Polynômes et racines

Dans ce chapitre, $\mathbb{K}$ est l'un des corps $\mathbb{R}$ ou $\mathbb{C}$, sauf dans la dernière partie où une généralisation au cas des corps finis $\mathbb{F}_p$ est abordée.

1 Construction des anneaux de polynômes

Jusqu'ici, vous n'avez jamais distingué les polynômes des fonctions polynomiales, et pourtant nous allons faire en sorte que les polynômes ne soient plus des fonctions.

La fonction polynomiale $x \overset{P}{\longmapsto} 3x^2+4x+1$ est définie de $\mathbb{R}$ dans $\mathbb{R}$, mais de nouvelles formes d'objets $P(x)$ apparaîtront dans les prochains mois dans lesquels x est autre chose qu'un nombre. Par exemple, pour tout $M \in \mathcal{M}_n(\mathbb{R})$, on aura envie et besoin de noter $P(M)$ la matrice $3M^2+4M+I_n$. Plus généralement, dans tout anneau A dans lequel on sait multiplier par un réel, l'objet $3a^2+4a+1_A$ mérite d'être noté $P(a)$ pour tout $a \in A$. Le polynôme P gagnerait ainsi à être défini comme une recette de calcul dont les ingrédients sont les coefficients 1, 4 et 3 plutôt que comme une fonction.

Quand on est habitué aux fonctions polynomiales, on a l'impression qu'en connaissant la fonction $x \longmapsto 3x^2+4x+1$, on connaît ses coefficients, mais c'est l'inverse qui est vrai. Si vous connaissez 1, 4 et 3, vous pouvez calculer $3x^2+4x+1$ pour tout $x \in \mathbb{R}$, mais si on vous donne la fonction à proprement parler et non pas son expression, vous savez juste quel réel elle associe à tout x. Cela ne vous donne en aucun cas l'expression de la fonction et ses coefficients. Bref, renonçons aux fonctions et tâchons de créer un objet polynôme ouvert sur l'extérieur dont les coefficients seront l'essentiel.

Bien sûr, nos polynômes ne seront intéressants que si nous pouvons les additionner et les multiplier. Nous donnerons bientôt un sens rigoureux aux calculs suivants, qui relèvent seulement du désir pour l'instant. Étant donnés deux polynômes $P = a_nX^n+\ldots+a_1X+a_0$ et $Q = b_nX^n+\ldots+b_1X+b_0$ de degré au plus n, voilà ce que nous voulons pouvoir écrire :

$$P+Q = \sum_{k=0}^{2n}(a_k+b_k)X^k \qquad \text{et} \qquad P\times Q = \sum_{k=0}^{2n}\big(a_0b_k+\ldots+a_kb_0\big)X^k = \sum_{k=0}^{2n}\left(\sum_{i=0}^{k}a_ib_{k-i}\right)X^k,$$

où les termes ont été regroupés par degré. Y a plus qu'à forcer le destin.

1.1 Définition de l'anneau $\mathbb{K}[X]$

Définition **(Polynôme à une indéterminée à coefficients dans $\mathbb{K}$)** On appelle *polynôme (à une indéterminée) à coefficients dans* $\mathbb{K}$ toute suite *presque nulle* d'éléments de $\mathbb{K}$, i.e. toute suite $(a_k)_{k\in\mathbb{N}}$ d'éléments de $\mathbb{K}$ dont les éléments sont nuls à partir d'un certain rang. Pour tout $k \in \mathbb{N}$, le coefficient a_k est appelé le *coefficient de degré k* du polynôme.

On note généralement X l'*indéterminée* $X = (0,1,0,0,\ldots)$ et $\mathbb{K}[X]$ l'ensemble des polynômes à coefficients dans $\mathbb{K}$, mais parfois on préférera les noter T et $\mathbb{K}[T]$, etc.

Identification des coefficients : Deux polynômes sont égaux si et seulement si leurs coefficients sont égaux.

Avec cette définition, un polynôme est une suite $(a_0,a_1,\ldots,a_n,0,0,\ldots)$ à coefficients dans $\mathbb{K}$ que nous noterons bientôt $a_nX^n+\ldots+a_1X+a_0$. Gardez cet objectif en tête, il vous aidera à comprendre les prochaines définitions. Pour tout $\lambda \in \mathbb{K}$, notons au moins dès à présent λ le *polynôme constant* $(\lambda,0,0,\ldots)$. En particulier, le polynôme 0 est appelé le *polynôme nul*.

Le principe d'identification des coefficients est une trivialité maintenant que les polynômes sont des suites. Au lycée, on vous a énoncé un principe plus délicat d'identification des coefficients propre aux fonctions polynomiales, plus délicat car la connaissance d'une fonction polynomiale n'équivaut pas à la connaissance de son expression, i.e. de ses coefficients. Nous prouverons ce résultat plus loin.

Définition **(Degré d'un polynôme, coefficient dominant, polynôme unitaire)**

- **Degré du polynôme nul :** On convient que $\deg(0) = -\infty$.
- **Degré d'un polynôme non nul :** Soit $P = (a_k)_{k\in\mathbb{N}} \in \mathbb{K}[X]$ un polynôme non nul. Le plus grand indice k pour lequel $a_k \neq 0$ est appelé le *degré de P* et noté $\deg(P)$.

  Le coefficient de degré $\deg(P)$ de P est appelé son *coefficient dominant*. S'il est égal à 1, on dit que P est *unitaire*.

Pour tout $n \in \mathbb{N}$, l'ensemble des polynômes de degré AU PLUS n de $\mathbb{K}[X]$ est noté $\mathbb{K}_n[X]$.

Attention, le polynôme nul n'a pas de coefficient dominant. Quand on veut utiliser le coefficient dominant d'un polynôme P dans un raisonnement, il faut supposer P non nul.

Exemple $7X^3+2X^2-X+5$ a pour degré 3 et coefficient dominant 7, tandis que X^2+3X-1 est unitaire.

Il est facile à présent de définir une loi d'addition sur $\mathbb{K}[X]$. L'ensemble $\mathbb{K}^{\mathbb{N}}$ des suites à valeurs dans $\mathbb{K}$ est un groupe pour la loi héritée de $\mathbb{K}$ définie par la relation $(u_k)_{k\in\mathbb{N}}+(v_k)_{k\in\mathbb{N}}=(u_k+v_k)_{k\in\mathbb{N}}$ pour tous $(u_k)_{k\in\mathbb{N}},(v_k)_{k\in\mathbb{N}}\in\mathbb{K}^{\mathbb{N}}$. Montrons que $\mathbb{K}[X]$ en est un sous-groupe. Or déjà, $0=(0)_{n\in\mathbb{N}}\in\mathbb{K}[X]$. Pour la stabilité par différence, soient $P=(a_k)_{k\in\mathbb{N}},Q=(b_k)_{k\in\mathbb{N}}\in\mathbb{K}[X]$. Ainsi, $a_k=b_k=0$ à partir d'un certain rang N, donc $a_k-b_k=0$ pour tout $k\geqslant N$, donc $P-Q\in\mathbb{K}[X]$.

Pour tous $P=(a_k)_{k\in\mathbb{N}},Q=(b_k)_{k\in\mathbb{N}}\in\mathbb{K}[X]$, posons maintenant $P\times Q=\left(\sum_{i=0}^{k}a_ib_{k-i}\right)_{k\in\mathbb{N}}=\left(a_0b_k+\ldots+a_kb_0\right)_{k\in\mathbb{N}}$. S'agit-il bien d'un polynôme ? Autrement dit, la famille $P\times Q$ est-elle bien presque nulle ? Eh bien oui, car si on note N un rang à partir duquel $a_k=b_k=0$, alors pour tout $k\geqslant 2N$: $\displaystyle\sum_{i=0}^{k}a_ib_{k-i}=\sum_{i=0}^{N-1}a_i\underbrace{b_{k-i}}_{\substack{=0\text{ car}\\k-i>k-N\geqslant N}}+\sum_{i=N}^{k}\underbrace{a_i}_{=0}b_{k-i}=0$.

Théorème (Anneau $\mathbb{K}[X]$ et notation polynomiale) Le triplet $\left(\mathbb{K}[X],+,\times\right)$ est un anneau commutatif d'éléments neutres additif le polynôme nul 0 et multiplicatif le polynôme constant 1.

Pour tout $P=(a_k)_{k\in\mathbb{N}}\in\mathbb{K}[X]$: $P=\displaystyle\sum_{k=0}^{+\infty}a_kX^k$, écriture unique dans laquelle la somme est faussement infinie car P est une suite presque nulle. En outre, pour tout $\lambda\in\mathbb{K}$: $\lambda P=\displaystyle\sum_{k=0}^{+\infty}\lambda a_kX^k$.

Attention ! X N'EST PAS UN NOMBRE ! Ôtez-vous une fois pour toutes cette idée de la tête.

Démonstration Fixons une fois pour toutes $P=(a_k)_{k\in\mathbb{N}},Q=(b_k)_{k\in\mathbb{N}},R=(c_k)_{k\in\mathbb{N}}\in\mathbb{K}[X]$.

- **Commutativité de $\times$:** Pour tout $k\in\mathbb{N}$: $\displaystyle\sum_{i=0}^{k}a_ib_{k-i}\overset{j=k-i}{=}\sum_{j=0}^{k}b_ja_{k-j}$, donc $PQ=QP$.
- **Multiplication par un scalaire et élément neutre pour $\times$:** Soit $\lambda\in\mathbb{K}$. Pour tout $k\in\mathbb{N}$, le coefficient de degré k de $\lambda\times P=(\lambda,0,0,\ldots)\times(a_k)_{k\in\mathbb{N}}$ vaut $\lambda a_k+0.a_{k-1}+\ldots+0.a_0=\lambda a_k$, donc $\lambda P=(\lambda a_k)_{k\in\mathbb{N}}$. En particulier, $1\times P=P\times 1=P$.
- **Associativité de $\times$:** Pour tout $k\in\mathbb{N}$, le coefficient de degré k de $(PQ)R$ vaut :
$$\sum_{i=0}^{k}\left(\sum_{j=0}^{i}a_jb_{i-j}\right)c_{k-i}=\sum_{0\leqslant j\leqslant i\leqslant k}a_jb_{i-j}c_{k-i}=\sum_{j=0}^{k}a_j\left(\sum_{i=j}^{k}b_{i-j}c_{k-i}\right)\overset{l=i-j}{=}\sum_{j=0}^{k}a_j\left(\sum_{l=0}^{k-j}b_lc_{(k-j)-l}\right),$$
donc est égal au coefficient de degré k de $P(QR)$. Par conséquent, $(PQ)R=P(QR)$.
- **Distributivité de $\times$ sur $+$:** Pour tout $k\in\mathbb{N}$, le coefficient de degré k de $P(Q+R)$ vaut :
$$\sum_{i=0}^{k}a_i(b_{k-i}+c_{k-i})=\sum_{i=0}^{k}a_ib_{k-i}+\sum_{i=0}^{k}a_ic_{k-i},$$
donc est égal au coefficient de degré k de $(PQ)+(PR)$. Par conséquent, $P(Q+R)=(PQ)+(PR)$.
- **Notation polynomiale :** Sachant que $X=(0,1,0,0,\ldots)$, il n'est pas dur de montrer par récurrence que $X^k=(0,\ldots,0,1,0,0,\ldots)$ pour tout $k\in\mathbb{N}$, où le 1 apparaît au degré k. Ainsi, pour tout $P=(a_k)_{k\in\mathbb{N}}\in\mathbb{K}[X]$:
$$P=\sum_{k=0}^{+\infty}(0,\ldots,0,a_k,0,0,\ldots)=\sum_{k=0}^{+\infty}(a_k,0,0,\ldots)\times(0,\ldots,0,1,0,0,\ldots)=\sum_{k=0}^{+\infty}a_kX^k.$$

En dépit de la construction qui précède, abstenez-vous impérativement de penser que les polynômes sont des suites presque nulles. Nous les avons construits ainsi, mais d'autres constructions sont possibles. Soyez juste convaincus que les fonctions polynomiales sont une chose et les polynômes une autre.

