle-même, avec une liste de vecteurs fixée, qu'on décrit ainsi. Fixons des vecteurs $\vec{v}_1, \dots, \vec{v}_l$, dans un K-espace qu'on appelle E' car ce sera l'espace d'arrivée. L'espace de départ sera K^l, dont les éléments sont les l-uplets $(c_1, \dots, c_l)$ de scalaires ; comme on le sait bien, c'est un K-espace vectoriel avec addition et multiplication scalaire composante par composante :

$$\begin{aligned} (c_1, \dots, c_l) + (d_1, \dots, d_l) &= (c_1 + d_1, \dots, c_l + d_l) \quad \text{et} \\ \lambda(c_1, \dots, c_l) &= (\lambda c_1, \dots, \lambda c_l). \end{aligned} \tag{3}$$

L'application de combinaison linéaire $f : K^l \to E'$ est définie par $f : (c_1, \dots, c_l) \mapsto c_1 \vec{v}_1 + \cdots + c_l \vec{v}_l$. On vérifie facilement qu'elle vérifie les identités de (1) ; par exemple pour la première identité on a

$$\begin{aligned} f : ((c_1, \dots, c_l) + (d_1, \dots, d_l)) &= f(c_1 + d_1, \dots, c_l + d_l) = (c_1 + d_1)\vec{v}_1 + \cdots + (c_l + d_l)\vec{v}_l \\ &= (c_1 \vec{v}_1 + \cdots + c_l \vec{v}_l) + (d_1 \vec{v}_1 + \cdots + d_l \vec{v}_l) = f(c_1, \dots, c_l) + f(d_1, \dots, d_l). \end{aligned}$$

Dans les identités qui caractérisent les applications linéaires $E \to E'$, les additions et les multiplications scalaires dans les seconds membres ont lieu dans l'espace vectoriel E'. La notion d'application linéaire n'a besoin d'aucune relation particulière entre les espaces E et E' (sauf qu'ils utilisent le même corps de scalaires K), mais il est très bien possible que les deux espaces soient en fait le même. Ce cas particulier $E = E'$ est d'un intérêt particulier, et on appelle alors f un *endomorphisme* du K-espace vectoriel E. L'étude des endomorphismes de E est un sujet majeur de ce cours. Par rapport aux applications linéaires en général, le cas des endomorphismes est relativement simple dans la mesure où il n'y a qu'un seul type de vecteurs à considérer, mais on verra qu'il est aussi plus riche, à cause des relations qui sont possibles entre les vecteurs avant et après l'application de f.

1.2. *Familles génératrices (d'un sous-espace), liées ou libres ; bases, dimension (finie).*

On appellera $[v_1, \ldots, v_l]$ une famille génératrice du sous-espace S si on a $\mathrm{Vect}(v_1, \ldots, v_l) = S$ (et en particulier c'est une famille génératrice de l'espace E tout entier si $\mathrm{Vect}(v_1, \ldots, v_l) = E$: tout vecteur s'écrit comme combinaison linéaire de vecteurs de la famille). (Un mot sur la terminologie: une famille de vecteurs se distingue d'un ensemble de vecteurs par le fait que ses membres ont chacun une place fixe dans la famille, de sorte que $[x, y, z]$ ne soit pas la même famille que $[y, z, x]$, et par le fait que plusieurs membres peuvent être égaux. Pour une famille finie (le seul type considéré dans ce cours) on pourrait lire "liste" au lieu de "famille". On a choisi de noter les familles de vecteurs entre crochets, pour éviter la confusion possible avec les éléments de K^n notés avec parenthèses.)

La description $\mathrm{Vect}(v_1, \ldots, v_l)$ d'un sous-espace par générateurs n'est pas unique. Si l'on permute les vecteurs dans la liste, ou si on remplace un vecteur v_i par un multiple λv_i avec $\lambda \neq 0$, ou si on remplace v_i par celui $v_i + \lambda v_j$ obtenu par l'addition d'un multiple d'un *autre* membre v_j de la famille, le sous-espace $\mathrm{Vect}(v_1, \ldots, v_l)$ engendré par la famille ne change pas; en itérant ces opérations on peut obtenir des familles totalement différentes de celle du départ. On ne cherchera donc pas *la (meilleure)* description d'un sous-espace par générateurs.

Mais dans une expression $\mathrm{Vect}(v_1, \ldots, v_l)$, on peut au moins se demander si l'un des générateurs n'est pas redondant, c'est-à-dire que l'espace engendré ne changerait pas si l'on enlevait ce générateur. Si l'un des générateur s'écrit comme combinaison linéaire des *autres* générateurs, par exemple $v_l \in \mathrm{Vect}(v_1, \ldots, v_{l-1})$, alors ce générateur est redondant: $\mathrm{Vect}(v_1, \ldots, v_l) = \mathrm{Vect}(v_1, \ldots, v_{l-1})$. La raison est que dans une combinaison linéaire L de $v_1, \ldots, v_l$ on pourra remplacer v_l par une certaine combinaison linéaire de $v_1, \ldots, v_{l-1}$, et en simplifiant l'expression obtenir une combinaison linéaire de $v_1, \ldots, v_{l-1}$ qui donne le même vecteur que L. Si c'est la cas pour l'un (au moins) des générateurs v_i, on dira que la famille $[v_1, \ldots, v_l]$ de vecteurs est *liée*.

Supposons que le générateur v_i soit ainsi redondant. On peut alors prendre son expression comme combinaison linéaire des autres générateurs, et y rajouter un terme "$-v_i$", pour trouver une combinaison linéaire de $v_1, \ldots, v_l$ dont la valeur est le vecteur nul $\vec{0}$, mais quelle combinaison linéaire est non-triviale (car au moins le coefficient -1 de v_i n'est pas nul). Une telle combinaison linéaire est appelée un *relation linéaire non triviale* entre les vecteurs $v_1, \ldots, v_l$. Si réciproquement est donnée une telle relation linéaire non triviale R, alors on peut prendre l'un de ses termes dont le coefficient n'est pas nul (il y en a au moins un), disons $\lambda_i v_i$ avec $\lambda_i \neq 0$, écrire ensuite $\lambda_i v_i = -S$ où S est la somme des autres termes de R, et finalement $v_i = \frac{-1}{\lambda_i} S$ ce qui montre que $v_i \in \mathrm{Vect}(v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n)$, et donc que la famille $[v_1, \ldots, v_l]$ est liée. On a donc trouvé une autre caractérisation des familles liées (souvent donnée comme définition): une famille de vecteurs est liée s'il existe une relation linéaire non triviale entre ses vecteurs.

Le cas contraire, où aucun des vecteurs $v_1, \ldots, v_l$ ne s'exprime comme combinaison linéaire des autres vecteurs de la famille, est plus intéressant. Dans ce cas on dit que $[v_1, \ldots, v_l]$ est une *famille libre*. La deuxième description des famille liées permet de caractériser les familles libres $[v_1, \ldots, v_l]$ par la propriété que *la seule* combinaison linéaire $\lambda_1 v_1 + \cdots + \lambda_l v_l$ dont la valeur est $\vec{0}$ est la combinaison triviale, celle avec $\lambda_1 = \cdots = \lambda_l = 0$. Une description $S = \mathrm{Vect}(v_1, \ldots, v_l)$ d'un sous-espace par générateurs est donc non redondante si $[v_1, \ldots, v_l]$ est une famille libre. Un vecteur v ne peut alors s'écrire que d'une façon au plus comme combinaison linéaire de vecteurs d'une famille libre donnée : s'il y avait deux écritures *distinctes*, leur soustraction produirait une combinaison *non-triviale* valant $\vec{0}$, ce qui n'existe pas.

Une *base* d'un sous-espace S est une famille libre et génératrice de S. C'est donc précisément une famille $[v_1, \ldots, v_l]$ telle que $S = \mathrm{Vect}(v_1, \ldots, v_l)$ soit une description non redondante de S par générateurs. En particulier, une base de E est une famille libre $[v_1, \ldots, v_l]$ de vecteurs qui est aussi génératrice de E. La famille étant génératrice, tout vecteur $x \in E$ s'écrit comme combinaison linéaire $x = \lambda_1 v_1 + \cdots + \lambda_l v_l$ de vecteurs de la famille, et la famille étant libre, cette écriture sera unique. Les scalaires λ_i qui figurent dans cette écriture s'appellent les *coordonnées* de x dans la base $[v_1, \ldots, v_l]$.

Si E admet une famille finie de générateurs, alors c'est un résultat fondamental de l'algèbre linéaire que *toute* description non redondante de E par générateurs fait intervenir *le même nombre* de générateurs. Ce nombre est appelé la *dimension* de E et noté $\dim(E)$. Ces générateurs forment précisément une base de E ; le résultat dit donc que toutes les bases de E contiennent le même nombre $\dim(E)$ de vecteurs.

(Il faut savoir qu'il existe des espaces vectoriels qui n'admettent aucune famille finie de générateurs, et donc en particulier pas de base finie. Ceci arrive notamment pour des espaces définis de façon très générale, comme ceux de toutes les fonctions différentiables $\mathbf{C} \to \mathbf{C}$, de toutes le suites infinies de scalaires, ou encore de tous les polynômes en X (à coefficients dans K). Ces espaces, dits de dimension infinie, ont des propriétés bien plus compliquées que ceux de dimension finie, et ils ne seront pas étudiés dans ce cours. Mais ils ont néanmoins leur utilité comme source de sous-espaces de dimension finie, comme par exemple ceux formés des solutions de certaines équations.)

Rappelons comment le résultat fondamental mentionné est démontré, car cela révèle à la fois une technique générale très utile, et un théorème supplémentaire.

L'observation de départ est que, si $[v_1, \ldots, v_i]$ est une famille libre et v_{i+1} un vecteur supplémentaire (candidat pour rejoindre la famille), précisément une des deux situations suivantes se produit : soit $[v_1, \ldots, v_i, v_{i+1}]$ est toujours un famille libre (on peut inclure v_{i+1}), soit $v_{i+1} \in \mathrm{Vect}(v_1, \ldots, v_i)$ ce qui entraîne $\mathrm{Vect}(v_1, \ldots, v_i) = \mathrm{Vect}(v_1, \ldots, v_i, v_{i+1})$ (donc v_{i+1} est redondant comme générateur). Car s'il existe une relation linéaire non triviale $\lambda_1 v_1 + \cdots + \lambda_i v_i + \lambda_{i+1} v_{i+1} = 0$, alors forcément $\lambda_{i+1} \neq 0$ (sinon on obtiendrait une relation linéaire non triviale entre les vecteurs de la famille $[v_1, \ldots, v_i]$, ce qui est impossible car celle-ci est libre) et on a vu qu'un vecteur (ici v_{i+1}) dont le coefficient dans une relation linéaire est non nul est une combinaison linéaire des autres vecteurs intervenant dans la relation.

Ainsi on a une méthode systématique pour étendre une famille libre de façon qu'elle engendre un plus grand sous-espace, tout en restant libre : on considère de nouveaux candidats un à un, en ne gardant que ceux qui ne sont pas (déjà) dans le sous-espace engendré par la famille en son état actuel. Ainsi on montre le résultat suivant, qui est une forme (en dimension finie) du théorème "de la base incomplète".

1.2.1. Théorème. *Soit E un espace vectoriel possédant une famille génératrice $G = [w_1, \ldots, w_g]$ (donc $E = \mathrm{Vect}(w_1, \ldots, w_g)$), et $L = [v_1, \ldots, v_l]$ une famille libre de vecteurs de E, alors il existe une base de E de la forme $B = [v_1, \ldots, v_l, w_{i_1}, \ldots, w_{i_k}]$, c'est-à-dire obtenue en rajoutant à L certains vecteurs de G.*

Preuve. L'ensemble $S = \{i_1, \ldots, i_k\}$ des indices des vecteurs de G retenus dans la base est donné par $i \in S \iff w_i \notin \mathrm{Vect}(v_1, \ldots, v_l, w_1, \ldots, w_{i-1})$. On montrera pour $i = 0, 1, \ldots, k$ que la partie L_i de la famille B, contenant tous les vecteurs v ainsi que les vecteurs w dont l'indice est $\leq i$, est une base de $V_i = \mathrm{Vect}(v_1, \ldots, v_l, w_1, \ldots, w_i)$. Pour $k = g$ on obtiendra le résultat cherché, car $L_g = B$ et $V_g = E$. Raisonnant par récurrence, c'est vrai pour $i = 0$ car $L_0 = L$ est libre et engendre V_0. Pour $i > 0$ on a soit $w_i \in V_{i-1}$, auquel cas $i \notin S$ donc $L_i = L_{i-1}$ et $V_i = V_{i-1}$, soit on a $w_i \notin V_{i-1}$, auquel cas $i \in S$ donc $L_i = L_{i-1} \cup \{w_i\}$ est libre, et une base de $\mathrm{Vect}(L_{i-1}) + \mathrm{Vect}(w_i) = V_{i-1} + \mathrm{Vect}(w_i) = V_i$. □

Plusieurs de choses découlent directement de ce théorème. D'abord, si un espace vectoriel admet une famille génératrice finie G, il admet aussi une base finie (car dans tout espace vectoriel la famille vide $L = [\,]$ est libre). Alors, dire que E admet une famille génératrice finie revient au même que dire que E est "de dimension finie" (admet un base finie), et c'est cette dernière formulation qui est le plus souvent employé (même si au moment de prouver ce théorème on ne sait pas encore que la dimension est bien définie). Ensuite, du moins si on se limite aux espaces vectoriels de dimension finie, le théorème dit que toute famille libre peut être complétée pour devenir une base de l'espace (d'où le nom du théorème). Dans cette formulation la famille génératrice G n'est plus mentionnée explicitement, il suffit qu'une telle famille existe. Comme toutefois la base concrète obtenue dépend bel et bien du choix de G, le fait d'invoquer le théorème sous cette forme revient à choisir une parmi les multiples façons de compléter L. C'est justement souvent son intérêt, pour surmonter "l'embarras du choix" dans certaines constructions : étendre une famille libre à une base permet souvent de spécifier ensuite une construction de façon précise.

1.2.2. Lemme. *Supposons la situation du théorème 1.2.1, où on suppose en plus que G soit (une famille libre donc) une base de E. Alors la base B obtenue contient le même nombre g de vecteurs que G.*

Preuve. Par récurrence sur le nombre l d'éléments de L. Si $l = 0$ alors $B = G$, parce que la condition $w_i \notin \mathrm{Vect}(w_1, \ldots, w_{i-1})$ est toujours satisfaite (car G est libre), donc $S = \{1, \ldots, g\}$. Soit $l > 0$, et comparons l'ensemble S dans la démonstration avec celui, appelons le S', obtenu dans le cas où L est remplacé par $L' = [v_1, \ldots, v_{l-1}]$, et dont on appelle B' la base obtenue (par hypothèse de

récurrence B' contient g vecteurs). On a $S \subseteq S'$ car $\mathrm{Vect}(L, w_1, \ldots, w_{i-1}) \supseteq \mathrm{Vect}(L', w_1, \ldots, w_{i-1})$ pour tout $i > 0$. Il reste à montrer que S' contient précisément un indice i qui est absent de S (c'est-à-dire $i \in S' \setminus S$), car cela indique la disparition dans B d'un vecteur de G qui compense alors l'apparition du vecteur v_l dans L (on aura $|B| = |B'| = g$). Pour $i \in S' \setminus S$ on a $w_i \notin \mathrm{Vect}(L', w_1, \ldots, w_{i-1})$ mais $w_i \in \mathrm{Vect}(L, w_1, \ldots, w_{i-1}) = V_{i-1}$, donc w_i est combinaison linéaire des vecteurs de la base L_{i-1} de V_{i-1}, et dans cette combinaison v_l doit avoir un coefficient non nul. Cela permet d'écrire à son tour v_l comme combinaison linéaire des *autres* vecteurs de $L_{i-1} \cup \{w_i\}$, combinaison dans laquelle w_i a un coefficient non nul. Mais ces vecteurs appartiennent tous à la base B' de E, et il existe une expression *unique* de v_l comme une combinaison linéaire C des vecteurs de B', et c'est celle-ci qu'on a trouvée. En comparant, on voit que i doit être égal au dernier indice j tel que w_j ait un coefficient non nul dans C. En plus le fait que L est libre montre qu'un tel indice existe : les coefficients non nuls de C ne peuvent pas tous être associés à un $v_j \in L'$. On a ainsi établi l'existence et l'unicité de $i \in S' \setminus S$. □

1.2.3. Théorème/définition. *Si un espace vectoriel admet une famille génératrice finie, alors E admet aussi une base finie ; dans ce cas E est dit de dimension finie. Il existe alors un nombre $d \in \mathbf{N}$, appelé dimension de E et noté $d = \dim(E)$, tel que toute base de E est constituée de d vecteurs.*

Preuve. Montrons que si G est une base de E formée de d vecteurs, alors toute base B de E est constituée de d vecteurs. Si on applique le théorème 1.2.1 avec $L = B$ et G, il rend B comme base (la condition pour avoir $i \in S$ est impossible à vérifier, donc $S = \emptyset$). Le lemme 1.2.2 dit alors que B contient d vecteurs. □

En comparant les théorèmes 1.2.1 et 1.2.3, on voit qu'aucune famille libre ne peut comprendre *plus* de $\dim(E)$ vecteurs (car elle peut être complétée à une base de E), et qu'aucune famille génératrice de E ne peut comprendre *moins* de $\dim(E)$ vecteurs (car on peut en extraire une base de E).

1.3. *Somme de sous-espaces, somme directe.*

1.3.1. Définition. *La somme de deux sous-espaces V, W d'un espace vectoriel E est*

$$V + W = \{\, v + w \mid v \in V, w \in W \,\},$$

et plus généralement la somme d'une collection de sous-espaces de E est l'ensemble de vecteurs de E qui peuvent être écrits comme une somme de vecteurs, un vecteur étant pris dans chacun de ces sous-espaces.

On voit facilement qu'une somme de sous-espaces de E est de nouveau un sous-espace de E. Il découle aussi directement de la définition que $\mathrm{Vect}(v_1, \ldots, v_k) + \mathrm{Vect}(w_1, \ldots, w_l) = \mathrm{Vect}(v_1, \ldots, v_k, w_1, \ldots, w_l)$. (On a déjà clandestinement utilisé une somme de sous-espaces à la fin de la preuve du théorème 1.2.1.)

Supposons que les familles $[v_1, \ldots, v_k]$ et $[w_1, \ldots, w_l]$ soient libres, c'est-à-dire qu'elles soient bases de $V = \mathrm{Vect}(v_1, \ldots, v_k)$ respectivement de $W = \mathrm{Vect}(w_1, \ldots, w_l)$. Il n'en résulte néanmoins pas que $[v_1, \ldots, v_k, w_1, \ldots, w_l]$ est libre, car la définition d'une famille libre porte sur l'ensemble de ses membres, et cette propriété ne peut pas être validée par petit groupe (ni d'ailleurs "deux à deux" pour toutes les paires de vecteurs). On peut bien sûr obtenir une base de $V + W$ en supprimant dans la famille $[v_1, \ldots, v_k, w_1, \ldots, w_l]$ les vecteurs w_i pour lesquels $w_i \in \mathrm{Vect}(v_1, \ldots, v_k, w_1, \ldots, w_{i-1})$. Mais en faisant cela, le nombre de vecteurs restants, et donc $\dim(V + W)$, dépendra de la situation précise ; selon les cas on pourra trouver pour $\dim(V + W)$ tout nombre entre $\max(k, l)$ et $k + l$. Cette incertitude motive la définition d'un cas spécial d'une somme de sous-espaces qui sera d'une utilité particulière.

1.3.2. Définition. *Une somme de sous-espaces de E est dite directe quand chaque vecteur de la somme s'écrit d'une façon unique comme une somme de vecteurs pris dans chacun de ces sous-espaces.*

La notation $V_1 \oplus V_2 \oplus \cdots \oplus V_k$ pour une somme indique qu'on affirme que cette somme est directe.

1.3.3. Théorème. *Une somme $V = V_1 + V_2 + \cdots + V_k$ de sous-espaces d'un espace E de dimension finie est directe si et seulement si $\dim V = \dim(V_1) + \dim(V_2) + \cdots + \dim(V_k)$.*

Preuve. On observe d'abord que si un vecteur s'écrit de deux manières différentes comme une somme de vecteurs pris dans chacun de ces sous-espaces, la différence de ces expressions donne une écriture

$v_1 + \cdots + v_k = \vec{0}$ avec $v_i \in V_i$ pour tout i, dans lequel $v_i \neq \vec{0}$ pour au moins un indice i. La somme de V_i sera donc directe si une telle écriture "non triviale" de $\vec{0}$ n'existe pas. On choisit pour chaque i des base $\mathcal{B}_i$ de V_i, et on enchaîne ces bases pour former une seule famille F de $\dim(V_1) + \cdots + \dim(V_k)$ vecteurs, qui est génératrice de V. On montrera que F est libre (et donc un base de V) si et seulement si la somme est directe. Supposons d'abord qu'il existe une écriture non triviale $v_1 + \cdots + v_k = \vec{0}$; en exprimant chaque v_i dans la base $\mathcal{B}_i$ et en faisant leur somme, on obtient une relation linéaire non triviale entre les vecteurs de F, donc F est lié. Réciproquement si F est lié, on prend une relation linéaire non triviale entre ces vecteurs, dans lequel on regroupe les termes concernant les vecteurs de $\mathcal{B}_i$ pour tout i. Le somme $v_i \in V_i$ du groupe de termes associé à $\mathcal{B}_i$ est non nul dès que l'un de ses coefficients est non nul (car $\mathcal{B}_i$ est une famille libre), et cala se produit au moins une fois ; l'écriture $\vec{0} = v_1 + \cdots + v_k$ est donc non triviale. En conclusion, si la somme est directe, alors $\dim V = \dim(V_1) + \cdots + \dim(V_k)$, et si elle n'est pas directe alors $\dim V < \dim(V_1) + \cdots + \dim(V_k)$ (car au moins un des vecteurs de la famille génératrice F de V est redondant) ; le théorème s'ensuit. □

On fera attention au fait que la notion de somme directe, à l'instar de celle de famille libre, porte sur l'ensemble (de sous-espace concernés), et en particulier ne se vérifie pas deux à deux. Par exemple l'espace $\mathbf{R}^2$ contient une infinité de sous-espaces différents de dimension 1 (ce sont toutes les droites passant par l'origine), qui forment des sommes directes deux à deux (en prenant une base à un seul vecteur dans chacune de deux droites, les deux forment toujours une famille libre), mais la somme d'au moins trois d'entre eux n'est jamais directe, car la dimension de la somme ne peut pas dépasser la dimension 2 de $\mathbf{R}^2$.

En vue de cela il est un peu surprenant que la notation pour les somme directes est bien associative. Pas seulement on a pour disons trois sous-espaces U, V, W de E que $(U + V) + W = U + V + W = U + (V + W)$ comme on vérifie facilement (que les sommes soient directes ou non), mais aussi on a le droit d'écrire cette somme comme $(U \oplus V) \oplus W$ si et seulement si on peut l'écrire $U \oplus V \oplus W$ si et seulement si on peut l'écrire $U \oplus (V \oplus W)$. L'explication est que $(U \oplus V) \oplus W$ affirme que la somme de U et de V est directe, et que la somme de $U + V$ (tout entier!) et de W est directe ; cette dernière affirmation est bien plus forte que de dire que les sommes $U + W$ et $V + W$ sont chacune directes. Que $(U \oplus V) \oplus W$ affirme la même chose que $U \oplus V \oplus W$ est facile à montrer : s'il y avait une expression $u + v + w = \vec{0}$ avec au moins un de u, v, w non nul, alors $(u + v) + w = \vec{0}$, et ou bien l'un de $(u + v)$ et w est non nul contredisant que la somme de $U + V$ et W est directe, ou bien ils sont nuls mais l'un du u, v est non nul contredisant que la somme de U et V est directe ; l'argument dans le sens opposé est similaire.

Grâce à cette associativité, le cas particulier d'une somme de deux sous-espaces seulement est assez intéressant pour justifier qu'on énonce la caractérisation suivante de quand une telle somme est directe.

1.3.4. Proposition. *La somme $V + W$ de deux sous-espaces est directe si et seulement si $V \cap W = \{\mathbf{0}\}$.*

Preuve. Si $V \cap W = \{\vec{0}\}$ et $v + w = 0$ avec $v \in V, w \in W$, alors $v = -w \in V \cap W$; donc $v, w = 0$, la somme est directe. Si $V \cap W \neq \{\vec{0}\}$ on écrit $0 = x + -x$ avec $0 \neq x \in V \cap W$, la somme n'est pas directe. □

On peut affiner ce résultat en décrivant $\dim(V + W)$ même pour les sommes qui ne sont pas directe.

1.3.5. Proposition. *Pour V, W de dimension finie on a* $\dim(V) + \dim(W) = \dim(V + W) + \dim(V \cap W)$.

Preuve. C'est une première occasion d'utiliser le théorème de la base incomplète. On commence à choisir une base $[u_1, \ldots, u_a]$ de $V \cap W$, qu'on étend (à l'aide d'une famille génératrice de V) à une base $[u_1, \ldots, u_a, v_1, \ldots, v_b]$ de V, et aussi (à l'aide d'une famille génératrice de W) à une base $[u_1, \ldots, u_a, w_1, \ldots, w_c]$ de W. La famille $[u_1, \ldots, u_a, v_1, \ldots, v_b, w_1, \ldots, w_c]$ est certainement génératrice de $V + W$; si on la montre libre, la proposition sera démontré car $(a + b) + (a + c) = (a + b + c) + a$. Considérons une relation linéaire entre les vecteurs $u_1, \ldots, u_a, v_1, \ldots, v_b, w_1, \ldots, w_c$. Regroupons la termes de ce combinaison linéaire en trois groupes, donnant une combinaison linéaire u de $u_1, \ldots, u_a$, une combinaison linéaire v de $v_1, \ldots, v_b$, et une combinaison linéaire w de $w_1, \ldots, w_c$, qui vérifient $u + v + w = 0$. On a $w = -(u + v) \in W \cap V$, donc w est (aussi) combinaison linéaire de $u_1, \ldots, u_a$ (une base de $V \cap W$). Mais $[u_1, \ldots, u_a, w_1, \ldots, w_c]$ est une famille libre, donc les deux combinaisons linéaires doivent

être nuls, et donc $w = 0$. Mais alors $u + v = 0$ est une relation linéaire entre les membres de la famille libre $[u_1, \ldots, u_a, v_1, \ldots, v_b]$, donc elle est triviale, et par conséquent la relation linéaire initiale aussi. □

1.4. *Expression dans une base, matrices d'applications linéaires.*

On a déjà observé que pour une famille de vecteurs $[v_1, \ldots, v_n]$ quelconque, on a une application linéaire $f : K^n \to E$ définie par la formation de combinaisons linéaires, c'est-à-dire $f(c_1, \ldots, c_n) = c_1 v_1 + \cdots + c_n v_n$.

1.4.1. Proposition. *Cette formation de combinaisons linéaires d'une famille finie de vecteurs est*
(1) une application surjective si et seulement si la famille est génératrice de E,
(2) une application injective si et seulement si la famille est libre,
(3) une application bijective si et seulement si la famille est une base de E.

Preuve. Le point (1) traduit directement la définition d'une famille génératrice. Pour le point (2), si l'application est injective, il ne peut pas y avoir de relation linéaire non triviale (avec la relation triviale elle donnerait deux combinaisons linéaires avec la même valeur nulle). Et si au contraire l'application est non injective, on pourra (comme indiqué avant) soustraire deux combinaisons linéaires avec la même valeur, pour trouver une relation linéaire non triviale, et la famille sera liée. Le point (3) résulte de (1) et (2). □

Pour une base $\mathcal{B} = [b_1, \ldots, b_n]$, l'application de formation de combinaisons linéaires est donc inversible, et c'est la raison principale pour l'importance de la notion d'une base. Il sera commode d'avoir dans ce cas une notation raccourcie pour ces combinaisons linéaires, donc on définit

$$(x_1, \ldots, x_n)_{\mathcal{B}} = x_1 b_1 + \cdots + x_n b_n \qquad \text{si } \mathcal{B} \text{ est une base de } E.$$

L'application réciproque $E \to K^n$, qui à $v \in E$ associe l'unique n-uplet $(\lambda_1, \ldots, \lambda_n)$ telle que $v = (x_1, \ldots, x_n)_{\mathcal{B}}$, est appelée l'expression de vecteurs dans la base $\mathcal{B}$, et les scalaires $x_1, \ldots, x_n$ s'appellent les *coordonnées* de v dans la base $\mathcal{B}$. Les applications linéaires et bijectives sont aussi appelées *isomorphismes* de K-espaces vectoriels ; ici on a donc un isomorphisme $K^n \to E$ et un isomorphisme réciproque $E \to K^n$. L'importance des isomorphismes vient du fait que, à l'aide des isomorphismes réciproques $E \to E'$ et $E' \to E$, toute question d'algèbre linéaire concernant l'espace E peut être traduit en une question équivalente concernant E', et vice versa. Par définition tout K-espace E de dimension n possède (au moins) une base de n éléments, qui donne lieu à une paire d'isomorphismes réciproques $E \leftrightarrow K^n$. Ce fait explique l'importance des espaces K^n comme exemples archétypiques de K-espaces de dimension finie.

Dans un espace vectoriel K^n, on n'a pas besoin d'une base pour associer à chaque vecteur un n-uplet des scalaires, car dans ce cas les vecteurs *sont* déjà des n-uplets des scalaires. Toutefois, on peut considérer ces scalaires comme les coordonnées du vecteur dans une base bien particulière appelée la *base canonique* $[e_1, \ldots, e_n]$ de K^n. Pour savoir quelle est cette base, on peut observer qu'en général pour un vecteur b_i pris dans la base $\mathcal{B}$ elle même, ses coordonnées dans la base $\mathcal{B}$ sont toutes nulles sauf la coordonnée à la position i qui est 1. Comme chaque vecteur de K^n doit être égal au n-uplet de ces coordonnées dans la base canonique, on a $e_i = (0, \ldots, 0, 1, 0, \ldots, 0)$, le n-uplet dont toutes les composantes sont nuls sauf celle à la position i qui est 1. Par exemple pour $n = 3$ on a $e_1 = (1, 0, 0)$, $e_2 = (0, 1, 0)$, et $e_3 = (0, 0, 1)$.

Contrairement aux espaces K^n, les autres espaces vectoriels de dimension finie ne sont pas en général munis d'une base préférée. Pour avoir une représentation numérique des vecteurs dans ces espaces, on est obligé de fixer une base d'abord. Mais cette obligation peut aussi être un avantage pour résoudre certains problèmes d'algèbre linéaire, car la possibilité de choisir une base bien adaptée au problème spécifique permet souvent de rendre la description plus transparente.

Considérons maintenant les applications linéaires $f : E \to E'$. Selon les cas, elles peuvent être définies d'une manière qui utilise la nature particulière (fonction, suite, ...) des éléments de E et de E'. À titre d'exemple, pour $K = \mathbf{R}$ et avec E égal l'espace de polynômes réels quadratiques en X, et $E' = \mathbf{R}$ (un espace vectoriel de dimension 1), on pourra considérer l'application linéaire $f : E \to \mathbf{R}$ qui consiste à évaluer les polynômes en le nombre $\sqrt{2}$, c'est-à-dire $f : P[X] \mapsto P[\sqrt{2}]$ (c'est un exercice facile de montrer que c'est effectivement une application linéaire). Mais pour pouvoir appliquer les méthodes de l'algèbre linéaire, on a besoin d'un type de description qui soit indépendante de la nature des espaces E, E'. Pour cela on pourra se servir de bases de ces espaces (s'il sont de dimension finie). La propriété suivante est d'une importance fondamentale pour la description des applications linéaires.

1.4.2. Proposition. *Si $\mathcal{B} = [b_1, \dots, b_m]$ est une base de E, alors une application linéaire $f : E \to E'$ est entièrement déterminée par la donnée de ses valeurs dans les vecteurs de la base $\mathcal{B}$, c'est-à-dire par la famille $[f(b_1), \dots, f(b_m)]$ de vecteurs de E'. Réciproquement, si l'on donne une famille $[v_1, \dots, v_m]$ de vecteurs de E', il existe une application linéaire $f : E \to E'$ telle que $f(b_i) = v_i$ pour $i = 1, \dots, n$.*

Preuve. Il suffit d'utiliser la compatibilité d'une application linéaire avec les combinaisons linéaires : l'égalité $f(\lambda_1 b_1 + \cdots + \lambda_m b_m) = \lambda_1 f(b_1) + \cdots + \lambda_m f(b_m)$ donne la valeur de $f(v)$ d'un vecteur quelconque $v = (\lambda_1, \dots, \lambda_n)_{\mathcal{B}} \in E$. Pour la réciproque, l'application $(\lambda_1, \dots, \lambda_n)_{\mathcal{B}} \mapsto \lambda_1 v_1 + \cdots + \lambda_m v_m$ convient. □

Pour une description explicite d'une application linéaire $f : E \to E'$ en termes de scalaires seulement, on aura, en plus de la base $\mathcal{B}$ de E, aussi besoin d'une base $\mathcal{B}' = [b'_1, \dots, b'_n]$ de E'. On pourra alors exprimer les vecteurs $f(b_1), \dots, f(b_m)$ en coordonnées dans cette base, le tout formant une matrice.

1.4.3. Définition. *La matrice d'une application linéaire $f : E \to E'$, par rapport à un couple de bases $\mathcal{B} = [b_1, \dots, b_m]$ de E et $\mathcal{B}' = [b'_1, \dots, b'_n]$, est la matrice $n \times m$*

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,m} \\ a_{2,1} & a_{2,2} & \dots & a_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \dots & a_{n,m} \end{pmatrix},$$

dont la colonne j contient les coordonnées dans la base $\mathcal{B}'$ de l'image par f du j-ème vecteur de base de $\mathcal{B}$, c'est-à-dire $f(b_j) = (a_{1,j}, \dots, a_{m,j})_{\mathcal{B}'}$, pour $j = 1, \dots, m$.

On fera attention au fait qu'en donnant la taille d'une matrice, on mentionne le nombre de lignes avant le nombre de colonnes, et que ce nombre de lignes est la dimension de l'espace *d'arrivée*.

La définition décrit la matrice d'une application linéaire f donnée, mais la pratique est également importante pourvoir décrire réciproquement l'application f à partir de sa matrice A (toujours par rapport à des bases de E et E' fixées). On trouve (en utilisant la linéarité de f) que

$$f\big((x_1, \dots, x_m)_{\mathcal{B}}\big) = (y_1, \dots, y_n)_{\mathcal{B}'} \qquad \text{où } y_i = a_{i,1}x_1 + \cdots + a_{i,m}x_m \text{ pour } i = 1, \dots, n. \tag{4}$$

L'opération de cette application f peut être réalisée en trois étapes : d'abord on exprime le vecteur $v \in E$ auquel f est appliqué en coordonnées (l'isomorphisme $E \to K^m$ déterminé par la base $\mathcal{B}$), puis on transforme le m-uplet de coordonnées $(x_1, \dots, x_m)$ en le n-uplet de coordonnées $(y_1, \dots, y_n)$ à l'aide de la matrice A, et on transforme finalement ce n-uplet en un vecteur de E' par combinaison linéaire des vecteurs de la base $\mathcal{B}'$ (l'isomorphisme $K^n \to E'$ déterminé par la base $\mathcal{B}'$). Seulement l'étape du milieu dépend de A, et elle ne dépend de rien d'autre. En effet, cette application linéaire $L_A : K^m \to K^n$ est celle dont A est la matrice *par rapport aux bases canoniques* de K^m et de K^n.

