CPI1

Algèbre

Cours

jean-michel.masereel@eisti.eu

La vie n'est bonne qu'à deux choses :
Découvrir les mathématiques et enseigner les mathématiques.

Siméon Denis Poisson (1781 - 1840)

Sommaire

Première partie

S1

1. Logique et raisonnement

Tous les hommes sont mortels, or Socrate est un homme donc Socrate est mortel.

(Aristote.)

Autre syllogisme : tous les chats sont mortels. Socrate est mortel. Donc Socrate est un chat.

(Le logicien, dans Le rhinocéros*, E. Ionesco.)*

Ce chapitre a pour objectif de vous apprendre à raisonner correctement, c'est-à-dire pas comme le logicien de Ionesco. Vous utilisez déjà, peut-être sans le savoir, le syllogisme formalisé par Aristote. En effet, vous avez déjà l'habitude de démontrer qu'une implication « $P \Rightarrow Q$ » est vraie, puis de vérifier que la proposition P est vraie pour en déduire que la proposition Q est vraie.

Nous allons voir dans ce chapitre d'autres types de raisonnements.

1.1. Rudiments de logique

1.1.1. Première définition

Définition 1.1 (Proposition). On appelle *proposition* un énoncé (mathématiques) qui peut être soit vrai (noté V ou 1) soit faux (noté F ou 0).

Les énoncés ne sont pas forcément mathématiques. Durant votre parcours à l'EISTI, nous ne nous intéresserons qu'aux énoncés mathématiques et physique, même si, pour expliquer un concept, nous pourrons utiliser des exemples issus d'autres disciplines.

Exemple 1.1.
- « 1+1=2 » est une proposition vraie.
- « L'entier 2014 est impair » est une proposition fausse.
- « Tous les élèves de la classe auront leur diplôme d'ici 5 ans ». Nous ne pouvons pas encore déterminer si cet énoncé est vrai ou faux, il reste néanmoins soit vrai soit faux, c'est donc bien une proposition.
- « Quelle heure est-il ? » est un énoncé, mais il n'a pas de valeurs de vérité : ce n'est ni vrai ni faux. Ce n'est donc pas une proposition.

1.1.2. Opérations

À partir d'une ou plusieurs propositions, il est possible d'en créer de nouvelles.

Définition 1.2 (Négation). Soit une proposition P. On définit la *négation de* P que l'on note « $\neg P$ » ou « $\overline{P}$ » ou encore « $\text{NON}\,(P)$ », comme la proposition qui est vraie lorsque P est fausse et est fausse lorsque P est vraie.

Remarque 1.1. Les deux propositions P et $\overline{P}$ ne peuvent pas être vraies en même temps.

Exemple 1.2. Soit P la proposition « Il fait plus de 20 °C aujourd'hui ». La négation de P est « Il fait moins de 20 °C aujourd'hui ». Il ne peut pas faire à la fois plus de 20 °C et moins de 20 °C.

Définition 1.3 (Conjonction). Soit deux propositions P et Q. On définit la *conjonction de* P *et* Q que l'on note « $P \wedge Q$ » ou « P ET Q », comme la proposition qui est vraie lorsque P et Q sont toutes les deux vraies, et fausse lorsque au moins l'une des deux propositions P ou Q est fausse.

Exemple 1.3. Soit P la proposition « Il fait plus de 20 °C aujourd'hui » et Q la proposition « Il va pleuvoir demain ». La conjonction de P et Q est « Il fait plus de 20 °C aujourd'hui et il va pleuvoir demain ».

Définition 1.4 (Disjonction). Soit deux propositions P et Q. On définit la *disjonction de* P *et* Q que l'on note « $P \vee Q$ » ou « P OU Q », comme la proposition qui est vraie lorsque au moins l'une des deux propositions P ou Q est vraie, et fausse lorsque les deux propositions P et Q sont fausses.

Exemple 1.4. Soit P la proposition « Il fait plus de 20 °C aujourd'hui » et Q la proposition « Il va pleuvoir demain ». La disjonction de P et Q est « Il fait plus de 20 °C aujourd'hui ou il va pleuvoir demain ».

Remarque 1.2. Attention, le « ou » mathématiques n'a pas exactement le même sens que le « ou » français. En effet, le « ou » mathématique est *non exclusif*, c'est-à-dire que l'on n'interdit pas à P et Q d'être vrais tous les deux. En français, le « ou » est plutôt considéré comme *exclusif*. Lorsque dans un restaurant par exemple, il est écrit sur la carte « fromage ou dessert », il ne viendrait à l'idée de personne (sauf peut-être d'un prof de maths) de commander les deux.

Dans les définitions précédentes, il n'y avait pas de lien entre les deux propositions P et Q. Nous allons voir maintenant des opérations qui sous tendent un lien entre P et Q.

Définition 1.5 (Implication). Soit P et Q deux propositions. On définit la proposition P *implique* Q, notée $P \Rightarrow Q$, comme la proposition qui est vraie si P est fausse ou si P et Q sont toutes les deux vraies ; et qui est fausse si P est vraie et Q fausse.

Exemple 1.5.
- Soit P la proposition « je ne me suis pas réveillé ce matin » et Q la proposition « je suis en retard à l'école ». La proposition $P \Rightarrow Q$ est vraie. Remarquez que la proposition $Q \Rightarrow P$ est fausse.
- La proposition $1 = 2 \Rightarrow \pi = e$ est-elle vraie ou fausse ?

Définition 1.6 (Réciproque). Soit P et Q deux propositions. La *réciproque* de la proposition « $P \Rightarrow Q$ » est la proposition « $Q \Rightarrow P$ »

Tout comme nous l'avons vu sur l'exemple précédent, connaître la valeur de vérité de $P \Rightarrow Q$ ne donne aucune information sur la valeur de vérité de $Q \Rightarrow P$.

Définition 1.7 (Équivalence). Soit P et Q deux propositions. On définit la proposition P *équivalent à* Q, notée $P \Leftrightarrow Q$, comme la proposition qui est vraie si P et Q sont vraies ou si P et Q sont fausses ; et qui est fausse si l'une est fausse et l'autre vraie.

Remarque 1.3. On démontrera en TD que la proposition $P \Leftrightarrow Q$ a la même valeur de vérité que $(P \Rightarrow Q)$ ET $(Q \Rightarrow P)$. Cela donne une certaine cohérence dans les notations $\Rightarrow$ et $\Leftrightarrow$.