Exemple Pour tous $a,b,n\in\mathbb{N}$: $\displaystyle\sum_{k=0}^{n}\binom{a}{k}\binom{b}{n-k}=\binom{a+b}{n}$ (*formule de Vandermonde*).

En particulier, $\displaystyle\sum_{k=0}^{n}\binom{n}{k}^2=\binom{2n}{n}$ pour $a=b=n$.

Nous avons déjà prouvé ces formules par double comptage au chapitre « Dénombrement ».

Démonstration $\displaystyle\sum_{k=0}^{a+b}\binom{a+b}{k}X^k=(X+1)^{a+b}=(X+1)^a(X+1)^b=\sum_{i=0}^{a}\binom{a}{i}X^i\times\sum_{j=0}^{b}\binom{b}{j}X^j.$

Or que vaut le coefficient de degré n des deux côtés de cette identité ?

Il vaut $\binom{a+b}{n}$ à gauche et $\displaystyle\sum_{i=0}^{n}\binom{a}{i}\binom{b}{n-i}$ à droite par définition du produit de deux polynômes.

Théorème (Degré d'une somme et d'un produit) Soient $P,Q\in\mathbb{K}[X]$.

(i) **Degré d'une somme :** $\deg(P+Q)\leqslant\max\{\deg(P),\deg(Q)\}$.

Cette inégalité est une égalité notamment quand $\deg(P)>\deg(Q)$ ou $\deg(Q)>\deg(P)$.

(ii) **Degré d'un produit :** $\deg(PQ)=\deg(P)+\deg(Q)$. En particulier, $\deg(\lambda P)=\deg(P)$ pour tout $\lambda\in\mathbb{K}^*$.

Démonstration Le résultat est évident si P ou Q est nul. Supposons-les non nuls et notons m le degré de P et n celui de Q. Introduisons aussi leurs coefficients : $P=\displaystyle\sum_{k=0}^{+\infty}a_kX^k$, $Q=\sum_{k=0}^{+\infty}b_kX^k$ et $PQ=\sum_{k=0}^{+\infty}c_kX^k$.

(i) $a_k+b_k=0$ pour tout $k>\max\{m,n\}$, donc $\deg(P+Q)\leqslant\max\{m,n\}=\max\{\deg(P),\deg(Q)\}$.

(ii) Pour tout $k>m+n$: $c_k=\displaystyle\sum_{i=0}^{k}a_ib_{k-i}=\sum_{i=0}^{m}a_i\underbrace{b_{k-i}}_{=0}+\sum_{i=m+1}^{k}\underbrace{a_i}_{=0}b_{k-i}=0$, donc $\deg(PQ)\leqslant m+n$.

Inversement : $c_{m+n}=\displaystyle\sum_{i=0}^{m-1}a_i\overbrace{b_{m+n-i}}^{=0}+a_mb_n+\sum_{i=m+1}^{m+n}\overbrace{a_i}^{=0}b_{m+n-i}=a_mb_n$ avec $a_m\neq0$ et $b_n\neq0$, donc $c_{m+n}\neq0$, donc $\deg(PQ)\geqslant m+n$.

Théorème (Intégrité de $\mathbb{K}[X]$) L'anneau $\mathbb{K}[X]$ est intègre : $\forall P,Q\in\mathbb{K}[X],\ \big(PQ=0\implies P=0\text{ ou }Q=0\big)$.

Démonstration Pour commencer, $\mathbb{K}[X]$ est un anneau non nul. Ensuite, pour tous $P,Q\in\mathbb{K}[X]$, si $PQ=0$, alors $\deg(P)+\deg(Q)=\deg(PQ)=-\infty$, donc $\deg(P)$ ou $\deg(Q)$ vaut $-\infty$, donc P ou Q est nul.

Ce résultat serait nettement plus difficile à prouver si on travaillait avec des fonctions polynomiales et non avec des polynômes. En effet, si $P(x)Q(x)=0$ pour tout $x\in\mathbb{R}$, alors en tout point l'une des fonctions P et Q s'annule, mais qui nous dit que l'une des deux s'annule tout le temps ? Rien a priori.

1.2 Évaluation polynomiale

Définition-théorème (Évaluation polynomiale, fonction polynomiale)

- **Évaluation :** On pose $P(\lambda)=\displaystyle\sum_{k=0}^{+\infty}a_k\lambda^k$ pour tous $P=\sum_{k=0}^{+\infty}a_kX^k\in\mathbb{K}[X]$ et $\lambda\in\mathbb{K}$.

  Le résultat $P(\lambda)$ obtenu est un élément de $\mathbb{K}$, et comme toujours, la somme est faussement infinie.
- **Fonction polynomiale :** Pour tout $P\in\mathbb{K}[X]$, la fonction $x\longmapsto P(x)$ de $\mathbb{K}$ dans $\mathbb{K}$ est appelée la *fonction polynomiale associée à* P. On la note souvent P par abus et parfois $\widetilde{P}$ quand on veut la distinguer du polynôme P.
- **Deux morphismes d'anneaux importants :** Pour tout $x\in\mathbb{K}$, l'application d'évaluation $P\longmapsto P(x)$ est un morphisme d'anneaux de $\mathbb{K}[X]$ dans $\mathbb{K}$.

  Également, l'application $P\longmapsto\widetilde{P}$ est un morphisme d'anneaux de $\mathbb{K}[X]$ dans l'anneau $\mathbb{K}^{\mathbb{K}}$.

Concrètement, dire que l'application $P\longmapsto P(x)$ est un morphisme d'anneaux, c'est dire qu'elle envoie le polynôme 1 sur le nombre 1, et surtout que pour tous $P,Q\in\mathbb{K}[X]$: $(P+Q)(x)=P(x)+Q(x)$ et $(PQ)(x)=P(x)Q(x)$.

Même chose avec l'autre morphisme : $\widetilde{P+Q}=\widetilde{P}+\widetilde{Q}$ et $\widetilde{PQ}=\widetilde{P}\,\widetilde{Q}$.

Attention ! X N'EST PAS UN NOMBRE ! On ne dit pas « Posons $X=1$ », mais « Évaluons en 1 ».

1.3 Composition, dérivation, conjugaison

Dans ce paragraphe, on apprend à composer les polynômes et à les dériver, mais les polynômes ne sont pas des fonctions !

Définition-théorème (Composition) On pose $P\circ Q=\sum_{k=0}^{+\infty}a_kQ^k$ pour tous $P=\sum_{k=0}^{+\infty}a_kX^k\in\mathbb{K}[X]$ et $Q\in\mathbb{K}[X]$.

(i) **Degré :** Si Q n'est pas constant, alors $\deg(P\circ Q)=\deg(P)\times\deg(Q)$.

(ii) **Lien avec la composition des fonctions :** $\widetilde{P\circ Q}=\widetilde{P}\circ\widetilde{Q}$. Le symbole $\circ$ a deux sens très différents ici !

Démonstration Pour (i), supposons Q non constant et posons $m=\deg(P)$. Par produit, $\deg\left(Q^k\right)=k\deg(Q)$ pour tout $k\in\mathbb{N}$, donc comme $\deg(Q)\geqslant 1$ et $a_m\neq 0$:

$$\deg\left(a_{m-1}Q^{m-1}+\ldots+a_1Q+a_0\right)\leqslant(m-1)\deg(Q)\underset{\text{Inégalité STRICTE !}}{<}m\deg(Q)=\deg\left(a_mQ^m\right).$$

donc : $\deg(P\circ Q)=\deg\left(a_mQ^m+\ldots+a_1Q+a_0\right)=\max\left\{\deg\left(a_mQ^m\right),\deg\left(a_{m-1}Q^{m-1}+\ldots+a_1Q+a_0\right)\right\}$, et enfin $\deg(P\circ Q)=\deg(P)\times\deg(Q)$. ∎

Définition-théorème (Dérivation) On pose $P'=\sum_{k=0}^{+\infty}ka_kX^{k-1}$ pour tout $P=\sum_{k=0}^{+\infty}a_kX^k\in\mathbb{K}[X]$.

Attention, on convient ici que $0X^{-1}=0$ pour $k=0$ — fausse apparition de X^{-1}, donc.

À partir de là, on définit les polynômes dérivés successifs de P en posant $P^{(0)}=P$ et $P^{(n+1)}=\left(P^{(n)}\right)'$ pour tout $n\in\mathbb{N}$.

Soient $P,Q\in\mathbb{K}[X]$, $\lambda\in\mathbb{K}$ et $n\in\mathbb{N}$.

(i) **Degré :** Si $n\leqslant\deg(P)$, alors $\deg\left(P^{(n)}\right)=\deg(P)-n$, et si au contraire $n>\deg(P)$, alors $P^{(n)}=0$.

(ii) **Opérations usuelles :** $(\lambda P+Q)'=\lambda P'+Q'$, $(PQ)'=P'Q+PQ'$ et $(P\circ Q)'=Q'\times P'\circ Q$.

(iii) **Lien avec la dérivation des fonctions :** $\widetilde{P'}=\widetilde{P}'$. Le symbole $'$ a deux sens très différents ici !

(iv) **Formule de Taylor polynomiale :** Pour tous $P\in\mathbb{K}[X]$ et $\lambda\in\mathbb{K}$: $P=\sum_{k=0}^{+\infty}\frac{P^{(k)}(\lambda)}{k!}(X-\lambda)^k$.