On peut illustrer la situation par un diagramme, d'un type que s'avère très utile. Toute flèche désigne une application linéaire, et s'il existe plusieurs façons de passer d'un endroit à un autre en suivant des flèches, les résultats des compositions d'applications linéaires correspondantes seront égaux (on parle de diagrammes commutatifs). Le choix d'orienter les flèches horizontales de droite à gauche est lié au fait qu'on écrit les applications à gauche de leurs arguments.

$$\begin{array}{ccc} f(v) \in E' & \xleftarrow{f} & v \in E \\ \Big\updownarrow{\scriptstyle \mathcal{B}'} & & \Big\updownarrow{\scriptstyle \mathcal{B}} \\ (y_1, \dots, y_n) \in K^n & \xleftarrow{L_A} & (x_1, \dots, x_m) \in K^m \end{array} \tag{5}$$

L'application L_A est déterminée par (4) ; en écrivant les listes d'éléments de K verticalement on a

$$L_A : \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} \mapsto \begin{pmatrix} a_{1,1}x_1 + \cdots + a_{1,m}x_m \\ a_{2,1}x_1 + \cdots + a_{2,m}x_m \\ \vdots \\ a_{n,1}x_1 + \cdots + a_{n,m}x_m \end{pmatrix}. \tag{6}$$

Cette opération est prise comme définition de la multiplication à gauche d'une "colonne" par A :

$$A \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} = \begin{pmatrix} a_{1,1} & \dots & a_{1,m} \\ a_{2,1} & \dots & a_{2,m} \\ \vdots & \ddots & \vdots \\ a_{n,1} & \dots & a_{n,m} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} \overset{\text{déf}}{=} \begin{pmatrix} a_{1,1}x_1 + \dots + a_{1,m}x_m \\ a_{2,1}x_1 + \dots + a_{2,m}x_m \\ \vdots \\ a_{n,1}x_1 + \dots + a_{n,m}x_m \end{pmatrix}. \tag{7}$$

La raison de présenter L_A comme opérant sur une colonne, par la multiplication à gauche de celle-ci par A, est que ce produit se généralise au produit matriciel. Celui-ci est définie de façon à correspondre à la composition d'applications linéaire. Concrètement, supposons qu'en plus de cette matrice A de taille $n \times m$ on ait une autre matrice B de taille $m \times l$ pour un certain l, qui détermine une application linéaire $L_B : K^l \to K^m$. On peut alors former la composée $L_A \circ L_B : K^l \to K^n$ (dans une composée c'est toujours l'application écrite à droite, ici L_B, qui agit en premier : $(L_A \circ L_B)(x) = L_A(L_B(x))$). La composition d'applications linéaires donne toujours une application linéaire, et on peut donc exprimer $L_A \circ L_B$ par une matrice $n \times l$ par rapport aux bases canoniques de K^l et de K^n, quelle matrice sera par définition le produit matriciel $A \cdot B$, autrement dit on aura $L_{A \cdot B} = L_A \circ L_B$.

Pour déterminer les coefficients de ce produit $A \cdot B$, on utilise la définition 1.4.3 : ses colonnes sont formées par les images des vecteurs de la base de départ, exprimés dans la base d'arrivée. Comme les bases considérées ici sont toutes canoniques, l'expression sur la base d'arrivée est une opération sans effet, et on peut simplement dire que la k-ème colonne d'une matrice M *est égale à* l'image $L_M(e_k)$ du k-ème vecteur de la base canonique de l'espace de départ. En particulier k-ème colonne de $A \cdot B$ est égale à l'image $L_{A \cdot B}(e_k) = (L_A \circ L_B)(e_k) = L_A(L_B(e_k))$ du vecteur e_k de la base canonique de K^l. Or, le vecteur $L_B(e_k) \in K^m$ est égal à la k-ème colonne de la matrice B. Il suffit donc d'utiliser (6) pour L_A appliqué à $L_B(e_k)$, c'est-à-dire en prenant pour les coefficients x_j les coefficients $b_{j,k}$ de la k-ème colonne de la matrice B, pour $j = 1, \dots, m$. On obtient la définition du produit matriciel :

$$\begin{pmatrix} a_{1,1} & \dots & a_{1,m} \\ a_{2,1} & \dots & a_{2,m} \\ \vdots & \ddots & \vdots \\ a_{n,1} & \dots & a_{n,m} \end{pmatrix} \cdot \begin{pmatrix} b_{1,1} & b_{1,2} & b_{1,3} & \dots & b_{1,l} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c_{m,1} & c_{m,2} & c_{m,3} & \dots & c_{m,l} \end{pmatrix} \overset{\text{déf}}{=} \begin{pmatrix} c_{1,1} & c_{1,2} & c_{1,3} & \dots & c_{1,l} \\ c_{2,1} & c_{2,2} & c_{2,3} & \dots & c_{2,l} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c_{n,1} & c_{n,2} & c_{n,3} & \dots & c_{n,l} \end{pmatrix}, \tag{8}$$

où les coefficients du produit sont donnés par $c_{i,k} = a_{i,1}b_{1,k} + \dots + a_{i,m}b_{m,k} = \sum_{j=1}^{m} a_{i,j}b_{j,k}$ pour $i = 1, 2, \dots, n$ et $k = 1, 2, 3, \dots, l$. Dans $A \cdot B = C$, chaque colonne de C ne dépend que de la colonne correspondante de B et que chaque ligne de C ne dépend que de la ligne correspondante de A.

Le produit matriciel peut également être utilisé pour décrire la composition d'applications linéaires dont les matrices sont données par rapport à des bases autres que les bases canoniques d'espaces de la forme K^n. Pour cela il est essentiel que pour l'expression des matrices on se serve dans chacun des espaces concernés toujours d'une même base. Le diagramme est alors

$$\begin{array}{ccccc} f(g(v)) \in E' & \xleftarrow{f} & g(v) \in E' & \xleftarrow{g} & v \in E \\ \updownarrow \mathcal{B}'' & & \updownarrow \mathcal{B}' & & \updownarrow \mathcal{B} \\ (z_1, \dots, z_n) \in K^n & \xleftarrow{L_A} & (y_1, \dots, y_n) \in K^n & \xleftarrow{L_B} & (x_1, \dots, x_m) \in K^m \end{array} \tag{9}$$

L'utilisation de deux bases différentes dans un même espace est discutée dans la section suivante.

La matrice de l'identité $\mathrm{id}_E : E \to E$ (donnée par $\mathrm{id}_E(x) = x$ pour tout $x \in E$) par rapport à une base $\mathcal{B} = (b_1, \dots, b_n)$ de E utilisée au départ comma à l'arrivée, est toujours de la même forme. En effet sa colonne j contient les coordonnées de $\mathrm{id}_E(b_j) = b_j$ dans la base $\mathcal{B}$, et est donc égale au vecteur e_j de la base canonique de K^n. On trouve donc la "matrice identité" I_n, de taille $n \times n$ et avec des coefficients 1 partout sur la diagonale principale, et des coefficients 0 partout ailleurs.

Si $f : E \to E'$ est un isomorphisme d'espaces vectoriels (ce qui nécessite $\dim(E) = \dim(E')$) et $g : E' \to E$ l'isomorphisme réciproque, on a $g \circ f = \mathrm{id}_E$ et $f \circ g = \mathrm{id}_{E'}$. Si $\mathcal{B}$ est une base de E et $\mathcal{B}'$ une base de $\mathcal{B}'$, les matrices P de f par rapport à $\mathcal{B}, \mathcal{B}'$ et Q de g par rapport à $\mathcal{B}', \mathcal{B}$ vérifieront donc $Q \cdot P = I_n$ et $P \cdot Q = I_n$. On appelle ces matrices inverses l'une de l'autre, et qu'on note $Q = P^{-1}$ ou $P = Q^{-1}$. Une matrice qui possède un inverse est dit *inversible*. Pour qu'une matrice soit inversible il est nécessaire qu'elle soit carrée, mais ce n'est pas suffisant. Mais il est utile de savoir que si les matrices P et Q sont carrées et $P \cdot Q = I_n$, alors on aura automatiquement $Q \cdot P = I_n$, et donc $Q = P^{-1}$.

1.5. *Changement de base.*

L'existence de bases montre que tout K-espace vectoriel de dimension finie possède une structure assez simple, et que la dimension est la seule propriété distinctive de tels espaces. On dit que ces espaces sont classifiés (à isomorphisme près) par leur dimension: E est isomorphe à K^n pour $n = \dim(E)$. La contrepartie de la facilité de trouver des bases est que l'espaces ont beaucoup de symétrie : il existe plein de bases différentes pour un même espace E, et chacune apporte son propre isomorphisme $E \to K^n$.

La question se pose alors comment convertir des informations numériques données par rapport à des bases différentes. Le cas le plus simple est celui d'un vecteur donné par ses coordonnées dans une base $\mathcal{B}$, et dont on cherche les coordonnées dans une autre base $\mathcal{B}'$. Le diagramme suivant décrit la situation.

$$\begin{array}{ccc} E & \overset{\mathrm{id}_E}{\longleftrightarrow} & E \\ \updownarrow{\scriptstyle \mathcal{B}} & & \updownarrow{\scriptstyle \mathcal{B}'} \\ K^n & \underset{d}{\overset{c}{\rightleftarrows}} & K^n \end{array} \tag{10}$$

Les flèches en haut ne font rien, et sont là juste pour que le diagramme prenne la forme du diagramme (5). Les flèches c, d en bas donnent les conversions de coordonnées qui nous intéressent ; leur effet est par définition celle de passer passer d'abord en haut vers E (interprétation des coordonnées dans la base indiquée), puis repasser en bas en utilisant l'autre base (expression du vecteur obtenu en coordonnées). Il s'agit d'applications linéaires mutuellement réciproques (donc des isomorphismes) de l'espace K^n vers lui-même. Ces flèches correspondent donc (directement) à des matrices, inverses l'une de l'autre.

Une première question est laquelle des deux matrices associer au passage de l'ancienne base $\mathcal{B}$ au la nouvelle base $\mathcal{B}'$. Le diagramme suggère que c'est c, mais malheureusement la convention veut que c'est d. La conversion entre différentes d'unités de mesure peut servir comme modèle de la situation: la quantité reste la même, mais sa représentation numérique change. En effet on peut considérer une unité de mesure comme une base (à un seul vecteur) dans un espace de dimension 1 de quantités physiques (par exemples les longueurs), et la conversion d'unité est alors un changement de base. Dans une conversion d'unités, le facteur f le plus facilement associé à la conversion est celui qui exprime la nouvelle unité comme f anciennes unités. Par exemple au passage du franc français à l'euro on a dit que ce dernier valait 6,55957 francs (et non pas que le franc allait valoir 0,152449 euros). Mais pour convertir une quantité donnée comme x anciennes unités, il faut *diviser* x par de f, pour donner x/f nouvelles unités. Ainsi la matrice correspondant à la flèche c dans cette situation est $(1/f)$ (une matrice 1×1), et (f) correspond à la flèche d. Le cas de dimension n est similaire : la *matrice de passage* de la base $\mathcal{B}$ à la base $\mathcal{B}'$ est celle qui exprime les vecteurs de $\mathcal{B}'$ dans la base $\mathcal{B}$. Elle correspond a la flèche d ci-dessus.

1.5.1. Définition/Proposition. *Si $\mathcal{B}, \mathcal{B}'$ sont deux bases d'un espace vectoriel E de dimension n, la matrice de passage de $\mathcal{B}$ vers $\mathcal{B}'$ est la matrice $n \times n$ dont la colonne j contient les coordonnées dans la base $\mathcal{B}$ du j-ème vecteur de la base $\mathcal{B}'$. En multipliant une colonne contenant les coordonnées d'un vecteur $x \in E$ dans la base $\mathcal{B}'$ à gauche par cette matrice, on obtient les coordonnées de x dans la base $\mathcal{B}$.*

Sachant comment convertir les coordonnées d'un vecteur pour exprimer ce vecteur dans une autre base, l'effet d'un changement de base sur des matrices exprimées par rapport à cette bases est facile à décrire. Supposons qu'on connaisse la matrice A d'une application $f : E \to E'$ par rapport aux bases $\mathcal{B}$ de E et $\mathcal{B}'$ de E', et que $\mathcal{C}$ soit une autre base de E, telle que la matrice (inversible) P soit la matrice de passage de $\mathcal{B}$ à $\mathcal{C}$. Alors si l'on cherche la matrice A' de f par rapport aux bases $\mathcal{C}$ et $\mathcal{B}'$, on remarque que A' doit opérer sur une colonne de coordonnées exprimés dans la base $\mathcal{C}$. Il convient donc de convertir ces coordonnées en coordonnées sur la base $\mathcal{B}$, en multipliant la colonne par P, à quel point on peut appliquer la matrice A pour obtenir l'image par f, exprimée dans la base $\mathcal{B}'$; au total on a $A' = A \cdot P$. Par un même type de raisonnement, on voit qu'un changement de base à l'arrivée avec matrice de passage Q a pour effet sur la matrice une multiplication à gauche par la matrice Q^{-1} (car cette fois on obtient d'abord les coordonnées sur l'ancienne base de E', qu'il faut convertir en coordonnées sur la nouvelle base).

1.5.2. Proposition. *Si A est la matrice de $f : E \to E'$ par rapport aux bases $\mathcal{B}$ de E et $\mathcal{B}'$ de E', et si $\mathcal{C}$ et $\mathcal{C}'$ sont d'autres bases de E respectivement de E', et si P et Q sont les matrices de passage de $\mathcal{B}$ vers $\mathcal{C}$ respectivement de $\mathcal{B}'$ vers $\mathcal{C}'$, alors la matrice de f rapport aux bases $\mathcal{C}$ et $\mathcal{C}'$ sera $Q^{-1} \cdot A \cdot P$.* □

1.6. *Équivalence de matrices rectangulaires, image, noyau, et rang.*

Si une application linéaire est représentée par une matrice A par rapport à un couple de bases (une au départ et une à l'arrivée), et par une matrice B par rapport à un autre couple de bases, on dit que A et B sont des "matrices équivalentes". En vue de la proposition cela veut dire qu'il existe des matrices inversibles P, Q telles que $B = Q^{-1} \cdot A \cdot P$, et on en déduit qu'il s'agit en effet d'une relation d'équivalence. Deux matrices équivalentes sont évidemment de la même taille ; on verra ci-dessous que parmi les matrices d'une taille donnée il n'y a qu'un nombre fini de classes d'équivalence de matrices pour cette relation.

On peut associer à chaque application linéaire $f : E \to E'$ deux sous-espaces, l'un dans E' et l'autre dans E, qui donnent des renseignements importants concernant f. Dans E', il s'agit de l'image $\mathrm{Im}(f)$ se f, le sous-espace des vecteurs sont l'image par f d'au moins une vecteur de E ; en formule

$$\mathrm{Im}(f) = \{\, f(v) \mid v \in E \,\}.$$

Dans E il s'agit du noyau $\mathrm{Ker}(f)$ de f, le sous-espace des vecteurs que f envoie sur $\vec{0} \in E'$; en formule

$$\mathrm{Ker}(f) = \{\, v \in E \mid f(v) = \vec{0} \,\}.$$

Les caractérisations suivantes sont immédiates.

1.6.1. Proposition. *Soit $f : E \to E'$ une application linéaire, alors*
(1) f est surjectif si et seulement si $\mathrm{Im}(f) = E'$,
(2) f est injectif si et seulement si $\mathrm{Ker}(f) = \{0\}$. □

L'image et le noyau de f ne sont pas liés de façon directe, après tout ce sont des sous-espaces de différents espaces E', E. Mais il existe une relation fondamentale entre leurs *dimensions* et celle de E.

1.6.2. Théorème du rang. *Soit E un espace vectoriel de dimension finie, et $f : E \to E'$ linéaire, alors* $\dim(\mathrm{Ker}(f)) + \dim(\mathrm{Im}(f)) = \dim(E)$. *Le nombre* $\dim(\mathrm{Im}(f))$ *est appelé le rang* $\mathrm{rg}(f)$ *de f.*

Une méthode naturelle d'approcher cette preuve serait de choisir une base de $\mathrm{Ker}(f)$, la compléter à l'aide du théorème de la base incomplète à une base, et de montrer que l'image par f de la famille des vecteurs rajoutés forme une base de $\mathrm{Im}(f)$. Bien que cette méthode marche bien, on peut éviter certaines vérifications avec une approche astucieuse qui commence avec le choix d'une base à l'arrivée dans $\mathrm{Im}(f)$.

Preuve. On choisit une base $L = [w_1, \ldots, w_r]$ du sous-espace $\mathrm{Im}(f)$ de E', et ensuite des vecteurs $v_1, \ldots, v_r \in E$ tels que $f(v_i) = w_i$ pour tout i (que qui est possible d'après la définition de $\mathrm{Im}(f)$). On peur définir d'après la proposition 1.4.2 une application linéaire $g : \mathrm{Im}(f) \to E$ par $g(w_i) = v_i$, de sorte que $f(g(w)) = w$ pour tout $w \in \mathrm{Im}(f)$, et on pose $C = \mathrm{Im}(g)$. Comme $g(w) = g(w')$ entraîne $w = f(g(w)) = f(g(w')) = w'$, l'application g est injectif, et donne donc une bijection $\mathrm{Im}(f) \to C$ dont l'inverse est et la restriction de f à C. Le noyau de cet inverse est $\{0\}$ mais aussi $C \cap \mathrm{Ker}(f)$, donc la somme $\mathrm{Ker}(f) + C$ est directe (proposition 1.3.4). Or pour tout $v \in E$ on a $f(v) = f(g(f(v)))$ donc $v - g(f(v)) \in \mathrm{Ker}(f)$; l'écriture $v = (v - g(f(v))) + g(f(v))$ montre que $E = \mathrm{Ker}(f) \oplus C$, et le théorème 1.3.3 donne l'égalité $\dim(E) = \dim(\mathrm{Ker}(f)) + \dim(\mathrm{Im}(f))$ cherchée. □

Si A est la matrice de f par rapport aux bases $\mathcal{B}$ de E et $\mathcal{B}'$ de E', alors $\mathrm{rg}(f) = \mathrm{rg}(L_A)$: les isomorphismes réciproques associés à la base $\mathcal{B}'$ dans le diagramme

$$\begin{array}{ccc} E' & \xleftarrow{f} & E \\ \updownarrow{\scriptstyle \mathcal{B}'} & & \updownarrow{\scriptstyle \mathcal{B}} \\ K^n & \xleftarrow{L_A} & K^m \end{array}$$

font correspondre les sous-espaces $\mathrm{Im}(f)$ de E et $\mathrm{Im}(L_A)$ de K^n, quels sous-espaces ont donc la même dimension $\mathrm{rg}(f) = \mathrm{rg}(L_A)$. Or $\mathrm{Im}(L_A)$ ne dépend que de la matrice A, ce qui permet de définir $\mathrm{rg}(A) = \dim(\mathrm{Im}(L_A))$; ce nombre est égal à $\mathrm{rg}(f)$. Il est aussi égal au nombre d'une famille maximale *libre* de colonnes choisies parmi celle de A (car cela donne une base de $\mathrm{Im}(L_A) \subseteq K^n$). Le rang de deux matrices équivalentes est le même : c'est $\mathrm{rg}(f)$ si l'une et l'autre représentent f, chacune sur une base convenable.

1.6.3. Théorème. *Deux matrices sont équivalentes si et seulement si elles ont la même taille et la même rang. Pour $n, m, r \in \mathbf{N}$, il existe des matrices de taille $n \times m$ et de rang r si et seulement si $r \leq \min(n, m)$, dont un exemple est la matrice $M_{n,m,r}$ de taille $n \times m$ dont les seuls coefficients non nuls sont r coefficients* 1*, situés sur les r premières positions de la diagonale principale.*

Preuve. On a déjà vu que pour que deux matrices soient équivalentes il est nécessaire qu'elles aient la même taille $n \times m$ et le même rang r ; or on aura $r \leq \min(n, m)$ car $r = \dim(\mathrm{Im}(L_A))$ ne peut dépasser ni la dimension m de l'espace de départ de L_A (car L_A envoie une base de celui-ci sur une famille génératrice de $\mathrm{Im}(L_A)$), ni la dimension n de son espace d'arrivée (qui contient $\mathrm{Im}(L_A)$). Pour conclure, il suffit de montrer qu'une telle matrice A est équivalente à la matrice $M_{n,m,r}$. Soit donc $f : E \to E'$ une application linéaire dont A est la matrice par rapport à un certain couple de bases ; pour montrer que A est équivalente à la matrice $M_{n,m,r}$ il convient de trouver un autre couple de bases par rapport auquel la matrice de f est $M_{n,m,r}$. Dans la preuve du théorème 1.6.2 on a construit des familles libres $[v_1, \ldots, v_r]$ dans E et $[w_1, \ldots, w_r]$ dans E' (en fait une base de $\mathrm{Im}(f)$) telles que $f(v_i) = w_i$ pour tout i. Il suffit de compléter la première famille à une base de E en rajoutant une base de $\mathrm{Ker}(f)$ (on rappelle que $E = C \oplus \mathrm{Ker}(f)$) et de compléter la seconde famille (de façon quelconque) à une base de E' ; on vérifie que la matrice de A par rapport à de telles bases est $M_{n,m,r}$. □

1.7. *Endomorphismes, similitude de matrices carrées.*

La preuve du théorème 1.6.3 montre combien la liberté ce choisir des bases adaptées à une situation donnée est utile. La problématique principale de ce cours concerne, au lieu d'une application linéaire $f : E \to E'$ entre des espaces distincts, un endomorphisme de E : une application linéaire $f : E \to E$ vers l'espace E *lui-même*. On note $\mathrm{End}(E)$ l'ensemble des endomorphismes de l'espace vectoriel E. Dans ce contexte, il n'est plus raisonnable de choisir séparément une base au départ et à l'arrivée de f (les deux espaces étant confondus) : quand on parle d'une matrice d'un endomorphisme, on suppose toujours (sans mention explicite du contraire) que *la même* base est utilisée au départ et à l'arrivée. Cela est important entre autres si on veut itérer f pour obtenir ses puissances comme $f^3 : v \mapsto f(f(f(v)))$; si l'on veut que la matrice de f^n soit le produit matriciel A^n de n copies de la matrice A de f, il est essentiel (comme pour tout produit matriciel) que la base utilisée à l'arrivée d'une application linéaire (une des instances de f) coïncide avec celle utilisée au départ de la suivante.

On est donc amené à définir, pour les matrices carrées uniquement, une relation de *similitude* (qui est plus fine que celle d'équivalence) : deux matrices carrées A, B sont dites *semblables* si elles peuvent être obtenues comme les matrices d'un même endomorphisme de E, chacune par rapport à une base de E. Si la matrice de passage de la première base vers la seconde est P, on aura donc d'après la proposition 1.5.2 que $B = P^{-1} \cdot A \cdot P$. La relation d'être des matrices semblables est (aussi) une relation d'équivalence.

Pour illustrer que cette relation est plus fine que l'équivalence, considérons le cas $A = I_n$. Comme $P^{-1} \cdot I_n \cdot P = P^{-1} \cdot P = I_n$ pour toute matrice $n \times n$ inversible P, la matrice I_n n'est semblable à *aucune* autre matrice, bien qu'elle soit équivalente (d'après le théorème 1.6.3) à toute matrice $n \times n$ de rang n, c'est-à-dire à toute matrice inversible. Ce cas est bien sûr extrême (la plupart des classes de similitude contiennent plus qu'une seule matrice), mais il est indicatif de la différence entre les deux relations.

En fait on n'arrivera pas dans ce cours à décrire complètement la relation de similitude sur les matrices carrées. Mais la notions importantes du polynôme caractéristique d'une matrice (dont la définition est un but important de ce cours) et du polynôme minimal sont telles que la définition donne le même polynôme pour des matrices semblables. Même si l'égalité de leurs polynômes caractéristiques et de leurs polynômes minimaux ne sera pas une garantie de la similitude de deux matrices, il s'agit dans les deux cas de conditions nécessaires importantes.

Chapitre 2. Vecteurs propres, valeurs propres.

Dans cette section on supposera donné un espace vectoriel E de dimension finie, et un endomorphisme particulier ϕ de E. Il est possible que ϕ soit donné par une matrice par rapport à une base particulière de E, ou peut-être il est donné d'une autre façon. En tout cas, si on cherche a comprendre les propriétés de ϕ, elles ne sont pas forcément faciles à déduire de sa matrice sur une base quelconque, et une méthode importante consiste à chercher une description de ϕ qui rend ses propriétés plus transparentes.

Très généralement, la compréhension de ϕ sera facilité si on trouve des vecteurs et des sous-espaces qui ont des propriétés spéciales par rapport à ϕ. On connaît les sous-espaces $\operatorname{Ker}\phi$ et $\operatorname{Im}\phi$, tous deux de E, qui sont certainement spéciaux par rapport à ϕ. Mais cela ne donne rien d'intéressant pour la plupart des endomorphismes qui sont inversibles, car pour eux on a $\operatorname{Ker}\phi = \{0\}$ et $\operatorname{Im}\phi = E$. En revanche, le fait que ϕ est un endomorphisme permet de comparer les vecteurs v avec leur propre image $\phi(v)$, et de replacer le test $\phi(v) = 0$ fait dans la définition de $\operatorname{Ker}\phi$ par un test de dépendance linéaire du vecteur $\phi(v)$ avec v lui-même. Cela reste une condition exceptionnelle pour $v \neq 0$, mais on verra que souvent certain vecteurs la vérifient néanmoins. La dépendance linéaire de $\phi(v)$ avec $v \neq 0$ signifie l'existence d'un scalaire $\lambda \in K$ tel que $\phi(v) = \lambda v$. Cela nous me aux notions de vecteur propre et de valeur propre.

2.1. *Définition de vecteur propres et de valeur propres ; premières propriétés.*

2.1.1. Définition. *Soit E un K-espace vectoriel et $\phi \in \operatorname{End}(E)$. Si $v \in E$ est non nul et vérifie $\phi(v) = \lambda v$ pour un scalaire $\lambda \in K$, on appelle v un vecteur propre de ϕ et λ une valeur propre de ϕ.*

Pour un vecteur propre v, il existe précisément une valeur propre λ tel que $\phi(v) = \lambda v$, la valeur propre associé à v. Réciproquement v s'appelle un vecteur propre pour la valeur propre λ. Pour $\lambda \in K$ donné, les vecteurs propres pour λ, s'ils existent, sont les vecteurs non nuls dans le sous-espace $\operatorname{Ker}(\phi - \lambda \operatorname{id}_E)$ de E (en particulier ils ne sont pas uniques). Cet espace $\operatorname{Ker}(\phi - \lambda \operatorname{id}_E)$ est appelé l'*espace propre* de λ (il contient évidemment le vecteur nul, même si 0 n'est pas un vecteur propre). Il est parfois utile d'admettre la phrase "espace propre de λ" même si λ n'est peut-être pas une valeur propre; si en fait il ne l'est pas, son "espace propre" $\operatorname{Ker}(\phi - \lambda \operatorname{id}_E)$ est réduit à $\{0\}$, ne contenant donc aucun vecteur propre.

On voit que la notion d'espace propre est relatée à celle du noyau, mais qu'on soustrait de ϕ le multiple $\lambda \operatorname{id}_E$ de l'identité avant de prendre le noyau. (On appelle l'endomorphisme $\lambda \operatorname{id}_E$ l'*homothétie* de E de facteur λ.) La possibilité de choisir λ convenablement augmente nos chances de trouver un noyau non réduit à $\{0\}$, et la recherche de tels valeurs (propres) λ sera importante dans l'investigation de ϕ.

Les espaces propres sont des exemples de sous-espaces ϕ-stables, qui sont des sous-espaces avec la propriété spéciale suivante par rapport à ϕ.

2.1.2. Définition. *Un sous-espace F de E est ϕ-stable pour $\phi \in \operatorname{End}(E)$ si $\phi(F) \subseteq F$.*

Les sous-espaces ϕ-stables F sont ceux pour lesquels on peut considérer l'effet de ϕ juste sur F (on parle de restreindre ϕ à un endomorphisme de F). Si F est un sous-espace non ϕ-stable, on peut bien restreindre ϕ à une application linéaire $F \to E$, mais cela ne donne pas un endomorphisme de F ni de E. La restriction de ϕ à l'espace propre de ϕ pour λ est l'homothétie de cet espace de facteur λ. Réciproquement tout sous-espace ϕ-stable F tel que la restriction de λ à F soit une homothétie de facteur λ est contenu dans l'espace propre de ϕ pour λ. Et tout sous-espace ϕ-stable de dimension 1 est contenu dans (et souvent égal à) un espace propre, car en dimension 1 les seuls endomorphismes possibles sont les homothéties. Donc si tous les espaces propres sont de dimension 1, ils forment aussi l'ensemble des sous-espaces ϕ-stables de E de dimension 1. Mais si ϕ est une homothétie (une possibilité à l'extrême opposé), *tous* les sous-espaces (qu'ils soient de dimension 1 ou non) seront ϕ-stables.

2.1.3. Proposition. *Si $\lambda_1, \dots, \lambda_k \in K$ sont des valeurs distinctes, la somme de leurs espaces propres $\operatorname{Ker}(\phi - \lambda_1 \operatorname{id}_E) + \cdots + \operatorname{Ker}(\phi - \lambda_k \operatorname{id}_E)$ est toujours directe.*

C'est un premier résultat important, même si on le remplacera plus tard par un énoncé plus fort ; il montre d'emblée, en utilisant le théorème 1.3.3, que le nombre de valeurs propres de ϕ ne peut jamais dépasser $\dim E$. On en donnera deux preuve différentes ; chacune exploite le fait que les λ_i sont distincts.

Preuve. On raisonne par récurrence sur k. Pour $k \leq 1$ il n'y a rien à montrer. Soit donc $k \geq 2$, et posons $V = \mathrm{Ker}(\phi - \lambda_1 \mathrm{id}_E) \oplus \cdots \oplus \mathrm{Ker}(\phi - \lambda_{k-1} \mathrm{id}_E)$ (somme directe par hypothèse de récurrence); par associativité de '$\oplus$' il suffira de démontrer que la somme $V + \mathrm{Ker}(\phi - \lambda_k \mathrm{id}_E)$ est directe. Pour cela on montrera $V \cap \mathrm{Ker}(\phi - \lambda_k \mathrm{id}_E) = \{0\}$; supposons donc que $v \in V \cap \mathrm{Ker}(\phi - \lambda_k \mathrm{id}_E)$. Comme élément de V on peut écrire $v = v_1 + \cdots + v_{k-1}$ avec $v_i \in \mathrm{Ker}(\phi - \lambda_i \mathrm{id}_E)$, et comme $v \in \mathrm{Ker}(\phi - \lambda_k \mathrm{id}_E)$ on a

$$\lambda_k v_1 + \cdots + \lambda_k v_{k-1} = \lambda_k v = \phi(v) = \phi(v_1 + \cdots + v_{k-1}) = \lambda_1 v_1 + \cdots + \lambda_{k-1} v_{k-1}.$$

La somme $\mathrm{Ker}(\phi - \lambda_1 \mathrm{id}_E) \oplus \cdots \oplus \mathrm{Ker}(\phi - \lambda_{k-1} \mathrm{id}_E)$ étant directe les deux sommes aux extrémités doivent être égales terme par terme, c'est-à-dire $\lambda_k v_i = \lambda_i v_i$ pour $i = 1, 2, \ldots, k-1$. Cela s'écrit $(\lambda_k - \lambda_i)v_i = 0$, et comme $\lambda_i \neq \lambda_k$ cela n'est possible que si $v_i = 0$, et cela pour tout i, donc $v = 0$. □

Preuve. Par l'absurde. Supposons qu'il existe une écriture

$$v_1 + \cdots + v_k = 0$$

avec $v_i \in \mathrm{Ker}(\phi - \lambda_i \mathrm{id}_E)$ pour $i = 1, \ldots, k$ et les vecteurs v_i non tous nuls. On pourra alors choisir une telle écriture avec le plus petit nombre > 0 possible de termes non nuls. Ce nombre de termes non nuls ne peut clairement pas être 1, car $0 + \cdots + 0 + v_i + 0 + \cdots = 0$ entraînerait $v_i = 0$. En réduisant ce nombre de termes non nuls, sans le réduire à 0, on obtiendra une contradiction avec la minimalité. Choisissons un terme non nul $v_i \neq 0$, et appliquons $\phi - \lambda_i \mathrm{id}_E$ à l'écriture. Comme $(\phi - \lambda_i \mathrm{id}_E)(v_j) = \phi(v_j) - \lambda_i v_j = (\lambda_j - \lambda_i)v_j$, le résultat est

$$(\lambda_1 - \lambda_i)v_1 + \cdots + (\lambda_k - \lambda_i)v_k = 0.$$

Le terme non nul v_i est devenu $(\lambda_i - \lambda_i)v_i = 0$, et pour les autres termes $v_j \neq 0 \iff (\lambda_j - \lambda_i)v_j \neq 0$. On a donc réduit de 1 le nombre de termes non nuls dans la somme, c'est la contradiction cherchée. □

Comme un exemple simple de vecteurs et valeurs propres, considérons dans un $\mathbf{R}$-espace de dimension 2, un endomorphisme ϕ dont la matrice par rapport à une certaine base $\mathcal{B} = [b_1, b_2]$ est $\left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$, autrement dit on a $\phi(b_1) = b_2$ et $\phi(b_2) = b_1$. Alors b_1 et b_2 ne sont visiblement pas des vecteurs propres, mais $b_1 + b_2$ en est un, avec valeur propre 1, car $\phi(b_1 + b_2) = b_2 + b_1 = 1(b_1 + b_2)$. Un autre vecteur propre est $b_1 - b_2$, avec valeur propre -1 cette fois-ci, car $\phi(b_1 - b_2) = b_2 - b_1 = -1(b_1 - b_2)$. On vérifie facilement que les espaces propres pour $\lambda = 1$ et pour $\lambda = -1$ sont respectivement les droites vectorielles $\mathrm{Vect}(b_1 + b_2)$ et $\mathrm{Vect}(b_1 - b_2)$. On peut décrire géométriquement ϕ comme une réflexion dans la droite $\mathrm{Vect}(\left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right))$, et on a trouvé cet axe ainsi qu'une droite perpendiculaire à l'axe comme deux sous-espaces ϕ-stables. Le nombre de valeurs propres ne pouvant pas dépasser la dimension de l'espace, on a trouvé toutes les valeurs propres de ϕ, et les espaces propres correspondants.

Si on change cet exemple légèrement, en prenant l'endomorphisme ρ dont la matrice est $R = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$, la situation devient différente. Il s'agit ici d'une rotation par un quart de tour, et il est géométriquement évident que aucune droite n'est ρ-stable (sa direction change), ce qui exclut la possibilité d'avoir des vecteurs propres. On peut montrer cela algébriquement ; pour qu'un vecteur $xb_1 + yb_2$ soit vecteur propre avec valeur propre λ, il faut avoir $\rho(xb_1 + yb_2) = -yb_1 + xb_2 = \lambda x b_1 + \lambda y b_2$, et donc $\lambda x = -y$ et $\lambda y = x$, dont on déduit $\lambda^2 x = -x$ et $\lambda^2 y = -y$. Comme au moins un de x, y doit être non nul, il est nécessaire que $\lambda^2 = -1$, ce qui est impossible car $K = \mathbf{R}$. Si par contre on considérait un espace vectoriel complexe ($K = \mathbf{C}$), on aurait les possibilités $\lambda = \mathbf{i}$ et $\lambda = -\mathbf{i}$; ils sont alors effectivement des valeurs propres de ρ, avec pour vecteurs propres $(1, -\mathbf{i})_{\mathcal{B}} = b_1 - \mathbf{i}b_2$ pour $\lambda = \mathbf{i}$, et $(1, \mathbf{i})_{\mathcal{B}} = b_1 + \mathbf{i}b_2$ pour $\lambda = -\mathbf{i}$.