Exemple 1.6.
— Il semble difficile de trouver un exemple de la vie courante. Et vous ? Réussirez-vous à en proposer une ?
— Soit a et b deux réels. L'équivalence $a = b \Leftrightarrow \mathrm{e}^a = \mathrm{e}^b$ est vraie.
— La proposition $1 = 2 \Leftrightarrow 3^2 = 9$ est elle vraie ou fausse ?

Attention de ne pas confondre l'implication et l'équivalence, ce que vous faites souvent. Lorsque l'on rédige une démonstration il est plus naturel de raisonner en implication. Or on a plus souvent besoin d'avoir des équivalences. Lorsque l'on écrit une équivalence, il faut donc toujours penser à vérifier que la réciproque est bien vraie.

Définition 1.8 (Condition nécessaire/suffisante).

— Soit $P \Rightarrow Q$ une implication. On dit que P est une *condition suffisante* à Q, et que Q est une *condition nécessaire* à P.
— Soit $P \Leftrightarrow Q$ une équivalence. On dit que P est une *condition nécessaire et suffisante* à Q (ou que Q est une condition nécessaire et suffisante à P).

Remarque 1.4. Supposons que $P \Rightarrow Q$.
— Le fait que P soit vraie <u>suffit</u> à ce que Q soit vraie.
— Par contre le fait que P soit vraie n'est <u>pas nécessaire</u> à Q puisque Q peut très bien être vraie alors que P ne l'est pas.
— Lorsque P est vraie, Q est <u>nécessairement</u> vraie aussi.
— Le fait que Q soit vraie ne <u>suffit pas</u> à ce que P soit vraie puisque, comme nous l'avons déjà signalé, Q peut très bien être vraie alors que P ne l'est pas.

Remarque 1.5. Lorsque $P \Leftrightarrow Q$ chaque proposition est à la fois <u>nécessaire</u> et <u>suffisante</u> à l'autre proposition.

1.1.3. Tables de vérité

Lorsque l'on fait des opérations sur des propositions, il est parfois plus aisé d'utiliser ce que l'on appelle une *table de vérité*. Celle-ci résume les différentes valeurs prise par une proposition composée. Ci dessous la table de vérité des opérations définies dans la section précédente.

P	NON (P)
V	F
F	V

P	Q	P OU Q	P ET Q	$P \Rightarrow Q$	$P \Leftrightarrow Q$
V	V	V	V	V	V
V	F	V	F	F	F
F	V	V	F	V	F
F	F	F	F	V	V

Propriété 1.1. Deux propositions ont la même table de vérité si et seulement si elles sont équivalentes.

Démonstration. Cela découle directement de la définition d'une équivalence : Si les deux propositions prennent les mêmes valeurs de vérité en même temps, alors l'équivalence est vraie. Réciproquement, si elles sont équivalentes alors elles prennent forcément les mêmes valeurs de vérité en même temps. □

Exercice 1.1. À l'aide d'une table de vérité démontrer que les deux propositions suivantes sont vraies.

1. $\big((P \text{ OU } Q) \text{ ET } R\big) \Leftrightarrow \big((P \text{ ET } R) \text{ OU } (Q \text{ ET } R)\big)$
2. $\big((P \text{ ET } Q) \text{ OU } R\big) \Leftrightarrow \big((P \text{ OU } R) \text{ ET } (Q \text{ OU } R)\big)$

Remarque 1.6. On dit que la conjonction (resp. la disjonction) est distributive sur la disjonction (resp. conjonction).

1.1.4. Quantificateurs

Parfois une propriété peut dépendre d'un élément (une variable) et suivant la valeur de cet élément, la propriété est vraie ou fausse. Afin de noter simplement des propositions faisant intervenir de telles propriétés, nous avons besoin des *quantificateurs*.

Définition 1.9 (Quantificateur existentiel). Soit $\mathscr{P}$ une propriété portant sur les éléments d'un ensemble E. La proposition « au moins un élément x de E vérifie $\mathscr{P}$ » se note « $\exists x \in E,\ \mathscr{P}(x)$ ». On appelle *quantificateur existentiel* le symbole « $\exists$ ».

Exemple 1.7. Soit E l'ensemble des élèves de la classe et $\mathscr{P}$ la propriété « L'élève x possède des lunettes ». La proposition $\exists x \in E,\ \mathscr{P}(x)$ se lit « au moins un élève de la classe possède des lunettes » ou encore, « il existe un élève de la classe qui possède des lunettes ».

Définition 1.10 (Quantificateur universel). Soit $\mathscr{P}$ une propriété portant sur les éléments d'un ensemble E. La proposition « tout élément x de E vérifie $\mathscr{P}$ » se note « $\forall x \in E,\ \mathscr{P}(x)$ ». On appelle *quantificateur universel* le symbole « $\forall$ ».

Exemple 1.8. Soit E l'ensemble des élèves de la classe et $\mathscr{P}$ la propriété « L'élève x possède des lunettes ». La proposition $\forall x \in E,\ \mathscr{P}(x)$ se lit « tout élève de la classe possède des lunettes ».

Définition 1.11. Soit $\mathscr{P}$ une propriété portant sur les éléments d'un ensemble E. La proposition « un et un seul élément x de E vérifie $\mathscr{P}$ » se note « $\exists! x \in E,\ \mathscr{P}(x)$ ».

Exemple 1.9. Soit E l'ensemble des élèves de la classe et $\mathscr{P}$ la propriété « L'élève x possède des lunettes ». La proposition $\exists! x \in E,\ \mathscr{P}(x)$ se lit « il existe un unique élève de la classe qui possède des lunettes ».

Il est possible d'utiliser plusieurs quantificateurs dans une même propositions. Il faut alors faire très attention à l'ordre dans lequel ils sont utilisés. Comparez le sens des deux phrases suivantes :

— Pour toute serrure, il existe une clef qui l'ouvre $(\forall x,\ \exists y,\ \mathscr{P}(x,y))$.

— Il existe une clef qui ouvre toutes les serrures $(\exists y,\ \forall x,\ \mathscr{P}(x,y))$.

1.1.5. Négations

Nous avons vu la négation d'une simple proposition, mais qu'en est-il de la négation d'une proposition complexe faisant intervenir des opérations et des quantificateurs ?

Propriété 1.2. Soit P et Q deux propositions. On a alors

— $\text{NON}(\text{NON}(P)) \Leftrightarrow P$

— $\text{NON}(P \text{ ET } Q) \Leftrightarrow \big((\text{NON}(P)) \text{ OU } (\text{NON}(Q))\big)$ (Première loi de De Morgan)

— $\text{NON}(P \text{ OU } Q) \Leftrightarrow \big((\text{NON}(P)) \text{ ET } (\text{NON}(Q))\big)$ (Deuxième loi de De Morgan)

— $\text{NON}(P \Rightarrow Q) \Leftrightarrow (P \text{ ET } (\text{NON}(Q)))$

Démonstration. La démonstration se fait très simplement avec les tables de vérités. □

Propriété 1.3. Soit $\mathscr{P}$ une propriété portant sur les éléments d'un ensemble E. On a alors

— $\text{NON}(\exists x \in E,\ \mathscr{P}(x))$ est $\forall x \in E,\ \text{NON}(\mathscr{P}(x))$.