En particulier, le coefficient de degré k de P vaut $\frac{P^{(k)}(0)}{k!}$ pour tout $k\in\mathbb{N}$.

Le seuil $n=\deg(P)$ de l'assertion (i) doit être bien compris. Sur l'exemple du polynôme $P=X^n$, voilà ce que ça donne : $P'=nX^{n-1}$, $P''=n(n-1)X^{n-2}$, …, $P^{(n)}=n!$, puis $P^{(n+1)}=0$.

La formule de Taylor montre qu'on connaît tout d'un polynôme quand on le connaît « profondément » en un point, i.e. si on connaît la valeur de toutes ses dérivées en ce seul point.

Démonstration Introduisons les coefficients de P et Q : $P=\sum_{k=0}^{+\infty}a_kX^k$ et $Q=\sum_{k=0}^{+\infty}b_kX^k$.

(i) Posons $d=\deg(P)$. Si $d\leqslant 0$, alors $P'=0$. Si au contraire $d\geqslant 1$, alors $P'=\sum_{k=0}^{d}ka_kX^{k-1}$ avec $da_d\neq 0$, donc $\deg(P')=d-1$. Pour les dérivées suivantes, récurrence !

(iii) Montrons que $(PQ)'=P'Q+PQ'$. Soit $k\in\mathbb{N}$. Le coefficient de degré k de $(PQ)'$ vaut $(k+1)\sum_{i=0}^{k+1}a_ib_{k+1-i}$ et celui de $P'Q+PQ'$ vaut $\sum_{j=0}^{k}(j+1)a_{j+1}b_{k-j}+\sum_{i=0}^{k}a_i(k-i+1)b_{k-i+1}$. Ces coefficients sont égaux car :

$$\sum_{j=0}^{k}(j+1)a_{j+1}b_{k-j}+\sum_{i=0}^{k}a_i(k-i+1)b_{k-i+1}\overset{i=j+1}{=}\sum_{i=1}^{k+1}ia_ib_{k+1-i}+\sum_{i=0}^{k}a_i(k-i+1)b_{k-i+1}$$
$$=\sum_{i=0}^{k+1}ia_ib_{k+1-i}+\sum_{i=0}^{k+1}a_i(k-i+1)b_{k-i+1}=\sum_{i=0}^{k+1}\left(ia_ib_{k+1-i}+a_i(k-i+1)b_{k+1-i}\right)=(k+1)\sum_{i=0}^{k+1}a_ib_{k+1-i}.$$

Montrons que $(P\circ Q)'=Q'\times P'\circ Q$. Par récurrence à partir de la relation précédente, $\left(Q^k\right)'=kQ'Q^{k-1}$ pour tout $k\in\mathbb{N}$, donc $(P\circ Q)'=\sum_{k=0}^{+\infty}a_k\left(Q^k\right)'=Q'\sum_{k=0}^{+\infty}ka_kQ^{k-1}=Q'\times P'\circ Q$.

(iv) **Cas où $\lambda = 0$:** En notant $P = \sum_{i=0}^{+\infty} a_i X^i$, dérivons k fois pour tout $k \in \mathbb{N}$: $P^{(k)} = \sum_{i=k}^{+\infty} a_i \frac{i!}{(i-k)!} X^{i-k}$, puis évaluons en 0 : $P^{(k)}(0) = \underbrace{a_k}_{i=k} k!$. Aussitôt, $a_k = \frac{P^{(k)}(0)}{k!}$ donc $P = \sum_{k=0}^{+\infty} \frac{P^{(k)}(0)}{k!} X^k$.

Cas général : Posons $Q = P(X+\lambda)$. Pour tout $k \in \mathbb{N}$: $Q^{(k)} = P^{(k)}(X+\lambda)$, donc :

$$Q = \sum_{k=0}^{+\infty} \frac{Q^{(k)}(0)}{k!} X^k = \sum_{k=0}^{+\infty} \frac{P^{(k)}(\lambda)}{k!} X^k, \qquad \text{et on compose à droite par } X-\lambda \text{ pour terminer.}$$

Pour la conjugaison des polynômes, on suppose que $\mathbb{K} = \mathbb{C}$.

Définition-théorème **(Conjugaison)** On pose $\overline{P} = \sum_{k=0}^{+\infty} \overline{a_k} X^k$ pour tout $P = \sum_{k=0}^{+\infty} a_k X^k \in \mathbb{C}[X]$.

Pour tous $P, Q \in \mathbb{C}[X]$: $\overline{\lambda P + Q} = \overline{\lambda}\,\overline{P} + \overline{Q}$, $\overline{PQ} = \overline{P}\,\overline{Q}$ et $\overline{P \circ Q} = \overline{P} \circ \overline{Q}$.

Seuls les coefficients sont vraiment conjugués. En particulier, $\overline{X} = X$.

2 Divisibilité et division euclidienne dans $\mathbb{K}[X]$

Si on veut une définition simple et générale, l'arithmétique est la théorie de la divisibilité dans les anneaux. La notion de divisibilité peut être définie dans n'importe quel anneau, et pas seulement dans $\mathbb{Z}$, et les anneaux sont plus ou moins semblables ou dissemblables selon les propriétés de leur relation de divisibilité. Nous nous contenterons quant à nous des anneaux $\mathbb{Z}$ et $\mathbb{K}[X]$ et verrons peu à peu qu'ils sont quasiment identiques d'un point de vue arithmétique. À quoi cela tient-il ? Ces deux anneaux partagent un même théorème de la division euclidienne.

Définition-théorème **(Relation de divisibilité)** Soient $A, B \in \mathbb{K}[X]$. On dit que A *divise* B, ce qu'on note $A \mid B$, si $B = AP$ pour un certain $P \in \mathbb{K}[X]$. On dit aussi que A est un *diviseur de* B, que B est *divisible par* A ou que B est un *multiple de* A.

- **Relation d'ordre :** La relation de divisibilité $\mid$ est réflexive et transitive sur $\mathbb{K}[X]$, mais pas antisymétrique car pour tous $A, B \in \mathbb{K}[X]$:

  $$A \mid B \quad \text{et} \quad B \mid A \iff \exists \lambda \in \mathbb{K}^*,\ A = \lambda B.$$

  Le cas échéant, on dit que A et B sont *associés*.

  Cela dit, la relation de divisibilité $\mid$ est une relation sur l'ensemble des polynômes unitaires de $\mathbb{K}[X]$.
- **Combinaisons linéaires et produits :** Soient $A, B, C, D \in \mathbb{K}[X]$.
  - Si $D \mid A$ et $D \mid B$, alors $D \mid (AU + BV)$ pour tous $U, V \in \mathbb{K}[X]$.
  - Si $A \mid B$ et $C \mid D$, alors $AC \mid BD$, et en particulier, $A^k \mid B^k$ pour tout $k \in \mathbb{N}$.

Démonstration Pour le défaut d'antisymétrie, si $A = \lambda B$ pour un certain $\lambda \in \mathbb{K}^*$, alors $B = \frac{1}{\lambda} A$, donc $A \mid B$ et $B \mid A$. Réciproquement, faisons l'hypothèse que $A \mid B$ et $B \mid A$. Ainsi, $A = BP$ et $B = AQ$ pour certains $P, Q \in \mathbb{K}[X]$, donc $B = B(PQ)$.

- Si $B = 0$, alors $A = BP = 0$, donc $A = \lambda B$ pour $\lambda = 1$.
- Si $B \neq 0$, alors $PQ = 1$ par intégrité de $\mathbb{K}[X]$, donc $\deg(P) + \deg(Q) = 0$, donc $\deg(P) = \deg(Q) = 0$ car $\deg(P)$ et $\deg(Q)$ appartiennent à $\mathbb{N} \sqcup \{-\infty\}$. Ainsi, P est une constante non nulle λ et $A = \lambda B$.

Nous pratiquons la division euclidienne des polynômes depuis la fin du chapitre « Techniques élémentaires de calcul intégral », mais nous n'avions rien démontré à ce moment-là.

Théorème **(Théorème de la division euclidienne)** Soient $A, B \in \mathbb{K}[X]$ avec $B \neq 0$. Il existe un et un seul couple $(Q, R) \in \mathbb{K}[X] \times \mathbb{K}[X]$ pour lequel $A = BQ + R$ et $\deg(R) < \deg(B)$. On appelle A le *dividende* de la division euclidienne de A par B, B son *diviseur*, Q son *quotient* et R son *reste*.

Démonstration

- **Existence :** Notons b le degré de B et $\beta \neq 0$ son coefficient dominant. Si B divise A, alors $A = BQ$ pour un certain $Q \in \mathbb{K}[X]$ et on peut poser $R = 0$. Supposons maintenant que B ne divise pas A. L'ensemble $\mathscr{D} = \left\{\deg(A-BK) \mid K \in \mathbb{K}[X]\right\}$ est une partie non vide de $\mathbb{N}$ — valeur $-\infty$ exclue par hypothèse — donc possède un plus petit élément r. Notons $Q \in \mathbb{K}[X]$ un polynôme pour lequel $\deg(A-BQ) = r$, puis posons $R = A-BQ$ et notons ρ le coefficient dominant de R. Est-il vrai que $\deg(R) < \deg(B)$?

  Supposons par l'absurde que $r \geqslant b$. Dans ce cas, $\deg\left(R - \dfrac{\rho}{\beta} X^{r-b} B\right) < r$ car la soustraction de $\dfrac{\rho}{\beta} X^{r-b} B$ tue le terme dominant ρX^r de R. Pourtant, $\deg\left(R - \dfrac{\rho}{\beta} X^{r-b} B\right) = \deg(A-BK) \in \mathscr{D}$ pour $K = Q + \dfrac{\rho}{\beta} X^{r-b}$, donc la minimalité de r est contredite et $r < b$ comme voulu.
- **Unicité :** Soient (Q_1, R_1) et (Q_2, R_2) deux couples de division euclidienne de A par B. Si $Q_1 \neq Q_2$, alors $\deg(Q_1 - Q_2) \geqslant 0$, donc :

  $$\deg(B) \leqslant \deg(B) + \deg(Q_1 - Q_2) \leqslant \deg\big(B(Q_1 - Q_2)\big) = \deg(R_1 - R_2) \leqslant \max\left\{\deg(R_1), \deg(R_2)\right\} < \deg(B)$$

  — contradiction. Ainsi $Q_1 = Q_2$, donc $R_1 = A - BQ_1 = A - BQ_2 = R_2$. ∎

Telle qu'elle est définie, la notion de divisibilité aurait pu dépendre du corps $\mathbb{K}$, mais en fait non.