Cet exemple est une première indication que dans la recherche des valeurs propres, on est confronté à des équations ($\lambda^2 = -1$) qui sont polynomiales, et non pas linéaires, malgré leur origine dans un problème d'algèbre linéaire. Cela expliquera que (une partie de) l'algèbre des polynômes sera traitée dans ce cours.

On peut observer que dans de premier exemple l'endomorphisme vérifie $\phi^2 = \mathrm{id}_E$ et les valeurs propres vérifient $\lambda^2 = 1$, pendant que dans le second exemple on a $\rho^2 = -\mathrm{id}_E$ et les valeurs propres doivent vérifier $\lambda^2 = -1$. Cette coïncidence n'est pas un accident, car on montre facilement:

2.1.4. Proposition. *Si un endomorphisme ϕ vérifie une équation $a_n\phi^n + \cdots + a_2\phi^2 + a_1\phi + a_0 \mathrm{id}_E = 0_E$, alors toutes les valeurs propres λ de ϕ vérifient l'équation correspondante $a_n\lambda^n + \cdots + a_2\lambda^2 + a_1\lambda + a_0 = 0$.*

Preuve. Il suffit d'appliquer les deux membres (des endomorphismes) de l'équation donnée à un vecteur propre v de ϕ ce qui donne $(a_n\lambda^n + \cdots + a_2\lambda^2 + a_1\lambda + a_0)v = 0$. Comme $v \neq 0$, le scalaire doit être nul. □

2.2. *Diagonalisation.*

2.2.1. Définition. *Une matrice diagonale est une matrice carrée A dont tous les coefficients hors de la diagonale principale sont nuls : on a $A_{i,j} = 0$ si $i \neq j$. Une matrice diagonalisable (sur K) est une matrice semblable à une matrice diagonale à coefficients dans K. Un endomorphisme ϕ d'un K-espace vectoriel E de dimension finie est diagonalisable si sa matrice par rapport à une base convenable de E est diagonale. Dans ce cas la matrice de ϕ par rapport à une base quelconque est diagonalisable sur K.*

Comme dans la matrice A d'un endomorphisme ϕ par rapport à une base $\mathcal{B} = [b_1, \ldots, b_n]$, la colonne j contient les coordonnées de $\phi(b_j)$ dans la base $\mathcal{B}$, la condition que les coefficients de cette colonne qui ne sont pas sur la diagonale principale soient nuls veut dire que b_j est un vecteur propre de ϕ, pour la valeur propre égale au coefficient $A_{j,j}$ sur la diagonale. Par conséquent, la condition que ϕ soit diagonalisable est équivalente à la condition que E admette une base formée de vecteurs propres de ϕ. Dans ce cas, comme tout vecteur s'exprime sur cette base, la somme (directe) des espaces propres de ϕ est égale à E tout entier, et réciproquement si cette somme rempli E, on peut choisir une base de E formée de vecteurs propres en mettant ensemble des bases choisies séparément dans chaque espace propre. On obtient donc :

2.2.2. Proposition. *Pour $\phi \in \mathrm{End}(E)$, avec $\dim(E) < \infty$, les conditions suivantes sont équivalentes :*
(i) ϕ est diagonalisable,
(ii) E possède une base entièrement constituée de vecteurs propres pour ϕ,
(iii) E est la somme (toujours directe) des différents espaces propres de ϕ. □

Si une base de vecteurs propres (aussi appelée base de diagonalisation) de ϕ est connue, les valeurs propres de ϕ sont celles associées aux vecteurs dans la base, et ce sont ces valeurs qui figurent comme coefficients diagonaux dans la matrice diagonale de ϕ par rapport à $\mathcal{B}$. Pour une telle valeur λ, l'espace propre associé est engendré par les vecteurs dans la base qui ont λ comme vecteur propre.

Une matrice diagonalisable est semblable à une matrice diagonale, mais celle-ci n'est pas (en général) unique. Ceci dit, les choix pour une telle matrice sont très limités : deux matrices diagonales qui sont semblables ont le même ensemble de coefficients diagonaux (c'est l'ensemble des valeurs propres de l'endomorphisme) chaque valeur λ figure autant de fois sur la diagonale de l'une que sur celle de l'autre (c'est la dimension de l'espace propre pour λ). Deux matrices diagonales sont donc semblables seulement si l'une est obtenue de l'autre par une permutation de ses coefficients diagonaux ; la permutation correspondante de vecteurs de la base montre que deux telles matrices sont effectivement semblables.

Si un endomorphisme ϕ est diagonalisable, sur une base $\mathcal{B} = [b_1, \ldots, b_n]$ de vecteur propres et avec des valeurs propres correspondantes $\lambda_1, \ldots, \lambda_n$, alors l'effet d'appliquer ϕ sur un vecteur en termes de ses coordonnées dans $\mathcal{B}$ est de multiplier chaque coordonnée i par la valeur propre correspondante λ_i, c'est-à-dire

$$\phi\big((x_1, \ldots, x_n)_{\mathcal{B}}\big) = (\lambda_1 x_1, \ldots, \lambda_n x_n)_{\mathcal{B}}. \tag{11}$$

Ceci est très utile pour décrire les puissances ϕ^k de l'endomorphisme, car il en découle que

$$\phi^k\big((x_1, \ldots, x_n)_{\mathcal{B}}\big) = (\lambda_1^k x_1, \ldots, \lambda_n^k x_n)_{\mathcal{B}} \qquad \text{pour tout } k \in \mathbf{N}. \tag{12}$$

Autrement dit, la puissance D^k d'une matrice diagonale D est la matrice diagonale obtenue de D en remplaçant chaque coefficient diagonal par sa puissance k-ème. Pour une matrice générale, l'expression pour sa puissance k-ème qu'on peut déduire de celle du produit matriciel est *beaucoup* plus compliquée.

Considérons un exemple concret. Dans un $\mathbf{Q}$-espace E de dimension 2, muni d'une base $\mathcal{B} = [b_1, b_2]$, on considère un endomorphisme ϕ aux valeurs propres -3 et 2, mais avec vecteurs propres correspondants $v_1 = (1,1)_{\mathcal{B}} = b_1 + b_2$ et $v_2 = (2,1)_{\mathcal{B}} = 2b_1 + b_2$. Le matrice de passage de la base $\mathcal{B}$ à la base $[v_1, v_2]$ est donc $P = \left(\begin{smallmatrix} 1 & 2 \\ 1 & 1 \end{smallmatrix}\right)$, dont l'inverse est $P^{-1} = \left(\begin{smallmatrix} -1 & 2 \\ 1 & -1 \end{smallmatrix}\right)$, et l'endomorphisme ϕ, dont la matrice par rapport à la base $[v_1, v_2]$ est par construction $\mathrm{Mat}_{[v_1,v_2]}(\phi) = D = \left(\begin{smallmatrix} -3 & 0 \\ 0 & 2 \end{smallmatrix}\right)$, aura comme matrice A par rapport à la base $\mathcal{B}$:

$$A = \mathrm{Mat}_{\mathcal{B}}(\phi) = P \cdot D \cdot P^{-1} = \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} -3 & 0 \\ 0 & 2 \end{pmatrix} \cdot \begin{pmatrix} -1 & 2 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 7 & -10 \\ 5 & -8 \end{pmatrix}.$$

[La transformation de la matrice par rapport à la nouvelle base $[v_1, v_2]$ vers la matrice par rapport à l'ancienne base $\mathcal{B}$ est l'inverse de celle décrite dans la proposition 1.5.2, d'où on utilise P^{-1} *à droite*. Notre démarche est en fait un peu inhabituel, en imposant à certains vecteur d'être des vecteurs propres, et en construisant la matrice dans l'ancienne base après coup. Cela est fait pour éviter la recherche (pas encore abordée dans ce cours) des valeurs propres et des vecteurs propres d'une matrice donnée. Le lecteur prendra néanmoins soin de vérifier qu'on a effectivement obtenu $\phi(v_1) = -3v_1$ et $\phi(v_2) = 2v_2$.] On a pour $n \in \mathbf{N}$:

$$\mathrm{Mat}_{[v_1,v_2]}(\phi^n) = D^n = \begin{pmatrix} (-3)^n & 0 \\ 0 & 2^n \end{pmatrix}$$

et donc

$$A^n = \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} (-3)^n & 0 \\ 0 & 2^n \end{pmatrix} \cdot \begin{pmatrix} -1 & 2 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} -(-3)^n + 2 \times 2^n & 2 \times (-3)^n - 2 \times 2^n \\ -(-3)^n + 2^n & 2 \times (-3)^n - 2^n \end{pmatrix}.$$

On peut vérifier cette formule facilement par récurrence, mais elle se laisse difficilement deviner à partir de quelques valeurs explicites de A^n : pour la trouver il était nécessaire de connaître les valeurs propres et les vecteurs propres de A. Par contre, ce qu'on peut constater facilement par le calcul explicite des puissances A^n, est que ses colonnes tendent vers des multiples du vecteur propre $v_1 = (1, 1)_{\mathcal{B}} = b_1 + b_2$. On a par exemple $A^{15} = \left(\begin{smallmatrix} 14414443 & -28763350 \\ 14381675 & -28730582 \end{smallmatrix}\right)$; la proximité de ses colonnes à $-(-3)^{15}v_1$ respectivement à $2 \times (-3)^{15}v_1$ s'explique par le fait que le facteur $(-3)^{15} = -14\,348\,907$ est beaucoup plus grand en valeur absolue que $2^{15} = 32\,768$. Il est un phénomène général que, si A est une matrice diagonalisable sur le nombres réels, alors pour presque tous les vecteurs v l'image itérée $A^n \cdot v$ s'«approche» (dans un sens qui reste à préciser) de l'espace propre pour la valeur propre qui est la plus grande en valeur absolue. Ce constat est très important en analyse numérique, et est à la base de nombreux algorithmes qui veulent éviter (car trop coûteuse) notamment la détermination d'une base de diagonalisation. Mais ces considérations étant d'une nature différente de l'approche algébrique de ce cours, on en dira pas plus.

Un autre exemple illustrera ce qui se passe quand une matrice n'a pas de valeurs propres réelles, mais est diagonalisable sur les nombres complexes. En reprenant la matrice $R = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$ de ce type qu'on a vue avant, on peut observer qu'il s'agit d'un quart de tour, d'où $R^2 = \left(\begin{smallmatrix} -1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$, $R^3 = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$, $R^4 = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$ est la matrice identité, et pour toute autre puissance de R on trouvera la même matrice qu'en remplaçant l'exposant par son reste modulo 4. Mais même si dans ce cas on voit les puissances de la matrice tout de suite, on peut aussi (en supposant que la matrice corresponde à un endomorphisme d'un espace *complexe* de dimension 2) les déterminer en utilisant la base de vecteurs propres $v_1 = (1, -\mathbf{i})_{\mathcal{B}}$ et $v_2 = (1, -\mathbf{i})_{\mathcal{B}}$ (où $\mathcal{B}$ est la base par rapport à laquelle la matrice est R). La matrice de passage de $\mathcal{B}$ à $[v_1, v_2]$ est $P = \left(\begin{smallmatrix} 1 & 1 \\ -\mathbf{i} & \mathbf{i} \end{smallmatrix}\right)$ avec inverse $P^{-1} = \frac{1}{2}\left(\begin{smallmatrix} 1 & \mathbf{i} \\ 1 & -\mathbf{i} \end{smallmatrix}\right)$; alors

$$\mathrm{Mat}_{[v_1,v_2]}(\phi) = D = \begin{pmatrix} \mathbf{i} & 0 \\ 0 & -\mathbf{i} \end{pmatrix} = P^{-1} \cdot R \cdot P = \frac{1}{2} \begin{pmatrix} 1 & \mathbf{i} \\ 1 & -\mathbf{i} \end{pmatrix} \cdot \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 \\ -\mathbf{i} & \mathbf{i} \end{pmatrix}$$

et donc

$$\begin{aligned} R^n = P \cdot D^n \cdot P^{-1} &= \frac{1}{2} \begin{pmatrix} 1 & 1 \\ -\mathbf{i} & \mathbf{i} \end{pmatrix} \cdot \begin{pmatrix} \mathbf{i}^n & 0 \\ 0 & (-\mathbf{i})^n \end{pmatrix} \cdot \begin{pmatrix} 1 & \mathbf{i} \\ 1 & -\mathbf{i} \end{pmatrix} \\ &= \frac{1}{2} \begin{pmatrix} \mathbf{i}^n + (-\mathbf{i})^n & \mathbf{i}^{n+1} + (-\mathbf{i})^{n+1} \\ -(\mathbf{i}^{n+1} + (-\mathbf{i})^{n+1}) & \mathbf{i}^n + (-\mathbf{i})^n \end{pmatrix} \end{aligned}$$

On peut observer que ces matrices sont réelles pour tout n, comme elles doivent l'être. En plus, les termes sur le diagonal s'annulent pour n impair, et les autres termes pour n pair, et on retrouve les puissances de R mentionnées ci-dessus, qui sont périodiques modulo 4 en fonction de n. C'est périodicité est liée au fait que la puissance 4 de toutes les valeurs propres est 1 (on dit qu'elles sont des 4-èmes racines de l'unité). La relation entre la matrice diagonale et la matrice correspondante sur la base $\mathcal{B}$ devient plus claire si l'on remplace le couple de valeurs propres par un couple de conjugués complexes plus général, qu'on écrira sous la forme exponentielle $(re^{\mathbf{i}\theta}, re^{-\mathbf{i}\theta})$ avec $r, \theta \in \mathbf{R}$. On aura alors

$$D = \begin{pmatrix} re^{\mathbf{i}\theta} & 0 \\ 0 & re^{-\mathbf{i}\theta} \end{pmatrix}$$

et

$$P\cdot D^n\cdot P^{-1} = \frac{1}{2}\begin{pmatrix}1 & 1\\ -\mathbf{i} & \mathbf{i}\end{pmatrix}\cdot\begin{pmatrix}r^n e^{n\mathbf{i}\theta} & 0\\ 0 & r^n e^{-n\mathbf{i}\theta}\end{pmatrix}\cdot\begin{pmatrix}1 & \mathbf{i}\\ 1 & -\mathbf{i}\end{pmatrix}$$
$$= \frac{1}{2}\begin{pmatrix} r^n(e^{n\mathbf{i}\theta}+e^{-n\mathbf{i}\theta}) & r^n(\mathbf{i}e^{n\mathbf{i}\theta}-\mathbf{i}e^{-n\mathbf{i}\theta})\\ -r^n(\mathbf{i}e^{n\mathbf{i}\theta}-\mathbf{i}e^{-n\mathbf{i}\theta}) & r^n(e^{n\mathbf{i}\theta}+e^{-n\mathbf{i}\theta})\end{pmatrix} = r^n\begin{pmatrix}\cos n\theta & -\sin n\theta\\ \sin n\theta & \cos n\theta\end{pmatrix}.$$

Ainsi les puissances d'un couple de conjugués complexes sur cette base particulière de vecteurs propres se traduit sur la base $\mathcal{B}$ en la puissance correspondante de la similitude composée d'une rotation par θ et une homothétie de facteur r. En règle générale, si les puissances d'une matrice réelle (appliquée à un vecteur) manifestent un comportement de rotation ou d'oscillation, éventuellement en combinaison avec une croissance ou décroissance exponentielle, cela est souvent une indication d'un couple de conjugués complexes comme valeurs propres, si l'on interprète la matrice comme une matrice complexe.

2.3. *Existence de valeurs propres.*

On a vu que les vecteurs propres peuvent rendre la description d'un endomorphisme ϕ plus transparente, surtout quand on peut trouver une base constituée de vecteurs propres (le cas où ϕ est diagonalisable). Mais on a aussi vu des exemples d'endomorphismes de $\mathbf{R}$-espaces vectoriels pour lequel aucun vecteur propre existe. Dans cette section on montrera que pour les endomorphismes de $\mathbf{C}$-espaces vectoriels (de dimension finie et non nulle), au moins l'existence de vecteurs propres est assurée (ce qui ne veut pas forcément dire qu'il existe une base formée de vecteurs propres). La propriété du corps $\mathbf{C}$ qui le distingue à cet égard de $\mathbf{R}$ (ou de $\mathbf{Q}$) est le fait que $\mathbf{C}$ est *algébriquement clos* : tout polynôme à coefficients dans $\mathbf{C}$ et de degré >0 possède au moins une racine dans $\mathbf{C}$ (c'est l'énoncé du Théorème de d'Alembert–Gauss, prouvé en 1814 par le mathématicien suisse Jean-Robert Argand, et qu'on admettra ici).

Il peut paraître étrange que cette propriété de polynômes joue un rôle pour la question d'existence de vecteurs propres, car rien dans la définition des vecteurs ou valeurs propres fait référence à des polynômes. La proposition 2.1.4 établit un lien avec des polynômes, mais ne dit pas que *toute* racine du polynôme mentionné est une value propre. Néanmoins, on verra dans la suite qu'on peut associer de façon naturelle un polynôme à ϕ, non constant si $\dim E > 0$, dont les racines sont précisément les valeurs propres de ϕ. Il y aura même *deux* façons de le faire, à savoir d'associer à ϕ son polynôme caractéristique χ_ϕ ou son polynôme minimal μ_ϕ. En attendant ces constructions, on peut déjà prouver l'existence de valeurs propres directement. Dans cet argument, on fera correspondre (comme dans la proposition 2.1.4) aux monômes X^n les puissances ϕ^n de ϕ. On remarque en passant, qu'au polynôme constant $1 = X^0$ correspondra l'endomorphisme ϕ^0, défini par convention comme $\phi^0 = \mathrm{id}_E$.

L'ensemble de polynômes à coefficients dans K sera noté $K[X]$, où X désigne l'indéterminée utilisée dans les polynômes. Cet ensemble est muni d'opérations de addition, soustraction, multiplication (on dit que c'est une *anneau*). On peut aussi substituer une valeur $a \in K$ pour X dans $P \in K[X]$, et le résultat est une valeur notée $P[a] \in K$. Si $P[a] = 0$ on appelle a une racine de P ; comme on montrera plus tard, on peut dans ce cas diviser P par $X - a$, et donc écrire $P = (X - a)Q$ avec $Q \in K[X]$ (le quotient).

On peut aussi substituer pour X l'endomorphisme ϕ, et le résultat est alors (comme dans la proposition 2.1.4) un endomorphisme qui sera noté $P[\phi]$ (en accord avec la convention $\phi^0 = \mathrm{id}_E$, un terme constant c de P devient l'homothétie $c\,\mathrm{id}_E$ après cette substitution). La substitution est compatible avec addition et soustraction de polynômes, et la multiplication de polynômes correspond à *composition* d'endomorphismes, c'est-à-dire $(PQ)[\phi] = P[\phi] \circ Q[\phi]$; comme la multiplication de polynômes est commutatif, ce résultat peut est aussi égal à la composée $Q[\phi] \circ P[\phi]$, et cela malgré le fait que la composition d'endomorphismes n'est pas en général commutatif. Souvent on appliquera $P[\phi] \in \mathrm{End}(E)$ tout de suite à un vecteur $v \in E$, et pour alléger la notation on écrira $P \cdot_\phi v$ pour $P[\phi](v)$; ici chaque terme cX^k de P est devenu $c\phi^k(v)$. La compatibilité de substitution avec multiplication de polynômes s'exprime alors par

$$(PQ)\cdot_\phi v = P\cdot_\phi (Q\cdot_\phi v). \tag{13}$$

L'idée pour obtenir un polynôme à partir de ϕ et un vecteur v choisi, est de considérer la suite de vecteurs $v = \phi^0(v), \phi^1(v), \phi^2(v), \ldots$ de E. Comme une famille libre ne peut avoir plus de $\dim E$ membres, on trouvera une première famille $[\phi^0(v), \cdots, \phi^d(v)]$ qui est liée, et la relation $a_0\phi^0(v) + \cdots + a_d\phi^d(v) = 0$ veut dire $P \cdot_\phi v = 0$ pour $P = a_0X^0 + \cdots + a_dX^d \neq 0$, qui est de degré minimal pour cette propriété.

2.3.1. Proposition. *Soit ϕ un endomorphisme d'un* **C***-espace vectoriel E de dimension finie et non nulle. Alors E contient (au moins) un vecteur propre de ϕ.*

Preuve. Soit v un vecteur non nul (qui existe car $\dim(E) \neq 0$). Comme indiqué ci-dessus il existe $P \in K[X]$ non nul et de degré minimal tel que $P \cdot_\phi v = 0$. D'après le théorème de d'Alembert–Gauss P possède (au moins) une racine $\lambda \in \mathbf{C}$, et on aura donc $P = (X-\lambda)Q$ pour un certain polynôme $Q \in \mathbf{C}[X]$, qui sera de degré $d-1$. Alors, en utilisant (13) on a $0 = P\cdot_\phi v = (X-\lambda)\cdot_\phi(Q\cdot_\phi v) = (\phi - \lambda\,\mathrm{id}_E)(Q\cdot_\phi v)$. Donc le vecteur $w = Q\cdot_\phi v$ est dans l'espace propre $\mathrm{Ker}(\phi - \lambda\,\mathrm{id}_E)$ de ϕ pour λ ; montrons que $w \neq 0$. Ce w est une combinaison linéaire des vecteurs $\phi^0(v), \ldots, \phi^{d-1}(v)$, et même une combinaison linéaire non-triviale (car $Q \neq 0 \in K[X]$) ; or la famille $[\phi^0(v), \ldots, \phi^{d-1}(v)]$ est libre par minimalité de $\deg(P)$, donc $w \neq 0$. □

Remarquons d'emblée que cette proposition ne permettra pas de trouver dans tous les cas *une base* de vecteurs propres : une telle base n'existe pas toujours si $\dim(E) \geq 2$. Si $\mathcal{B} = [b_1, \ldots, b_n]$ est une base de E, on peut définir $\phi \in \mathrm{End}(E)$ par $\phi(b_1) = 0$ et $\phi(b_{i+1}) = b_i$ pour $i = 2, \ldots, n$. Alors $\phi \neq 0$ (si $n \geq 2$) mais $\phi^n = 0$, donc d'après la proposition 2.1.4 la seul valeur propre possible est $\lambda = 0$. Mais l'espace propre $\mathrm{Ker}(\phi)$ pour $\lambda = 0$ est $\mathrm{Vect}(b_1)$ qui n'est que de dimension 1, donc il n'y a pas de base de E formée de vecteurs propres (pour $\lambda = 0$). On voit de la même façon qu'aucun endomorphisme nilpotent (c'est-à-dire vérifiant $\phi^k = 0$ pour un certain k) n'est diagonalisable, sauf l'endomorphisme nul. Ces exemples d'endomorphismes non diagonalisables se généralisent facilement à d'autres valeurs propres que $\lambda = 0$, en ajoutant $\lambda\,\mathrm{id}_E$ à l'endomorphisme ϕ ; l'endomorphisme ainsi obtenu aura λ comme unique valeur propre, sans que son espace propre $\mathrm{Ker}(\phi - \lambda\,\mathrm{id}_E)$ ne soit l'espace entier.

Malgré ces exemples, les endomorphismes complexes non diagonalisables sont assez rares. Dans la démonstration de la proposition 2.3.1, on aura le plus souvent $d = n$ (la dépendance linéaire n'apparaît que quand elle est inévitable). En itérant l'application du théorème d'Alembert–Gauss on peut factoriser P comme produit $P = c(X-\lambda_1)\cdots(X-\lambda_d)$ de facteurs de degré 1, pour certains $\lambda_1, \ldots, \lambda_d \in \mathbf{C}$. Si ces d racines λ_i sont toutes *distinctes*, on peut appliquer l'argument donné pour chacune, et ainsi trouver d vecteurs propres pour des valeurs propres distincts, et qui formeront une base de $\mathrm{Vect}(\phi^0(v), \ldots, \phi^{d-1}(v))$, espace qui est égal à E si $d = n$. Ce n'est donc que dans les cas exceptionnels où soit $d < n$, soit le polynôme P possède au moins une racine multiple, que ϕ peut ne pas être diagonalisable. (Mais en revanche, si P possède une ou plusieurs racines multiples, on verra que ϕ ne sera jamais diagonalisable.)

2.4. *Exemples d'application des vecteurs propres.*

Dans cette section on donnera quelques exemples dans lesquels il sera possible de trouver toutes les valeurs propres d'un endomorphisme, et qui illustrent l'utilité de cette démarche.

Le premier exemple concerne des *suites récurrentes linéaires*. Ce sont des suites infinies $(a_i)_{i\in\mathbf{N}}$ de scalaires $a_i \in K$ qui vérifient une relation dite de récurrence linéaire, ce qui est une condition de la forme

$$a_{i+d} = c_0 a_i + c_1 a_{i+1} + \cdots + c_{d-1} a_{i+d-1} \qquad \text{pour tout } i \in \mathbf{N}, \tag{14}$$

avec $c_0, \ldots, c_{d-1} \in K$ fixés ; on appelle d l'ordre de la récurrence. Pour spécifier une suite récurrente particulière il faut en plus donner les d premiers termes $(a_0, \ldots, a_{d-1})$ de la suite, après laquelle les autres termes sont successivement uniquement déterminés par (14). Il sera néanmoins être utile de considérer l'ensemble E de toutes les suites qui satisfont (14) (pour les mêmes $c_0, \ldots, c_{d-1}$), quels que soient leurs termes initiaux ; on voit facilement que E forme un K-espace vectoriel (pour les opérations d'addition et de multiplication scalaire habituelles de suites). L'exemple le plus célèbre d'une suite récurrente linéaire est la suite de Fibonacci $(F_i)_{i\in\mathbf{N}}$, définie par ses termes initiaux $F_0 = 0$ et $F_1 = 1$ et la relation de récurrence $F_{i+2} = F_i + F_{i+1}$.

Le problème qu'on considère ici est de trouver une formule pour F_i, ou plus généralement pour le terme général d'une suite récurrente linéaire. Il existe plusieurs façons d'approcher ce problème à l'aide de l'algèbre linéaire ; bien qu'elles utilisent des formulations différentes, elles mènent toutes à des problèmes équivalents de recherche de vecteurs propres.

Une approche simple qui évite de considérer explicitement l'espace E mentionné de suites infinies, est de considérer des "fenêtres" $(a_i, \ldots, a_{i+d-1})_n \in K^d$ de d termes consécutifs pris dans une suite qui

vérifie (14), et d'observer que cette relation prescrit que le passage $(a_i, \ldots, a_{i+d-1}) \mapsto (a_{i+1}, \ldots, a_{i+d})$ d'une fenêtre à la suivante est l'application linéaire de matrice (dans la base canonique)

$$A = \begin{pmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ddots & \vdots \\ 0 & 0 & 0 & \ddots & 0 \\ \vdots & \vdots & \vdots & \ddots & 1 \\ c_0 & c_1 & c_2 & \ldots & c_{d-1} \end{pmatrix}, \tag{15}$$

dont les $d-1$ premières lignes décrivent que les $d-1$ premières composantes de la nouvelle fenêtre sont juste les $d-1$ dernières composantes de l'ancienne déplacées d'un cran, et que la dernière composante des donnée par (14) en termes des composantes de l'ancienne fenêtre. Si $v = (a_0, \ldots, a_{d-1}) \in K^d$ est le vecteur des termes initiaux, qu'on considère comme une colonne, alors le terme a^i est la première composante du vecteur $A^i \cdot v$. Ceci montre que notre problème est lié au calcul des puissances de A.

Une autre approche est de travailler dans le sous-espace $E \subseteq K^{\mathbf{N}}$ des suites vérifiant la relation de récurrence. Comme les d premiers termes de la suite peuvent être choisis librement et déterminent le reste de la suite, l'application $E \to K^d$ qui associe à une suite infinie ses d premiers termes, c'est-à-dire $f : (a_i)_{i\in\mathbf{N}} \mapsto (a_0, \ldots, a_{d-1})$, est une bijection. Aussi f est clairement une application linéaire, donc c'est un isomorphisme de K-espaces vectoriels, ce qui montre en particulier que $\dim(E) = d$. La base canonique de K^d correspond par f^{-1} une certaine base $\mathcal{B} = [\mathbf{b}_0, \ldots, \mathbf{b}_{d-1}]$ de E, pour laquelle f associe leurs coordonnées aux vecteurs. Explicitement, le vecteur $\mathbf{b}_k$, pour $0 \leq k < d$, est la suite $(a_i)_{i\in\mathbf{N}}$ dont les d premiers termes sont tous nuls, sauf a_k qui est 1, et dont les autres termes sont déterminés par (14).

Concrètement, pour la relation de récurrence de Fibonacci, les deux suites suivantes qui constituent la base $\mathcal{B}$:

$$\mathbf{b}_0 = (1, 0, 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, \ldots),$$
$$\mathbf{b}_1 = (0, 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, 233, \ldots).$$

Le fait que $\mathbf{b}_1$ est obtenue par un décalage de $\mathbf{b}_0$ est particulier à cet exemple, comme on peut voir en changeant le récurrence en $a_{i+2} = 2a_i + a_{i+1}$, auquel cas on obtiendrait comme base le couple de suites

$$\mathbf{b}'_0 = (1, 0, 2, 2, 6, 10, 22, 42, 86, 170, 342, 682, \ldots),$$
$$\mathbf{b}'_1 = (0, 1, 1, 3, 5, 11, 21, 43, 85, 171, 341, 683, \ldots).$$

En revanche, le fait que le décalage $(a_1, a_2, a_3, \ldots)$ d'une suite $(a_0, a_1, a_2, \ldots)$ de E appartient à E est vrai en général, et facile à comprendre : pour que la suite décalée vérifie (14) à l'indice i, il suffit que la suite originale vérifie (14) à l'indice $i+1$, ce qui est le cas par hypothèse. L'opération de décalage définit donc une application linéaire $\delta : K^{\mathbf{N}} \to K^{\mathbf{N}}$ pour laquelle E est δ-stable, et par restriction on obtient $\delta_E \in \mathrm{End}(E)$. Dans l'exemple de le récurrence de Fibonacci, on a $\delta(\mathbf{b}_0) = \mathbf{b}_1$ et $\delta(\mathbf{b}_1) = \mathbf{b}_0 + \mathbf{b}_1$. En général, l'équation (14) dit que le terme à la position d de la suite $\mathbf{b}_k$ est c_k pour $0 \leq k < d$, et il s'ensuit que $\delta(\mathbf{b}_0) = c_0\mathbf{b}_{d-1}$ et $\delta(\mathbf{b}_i) = \mathbf{b}_{i-1} + c_i\mathbf{b}_{d-1}$ pour $0 < i < d$. Cela veut dire que $\mathrm{Mat}_{\mathcal{B}}(\delta_E) = A$, la même matrice que donnée dans (15). Cela n'est pas étonnant, car le "transport" de $\delta_E : E \to E$ en une application $K^d \to K^d$ via l'isomorphisme $f : E \to K^d$ d'expression dans la base $\mathcal{B}$, donne l'application linéaire $f(v) \mapsto f(\delta(v))$ de "avancement de fenêtre" $(a_0, \ldots, a_{d-1}) \mapsto (a_1, \ldots, a_d)$ décrit par la matrice A.

On a déjà vu que le fait de connaître une base de vecteurs propres aide à trouver une expression explicite pour les puissances A^n. Le point de vue de l'espace E de suites permet une interprétation de ces vecteurs propres éventuels. L'opération de décalage est bien définie dans l'espace $K^{\mathbf{N}}$ entier (donc pour des suites ne vérifiant pas forcément une relation de récurrence), et dire pour $\mathbf{s} = (a_i)_{i\in\mathbf{N}}$ que sa suite décalée $(a_{i+1})_{i\in\mathbf{N}}$ est égale à un multiple $\lambda\mathbf{s}$ de $\mathbf{s}$ veut dire que les termes de $\mathbf{s}$ vérifient $a_{i+1} = \lambda a_i$ pour tout $i \in \mathbf{N}$, une récurrence linéaire d'ordre 1. Alors $\mathbf{s}$ est une suite géométrique de raison λ, et pour de telles suites il est facile de donner le terme général : $a_i = a_0\lambda^i$. Si on sait trouver dans E une base de vecteurs propres pour δ_E, c'est-à-dire de suites géométriques, cela permettra d'exprimer le terme général de toutes les suites dans E, en les exprimant dans cette base.

La question devient alors de chercher les valeurs λ pour lesquelles la suite géométrique $(\lambda^i)_{i\in\mathbf{N}}$ appartient à E, c'est-à-dire les valeurs propres de δ_E. Substitution dans (14) donne pour $i=0$ l'équation

$$\lambda^d = c_0 + c_1\lambda^1 + \cdots + c_{d-1}\lambda^{d-1}, \tag{16}$$

et pour d'autres indices i on retrouve la même équation multipliée par λ^i. On trouve donc que λ est valeur propre de δ_E si et seulement si λ est racine du polynôme $Q = X^d - c_{d-1}X^{d-1} - \cdots - c_1X^1 - c_0$, dit polynôme caractéristique de la relation de récurrence. Si ce polynôme possède d racines *distinctes* $\lambda_1,\ldots,\lambda_d$ dans K, alors l'endomorphisme δ_E admet une famille de d vecteurs propres associées à ces d valeurs propres, à savoir les d suites géométriques $\mathbf{g}_{\lambda_i} = (1,\lambda_i,\lambda_i^2,\lambda_i^3,\ldots) = (\lambda_i^k)_{k\in\mathbf{N}}$ pour $i=1,\ldots,d$. Une telle famille étant libre (proposition 2.1.3), elle formera une base de l'espace E (qui est de dimension d). L'endomorphisme δ est donc diagonalisable si le polynôme Q admet d racines distinctes dans K.

Considérons le cas concret de la récurrence de Fibonacci $a_{i+2} = a_i + a_{i+1}$, dont le polynôme caractéristique est $Q = X^2 - X - 1$. Il a deux racines réelles, à savoir $\frac{1+\sqrt5}{2}$ (le nombre d'or, souvent noté φ, qui vaut approximativement 1,618) et $\frac{1-\sqrt5}{2} = 1-\varphi = -\frac{1}{\varphi} \approx -0{,}618$. Alors si $K=\mathbf{R}$ (ou $K=\mathbf{C}$), les deux suites géométriques $\mathbf{g}_1 = \big((\frac{1+\sqrt5}{2})^k\big)_{k\in\mathbf{N}}$ et $\mathbf{g}_2 = \big((\frac{1-\sqrt5}{2})^k\big)_{k\in\mathbf{N}}$ forment une base de E. Pour exprimer ces deux vecteurs en coordonnées dans l'ancienne base $\mathcal{B} = [\mathbf{b}_0,\mathbf{b}_1]$, il suffit de prendre les deux premiers termes de ces suites ; on a donc $\mathbf{g}_1 = \mathbf{b}_0 + (\frac{1+\sqrt5}{2})\mathbf{b}_1 = (1,\frac{1+\sqrt5}{2})_{\mathcal{B}}$ et $\mathbf{g}_2 = \mathbf{b}_0 + (\frac{1-\sqrt5}{2})\mathbf{b}_1 = (1,\frac{1-\sqrt5}{2})_{\mathcal{B}}$. La matrice de passage P de $\mathcal{B}$ vers la base $\mathcal{G} = [\mathbf{g}_1,\mathbf{g}_2]$ des vecteurs propres est

$$P = \begin{pmatrix} 1 & 1 \\ \frac{1+\sqrt5}{2} & \frac{1-\sqrt5}{2} \end{pmatrix}, \qquad \text{dont la matrice l'inverse est} \qquad P^{-1} = \frac{1}{\sqrt5}\begin{pmatrix} \frac{-1+\sqrt5}{2} & 1 \\ \frac{1+\sqrt5}{2} & -1 \end{pmatrix}.$$

Pour trouver une expression pour le nombres de Fibonacci, exprimons la suite $F = (F_i)_{i\in\mathbf{N}}$ de Fibonacci dans la base $\mathcal{G}$. Son expression $F = (0,1)_B$ dons l'ancienne base est donnée pas ces deux premiers termes, et pour convertir ces coordonnées vers la base $\mathcal{G}$ on applique la proposition 1.5.1 (dans le sens opposé) : $P^{-1}\binom{0}{1} = \binom{1/\sqrt5}{-1/\sqrt5}$, donc $F = (1/\sqrt5, -1/\sqrt5)_{\mathcal{G}}$. Il suffit de prendre dans cette combinaison linéaire de $\mathbf{g}_1$ et $\mathbf{g}_2$ le terme au rang n :

$$F_n = \frac{1}{\sqrt5}\left(\frac{1+\sqrt5}{2}\right)^n - \frac{1}{\sqrt5}\left(\frac{1-\sqrt5}{2}\right)^n. \tag{17}$$

On trouve dans cette expression les puissances des deux valeurs propres de δ, dont le première φ est plus grand que 1, pendant que la seconde $-\frac{1}{\varphi}$ est (négative et) plus petite en valeur absolue que 1. Par conséquent, le premier terme dominera largement le second quand n devient grand, et on pourra conclure qu'alors $F_n \approx \varphi^n/\sqrt5$ est une très bonne approximation, et que $\lim_{n\to\infty} F_{n+1}/F_n = \varphi = \frac{1+\sqrt5}{2}$; la suite de Fibonacci n'est pas une suite géométrique, mais elle ressemble à une suite géométrique pour grand n.