— $\text{NON}(\forall x \in E,\ \mathscr{P}(x))$ est $\exists x \in E,\ \text{NON}(\mathscr{P}(x))$.

Remarque 1.7. Il suffit de retenir que pour écrire la négation d'une proposition contenant des quantificateurs, il faut remplacer les $\exists$ par des $\forall$ et réciproquement, et prendre la négation de la proposition.

Exemple 1.10. La négation de « tous les élèves ont des lunettes » **n'est pas** « aucun élève n'a de lunette » mais « il existe un élève qui n'a pas de lunettes ».

1.2. Différents types de raisonnements

1.2.1. Raisonnement par contraposition

Définition 1.12 (Contraposée). Soit P et Q deux propositions. La *contraposée* de la proposition « $P \Rightarrow Q$ » est la proposition « $\text{NON}(Q) \Rightarrow \text{NON}(P)$ »

Propriété 1.4. Une implication et sa contraposée sont équivalentes. C'est-à-dire que

$$(P \Rightarrow Q) \Leftrightarrow (\text{NON}(Q) \Rightarrow \text{NON}(P))$$

Démonstration. Vous avez déjà démontré cela à l'aide des tables de vérité lors de la question 15 de l'exercice 1 du polycopié de TD du chapitre 1 « Logique et raisonnement ». Nous allons cependant la redémontrer en utilisant les propriétés de logique que l'on connaît déjà.

$$\begin{aligned}(P \Rightarrow Q) &\Leftrightarrow (\text{NON}(P) \ \text{ OU } \ Q)\\ &\Leftrightarrow (Q \ \text{ OU } \ \text{NON}(P))\\ &\Leftrightarrow (\text{NON}(\text{NON}(Q)) \ \text{ OU } \ \text{NON}(P))\\ &\Leftrightarrow (\text{NON}(Q) \Rightarrow \text{NON}(P))\end{aligned}$$

Remarque 1.8.

— Attention ! Ne confondez pas contraposée et réciproque. Si une implication est vraie, sa contraposée est toujours vraie, par contre sa réciproque ne l'est pas forcément.

— Ainsi, lorsque l'on doit démontrer une implication du type $P \Rightarrow Q$, on peut soit la démontrer directement, soit démontrer sa contraposée : on suppose que Q est faux et on démontre que P est faux aussi.

Exemple 1.11. Afin de démontrer qu'un triangle dont on connaît les longueurs des côtés n'est pas rectangle, on utilise la contraposée du théorème de Pythagore. En effet, le théorème de Pythagore « Si le triangle ABC est rectangle en A, alors $BC^2 = AB^2 + AC^2$ » est équivalent à sa contraposée « Si $BC^2 \neq AB^2 + AC^2$ alors le triangle ABC n'est pas rectangle en A »

Remarque 1.9. Notez bien qu'un raisonnement par contraposé s'applique exclusivement à des implications, alors qu'un raisonnement par l'absurde peut s'appliquer à différents types de propriétés.

1.2.2. Raisonnement par l'absurde

On veut montrer une proposition P. Pour cela on suppose que cette proposition est fausse, et on tente d'en déduire une absurdité. Cette absurdité peut être soit une proposition Q qui est clairement fausse, soit une proposition de type $Q = (R \ \text{ ET } \ \text{NON}(R))$ (qui est clairement fausse aussi).
Intuitivement on voit bien que si l'on aboutit à une absurdité, c'est que l'hypothèse de départ (P fausse) était fausse (et donc que P est vraie). Cela se démontre aussi en utilisant les propriétés de logique :

Nous avons montré que $(\text{NON}(P) \Rightarrow Q)$, or par contraposition cela revient à dire que
$(\text{NON}(Q) \Rightarrow \text{NON}(\text{NON}(P)))$ ou encore $(\text{NON}(Q) \Rightarrow P)$. Mais puisque Q est fausse, $\text{NON}(Q)$ est vraie et donc P aussi.

Exemple 1.12. Un exemple fréquent consiste à vouloir démontrer une proposition P de type $(A \Rightarrow B)$ par l'absurde. On part donc de $\text{NON}\,(P)$, c'est-à-dire de $(A \text{ ET } \text{NON}\,(B))$. Par divers raisonnement et calculs on aboutit à une absurdité. En général cette absurdité est la proposition $\text{NON}\,(A)$, qui est effectivement absurde puisque l'on est parti de A.

Remarque 1.10. [Attention ! !] L'exemple fréquent 1.12 que l'on vient de voir conduit aussi malheureusement à une erreur tout aussi fréquente : confondre le raisonnement par contraposition et le raisonnement par l'absurde. Beaucoup de personnes pensent démontrer $(A \Rightarrow B)$ par l'absurde en écrivant des choses du type :

Supposons que l'on a A et pas B. Puisque l'on n'a pas B . . .

$\vdots$

Et donc on n'a pas A. Ce qui est absurde puisque l'on a supposé que l'on avait A.

Si à aucun moment de la démonstration (dans la partie en pointillée), la proposition A n'est utilisée, alors ce raisonnement est tout bêtement un raisonnement par contraposition : $(\text{NON}\,(B) \Rightarrow \text{NON}\,(A))$.

1.2.3. Raisonnement par analyse-synthèse

Nous devons parfois démontrer des choses trop complexes pour être attaquées de front. Il nous faut d'abord analyser le problème afin de le réduire. C'est le cas lorsque l'on cherche à montrer l'existence d'un objet (un nombre par exemple) devant satisfaire certaines conditions. Une analyse du problème permet de réduire le nombre de candidats possibles. La synthèse consiste alors à déterminer parmi les candidats ceux qui sont effectivement solutions du problème.

Le schéma d'un raisonnement par analyse-synthèse est le suivant :

Analyse : On suppose que l'objet existe et on essaie de trouver des conditions nécessaires que doit vérifier cet objet. Ce faisant, on prouve que si l'objet existe, alors il est nécessairement dans un ensemble restreint.

Synthèse : On considère les objets identifiés dans la partie analyse, et on vérifie que certains ont bien les propriétés voulues (ceci assure l'existence).