Théorème (Divisibilité dans $\mathbb{C}[X]$ de polynômes réels) Soient $A, B \in \mathbb{R}[X]$ — à coefficients réels, donc.

$$A \text{ divise } B \text{ dans } \mathbb{C}[X] \text{ si et seulement si } A \text{ divise } B \text{ dans } \mathbb{R}[X].$$

Démonstration On peut effectuer a priori deux divisions euclidiennes de B par A, une dans $\mathbb{R}[X]$ et une dans $\mathbb{C}[X]$. Cela dit, la division euclidienne dans $\mathbb{R}[X]$ est aussi une division euclidienne dans $\mathbb{C}[X]$, donc par unicité de la division euclidienne dans $\mathbb{C}[X]$, les deux divisions euclidiennes coïncident et le résultat en découle. ∎

À partir de ce théorème dans $\mathbb{K}[X]$ identique à celui que nous avons démontré dans $\mathbb{Z}$, deux chemins naturels d'étude de $\mathbb{K}[X]$ se présentent à nous. Nous pourrions suivre la piste arithmétique des PGCD, relations de Bézout, lemme d'Euclide et prouver que les polynômes possèdent une et une seule *factorisation irréductible*, cet analogue polynomial de la factorisation première dans $\mathbb{Z}$. Le long de ce chemin, les preuves sont essentiellement les mêmes que dans $\mathbb{Z}$, mais nous verrons cela plus tard au chapitre « Arithmétique des polynômes et fractions rationnelles ». L'autre chemin que nous suivons à présent, c'est celui des racines, très spécifique aux polynômes et qui nous permet de contourner l'arithmétique encore quelques temps.

3 Racines

3.1 Racines et multiplicités

Théorème (Division euclidienne par $X - \lambda$) Soient $\lambda \in \mathbb{K}$ et $P \in \mathbb{K}[X]$. Le reste de la division euclidienne de P par $X - \lambda$ est $P(\lambda)$.

Démonstration La division de P par $X-\lambda$ s'écrit $P = (X-\lambda)Q + a$ pour certains $Q \in \mathbb{K}[X]$ et $a \in \mathbb{K}$. L'application d'évaluation $T \longmapsto T(\lambda)$ étant un morphisme d'anneaux : $P(\lambda) = (\lambda - \lambda)Q(\lambda) + a = a$. ∎

De ce résultat découle la double définition suivante.

Définition (Racine) Soient $P \in \mathbb{K}[X]$ et $\lambda \in \mathbb{K}$. On dit que λ est une *racine de P (dans $\mathbb{K}$)* si l'une des deux assertions équivalentes suivantes est vraie : $P(\lambda) = 0$ ou bien : P est divisible par $X - \lambda$.

✖ **Attention !** La précision « racine DANS $\mathbb{K}$ » n'est pas superflue. Le polynôme $X^2 + 1$ n'a pas de racine dans $\mathbb{R}$, mais il en a deux dans $\mathbb{C}$, à savoir i et $-$i.

Via la notion de racine, on ramène souvent les problèmes de divisibilité à des problèmes d'évaluation — et vice versa.

Exemple Pour tout $n \in \mathbb{N}$, le reste de la division euclidienne de X^n par X^2-3X+2 vaut $(2^n-1)X-(2^n-2)$.

Démonstration Soit $n \in \mathbb{N}$. La division euclidienne de X^n par X^2-3X+2 s'écrit $X^n = (X-1)(X-2)Q+aX+b$ pour certains $Q \in \mathbb{R}[X]$ et $a, b \in \mathbb{R}$. Évaluons en 1 : $1 = a+b$, puis en 2 : $2^n = 2a+b$. Après calcul : $a = 2^n-1$ et $b = 2-2^n$.

Définition-théorème (Multiplicité) Soient $P \in \mathbb{K}[X]$ non nul et $\lambda \in \mathbb{K}$.

- **Définition :** L'ensemble $\left\{k \in \mathbb{N} \mid (X-\lambda)^k \text{ divise } P\right\}$ possède un plus grand élément m appelé la *multiplicité de λ dans P*. Pour résumer, on dit souvent que m est la plus grande puissance de $X-\lambda$ qui divise P.

  En particulier, dire que λ est de multiplicité 0 dans P, c'est dire que λ n'est pas racine de P. Une racine est dite *simple* si elle est de multiplicité 1, *double* si elle est de multiplicité 2, etc.

  Plus concrètement, m est caractérisé par les deux propositions équivalentes suivantes :
  - P est divisible par $(X-\lambda)^m$ mais pas par $(X-\lambda)^{m+1}$.
  - Il existe $Q \in \mathbb{K}[X]$ pour lequel $P = (X-\lambda)^m Q$ et $Q(\lambda) \neq 0$.

(ii) **Caractérisation de la multiplicité par les dérivées successives :** Soit $m \in \mathbb{N}$.

λ est de multiplicité m dans P si et seulement si $P^{(i)}(\lambda) = 0$ pour tout $i \in [\![0, m-1]\!]$ mais $P^{(m)}(\lambda) \neq 0$.

Il en découle que si λ est de multiplicité $m \geqslant 1$ dans P, λ est de multiplicité $m-1$ dans P'.

Démonstration

(i) Pour montrer que l'ensemble $\mathscr{M} = \left\{k \in \mathbb{N} \mid (X-\lambda)^k \text{ divise } P\right\}$ possède un plus grand élément, montrons que c'est une partie non vide majorée de $\mathbb{N}$. Or déjà, $\mathscr{M}$ contient 0. Montrons maintenant que $\deg(P)$ majore $\mathscr{M}$. Pour tout $k \in \mathscr{M}$, $P = (X-\lambda)^k Q$ pour un certain $Q \in \mathbb{K}[X]$, or $Q \neq 0$ car $P \neq 0$, donc $\deg(Q) \geqslant 0$, et enfin $k \leqslant \deg(Q)+k = \deg\big((X-\lambda)^k Q\big) = \deg(P)$.

(ii) Montrons seulement l'équivalence, le résultat sur la dérivation en découle directement. L'air de rien, la formule de Taylor nous fournit facilement la division euclidienne de P par $(X-\lambda)^m$:

$$P \overset{\text{Taylor}}{=} \sum_{i=0}^{+\infty} \frac{P^{(i)}(\lambda)}{i!}(X-\lambda)^i = (X-\lambda)^m \sum_{i=m}^{+\infty} \frac{P^{(i)}(\lambda)}{i!}(X-\lambda)^{i-m} + \underbrace{\sum_{i=0}^{m-1} \frac{P^{(i)}(\lambda)}{i!}(X-\lambda)^i}_{\text{Degré strictement inférieur à } m}.$$

On en tire les équivalences suivantes :

$$\begin{aligned}
\lambda \text{ est de multiplicité au moins } m \text{ dans } P &\iff (X-\lambda)^m \text{ divise } P \iff \sum_{i=0}^{m-1} \frac{P^{(i)}(\lambda)}{i!}(X-\lambda)^i = 0 \\
&\iff \sum_{i=0}^{m-1} \frac{P^{(i)}(\lambda)}{i!} X^i = 0 \quad \text{après composition à droite par } X+\lambda \\
&\iff \forall i \in [\![0, m-1]\!], \quad P^{(i)}(\lambda) = 0.
\end{aligned}$$

A fortiori : λ est de multiplicité au moins $m+1$ dans $P \iff \forall i \in [\![0, m]\!], \quad P^{(i)}(\lambda) = 0$. Il en découle que λ est de multiplicité exactement m dans P si et seulement si $P^{(i)}(\lambda) = 0$ pour tout $i \in [\![0, m-1]\!]$ mais $P^{(m)}(\lambda) \neq 0$.

Exemple 1 est de multiplicité 2 dans le polynôme $P = X^4+3X^3-3X^2-7X+6$, car d'abord $P(1) = 1+3-3-7+6 = 0$, ensuite $P' = 4X^3+9X^2-6X-7$ donc $P'(1) = 0$, et enfin $P'' = 12X^2+18X-6$ donc $P''(1) = 24 \neq 0$.

Soient $A, B \in \mathbb{K}[X]$ avec $B \neq 0$. On l'a vu, le reste de la division euclidienne de A par B peut être calculé à partir des racines de B. Pour $B = (X-2)^3(X+3)$, cette division s'écrit $A = (X-2)^3(X+3)Q + aX^3+bX^2+cX+d$ pour certains $Q \in \mathbb{R}[X]$ et $a, b, c, d \in \mathbb{R}$. On n'obtient que deux équations linéaires en évaluant en 2 et -3 alors qu'il nous en faut quatre pour calculer a, b, c et d, mais la multiplicité de 2 dans B cache deux équations supplémentaires. Pour les faire surgir, on dérive ! Le réel 2 est de multiplicité au moins 3 dans $BQ = (X-2)^3(X+3)Q$, donc au moins 2 dans $(BQ)'$ et au moins 1 dans $(BQ)''$. Deux nouvelles équations en découlent : $A'(2) = 12a+4b+c$ et $A''(2) = 12a+2b$, et on conclut en résolvant un système linéaire 4×4.

Exemple Pour tout $n \in \mathbb{N}^*$, le reste de la division euclidienne de X^n par $X(X-1)^2$ vaut $(n-1)X^2-(n-2)X$.

Démonstration La division euclidienne étudiée s'écrit $X^n = X(X-1)^2Q+aX^2+bX+c$ pour certains $Q \in \mathbb{R}[X]$ et $a, b, c \in \mathbb{R}$. Évaluons en 0 : $c = 0$, puis en 1 : $a+b+c = 1$. Ainsi $a+b = 1$ et $c = 0$, mais il nous manque une équation. Dérivons puis évaluons en 1, sachant que 1 est de multiplicité 2 dans $X(X-1)^2Q$: $2a+b = n$. Après calcul : $a = n-1$, $b = 2-n$ et $c = 0$.

Théorème (Racines complexes d'un polynôme réel) Soient $P \in \mathbb{R}[X]$ — à coefficients réels, donc — et $\lambda \in \mathbb{C}$. Alors λ et $\overline{\lambda}$ ont la même multiplicité dans P.