On observe que la définition de la suite de Fibonacci ne fait intervenir que des entiers, et qu'il est donc *a priori* évident que $F_n \in \mathbf{Z}$; cependant ceci n'est pas tellement clair dans la formule (17) dans laquelle les nombres irrationnels φ et $-\frac{1}{\varphi}$ jouent un rôle central. Cela est typique pour des problèmes de valeurs propres : souvent on est amené à chercher des valeurs propres dans un corps plus grand que celui de départ. Il arrive également que les valeurs propres pour une matrices réelle sont complexes.

Avec la base de vecteurs propres, on peut aussi calculer les puissances A^i de la matrice de (15) explicitement. La matrice est $A = \mathrm{Mat}_{\mathcal{B}}(\delta_E) = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 1 \end{smallmatrix}\right)$ et on a

$$P^{-1}\cdot A\cdot P = \begin{pmatrix} \frac{1+\sqrt5}{2} & 0 \\ 0 & \frac{1-\sqrt5}{2} \end{pmatrix}$$

et on peut alors calculer (en comparant les expressions obtenues à celle de F_n)

$$A^i = \begin{pmatrix} 1 & 1 \\ \frac{1+\sqrt5}{2} & \frac{1-\sqrt5}{2} \end{pmatrix}\cdot\begin{pmatrix} \frac{(1+\sqrt5}{2)^i} & 0 \\ 0 & \frac{(1-\sqrt5}{2)^i} \end{pmatrix}\cdot\frac{1}{\sqrt5}\begin{pmatrix} \frac{-1+\sqrt5}{2} & 1 \\ \frac{1+\sqrt5}{2} & -1 \end{pmatrix} = \begin{pmatrix} F_{i-1} & F_i \\ F_i & F_{i+1} \end{pmatrix}$$

(où on pose $F_{-1}=1$). Application de cette matrice au vecteur $v = \binom{F_0}{F_1} = \binom{0}{1}$ donne sa dernière colonne, et on retrouve F_i comme la première composante de $A^i\cdot v$, ce qui confirme ce qu'on avait dit initialement.

Le cas général d'une suite récurrente linéaire peut être traité de la même façon, sous la condition pratique qu'on puisse factoriser le polynôme caractéristique (théoriquement cela est toujours possible si $K = \mathbf{C}$), et la condition plus fondamentale que cette factorisation ne révèle pas de racines multiples. Pour une valeur (propre) donnée λ, les seuls vecteurs propres de δ dans $K^{\mathbf{N}}$ sont les suites géométriques de raison λ (qui forment un sous-espace de dimension 1), donc si Q possède moins de d racines distinctes, il n'y a pas d'espoir de trouver une base de vecteurs propres dans E, et δ_E ne sera pas diagonalisable. Dans ces cas, on aura besoin d'autres solutions particulières que les suites géométriques pour former une base dans laquelle tout suite dans E peut être exprimée.

Comme autre exemple on peut considérer un problème d'équations différentielles en analyse ; même si le contexte est assez différent, sa nature algébrique sera assez semblable à celle de l'exemple précédent. On cherche à connaître l'espace E des fonctions $f : \mathbf{R} \to K$, avec $K = \mathbf{R}$ ou $K = \mathbf{C}$, suffisamment dérivables, qui vérifient une équation différentielle linéaire d'ordre d à coefficients constants :

$$f^{(d)}(t) = c_o f(t) + c_1 f'(t) + c_2 f^{(2)}(t) + \cdots + c_{d-1} f^{(d-1)}(t) \qquad \text{pour tout } t \in \mathbf{R}, \tag{18}$$

dont on voit facilement que c'est une espace K-vectorielle. Une telle équation avec $d = 2$ est bien connue en mécanique : elle exprime l'accélération $f''(t)$ instantanée d'un particule comme une fonction linéaire de la position $f(t)$ et de la vitesse $f'(t)$ instantanées. Dans cette application les constantes c_0 et c_1 seront négatives en règle générale : $-c_0$ correspond à une force élastique qui repousse le particule vers la position d'équilibre (la valeur 0), quelle force est proportionnelle à la déviation $f(x)$ de cette position, et $-c_1$ correspond à une force de frottement opposée à la vitesse $f'(t)$ qui tend à amortir le mouvement (l'hypothèse que cette force varie de façon *linéaire* avec $f'(t)$ n'est réaliste que pour certains mécanismes de frottement, mais si l'amortissement est faible l'approximation peut néanmoins être bonne). Pour les deux coefficients, la constante donnant la force par unité de distance, respectivement par unité de vitesse, doit encore être divisée par la masse du particule pour donner l'accélération nécessaire dans l'équation (18). Dans le cas ou $c_0 < 0$ et $c_1 = 0$ (absence de frottement) c'est l'équation d'un oscillateur harmonique, et dans le cas $c_0 < 0$ et $c_1 < 0$ c'est l'équation d'un oscillateur harmonique amorti.

On admettra ici le fait que pour t_0 fixé les valeurs $f(t_0)$ ainsi que $f^{(i)}(t_0)$ pour $i = 1, \ldots d-1$ déterminent une solution unique de l'équation différentielle, ce qu'on sait montrer avec des méthodes d'analyse. L'application linéaire $T_{t_0} : E \to K^d$ donnée par $f \mapsto \big(f(t_0), f'(t_0), \ldots, f^{(d-1)}(t_0)\big)$ est donc un isomorphisme, et en particulier $\dim(E) = d$. En prenant la dérivée de l'équation (18), on voit que si f en est une solution, f' en est une solution aussi. L'opération de dérivation définit donc un endomorphisme δ_E de E. La matrice A de δ_E par rapport à la base $\mathcal{B}$ correspondant par l'isomorphisme linéaire $T_0 : E \to K^d$ à la base canonique de K^n est donnée par la même expression (15) que dans le problème précédent. Comme dans ce problème les vecteurs propres admettent une description simple, d'où une expression d'une solution quelconque comme somme de vecteurs propres résoudra l'équation différentielle. Un vecteur propre de δ pour λ est une fonction f qui vérifie $f' = \lambda f$ (autrement dit, elle vérifie (18) pour $d = 1$ et $c_0 = \lambda$) ; or les solutions de cette équation sont les multiples scalaires de la fonction exponentielle $t \mapsto e^{\lambda t}$. Les valeurs λ pour lesquelles que cette fonction appartient à E sont les valeurs propres de δ. Comme δ s'exprime par la même matrice A qu'avant, on s'attend à trouver les mêmes valeurs propres ; en effet, dire que $f : t \mapsto e^{\lambda t}$ vérifie (18) veut dire que λ vérifie l'équation (16) : il est racine du polynôme $Q = X^d - c_{d-1}X^{d-1} - \cdots - c_1 X^1 - c_0$. Comme dans le problème précédent on peut conclure que δ sera diagonalisable si et seulement si Q possède d racines distinctes dans K.

Considérons le cas concret de oscillateur harmonique. On aura $Q = X^2 - c_0$ avec $c_0 < 0$; c'est un polynôme quadratique à deux racines purement imaginaires, distinctes et conjuguées complexes, à savoir $\lambda = \sqrt{-c_0}\,\mathbf{i}$ et $\overline{\lambda} = -\sqrt{-c_0}\,\mathbf{i} = -\lambda$. Pour que δ soit diagonalisable il faut donc dans ce prendre $K = \mathbf{C}$, c'est-à-dire considérer des fonctions à valeurs complexes. On aura alors comme base de vecteurs propres $\mathcal{E} = [e_\lambda, e_{-\lambda}]$ où $e_\lambda : t \mapsto e^{\lambda t}$ et $e_{-\lambda} : t \mapsto e^{-\lambda t}$, avec comme matrice de passage de $\mathcal{B}$ vers cette base

$$P = \begin{pmatrix} 1 & 1 \\ \lambda & -\lambda \end{pmatrix}, \qquad \text{dont la matrice l'inverse est} \qquad P^{-1} = \frac{1}{2}\begin{pmatrix} 1 & \lambda^{-1} \\ 1 & -\lambda^{-1} \end{pmatrix}.$$

Pour décrire les solutions de l'oscillateur harmonique, écrivons $c = \sqrt{-c_0}$ de façon que l'équation différentielle s'écrive $f''(t) = -c^2 f(t)$, et $\lambda = c\mathbf{i}$. La solution f avec $f(0) = a$ et $f'(0) = b$, s'écrit

$f = (a, b)_{\mathcal{B}}$, et pour la décrire explicitement il suffit de convertir ses coordonnées sur la base $\mathcal{E}$ des vecteurs propres. Cela se fait en appliquant la matrice P^{-1}, ce qui donne $f = \frac{1}{2}(a + \frac{b}{c\mathbf{i}}, a - \frac{b}{c\mathbf{i}})_{\mathcal{E}}$. Avec les expressions $e^{ct\mathbf{i}}$ et $e^{-ct\mathbf{i}}$ pour les vecteurs de la base $\mathcal{E}$ on trouve

$$f = \frac{1}{2}\big((a + \frac{b}{c\mathbf{i}})e^{ct\mathbf{i}} + (a - \frac{b}{c\mathbf{i}})e^{-ct\mathbf{i}}\big) = a\,\frac{e_\lambda + e_{-\lambda}}{2} + \frac{b}{c}\,\frac{e^{ct\mathbf{i}} - e^{-ct\mathbf{i}}}{2\mathbf{i}} = a\cos(ct) + \frac{b}{c}\sin(ct).$$

Donc les solutions à valeurs réelles de l'équation de l'oscillateur harmonique sont les combinaisons linéaires réelles des fonctions périodiques $t \mapsto \cos(ct)$ et $t \mapsto \sin(ct)$.

Dans le cas de l'oscillateur harmonique amorti, on aura $Q = X^2 - c_1X - c_0$ avec $c_0 < 0$ et $c_1 < 0$, un polynôme quadratique de discriminant $\Delta = c_1^2 + 4c_0$. Si $|\frac{c_1}{2}| < \sqrt{-c_0}$, ce discriminant sera négatif comme pour oscillateur harmonique, et les racines de Q seront encre deux nombres conjugués complexes, mais elles ne sont plus purement imaginaires, car elles ont un partie réelle $\frac{c_1}{2} < 0$. Comme vecteurs propres on aura dans ce cas les fonctions exponentielles complexes $t \mapsto \exp(\frac{c_1 \pm \sqrt{-\Delta}\,\mathbf{i}}{2}t)$, dont on peut former les combinaisons linéaires à valeurs réelles $t \mapsto \exp(\frac{c_1}{2}t)\cos(\frac{\sqrt{-\Delta}}{2}t)$ et $t \mapsto \exp(\frac{c_1}{2}t)\sin(\frac{\sqrt{-\Delta}}{2}t)$. Par rapport à l'oscillateur harmonique non amorti, on observe la présence d'un facteur décroissante exponentielle $\exp(\frac{c_1}{2}t)$, ainsi qu'un facteur $\frac{\sqrt{-\Delta}}{2}$ qui détermine la fréquence des oscillations, et qui est légèrement plus petit que le facteur $c = \sqrt{-c_0}$ vu avant (le fait d'amortir l'oscillateur baisse aussi sa fréquence un peu).

Si par contre $|\frac{c_1}{2}| > \sqrt{-c_0}$, on aura $\Delta > 0$ et les racines seront deux nombres réels distincts qu'on désignera $\lambda_0 = \frac{c_1 + \sqrt{\Delta}}{2}$ et $\lambda_1 = \frac{c_1 - \sqrt{\Delta}}{2}$; comme c_0 et c_1 sont négatifs on aura $\lambda_1 < \lambda_0 < 0$. La solution générale sera donc de la forme $t \mapsto ae^{\lambda_0 t} + be^{\lambda_1 t}$ pour des constantes réelles a, b, donc il n'y aura pas d'oscillation (la solution pas au plus une fois par 0). Dans une telle solution le terme $ae^{\lambda_0 t}$ dominera quand t est grand (sauf si $a = 0$) donc l'évolution ressemble beaucoup à une décroissance exponentielle.

Le cas $|\frac{c_1}{2}| = \sqrt{-c_0}$ est spécial car $\Delta = 0$, et Q possède une racine double $\frac{c_1}{2}$. Cela veut dire que δ n'est pas diagonalisable, même avec $K = \mathbf{C}$. Dans ce cas spécial on parle d'un amortissement critique. On peut vérifier que, à part de la solution exponentielle $t \mapsto \exp(\frac{c_1}{2}t)$, on a comme solution indépendante dans ce cas $t \mapsto t\exp(\frac{c_1}{2}t)$, donc la solution générale sera de la forme $t \mapsto (a + bt)\exp(\frac{c_1}{2}t)$ avec $a, b \in K$. Ceci montre qu'on arrive bien à résoudre l'équation différentielle, même en présence d'une racine double du polynôme Q, mais que les seuls vecteurs propres de δ ne suffisent pas pour cela. La situation dans les cas d'une équation d'ordre supérieur, et celle pour le problème précédent des suites récurrentes, ont des caractéristiques similaires.

Chapitre 3. Corps et anneaux, polynômes.

Dans ce chapitre on introduira la structure algébrique d'un anneau commutatif, une notion obtenue en relaxant les exigences qu'on impose aux corps commutatifs, et on en discutera les exemples les plus fondamentaux: l'anneau des entiers et les anneaux de polynômes. On étudiera les propriétés de ces derniers en particulier lorsque les coefficients des polynômes sont pris dans un corps, car c'est l'anneau $K[X]$ des polynômes à coefficients dans K qui va s'avérer particulièrement utile dans l'étude des endomorphismes de K-espaces vectoriels de dimension finie.

3.1. *Définition de corps et anneaux.*

On a déjà évoqué la notion de corps commutatif au début de ce cours, car elle est un préalable à la notion d'espace vectoriel. On on rappelle la définition, tout en donnant en même temps celle d'un anneau commutatif ; la seul distinction entre les deux est la présence d'un dernier axiome dans la définition d'un corps commutatif. L'axiome précédent stipulant la commutativité de la multiplication peut aussi être omis pour obtenir des notions encore plus générales de corps et anneaux (sans qualification "commutatif").

3.1.1. Définition. *Un anneau est un ensemble R muni d'opérations '$+$' : $R\times R\to R$, '$\times$' : $R\times R\to R$, ainsi que des constantes notées $0,1\in R$, tel que, pour tout $a,b,c\in R$:*
(1) $0+a=a=a+0$ *(0 est élément neutre pour l'addition);*
(2) il existe un élément, noté $-a$, tel que $a+-a=0=-a+a$ (existence de symétriques pour l'addition);
(3) $a+(b+c)=(a+b)+c$ *(associativité de l'addition);*
(4) $a+b=b+a$ *(commutativité de l'addition);*
(5) $1\times a=a=a\times 1$ *(1 est élément neutre pour la multiplication);*
(6) $a\times(b\times c)=(a\times b)\times c$ *(associativité de la multiplication);*
(7) $a\times(b+c)=(a\times b)+(a\times c)$ *et* $(a+b)\times c=(a\times c)+(b\times c)$ *(distributivité à gauche et à droite).*
On appelle R un anneau commutatif si en plus pour tout $a,b\in R$:
(8) $a\times b=b\times a$ *(commutativité de la multiplication).*
On appelle un anneau R (commutatif ou non) un corps si
(9) $1\neq 0$, *et pour tout $a\in R$ avec $a\neq 0$ il existe $b\in R$ tel que $a\times b=1=b\times a$ (existence de symétriques pour la multiplication).*

Certaines règles de calcul habituelles ne se trouvent pas parmi les axiomes, mais sont néanmoins valables dans tous les anneaux car elles peuvent être déduites des axiomes ; notamment la multiplication (à gauche ou à droite) par 0 donne toujours 0. En fait toutes les règles habituelles qui ont la forme d'une égalité, comme par exemple $(x+y)(x-y)=x^2-y^2$ (ou x^2 est une abréviation pour xx) sont valables dans tout anneau commutatif. On se servira librement de ces identités sans les mentionner explicitement.

Parmi les anneaux commutatifs qui ne sont pas des corps, l'exemple le plus basique et le plus important est l'anneau $\mathbf{Z}$ des entiers (dits relatifs, en France) ; dans $\mathbf{Z}$ les seuls éléments possédant un inverse (symétrique pour la multiplication) sont 1 and -1. Il en existe de très nombreux autres exemples d'anneaux commutatifs qui ne sont pas des corps, dont le plus important pour nous sera celui des anneaux de polynômes $K[X]$ de polynômes en X à coefficients dans le corps commutatifs K. On peut remarquer que pour un K-espace vectoriel E l'ensemble $\mathrm{End}(E)$ de ses endomorphismes est une anneau pour les opérations de addition et composition (la constante 1 désigne dans ce cas l'endomorphisme identité) ; cet anneau n'est pas commutatif (dès que $\dim E\geq 2$). Cette structure de $\mathrm{End}(E)$ a déjà été utilisée pour former des "polynômes en $\phi\in\mathrm{End}\,E$" (proposition 2.1.4), mais à part cela elle ne sera pas étudiée. Comme exemple d'un corps non commutatif on mentionne $\mathbf{H}$, le corps des quaternions de Hamilton (une extension des nombres complexes qui est de dimension 4 comme $\mathbf{R}$-espace vectoriel, d'où le nom).

On peut à partir de $\mathbf{Z}$ construire de nouveaux anneaux commutatifs par ce qu'on appelle la réduction modulo un entier particulier $n>0$. Dans ce cas on stipule que les membres de toute paire d'entiers qui diffèrent par un multiple (entier) de n seront considérés équivalents. Ainsi les éléments de la nouvelle structure, qui sera notée $\mathbf{Z}/n\mathbf{Z}$, sont formellement des *classes* d'entiers dont les membres diffèrent entre eux tous par des multiples de n ; dans la pratique une telle classe sera toujours désignée par l'un de ses membres, dit un *représentant* de la classe. Comme tout entier diffère par un multiple de n de son reste r

après division par n, et let n différentes valeurs $0, 1, \ldots, n-1$ parmi lesquelles doit se trouver r sont toutes non équivalentes entre elles, $\mathbf{Z}/n\mathbf{Z}$ contient précisément n éléments (classes d'entiers) dont chacun possède un représentant r unique avec $0 \leq r < n$. On peut munir l'ensemble $\mathbf{Z}/n\mathbf{Z}$ des opérations '+', '−', et '×' par le procédé suivant : pour faire l'opération sur deux classes données, on choisit des représentants r_1, r_2 de ces classes, on applique l'opération correspondante dans $\mathbf{Z}$ sur r_1 et r_2, et finalement on prend la classe du résultat comme résultat dans $\mathbf{Z}/n\mathbf{Z}$.

Pour juste dresser les tables (finies) d'addition, soustraction et multiplication dans $\mathbf{Z}/n\mathbf{Z}$, il suffit de répéter ce procédé en utilisant toute paire $r_1, r_2 \in \{0, \ldots, n-1\}$ comme représentants de leurs classes respectives. Mais il est essentiel de vérifier que d'autres choix de représentants ne contrediront pas le valeurs trouvées : si au lieu de r_1, r_2 on prend les représentants $r_1' = r_1 + na$ et $r_2' = r_2 + nb$ (avec $a, b \in \mathbf{Z}$) des mêmes classes, l'opération effectuée sur r_1', r_2' doit donner un résultat *dans la même classe* que celle effectuée sur r_1, r_2. Cette vérification est facile pour l'addition et soustraction ($r_1' + r_2' = r_1 + r_2 + n(a+b)$ et $r_1' - r_2' = r_1 - r_2 + n(a-b)$), et un peu moins facile pour la multiplication : $r_1' r_2' = r_1 r_2 + n(ar_2 + br_1 + abn)$. On notera que pour cette dernière vérification il est essentiel que r_1, r_2 soient *entiers*, et qu'il n'est donc pas possible de définir de la même façon une multiplication dans les ensembles similaires $\mathbf{R}/n\mathbf{Z}$ ou $\mathbf{Q}/n\mathbf{Z}$. Une fois cette vérification faite, il est simple à montrer que les opérations munissent $\mathbf{Z}/n\mathbf{Z}$ d'une structure d'anneau, avec comme éléments $0, 1 \in \mathbf{Z}/n\mathbf{Z}$ les classes de 0 respectivement de 1.

Pour certaines valeurs de n l'anneau $\mathbf{Z}/n\mathbf{Z}$ est même un corps. Ceci est en fait le cas si et seulement si n est un nombre premier (la preuve n'est pas très difficile, mais on l'admet pour le moment). Pour $n = 2$ c'est particulièrement facile à voir, car (la classe de) 1 est le seul élément non nul, et il est évidemment inversible. Par contre $\mathbf{Z}/1\mathbf{Z}$ n'a qu'un seul élément, qui est donc à la fois 0 et 1 (on dit que c'est l'anneau trivial), et il n'est donc pas un corps car la condition $0 \neq 1$ n'est pas vérifiée.

Les anneaux $\mathbf{Z}/p\mathbf{Z}$ avec p premier nous fournissent une infinité de corps (car il y a une infinité de nombres premiers, comme Euclide le savait déjà). Ces corps, tous finis, sont intéressants par le fait que leurs opérations peuvent être réalisés de façon exacte et très efficace. En informatique on s'intéresse particulièrement au corps $\mathbf{Z}/2\mathbf{Z}$ et aux espaces vectoriels sur ce corps ; on remarque que les seuls multiplications scalaires seront alors par 0 et par 1, ce qui ne semble guère intéressant, mais le fait que tout vecteur v dans un tel espace doit vérifier $v + v = 2v = 0v = \vec{0}$ rend ces espaces vectoriels assez particuliers.

3.2. *Anneaux de polynômes.*

On considéra maintenant l'anneau $K[X]$ des polynômes en X et à coefficients dans K. On connaît les polynômes d'abord dans le manipulation des expressions contenant une inconnue x : de telles expressions formées en utilisant des constantes, addition, soustraction, et multiplication son tes expressions polynomiales. Puis une fonction dont la valeur est donnée par une expression polynomiale en l'argument x de la fonction est une fonction polynomiale. Pour une considération algébrique des polynômes, le statut de x change en celui d'un élément comme les autres, qui ne cache pas une valeur inconnue ou variable ; on l'appelle *indéterminée*, et on l'écrit en majuscule pour marquer son statut. Si $x^2 = 3x + 1$ est une équation qui peut être vérifiée pour certains valeur concrètes de x, l'équation $X^2 = 3X + 1$ est tout simplement *fausse* dans $K[X]$: les polynômes X^2 et $3X + 1$ sont des éléments distincts de $K[X]$. Sans définir formellement $K[X]$, on peut le caractériser par la propriété suivante.

3.2.1. Caractérisation. *$K[X]$ est anneau commutatif contenant le corps K, et un élément X ; ceci étant, l'addition et la multiplication par des éléments de $K \subseteq K[X]$ le munissent d'une structure de K-espace vectoriel, et on peut définir les monômes X^i pour $i \in \mathbf{N}$ par récurrence : $X^0 = 1 \in K$ et $X^{n+1} = XX^n$ pour $n \in \mathbf{N}$. Alors $K[X]$ est caractérisé par le fait que $[(X^i)_{i \in \mathbf{N}}]$ est une K-base de $K[X]$.*

D'après cette propriété, chaque $P \in K[X]$ est de façon unique une combinaison linéaire de monômes. Comme une combinaison linéaire ne peut par définition contenir qu'un nombre fini de termes non nuls, les exposants des monômes ayant un coefficient non nul est borné, et on peut écrire $P = \sum_{i=0}^{d} p_i X^i$ avec $p_i \in K$ pour tout i. Dans cette écriture les p_i sont appelés les coefficients de P. Ils peuvent être nuls, ce qui nous permet d'utiliser toujours une ensemble contigu $1, X, X^2, \ldots, X^d$ de monômes. On peut toujours augmenter d en rajoutant des termes à coefficient nul, ce qui est souvent pratique ; ainsi cette écriture d'un polynôme n'est pas totalement unique, mais c'est la seule liberté qu'on a dans cette écriture.

L'addition de polynômes est donnée par l'addition des coefficients pour chacun des monômes, et la multiplication est dictée par la loi distributive et le fait que $X^iX^j = X^{i+j}$ pour tout i, j :

$$\left(\sum_{i=0}^{d} p_iX^i\right) + \left(\sum_{i=0}^{d} q_iX^i\right) = \sum_{i=0}^{d}(p_i + q_i)X^i \tag{19}$$

$$\left(\sum_{i=0}^{d_1} p_iX^i\right) \times \left(\sum_{j=0}^{d_2} q_jX^j\right) = \sum_{i=0}^{d_1}\sum_{j=0}^{d_2} p_iq_jX^{i+j} = \sum_{k=0}^{d_1+d_2}\left(\sum_{i+j=k} p_iq_j\right)X^k \tag{20}$$

3.2.2. Définition. *Pour un polynôme non nul $P = \sum_{i=0}^{d} p_iX^i \in K[X]$ on définit son degré comme* $\deg(P) = \max\{\, i \mid p_i \neq 0 \,\}$. *Pour le polynôme nul on convient que* $\deg(0) = -\infty$.

Dans l'écriture d'un polynôme on peut arrêter la sommation au degré du polynôme (c'est-à-dire prendre $d = \deg(P)$) mais on n'y est pas obligé ; c'est la raison pour laquelle la définition de degré prend le plus grand indice dont le coefficient est *non nul*, l'indice avec cette propriété ne dépend pas de l'écriture choisie. Si $P = \sum_{i=0}^{N} p_iX^i \in K[X]$ est non nul et $d = \deg(P)$, on appellera le terme p_dX^d le *terme dominant*, et p_d le *coefficient dominant* de P (ils sont non nuls par définition de $\deg(P)$). Un *polynôme unitaire* est un polynôme non nul dont le coefficient dominant est 1.

3.2.3. Proposition. *Pour tout $P, Q \in K[X]$ on a*
(1) $\deg(P + Q) \leq \max(\deg(P), \deg(Q))$, *avec égalité si* $\deg(P) \neq \deg(Q)$, *et*
(2) $\deg(PQ) = \deg(P) + \deg(Q)$.

Preuve. La partie (1) découle de (19) où on peut prendre $d = \max(\deg(P), \deg(Q))$; comme $p_d \neq 0$ ou $q_d \neq 0$, on ne peut avoir $p_d + q_d = 0$ que si $p_d = -q_d \neq 0$, et donc $\deg(P) = \deg(Q)$. La partie (2) découle de (20), car la somme donnant le coefficient dominant du produit, $\sum_{i+j=\deg(P)+\deg(Q)} p_iq_j$, est réduite à $p_{\deg(P)}q_{\deg(Q)}$, le produit non nul (car $p_{\deg(P)} \neq 0$ et $q_{\deg(Q)} \neq 0$) des coefficients dominants. □

La valeur $\deg(0) = -\infty$ a été choisi pour rendre cet énoncé valable quand l'un ou l'autre des polynômes est nul, avec des conventions évidentes pour $-\infty$ (notamment que $-\infty + d = -\infty$ pour $d \in \mathbf{N} \cup \{-\infty\}$). Une conséquence du second point est que $K[X]$ est un anneau intègre, c'est-à-dire qu'un produit de polynômes ne saurait être nul (de degré $-\infty$) que si l'un au moins des polynômes est nul. Il est aussi clair que pour un polynôme P, tout multiple PQ de P par un polynôme $Q \neq 0$ est de degré au moins $\deg(P)$. On note l'ensemble $\{\, PQ \mid Q \in K[X] \,\}$ des multiples de P par $P\,K[X]$, comme on écrit $n\mathbf{Z}$ pour l'ensemble des multiples de n dans $\mathbf{Z}$. On dit que Q et Q' sont congruents modulo P si $Q - Q' \in P\,K[X]$ (c'est une relation d'équivalence), et on note $Q + P\,K[X]$ la classe de tous les polynômes congruents à Q modulo P. Ces notions sont utiles pour comprendre la division euclidienne dans $K[X]$:

3.2.4. Proposition. *Soit $A, B \in K[X]$ avec $B \neq 0$. Alors il existe $Q, R \in K[X]$ avec* $\deg(R) < \deg(B)$ *tels que $A = QB + R$, et le couple (Q, R) est unique.*

Preuve. On montre d'abord qu'une solution (Q, R) sera forcément unique. Si (Q', R') est une autre solution, alors R et R' sont congruents modulo B, car les deux sont congruents à A. Mais on a $\deg(R) < \deg(B)$, et si on avait $R' = R + SB$ avec $S \neq 0$ alors on aurait, d'après la proposition 3.2.3, $\deg(R') = \deg(S) + \deg(B) \geq \deg(B)$, ce qui n'est pas le cas ; par conséquent $R' = R$. Ensuite on a aussi $QB = A - R = A - R' = Q'B$, et comme $K[X]$ est intègre, on obtient $Q' = Q$ aussi: $(Q - Q')B = 0$ implique $Q = Q'$ car $B \neq 0$. Pour l'existence de (Q, R), on prend pour R un élément de plus petit degré possible dans la classe de congruence $A + B\,K[X]$, qui admet donc par définition un Q tel que $A - QB = R$; il suffira de vérifier que $\deg(R) < \deg(B)$. Posons $r = \deg(R)$ et $b = \deg(B)$, et supposons $r \geq b$ pour en tirer une contradiction. Le terme principal de R est de la forme cX^r, et celui de B de la forme $c'X^b$. Le polynôme $R' = R - \frac{c}{c'}X^{r-b}B$ vérifie alors $\deg(R') \leq r$ d'après la proposition 3.2.3(1), et on a fait en sorte que le coefficient de X^r dans R' soit nul, donc en fait $\deg(R') < r$. Comme on a aussi $R' \in R + B\,K[X] = A + B\,K[X]$, ce R' contredit la minimalité supposée de $\deg(R)$. □

La preuve ne donne pas explicitement un algorithme pour trouver (Q, R), car prendre un élément de plus petit degré dans la classe $A + B\,K[X]$, qui est un ensemble infini, n'est pas une opération effective. Ceci dit, la preuve montre comment on peut procéder : en utilisant un polynôme *variable* R, qui sera toujours un représentant de la classe $A + B\,K[X]$, on initialise $R := A$, et tant que $\deg(R) \geq \deg(B)$ on replace R par un représentant de plus petit degré, comme indiqué dans la transition de R vers R' dans la preuve. Comme le degré baisse chaque fois d'au moins 1, on obtiendra $\deg(R) < \deg(B)$ après au plus $\deg(A) - \deg(B) + 1$ itérations, et la valeur finale de R sera le polynôme «reste» cherché ; le quotient Q est la somme des facteurs $\frac{c}{c'}X^{r-b}$ dans les multiples de B qui ont été soustraits dans les différentes étapes.

3.3. *Substitution dans $K[X]$, racines, polynômes annulateurs.*

Une des propriétés fondamentales des polynômes est la possibilité de remplacer X par une constante a quelconque, et d'évaluer l'expression obtenue pour trouver une valeur dans K (on dit qu'on évalue le polynôme en a). En fait, c'est de cette manière qu'on utilise les polynômes dans les équations et fonctions polynomiales. Plus précisément ce n'est pas tellement la possibilité d'évaluer en a qui est importante, mais le fait que ceci est compatible avec les opérations d'anneau : l'évaluation en a d'une somme ou d'un produit de polynômes P, Q donne la somme respectivement le produit de valeurs obtenues en évaluant P et Q individuellement en a. On appelle pour cette raison l'opération d'évaluation $K[X] \to K$ un *homomorphisme* d'anneaux (le terme homomorphisme est utilisé dans beaucoup d'autres contextes, pour des applications qui sont compatibles avec une certaine structure, ici celle d'un anneau commutatif).

La substitution de a pour X sera noté dans ce cours, avec une notation emprunté à l'informatique, "$X := a$" (ce qui prononcé "X devient a"), et la valeur obtenue en effectuant cette opération sur le polynôme $P \in K[X]$ par $P[X := a]$, parfois raccourci à $P[a]$. Dans la plupart des texte cette valeur est notée simplement $P(a)$, mais cette notation qui gomme la distinction entre un polynôme et sa fonction polynomiale n'est pas sans problèmes. D'après la proposition 3.2.4, on peut écrire $P = (X - a)Q + R$ avec $\deg(R) \leq 0$, autrement dit R est une constante. En appliquant l'homomorphisme $X := a$ à cette égalité on obtient $P[X := a] = R$: la valeur constante du reste n'est autre que l'évaluation en a de P. (Si on étudie en détail le calcul de la division euclidienne par $X - a$, on verra qu'il consiste essentiellement en une façon particulière de substituer au fur et à mesure $X := a$ dans le polynôme P.) En particulier on aura $R = 0$ (la division est exacte, et P est un multiple de $X - a$) si et seulement si $p[X := a] = 0$.

3.3.1. Définition. *Quand un polynôme A est un multiple QB de B, on dira également que B divise A, que A est divisible par B, ou que B est un diviseur ou un facteur de A (il faut admettre qu'il y a a beaucoup de façons de dire la même chose). Dans ce cas, si $B \neq 0$, on écrit $Q = A/B$ pour le quotient.*

3.3.2. Définition. *On dit qu'un élément $a \in K[X]$ est une racine du polynôme $P \in K[X]$, et que P est un polynôme annulateur de a, si $P[X := a] = 0$.*

3.3.3. Proposition. *On a $P[X := a] = 0$ si et seulement si P est un multiple $(X - a)Q$ de $X - a$.* □

La possibilité de substituer des valeurs pour X n'est pas limité aux seuls éléments du corps de base K. On pourra aussi substituer des valeurs dans un anneau commutatif A, qui contient K (cette dernière condition sert à pouvoir interpréter les coefficients des polynômes). Par exemple pour les polynômes $P \in \mathbf{R}[X]$, il est possible de prendre $A = \mathbf{C}$ et de substituer un nombre complexe z pour X ; le résultat $P[X := z]$ sera en général un nombre complexe. S'il arrive que $P[X := z] = 0$, on appelle toujours z une racine (complexe) de P, et P un polynôme annulateur de z. Si $z \in \mathbf{C} \setminus \mathbf{R}$, aucun polynôme non nul de degré < 2 dans $\mathbf{R}[X]$ ne peut être annulateur de z, car $c_0 z^0 + c_1 z^1 = 0$ avec $c_0, c_1 \in \mathbf{R}$ voudrait dire que $1 = z^0$ et $z = z^1$ sont linéairement dépendants dans le $\mathbf{R}$-espace vectoriel $\mathbf{C}$, ce qui contredit $z \notin \mathbf{R}$. Par contre $[1, z, z^2]$ est toujours une famille liée dans ce $\mathbf{R}$-espace (car $\dim_{\mathbf{R}}(\mathbf{C}) = 2$), donc il existe un polynôme annulateur de z de degré 2, et même un unique tel polynôme unitaire. Par exemple pour $z = \mathbf{i}$ ce polynôme annulateur unitaire est $X^2 + 1$, et plus généralement le polynôme $U = X^2 - 2aX + a^2 + b^2$ est annulateur de $z = a + b\mathbf{i}$ avec $a, b \in \mathbf{R}$ et $b \neq 0$. Si $P \in \mathbf{R}[X]$ est un polynôme annulateur de z quelconque, on peut écrire $P = QU + R$ avec $\deg(R) < \deg(U) = 2$, et comme $X := z$ annule P et U, cet homomorphisme annule aussi R, ce qui entraîne donc $R = 0$: les polynômes annulateurs de z sont précisément les multiples de U.