Remarque 1.11. Si l'analyse n'exhibe qu'un seul candidat possible, on parle alors d'*unicité si existence*.

Exemple 1.13. [Le BN géant]
Commençons par rendre à César ce qui appartient à César. L'exemple qui suit est tiré de cette page : `http://la-bnbox.fr/c541-Le-raisonnement-par-analyse-synthese.cahier`

Un jour, on vous demande de prouver l'existence d'un BN au chocolat géant vivant !
À première vue, cela semble assez difficile à faire. . . Et vous n'avez aucune idée de la manière de procéder.
Alors procédons par Analyse-Synthèse.

Analyse : Supposons qu'il existe un BN au chocolat géant vivant quelque part dans le monde.

— Si un BN de ce genre existe, il est évident qu'il vivra nécessairement loin de l'eau, parce qu'un BN dans l'eau devient tout mou et se dissous. . .

— Si ce type de BN existe, il se trouvera nécessairement loin des régions chaudes, sinon son chocolat fondrait et il disparaitrait.

— Il sera aussi nécessairement loin des régions très froides, pour ne pas geler. *(Je laisse l'auteur responsable de cette affirmation, personnellement je pense que rien ne prouve qu'il ne survivrait pas à des régions très froides.)*

Ces conditions nécessaires réduisent déjà notre champ de recherche. On sait que maintenant, le seul endroit où on peut trouver un BN de ce type, c'est en France. *(Là encore je laisse l'auteur responsable de cette conclusion un peu hâtive. . .)*
Mais la France c'est toujours assez grand. On va donc chercher d'autres conditions nécessaires encore plus restrictives.

Un grand BN comme ça, ça a besoin de beaucoup de chocolat pour tenir ensemble... Il doit donc nécessairement vivre près d'une chocolaterie, ou d'une biscuiterie.
De plus, les BN sont créés à Nantes. En admettant le fait qu'un grand gaillard comme lui habite encore près de chez ses parents, il habite nécessairement près de Nantes.
Ce qui nous amène directement à la conclusion que le BN géant habite dans la biscuiterie BN. *(Encore une conclusion bien hâtive, comment de « près de Nantes » on aboutit à « dans la biscuiterie BN » ? Est-ce la seule biscuiterie de la région ? Passons, passons, c'est juste pour l'exemple...)*

Synthèse : Nous devons vérifier notre conclusion, c'est-à-dire que nous devons prendre le premier avion ou le premier TGV *(ou tout autre moyen de transport, même si ce n'est pas le premier)* pour Nantes et nous rendre à la biscuiterie pour vérifier que le BN géant s'y trouve.

- Soit on le trouve, et on a bien prouvé qu'il existe.
- Soit on ne le trouve pas, et on a prouvé qu'il n'existe pas, puisqu'il n'est pas à l'endroit où il devait nécessairement être.

Exemple 1.14. Soit (x_n) la suite réelle définie par récurrence par $x_0 = 1$ et $x_{n+1} = \sqrt[5]{3x_n}$. Déterminer un majorant de (x_n).

Le problème semble bien complexe. Une analyse s'impose !

Analyse : Nous ne voyons pas du tout à quoi pourrait ressembler un majorant de (x_n). Supposons donc que nous soyons en présence d'un tel majorant et étudions les propriétés qu'il possède nécessairement.
Tout d'abord, si un majorant existe, alors l'ensemble des majorants possède un plus petit élément (c'est une propriété qui est hors programme). Notons M_{min} ce plus petit des majorants.
Puisque M_{min} est un majorant, nous pouvons affirmer que pour tout $n \in \mathbb{N}$, $x_n \leq M_{min}$. En multipliant par 3 l'inégalité, nous avons $3x_n \leq 3M_{min}$. La fonction « racine-5ᵉ » étant croissante (admis), nous pouvons prendre la racine-5ᵉ de l'inégalité pour obtenir $\sqrt[5]{3x_n} \leq \sqrt[5]{3M_{min}}$, c'est-à-dire :

$$x_{n+1} \leq \sqrt[5]{3M_{min}}$$

Ceci étant vrai pour tout $n \in \mathbb{N}$, on en conclut que pour tout $n \in \mathbb{N}^*$, $x_n \leq \sqrt[5]{3M_{min}}$.
Ainsi $\sqrt[5]{3M_{min}}$ est un majorant, il est donc plus grand que le plus petit des majorants :

$$M_{min} \leq \sqrt[5]{3M_{min}}$$

Étant donné que $u_0 = 1$, nous pouvons affirmer que tous les majorants sont supérieurs ou égaux à 1, en particulier $M_{min} > 0$. Ainsi l'inégalité précédente est équivalente à

$$M_{min}^5 \leq 3M_{min} \Leftrightarrow M_{min}^4 \leq 3 \Leftrightarrow M_{min} \leq \sqrt[4]{3}$$

On peut donc conclure que si des majorants existent, alors $\sqrt[4]{3}$ en est un (ce n'est pas forcément M_{min}).

Synthèse : On vérifie que $\sqrt[4]{3}$ est bien un majorant en faisant par exemple une démonstration par récurrence. À vos stylos !

Vous voyez que sans le raisonnement par analyse-synthèse vous auriez été obligé de vous reposer, soit sur une indication de l'énoncé vous demandant de montrer que $\sqrt[4]{3}$ était un majorant, soit sur votre calculatrice pour conjecturer que 2 faisait un bon candidat.

1.3. Propriétés de $\mathbb{N}$

Vous connaissez déjà très bien $\mathbb{N}$, l'addition, la multiplication sur $\mathbb{N}$, ainsi que les relations $<$, $\leq$, $>$ et $\geq$. Nous n'allons donc pas revenir dessus.

1.3.1. Définitions générales

Commençons par des définitions qui ne sont pas propres à $\mathbb{N}$.

Définition 1.13 (Plus grand/petit élément). Soit E un ensemble possédant une relation « inférieur ou égal à », notée « $\leq$ », permettant de comparer les éléments entre eux. On appelle *plus grand (resp. petit) élément de E*, un élément x de E (s'il existe) tel que pour tout élément y de E, y est inférieur (resp. supérieur) ou égal à x. En notation mathématiques, cela donne « $\forall y \in E,\ y \leq x$ » (resp. « $\forall y \in E,\ x \leq y$ »).

Remarque 1.12. Pourquoi préciser « s'il existe » ? Tout simplement parce que l'ensemble E ne possède pas forcément de plus grand élément. Considérez par exemple $E = [0;1[$. Le nombre 1 est effectivement plus grand que tous les éléments de E, mais il n'appartient pas à E.