Démonstration L'application $T \longmapsto \overline{T}$ est un automorphisme d'anneau de $\mathbb{C}[X]$, donc pour tout $k \in \mathbb{N}$:

$$\begin{aligned}(X-\lambda)^k \text{ divise } P &\iff \exists Q \in \mathbb{C}[X],\quad P = (X-\lambda)^k Q &&\iff \exists Q \in \mathbb{C}[X],\quad \overline{P} = \left(X-\overline{\lambda}\right)^k \overline{Q}\\ &\iff \exists Q \in \mathbb{C}[X],\quad \overline{P} = \left(X-\overline{\lambda}\right)^k \overline{Q} &&\iff \left(X-\overline{\lambda}\right)^k \text{ divise } \overline{P},\end{aligned}$$

donc en effet, λ et $\overline{\lambda}$ ont la même multiplicité dans P. ∎

3.2 Nombre maximal de racines

On dit souvent qu'on factorise *par les racines*, mais en réalité, ce ne sont pas les racines λ qu'on met en facteur, ce sont les polynômes $X-\lambda$.

Théorème **(Factorisation par les racines)** Soient $P \in \mathbb{K}[X]$ non nul et $\lambda_1, \dots, \lambda_r$ des racines distinctes de P de multiplicités respectives $m_1, \dots, m_r$. Alors $(X-\lambda_1)^{m_1} \dots (X-\lambda_r)^{m_r}$ divise P. En particulier, $m_1 + \dots + m_r \leqslant \deg(P)$.

Démonstration Montrons par récurrence que pour tout $k \in [\![1, r]\!]$, $(X-\lambda_1)^{m_1} \dots (X-\lambda_k)^{m_k}$ divise P.

Initialisation : λ_1 est racine de P de multiplicité m_1, donc $(X-\lambda_1)^{m_1}$ divise P.

Hérédité : Soit $k \in [\![1, r-1]\!]$. Faisons l'hypothèse que $(X-\lambda_1)^{m_1} \dots (X-\lambda_k)^{m_k}$ divise P.

— Dans ces conditions, $P = (X-\lambda_1)^{m_1} \dots (X-\lambda_k)^{m_k} A$ pour un certain $A \in \mathbb{K}[X]$.

— Ensuite, si on note α la multiplicité de λ_{k+1} dans A, $A = (X-\lambda_{k+1})^\alpha B$ pour un certain $B \in \mathbb{K}[X]$ avec $B(\lambda_{k+1}) \neq 0$. Et comme $(X-\lambda_{k+1})^\alpha$ divise A, il divise aussi P, donc $\alpha \leqslant m_{k+1}$.

— Enfin, $P = (X-\lambda_{k+1})^{m_{k+1}} C$ pour un certain $C \in \mathbb{K}[X]$ avec $C(\lambda_{k+1}) \neq 0$.

Il découle de ces trois points que : $(X-\lambda_1)^{m_1} \dots (X-\lambda_k)^{m_k} (X-\lambda_{k+1})^\alpha B = (X-\lambda_{k+1})^{m_{k+1}} C$. Divisons par $(X-\lambda_{k+1})^\alpha$ grâce à l'intégrité de $\mathbb{K}[X]$: $(X-\lambda_1)^{m_1} \dots (X-\lambda_k)^{m_k} B = (X-\lambda_{k+1})^{m_{k+1}-\alpha} C$. Or le polynôme de gauche n'admet pas λ_{k+1} pour racine, donc celui de droite non plus, donc $\alpha = m_{k+1}$. Conclusion : $(X-\lambda_{k+1})^{m_{k+1}}$ divise A, donc $(X-\lambda_1)^{m_1} \dots (X-\lambda_{k+1})^{m_{k+1}}$ divise P. ∎

Exemple À quelle condition nécessaire et suffisante sur $n \in \mathbb{N}$ le polynôme X^2+1 divise-t-il X^n+1 ? Réponse : $n \equiv 2\ [4]$.

Démonstration

$$\begin{aligned}X^2+1 \text{ divise } X^n+1 &\iff \mathrm{i} \text{ et } -\mathrm{i} \text{ sont racines de } X^n+1\\ &\iff \mathrm{i} \text{ est racine de } X^n+1 \text{ car } X^n+1 \text{ est à coefficients réels} \iff \mathrm{i}^n+1=0\\ &\iff \mathrm{e}^{\frac{\mathrm{i}n\pi}{2}} = \mathrm{e}^{\mathrm{i}\pi} \iff \frac{n\pi}{2} \equiv \pi\ [2\pi] \iff n \equiv 2\ [4].\end{aligned}$$

Dans cette chaîne d'équivalences, le théorème de factorisation par les racines justifie l'implication :

$$\mathrm{i} \text{ et } -\mathrm{i} \text{ sont racines de } X^n+1 \implies X^2+1 \text{ divise } X^n+1.$$

Exemple Le polynôme $(X-1)^4 X^2 (X+2)$ possède en tout trois racines distinctes, à savoir 1 de multiplicité 4, 0 qui est double et -2 qui est simple. On dit en revanche qu'il possède $7 = 4+2+1$ racines comptées avec multiplicité.

Théorème **(Rigidité de $\mathbb{K}[X]$)** Un polynôme NON NUL $P \in \mathbb{K}[X]$ possède au plus $\deg(P)$ racines comptées avec multiplicité. En particulier, seul le polynôme nul possède une infinité de racines.

Attention ! En dépit des apparences, ce théorème est l'un des plus importants du chapitre !

Un polynôme de degré n ne possède pas forcément n racines comptées avec multiplicité. Nous verrons que c'est le cas si $\mathbb{K} = \mathbb{C}$, mais pas si $\mathbb{K} = \mathbb{R}$. Par exemple, le polynôme réel X^2+1 est de degré 2 mais n'a pas de racine réelle.

Exemple Soit $P \in \mathbb{C}[X]$. On suppose que $P(n) = n^3 - n^2 + 1$ pour tout $n \in \mathbb{N}$. Alors $P = X^3 - X^2 + 1$, donc après coup, $P(z) = z^3 - z^2 + 1$ pour tout $z \in \mathbb{C}$!

Démonstration On connaît les valeurs et même une expression polynomiale de P sur $\mathbb{N}$, mais un polynôme n'est pas une fonction. Les coefficients de l'expression de P sur $\mathbb{N}$ sont-ils les coefficients de P tout court ? La réponse est oui par rigidité. Par hypothèse, le polynôme $P - X^3 + X^2 - 1$ admet tout entier naturel pour racine, il possède donc une infinité de racines, donc il est nul par rigidité. En d'autres termes, $P = X^3 - X^2 + 1$.

On le voit bien sur cet exemple, la rigidité de $\mathbb{K}[X]$ est un principe de « dés-évaluation » — le terme n'existe pas, je l'invente. Évaluer, c'est passer d'une égalité polynomiale à une égalité de nombres. Dés-évaluer, c'est faire le contraire et remonter d'une collection d'égalités de nombres à une égalité polynomiale. Concrètement, quand on connaît certaines valeurs d'un polynôme P, il est souvent fructueux de les interpréter en termes de racines d'un autre polynôme Q. Si Q a trop de racines, il est nul par rigidité et on en tire des informations sur P.

Exemple La fonction $x \longmapsto \sqrt[3]{x^2+1}$ n'est pas polynomiale sur $\mathbb{R}$.

Démonstration Supposons par l'absurde qu'il existe un polynôme $P \in \mathbb{R}[X]$ pour lequel $P(x) = \sqrt[3]{x^2+1}$ pour tout $x \in \mathbb{R}$. Le polynôme $P^3 - X^2 - 1$ possède alors une infinité de racines, en l'occurrence tous les réels, il est donc nul par rigidité, et ainsi $P^3 = X^2 + 1$. En particulier, $3\deg(P) = 2$ donc $\deg(P) = \dfrac{2}{3}$ — contradiction.

Exemple Soit $P \in \mathbb{R}[X]$. On suppose que P est de degré n entier et que $P(k) = \dfrac{1}{k}$ pour tout $k \in [\![1, n+1]\!]$.
Dans ces conditions : $P(-1) = n+1$.

Démonstration

- **Analyse des hypothèses :** Le polynôme $XP(X)-1$ admet $1, 2, \dots, n+1$ pour racines, soit déjà $n+1$ racines distinctes. Cela dit, il est justement de degré $n+1$, donc $XP(X)-1 = \lambda \prod_{k=1}^{n+1}(X-k)$ pour un certain $\lambda \in \mathbb{R}^*$. Évaluons en 0 : $-1 = \lambda \prod_{k=1}^{n+1}(-k)$, i.e. $\lambda = \dfrac{(-1)^n}{(n+1)!}$. Conclusion : $XP(X) = 1 + \dfrac{(-1)^n}{(n+1)!}\prod_{k=1}^{n+1}(X-k)$.
- **Calcul de $P(-1)$:** Évaluons ce résultat en -1 :
$$-P(-1) = 1 + \frac{(-1)^n}{(n+1)!}\prod_{k=1}^{n+1}\big(-(k+1)\big) = 1 + \frac{(-1)^n}{(n+1)!} \times (-1)^{n+1}(n+2)! = 1-(n+2), \qquad \text{donc } P(-1) = n+1.$$

Théorème (Identification polynôme/fonction polynomiale) Le morphisme d'anneaux $P \longmapsto \widetilde{P}$ est injectif de $\mathbb{K}[X]$ dans $\mathbb{K}^{\mathbb{K}}$. En d'autres termes, pour tous $P, Q \in \mathbb{K}[X]$, si les fonctions polynomiales $\widetilde{P}$ et $\widetilde{Q}$ sont égales, les polynômes P et Q eux-mêmes le sont, i.e. ont les mêmes coefficients.

Démonstration Montrons que $\operatorname{Ker} f = \{0\}$. Soit $P \in \operatorname{Ker} f$. La fonction $f(P) = \widetilde{P}$ est identiquement nulle sur $\mathbb{K}$, donc P possède une infinité de racines car ici $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$, donc $P = 0$ par rigidité.

3.3 Polynômes scindés et théorème de d'Alembert-Gauss

Définition (Polynôme scindé) Soit $P \in \mathbb{K}[X]$. On dit que P est *scindé* (*sur* $\mathbb{K}$) s'il n'est pas constant et possède exactement $\deg(P)$ racines (dans $\mathbb{K}$) comptées avec multiplicité.

Il est équivalent d'exiger que P puisse être écrit sous la forme $P = a(X-\lambda_1)^{m_1}\dots(X-\lambda_r)^{m_r}$ avec a le coefficient dominant de P, $\lambda_1, \dots, \lambda_r$ ses racines distinctes dans $\mathbb{K}$ et $m_1, \dots, m_r$ leurs multiplicités respectives.

Forme scindée = **3 informations** (racines, multiplicités, coefficient dominant)

Attention ! La précision « scindé **sur** $\mathbb{K}$ » n'est pas superflue car un polynôme peut avoir des racines complexes mais aucune réelle. Par exemple, $X^2+1 = (X+\mathrm{i})(X-\mathrm{i})$ est scindé sur $\mathbb{C}$, mais pas sur $\mathbb{R}$.