Dans ce cours il sera surtout utile de substituer pour X un endomorphisme ϕ d'un K-espace vectoriel E, comme dans la proposition 2.1.4. Si $P = \sum_{i=0}^{d} c_i X^i$ on définit $P[X := \phi] = P[\phi] = \sum_{i=0}^{d} c_i \phi^i$, où on convient que $\phi^0 = \mathrm{id}_E$ pour tout $\phi \in \mathrm{End}(E)$. Si $P[\phi] = 0$ on appelle P un polynôme annulateur de ϕ (mais on n'appelle pas ϕ une racine de P).

Il est important de se rendre compte que l'expression $\sum_{i=0}^{d} c_i \phi^i$ est calculée dans l'anneau non commutatif $\mathrm{End}(E)$, ce qui rend moins évident que la substitution transforme la multiplication (commutative!) de polynômes en la composition d'endomorphismes. Mais la seule forme de commutation vraiment utilisée dans la multiplication de polynômes est celle entre coefficients et monômes, pour récrire des produits de termes $c_i X^i d_j X^j$ comme $c_i d_j X^{i+j}$, et cela reste valable après la substitution $X := \phi$, car ϕ^i commute bien avec la multiplication *scalaire* par d_j (car ϕ^i est linéaire), d'où $c_i \phi^i d_j \phi^j = c_i d_j \phi^{i+j}$. On peut alors vérifier facilement que $X := \phi$ définit un homomorphisme d'anneaux $K[X] \to \mathrm{End}(E)$:

$$(P + Q)[\phi] = P[\phi] + Q[\phi], \tag{21}$$

$$(PQ)[\phi] = P[\phi] \circ Q[\phi] = Q[\phi] \circ P[\phi]. \tag{22}$$

3.3.4. Proposition. *Pour $\phi \in \mathrm{End}(E)$, il existe un polynôme U annulateur de ϕ (et qu'on peut choisir unitaire), tel que pour $P \in K[X]$ on ait $P[\phi] = 0$ si et seulement si P est un multiple QU de U.*

Preuve. Comme $\dim(\mathrm{End}(E)) = \dim(E)^2$ est fini, la famille d'endomorphismes $\mathrm{id}_E = \phi^0, \phi, \phi^2, \ldots$ ne peut pas rester libre. La première dépendance linéaire entre ces puissances de ϕ donne un polynôme annulateur U de degré minimal. Pour $P \in K[X]$ on peut d'après la proposition 3.2.4 trouver Q, R tels que $P = QU + R$ et $\deg(R) < \deg(U)$. Si P est annulateur de ϕ, alors $0 = P[\phi] = Q[\phi]U[\phi] + R[\phi] = R[\phi]$ en utilisant (21),(22), ce qui compte de $\deg(R) < \deg(U)$ entraîne $R = 0$, et donc $P = QU$. □

Si $P \in K[X]$ possède une racine $a \in K$, on a vu que P est multiple de $X - a$: il existe $P' \in K[X]$ tel que $P = (X - a)P'$. Alors toute *autre* racine a' de P doit aussi être racine de P' : il suffit d'appliquer $X := a'$, ce qui donne $0 = P[a'] = (X - a)[X := a']\, P'[a'] = (a' - a)P'[a']$, et d'observer que $a' - a \neq 0$. Réciproquement toute racine de P' est racine de P. Il n'est pas exclus que la première racine a soit également racine de P' ; dans ce cas on dit qu'elle est *racine multiple* de P. On peut répéter cette opération de mettre en facteur des polynômes $X - a_i$ pour les racines de P', et on trouvera ainsi donc toutes les racines de P dans K. Il reste un facteur Q qui n'a aucune racine dans K. Dans le cas $K = \mathbf{C}$ le théorème d'Alembert–Gauss dit que Q doit être un polynôme constant, et comme les autres facteurs sont tous des polynômes unitaires, la valeur constante de Q est le coefficient dominant de P.

Dans le cas $K = \mathbf{R}$, le degré de Q peut encore être élevé, mais dans ce cas Q possède au moins une racine *complexe* z. On a vu que z possède un polynôme annulateur unitaire U dans $\mathbf{R}[X]$ de degré au plus 2 et qui divise tout polynôme annulateur de z, donc en particulier Q ; alors U n'a pas de racines réelles (car elles seraient aussi racines de Q qui n'en a pas) donc U est un polynôme quadratique de discriminant négatif. La division de Q par U est sans reste, et $Q/U \in \mathbf{R}[X]$. En itérant, on trouve :

3.3.5. Proposition. *Soit K un corps commutatif et $P \in K[X]$ un polynôme non nul.*

(1) On peut écrire, de façon unique à l'ordre des $a_i \in K$ près, $P = (X - a_1) \ldots (X - a_k)Q$ avec $k \geq 0$, où $Q \in K[X]$ est sans racine dans K. Dans ce cas $\{a_1, \ldots, a_k\}$ est l'ensemble de toutes les racines de P dans K, et une racine de multiplicité m dans P est présente m fois parmi $a_1, \ldots, a_k$.

(2) Le nombre k des racines comptées avec multiplicité vérifie $k + \deg(Q) = \deg(P)$, et donc $k \leq \deg(P)$.

(3) Si $K = \mathbf{C}$, le polynôme Q du point (1) est une constante, le coefficient dominant de P, et $k = \deg(P)$.

(4) Si $K = \mathbf{R}$, le polynôme Q se décompose en un produit d'une constante (le coefficient dominant de P) et un nombre $s \geq 0$ de polynômes quadratiques unitaires à discriminant strictement négatif. □

3.4. *Quelques éléments d'arithmétique dans $K[X]$.*

Les polynômes dans $\mathbf{C}[X]$ et dans $\mathbf{R}[X]$ admettent une décomposition en facteurs qui ressemble celle des entiers strictement positifs en facteurs premiers. Le but de cette section est de montrer qu'une telle décomposition existe pour tout corps commutatif K (mais sans que les facteurs soient limités à être de degré 1 ou 2), et que d'autres propriétés de $\mathbf{Z}$ ont un pendant dans les anneaux $K[X]$.

On poursuit sur la notion de divisibilité, introduite dans la définition 3.3.1. Même si selon cette définition tout polynôme divise 0, qui lui ne divise aucun autre polynôme, on exclut en général le polynôme nul des considérations de divisibilité, par exemple dans la proposition 3.3.5. Les polynômes constants non nuls sont aussi exceptionnels, à l'autre extrémité : ils divisent tout polynôme, et ils ne sont divisibles par aucun polynôme non constant. Cela explique leur rôle particulier dans la définition suivante.

3.4.1. Définition. *Un polynôme non nul $P \in K[X]$ est réductible s'il s'écrit comme le produit de deux polynômes non constants. Un polynôme non constant et non réductible est un polynôme irréductible.*

Ces notions correspondent à celles de nombres composés et premiers dans $\mathbf{Z}$. Les polynômes constants dans $K[X]$ ne sont ni réductibles ni irréductibles, tout comme les nombres ± 1 ne sont ni composés ni premiers. On peut toujours modifier une décomposition multiplicative d'un polynôme en multipliant l'un des facteurs par une constante non nulle, et un autre par la constante inverse. Ceci nécessite des formulations prudentes pour parler de l'unicité de telles décompositions. Parfois on peut éviter des formulations compliquées en ne considérant que des facteurs irréductibles *unitaires*, comme on simplifie les formulations concernant la factorisation dans $\mathbf{Z}$ en se limitant aux nombres strictement positifs.

3.4.2. Proposition. *Tout polynôme non constant s'écrit comme un produit de polynômes irréductibles.*

Preuve. Par récurrence sur le degré : soit le polynôme est irréductible et donc le produit d'un seul facteur, soit il est réductible, et on l'écrit comme produit de deux facteurs non constants de degré plus bas, lesquels peuvent être écrits chacun comme produit de polynômes irréductibles par hypothèse de récurrence. □

Si cette décomposition est facile à obtenir du point de vue théorique, ce résultat n'est pas effectif, et il n'existe pas de méthode générale et effective de trouver une telle décomposition.

3.4.3. Définition. *Un polynôme $P \in K[X]$ est scindé (sur K) s'il s'écrit comme un produit d'une constante non nulle et des facteurs de la forme $X - \alpha$ avec $\alpha \in K$.*

Un polynôme de degré 1 est toujours irréductible (le degré du produit de deux polynômes non constants est au moins 2) et scindé, mais un polynôme irréductible de degré > 1 n'est pas scindé. D'après la proposition 3.3.5(3), *tout* polynôme non nul dans $\mathbf{C}[X]$ est scindé (sur $\mathbf{C}$). En fait un corps K est algébriquement clos si et seulement si tout polynôme non nul dans $K[X]$ est scindé : un polynôme scindé non constant possède évidemment au moins une racine, et réciproquement si K est algébriquement clos on montre par récurrence sur le degré que tout polynôme non nul est scindé. Sur un corps algébriquement clos tout polynôme irréductible est donc de degré 1, et d'après la proposition 3.4.2, la réciproque est aussi vraie. Pour certains problèmes, la possibilité d'avoir des polynômes avec des facteurs irréductibles de degré > 1 (donc sans racines) peut poser une difficulté ; on peut la contourner en supposant $K = \mathbf{C}$, mais souvent il suffit de supposer simplement qu'un polynôme spécifique soit scindé sur K.

Pour un polynôme P de degré > 1, la condition que P n'ait pas de racines est nécessaire pour que P soit irréductible. Mais attention, cette condition n'est pas suffisante : en multipliant deux (voire plus) polynômes irréductibles, tous de degré ≥ 2, le résultat sera réductible sans avoir des racines. La condition est suffisante seulement si $\deg(P) = 2$ ou $\deg(P) = 3$, car dans ces cas si P est réductible dans $K[X]$, il doit forcément avoir un facteur de degré 1, et donc une racine dans K.

3.4.4. Théorème/définition. *Pour $A, B \in K[X]$ non nuls, il existe un polynôme unitaire unique D tel que D divise à la fois A et B, et tel qu'il existe $U, V \in K[X]$ avec $D = UA+VB$. On écrit $D = \operatorname{pgcd}(A, B)$.*

Preuve. Soit D_0 un polynôme non nul de degré minimal qui s'écrive $D_0 = SA + TB$; on prend alors $D = \alpha^{-1}D_0$, où $\alpha \in K$ est le coefficient dominant de D_0 (pour que D soit unitaire). On a bien $D = UA + VB$, avec $U = \alpha^{-1}S$ et $V = \alpha^{-1}T$. Pour la première condition, si D ne divisait pas A, il y aurait un reste $R \neq 0$ dans la division de A par D, pour lequel $R = (S - Q)A + TB$ où Q est le quotient de cette division. Mais $\deg(R) < \deg(D) = \deg(D_0)$ contredit alors le choix minimal de D_0. Donc D divise A, et par un argument similaire il divise aussi B. Cela montre l'existence de D ; pour l'unicité supposons que D' vérifie les mêmes conditions. Comme D' divise A et B, il divise aussi $D = UA + VB$, et par symétrie D divise aussi D' ; les polynômes D, D' étant unitaires, ces conditions forcent $D = D'$. □

On appelle $D = \text{pgcd}(A, B)$ le "plus grand commun diviseur" de A et de B. Il s'agit clairement d'un diviseur commun de A et de B, et dans la démonstration on a vu que tout diviseur commun D' de A et de B divise aussi D. Cela donne certainement $\deg(D') \leq \deg(D)$, et dans le cas $\deg(D') = \deg(D)$ la division D/D' est exacte et donne comme quotient un polynôme constant non nul. Donc D a le plus grand degré possible pour un commun diviseur, et à un scalaire près il est le seul commun diviseur de degré $\deg(D)$. Le nom "plus grand commun diviseur" est donc justifié, même s'il ne transmet pas l'information pourtant importante que les autres communs diviseurs *divisent* tous D.

Avec cette définition du pgcd, le "théorème de Bezout", qui affirme l'existence des coefficients S, T dits de Bezout tel que $\text{pgcd}(A, B) = SA + TB$ pour tout $A, B \in K[X]$, est une évidence. Ce qui n'est pas évident, mais néanmoins vrai, est qu'il est possible de calculer explicitement $\text{pgcd}(A, B)$ ainsi que de tels coefficients de Bezout. Cela se fait par une version étendue de l'algorithme d'Euclide : on maintient un couple de polynômes (P, Q) chacun de la forme $x = S_x A + T_x B$, initialisé $(P, Q) = (A, B)$; tant que $Q \neq 0$ on remplace (P, Q) par (Q, R), où R est le reste de la division de P par Q, et une fois $Q = 0$ est atteint on aura $\text{pgcd}(A, B) = c^{-1}P$ où c est le coefficient dominant de P (pour que le résultat soit unitaire). Le reste est trouvé comme $R = P - FQ$ pour un certain (quotient) $F \in K[X]$, et on trouve le coefficient S_R de A dans R en termes des coefficients correspondants S_P, S_Q de P et de Q en calculant $S_R = S_P - FS_Q$, et pareillement pour le coefficient T_R de B dans R.

On illustre le procédé par le calcul des coefficients de Bezout pour le pgcd de deux nombres entiers, pour lequel l'algorithme est identique, mais les valeurs qui interviennent, des entiers, sont plus simples que les polynômes qu'il faudrait manipuler pour calculer des coefficients de Bezout dans $K[X]$. Voici une trace détaillée de l'algorithme d'Euclide pour calculer $\text{pgcd}(267, 93)$ et ses coefficients de Bezout s, t :

$$\begin{aligned}
267 &= 1 \times 267 \; -0 \times 93 \\
93 &= 0 \times 267 \; +1 \times 93 \\
267 \bmod 93 = 81 \quad &= 267 - 2 \times 93 = 1 \times 267 \; -2 \times 93 \\
93 \bmod 81 = 12 \quad &= 93 - 1 \times 81 = -1 \times 267 \; +3 \times 93 \\
81 \bmod 12 = 9 \quad &= 81 - 6 \times 12 = 7 \times 267 \; -20 \times 93 \\
12 \bmod 9 = 3 \quad &= 12 - 1 \times 9 = -8 \times 267 \; +23 \times 93 \\
9 \bmod 3 = 0 \quad &= 9 - 3 \times 3 = 31 \times 267 \; -89 \times 93 \\
\text{pgcd}(297, 93) = 3 \quad &= s \times 267 \; +t \times 93 \quad \text{avec } (s, t) = (-8, 23).
\end{aligned}$$

La justification que $\text{pgcd}(A, B)$ ainsi calculé divise effectivement A et B (c'est le seul point non évident dans l'algorithme) peut être donnée en utilsant un «invariant» de l'algorithme, une quantité associée au couple (P, Q) qui ne change pas quand un passe au couple suivant. Le $\text{pgcd}(P, Q)$ est un tel invariant, mais pour le voir il est le plus simple de s'appuyer sur l'invariant plus abstrait qu'est l'*ensemble* des polynômes de la forme $SP + TQ$ avec $S, T \in K[X]$ (avec un terme qu'on a évité d'introduire dans ce cours on dit aussi "l'idéal de $K[X]$ engendré par P, Q") dont $\text{pgcd}(P, Q)$ est par définition l'élément unitaire de plus petit degré ; l'invariance de cet ensemble se montre en exprimant les nouveaux P, Q en termes des anciens et vice versa. Si on avait défini $\text{pgcd}(A, B)$ comme le commun diviseur unitaire de A, B de plus grand degré, il serait naturel d'utilser comme invariant l'ensemble des diviseurs communs de P, Q, dont l'invariance se montre par un argument très similaire. En tout état de cause, ces invariants abstraits ne figurent pas dans l'algorithme, mais ils sont seulement invoqués pour justifier celui-ci.

3.4.5. Définition. *Deux polynômes $A, B \in K[X]$ sont premiers entre eux si* $\text{pgcd}(A, B) = 1$.

On verra que pour une famille de polynômes, la condition d'être premiers entre eux deux à deux permettra certaines décompositions qui seront utiles pour la réduction d'endomorphismes d'un espace vectoriel. Par exemple la proposition 2.1.3 est liée au fait que deux polynômes $X - \lambda$ et $X - \mu$ sont premiers entre eux dès que $\lambda \neq \mu$, un fait qui est une conséquence évidente de l'absence de diviseurs communs de degré ≥ 1 de ces polynômes. Ce qui est remarquable, est que dans $K[X]$ la condition que A, B sont premiers entre eux soit à la fois caractérisée de façon négative par l'absence de diviseurs communs non constants, et de façon positive par l'existence de coefficients de Bezout S, T tels que $1 = SA + TB$, quelle propriété est souvent exploité dans les démonstrations.

3.4.6. Lemme de Gauss. *Si $P \in K[X]$ divise AB et P et A sont premiers entre eux, alors P divise B.*

Preuve. On utilise une relation de Bezout $\text{pgcd}(P, A) = 1 = UP + VA$, ce qui permet d'écrire $B = 1B = UPB + VAB$; alors par hypothèse P divise chaque terme de la somme, d'où le résultat. □

3.4.7. Corollaire. *Si $P_1, \ldots, P_k \in K[X]$ sont premiers entre eux deux à deux, alors chaque polynôme P_i et le produit $\prod_{j \neq i} P_j = P_1 \ldots P_{i-1} P_{i+1} \ldots P_k$ des autres polynômes sont aussi premiers entre eux.*

Preuve. Par symétrie il suffit de prouver ceci pour $i = 1$, ce qui se fait facilement par récurrence sur k. Pour $k \leq 2$ le résultat est évident. Pour $k > 2$ montrons que tout diviseur commun D de P_1 et $P_2 \ldots P_k$ est constant. On a $\text{pgcd}(D, P_2) = 1$ car D divise P_1 et $\text{pgcd}(P_1, P_2) = 1$ (tout diviseur commun de D et P_2 est aussi diviseur commun de P_1 et P_2). Donc d'après le lemme 3.4.6, D divise $P_3 \ldots P_k$, et comme D est alors diviseur commun de P_1 et de $P_3 \ldots P_k$, il doit être constant d'après l'hypothèse de récurrence. □

3.4.8. Lemme d'Euclide. *Si un polynôme irréductible P divise un produit AB de polynômes, il divise au moins un des facteurs A, B.*

Preuve. Montrons que si P ne divise pas A, alors il divise B. Comme P est irréductible, ses seuls diviseurs non constants sont ses propres multiples cP par une constante non nulle, mais ceux-ci ne divisent pas A par hypothèse ; on a donc $\text{pgcd}(P, A) = 1$. Alors il suffit d'appliquer le lemme 3.4.6. □

Le lemme d'Euclide se généralise par une récurrence facile aux produits de plusieurs facteurs (qu'on peut voir comme le premier facteur multiplié par le produit des autres) : si un polynôme irréductible P divise un tel produit, il divise au moins un de ses facteurs (c'est pareil pour un nombre premier dans $\mathbf{Z}$).

On peut maintenant montrer que la décomposition de polynômes en facteurs irréductibles est essentiellement unique. C'est l'équivalent pour les polynômes de la factorisation unique des entiers $n > 0$ en nombres premiers, et la démonstration est presque identique à celle utilisée dans le case de $\mathbf{Z}$.

3.4.9. Théorème de factorisation unique dans $K[X]$. *Tout polynôme $Q \neq 0$ dans $K[X]$ s'écrit comme le produit de son coefficient dominant c et d'un produit de $k \geq 0$ polynômes irréductibles unitaires. Cette écriture est unique à l'ordre des facteurs irréductibles près.*

Preuve. Pour l'existence on applique la proposition 3.4.2 au polynôme Q, on rend unitaire chaque facteur irréductible en le divisant par son coefficient dominant, et le produit de ces coefficients dominants donne c. Il reste à montrer l'unicité de la factorisation dans le cas $c = 1$. Supposons donc qu'on ait deux factorisations $P_1 \ldots P_k = Q = P'_1 \ldots P'_l$ avec les P_i et P'_i unitaires et irréductibles. On prouvera par récurrence sur k que $(P'_1, \ldots, P'_l)$ est une permutation de $(P_1, \ldots, P_k)$ (et en particulier $l = k$). Si $k = 0$ alors $Q = 1$, et donc $l = 0$. Sinon P_1 divise Q, donc d'après le lemme d'Euclide (généralisé comme ci-dessus), P_1 divise au moins un des facteurs P'_j, et comme P'_j est irréductible cela force $P_1 = P'_j$. Alors $P_2 \ldots P_k = Q/P_1 = P'_1 \ldots P'_{i-1} P'_{i+1} \ldots P'_l$ et on termine en appliquant l'hypothèse de récurrence. □

Si $Q = P_1 \ldots P_k$ est une factorisation, et si S divise Q, alors S est à un facteur constant près le produit d'une partie des facteurs irréductibles P_i. En particulier si Q est scindé, S est aussi scindé.

3.5. *Décomposition des noyaux.*

Dans cette section on donne la principale application de l'arithmétique des polynômes pour les questions de réduction d'endomorphisme. On fixe $\phi \in \text{End}(E)$ dans toute la section, avec E un espace vectoriel de dimension finie. On a vu dans la proposition 3.3.4 que ϕ possède un polynôme annulateur non nul, c'est-à-dire $P \neq 0$ avec $P[\phi] = 0$. Ici on regarde la situation où $P[\phi]$ n'est peut-être pas entièrement nul, mais s'annule sur un sous-espace de E.

Une première observation importante est que de tels sous-espaces sont toujours ϕ-stables, ce qui permettra de considérer l'endomorphisme du sous-espace obtenu par restriction de ϕ (et dont par définition P est un polynôme annulateur). Pour l'équité entre noyaux et images, on rajoutera un second cas qui se démontre tout aussi facilement. On pourrait encore généraliser l'énoncé en remplaçant $P[\phi]$ par tout endomorphisme qui *commute* avec ϕ, mais cette généralisation ne sera pas utilisée dans ce cours.

3.5.1. Proposition. *Pour $P \in K[X]$ les sous-espaces $\mathrm{Ker}(P[\phi])$ et $\mathrm{Im}(P[\phi])$ de E sont ϕ-stables.*

Preuve. Soit, $v \in \mathrm{Ker}(P[\phi])$; vérifions si $\phi(v) \in \mathrm{Ker}(P[\phi])$. Cela découle de $P \cdot_\phi \phi(v) = (PX) \cdot_\phi v = (XP) \cdot_\phi v = \phi(P \cdot_\phi v) = \phi(\vec{0}) = \vec{0}$. Ensuite vérifions si $v \in \mathrm{Im}(P[\phi])$ que $\phi(v) \in \mathrm{Im}(P[\phi])$. Par hypothèse on peut écrire $v = P \cdot_\phi w$, et on aura $\phi(v) = \phi(P \cdot_\phi w) = (XP) \cdot_\phi w = (PX) \cdot_\phi w = P \cdot_\phi \phi(w) \in \mathrm{Im}(P[\phi])$. □

3.5.2. Proposition. *Si D divise P dans $K[X]$, alors $\mathrm{Ker}(D[\phi]) \subseteq \mathrm{Ker}(P[\phi])$.*

Preuve. Si $P = QD$ et $v \in \mathrm{Ker}(D[\phi])$, alors $P \cdot_\phi v = Q \cdot_\phi (D \cdot_\phi v) = Q \cdot_\phi (0) = 0$, donc $v \in \mathrm{Ker}(P[\phi])$. □

Notre résultat principal dira que le noyau associé à un produit de polynômes se décompose comme *somme directe* de noyaux associés aux polynômes individuels, à condition que ces polynômes sont premiers entre eux deux à deux. Il n'est pas surprenant qu'on ait besoin d'une telle condition, car pour tout facteur commun D de deux de ces polynômes P, Q qu'on peut avoir, on a $\mathrm{Ker}(D[\phi]) \subseteq \mathrm{Ker}(P[\phi]) \cap \mathrm{Ker}(Q[\phi])$ d'après la proposition 3.5.2 ; ceci empêche la somme des noyaux d'être directe dès que $\mathrm{Ker}(D[\phi])$ n'est pas réduit à $\{0\}$. Il est remarquable que cette simple condition suffisse. Le lemme suivant montre qu'on obtient une décomposition du noyau en somme directe dans le cas de deux facteurs premiers entre eux.

3.5.3. Lemme. *Si $P, Q \in K[X]$ sont premiers entre eux, alors $\mathrm{Ker}((PQ)[\phi]) = \mathrm{Ker}(P[\phi]) \oplus \mathrm{Ker}(Q[\phi])$, et les projections de $\mathrm{Ker}((PQ)[\phi])$ sur $\mathrm{Ker}(P[\phi])$ et sur $\mathrm{Ker}(Q[\phi])$ selon la somme directe peuvent être écrites comme des restrictions à $\mathrm{Ker}((PQ)[\phi])$ de certains polynômes en ϕ.*

Preuve. Comme l'énoncé ne parle que de la restriction de ϕ à $\mathrm{Ker}((PQ)[\phi])$ (qui est ϕ-stabled'après la proposition 3.5.1 et qui contient $\mathrm{Ker}(P[\phi])$ et $\mathrm{Ker}(Q[\phi])$ d'après la proposition 3.5.2), on pourra sans perte de généralité remplacer E par $\mathrm{Ker}((PQ)[\phi])$ et ϕ par sa restriction à ce noyau. On suppose donc désormais que $(PQ)[\phi] = 0$. Comme P, Q sont premiers entre eux il existe des coefficients de Bezout $U, V \in K[X]$ tels que $\mathrm{pgcd}(P, Q) = 1 = UP + VQ$; choisissons les, et posons $\pi_1 = (VQ)[\phi]$ et $\pi_2 = (UP)[\phi]$. On a donc $\pi_1 + \pi_2 = \mathrm{id}_E$, ainsi que $\mathrm{Im}(\pi_1) \subseteq \mathrm{Ker}(P[\phi])$ et $\mathrm{Im}(\pi_2) \subseteq \mathrm{Ker}(Q[\phi])$, car PQ divise PVQ et QUP, pendant que $(PQ)[\phi] = 0$. Alors d'une part pour $v \in E$ on a $v = \pi_1(v) + \pi_2(v) \in \mathrm{Ker}(P[\phi]) + \mathrm{Ker}(Q[\phi])$, ce qui montre que $\mathrm{Ker}(P[\phi]) + \mathrm{Ker}(Q[\phi]) = E$, et d'autre part pour $v_1 \in \mathrm{Ker}(P[\phi]) \subseteq \mathrm{Ker}(\pi_2)$ et $v_2 \in \mathrm{Ker}(Q[\phi]) \subseteq \mathrm{Ker}(\pi_1)$ on a $\pi_1(v_1 + v_2) = \pi_1(v_1) = (\pi_1 + \pi_2)(v_1) = v_1$ et $\pi_2(v_1 + v_2) = \pi_2(v_2) = (\pi_1 + \pi_2)(v_2) = v_2$, donc la somme $E = \mathrm{Ker}(P[\phi]) + \mathrm{Ker}(Q[\phi])$ est directe, avec π_1, π_2 comme les projections sur ses facteurs. Or an a vu que $\pi_1 = (VQ)[\phi]$ et $\pi_2 = (UP)[\phi]$ sont des polynômes en ϕ. □

La restriction au sous-espace $\mathrm{Ker}((PQ)[\phi])$ faite au début de la démonstration facilite sa formulation sensiblement. Si on avait défini π_1, π_2 sur un espace où $(PQ)[\phi] \neq 0$, on n'aurait pas pu dire par exemple $\mathrm{Im}(\pi_1) \subseteq \mathrm{Ker}(P[\phi])$, mais seulement que $\pi_1(v) \in \mathrm{Ker}(P[\phi])$ pour tout $v \in \mathrm{Ker}((PQ)[\phi])$; la démonstration n'en serait pas très différente, mais moins transparente. On peut aussi noter que la démonstration donnée ne dépend pas de l'hypothèse que E ou $\mathrm{Ker}((PQ)[\phi])$ soient de dimension finie.

3.5.4. Théorème de décomposition des noyaux. *Si $P_1, \ldots, P_l \in K[X]$ sont premiers entre eux 2 à 2 et $\phi \in \mathrm{End}(E)$, alors*

$$\mathrm{Ker}((P_1 \cdots P_l)[\phi]) = \mathrm{Ker}(P_1[\phi]) \oplus \cdots \oplus \mathrm{Ker}(P_l[\phi]),$$

et les projections de la somme sur chacun des facteurs sont des restrictions de certains polynômes en ϕ.

Preuve. Par récurrence sur l ; pour $l \leq 1$ c'est évident. Pour $l \geq 2$ on applique le lemme 3.5.3 avec $P = P_1 \cdots P_{l-1}$ et $Q = P_l$, qui sont premiers entre eux, d'après le corollaire 3.4.7. On obtient

$$\mathrm{Ker}((P_1 \cdots P_l)[\phi]) = \mathrm{Ker}((P_1 \cdots P_{l-1})[\phi]) \oplus \mathrm{Ker}(P_l[\phi]) = \mathrm{Ker}(P_1[\phi]) \oplus \cdots \oplus \mathrm{Ker}(P_{l-1}[\phi]) \oplus \mathrm{Ker}(P_l[\phi]),$$

et la composition des projections réalise celle $\mathrm{Ker}((P_1 \cdots P_l)[\phi]) \to \mathrm{Ker}(P_j[\phi])$ par un polynôme en ϕ. □

Comme on avait déjà observé, ce théorème s'applique en particulier à des polynômes $P_i = X - \lambda_i$ de degré 1, pour $i = 1, \ldots, k$ où $\lambda_1, \ldots, \lambda_k \in K$ sont des valeurs distinctes, et dans ce cas la conclusion est l'énoncé de la proposition 2.1.3, avec le complément que la somme des espaces propres est égal au noyau de l'endomorphisme $(X - \lambda_1) \ldots (X - \lambda_k)[X := \phi]$. Le théorème montre que la primalité entre eux des polynômes λ_i est l'ingrédient essentiel de cette proposition, et donne en même temps un résultat plus largement applicable que la proposition, notamment pour des endomorphismes non diagonalisables.

Chapitre 4. Déterminants.

Considérons un système de k équations linéaires en $n \geq k$ inconnues. On sait que par l'échelonnement du système, on saura déterminer l'ensemble de solutions du système. Le système n'aura une solution unique que si $k = n$, et dans ce cas seulement si pendant l'échelonnement on réussit pour chaque inconnue successive a trouver une équation dans laquelle son coefficient est non nul, pour pouvoir servir en tant que "pivot de Gauss". La méthode de l'échelonnement est basée sur la possibilité de diviser par ce pivot dès qu'il est non nul. La méthode fonctionne bien quand tous les coefficients des inconnues dans le système sont des constantes, mais on peut rencontrer des problèmes quand ces coefficients dépendent d'un ou plusieurs paramètres. Dans ce cas le choix d'un pivot peut dépendre des valeurs de ces paramètres, on est amené à discuter plusieurs cas de figure, et la situation peut devenir rapidement très compliquée.

Le système définissant (en coordonnées) un espace propre d'un endomorphisme est justement de cette nature, car il s'agit du système $(A - \lambda I_n) \cdot v = 0$ où A est la matrice de ϕ dans une base, et $\lambda \in K$ est un paramètre. On s'intéresse notamment aux valeurs de λ pour laquelle la solution de ce système *n'est pas unique*, car pour une système homogène comme celui-ci on a toujours la solution triviale $v = 0$. Il s'avère que pour les systèmes carrés ($k = n$), l'unicité de leur solution dépend de la valeur d'*une seule expression* en les coefficients, appelé le *déterminant* du système. Si et seulement si ce déterminant a une valeur est non nulle, le système possède une solution unique ; il s'appelle alors un *système de Cramer*. Cette propriété du déterminant paraît assez magique du point de vue de l'algorithme de Gauss, et c'est le but de ce chapitre de motiver l'existence et la définition du déterminant. Pour l'équation pour les espaces propres, c'est l'annulation du déterminant du système $(A - \lambda I_n) \cdot v = 0$, quel déterminant donne un polynôme en λ appelé polynôme caractéristique de A, qui caractérisera les valeurs propres.

La question de l'unicité de la solution $x = 0$ d'un système linéaire homogène écrit sous forme d'une équation vectorielle $A \cdot x = 0$, est équivalente à celle qui demande si les colonnes de A sont linéairement indépendantes, car $A \cdot x$ est juste la combinaison linéaire de ces colonnes avec les composantes de x comme coefficients. Une autre façon de dire ce qui précède est donc que pour une famille de n vecteurs dans K^n (les colonnnes de A), le déterminant est une expression en leur n^2 coordonnées qui s'annule si et seulement si la famille est liée.

4.1. *Déterminants en dimension $n \leq 3$, formes linéaires.*

Il est instructif de considérer d'abord la situation quand la dimension n est petite. Les cas $n = 0, 1$ n'étant guère intéressants, prenons $n = 2$. Il s'agit de décider si une famille $[(a, b), (x, y)]$ de vecteurs de K^2 est libre. Supposons que (a, b) soit fixé et non nul, alors la famille sera liée seulement si (x, y) est un multiple scalaire de (a, b). En fonction de (x, y), le déterminant doit donc être une fonction qui s'annule précisément sur ces multiples scalaires, et comme ils forment un sous-espace de dimension $1 = 2 - 1$, on pourra trouver une fonction *linéaire* qui fait l'affaire. En effet il suffit de prendre une fonction non nulle qui s'annule en (a, b), comme celle de matrice $(-b \quad a)$, c'est-à-dire la fonction $(x, y) \mapsto -bx + ay$. Ce qui est remarquable dans ce choix d'application linéaire, est que la fonction choisie dépende elle même de façon linéaire du vecteur (a, b), c'est-à-dire que l'application $K^2 \to \mathcal{L}(K^2, K)$ qui associe à (a, b) cette fonction $(x, y) \mapsto -bx+ay$ est une application linéaire (or rappelle que l'ensemble $\mathcal{L}(E, F)$ des applications linéaires $E \to F$ est lui-même un K-espace vectoriel, quels que soient les K-espaces vectoriels E, F). Cela implique en particulier que cette application est nulle quand on prend $(a, b) = (0, 0)$, et c'est juste ce qu'on veut : dans ce cas la famille sera liée, quel que soit (x, y). Le déterminant d'une matrice 2×2, défini par

$$\begin{vmatrix} a & x \\ b & y \end{vmatrix} \overset{\text{déf}}{=} \det \begin{pmatrix} a & x \\ b & y \end{pmatrix} = (-b \quad a) \cdot \begin{pmatrix} x \\ y \end{pmatrix} = ay - bx, \tag{23}$$

a donc la propriété de s'annuler si et seulement si (a, b) et (x, y) sont linéairement dépendants. L'idée d'avoir une forme linéaire $(-b \quad a)$ qui dépend de façon linéaire d'un vecteur (a, b) donnera lieu à la notion d'une forme *bilinéaire*, qu'on définira ci-dessous. L'expression $ay - bx$ est bilinéaire en (a, b) et (x, y).