Définition 1.14 (Ensemble des majorants/Minorants). Soit E et F deux ensembles possédant une relation « inférieur ou égal à », notée « $\leq$ », permettant de comparer les éléments entre eux et tels que $E \subset F$. On appelle *ensemble des majorants (resp. minorants) de E* (dans F), tous les éléments de F qui sont supérieurs (resp. inférieurs) à tous les éléments de E. Un tel élément est simplement appelé *majorant (resp. minorant) de E*. En notation mathématiques, dire que x est un majorant (resp. minorant) de E, revient à dire « $\forall y \in E,\ y \leq x$ » (resp. « $\forall y \in E,\ x \leq y$ »).
Si l'ensemble des majorants (resp. minorants) est non vide, on dit que E est une partie majorée (resp. minorée) de F.

Remarque 1.13. Comme le laisse entrevoir la définition, l'ensemble des majorants ou minorants peut être vide. Prenez par exemple $E = [0;+\infty[$. Il n'existe aucun élément réel qui soit supérieur à tous les éléments de E.
Remarquez qu'un majorant ou un minorant n'appartient pas forcément à l'ensemble E.

Exemple 1.15. Si on reprend l'ensemble évoqué plus haut, $E = [0;1[$:
- E ne possède pas de plus grand élément.
- 0 est le plus petit élément de E.
- $]-\infty;0]$ est l'ensemble des minorants de E (remarquez que 0 en fait parti).
- $[1;+\infty[$ est l'ensemble des majorants de E.

1.3.2. Parties de $\mathbb{N}$

Nous nous intéressons maintenant plus particulièrement à $\mathbb{N}$ et à ses sous-parties.

Axiome 1.1. Toute partie non vide de $\mathbb{N}$ possède un (unique) plus petit élément.

Démonstration. Un axiome, par définition, n'a pas besoin de démonstration. Cependant, l'unicité du plus petit élément ne fait en réalité pas partie de l'axiome. Nous allons donc le démontrer.
Pour démontrer l'unicité d'un objet, la méthode classique consiste à supposer qu'il en existe deux et à démontrer qu'ils sont nécessairement égaux.

Supposons donc que l'on a A une partie non vide de $\mathbb{N}$. Puisque A est non vide, elle possède au moins un élément a. Si A ne possède qu'un seul élément, la démonstration s'arrête là ! On suppose donc de plus que A possède au moins deux éléments et qu'ils sont tout deux un « plus petit élément de A ». Notons les a et b. Puisque a est un plus petit élément de A et que $b \in A$, nous avons $a \leq b$. De même nous avons nécessairement $b \leq a$. Cela revient à dire que $a = b$ et que le plus petit élément est forcément unique. □

Propriété 1.5. Toute partie majorée non vide de $\mathbb{N}$ possède un (unique) plus grand élément.

Démonstration. La démonstration de l'unicité du plus grand élément est la même que pour le plus petit élément. Focalisons nous donc sur l'existence.

Soit A une partie majorée non vide de $\mathbb{N}$. Puisque A est majorée, elle possède des majorants(!), notons B l'ensemble de ses majorants, qui est non vide. D'après l'axiome 1.1, B admet un plus petit élément que l'on note m. Montrons que m appartient à A. Nous allons faire un raisonnement par l'absurde.
Supposons que m n'appartient pas à A et montrons que l'on aboutit à une absurdité (contradiction). Comme m est un majorant de A, on a

$$\forall x \in A, \ x \leq m$$

Or x appartient à A et m n'appartient pas à A, ils sont donc distincts et par conséquent

$$\forall x \in A, \ x < m$$

Cela impose que $m \geq 1$ (sinon on aurait : $\forall x \in A, \ x < 0$, ce qui est impossible puisque nous sommes dans $\mathbb{N}$ et que A est non vide) et que $m-1$ appartient à $\mathbb{N}$. Nous avons donc

$$\forall x \in A, \ x \leq m-1$$

C'est la définition d'un majorant, donc $m-1$ est un majorant de A et appartient à B. Or m était censé être le plus petit élément de B. Nous arrivons ainsi à une absurdité et pouvons conclure que notre supposition d'origine ($m \notin A$) était fausse. D'où $m \in A$, m majore A, c'est bien le plus grand élément de A.
Nous avons donc montré l'existence d'un plus grand élément de A. □

1.3.3. Raisonnement par récurrence

Vous avez déjà vu le raisonnement par récurrence au lycée. Nous allons le démontrer.

Théorème 1.1 (Principe de récurrence). Soit A une partie de $\mathbb{N}$ telle que,

1. A contient 0.
2. Si A contient un entier n, alors A contient $n+1$.

Alors $A = \mathbb{N}$.

Démonstration. Nous devons montrer que $A = \mathbb{N}$. La méthode classique consiste à faire une double inclusion. Ce n'est pas ce que nous allons faire ici... Nous allons montrer que $\mathbb{N} \backslash A$ ($\mathbb{N}$ privé de A) est l'ensemble vide par un raisonnement par l'absurde et en conclure que $A = \mathbb{N}$.

Soit une partie A de $\mathbb{N}$ vérifiant les conditions de la propriété. Supposons que $\mathbb{N}\backslash A$ est non vide. En vertu de l'axiome 1.1, il possède un plus petit élément que nous noterons m. Par définition, m n'appartient pas à A. Or d'après la condition 1, 0 appartient à A, donc nécessairement, m est différent de 0 et $m-1 \in \mathbb{N}$. Un élément de $\mathbb{N}$ appartient soit à A soit à $\mathbb{N}\backslash A$, de plus $m-1$ ne peut pas appartenir à $\mathbb{N}\backslash A$, sinon m ne serait pas le plus petit élément. D'où $m-1$ appartient à A. D'après la condition 2 nous avons nécessairement $m-1+1=m$ qui appartient à A. Or m ne peut appartenir à la fois à A et à $\mathbb{N}\backslash A$. D'où l'absurdité et $\mathbb{N}\backslash A$ est l'ensemble vide. Nous en concluons que $A=\mathbb{N}$. □

Remarque 1.14. L'énoncé du théorème 1.1 peut-être légèrement modifié pour « démarrer » à partir d'un entier n_0 :
Soit A une partie de $\mathbb{N}$ telle que

1. A contient n_0.
2. Si A contient un entier $n \geq n_0$, alors A contient $n+1$.

Alors $A = [\![n_0\,; +\infty [\![$.