Exemple Le polynôme $X^n - 1$ est scindé sur $\mathbb{C}$ pour tout $n \in \mathbb{N}^*$: $X^n - 1 = \prod_{\omega \in \mathbb{U}_n}(X-\omega) = \prod_{k=0}^{n-1}\left(X - \mathrm{e}^{\frac{2\mathrm{i}k\pi}{n}}\right)$.

Démonstration X^n-1 n'est pas constant et admet les n éléments de $\mathbb{U}_n$ pour racines distinctes. Cela dit, X^n-1 possède au plus n racines comptées avec multiplicité pour une raison de degré, donc forcément, il en possède exactement n, donc est scindé sur $\mathbb{C}$.

Exemple Le polynôme X^3+27 est scindé sur $\mathbb{C}$ et sa forme scindée est $X^3+27 = (X-3)\left(X-3\mathrm{e}^{\frac{\mathrm{i}\pi}{3}}\right)\left(X-3\mathrm{e}^{-\frac{\mathrm{i}\pi}{3}}\right)$.

Démonstration Pour tout $r \in \mathbb{C}$:
$$r^3+27 = 0 \iff r^3 = -27 = \left(3\mathrm{e}^{\frac{\mathrm{i}\pi}{3}}\right)^3 \iff \exists k \in [\![0,2]\!], \quad r = 3\mathrm{e}^{\frac{\mathrm{i}\pi}{3}+\frac{2\mathrm{i}k\pi}{3}}.$$
Les racines complexes de X^3+27 sont ainsi $3\mathrm{e}^{\frac{\mathrm{i}\pi}{3}}$ pour $k=0$, -3 pour $k=1$ et $3\mathrm{e}^{-\frac{\mathrm{i}\pi}{3}}$ pour $k=2$. Ces trois racines sont distinctes et X^3+27 est de degré 3, donc scindé sur $\mathbb{C}$ à racines simples, et par ailleurs unitaire.

Mais finalement, tout polynôme possède-t-il une racine ? Question essentielle à laquelle nous n'avons pas encore répondu. La réponse affirmative suivante est un théorème majeur de l'algèbre.

Théorème **(Théorème de d'Alembert-Gauss)** Tout polynôme non constant de $\mathbb{C}[X]$ possède une racine COMPLEXE. A fortiori, tout polynôme non constant de $\mathbb{C}[X]$ est scindé sur $\mathbb{C}$.

Démonstration Nous démontrerons la première partie du théorème au chapitre « Topologie de $\mathbb{R}$ et $\mathbb{C}$ ». Le caractère scindé de tout polynôme non constant de $\mathbb{C}[X]$ en découle aisément par récurrence. ∎

Attention ! Le théorème est faux sur $\mathbb{R}$ — un polynôme non constant de $\mathbb{R}[X]$ peut ne pas avoir de racine RÉELLE, par exemple le polynôme X^2+1.

Les multiplicités sont aux polynômes ce que les valuations p-adiques sont aux entiers. Le critère de divisibilité qui suit se démontre ainsi dans $\mathbb{C}[X]$ comme son analogue dans $\mathbb{Z}$.

Théorème **(Divisibilité dans $\mathbb{C}[X]$ et racines)** Soient $A,B\in\mathbb{C}[X]$ non nuls. Alors A divise B si et seulement si pour tout $\lambda\in\mathbb{C}$, la multiplicité de λ dans A est inférieure à sa multiplicité dans B.

Ce théorème ramène toute question de divisibilité à une comparaison de multiplicités, mais une telle réduction n'est possible que dans $\mathbb{C}[X]$ car grâce au théorème de d'Alembert-Gauss, on connaît tout d'un polynôme complexe quand on connaît son coefficient dominant, ses racines et leurs multiplicités. Ici, les coefficients dominants ne servent à rien car la relation de divisibilité n'y est pas sensible.

3.4 Relations coefficients-racines

On travaille dans ce paragraphe avec des polynômes non constants de $\mathbb{C}[X]$, donc avec des polynômes scindés sur $\mathbb{C}$ d'après le théorème de d'Alembert-Gauss.

- **Polynômes de degré 2 :** Soit $P=a_2X^2+a_1X+a_0\in\mathbb{C}[X]$ de racines λ_1 et λ_2 comptées avec multiplicité. Alors : $P=a_2(X-\lambda_1)(X-\lambda_2)=a_2X^2-a_2(\lambda_1+\lambda_2)X+a_2\lambda_1\lambda_2$, donc après identification : $\lambda_1+\lambda_2=-\dfrac{a_1}{a_2}$ et $\lambda_1\lambda_2=\dfrac{a_0}{a_2}$.
- **Polynômes de degré 3 :** Soit $P=a_3X^3+a_2X^2+a_1X+a_0\in\mathbb{C}[X]$ de racines $\lambda_1,\lambda_2,\lambda_3$ comptées avec multiplicité. Alors : $P=a_3(X-\lambda_1)(X-\lambda_2)(X-\lambda_3)=a_3X^3-a_3\big(\lambda_1+\lambda_2+\lambda_3\big)X^2+a_3\big(\lambda_1\lambda_2+\lambda_1\lambda_3+\lambda_2\lambda_3\big)X-a_3\lambda_1\lambda_2\lambda_3$, donc après identification : $\lambda_1+\lambda_2+\lambda_3=-\dfrac{a_2}{a_3}$, $\lambda_1\lambda_2\lambda_3=-\dfrac{a_0}{a_3}$ et $\lambda_1\lambda_2+\lambda_1\lambda_3+\lambda_2\lambda_3=\dfrac{a_1}{a_3}$.

Le résultat qui suit généralise les calculs précédents par « simple » développement du produit $(X-\lambda_1)\dots(X-\lambda_n)$.

Théorème **(Relations coefficients-racines)** Soit $P=a_nX^n+\dots+a_1X+a_0\in\mathbb{C}[X]$ de degré $n\geqslant 1$ et de racines $\lambda_1,\dots,\lambda_n$ comptées avec multiplicité. Si on pose $\sigma_k=\displaystyle\sum_{1\leqslant i_1<\dots<i_k\leqslant n}\lambda_{i_1}\dots\lambda_{i_k}$ pour tout $k\in[\![1,n]\!]$, alors :

$$P=a_n\prod_{i=1}^{n}(X-\lambda_i)=a_n\Big(X^n-\sigma_1X^{n-1}+\sigma_2X^{n-2}-\sigma_3X^{n-3}+\dots+(-1)^n\sigma_n\Big).$$

Démonstration En développant $(X-\lambda_1)\dots(X-\lambda_n)$, on obtient 2^n termes dont chacun repose sur un choix de X ou $-\lambda_1$, puis X ou $-\lambda_2$, …, et enfin X ou $-\lambda_n$. Un terme de degré $k\in[\![1,n]\!]$ apparaît chaque fois qu'on choisit k fois l'indéterminée X et $n-k$ coefficients $-\lambda_i$, donc le coefficient de degré k de $(X-\lambda_1)\dots(X-\lambda_n)$ est la somme des mini-contributions $(-1)^k\displaystyle\sum_{1\leqslant i_1<\dots<i_k\leqslant n}\lambda_{i_1}\dots\lambda_{i_k}$. Le résultat en découle. ∎

La relation obtenue n'exprime pas les racines $\lambda_1,\dots,\lambda_n$ de P en fonction de ses coefficients $a_0,\dots,a_n$, mais par identification, nous pouvons en tirer σ_k en fonction de $a_0,\dots,a_n$ pour tout $k\in[\![1,n]\!]$: $\sigma_k=(-1)^k\dfrac{a_{n-k}}{a_n}$. Vues comme fonctions de $\lambda_1,\dots,\lambda_n$, les expressions $\sigma_1,\dots,\sigma_n$ sont appelées les *fonctions symétriques élémentaires de* $\lambda_1,\dots,\lambda_n$ — symétriques parce qu'elles ne dépendent pas de l'ordre dans lequel on a rangé $\lambda_1,\dots,\lambda_n$. Deux d'entre elles sont particulièrement naturelles :

$$\sigma_1=\sum_{k=1}^{n}\lambda_k\quad\text{(somme des racines)}\qquad\text{et}\qquad\sigma_n=\prod_{k=1}^{n}\lambda_k\quad\text{(produit des racines)}.$$

Mais pour que tout soit bien clair, détaillons σ_1, σ_2, σ_3 et σ_4 dans le cas où $n=4$: $\sigma_1=\lambda_1+\lambda_2+\lambda_3+\lambda_4$, $\sigma_2=\lambda_1\lambda_2+\lambda_1\lambda_3+\lambda_1\lambda_4+\lambda_2\lambda_3+\lambda_2\lambda_4+\lambda_3\lambda_4$, $\sigma_3=\lambda_1\lambda_2\lambda_3+\lambda_1\lambda_2\lambda_4+\lambda_1\lambda_3\lambda_4+\lambda_2\lambda_3\lambda_4$ et $\sigma_4=\lambda_1\lambda_2\lambda_3\lambda_4$.

Exemple Pour tout $n\geqslant 2$: $\displaystyle\sum_{k=0}^{n-1}\mathrm{e}^{\frac{2ik\pi}{n}}=\sum_{\omega\in\mathbb{U}_n}\omega=0$ et $\displaystyle\prod_{k=0}^{n-1}\mathrm{e}^{\frac{2ik\pi}{n}}=\prod_{\omega\in\mathbb{U}_n}\omega=(-1)^{n+1}$.

Démonstration Nous avons déjà calculé ces deux quantités à la main, mais notre but est différent ici. Dans le contexte du polynôme scindé X^n-1 : $\sigma_1=\displaystyle\sum_{\omega\in\mathbb{U}_n}\omega$ et $\sigma_n=\prod_{\omega\in\mathbb{U}_n}\omega$ et les relations coefficients-racines s'écrivent : $X^n-1=X^n-\sigma_1X^{n-1}+\ldots+(-1)^n\sigma_n$, donc $\sigma_1=0$ et $\sigma_n=(-1)^{n+1}$.

Exemple Le polynôme non constant X^3-3X+4 est scindé sur $\mathbb{C}$ d'après le théorème de d'Alembert-Gauss — mais pas forcément sur $\mathbb{R}$ — et nous pouvons noter x, y et z ses trois racines complexes comptées avec multiplicité. L'unique polynôme unitaire de degré 3 dont les racines sont xy, xz et yz est alors le polynôme X^3+3X^2-16.

Vous verrez ci-dessous qu'on arrive au résultat sans jamais avoir eu la moindre idée de ce que valent x, y et z.