En général on appellera *forme linéaire* sur E un élément de $E^* \overset{\text{déf}}{=} \mathcal{L}(E, K)$, l'espace des applications linéaires sur K à valeur scalaire, espace qu'on appelle aussi l'*espace dual* de E. Sa dimension est donnée par $\dim(E^*) = n = \dim(E)$, car par rapport à une base de E et la base canonique de K, les formes

linéaires sont données par une matrice $1 \times n$, et l'espace de telles matrices est de dimension n. Des exemples typiques de formes linéaires sont les n *fonctions coordonnées* par rapport à une base $\mathcal{B}$, à savoir les formes $(x_1, \ldots, x_n)_{\mathcal{B}} \mapsto x_i$ pour $i = 1, \ldots, n$. La représentation matricielle montre que toute forme linéaire est des façon unique une combinaison linéaire de ces fonctions coordonnées, qui forment donc une base de E^*, appelée la *base duale* de $\mathcal{B}$. Même si $\dim(E^*) = \dim(E)$, et s'il existe donc des isomorphismes entre E et E^*, il n'y a pas une manière *naturelle* (donc sans utiliser une donnée supplémentaire, comme le choix d'une base) d'associer un vecteur de E à une forme linéaire. Par contre on peut associer à une forme linéaire un sous-espace de E, à savoir son noyau ; ce sera toujours un sous-espace de dimension $n-1$, ce qu'on appelle un *hyperplan vectoriel* de E, sauf si la forme est nulle (le noyau de la forme nulle est évidemment tout l'espace E, et donc de dimension n). Chaque hyperplan vectoriel H de E est le noyau d'une forme linéaire (il suffit de choisir une base de H, et de la compléter par un vecteur à une base de E ; la dernière fonction coordonnée pour cette base a pour noyau H), et deux formes linéaires avec le même noyau H sont égales à un facteur scalaire non nul près.

Revenons à la question de l'indépendance linéaire, pour $n = 3$. Si $v_1 = (a, b, c)$ et $v_2 = (p, q, r)$ sont des vecteurs indépendants de K^3, ils engendrent un (hyper)plan vectoriel $P = \mathrm{Vect}(v_1, v_2)$ dans K^3. Il existe donc une forme linéaire non nulle f sur K^3, unique à un scalaire près, telle que $f(v_1) = f(v_2) = 0$, et avec donc $\mathrm{Ker}(f) = P$. Un troisième vecteur $v_3 = (x, y, z)$ formera alors une famille libre avec v_1, v_2 si et seulement si $f(v_3) \neq 0$. Ce que le déterminant 3×3 nous fournira est un choix d'une telle forme linéaire *en fonction de* v_1, v_2, à savoir la forme dont la matrice est $(br - cq \quad -ar + cp \quad aq - bp)$, donc $f : (x, y, z) \mapsto brx - cqx - ary + cpy + aqz - bpz$. On vérifie facilement que $f((x, y, z)) = 0$ quand $(x, y, z) = (a, b, c)$ et quand $(x, y, z) = (p, q, r)$. En plus, cette forme linéaire f dépend de façon linéaire de chacun des vecteurs $v_1 = (a, b, c)$ et $v_2 = (p, q, r)$ quand l'autre est fixé, ce qui mènera à la notion d'une *forme multilinéaire*. Et finalement, on peut vérifier que f est la forme nulle si (et seulement si) v_1 et v_2 sont déjà liés. Ainsi le déterminant de matrice 3×3

$$\begin{vmatrix} a & p & x \\ b & q & y \\ c & r & z \end{vmatrix} \quad \stackrel{\text{déf}}{=} \quad \det \begin{pmatrix} a & p & x \\ b & q & y \\ c & r & z \end{pmatrix} \quad = \quad aqz - ary - bpz + brx + cpy - cqx \tag{24}$$

détermine une forme 3-linéaire (ou trilinéaire) en les vecteurs $v_1, v_2, v_3 \in K^3$ qui forment les trois colonnes de la matrice, et cette forme s'annule précisément quand ces vecteurs sont linéairement dépendants.

4.2. *Formes multilinéaires alternées.*

En vue de ce qu'on vient de discuter, on pourrait définir récursivement les espaces vectoriels $\mathcal{L}_k(E)$ des formes k-linéaires sur E, en prenant $\mathcal{L}_0(E) = K$ comme cas de base, et en définissant $\mathcal{L}_{k+1}(E)$ comme $\mathcal{L}(E, \mathcal{L}_k(E))$ pour $k \in \mathbf{N}$; autrement dit une forme k-linéaire sur E est une constante si $k = 0$, et sinon une application qui produit une forme $k-1$-linéaire sur E, de manière linéaire en fonction d'un vecteur de E. C'est essentiellement comme cela que les formes k-linéaires sont définies, mais l'idée d'une application qui à $v_1 \in E$ associe une application qui à $v_2 \in E$ associe... une application qui à $v_k \in E$ associe un scalaire est un peu difficile pour la discussion et pour la notation. On définira donc une forme k-linéaire comme une seule fonction qui à un k-uplet de vecteurs $[v_1, \ldots, v_k]$ associe directement un scalaire. Mais les propriétés de la notion récursive seront préservées, notamment les fait que $\mathcal{L}_k(E)$ est un K-espace vectoriel, et que sa dimension est n^k où $n = \dim(E)$ (ce qu'on déduit de la règle générale $\dim(\mathcal{L}(E, F)) = \dim(E) \times \dim(F)$, qui résulte de la représentation matricielle des applications linéaires).

4.2.1. Définition. *Soit E un K-espace vectoriel, et $k \in \mathbf{N}$. Une forme k-linéaire sur E est une application $f : E^k \to K$ telle que pour tout indice $1 \leq i \leq k$ et tout choix de $k-1$ vecteurs fixés $v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_k$, l'application $x \mapsto f(v_1, \ldots, v_{i-1}, x, v_{i+1}, \ldots, v_k)$ soit une forme linéaire sur E. Les formes k-linéaires sur E forment un K espace vectoriel noté $\mathcal{L}_k(E)$.*

4.2.2. Proposition. *Si E est de dimension finie n, la dimension de $\mathcal{L}_k(E)$ est n^k.* □

Concrètement, si $\ell_1, \ldots, \ell_n \in E^*$ sont les n fonctions coordonnées pour une base $\mathcal{B}$ de E, une base de $\mathcal{L}_k(E)$ est donnée par les n^k formes k-linéaires $m_{i_1,\ldots,i_k} : (v_1, \ldots, v_k) \mapsto \ell_{i_1}(v_1) \ldots \ell_{i_k}(v_k)$, où le k-uplet $(i_1, \ldots, i_k)$ d'indices parcourt $\{1, \ldots, n\}^k$ (c'est-à-dire chaque i_j parcourt indépendamment $\{1, \ldots, n\}$).

Si $f \in \mathcal{L}_k(E)$, si $\mathcal{B} = [b_1, \dots, b_n]$ est une base de E, et $v_1, \dots, v_k \in E$ sont des vecteurs quelconques, on pourra prendre l'expression $f(v_1, \dots, v_k)$, exprimer chaque vecteur v_i dans la base $\mathcal{B}$, et appliquer successivement la linéarité de f par rapport à chacun de ses k arguments, pour obtenir une de cette expression comme une grande somme

$$f(v_1, \dots, v_k) = \sum_{i_1=1}^{n} \dots \sum_{i_k=1}^{n} c_{i_1,\dots,i_k} f(b_{i_1}, \dots, b_{i_k}),$$

dont les coefficients $c_{i_1,\dots,i_k}$ dépendent seulement des coordonnées des vecteurs v_i dans la base $\mathcal{B}$ (d'une façon sans grande importance, mais si $v_j = (x_{1,j}, \dots, x_{n,j})_{\mathcal{B}}$, alors $c_{i_1,\dots,i_k} = x_{i_1,1} x_{i_2,2} \dots x_{i_k,k}$). Dans cette multiple somme, toutes les applications de la forme k-linéaire f ont tous leurs arguments parmi les vecteurs de la base $\mathcal{B}$, donc on voit que f est entièrement déterminé par ces valeurs particulières.

4.2.3. Proposition. *Si $\mathcal{B} = [b_1, \dots, b_n]$ est une base de E, une forme k-linéaire f est déterminée par les n^k valeurs $f(b_{i_1}, \dots, b_{i_k}) \in K$, qui peuvent être quelconques (un $f \in \mathcal{L}_k(E)$ correspondant existera).*

Preuve. Une combinaison linéaire des n^k formes k-linéaires $m_{i_1,\dots,i_k}$ mentionnées ci-dessus donnera, quand elle est appliquée au k-uplet $(b_{i_1}, \dots, b_{i_k})$, comme valeur le coefficient qu'on avait affecté à $m_{i_1,\dots,i_k}$; ceci montre qu'on peut choisir ces valeurs librement. Par conséquent, la forme donnée $f \in \mathcal{L}_k(E)$ et la combinaison linéaire $\sum_{i_1,\dots,i_k \in \{1,\dots,n\}} f(b_{i_1}, \dots, b_{i_k}) m_{i_1,\dots,i_k} \in \mathcal{L}_k(E)$ donnent les mêmes valeurs quand leurs arguments sont tous parmi les vecteurs de la base $\mathcal{B}$, et d'après l'argument donné ci-dessus cette combinaison linéaire est donc égale à f. (On a démontré que ces $m_{i_1,\dots,i_k}$ forment une base de $\mathcal{L}_k(E)$.) □

Pour l'espace $\mathcal{L}_2(K^2)$ des formes bilinéaires (c'est-à-dire 2-linéaires) sur K^2, la base des $2^2 = 4$ formes m_{i_1,i_2}, écrites en fonction de leurs deux arguments $v_1 = (a, b)$ et $v_2(x, y)$, est donnée par $m_{1,1} = ax$, $m_{1,2} = ay$, $m_{2,1} = bx$, $m_{2,2} = by$. Le déterminant de (23) en est une combinaison linéaire $ay - bx$. Une autre forme bilinéaire importante, dans le cas $K = \mathbf{R}$, est la combinaison linéaire $ax + by$ connue comme le *produit scalaire* qui munit $\mathbf{R}^2$ d'une structure d'*espace euclidien*. (Des espaces vectoriels munis de ce type de structure forment le sujet d'un autre cours.) La propriété spéciale qui distinguera le déterminant des autre formes multilinéaires, est qu'il s'annule dès qu'il existe une relation de dépendance linéaire entre les vecteurs qui forment ses arguments. Cela sera assuré dans le cas d'une forme multilinéaire *alternée*.

4.2.4. Définition. *Une forme k-linéaire $f \in \mathcal{L}_k(E)$ est dite alternée si on a $f(v_1, \dots, v_k) = 0$ dès qu'il y a des vecteurs égaux parmi $v_1, \dots, v_k$, c'est-à-dire qu'il existe des indices $i \neq j$ tels que que $v_i = v_j$.*

La condition pour une forme k-linéaire f d'être alternée réduit considérablement ses possibilités. On voit tout de suite, pour une base $\mathcal{B} = [b_1, \dots, b_n]$ de E, que parmi les valeurs $f(b_{i_1}, \dots, b_{i_k})$ qui déterminent f, celles pour lesquelles deux indices au moins sont égaux seront toutes nulles. Cela nous apprend déjà que pour $k > n$ la seule forme k-linéaire alternée est la forme nulle. Mais on peut dire plus.

4.2.5. Proposition. *Soit f une forme k-linéaire alternée sur E, et $i, j \in \{1, \dots, k\}$ des indices distincts.*
(1) Pour tout $\lambda \in K$ on a $f(v_1, \dots, v_{i-1}, v_i + \lambda v_j, v_{i+1}, \dots, v_k) = f(v_1, \dots, v_k)$.
(2) Si la famille de vecteurs $[v_1, \dots, v_k]$ est liée, on a $f(v_1, \dots, v_k) = 0$.
(3) Si on suppose $i < j$, on a $f(v_1, \dots, v_{i-1}, v_j, v_{i+1}, \dots, v_{j-1}, v_i, v_{i+1}, \dots, v_k) = -f(v_1, \dots, v_k)$.

Preuve. Pour (1) on utilise la linéarité de f par rapport à v_i : $f(v_1, \dots, v_{i-1}, v_i + \lambda v_j, v_{i+1}, \dots, v_k) = f(v_1, \dots, v_k) + \lambda f(v_1, \dots, v_{i-1}, v_j, v_{i+1}, \dots, v_k)$, où on observe que le second terme est nul car f est alternée (le vecteur v_j est présent deux fois parmi les arguments). La partie (2) en découle, car par hypothèse l'un des vecteurs v_i est égal à une combinaison linéaire des *autres* vecteurs v_j, et en appliquant (1) successivement pour les différents v_j pour soustraire leur contribution dans cette combinaison linéaire de v_i, on réduit ce i-ème argument de f au vecteur nul, et ceci étant fait la valeur de f sera certainement nulle. Pour (3) les arguments autres que v_i et v_j ne changent pas, donc on simplifiera l'argument en considérant $g : E^2 \to K$ avec $g(x, y) = f(v_1, \dots, v_{i-1}, x, v_{i+1}, \dots, v_{j-1}, y, v_{i+1}, \dots, v_k)$, qui est une forme bilinéaire alternée. Alors on a $0 = g(x + y, x + y) = g(x, x) + g(x, y) + g(y, x) + g(y, y) = g(x, y) + g(y, x)$ pour tout $x, y \in E$, d'où $g(y, x) = -g(x, y)$ comme l'affirme l'énoncé (pour $x = v_i$ et $y = v_j$). □

4.2.6. Corollaire. *Si $\mathcal{B} = [b_1, \dots, b_n]$ est une base de E, une forme k-linéaire alternée f est déterminée par les $\binom{n}{k}$ valeurs $f(b_{i_1}, \dots, b_{i_k}) \in K$ avec indices strictement croissants : $1 \leq i_1 < i_2 < \cdots < i_k \leq n$.*

Preuve. On a vu qu'une forme k-linéaire est déterminée par de telles valeurs où $(i_1, \dots, i_k)$ parcourt $\{1, \dots, n\}^k$, et que pour une forme alternée ces valeurs sont nulles dès qu'il y a deux indices égaux. Il reste les valeurs prises par f pour les suites d'indices tous distincts (les *arrangements* de k indices choisis parmi $\{1, \dots, n\}$). Les valeurs obtenues pour l'ensemble de telles suites qui concernent toutes un même sous-ensemble (*combinaison*) S d'indices, sont reliées entre elles par une suite de relations de la proposition 4.2.5(3) qui intervertissent une paire d'indices. Il suffit donc de connaître l'une de ces valeurs, pour laquelle on peut prendre $f(b_{i_1}, \dots, b_{i_k})$, si $S = \{i_1, \dots, i_k\}$ avec $i_1 < \cdots < i_k$. □

Le corollaire implique que l'espace des formes n-linéaires alternées sur E est de dimension au plus $\binom{n}{k}$, et en particulier que l'espace des formes n-linéaires alternées est de dimension 1 au plus: une telle forme f est déterminée par la *seule valeur* $f(b_1, \dots, b_n)$. Pour être sûr que cette dimension est atteinte, et donc qu'il existe une forme n-linéaire alternée non nulle sur E, il nous faudra montrer qu'on peut choisir des valeurs de f sur les permutations des vecteurs $b_1, \dots, b_k$ en accord avec la proposition 4.2.5(3).

On appellera *permutation de n* tout arrangement $\sigma = (\sigma_1, \dots, \sigma_n)$ tel que $\{\sigma_1, \dots, \sigma_n\} = \{1, \dots, n\}$ (on trouve donc tous les nombres $1, \dots, n$ chacun une fois dans l'arrangement, mais pas dans un ordre particulier), et on notera $\mathbf{S}_n$ l'ensemble des permutations de n (elles sont $n! = n \times (n-1) \times \cdots \times 2 \times 1$ en nombre). L'unique permutation croissante $(1, 2, \dots, n)$ est appelée la permutation identique de n. Une *transposition simple* intervertit deux composentes voisines σ_i, σ_{i+1} de $\sigma \in \mathbf{S}_n$, et résulte donc en $\sigma' = (\sigma'_1, \dots, \sigma'_n) \in \mathbf{S}_n$ vérifiant $\sigma'_i = \sigma_{i+1}$, $\sigma'_{i+1} = \sigma_i$, et $\sigma'_j = \sigma_j$ pour tout $j \in \{1, \dots, n\} \setminus \{i, i+1\}$.

4.2.7. Proposition/Définition. *Pour tout $n \in \mathbf{N}$ il existe une application $f_n : \mathbf{S}_n \to \{1, -1\}$ telle que $f_n(\sigma) = -f_n(\sigma')$ pour tout $\sigma \in \mathbf{S}_n$ et toute permutation σ' obtenue de σ par une transposition simple, et $f_n(\sigma) = 1$ quand σ est la permutation identique de n. On appelle $f_n(\sigma)$ la signature de σ, notée* $\mathrm{sg}(\sigma)$.

Preuve. On définit f_n par récurrence sur n : pour $n \leq 1$ c'est la fonction constante à valeur 1 (sur la permutation identique, qui est la seule dans ces cas) ; pour $n > 1$ et $\sigma \in \mathbf{S}_n$ soit i l'indice tel que $\sigma_i = n$, alors on pose $f_n(\sigma) = (-1)^{n-i} f_{n-1}(\tau)$ où $\tau = (\sigma_1, \dots, \sigma_{i-1}, \sigma_{i+1}, \dots, \sigma_n) \in \mathbf{S}_{n-1}$. Pour σ la permutation identique de n on aura $i = n$ et τ est la permutation identique de $n-1$, donc $f_n(\sigma) = 1$ par l'hypothèse de récurrence. Pour vérifier $f_n(\sigma') = -f_n(\sigma)$ quand σ' est obtenue de σ par une transposition simple, on distingue deux cas : $\sigma'_i = n$ (la transposition n'a pas déplacé le terme $\sigma_i = n$), et le cas contraire ou une transposition simple a déplacé $\sigma_i = n$ vers $\sigma'_j = n$ pour un indice $j \in \{i-1, i+1\}$. Dans le premier cas on a $f_n(\sigma') = (-1)^{n-i} f_{n-1}(\tau')$ pour un τ' obtenu de τ par une transposition simple, et par l'hypothèse de récurrence $f_n(\sigma') = -(-1)^{n-i} f_{n-1}(\tau) = -f_n(\sigma)$. Dans le second cas la permutation $\tau \in \mathbf{S}_{n-1}$ qui reste après suppresion de $n = \sigma_i$ dans σ est la même que celle qui reste après suppresion de $n = \sigma'_j$ dans σ', donc $f_n(\sigma') = (-1)^{n-j} f_{n-1}(\tau) = -f_n(\sigma)$, car $(-1)^{n-j} = -(-1)^i$. □

On remarque dans la définition de f_n que $n - i$ est le nombre de termes de σ qui viennent *après* le terme $\sigma_i = n$ (et qui sont forcément plus petit que celui-ci). On voit alors facilement que $\mathrm{sg}(\sigma) = (-1)^N$, où N est le nombre d'*inversions* de σ, c'est-à-dire de couples d'indices (i, j) avec $i < j$ et $\sigma_i > \sigma_j$.

4.2.8. Corollaire. *Pour toute base $\mathcal{B} = [b_1, \dots, b_n]$ de E il existe une forme n linéaire alternée unique f sur E telle que $f(b_1, \dots, b_n) = 1$. Si $[\ell_1, \dots, \ell_n]$ est la base de E^* duale de $\mathcal{B}$, on a pour $v_1, \dots, v_n \in E$:*

$$f(v_1, \dots, v_n) = \sum_{\sigma \in \mathbf{S}_n} \mathrm{sg}(\sigma) \ell_{\sigma_1}(v_1) \dots \ell_{\sigma_n}(v_n).$$

La formule définit l'unique forme n-linéaire telle que $f(b_{\sigma_1}, \dots, b_{\sigma_n}) = \mathrm{sg}(\sigma)$ pour tout $\sigma \in \mathbf{S}_n$, et $f(b_{i_1}, \dots, b_{i_n}) = 0$ pour tout n-uplet d'indices $(i_1, \dots, i_n) \in \{1, \dots, n\}^n$ qui n'est pas permutation de n, c'est-à-dire avec au moins un indice qui apparaît deux fois. D'après la proposition 4.2.7, cette forme est en fait alternée (pour voir le changement de signe pour une transposition de (i, j) avec $j - i > 1$, on la réalise par une suite d'un nombre impair $2(j-i)-1$ de transpositions simples, à trouver comme exercice).

Par une expression similaire (mais encore plus compliquée à écrire) on montre que les $\binom{n}{k}$ valeurs dont parle le corollaire 4.2.6 peuvent toutes être choisies indépendamment, et la dimension de l'espace des formes k-linéaires alternées sur E est donc donnée par ce coefficient binomial $\binom{n}{k}$ (qui vaut 1 si $k = n$).

4.3. *Définitions de déterminant.*

Jusqu'ici on a travaillé dans un K-espace vectoriel E de dimension n pour motiver l'existence d'une forme n-linéaire alternée. Mais pour définir le déterminant de façon unique, le mieux est de prendre comme point de départ une matrice : d'une part on ne saura pas faire un choix d'une forme n-linéaire alternée particulière (elle forment un espace de dimension 1) sans utiliser une base de E, et d'autre part en considérant une matrice on n'est pas obligé de supposer qu'elle représente un n-uplet de vecteurs, et notamment pas que ses coefficients soient pris dans un corps. En fait on donnera une définition du déterminant dans trois cadres distincts : celui des matrices carrées à coefficients dans un anneau commutatif, celui des n-uplets de vecteurs de E, et celui des endomorphismes de E.

Déterminant d'une matrice à coefficients dans un anneau commutatif.

La définition du déterminant d'une matrice carrée ne fait intervenir que des additions, soustractions et des multiplications; il est donc possible de la formuler pour les matrices à coefficients dans un anneau, pas nécessairement dans un corps. Et ceci est très utile, car notre raison principale de considérer le déterminant et pouvoir définir le polynôme caractéristique, qui est le déterminant d'une matrice à coefficients dans $K[X]$. Cependant, pour avoir les propriétés les plus basiques des déterminants, il faut supposer que l'anneau soit *commutatif* (car la multilinéarité nécessite des scalaires qui commutent).

4.3.1. Définition. *Pour une matrice carrée $A = (A_{i,j})_{i,j=1}^n$ à coefficients dans un anneau commutatif R, son déterminant est :*

$$\det(A) = \sum_{\sigma\in\mathbf{S}_n} \left(\operatorname{sg}(\sigma) \prod_{j=1}^{n} A_{\sigma_j,j} \right) \in R. \tag{25}$$

Un nombre d'identités fondamentales, valables pour tout anneau commutatif R, suivent directement de cette définition. Nous formulons seulement les cas de base ; d'autres identités peuvent être déduites en les combinant (notamment (2) permet de conclure pour les lignes tout ce qui est dit pour les colonnes).

4.3.2. Théorème. *Le déterminant d'une matrice $n \times n$ sur R possède les propriétés suivantes.*

(1) Si $f : R \to S$ est un homomorphisme d'anneaux commutatifs, l'application de f au déterminant a le même effet que d'appliquer f à tous les coefficients: $f\big(\det\big((A_{i,j})_{i,j=1}^n\big)\big) = \det\big((f(A_{i,j}))_{i,j=1}^n\big)$.

(2) La transposition de la matrice ne change pas le déterminant: $\det(A) = \det({}^tA)$.

(3) Le déterminant de la matrice identité est 1.

(4) Comme fonction d'une seule colonne (donc en fixant les autres) le déterminant définit une application R-linéaire: si j et les colonnes de A autres que la colonne j sont fixés, et si $f : R^n \to R$ est défini par $f(C) = \det(A \leftarrow_j C)$ où $A \leftarrow_j C$ désigne la matrice obtenue de A en remplaçant la colonne j par C, on a $f(\sum r_i C_i) = \sum_i r_i f(C_i)$ pour toute combinaison R-linéaire $\sum r_i C_i$ avec $r_i \in R$ et $C_i \in R^n$.

(5) Le déterminant est alterné en les colonnes: si deux colonnes de A sont identiques, alors $\det(A) = 0$.

(6) Si A possède un bloc de zéros en bas à gauche qui touche à la diagonale principale, le déterminant se factorise comme le produit des deux déterminants des blocs carrés sur la diagonale principale:

$$\det\begin{pmatrix} A & C \\ \mathbf{0} & B \end{pmatrix} = \det(A)\det(B) \qquad \text{si } A \text{ et } B \text{ sont des blocs carrés.}$$

(7) Si A est triangulaire, alors le déterminant est égal au produit des coefficients diagonaux:

$$\det\begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} & \dots & a_{1,n} \\ 0 & a_{2,2} & a_{2,3} & \dots & a_{2,n} \\ 0 & 0 & a_{3,3} & \dots & a_{3,n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & a_{n,n} \end{pmatrix} = a_{1,1}\, a_{2,2}\, a_{3,3} \dots a_{n,n}$$

(8) Le déterminant est multiplicatif par rapport au produit matriciel de matrices carrées:

$$\det(A \cdot B) = \det(A)\det(B).$$

(9) Si $n > 0$, le déterminant peut être "développé" en termes de déterminants de matrices $(n-1)\times(n-1)$: pour tout $j \in \{1, \dots, n\}$ fixé, on a $\det(A) = \sum_{i=1}^n (-1)^{i-j} A_{i,j} \det(A_{\widehat{i,j}})$, où $A_{\widehat{i,j}}$ est la matrice $(n-1) \times (n-1)$ obtenue à partir de A en supprimant la ligne et la colonne du coefficient $A_{i,j}$.

Toutes ces identités sont obtenues à partir de la définition par des manipulations algébriques plus ou moins compliquées. On ne donnera ici seulement une indication du type de manipulation et du raisonnement qui mènent à ces identités, parfois accompagnée d'une explication de leur signification.

L'identité (1) est une conséquence directe de la notion de homomorphisme, en vue du fait que le déterminant est donné par une expression explicite (aussi grande qu'elle soit). Elle est d'un importance fondamentale, ne serait-ce que parce qu'elle dit que la valeur du déterminant ne change pas quand on interprète des coefficients dans une anneau commutatif plus grand (avec f l'inclusion des anneaux en question). Une utilisation moins évidente est avec f un morphisme non injectif: cela permet par exemple de calculer l'évaluation en $a \in K$ du déterminant d'une matrice de polynômes en évaluant tous les polynômes de la matrice en a, et de calculer le déterminant de la matrice de scalaires ainsi obtenue.

L'identité (2) est une conséquence du fait que chaque terme $\mathrm{sg}(\sigma)\prod_{i=1}^{n} A_{i,\sigma_i}$ de l'expression pour le déterminant de ${}^tA$ est en correspondance avec un terme du déterminant de A pour une autre permutation, appelée la permutation inverse σ^{-1} de σ, qui vérifie $\mathrm{sg}(\sigma^{-1}) = \mathrm{sg}(\sigma)$. Si on a appelle matrice de la permutation σ la matrice ayant pour colonne j le vecteur $\mathbf{e}_{\sigma_j}$ de la base canonique de K^n (pour $j = 1, \ldots, n$), la permutation σ^{-1} est par définition celle dont la matrice est la transposée de celle de σ. Il s'agit clairement d'une correspondance bijective (et même involutive : elle est sa propre réciproque). Pour voir que $\mathrm{sg}(\sigma^{-1}) = \mathrm{sg}(\sigma)$, on peut soit argumenter que les inversions de σ sont en bijection avec celles de σ^{-1}, soit argumenter que le fait d'effectuer une transposition simple sur σ résulte toujours en une transposition effectuée sur σ^{-1}, qui correspond à un nombre *impair* de transpositions simples.

L'identité (3) est évidente. Avec les identités (4) et (5) elle permet de calculer la valeur d'un déterminant quelconque sans faire référence à la définition (25).

L'identité (4) affirme le caractère n-linéaire en les colonnes du déterminant (et grâce à (2) il est aussi n-linéaire en les lignes), mais on a évité ce terme qui n'a été introduit que dans le cadre des espaces vectoriels. Elle est fondamentale pour le calcul pratique des déterminants (l'application directe de la définition étant en général trop compliquée), permettant le développement par une colonne, ou (avec l'identité (5)) la simplification par opérations sur les colonnes. Cette identité est une conséquence de la forme de l'expression définissant le déterminant : chaque terme $\mathrm{sg}(\sigma)\prod_{j=1}^{n} A_{\sigma_j,j}$ de (25) est une fonction R-linéaire de la colonne j de A (le reste de A étant fixe), car c'est un multiple de la coordonnée numéro σ_j de cette colonne, dont la valeur pour la matrice A est $A_{\sigma_j,j}$. La somme pour $\sigma \in \mathbf{S}_n$ de ces termes est donc aussi une fonction R-linéaire de la colonne j de A.

L'identité (5) affirme le caractère alterné en les colonnes du déterminant; c'est la raison d'être de la sommation alternée sur toutes les permutations dans la définition du déterminant. Si les colonnes i et j de A sont identiques, on considère les paires formées d'une permutation σ et de la permutation τ obtenu en effectuant la transposition (i, j) sur σ (comme la même transposition appliqué à τ redonne σ, il s'agit d'une partition de l'ensemble des permutation en paires). Les produits dans les termes associés à σ et τ dans (25) diffèrent seulement aux indices i et j, mais on a $A_{\sigma_i,i}A_{\sigma_j,j} = A_{\tau_j,i}A_{\tau_i,j} = A_{\tau_j,j}A_{\tau_i,i}$ (la dernière équation car $A_{k,i} = A_{k,j}$ pour tout k), donc ce produit de deux facteurs est le même pour les deux termes, et par conséquence le produit entier aussi. Mais $\mathrm{sg}(\sigma) = -\mathrm{sg}(\tau)$ les deux termes s'annulent ; en prenant la somme sur toutes les paires on obtient $\det(A) = 0$.

Quand l'identité (6) s'applique, elle permet une simplification considérable du calcul du déterminant, car le nombre combiné de termes dans les expressions pour $\det(A)$ et $\det(B)$ est en général nettement plus petit que le nombre de termes dans le déterminant global. Cette simplification est une conséquence du fait que, pour qu'une permutation σ donne une contribution non nulle au déterminant, il faut qu'elle permute les indices $(k+1, \ldots, n)$ entre eux (où k est la taille de A), car si $\sigma_j > k$ pour un indice $j \leq k$ on aura $A_{\sigma_j,j} = 0$. Par conséquent les indices $(1, \ldots, k)$ doivent aussi être permutés entre eux, et la permutation σ consiste effectivement en deux permutations indépendantes $\sigma^{(1)} \in \mathbf{S}_k$, $\sigma^{(2)} \in \mathbf{S}_{n-k}$ de ces deux parties. Aussi une inversion d'un tel σ ne peut concerner qu'une des deux parties, donc $\mathrm{sg}(\sigma)$ est le produit $\mathrm{sg}(\sigma^{(1)})\,\mathrm{sg}(\sigma^{(2)})$ des signatures des deux permutations. Du coup la somme dans (25) se factorise comme le produit de deux sommes, correspondants respectivement à $\det(A)$ et $\det(B)$.

L'identité (7) découle de (6) par récurrence sur n, ou se démontre directement car toute permutation non identique aura $j < \sigma_j$ pour au moins un indice j, ce qui annule sa contribution à cause de $A_{\sigma_j,j} = 0$.

L'identité (8) est une propriété fondamentale, qui dit que le déterminant est multiplicatif (ou de façon équivalente, qu'il est un morphisme du monoïde multiplicatif de $\mathcal{M}_n(R)$ vers celui de R). Elle est un peu plus difficile à démontrer que les autres identités, car l'expansion de (25) pour la matrice $A \cdot B$ est extrêmement compliquée. On verra que pour le cas où R est un corps, elle découle facilement de ce qu'on a dit sur les formes n-linéaires alternées. Pour la montrer pour le cas général d'un anneau commutatif, on pourra utiliser que chaque colonne du produit $A \cdot B$ est le produit de A et de la colonne correspondante de B, et pour A fixé elle est donc une fonction R-linéaire de cette colonne ; d'après (4) les expressions $\det(A \cdot B)$ et $\det(A)\det(B)$ sont donc des fonctions R-multilinéaires des n colonnes de B. Par conséquent, l'égalité des ces expressions sera assurée si elles prennent les mêmes valeurs dans les cas particuliers où chaque colonne de B est un des générateurs canoniques $\mathbf{e}_i$ de R^n (elle contient un seul coefficient non nul, qui est 1). Ensuite on peut utiliser l'identité (5), qui entraîne que $\det(A \cdot B)$ et $\det(A)\det(B)$ sont des fonctions alternées des colonnes de B et qu'elles s'annullent donc si deux colonnes de B sont égales et (comme dans la proposition 4.2.5(3)) changent de signe quand on intervertit deux colonnes, pour réduire le cas restant au cas $B = I_n$; pour ce cas $\det(A \cdot B) = \det(A)\det(B)$ est évident.

Finalement l'identité (9) est une conséquence de (4), qui permet de développer le déterminant comme combinaison R-linéaire de déterminants $n \times n$, à savoir $\det(A) = \sum_{i=1}^{n} A_{i,j} \det(A \leftarrow_j \mathbf{e}_i)$, où $\mathbf{e}_i$ est la colonne donnant le i-ème générateur canonique de R^n. Le fait que $\det(A \leftarrow_j \mathbf{e}_i) = (-1)^{i-j}\det(A_{\widehat{i,j}})$ peut être montré en utilisant le caractère alterné du déterminant par lignes et par colonnes, ou directement de la définition, ainsi. Seules les permutations σ avec $\sigma_j = i$ donnent une contribution à $\det(A \leftarrow_j \mathbf{e}_i)$, et elles sont en bijection avec les permutations de $n-1$, par l'opération de supprimer le terme $\sigma_j = i$, et de soustraire 1 de la valeur des termes restants qui sont $> i$ (rendant le résultat une permutation). On a $\#\{\, j' > j \mid \sigma_{j'} < i \,\} - \#\{\, j' < j \mid \sigma_{j'} > i \,\} = i - j$ par un simple argument de comptage, donc la parité du nombre d'inversions que cette opération enlève de la permutation est celle de $i - j$.

Déterminant dans une base d'un système de n vecteurs.

4.3.3. Définition. *Si E est un K-espace vectoriel de dimension n et $\mathcal{E} = [e_1, \ldots, e_n]$ une base de E. Le déterminant $\det_B(v_1, \ldots, v_n)$ par rapport à $\mathcal{E}$ d'un n-uplet de vecteurs $v_1, \ldots, v_n \in E$ est le déterminant de la matrice dont les colonnes donnent les coordonnées des vecteurs v_j dans la base $\mathcal{E}$:*

$$\det_{\mathcal{E}}(v_1, \ldots, v_n) = \det(M) \quad \text{où } M \in \mathrm{Mat}_n(K) \text{ vérifie } v_j = (M_{1,j}, \ldots, M_{n,j})_{\mathcal{E}} \text{ pour } 1 \leq j \leq n. \tag{26}$$

Une autre façon de décrire M est que, avec $[\ell_1, \ldots, \ell_n]$ la base duale de $\mathcal{E}$ (les fonctions coordonnées pour $\mathcal{E}$), on a $M = \big(\ell_i(v_j)\big)_{i,j=1,\ldots,n}$. Alors l'expression (25) appliquée à $\det(M)$ donne la formule du corollaire 4.2.8. Les propriétés fondamentales de ce déterminant de n vecteurs sont les suivantes.

4.3.4. Théorème.
(1) $\det_{\mathcal{E}} : E \times \cdots \times E \to K$ *est une forme n-linéaire alternée, et* $\det_{\mathcal{E}}(e_1, \ldots, e_n) = 1$.
(2) Pour tout forme n-linéaire alternée $f : E \times \cdots \times E \to K$ il existe $\lambda \in K$ tel que $f = \lambda \det_{\mathcal{E}}$.
(3) Si $\mathcal{B} = [b_1, \ldots, b_n]$ est une autre base de E, le facteur λ du point précédent pour $f = \det_{\mathcal{B}}$ est donné par $\lambda = \det_{\mathcal{E}}(b_1, \ldots, b_n)^{-1}$ (et en particulier $\det_{\mathcal{E}}(b_1, \ldots, b_n) \neq 0$).
(4) Dans ce cas on a $\det_{\mathcal{E}}(v_1, \ldots, v_n) = \det_{\mathcal{E}}(b_1, \ldots, b_n)\det_{\mathcal{B}}(v_1, \ldots, v_n)$ pour tout $v_1, \ldots, v_n \in E$.
(5) Pour $v_1, \ldots, v_n \in E$ on a $\det_{\mathcal{E}}(v_1, \ldots, v_n) \neq 0$ si et seulement si $[v_1, \ldots, v_n]$ forme une base de E.