Propriété 1.6 (Récurrence simple). Soit $\mathscr{P}$ une propriété portant sur les entiers naturels.
Soit n_0 un entier naturel.
On suppose que $\mathscr{P}(n_0)$ est vraie.
On suppose également que pour tout entier $n \geq n_0$, si $\mathscr{P}(n)$ est vraie, alors $\mathscr{P}(n+1)$ est vraie (cela revient à dire que l'implication $\mathscr{P}(n) \Rightarrow \mathscr{P}(n+1)$ est vraie).
Alors pour tout entier $n \geq n_0$, la proposition $\mathscr{P}(n)$ est vraie.

Démonstration. Soit $\mathscr{P}$ un propriété vérifiant les conditions de l'énoncé. Soit $A = \{n \in \mathbb{N},\ \mathscr{P}(n)\}$ (A est l'ensemble des entiers naturels tels que $\mathscr{P}$ est vraie). Alors A vérifie les conditions du théorème 1.1 (ou plutôt de sa version modifiée de la remarque 1.14). Donc $A = [\![n_0\,; +\infty [\![$ et tout entier $n \geq n_0$ vérifie $\mathscr{P}$. □

Voici comment doit se faire un raisonnement par récurrence et comment il doit être rédigé.

1. On énonce clairement la propriété $\mathscr{P}$ étudiée.
2. On vérifie que la propriété $\mathscr{P}$ est vraie pour n_0 (Initialisation).
3. On se donne un entier $n \geq n_0$ quelconque et on suppose que $\mathscr{P}(n)$ est vraie (Hypothèse de récurrence). On démontre alors que $\mathscr{P}(n+1)$ est vraie (Hérédité).
4. On conclut en annonçant que, par récurrence, la propriété est vraie pour tout entier $n \geq n_0$.

Il faut voir le principe de récurrence comme une échelle que l'on essaye de monter : Si l'on est capable de poser le pied sur le premier barreau et que l'on est capable de passer d'un barreau au suivant, alors on est capable de monter l'échelle en entier. Si l'une des deux conditions manque, alors nous ne pouvons pas monter l'échelle. D'où l'importance de vérifier l'initialisation **ET** l'hérédité. Une erreur classique est d'oublier l'initialisation. L'exemple 1.16 est un exemple de cette erreur classique.

Exemple 1.16. [FAUX !] Démontrer par récurrence que pour tout $n \in \mathbb{N}^*$, on a l'égalité

$$1+2+3+\cdots+n = \frac{n^2+n+1}{2}$$

1. Soit $\mathscr{P}(n)$ la propriété définie sur $\mathbb{N}^*$ par « $1+2+3+\cdots+n = \dfrac{n^2+n+1}{2}$ ».
2. L'élève oublie l'initialisation.
3. Soit un entier $n \in \mathbb{N}^*$ tel $\mathscr{P}(n)$ est vraie. Démontrons alors que $\mathscr{P}(n+1)$ est vraie aussi. Nous devons donc

montrer que

$$1+2+3+\cdots+n+(n+1)=\frac{(n+1)^2+(n+1)+1}{2}$$

La fraction est égale à $\frac{n^2+3n+3}{2}$. Nous avons supposé que $\mathscr{P}(n)$ est vrai, donc on peut écrire :

$$\begin{aligned}1+2+\cdots+n+(n+1) &= (1+2+\cdots+n)+(n+1)\\ &= \frac{n^2+n+1}{2}+(n+1)\\ &= \frac{n^2+n+1+2n+2}{2}\\ &= \frac{n^2+3n+3}{2}\\ &= \frac{(n+1)^2+(n+1)+1}{2}\end{aligned}$$

Donc l'implication $\mathscr{P}(n) \Rightarrow \mathscr{P}(n+1)$ est vraie.

4. Nous avons donc montré par récurrence que pour tout entier $n \in \mathbb{N}^*$, $1+2+3+\cdots+n=\frac{n^2+n+1}{2}$.

Vous l'aurez compris, il y a une erreur. Un calcul simple permet de montrer que $\mathscr{P}(1)$ est fausse puisqu'elle stipule que $1=\frac{3}{2}$. Le premier barreau étant cassé, nous ne pouvons pas monter à l'échelle.

Exemple 1.17. Démontrer par récurrence que pour tout $n \in \mathbb{N}^*$, on a l'égalité

$$1+2+3+\cdots+n=\frac{n(n+1)}{2}$$

1. Soit $\mathscr{P}(n)$ la propriété définie sur $\mathbb{N}^*$ par $1+2+3+\cdots+n=\frac{n(n+1)}{2}$.
2. Nous avons $1=\frac{1(1+1)}{2}$. Donc $\mathscr{P}(1)$ est vraie.
3. Soit un entier $n \in \mathbb{N}^*$ tel $\mathscr{P}(n)$ est vraie. Démontrons alors que $\mathscr{P}(n+1)$ est vraie aussi. Nous devons donc montrer que

$$1+2+3+\cdots+n+(n+1)=\frac{(n+1)(n+2)}{2}$$

Nous avons supposé que $\mathscr{P}(n)$ est vrai, donc on peut écrire :

$$\begin{aligned}1+2+\cdots+n+(n+1) &= (1+2+\cdots+n)+(n+1)\\ &= \frac{n(n+1)}{2}+(n+1)\\ &= \frac{n(n+1)+2(n+1)}{2}\\ &= \frac{(n+1)(n+2)}{2}\end{aligned}$$

Donc l'implication $\mathscr{P}(n) \Rightarrow \mathscr{P}(n+1)$ est vraie.

4. Nous avons donc montré par récurrence que pour tout entier $n \in \mathbb{N}^*$, $1+2+3+\cdots+n=\frac{n(n+1)}{2}$.

Il arrive parfois que supposer simplement $\mathscr{P}(n)$ vraie ne suffise par à démontrer que $\mathscr{P}(n+1)$ soit vraie aussi et qu'il faut supposer en plus que $\mathscr{P}(k)$ est vraie pour tout $k \leq n$. C'est le principe de la récurrence forte.

Propriété 1.7 (récurrence avec prédécesseurs). Soit $\mathscr{P}$ une propriété portant sur les entiers naturels.
Soit n_0 et m deux entiers naturels.
On suppose que $\mathscr{P}(n_0), \mathscr{P}(n_0+1), \ldots, \mathscr{P}(n_0+m)$ sont vraies.
On suppose également que pour tout entier $n \geq n_0$, l'implication

$$(\mathscr{P}(n) \text{ ET } \mathscr{P}(n+1) \ldots \text{ ET } \mathscr{P}(n+m)) \Rightarrow \mathscr{P}(n+m+1)$$

est vraie.
Alors pour tout entier $n \geq n_0$, la proposition $\mathscr{P}(n)$ est vraie.