Démonstration Il s'agit de calculer explicitement les coefficients du polynôme $(X-xy)(X-xz)(X-yz)$. Posons pour cela $\sigma_1=x+y+z$, $\sigma_2=xy+yz+zx$ et $\sigma_3=xyz$. Les relations coefficients-racines du polynôme X^3-2X+5 s'écrivent : $\sigma_1=0$, $\sigma_2=-3$ et $\sigma_3=-4$, donc :

$$\begin{aligned}(X-xy)(X-xz)(X-yz)&=X^3-(xy+xz+yz)X^2+\left(x^2yz+xy^2z+xyz^2\right)X-x^2y^2z^2\\&=X^3-\sigma_2X^2+\sigma_1\sigma_3X-\sigma_3^2=X^3+3X^2-16.\end{aligned}$$

4 Un polynôme, c'est combien d'informations ?

4.1 Résumé du chapitre et perspectives

Donnons-nous un polynôme non constant $P\in\mathbb{C}[X]$ de degré n. Que suffit-il de connaître pour connaître P entièrement ? On peut répondre de trois manières à cette question.

— Par définition, P est entièrement déterminé par ses coefficients : $P=a_nX^n+\ldots+a_1X+a_0$.

— D'après le théorème de d'Alembert-Gauss, P est entièrement déterminé par son coefficient dominant et de ses racines complexes comptées avec multiplicité : $P=a_n(X-\lambda_1)\ldots(X-\lambda_n)$.

— D'après la théorie des polynômes d'interpolation de Lagrange du prochain paragraphe, P est entièrement déterminé par la donnée de $n+1$ valeurs $P(x_1),\ldots,P(x_{n+1})$.

Dans les trois cas, P est entièrement déterminé par une collection de $n+1$ nombres — soit $n+1$ coefficients, soit le coefficient dominant et les n racines complexes comptées avec multiplicité, soit $n+1$ valeurs. À mes yeux, cette trinité de points de vue est très importante, car manipuler des polynômes, c'est jongler avec les points de vue.

Si on creuse un peu, pourtant, il est très curieux que les deux premiers points de vue soient équivalents. Pour la clarté du propos, supposons P unitaire, i.e. que $a_n=1$. Connaître P revient alors connaître au choix soit ses coefficients $a_0,\ldots,a_{n-1}$, soit ses racines $\lambda_1,\ldots,\lambda_n$. La bizarrerie de l'affaire, c'est qu'on modifie $P=X^n+a_{n-1}X^{n-1}+\ldots+a_0=(X-\lambda_1)\ldots(X-\lambda_n)$ quand on modifie l'ordre des coefficients mais pas quand on modifie l'ordre des racines, indiscernables par commutativité de $\mathbb{K}[X]$. Il y a ainsi autant d'information dans la collection ordonnée des coefficients que dans la collection non ordonnée des racines alors qu'on manipule n nombres dans les deux cas.

Le point intéressant, c'est que $a_0,\ldots,a_{n-1}$ et $\lambda_1,\ldots,\lambda_n$ ont beau déterminer entièrement P, on ne connaît pas P de la même manière selon qu'on connaît les uns ou les autres. Quand on connaît $\lambda_1,\ldots,\lambda_n$, on n'a qu'à développer le produit $P=(X-\lambda_1)\ldots(X-\lambda_n)$ pour calculer $a_0,\ldots,a_{n-1}$. Le résultat, ce sont les relations coefficients-racines, mais à défaut d'exprimer chaque racine individuellement en fonction des coefficients, ces relations donnent à toutes les racines le même rôle. En sens inverse, on ne sait pas faire, on ne sait pas calculer les racines à partir des coefficients et cette affirmation est fortement liée à l'indiscernabilité des racines. On aurait bien aimé pouvoir exprimer les racines à partir des coefficients à l'aide des symboles $+$, $\times$, $\div$ et $\sqrt[k]{\cdot}$ comme on sait le faire avec les polynômes de degré 2. Des formules sont disponibles jusqu'au degré $n=4$, mais à partir de $n=5$, un théorème profond énonce qu'aucune formule ne saurait convenir.

4.2 Interpolation de Lagrange

Étant donnés des réels $x_1,\dots,x_n$ pour lesquels $x_1 < \dots < x_n$ et des réels $y_1,\dots,y_n$ quelconques, le problème de l'*interpolation* consiste à construire des fonctions $f : [x_1,x_n] \longrightarrow \mathbb{R}$ pour lesquelles $f(x_i) = y_i$ pour tout $i \in [\![1,n]\!]$. Il existe bien sûr beaucoup de fonctions de ce genre, on peut par exemple en construire une en reliant linéairement les points de coordonnées $(x_1,y_1),\dots,(x_n,y_n)$. La méthode d'*interpolation de Lagrange* de ce paragraphe propose une autre approche.

Définition **(Polynômes de Lagrange d'une famille de points distincts)** Soient $x_1,\dots,x_n \in \mathbb{K}$ distincts. Pour tout $i \in [\![1,n]\!]$, on pose $L_i = \prod_{\substack{1\leqslant k\leqslant n\\ k\neq i}} \frac{X-x_k}{x_i-x_k}$. Les polynômes $L_1,\dots,L_n$ sont appelés les *polynômes de Lagrange de* $x_1,\dots,x_n$.

Propriété fondamentale : Pour tous $i,j \in [\![1,n]\!]$: $L_i(x_j) = \delta_{ij}$.
En particulier, L_i admet $x_1,\dots,x_{i-1},x_{i+1},\dots,x_n$ pour racines.

Pour $n=3$: $L_1 = \dfrac{(X-x_2)(X-x_3)}{(x_1-x_2)(x_1-x_3)}$, $L_2 = \dfrac{(X-x_1)(X-x_3)}{(x_2-x_1)(x_2-x_3)}$ et $L_3 = \dfrac{(X-x_1)(X-x_2)}{(x_3-x_1)(x_3-x_2)}$.

Théorème **(Polynôme d'interpolation de Lagrange de degré minimal)** Soient $x_1,\dots,x_n \in \mathbb{K}$ distincts. On note $L_1,\dots,L_n$ leurs polynômes de Lagrange. Pour tous $y_1,\dots,y_n \in \mathbb{K}$, $y_1L_1+\dots+y_nL_n$ est le seul polynôme $P \in \mathbb{K}_{n-1}[X]$ pour lequel $P(x_i) = y_i$ pour tout $i \in [\![1,n]\!]$.

On rappelle que $\mathbb{K}_{n-1}[X]$ est l'ensemble des polynômes de degré AU PLUS $n-1$ de $\mathbb{K}[X]$.

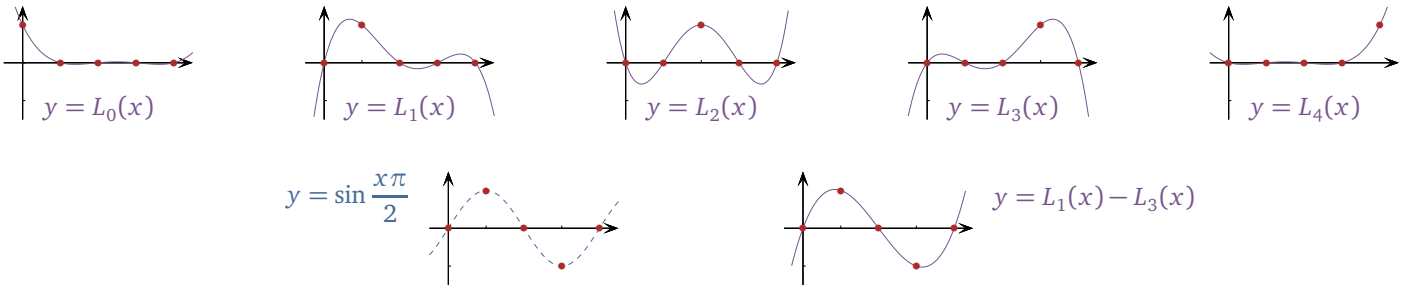
Démonstration Fixons $y_1,\dots,y_n \in \mathbb{K}$ quelconques.

- **Existence :** Posons $P = y_1L_1+\dots+y_nL_n$. Les polynômes $L_1,\dots,L_n$ sont de degré au plus $n-1$, donc P aussi. Ensuite, pour tout $j \in [\![1,n]\!]$: $P(x_j) = y_1L_1(x_j)+\dots+y_nL_n(x_j) = y_j$.
- **Unicité :** Soient $P,Q \in \mathbb{K}_{n-1}[X]$ deux polynômes pour lesquels $P(x_i) = Q(x_i) = y_i$ pour tout $i \in [\![1,n]\!]$. Le polynôme $P-Q$ est lui aussi de degré au plus $n-1$ et admet $x_1,\dots,x_n$ pour racines distinctes, il est donc nul par rigidité, autrement dit $P = Q$.

Exemple Notons f la fonction $x \longmapsto \sin\dfrac{x\pi}{2}$ sur $[0,4]$, pour laquelle : $f(0) = f(2) = f(4) = 0$, $f(1) = 1$ et $f(3) = -1$. Notons ensuite $L_0,\dots,L_4$ les polynômes de Lagrange de $0,\dots,4$. Le polynôme d'interpolation de Lagrange de f aux points $0,\dots,4$ vaut alors $f(0)L_0+\dots+f(4)L_4 = L_1-L_3$. Or :

$$L_1 = -\frac{1}{6}X(X-2)(X-3)(X-4) \quad \text{et} \quad L_3 = -\frac{1}{6}X(X-1)(X-2)(X-4),$$

donc $L_1-L_3 = -\dfrac{1}{6}X(X-2)(X-3)(X-4)+\dfrac{1}{6}X(X-1)(X-2)(X-4) = \dfrac{1}{3}X(X-2)(X-4)$.



Théorème **(Détermination d'un polynôme par ses valeurs)** Soient $x_1,\dots,x_n \in \mathbb{K}$ distincts. On note $L_1,\dots,L_n$ leurs polynômes de Lagrange.

Pour tout $P \in \mathbb{K}_{n-1}[X]$: $P = \sum_{i=1}^{n} P(x_i)L_i$, et aucun autre polynôme de la forme $\sum_{i=1}^{n} y_iL_i$ ne coïncide avec P.

Démonstration Pour tout $P \in \mathbb{K}_{n-1}[X]$, P envoie x_i sur $P(x_i)$ pour tout $i \in [\![1,n]\!]$, donc d'après le théorème précédent, $P = P(x_1)L_1+\dots+P(x_n)L_n$.

5 Polynômes annulateurs d'une matrice carrée

Soit $A \in \mathcal{M}_n(\mathbb{K})$. Pour tout $P = \sum_{k=0}^{+\infty} a_k X^k \in \mathbb{K}[X]$, on pose $P(A) = \sum_{k=0}^{+\infty} a_k A^k$. Le résultat de cette *évaluation en A* est une matrice carrée et non pas un polynôme, mais on dit que cette matrice carrée est un *polynôme en A*. L'ensemble des polynômes en A est quant à lui noté $\mathbb{K}[A]$.