Preuve. Le point (1) est évident, il découle aussi bien des propriétés (3)–(5) du théorème 4.3.2 que du fait qu'on vient de retrouver la formule du corollaire 4.2.8. Le point (2) exprime le fait que les formes n-linéaires alternées forment un espace vectoriel de dimension 1, et que $\det_{\mathcal{E}}$ n'est pas la forme nulle, ce qui est clair par ce qui précède. Pour (3) on sait que $\det_{\mathcal{B}}(b_1, \ldots, b_n) = 1$, ce qui donne $\lambda \det_{\mathcal{E}}(b_1, \ldots, b_n) = 1$. Le point (4) n'est qu'une réécriture de $\det_{\mathcal{B}} = \lambda \det_{\mathcal{E}}$ comme $\det_{\mathcal{E}} = \lambda^{-1}\det_{\mathcal{B}}$. La partie "si" du point (5) est mentionnée dans (3) ; réciproquement si $[v_1, \ldots, v_n]$ n'est pas une base, la famille est liée, et toute forme n-linéaire alternée, en particulier $\det_{\mathcal{E}}$, s'annule en $[v_1, \ldots, v_n]$ d'après la proposition 4.2.5(2). □

On remarque que si A est la matrice dont les colonnes expriment les vecteurs $v_1, \ldots, v_n$ dans la base $\mathcal{B}$, et P est la matrice de passage de $\mathcal{E}$ vers $\mathcal{B}$, alors $P \cdot A$ est la matrice dont les colonnes expriment les vecteurs $v_1, \ldots, v_n$ dans la base $\mathcal{E}$. Le point (4) se traduit alors par $\det(P \cdot A) = \det(P)\det(A)$.

Déterminant d'un endomorphisme.

Pour un endomorphisme de E, on pourrait définir son déterminant comme le déterminant de sa matrice par rapport à une base $\mathcal{B}$ de E. Mais une telle définition suggère une dépendance de la base $\mathcal{B}$, pendant que, contrairement au déterminant $\det_{\mathcal{B}}$ de n vecteurs, le déterminant d'un endomorphisme est en fait indépendant de la base. Pour cette raison on préfère de donner une définition qui rend cette indépendance manifeste, en évitant toute mention d'une base.

4.3.5. Proposition/Définition. *Pour tout endomorphisme ϕ d'un K-espace vectoriel E de dimension n, il existe un $\lambda \in K$ unique tel que, pour toute forme n-linéaire alternée f sur E et tout $v_1, \ldots, v_n \in E$ on ait $f(\phi(v_1), \ldots, \phi(v_n)) = \lambda f(v_1, \ldots, v_n)$. Ce scalaire λ est par définition le déterminant* $\det(\phi)$ *de ϕ.*

Preuve. Il découle de la linéarité de ϕ que si $f : E \times \cdots \times E \to K$ est n-linéaire alternée, alors l'application $(v_1, \ldots, v_n) \mapsto f(\phi(v_1), \ldots, \phi(v_n))$ l'est aussi. Pour $f = 0$ la condition donnée est vérifiée indépendamment de λ, donc on supposera $f \neq 0$ (ce qui est possible, par exemple en prenant $f = \det_{\mathcal{B}}$ pour une base quelconque $\mathcal{B}$ de E). Alors l'existence d'un unique scalaire λ qui vérifie la condition pour f est une conséquence du fait que les formes n-linéaires alternées forment un espace vectoriel de dimension 1. Or ce λ vérifie aussi la condition pour tout multiple scalaire μf de f, et encore par l'argument de dimension 1, cela montre que λ vérifie la condition pour toute forme n-linéaire alternée sur E. □

Les propriétés principales de cette notion de déterminant découlent facilement de la définition. Nous les résumons pour référence.

4.3.6. Théorème. *Le déterminant d'un endomorphisme ϕ d'un espace vectoriel E vérifie les propriétés :*
(1) $f(\phi(v_1), \ldots, \phi(v_n)) = \det(\phi) f(v_1, \ldots, v_n)$ *pour toute forme n-linéaire alternée f et $v_1, \ldots, v_n \in E$.*
(2) Si $\mathcal{B} = [b_1, \ldots, b_n]$ *est une base de E, alors* $\det_{\mathcal{B}}(\phi(v_1), \ldots, \phi(v_n)) = \det(\phi) \det_{\mathcal{B}}(v_1, \ldots, v_n)$.
(3) Pour toute base $\mathcal{B} = [b_1, \ldots, b_n]$ *de E on a* $\det(\phi) = \det_{\mathcal{B}}(\phi(b_1), \ldots, \phi(b_n)) = \det(\mathrm{Mat}_{\mathcal{B}}(\phi))$.
(4) Si ψ est un autre endomorphisme de E, on a $\det(\phi \circ \psi) = \det(\phi)\det(\psi)$.
(5) On a $\det(\phi) \neq 0$ *si et seulement si ϕ est un isomorphisme.*

Preuve. Le point (1) est juste la définition, (2) en est le cas particulier pour $f = \det_{\mathcal{B}}$, dont (3) découle en prenant $[v_\text{ı}, \ldots, v_n] = \mathcal{B}$, remarquant pour la dernière égalité reflète directement les définitions de $\det_{\mathcal{B}}$ et de $\mathrm{Mat}_{\mathcal{B}}(\phi)$. Pour (4) il suffit de remplacer chaque v_i par $\psi(v_i)$ dans (1). Pour (5) on sait que ϕ est un isomorphisme si et seulement si l'image par ϕ d'une base de E est de nouveau une base, et le point est donc une conséquence de théorème 4.3.4(5), compte tenu de (3) du théorème actuel. □

On voit en particulier que dans le point (3) que effectivement, le déterminant d'un endomorphisme peut être calculé comme le déterminant de sa matrice par rapport à une base quelconque. Et le point (4) nous donne la multiplicativité du déterminant, comme énoncée dans le théorème 4.3.2(8), pour le cas général des matrices carrées à coefficients dans K (pendant que la règle $\det(P \cdot A) = \det(P)\det(A)$ vue ci-dessus se limitait encore au cas où P est une matrice de passage, donc inversible).

4.4. *Déterminants et matrices inverses: la règle de Cramer.*

Fixons l'indice j d'une colonne, pour les matrices $n \times n$ à coefficients dans un anneau commutatif R (on peut penser au cas particulier d'un corps, mais le cas général nous sera utile). Le théorème 4.3.2(4) dit que pour une telle matrice A donnée, l'application $f : R^n \to R$ donnée par $f(C) = \det(A \leftarrow_j C)$ est R-linéaire, donc elle est donnée par une matrice $1 \times n$ à coefficients dans R qu'on appellera $M^{(j)}$. En fait (9) du théorème en donne les coefficients : le coefficient $M_i^{(j)}$ est $(-1)^{i-j}\det(A_{\widehat{i,j}})$ pour $i = 1, \ldots, n$. On a évidemment $f(C_j(A)) = \det(A)$ si $C_j(A)$ est la colonne j de A, et par le caractère alterné du déterminant f s'annule pour les autres colonnes de A, c'est-à-dire $f(C_k(A)) = 0$ si $k \neq j$. Sous forme matricielle, cela veut dire que $M^{(j)} \cdot A = \det(A)\,{}^t\mathbf{e}_j$, où comme avant $\mathbf{e}_j \in R^n$ est le j-ème élément de la base canonique, et ${}^t\mathbf{e}_j$ désigne qu'il est vu ici comme une matrice "ligne" $1 \times n$ (plutôt que "colonne" $n \times 1$). Il est clair qu'on peut également décrire ${}^t\mathbf{e}_j$ comme la ligne j de la matrice identité I_n.

L'identité $M^{(j)} \cdot A = \det(A)\,{}^t\mathbf{e}_j$ a une application au systèmes d'équations en n inconnues ayant A comme matrice de coefficients. Ces systèmes s'écrivent sous forme matriciel, si on appelle $x_1, \ldots, x_n$ les

inconnues, comme

$$\begin{pmatrix} A_{1,1} & A_{1,2} & \dots & A_{1,n} \\ A_{2,1} & A_{2,2} & \dots & A_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n,1} & A_{n,2} & \dots & A_{n,n} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}. \tag{27}$$

On utilisera les abréviations $\mathbf{x}$ et $\mathbf{b}$ pour les deux matrices "colonne" dans cette équation, de sorte qu'elle s'écrive $A \cdot \mathbf{x} = \mathbf{b}$. Si on suppose qu'on a une solution $\mathbf{x}$ de ce système, et on multiplie l'équation à gauche par $M^{(j)}$ on trouve pour le premier membre $M^{(j)} \cdot A \cdot \mathbf{x} = \det(A)\, {}^t\mathbf{e}_j \cdot \mathbf{x} = \det(A)x_j$, et pour le second membre $M^{(j)} \cdot \mathbf{b} = f(\mathbf{b}) = \det(A \leftarrow_j \mathbf{b}) \in R$. Ce qu'on a fait est de former une combinaison R-linéaire des équations dans laquelle toutes les inconnues autre que x_j ont disparu. Ce constat, même si ce n'est pas une méthode très efficace pour résoudre le système, est un important outil théorique.

4.4.1. Proposition. *Pour que $\mathbf{x} \in R^n$ soit solution de $A \cdot \mathbf{x} = \mathbf{b}$ avec $A \in \mathrm{Mat}_n(R)$ et $\mathbf{b} \in R^n$, il est nécessaire que $\det(A)x_j = \det(A \leftarrow_j \mathbf{b})$ pour $j = 1, \dots, n$, où $\leftarrow_j$ indique remplacement de la colonne j.* □

Rassemblons ensuite les n matrices $M^{(j)}$ pour $j = 1, \dots, n$, qui sont de taille $1 \times n$, verticalement en une matrice $M \in \mathrm{Mat}_n(R)$:

$$M = \big((-1)^{i-j} \det(A_{\widehat{i,j}})\big)_{j,i=1,\dots,n}. \tag{28}$$

Attention, par la force des choses, dans cette formule c'est j qui est l'indice des lignes $M^{(j)}$ de la matrice M, et i qui est l'indice des colonnes, ce qui est inhabituel. En prenant le produit matriciel avec A, les n résultats $\det(A)\, {}^t\mathbf{e}_j$ se rassemblent également en une matrice $n \times n$, qui est le multiple scalaire par $\det(A)$ de la matrice identité I_n :

$$M \cdot A = \det(A) I_n. \tag{29}$$

On voit que la matrice M possède une propriété bien particulière par rapport à la matrice A, pendant que ses coefficients sont (à signe près) des déterminants de matrices extraites de A (quel type de déterminant est connu plus généralement sous le nom *mineur* de A). Pour cette raison on appelle M la *comatrice* de A, ou plutôt c'est la transposée de M qui est appelée ainsi en terminologie française ; à tort, car ${}^tM$ ne bénéficie d'aucune propriété particulière par rapport à A qui ne passe pas par une transposition (par contre, dans la terminologie anglaise c'est bien M elle-même qui est appelée *adjugate matrix* pour A).

4.4.2. Définition/Proposition. *Pour une matrice carrée A de taille $n \times n$ à coefficients dans un anneau commutatif R, on appelle comatrice-transposée de A la matrice (du même type)*

$$\mathrm{comatr}(A) = \big((-1)^{j-i} \det(A_{\widehat{j,i}})\big)_{i,j=1,\dots,n}, \tag{30}$$

où la matrice $A_{\widehat{j,i}}$ est celle obtenue à partir de A en supprimant la ligne j et la colonne i. Elle vérifie les égalités

$$\mathrm{comatr}(A) \cdot A = \det(A) I_n = A \cdot \mathrm{comatr}(A). \tag{31}$$

La première égalité est celle qui nous a motivés de définir $\mathrm{comatr}(A)$. Or il découle de la formule définissant $\mathrm{comatr}(A)$ que $\mathrm{comatr}({}^tA) = {}^t\mathrm{comatr}(A)$ et donc ${}^t\mathrm{comatr}(A) \cdot {}^tA = \det({}^tA) I_n$; en utilisant dans cette équation ${}^tP {}^tQ = {}^t(QP)$ ainsi que $\det({}^tA) = \det(A)$, on obtient la seconde égalité (transposée).

4.4.3. Théorème. *Dans l'anneau $\mathrm{Mat}_n(R)$ des matrices $n \times n$ à coefficients dans un anneau commutatif R, une matrice A possède une matrice inverse (dans $\mathrm{Mat}_n(R)$) si et seulement si $\det(A)$ possède un inverse dans R. Si c'est la cas, la matrice inverse A^{-1} est donnée par $A^{-1} = \det(A)^{-1} \mathrm{comatr}(A)$.*

Preuve. Si B est une matrice inverse pour A, la multiplicativité du déterminant appliquée à $A \cdot B = I_n$ donne $\det(A)\det(B) = 1$ dans R, donc pour que A soit inversible dans $\mathrm{Mat}_n(R)$ il est nécessaire que $\det(A)$ soit inversible dans R. Mais si c'est le cas, proposition 4.4.2 montre que la matrice $\det(A)^{-1} \mathrm{comatr}(A)$ est bien une matrice inverse de A. Or, si un inverse existe, il est toujours unique. □

Ce théorème, appliqué à la matrice d'un endomorphisme $\phi \in \mathrm{End}(E)$ exprimé par rapport à une base quelconque, confirme que (la matrice et donc) ϕ est inversible (c'est-à-dire un isomorphisme) si et seulement si $\det(\phi) \neq 0$. Mais le théorème dit bien plus : il affirme par exemple aussi qu'une matrice $A \in \mathrm{Mat}_n(\mathbf{Z})$ possède une matrice inverse dans $\mathrm{Mat}_n(\mathbf{Z})$ si et seulement si $\det(A) \in \{1, -1\}$, car 1 et -1 sont les seuls éléments avec un inverse (multiplicatif) dans $\mathbf{Z}$. Et une matrice $A \in \mathrm{Mat}_n(\mathbf{Z})$ admet une autre matrice B qui lui est inverse modulo k (ce qui veut dire que AB et BA, sans forcément être égaux à I_n, ont des coefficients qui sont congruents modulo k aux coefficients correspondants de I_n, c'est-à-dire à 1 sur la diagonale et à 0 ailleurs) si et seulement si $\det(A)$ et k sont premiers entre eux (car c'est la condition pour que la classe de $\det(A)$ soit inversible dans $\mathbf{Z}/k\mathbf{Z}$).

4.4.4. Théorème [règle de Cramer]. *Un système d'équations de la forme (27) avec coefficients et inconnues dans un corps commutatif K possède une solution unique si et seulement si* $\det(A) \neq 0$. *Si c'est le cas, les solutions pour les inconnues sont données par* $x_j = \frac{\det(A \leftarrow_j \mathbf{b})}{\det(A)}$ *pour* $j = 1, \ldots, n$.

Preuve. Si $\det(A) \neq 0 \in K$ alors A est inversible, et $\mathbf{x} = A^{-1} \cdot \mathbf{b}$ est une solution. L'unicité de cette solution et l'expression pour les inconnues individuelles sont des conséquences de la proposition 4.4.1. Si par contre $\det(A) = 0$, alors l'équation $A \cdot \mathbf{x} = \vec{0}$ a des solutions non nulles, qui peuvent être ajoutées à toute solution éventuelle de $A \cdot \mathbf{x} = \mathbf{b}$, montrant qu'une telle solution ne peut jamais être unique. □

4.5. *Le polynôme caractéristique.*

Avec le déterminant comme outil pour tester si un endomorphisme est un isomorphisme par une *expression* en les coefficients de sa matrice (par rapport à une base fixée), nous pouvons caractériser les valeurs propres de l'endomorphisme comme les solutions d'une équation. Pour cela on constate d'abord :

4.5.1. Proposition. *Un scalaire $\lambda \in K$ est une valeur propre d'un endomorphisme $\phi \in \mathrm{End}(E)$ (avec E un K-espace vectoriel de dimension finie) si et seulement si* $\det(\lambda \,\mathrm{id}_E - \phi) = 0$.

Preuve. Un vecteur propre pour ϕ est par définition un vecteur non nul dans le noyau de l'application $\lambda \,\mathrm{id}_E - \phi$, et un tel vecteur existe si et seulement si cette application *n'est pas* injective, ce qui pour un endomorphisme de E est la même chose que de ne pas être un isomorphisme (grâce au théorème du rang). Or, d'après le théorème 4.3.6(5), c'est le cas précisément quand $\det(\lambda \,\mathrm{id}_E - \phi) = 0$. □

Si $A = \mathrm{Mat}_{\mathcal{B}}(\phi)$ pour une certaine base $\mathcal{B}$ de E, alors $\det(\lambda \,\mathrm{id}_E - \phi) = \det(\lambda I_n - A)$ pour tout $\lambda \in K$, d'après théorème 4.3.6(3). Pour trouver une équation pour les valeurs propres, il suffit donc de trouver une expression pour $\det(\lambda I_n - A)$ en termes de λ et des coefficients de A.

4.5.2. Définition. *Si $A \in \mathrm{Mat}_n(K)$, alors le polynôme caractéristique χ_A de A est*

$$\chi_A = \det(XI_n - A) \in K[X],$$

où $XI_n - A$ est la matrice $M \in \mathrm{Mat}_n(K[X])$ telle que $M_{i,i} = X - A_{i,i}$ et $M_{i,j} = -A_{i,j}$ si $i \neq j$.

Par exemple si $A = \left(\begin{smallmatrix} 3 & 5 \\ -1 & 7 \end{smallmatrix}\right)$ on a $\chi_A = \left|\begin{smallmatrix} X-3 & -5 \\ 1 & X-7 \end{smallmatrix}\right| = X^2 - 10X + 26$. En fait, le polynôme caractéristique peut être calculé sans connaître explicitement les coefficients de A ; par exemple pour une matrice 3×3 générique

$$A = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix},$$

le polynôme caractéristique est donné par

$$\chi_A = \begin{vmatrix} X-a & -b & -c \\ -d & X-e & -f \\ -g & -h & X-i \end{vmatrix} = X^3 - (a+e+i)X^2 + (ae - bd + ai - cg + ei - hf)X - \det(A).$$

Cela nous mène au constat suivant pour polynôme caractéristique des matrices $n \times n$ en général.

4.5.3. Proposition. *Pour tout $A \in \mathrm{Mat}_n(K)$, le polynôme caractéristique χ_A est un polynôme unitaire de degré n. En fonction de A, chaque coefficient de χ_A est donné par une expression en les coefficients de A, dont en particulier $-\mathrm{tr}(A)$ pour le coefficient de X^{n-1}, où $\mathrm{tr}(A) = \sum_{i=1}^{n} A_{i,i}$ est la trace de A, et $\det(-A) = (-1)^n \det(A)$ pour le coefficient de X^0 (le coefficient constant) dans χ_A.*

Preuve. Le fait que χ_A est de degré n et que ses coefficients sont donnés par des expressions en les coefficients de A est clair à partir des définitions. La description concrète du coefficient de X^i se déduit assez facilement pour $i = n, n-1, 0$ en regardant comment on trouve des termes de ces degrés dans l'expansion du déterminant (pour $i = 0$ on peut aussi considérer la substitution $X := 0$). Pour être plus précis, on obtient en général une contribution au terme de X^i en multipliant ensemble une combinaison de i facteurs X parmi les n qui sont présents sur le diagonal, et qui sont encore multipliés par un produit de $n-i$ coefficients de $-A$; pour une combinaison S fixée on obtient comme contribution le déterminant $(n-i) \times (n-i)$ (le mineur) extrait de $-A$ sur les lignes et les colonnes correspondantes au *complément* de S. Pour $i = n$ et $i = 0$ il n'y a qu'une telle combinaison (la combinaison pleine respectivement vide), et pour $i = n-1$ il y a $\binom{n}{n-1} = n$ combinaisons, qui contribuent chacun un coefficient diagonal de $-A$. □

4.5.4. Théorème. *Les valeurs propres de $\phi \in \mathrm{End}(E)$ sont les racines du polynôme caractéristique χ_A, où $A = \mathrm{Mat}_{\mathcal{B}}(\phi)$ pour une base quelconque $\mathcal{B}$ de E.*

Preuve. Pour $\lambda \in K$, la matrice de $\lambda \,\mathrm{id}_E - \phi$ par rapport à la base $\mathcal{B}$ est $\lambda I_n - A$ (car la matrice de l'homothétie $\lambda \,\mathrm{id}_E$ est toujours λI_n, quelle que soit la base). Son déterminant s'annule si et seulement si λ est une valeur propre de ϕ, et il vaut $\det(\lambda I_n - A) = \chi_A[X := \lambda]$ d'après le théorème 4.3.2(1). □

On voit que si ϕ possède n valeurs propres distinctes $\lambda_1, \ldots, \lambda_n$, alors avec $A = \mathrm{Mat}_{\mathcal{B}}(\phi)$ pour une base quelconque $\mathcal{B}$ de E, le polynôme χ_A est unitaire de degré n avec $\lambda_1, \ldots, \lambda_n$ pour racines, ce qui n'est possible que si $\chi_A = (X - \lambda_1) \ldots (X - \lambda_n)$. Donc dans ce cas le polynôme caractéristique ne dépend pas de la base utilisée pour exprimer la matrice de ϕ. Mais il n'est pas difficile de montrer directement que cela est le cas en général, ce qui permettra de parler simplement du polynôme caractéristique de ϕ.

4.5.5. Proposition/Définition. *Pour $\phi \in \mathrm{End}(E)$, le polynôme χ_A pour $A = \mathrm{Mat}_{\mathcal{B}}(\phi)$ ne dépend pas de la base $\mathcal{B}$, et est appelé le polynôme caractéristique χ_ϕ de ϕ.*

Preuve. Si $\mathcal{B}'$ est une autre base et $A' = \mathrm{Mat}_{\mathcal{B}'}(\phi)$, on aura $A' = P^{-1} \cdot A \cdot P$, où P est la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$. Comme $P^{-1} \cdot I_n \cdot P = I_n$, on aura $XI_n - A' = P^{-1} \cdot (XI_n - A) \cdot P$ dans $\mathrm{Mat}_n(K[X])$, et donc $\chi_{A'} = \det(P^{-1} \cdot (XI_n - A) \cdot P) = \det(P)\, \chi_A \det(P^{-1}) = \chi_A$ par multiplicativité du déterminant. □

Si ϕ est diagonalisable, cette proposition appliquée pour une base de diagonalisation montre que χ_ϕ est alors scindé, et la multiplicité de chaque racine λ de χ_ϕ est égale à la dimension de son espace propre. Réciproquement, si χ_ϕ est scindé et la dimension de chaque espace propre atteint la multiplicité de la racine correspondante dans χ_ϕ, alors la somme (toujours directe) des espace propres est de dimension $\deg(\chi_\phi) = \dim(E)$, donc elle est E tout entier, et ϕ est diagonalisable. On a donc

4.5.6. Critère. *Un endomorphisme ϕ est diagonalisable si et seulement si χ_ϕ est scindé, et la dimension de l'espace propre pour chaque valeur propre λ égale à la multiplicité de λ comme racine de χ_ϕ.*

Le calcul du polynôme caractéristique est simple en principe, mais peut être fastidieux quand la dimension n n'est pas petite (disons pour $n > 3$), car les méthodes habituelles pour simplifier une matrice dont on veut calculer le déterminant, en utilisant son caractère n-linéaire alterné, peuvent être frustrées par la présence de X (on ne peut pas inverser un polynôme non constant). Quelques principes simples peuvent néanmoins souvent simplifier le calcul du polynôme caractéristique.

4.5.7. Proposition. *Si pour une base $\mathcal{B}$ de E la matrice $M = \mathrm{Mat}_{\mathcal{B}}(\phi)$ est de la forme $\left(\begin{smallmatrix} A & C \\ \mathbf{0} & B \end{smallmatrix}\right)$ décrite dans théorème 4.3.2(6) avec $A \in \mathrm{Mat}_i(K)$ et $B \in \mathrm{Mat}_{n-i}(K)$, on a une décomposition $\chi_\phi = \chi_A\, \chi_B$.*

Preuve. Il suffit d'appliquer le théorème 4.3.2(6), compte tenu de la forme de $XI_n - M$. □

4.5.8. Corollaire. *Pour un sous-espace ϕ-stable V de E, le polynôme caractéristique $\chi_{\phi|_V}$ divise χ_ϕ.*

Preuve. En choisissant une base de E qui commence avec une base de V, la proposition précédente s'applique, et A est la matrice de la restriction $\phi|_V$, par rapport à la base de V, donc $\chi_{\phi|_V} = \chi_A$. □

4.5.9. Proposition. *Si $\phi \in \mathrm{End}(E)$ possède une matrice triangulaire T par rapport à une certaine base (on dit dans ce cas que ϕ est trigonalisable), alors on a une factorisation $\chi_\phi = \prod_{i=1}^{n}(X - a_i)$ du polynôme caractéristique, où $a_1, \ldots, a_n \in K$ sont les coefficients diagonaux de T.* □

La réciproque est aussi vraie (si χ_ϕ est scindé dans $K[X]$ alors ϕ est trigonalisable), mais on laisse cela pour le dernier chapitre (théorème 5.2.7), car il n'y a pas de rapport direct avec les déterminants.

Chapitre 5. Réduction d'endomorphismes.

Dans ce dernier chapitre nous allons approfondir l'étude des endomorphismes d'un espace vectoriel de dimension finie n. Rappelons (proposition 2.2.2 et la remarque qui suit) qu'un tel endomorphisme est certainement diagonalisable s'il admet n valeurs propres distinctes (ce qui est le maximum possible). On sait aussi (théorème 4.5.4) qu'on peut trouver les valeurs propres de ϕ comme les racines du polynôme caractéristique χ_ϕ. Par conséquent, si χ_ϕ est scindé *et a racines simples*, alors ϕ sera toujours diagonalisable (c'est aussi évident dans le critère 4.5.6). Si ce cas est le cas générique pour $K = \mathbf{C}$ (c'est-à-dire qu'il se produit presque toujours, en fonction des coefficients d'une matrice de ϕ), il n'est pas le seul cas possible, et les autres cas, même s'ils sont plus rares, sont d'autant plus intéressants. On va étudier notamment ce qu'on pourra dire dans le cas où χ_ϕ possède une ou plusieurs racines multiples. On verra qu'il est extrêmement rare que la dimension d'un espace propre pour d'une racine multiple de χ_ϕ atteigne sa multiplicité, et dans le cas contraire ϕ n'est pas diagonalisable, encore d'après le critère 4.5.6.

5.1. *Le polynôme minimal.*

Le fait que le polynôme caractéristique de $\phi \in \mathrm{End}(E)$ s'annule si on y substitue une valeur propre λ est basé sur la détection par le déterminant (en s'annulant) de la manque d'injectivité de $\lambda\,\mathrm{id}_E - \phi$, qui correspond précisément à l'existence de vecteurs propres. Mais le déterminant ne peut pas détecter plus une fois qu'il s'annule ; il serait notamment une erreur de croire que pour une racine multiple λ de χ_ϕ la dimension de $\mathrm{Ker}(\lambda\,\mathrm{id}_E - \phi)$ soit liée à la multiplicité de la racine (même si elle ne peut pas la dépasser). L'exemple du endomorphisme de K^2 de matrice $\left(\begin{smallmatrix}\lambda & x\\ 0 & \lambda\end{smallmatrix}\right)$, avec $\lambda, x \in K$, illustre cela : son polynôme caractéristique est $(X-\lambda)^2$ dans tous les cas, avec donc λ comme racine double, mais la condition $\dim \mathrm{Ker}(\lambda\,\mathrm{id}_E - \phi) = 2$ n'est vérifiée que dans le cas très particulier où $x = 0$, où ϕ est l'homothétie $\lambda\,\mathrm{id}_E$.

On obtiendra plus d'information par l'étude du polynôme minimal, un polynôme qui comme le polynôme caractéristique est défini en termes de ϕ et qui aura comme racines les valeurs propres, mais dont la définition est basée sur l'annulation de l'espace E tout entier, après la substitution $X := \phi$. Ainsi dans l'exemple le polynôme minimal ne sera $X - \lambda$ que dans le cas $x = 0$, pour indiquer que dans ce cas l'espace propre $\mathrm{Ker}(\lambda\,\mathrm{id}_E - \phi)$ est déjà E tout entier, mais pour $x \neq 0$ le polynôme minimal sera $(X-\lambda)^2$, indiquant que seulement $\mathrm{Ker}\big((\lambda\,\mathrm{id}_E - \phi)^2\big) = E$. Le polynôme minimal détecte donc la différence entre la situation diagonalisable $x = 0$ et le reste, mais il ne mesure pas non plus la dimension de l'espace propre ; en effet, plus la dimension de l'espace propre pour λ est grande, moins le polynôme minimal sera en général de degré élevé, car $\lambda\,\mathrm{id}_E - \phi$ annule déjà une grande partie de l'espace.

5.1.1. Définition. *Pour un endomorphisme ϕ d'un K-espace vectoriel de dimension finie, le polynôme minimal μ_ϕ de ϕ est le polynôme unitaire annulateur de ϕ du plus petit degré, dont l'existence est affirmée dans la proposition 3.3.4.*

On peut trouver le polynôme minimal μ_ϕ à l'aide de la matrice $A = \mathrm{Mat}_{\mathcal{B}}(\phi)$ par rapport une base $\mathcal{B}$ de E, en calculant successivement ses puissances $A^0 = I_n$, A, A^2, A^3, et en cherchant la première relation de dépendance linéaire entre ces matrices ; si cette relation est $A^d = c_0 A^0 + \cdots + c_{d-1}A^{d-1}$, alors on a $\mu_\phi = X^d - c_{d-1}X^{d-1} - \cdots - c_0 X^0$. Même ce conceptuellement cette méthode est simple, ce n'est pas forcément la meilleur méthode dans la pratique ; on mentionnera d'autres possibilités dans la suite.

Comme μ_ϕ est un polynôme annulateur de ϕ, la proposition 2.1.4 affirme que toute valeur propre de ϕ est une racine de μ_ϕ. Grâce à la minimalité on a ici également l'implication réciproque.

5.1.2. Théorème. *Les racines du polynôme minimal μ_ϕ forment l'ensemble des valeurs propres de ϕ.*

Preuve. Si λ est racine de μ_ϕ, celui-ci est divisible par $X-\lambda$: on peut écrire $\mu_\phi = (X-\lambda)Q$ avec $Q \in K[X]$. Comme $\deg(Q) < \deg(\mu_\phi)$ on a $Q[X := \phi] \neq \mathbf{0} \in \mathrm{End}(E)$. Il existe donc un vecteur $v \neq 0$ dans $\mathrm{Im}(Q[X := \phi])$, qu'on peut donc écrire $v = Q \cdot_\phi w$ pour un certain $w \in E$. On aura alors $(\phi - \lambda \mathrm{id}_E)(v) = (X-\lambda) \cdot_\phi (Q \cdot_\phi w) = \mu_\phi \cdot_\phi w = \mathbf{0}(w) = \vec{0}$. Ainsi λ est une valeur propre de ϕ. □

En comparant ce théorème avec le théorème 4.5.4, on voit que le polynôme minimal partage cette propriété avec le polynôme caractéristique. Pour beaucoup d'endomorphismes il s'agit d'un même polynôme, mais comme on a déjà vu ce n'est pas toujours le cas (par example pour les homothéties). Leurs définitions confèrent des propriétés assez différentes aux notions de polynôme caractéristique et de polynôme minimal. Basé sur le déterminant, le polynôme caractéristique dépend d'une façon régulière (car polynomiale) de (la matrice de) l'endomorphisme (proposition 4.5.3), ce qui n'est pas le cas du polynôme minimal. En fait le polynôme minimal de $\phi \in \mathrm{End}(E)$ est toujours untiaire, mais pas forcément de degré $n = \dim(E)$; cela implique que le degré et les coefficients de μ_ϕ "sautent" parfois quand on varie ϕ. Cette dépendance irrégulière rend le polynôme minimal un peu plus délicat à déterminer, mais lui permet en revanche de refléter des propriétés spécifiques de ϕ que le polynôme caractéristique ne saura pas détecter. Ce sera notamment le cas pour la propriété d'être diagonalisable. Donnons d'abord le sens facile de ce résultat.

5.1.3. Proposition. *Si ϕ est diagonalisable, alors le polynôme μ_ϕ est scindé et à racines simples.*

Preuve. Si ϕ est diagonalisable, avec comme valeurs propres *distinctes* $\lambda_1, \ldots, \lambda_l \in K$, alors μ_ϕ sera donné par le produit $P = (X-\lambda_1)\ldots(X-\lambda_l)$. D'une part on voit que $P[X := \phi]$ annule chaque vecteur propre de ϕ, et donc E tout entier qui possède une base de tels vecteurs ; d'autre part μ_ϕ doit avoir chacun des λ_i comme racine (d'après le théorème 5.1.2), et ne saura donc pas être de degré $< l$. □

5.1.4. Théorème. *Si $\phi \in \mathrm{End}(E)$ possède un polynôme annulateur $P \in K[X]$ qui est scindé et à racines simples, alors ϕ est diagonalisable.*

Preuve. Soit $P = (X-a_1)\ldots(X-a_l)$ un tel polynôme. Par hypothèse tous les a_i sont distincts, et les facteurs $X-a_i$ sont donc premier entre eux deux à deux. Le théorème 3.5.4 de décomposition des noyaux s'applique alors, et dit que l'espace E se décompose en somme directe de sous-espaces $\mathrm{Ker}(\phi - a_i \mathrm{id}_E)$ pour $i = 1, \ldots, l$, dont ceux qui ne sont pas nuls sont les espaces propres de ϕ. Alors ϕ est diagonalisable. □

5.1.5. Corollaire. *Le polynôme μ_ϕ est scindé à racines simples si et seulement si ϕ est diagonalisable.* □

Le théorème dit aussi par exemple que tout endomorphisme vérifiant $\phi^2 = \phi$ (qu'on appelle un *projecteur*) est diagonalisable, car $X^2 - X = X(X-1)$ est scindé à racines simples $0, 1$, et que c'est aussi le cas des endomorphismes vérifiant $\phi^2 = \mathrm{id}_E$ (dits *involutions*) car $X^2 - 1 = (X-1)(X+1)$, *sauf* si dans K on a l'égalité $2 = 0$, et donc $1 = -1$ (de tels corps sont dits de caractéristique 2, et $K = \mathbf{Z}/2\mathbf{Z}$ en est un exemple). Pour $K = \mathbf{C}$, un endomorphisme vérifiant $\phi^m = \mathrm{id}_E$ pour $m \in \mathbf{N}$ est aussi toujours diagonalisable, car $X^m - 1 \in \mathbf{C}[X]$ possède m racines distinctes $\exp(2k\pi\mathbf{i}/m)$ pour $k = 0, \ldots, m-1$.

On a vu que si χ_ϕ est scindé et à racines simples, cela entraîne que ϕ est diagonalisable (le critère 4.5.6, avec tous les multiplicités égales à 1), donc μ_ϕ est aussi scindé et à racines simples, et comme leurs ensembles de racines sont les mêmes (l'ensemble des valeurs propres de ϕ), on aura $\chi_\phi = \mu_\phi$ dans ce cas. Mais on n'a pas l'implication réciproque : quand μ_ϕ est scindé et à racines simples, certaines des ses racines (c'est-à-dire des valeurs propres de ϕ) peuvent être des racines multiples de χ_ϕ ; encore une fois les homothéties (en dimension > 1) en fournissent un exemple.