Démonstration. Admis. □

Remarque 1.15. Attention à l'étape d'initialisation : il faut vérifier que la propriété est vraie pour les $m+1$ premières valeurs !

Remarque 1.16. Le plus souvent $\mathscr{P}(n)$ vraie et $\mathscr{P}(n+1)$ vraie suffit à montrer que $\mathscr{P}(n+2)$ est vraie. Dans ce cas $(m=1)$ on parle de récurrence double.

Exemple 1.18. Soit (u_n) une suite réelle définie de la façon suivante : $u_0 = 1$; $u_1 = 3$ et pour tout entier n, $u_{n+2} = 3u_{n+1} - 2u_n$. Démontrer que $\forall n \in \mathbb{N},\ u_n = 2^{n+1} - 1$.

Nous allons faire la démonstration à l'aide d'une récurrence double.

1. Soit $\mathscr{P}(n)$ la propriété définie sur $\mathbb{N}$ par « $u_n = 2^{n+1} - 1$ ».
2. Nous avons $u_0 = 1 = 2^{0+1} - 1$ et $u_1 = 3 = 2^{1+1} - 1$. Donc $\mathscr{P}(0)$ et $\mathscr{P}(1)$ sont vraies.
3. Soit un entier $n \in \mathbb{N}$ tel que $\mathscr{P}(n)$ et $\mathscr{P}(n+1)$ sont vraies. Démontrons alors que $\mathscr{P}(n+2)$ est vraie aussi. Nous devons donc montrer que $u_{n+2} = 2^{n+3} - 1$. Nous avons supposé que $\mathscr{P}(n)$ et $\mathscr{P}(n+1)$ sont vraies, c'est-à-dire $u_n = 2^{n+1} - 1$ et $u_{n+1} = 2^{n+2} - 1$. Il s'en suit que :

$$\begin{aligned} u_{n+2} &= 3u_{n+1} - 2u_n \\ &= 3(2^{n+2} - 1) - 2(2^{n+1} - 1) \\ &= 3 \times 2^{n+2} - 3 - 2^{n+2} + 2 \\ &= 2 \times 2^{n+2} - 1 \\ &= 2^{n+3} - 1 \end{aligned}$$

   Donc $\mathscr{P}(n+2)$ est vraie et par conséquent l'implication $(\mathscr{P}(n) \text{ ET } \mathscr{P}(n+1)) \Rightarrow \mathscr{P}(n+2)$ aussi.
4. Nous avons donc montré par récurrence (double) que pour tout entier $n \in \mathbb{N}$, $u_n = 2^{n+1} - 1$.

2. Ensembles

> Il faut n'appeler Science que l'ensemble des recettes qui réussissent toujours. Tout le reste est littérature.
>
> *(Paul Valéry/Moralités.)*

2.1. Définitions

La définition rigoureuse d'un ensemble est hors-programme. Pourtant vous en avez déjà une connaissance intuitive et les manipulez déjà depuis longtemps. Nous allons pourtant en donner une définition, certes incorrecte, mais largement suffisante à notre niveau et bien meilleur qu'une simple « intuition ».

Définition 2.1 (Ensemble). — Soit un univers $\mathscr{U}$ contenant des objets, et des propositions vraies ou fausses à propos des éléments de $\mathscr{U}$.
Nous appelons *ensemble* la collection des éléments de $\mathscr{U}$ qui rendent vraie une proposition $\mathscr{P}$.
Si E est un ensemble et $\mathscr{P}$ la proposition permettant de construire E, on note

$$E = \{x \in \mathscr{U} \ / \ \mathscr{P}(x)\}$$

Les objets x qui vérifient $\mathscr{P}$ sont appelés *éléments* de E. On note alors $x \in E$.
Si un objet x n'est pas élément de E (il ne vérifie donc pas $\mathscr{P}$), on note $x \notin E$.

— L'ensemble qui ne contient aucun élément est appelé *ensemble vide* et est noté $\emptyset$.

— Un ensemble qui ne contient qu'un seul élément est appelé *singleton*.

Remarque 2.1. Si vous voulez comprendre pourquoi cette définition est inexacte, je vous invite à faire des recherches sur le paradoxe de Bertrand Russel et la collection $y = \{x \ / \ x \notin x\}$ dont on peut montrer que si elle existe, alors elle n'existe pas...

Remarque 2.2. Comme écrit en préambule, vous n'avez pas attendu cette définition tordue pour manipuler depuis longtemps des ensembles, en particulier les ensembles de nombre ($\mathscr{U}$ est l'univers contenant tous les nombres réels) :

— $\mathbb{N} = \{x \in \mathscr{U} \ / \ x \geq 0$ et l'écriture de x n'admet pas de décimales$\}$ (ensemble des entiers naturels).

— $\mathbb{Z} = \{x \in \mathscr{U} \ / \ $l'écriture de x n'admet pas de décimales$\}$ (ensemble des entiers relatifs).

— $\mathbb{D} = \{x \in \mathscr{U} \ / \ $la partie décimale de x est finie$\}$ (ensemble des décimaux).

— $\mathbb{Q} = \{x \in \mathscr{U} \ / \ x$ peut s'écrire comme le quotient de deux éléments de $\mathbb{Z}\}$ (ensemble des rationnels).

— $\mathbb{R} = \{x \in \mathscr{U} \ / \ x$ est dans $\mathscr{U}\}$ (ensembles des réels).

— $[-2\,;\pi[= \{x \in \mathscr{U} \ / \ -2 \leq x < \pi\}$ (intervalle).

— $\{x \in \mathscr{U} \ / \ x \in \mathbb{Z} \text{ ET } \exists n \in \mathbb{Z} \ / \ x = 2n\} = \{2n \ / \ n \in \mathbb{Z}\}$ (ensemble des nombres pairs).

Remarque 2.3. La définition 2.1 précédente est ce que l'on appelle une définition *en compréhension*. Les ensembles ayant un nombre fini d'éléments (et pas trop grand) peuvent être définis *en extension*, c'est-à-dire en énumérant tous leurs éléments (d'où la nécessité d'en avoir un nombre fini pas trop important !). Par exemple les nombres premiers (positifs) inférieurs à 20 : $\{2,3,5,7,11,13,17,19\}$. Dans ce cas, l'ordre dans lequel sont énuméré les éléments n'a pas d'importance.

Définition 2.2 (Égalité). Soit E et F deux ensembles. On dit que E et F sont *égaux* s'ils ont exactement les mêmes éléments, c'est-à-dire si :

$$\forall x,\ (x \in E \Leftrightarrow x \in F)$$

On note $E = F$.