Attention ! Évalué en A, le polynôme constant 1 se change en I_n. Par exemple, si $P = X^2 + 3$, alors $P(A) = A^2 + 3I_n$.

Définition-théorème (Polynômes annulateurs d'une matrice carrée) Soit $A \in \mathcal{M}_n(\mathbb{K})$. On appelle *polynôme annulateur de A* tout polynôme $P \in \mathbb{K}[X]$ pour lequel $P(A) = 0$.

On dit aussi que A *annule* P, mais jamais que A est « racine de P ». Les racines sont définitivement des éléments de $\mathbb{K}$.

Exemple Le polynôme $X^3 - 2X^2 - 1$ annule $A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 0 \end{pmatrix}$ car $A^3 - 2A^2 - I_3 = 0$ après calcul.

Exemple Soit $A \in \mathcal{M}_2(\mathbb{C})$. Il n'est pas dur de vérifier coefficient par coefficient que le polynôme $X^2 - \mathrm{tr}(A)X + \det(A)$ annule A, i.e. que $A^2 = \mathrm{tr}(A)A - \det(A)I_2$. Par exemple, $X^2 - 6X - 1$ annule $\begin{pmatrix} 1 & 3 \\ 2 & 5 \end{pmatrix}$.

Théorème (Deux remarques sur les polynômes annulateurs) Soit $A \in \mathcal{M}_n(\mathbb{K})$.

(i) Si A possède un polynôme annulateur de degré d, tout polynôme en A est combinaison linéaire des puissances $I_n, A, A^2, \dots, A^{d-1}$.

(ii) Si A possède un polynôme annulateur dont le coefficient constant est non nul, alors A est inversible.

Démonstration Faisons l'hypothèse que A possède un polynôme annulateur $\Pi = a_d X^d + \dots + a_1 X + a_0$ avec $a_0, \dots, a_{d-1} \in \mathbb{C}$ et $a_d \in \mathbb{C}^*$.

(i) Soit $P \in \mathbb{C}[X]$. Par division euclidienne, $P = \Pi Q + R$ pour certains $Q \in \mathbb{C}[X]$ et $R \in \mathbb{C}_{d-1}[X]$, donc après évaluation en A : $P(A) = \Pi(A)Q(A) + R(A) = R(A)$, donc $P(A)$ est combinaison linéaire de $I_n, A, \dots, A^{d-1}$.

(ii) Si $a_0 \neq 0$, alors : $-a_0 I_n = \sum_{k=1}^{n} a_k A^k = A \times \left(\sum_{k=1}^{n} a_k A^{k-1} \right) = \left(\sum_{k=1}^{n} a_k A^{k-1} \right) \times A$, donc A est inversible d'inverse $-\frac{1}{a_0} \sum_{k=1}^{n} a_k A^{k-1}$.

Exemple Revenons à la matrice $A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 0 \end{pmatrix}$ et à son polynôme annulateur $X^3 - 2X^2 - 1$. Comme $A^3 - 2A^2 = I_3$, alors $A \times (A^2 - 2A) = (A^2 - 2A) \times A = I_3$, donc A est inversible et $A^{-1} = A^2 - 2A = \begin{pmatrix} 0 & 1 & -1 \\ 0 & 0 & 1 \\ 1 & -1 & 0 \end{pmatrix}$.

Les polynômes annulateurs d'une matrice carrée servent aussi à calculer ses puissances. Deux mots d'ordre en la matière, division euclidienne et racines !

Exemple On pose $A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}$. Pour tout $k \in \mathbb{N}^*$: $A^k = \frac{1}{3} \begin{pmatrix} 2^{k+1} + (-1)^k & 2^k - (-1)^k & 2^k - (-1)^k \\ 2^k - (-1)^k & 2^{k-1} + (-1)^k & 2^{k-1} + (-1)^k \\ 2^k - (-1)^k & 2^{k-1} + (-1)^k & 2^{k-1} + (-1)^k \end{pmatrix}$.

Démonstration $A^2 = \begin{pmatrix} 3 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$ et $A^3 = \begin{pmatrix} 5 & 3 & 3 \\ 3 & 1 & 1 \\ 3 & 1 & 1 \end{pmatrix} = A^2 + 2A$,

donc le polynôme $P = X^3 - X^2 - 2X = (X+1)X(X-2)$ annule A.

À présent, soit $k \in \mathbb{N}^*$. La division euclidienne de X^k par P s'écrit $X^k = PQ + aX^2 + bX + c$ pour certains $Q \in \mathbb{R}[X]$ et $a, b, c \in \mathbb{R}$. Évaluons en les racines de P : $(-1)^k = a - b + c$, $0 = c$ et $2^k = 4a + 2b + c$. Après calcul : $a = \frac{2^{k-1} + (-1)^k}{3}$, $b = \frac{2^{k-1} - 2(-1)^k}{3}$ et $c = 0$, donc :

$$A^k = \underbrace{P(A)}_{=0} Q(A) + aA^2 + bA + cI_3 = \frac{2^{k-1} + (-1)^k}{3} \begin{pmatrix} 3 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} + \frac{2^{k-1} - 2(-1)^k}{3} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

6 Quelques remarques sur l'anneau $\mathbb{F}_p[X]$

Dans ce paragraphe, p est un nombre premier fixé une fois pour toutes. Nous avons construit $\mathbb{K}[X]$ pour $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$, mais n'importe quel corps aurait fait l'affaire en réalité. Pour tout corps $\mathbb{K}$, l'anneau $\mathbb{K}[X]$ est commutatif intègre et les résultats suivants sont conservés :

— le théorème de la division euclidienne,

— le théorème de factorisation par les racines,

— le théorème de rigidité,

— le théorème d'interpolation de Lagrange.

Certains phénomènes dépendent tout de même de $\mathbb{K}$. Jusqu'où $\mathbb{F}_p[X] = {}^{\mathbb{Z}}\!/_{p\mathbb{Z}}[X]$ ressemble-t-il à $\mathbb{R}[X]$ et $\mathbb{C}[X]$?

Exemple Dans $\mathbb{F}_3[X]$: $\left(X+\overline{1}\right)^2 = X^2 + \overline{2}X + \overline{1} = X^2 - X + \overline{1}$ et $\left(X+\overline{1}\right)^3 = X^3 + \overline{3}X^2 + \overline{3}X + \overline{1} = X^3 + \overline{1}$.

Attention ! La dérivation pose problème dans $\mathbb{F}_p[X]$. Par exemple, $\left(X^p\right)' = \overline{p}X^{p-1} = \overline{0}$ alors que X^p n'est pas constant. Il n'est donc pas vrai que $\deg(P') = \deg(P) - 1$ pour tout $P \in \mathbb{F}_p[X]$ non constant. La formule de Taylor tombe également, et avec elle la possibilité de calculer les multiplicités à partir des polynômes dérivés. En effet, p divise $k!$ pour tout $k \geqslant p$, donc $\overline{k!} = \overline{0}$ et on ne peut pas diviser par $\overline{0}$.

Exemple $X^p - X = \prod_{x \in \mathbb{F}_p} (X - x)$. A fortiori, le polynôme $X^p - X + 1$ n'a pas de racine dans $\mathbb{F}_p$.

Démonstration De degré p, $X^p - X$ possède au plus p racines comptées avec multiplicité. Cela dit, tout élément de $\mathbb{F}_p$ en est racine d'après le petit théorème de Fermat, donc $X^p - X$ est scindé sur $\mathbb{F}_p$ et le résultat proposé est sa forme scindée.

Nous l'avons vu, les fonctions de $\mathbb{R}$ dans $\mathbb{R}$ ne sont pas toutes polynomiales, mais au moins, tout polynôme peut être identifié à la fonction polynomiale qui lui est associée. Sur $\mathbb{F}_p$, c'est l'inverse ! Toute fonction est polynomiale, mais le théorème d'identification tombe. Par exemple, les polynômes X^p et X sont distincts, mais $x^p = x$ pour tout $x \in \mathbb{F}_p$ d'après le petit théorème de Fermat.

Théorème (Toute fonction est polynomiale) Le morphisme d'anneaux $P \longmapsto \widetilde{P}$ est surjectif de $\mathbb{F}_p[X]$ sur $\mathbb{F}_p^{\mathbb{F}_p}$. En d'autres termes, toute fonction de $\mathbb{F}_p$ dans $\mathbb{F}_p$ est polynomiale.

Démonstration Notons $L_0, \dots, L_{p-1}$ les polynômes de Lagrange de $\overline{0}, \dots, \overline{p-1}$. Toute fonction $f : \mathbb{F}_p \longrightarrow \mathbb{F}_p$ coïncide avec la fonction polynomiale $\widetilde{P}$ si on note P le polynôme $f\left(\overline{0}\right)L_0 + \dots + f\left(\overline{p-1}\right)L_{p-1}$.

Allez, un dernier pour la route. Grâce au résultat qui suit, nous pourrons parler des polynômes à coefficients entiers au chapitre « Arithmétique des polynômes et fractions rationnelles » en les observant dans $\mathbb{F}_p[X]$. Je ne le détaillerai pas, mais l'ensemble $\mathbb{Z}[X]$ des polynômes à coefficients dans $\mathbb{Z}$ est un sous-anneau de $\mathbb{R}[X]$, et même de $\mathbb{Q}[X]$.

Théorème (Réduction canonique modulo p) L'application $\sum_{k=0}^{+\infty} a_k X^k \longmapsto \sum_{k=0}^{+\infty} \overline{a_k} X^k$, dite *réduction canonique modulo p*, est un morphisme d'anneaux de $\mathbb{Z}[X]$ dans $\mathbb{F}_p[X]$.

Démonstration En notant f la réduction canonique modulo p : $f(1) = \overline{1}$.

Ensuite, pour tous $P = \sum_{k=0}^{+\infty} a_k X^k, Q = \sum_{k=0}^{+\infty} b_k X^k \in \mathbb{Z}[X]$:

$$f(P+Q) = f\left(\sum_{k=0}^{+\infty} (a_k + b_k) X^k\right) = \sum_{k=0}^{+\infty} \overline{(a_k + b_k)} X^k = \sum_{k=0}^{+\infty} \left(\overline{a_k} + \overline{b_k}\right) X^k = \sum_{k=0}^{+\infty} \overline{a_k} X^k + \sum_{k=0}^{+\infty} \overline{b_k} X^k = f(P) + f(Q)$$

et on montre de même que $f(PQ) = f(P)f(Q)$.