Dans le cas où ϕ n'est pas diagonalisable, on a soit que μ_ϕ n'est pas scindé (sur K), soit qu'il est scindé mais possède au moins une racine multiple. On étudiera ci-dessous le second cas en détail, mais pour le premier cas de figure on se contentera dans ce cours de dire qu'on peut étudier *à la place de* ϕ un automorphisme d'un espace vectoriel sur un corps $K' \supset K$ (par exemple $K' = \mathbf{C}$ si $K = \mathbf{R}$) dont la matrice est la même que celle de ϕ (dans une base convenable), pour rendre le polynôme minimal scindé sur K'. Il est important de savoir que cet "extension du corps de base" ne changera pas μ_ϕ lui-même.

5.1.6. Proposition. *Le polynôme minimal est déterminé par la matrice A de ϕ (dans une base donnée de E), et il ne change pas si A est interprétée comme élément de $\mathrm{Mat}_n(K')$ pour un corps $K' \supset K$.* □

Preuve. On a vu que qu'on saura $\mu_\phi = X^d - c_{d-1}X^{d-1} - \cdots - c_0X^0$ dès que deux conditions sont vérifiées : la famille de matrices $[A^0, \ldots, A^{d-1}]$ est libre, et $A^d = c_0A^0 + \cdots + c_{d-1}A^{d-1}$ (et la première condition assure que les coefficients $c_0, \ldots, c_{d-1} \in K$ sont entièrement fixés par la seconde condition). La seconde condition ne change pas si l'on remplace K par K', donc ils suffit de vérifier que la première reste valable par rapport à K' : une famille de matrice dans $\mathrm{Mat}_n(K)$ qui est libre sur K est aussi libre sur K' (où la famille est considérée comme étant dans le K'-espace $\mathrm{Mat}_n(K')$). Pour voir cela, on note d'abord que dire qu'une combinaison linéare d'une famille donnée de matrices vaut la matrice nulle équivaut à poser une système d'équations *linéaires* en les coefficients de la combinaison (une équation pour chaque position de la matrice). Or, une réflexion sur les méthodes de résoudre des systèmes linéaires (par exemple par l'échelonnement) nous apprendra qu'elles ne se servent que des tests d'égalité (à 0) et des opérations *arithmétiques* (pas de racines carrées par exemple) qui restent dans K ; si donc la conclusion est que le système n'a qu'une solution triviale sur K (tous les coefficients sont nuls, et la famille est donc libre), alors le rasinnement et la conclusion sont également valable pour l'indépendance linéaire sur K'. □

Quand on connaît une décomposition du polynôme minimal (ou tout autre polynôme annulateur), le théorème de décomposition des noyaux est d'une grande utilité. Mais son application est limitée au cas où les facteurs sont premiers entre eux. Le résultat suivant, qui est plus élémentaire, s'applique sans une telle limitation, mais exploite la minimalité de μ_ϕ comme polynôme annulateur de ϕ. Il caractérise le quotient de μ_ϕ par un diviseur P en termes du sous-espace *image* du polynôme $P[\phi]$ en ϕ.

5.1.7. Lemme. *Si $\mu_\phi = QP$ est une décomposition, où les polynômes P et Q sont unitaires et $P \neq 1$, alors $V = \mathrm{Im}(P[\phi])$ est un sous-espace strict de E, et Q est le polynôme minimal de la restriction $\phi|_V$.*

Preuve. Le sous-espace V est ϕ-stable d'après la proposition 3.5.1, donc la restriction $\phi|_V$ de ϕ à V est un endomorphisme de V. Un polynôme Q' est annulateur de la restriction $\phi|_V$ si et seulement si $Q'P$ est annulateur de ϕ : pour tout vecteur $v \in V$ on peut écrire $v = P[\phi](w)$ pour un certain $w \in E$, et donc $Q'[\phi|_V](v) = Q' \cdot_\phi (P \cdot_\phi w) = (Q'P)[\phi](w)$; les deux conditions disent que cette expression est nulle pour tout $w \in E$. Il est alors clair que le polynôme Q du lemme est un polynôme unitaire annulateur de $\phi|_V$ du plus petit degré possible, autrement dit son polynôme minimal. On a $V \neq E$, car le polynôme Q a degré $\deg(\mu_\phi) - \deg(P) < \deg(\mu_\phi)$, trop petit pour être annulateur de ϕ (sur E tout entier). □

On en déduit que pour que μ_ϕ soit scindé, il suffit que χ_ϕ le soit (la réciproque s'avérera vraie aussi).

5.1.8. Corollaire. *Si le polynôme caractéristique χ_ϕ de $\phi \in \mathrm{End}(E)$ est scindé, μ_ϕ est aussi scindé.*

Ce corollaire est évident quand on sait que μ_ϕ divise toujours χ_ϕ, comme le dit le théorème de Cayley–Hamilton (section 5.3). Mais on peut le montrer déjà sans faire appel à ce théorème.

Preuve. On montre la contraposée. Supposons μ_ϕ non scindé. Alors il possède un facteur Q de degré > 1 qui est unitaire et irréductible sur K, et qui n'a donc en particulier aucune racine dans K. Le lemme 5.1.7 montre que Q est le polynôme minimal d'une restriction de ϕ à un sous-espace ϕ-stable (et non réduit à $\{0\}$, car $Q \neq 1$) ; dans le cas $Q = \mu_\phi$ on prendra $V = E$. Que ce polynôme minimal Q n'a pas de racine dans K veut dire d'après le théorème 5.1.2 que $\phi|_V$ n'a pas de valeur propre. Mais $\chi_{\phi|_V}$ est alors aussi sans racine dans K, et il divise χ_ϕ d'après le corollaire 4.5.8. Par conséquent χ_ϕ n'est pas scindé. □

5.2. *Sous-espaces caractéristiques, trigonalisation.*

Considérons maintenant la situation où μ_ϕ est scindé, mais avec certain facteurs $X - \lambda_i$ présents plusieurs fois (donc μ_ϕ n'est pas diagonalisable, d'après la proposition 5.1.3). En regroupant ces facteurs, on a

$$\mu_\phi = (X - \lambda_1)^{m_1} \ldots (X - \lambda_k)^{m_k} \qquad \text{avec } \lambda_1, \ldots, \lambda_k \text{ distincts, et } m_1, \ldots, m_k \geq 1.$$

Contrairement aux facteurs individuels $X - \lambda_i$, les facteurs regroupés sont premiers entre eux deux à deux. Le théorème 3.5.4 s'applique alors, et l'espace E se décompose en somme directe de sous-espaces $\mathrm{Ker}((\phi - \lambda_i \,\mathrm{id}_E)^{m_i})$ pour $i = 1, \ldots, k$. Ces sous-espaces seront d'une importance fondamentale dans ce cas non diagonalisable, plus encore que les espace propres.

5.2.1. Définition. *Pour $\phi \in \mathrm{End}(E)$ et λ une valeur propre de ϕ, le sous-espace caractéristique de ϕ pour λ est $E_\lambda = \mathrm{Ker}\big((\phi - \lambda\,\mathrm{id}_E)^m\big)$, où m est la multiplicité de λ comme racine du polynôme minimal μ_ϕ.*

Si $m = 1$ il s'agit juste de l'espace propre pour λ. En général E_λ *contient* l'espace propre pour λ, car un vecteur (propre) annulé par $\phi - \lambda\,\mathrm{id}_E$ sera aussi annulé par $(\phi - \lambda\,\mathrm{id}_E)^m$ pour $m \geq 1$.

Comme annoncé ci-dessus, le théorème de décomposition des noyaux donne le théorème suivante, qui est la décomposition vectorielle ("réduction de l'endomorphisme") la plus importante de ce cours.

5.2.2. Théorème. *Si μ_ϕ est scindé, alors l'espace E se décompose en une somme directe $E = \bigoplus_{i=1}^k E_{\lambda_i}$, où $E_{\lambda_1}, \ldots, E_{\lambda_k}$ sont les espaces caractéristiques pour les valeurs propres distinctes $\lambda_1, \ldots, \lambda_k$ de ϕ.*

Preuve. La seule chose à vérifier pour que le théorème 3.5.4 donne ce résultat, est que deux polynômes de la forme $(X - \lambda_i)^{m_i}$ (dont E_{λ_i} est le noyau de l'évaluation en ϕ), pour des valeurs λ_i distinctes, sont premiers entre eux. Ceci est évident, compte tenu de l'unicité de la factorisation de ces polynômes. □

D'après le corollaire 5.1.8, l'hypothèse que χ_μ est scindé peut remplacer celle que μ_ϕ est scindé.

5.2.3. Proposition. *Tout espace caractéristique E_λ possède un sous-espace ϕ-stable supplémentaire W. Alors λ n'est pas valeur propre de la restriction $\phi|_W$, et $\phi|_W - \lambda\,\mathrm{id}_W \in \mathrm{End}(W)$ est un isomorphisme.*

Preuve. Avec $P = \mu_\phi/(X-\lambda)^m$, les polynômes $(X-\lambda)^m$ et P sont premiers entre eux, et $W = \mathrm{Ker}(P[\phi])$ est un tel supplémentaire d'après le lemme 3.5.3. L'espace propre pour λ est contenu dans E_λ, donc son intersection avec W est $\{0\}$, et λ n'est pas valeur propre de $\phi|_W$. Cela veut dire que $\phi|_W - \lambda\,\mathrm{id}_W \in \mathrm{End}(W)$ est injectif, donc bijectif. (Si μ_ϕ est scindé, W est la somme directe des autre espaces caractéristiques.) □

On remarque que dans le cas d'une racine multiple λ du polynôme minimal, l'espace propre pour λ n'admet (contrairement à E_λ) *aucun* sous-espace supplémentaire ϕ-stable. En effet, si W était un tel sous-espace, $\phi|_W$ n'aurait pas de valeur propre λ par le même argument que ci-dessus, donc le polynôme minimal $\mu_{\phi|_W}$ de la restriction n'a pas λ comme racine, et $(X - \lambda)\mu_{\phi|_W}$ serait un polynôme annulateur de ϕ (car annulateur de ses restrictions aux sous-espaces supplémentaires E_λ et W) sans être multiple de μ_ϕ, en contradiction avec la proposition 3.3.4. D'où l'importance de l'espace caractéristique E_λ.

Trigonalisation.

Considérons maintenant de plus proche un espace caractéristique $E_\lambda = \mathrm{Ker}\big((\phi - \lambda\,\mathrm{id}_E)^m\big)$. Le polynôme $(X - \lambda)^m$, qui est clairement un polynôme annulateur de la restriction $\phi|_{E_\lambda}$, est aussi son polynôme *minimal* : le polynôme minimal $\mu_{\phi|_W}$ de la restriction de ϕ au sous-espace W de la proposition 5.2.3 n'est pas divisible par $X - \lambda$, et si $(X - \lambda)^{m-1}$ était déjà annulateur de $\phi|_{E_\lambda}$, alors $(X - \lambda)^{m-1}\mu_{\phi|_W}$ serait annulateur de ϕ sans être multiple de μ_ϕ. Il existe donc un $v \in E_\lambda$ tel que $(\phi - \lambda\,\mathrm{id}_E)^{m-1}(v) \neq 0$. Si l'on pose $V_i = \mathrm{Ker}\big((\phi - \lambda\,\mathrm{id}_E)^i\big)$ pour $0 \leq i \leq m$, il est évident que $\{0\} = V_0 \subseteq \cdots \subseteq V_m = E_\lambda$. Mais comme en plus $v \notin V_{m-1}$, et plus généralement $(\phi - \lambda\,\mathrm{id}_E)^i(v) \in V_{m-i} \setminus V_{m-i-1}$ pour $i = 0, \ldots, m-1$, toutes les inclusions sont strictes : $\{0\} = V_0 \subset \cdots \subset V_m = E_\lambda$.

La dimension de $\mathrm{Ker}\big((\phi - \lambda\,\mathrm{id}_E)^i\big)$ augmente donc strictement avec i jusqu'à $i = m$, la multiplicité de λ comme racine de μ_ϕ, où elle atteint $\dim(E_\lambda)$. Après cela cette dimension n'augmentera plus : $\mathrm{Ker}\big((\phi - \lambda\,\mathrm{id}_E)^i\big) = E_\lambda$ pour tout $i > m$, car E_λ est déjà annulé par $(\phi - \lambda\,\mathrm{id}_E)^m$, et sur son sous espace supplémentaire W de la proposition 5.2.3, les puissances de $\phi - \lambda\,\mathrm{id}_E$ n'annulent aucun vecteur non nul.

5.2.4. Proposition. *Pour une valeur propre λ de ϕ, l'inclusion $\mathrm{Ker}\big((\phi - \lambda\,\mathrm{id}_E)^i\big) \subseteq \mathrm{Ker}\big((\phi - \lambda\,\mathrm{id}_E)^{i+1}\big)$ est stricte si et seulement si $i < m_\lambda$, où m_λ est la multiplicité de λ comme racine de μ_ϕ.* □

5.2.5. Corollaire. *Pour toute valeur propre λ de ϕ on a $m_\lambda \leq \dim(E_\lambda)$ dans la proposition 5.2.4.* □

On peut maintenant voir que si μ_ϕ est scindé, on peut trouver une base $\mathcal{B}$ de E telle que la matrice $\mathrm{Mat}_{\mathcal{B}}(\phi)$ est relativement simple, même si elle n'est par toujours diagonale. Grâce au théorème 5.2.2, on pourra faire en sorte que $\mathcal{B}$ soit une réunion de bases choisies séparément dans chaque espace caractéristique E_{λ_i} ; la matrice sera alors "diagonale en blocs" avec un bloc pour chaque E_{λ_i}.

Dans E_{λ_i} on pourra commencer à choisir une base de l'espace propre $V_1 = \mathrm{Ker}\big(\phi - \lambda_i \,\mathrm{id}_E\big)$, l'étendre (en appliquant le théorème de la base incomplète) à une base de $V_2 = \mathrm{Ker}\big((\phi - \lambda_i \,\mathrm{id}_E)^2\big)$, l'étendre encore à une base de V_3, et ainsi de suite jusqu'à obtenir un base de $V_m = E_{\lambda_i}$. Ainsi en appliquant $\phi - \lambda_i \,\mathrm{id}_E$ à un vecteur b de cette base, le résultat est toujours une combinaison linéaire de vecteurs de la base venant strictement avant b, et la matrice de $\phi|_{E_\lambda}$ dans cette base sera triangulaire supérieure avec des coefficients diagonaux tous égaux à λ. En combinant les blocs des E_{λ_i} individuels, cela prouve le résultat suivant.

5.2.6. Proposition. *Si E est la somme des espaces caractéristiques, alors ϕ est trigonalisable sur K.* □

La forme triangulaire du bloc de $\mathrm{Mat}_{\mathcal{B}}(\phi)$ correspondant à E_λ entraîne (selon la la proposition 4.5.9) que le polynôme caractéristique du bloc est $(X-\lambda)^d$ avec $d = \dim(E_\lambda)$. Le théorème suivant résume.

5.2.7. Théorème. *Pour $\phi \in \mathrm{End}(E)$ (avec E un K-espace vectoriel de dimension finie) sont équivalents :*
(1) le polynôme caractéristique χ_ϕ est scindé sur K,
(2) le polynôme minimal μ_ϕ est scindé sur K,
(3) l'espace E est la somme directe $\bigoplus_{i=1}^k E_{\lambda_i}$ des espaces caractéristiques, où $\lambda_1, \ldots, \lambda_k$ sont les valeurs propres distinctes de ϕ,
(4) l'endomorphisme ϕ est trigonalisable sur K.
Si ces conditions sont vérifiées, $\chi_\phi = (X-\lambda_1)^{d_1} \ldots (X-\lambda_k)^{d_k}$ où $d_i = \dim(E_{\lambda_i})$, et μ_ϕ divise χ_ϕ.

Preuve. Le corollaire 5.1.8 affirme que (1) implique (2), le théorème 5.2.2 que (2) implique (3), la proposition 5.2.6 que (3) implique (4), et la proposition 4.5.9 que (4) implique (1). Pour la formule pour χ_λ, on vient de voir que le polynôme caractéristique du bloc pour E_{λ_i} est $(X-\lambda_i)^{\dim(E_{\lambda_i})}$; il convient de les multiplier d'après la proposition 4.5.7. Que μ_ϕ divise ce produit découle du corollaire 5.2.5. □

Demander, comme le fait le critère 4.5.6, que la dimension de l'*espace propre* pour λ soit égale à la multiplicité de λ comme racine de χ_ϕ, demande donc que cet espace propre soit égal à E_λ, ou encore que la restriction $\phi|_{E_\lambda}$ soit l'homothétie de facteur λ. Si l'on fixe un sous-espace V pour être E_λ, il n'y a donc *qu'une seule possibilité* pour $\phi|_V$ qui permet d'avoir $V = E_\lambda$ à ϕ d'être diagonalisable. Mais si $\dim(V) \geq 2$, il y a plein d'autres possibilités qui donnent $V = E_\lambda$: pour une base quelconque de V, toute matrice triangulaire à coefficients diagonaux tous λ en donne une. Ceci justifie notre affirmation au début du chapitre qu'il est extrêmement rare que la dimension d'un espace propre pour une racine multiple de χ_ϕ atteigne sa multiplicité dans ce polynôme.

5.3. *Le théorème de Cayley–Hamilton.*

On mentionne maintenant un résultat célèbre qui rapproche encore plus le polynôme caractéristique au polynôme minimal, car il dit que χ_ϕ est un polynôme annulateur de ϕ. Des résultats déjà vus font soupçonner un tel lien étroit, notamment les théorèmes 4.5.4, 5.1.2, et 5.2.7. C'est un théorème très facile à retenir, ce qu'on conseille vivement de faire ; néanmoins, nous n'avons voulu montrer qu'on peut très bien se passer de son utilisation dans le développement de la théorie, parce que c'est un résultat finalement assez surprenant, et dont les démonstrations possibles sont très diverses et souvent subtiles. Dans cette section nous abordons quelques approches à une démonstration du théorème, comme illustration.

5.3.1. Théorème de Cayley–Hamilton. *Pour tout ϕ, son polynôme caractéristique vérifie $\chi_\phi[\phi] = 0$.*

Comme les polynômes annulateurs de ϕ sont les multiples du polynôme minimal μ_ϕ, un formulation équivalente est que μ_ϕ divise toujours χ_ϕ. Le théorème 5.2.7 affirme que c'est vrai dans le cas où ϕ est trigonalisable, quand χ_ϕ et μ_ϕ sont scindés. Ceci prouve déjà le théorème de Cayley–Hamilton quand K est un corps algébriquement clos comme $\mathbf{C}$, car dans ce cas χ_ϕ est toujours scindé.

Une méthode possible de démonstration dans le cas général consiste à trouver un corps K' contenant K et tel que χ_ϕ soit scindé dans $K'[X]$. La raison pour laquelle la validité du théorème pour K' entraîne sa validité pour K est que, si on représente ϕ par une matrice $A = \mathrm{Mat}_{\mathcal{B}}(\phi) \in \mathrm{Mat}_n(K)$ pour une base $\mathcal{B}$ quelconque, le polynôme caractéristique χ_A et ne change pas si on interprète A comme un élément de $\mathrm{Mat}_n(K')$ (ce qui est clair dans sa définition), et l'égalité $\chi_A[A] = 0$ sur K' reste valable sur K.

Par exemple pour $K = \mathbf{R}$ ou $K = \mathbf{Q}$ on peut prendre $K' = \mathbf{C}$, et le théorème est donc ainsi démontré pour ces cas. En fait on peut toujours *construire* un corps K' convenable à partir de K (mais pour cela on renvoie à un cours d'algèbre plus avancé), ce qui fournit une preuve dans le cas général.

Néanmoins, il est étrange de faire appel à un corps plus grand pour prouver une propriété qui ne concerne que des K-espaces. Voici une autre approche possible qui n'utilise que de l'algèbre linéaire sur K. Ne pouvant pas être sûr de l'existence d'un espace propre, l'idée est d'utiliser à sa place un sous-espace ϕ-stable V, non nul mais petit, à savoir celui engendré par les ϕ-images itérées d'un seul vecteur $v \neq 0$ (dont le choix reste libre). Si l'espace V est ainsi donné, il est naturel d'utiliser dans V la base $[v, \phi(v), \ldots, \phi^{d-1}(v)]$ où $\phi^d(v)$ est la première image à être linéairement dépendant des images précédentes. Comme on a vu déjà plusieurs fois, dès que ceci arrive, l'espace que les images engendrent est devenu ϕ-stable, et des applications supplémentaires de ϕ ne sortiront plus de ce sous-espace. La matrice de $\phi|_V$ par rapport à cette base est particulièrement simple : chaque vecteur de base est envoyé vers le suivant, sauf le dernier dont l'image est déterminé par le polynôme minimal de v pour ϕ.

5.3.2. Définition. *La matrice compagnon d'un polynôme unitaire* $P = X^n + c_{n-1}X^{n-1} + \cdots + c_0X^0$ *est*

$$\begin{pmatrix} 0 & 0 & 0 & \ldots & -c_0 \\ 1 & 0 & 0 & \ldots & -c_1 \\ 0 & 1 & 0 & \ldots & -c_2 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & -c_{d-1} \end{pmatrix} \tag{32}$$

(La matrice de (15) était la transposée d'une matrice compagnon.) Dans la situation mentionnée ci-dessus, la matrice de la restriction $\phi|_V$ de ϕ à l'espace engendré par les image itérées de v, par rapport à la base $[v, \phi(v), \ldots, \phi^{d-1}(v)]$, est la matrice compagnon du polynôme minimal P de v pour ϕ. De $P \cdot_\phi v$ on déduit $P \cdot_\phi \phi^i(v) = \phi^i(P \cdot_\phi v) = 0$ pour $0 \leq i < d$, donc P est même le polynôme minimal de $\phi|_V$.

5.3.3. Lemme. *Si A est la matrice compagnon d'un polynôme unitaire $P \in K[X]$, alors $\chi_A = P$.*

Preuve. C'est un simple calcul. Par définition du polynôme caractéristique, il faut montrer

$$\begin{vmatrix} X & 0 & 0 & \ldots & c_0 \\ -1 & X & 0 & \ldots & c_1 \\ 0 & -1 & X & \ldots & c_2 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & -1 & c_{n-1} + X \end{vmatrix} = X^n + \sum_{i=0}^{n-1} c_i X^i = P.$$

Plusiurs méthodes s'y prêtent. On peut commencer à modifier la matrice pour rendre nuls tous ses coefficients diagonaux à l'exception du dernier, en ajoutant la dernière ligne X fois de la précédente, puis cette ligne X fois de celle qui la précède, et ainsi de suite jusqu'à la première ligne. On obtient comme dernier coefficient de la première ligne $c_0 + X(c_1 + X(\cdots(c_{n-2} + X(c_{n-1} + X))\cdots)) = P$, et on vérifie facilement que

$$\begin{vmatrix} 0 & 0 & 0 & \ldots & P \\ -1 & 0 & 0 & \ldots & * \\ 0 & -1 & 0 & \ldots & * \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & -1 & c_{n-1} + X \end{vmatrix} = P.$$

Une autre méthode développe le déterminant de la matrice originale $M \in \mathrm{Mat}_n(K[X])$ directement par la dernière colonne. On remarque que pour $0 \leq i < n$ la matrice $M_{(i)}$ obtenue de M en supprimant la ligne contenant c_i et la dernière colonne est de la forme en blocs $\left(\begin{smallmatrix} L & \mathbf{0} \\ \mathbf{0} & U \end{smallmatrix}\right)$ où $L \in \mathrm{Mat}_i(K[X])$ est triangulaire inférieure avec des coefficients diagonaux X, et $U \in \mathrm{Mat}_{n-1-i}(K[X])$ est triangulaire supérieure avec des coefficients diagonaux -1. Donc $\det(M_i) = (-1)^{n-1-i}X^i$, ce qui donne $\det(M) = X^n + \sum_{i=0}^{n-1} c_i X^i = P$ comme voulu. Finalement on pourra aussi faire une démonstration par récurrence sur n : les cas $n \leq 1$ sont immédiats, et les cas $n > 1$ suit par récurrence après un développement par la première ligne. □

Preuve du théorème de Cayley–Hamilton. Par récurrence sur $n = \dim(E)$, le cas $n = 0$ étant évident (car alors $\mathrm{End}(E) = \{\mathbf{0}_E\}$). Si $n > 0$ on choisit un vecteur non nul v, et soit $P \in K[X]$ le polynôme minimal de v pour ϕ. Alors $\mathcal{B}_v = [v, \phi(v), \ldots, \phi^{d-1}(v)]$, où $d = \deg(P) > 0$, est une famille libre. C'est une base du sous-espace $V = \mathrm{Vect}(v, \ldots, \phi^{d-1}(v))$ qui est ϕ-stable, et la matrice $\mathrm{Mat}_{\mathcal{B}'}(\phi|_V)$ est la matrice compagnon de P. En étendant $\mathcal{B}_v$ à une base $\mathcal{B}$ de E, la matrice $\mathrm{Mat}_{\mathcal{B}}(\phi)$ sera de la forme en blocs $M = \left(\begin{smallmatrix} A & * \\ \mathbf{0} & B \end{smallmatrix}\right)$ où '$*$' désigne un bloc dont le contenu nous n'intéressera pas (et qui peut être différent pour chaque utilisation de '$*$'). On voit par une calcul direct que M^n est de la forme $\left(\begin{smallmatrix} A^n & * \\ \mathbf{0} & B^n \end{smallmatrix}\right)$ pour $n \in \mathbf{N}$, et donc $Q[M] = \left(\begin{smallmatrix} Q[A] & * \\ \mathbf{0} & Q[B] \end{smallmatrix}\right)$ pour tout $Q \in K[X]$. D'après la proposition 4.5.7 on a $\chi_\phi = \chi_A \chi_B$. Comme $A = \mathrm{Mat}_{\mathcal{B}'}(\phi|_V)$ est la matrice compagnon de P, on a $\chi_A = P$ (lemme 5.3.3) et $P[A] = \mathbf{0} \in \mathrm{Mat}_d(K)$. Or $B \in \mathrm{Mat}_{n-d}(K)$ où $n - d < n$, donc l'hypothèse de récurrence donne $\chi_B[B] = \mathbf{0} \in \mathrm{Mat}_{n-d}(K)$. On conclut par un calcul sous la forme en blocs : $\chi_\phi[M] = P[M] \cdot \chi_B[M] = \left(\begin{smallmatrix} \mathbf{0} & * \\ \mathbf{0} & * \end{smallmatrix}\right) \cdot \left(\begin{smallmatrix} * & * \\ \mathbf{0} & \mathbf{0} \end{smallmatrix}\right) = \left(\begin{smallmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{smallmatrix}\right)$. □

Une autre piste encore vise a démontrer $\chi_A[A] = \mathbf{0} \in \mathrm{Mat}_n(K)$ directement par des manipulations algébriques formelles. Une telle démonstration sera valable pour toute matrice à coefficients dans un anneau commutatif (car pour de telles matrices, ce type de manipulations sont valables). On en propose une, tout en mettant en garde pour le fait que la nature des expressions manipulées est inhabituelle (des matrices à coefficients polynomiaux) et nécessite prudence : certaines manipulations avec des polynômes ne sont valables que dans un contexte commutatif, et la justification de ce qui suit est subtile.

On commence avec l'identité (31) appliquée à la matrice $XI_n - A \in \mathrm{Mat}_n(K[X])$. Comme le déterminant de cette matrice est $\chi_A \in K[X]$, elle affirme l'existence d'une certaine matrice B telle que $(XI_n - A) \cdot B = \chi_A I_n$ (les coefficients de $B \in \mathrm{Mat}_n(K[X])$ sont explicitement donnés en termes de ceux de $XI_n - A$, et on peut voir que ce sont des polynômes de degré $< n$ en X, mais l'expression précise n'est pas utilisée dans la preuve.) Donc la combinaison linéaire $XB - A \cdot B$ donne la matrice $\chi_A I_n \in \mathrm{Mat}_n(K[X])$ dont les coefficients hors la diagonale principale sont tous nuls, est ceux sur la diagonle principale sont tous égaux à $\chi_A \in K[X]$. Pour mieux voir ce que cela veut dire, séparerons dans les coefficients de B, qui sont des polynômes en X, les termes selon leur degré en X, ce qui permet d'écrire $B = \sum_{i=0}^{n-1} X^i B_i$ où les matrices B_i sont dans $\mathrm{Mat}_n(K)$ (leurs coefficients sont constants). Si on écrit aussi $\chi_A = \sum_{i=0}^{n} c_i X^i$, on peut comparer dans l'équation $XB - A \cdot B = \chi_A I_n$ dans chacune des positions le coefficient du monôme X^i, ce qui donne les équations dans $\mathrm{Mat}_n(K)$

$$B_{i-1} - A \cdot B_i = c_i I_n \qquad \text{pour } 0 \leq i \leq n \tag{33}$$

où on a posé $B_{-1} = B_n = 0$ pour pouvoir écrire la même équation, les cas $i = 0$ et $i = n$ compris. Maintenant on prend l'instance de (33) pour chaque valeur de i, on la multiplie à gauche par A^i, et on forme la somme des équations matricielles ainsi obtenues. Le premier membre de la somme donne la matrice nulle, car les termes $A^i \cdot B_{i-1}$ et $-A^{i-1} \cdot A \cdot B_{i-1}$ s'annulent pour $i = 1, 2, \ldots, n$ et les termes restants $A^0 \cdot B_{-1}$ et $A^n \cdot A \cdot B_n$ sont nuls ; le second membre de la somme est $\sum_{i=0}^{n} c_i A^i = \chi_A[X := A]$. □

Finalement il y a un complément qu'on peut apporter au théorème, dans sa forme pour les endomorphismes ; non seulement μ_ϕ divise χ_ϕ, mais ils ont les *mêmes* facteurs irréductibles (la seule différence étant que la multiplicité d'un tel facteur peut être plus grand dans χ_ϕ). On l'a vu pour les facteurs de degré 1, correspondant aux racines. D'après le théorème de décomposition des noyaux on peut décomposer E en somme directe de sous-espaces V_i annulés chacun par une puissance d'un facteur irréductible différent P_i de μ_ϕ. Or on pourra montrer (pareillement à notre dernière preuve) que $\chi_{\phi|_{V_i}}$ est aussi une puissance de P_i, ce qui ne laisse pas de place pour d'autres facteurs irréductibles dans χ_ϕ.

Résumé des objectifs du cours.

Pour le premier chapitre toutes les notions doivent être bien comprises : combinaison linéaire, sous-espace, famille libre ou génératrice, base, coordonnées, dimension, application linéaire, matrice par rapport à des bases, isomorphisme, endomorphisme, rang, matrice de passage, changement de base. Dans les deux dernières sections il faut retenir surtout la proposition 1.5.2 (changement de base) et le théorème 1.6.2 du rang, qui avec le théorème 1.2.1 de la base incomplète forment le socle théorique de ce chapitre.

Le chapitre 2 sert surtout pour introduire et illustrer les problèmes de valeurs propres. Toutes les définitions ainsi que les propositions 2.1.3, 2.1.4, et 2.2.2 sont fondamentales et donc à retenir. La proposition 2.3.1 sera supplantée par des résultats plus précis, donc ni sa formulation ni sa démonstration sont nécessaires à retenir en tant que tel. Avec cela, les deux dernières sections de ce chapitre servent surtout de motivation, mais les exemples de la section 2.4 méritent d'être bien étudiés, car ils sont proche du type d'applications auquel on peut s'attendre dans les contrôles.

Le chapitre 3 est long et nécessaire pour servir de fondement pour la suite, mais une fois ses fondements assimilés, relativement peu de choses seront directement pertinentes pour le cours : la construction de $\mathbf{Z}/n\mathbf{Z}$ sert surtout pour sa relation avec la division euclidienne dans $\mathbf{Z}$ (mais il est utile de retenir l'exemple des corps $\mathbf{Z}/p\mathbf{Z}$ avec p premier, même s'il est rare qu'on ose les faire intervenir dans les contrôles), pour les polynômes le plus importante à retenir, mis à part les propriétés basiques du degré, est encore la division euclidienne, puis les notions de substitution et de racine. Sont importants à retenir dans ce chapitre : les propositions 3.2.3, 3.2.4, 3.3.3, 3.3.4, et 3.5.1, et les théorèmes 3.4.9 et 3.5.4.

Dans le chapitre 4, les deux premières sections sont préparatoires, les propriétés des formes multilinéaires alternées sont surtout à retenir dans leur application au déterminant d'un système de vecteurs, le seul exemple qu'on en verra dans la pratique. Pour le déterminant il est important de savoir qu'il est défini de façon naturelle pour les matrices et les endomorphismes, mais que le déterminant d'un système de n vecteurs a besoin d'être normalisé en spécifiant une base. Les énoncés contenus dans le théorème 4.3.2 servent continuellement et doivent être connus. Les propriétés des deux autres formes du déterminant (théorèmes 4.3.4 et 4.3.6) sont importantes également. La section 4.4 n'est pas beaucoup utilisée dans ce cours, mais on a intérêt à avoir vu ses résultats, surtout pour leur utilisation dans d'autres cours d'algèbre. La dernière section sur le polynôme caractéristique est fondamentale dans ce cours.

Le deux premières sections du chapitre 5 mènent aux résultats les plus complets de ce cours, mais certains des résultats sont préparatoires pour d'autres qui les supplantent. Pour le polynôme minimal on retiendra sa définition ainsi que les théorèmes 5.1.2 et 5.1.4, avec son corollaire 5.1.5 caractérisant les endomorphismes diagonalisables. Pour les sous-espaces caractéristiques on retiendra surtout leur définition, et les théorèmes 5.2.2 et 5.2.7. L'énoncé du théorème 5.3.1 de Cayley-Hamilton est joli et facile à mémoriser, donc il est a retenir, mais les preuves données sont optionnelles.

Table de matières.

Avant-propos 1
Introduction 2

1 Rappels de l'algèbre linéaire 2
1.1 Espaces vectoriels, sous-espaces, combinaisons linéaires, applications linéaires 2
1.2 Familles génératrices (d'un sous-espace), liées ou libres ; bases, dimension (finie) 4
1.3 Somme de sous-espaces, somme directe 6
1.4 Expression dans une base, matrices d'applications linéaires 8
1.5 Changement de base 11
1.6 Équivalence de matrices rectangulaires, image, noyau, et rang 12
1.7 Endomorphismes, similitude de matrices carrées 13

2 Vecteurs propres, valeurs propres 14
2.1 Définition de vecteur propres et de valeur propres ; premières propriétés 14
2.2 Diagonalisation 16
2.3 Existence de valeurs propres 18
2.4 Exemples d'application des vecteurs propres 19

3 Corps et anneaux, polynômes 24
3.1 Définition de corps et anneaux 24
3.2 Anneaux de polynômes 25
3.3 Substitution dans $K[X]$, racines, polynômes annulateurs 27
3.4 Quelques éléments d'arithmétique dans $K[X]$ 28
3.5 Décomposition des noyaux 31

4 Déterminants 33
4.1 Déterminants en dimension $n \leq 3$, formes linéaires 33
4.2 Formes multilinéaires alternées 34
4.3 Définitions de déterminant 37
Déterminant d'une matrice à coefficients dans un anneau commutatif 37
Déterminant dans une base d'un système de n vecteurs 39
Déterminant d'un endomorphisme 40
4.4 Déterminants et matrices inverses: la règle de Cramer 40
4.5 Le polynôme caractéristique 42

5 Réduction d'endomorphismes 44
5.1 Le polynôme minimal 44
5.2 Sous-espaces caractéristiques, trigonalisation 46
Trigonalisation 47
5.3 Le théorème de Cayley–Hamilton 48
Résumé des objectifs du cours 51