Définition 2.3 (Inclusion). Soit E et F deux ensembles. On dit que E est *inclus* dans F (ou que F *contient* E, ou encore que E est un *sous-ensemble* de F, ou encore que E est une *partie* de F) si tous les éléments de E appartiennent à F, c'est-à-dire si :

$$\forall x,\ (x \in E \Rightarrow x \in F)$$

On note $E \subset F$.

Remarque 2.4. Il est clair que $E = F \Leftrightarrow (E \subset F \text{ ET } F \subset E)$

Remarque 2.5. [Attention !] Vous confondez souvent l'inclusion et l'appartenance, ce qui n'a strictement rien à voir :

- Un <u>élément</u> **appartient à** un <u>ensemble</u> $(x \in E)$
- Un <u>ensemble</u> **est inclus dans** un <u>ensemble</u> $(E \subset F)$

Par exemple une bille appartient à un sac de billes, mais elle n'est pas incluse dans ce sac.
Un sac de billes est inclus dans un coffre à jouets, mais il n'appartient pas à ce coffre (une bille est un jouet à part entière, mais un sac de billes n'est pas un jouet).

Remarque 2.6. [Méthode]

- Afin de montrer une inclusion $E \subset F$, on doit montrer que tout élément x de E est dans F. La rédaction est donc forcément la suivante :

  Soit $x \in E$. Montrons que $x \in F$.

  $\vdots$

  D'où $x \in F$ et $E \subset F$.
- Afin de montrer une égalité, il y deux manière de faire :

  Par équivalence : Cette méthode n'est pas recommandée car vous écrivez souvent des équivalences qui sont en réalité des implications. La rédaction est alors une succession d'équivalences :

  Soit $x \in \mathscr{U}$.

$$\begin{aligned} x \in E &\Leftrightarrow \dots \\ &\Leftrightarrow \dots \\ &\vdots \\ &\Leftrightarrow x \in F \end{aligned}$$

Par double inclusion : Comme son nom l'indique, on utilise la remarque 2.4 en montrant d'abord que $E \subset F$ puis que $F \subset E$:

$\boxed{E \subset F}$ **:** Soit $x \in E$. ... Donc $x \in F$.

$\boxed{F \subset E}$ **:** Soit $x \in F$. ... Donc $x \in E$.

Définition 2.4 (Ensemble des parties). Soit E un ensemble. L'ensemble des parties de E est noté $\mathscr{P}(E)$. Ainsi, pour tout ensemble F : $F \in \mathscr{P}(E) \Leftrightarrow F \subset E$

Remarque 2.7. [Attention !] Ouch ! Nous avons utilisé le symbole « $\in$ » avec un ensemble ! Oui, car cette fois cet ensemble est considéré comme un objet (élément) de l'ensemble des parties.

Propriété 2.1. Pour tout ensemble, nous avons toujours $E \in \mathscr{P}(E)$ et $\emptyset \in \mathscr{P}(E)$.

Démonstration. —

$$\forall x,\ (x \in E \Rightarrow x \in E)$$

Donc par définition, $E \subset E$ et $E \in \mathscr{P}(E)$.

— De même :

$$\underbrace{\forall x,\ \left(\underbrace{\underbrace{x \in \emptyset}_{\text{Faux...}} \Rightarrow x \in E}_{\text{...donc vrai...}}\right)}_{\text{... donc vrai !}}$$

Donc par définition, $\emptyset \subset E$ et $\emptyset \in \mathscr{P}(E)$.

Exemple 2.1.

$$\mathscr{P}(\{0,1,2\}) = \left\{ \underbrace{\emptyset}_{0 \text{ élément}}, \underbrace{\{0\},\{1\},\{2\}}_{1 \text{ élément}}, \underbrace{\{0,1\},\{0,2\},\{1,2\}}_{2 \text{ éléments}}, \underbrace{\{0,1,2\}}_{3 \text{ éléments}} \right\}$$

2.2. Opérations

Définition 2.5 (Intersection et Union). Soit E et F deux ensembles.

— On appelle *intersection* de E et F l'ensemble noté $E \cap F$ des éléments qui sont à la fois dans E et dans F. C'est-à-dire

$$(x \in E \cap F) \Leftrightarrow (x \in E \ \text{ET}\ x \in F)$$

— On appelle *union* de E et F l'ensemble noté $E \cup F$ des éléments qui sont dans E ou dans F. C'est-à-dire

$$(x \in E \cup F) \Leftrightarrow (x \in E \ \text{OU}\ x \in F)$$

Remarque 2.8. Vous aurez noté l'analogie à la fois dans le sens et dans la notation entre

- l'intersection notée $\cap$ et le ET logique noté aussi $\wedge$.
- l'union notée $\cup$ et le OU logique noté aussi $\vee$.

Cette définition peut être généralisée à plus de deux ensembles.

Définition 2.6 (Intersection et Union d'une famille d'ensembles). Soit I un ensemble et $(E_i)_{i\in I}$ une famille d'ensembles.

- On appelle *intersection* des E_i, i décrivant I, notée $\bigcap\limits_{i\in I} E_i$, l'ensemble des éléments qui sont dans tous les E_i. C'est-à-dire
$$\left(x \in \bigcap_{i\in I} E_i\right) \Leftrightarrow (\forall i \in I,\ x \in E_i)$$
- On appelle *union* des E_i, i décrivant I, notée $\bigcup\limits_{i\in I} E_i$, l'ensemble des éléments qui sont dans au moins l'un des E_i. C'est-à-dire
$$\left(x \in \bigcup_{i\in I} E_i\right) \Leftrightarrow (\exists i \in I,\ x \in E_i)$$

Remarque 2.9.

- En général, l'ensemble I de la définition précédente est un ensemble d'entiers naturels, mais pas forcément. Il n'est même pas obligé d'être discret !
- On remarquera cette fois que l'intersection fait appel à la notion de « Pour tout » ($\forall$) et que l'union fait appel à la notion de « Il existe » ($\exists$).

Définition 2.7 (Ensembles disjoints). Deux ensembles E et F sont dit *disjoints* si leur intersection est vide ($E \cap F = \emptyset$). Dans ce cas, on dit que l'union $E \cup F$ est une *union disjointe* et on la note parfois $E \sqcup F$ ou encore $E \dot{\cup} F$.

Définition 2.8 (Complémentaire et Différence). Soit E un ensemble, et A et B deux sous-ensembles de E.

- On appelle *complémentaire de A dans E* l'ensemble des éléments de E qui ne sont pas dans A. Cet ensemble est noté $E\backslash A$ ou $\complement_E A$ ou $\overline{A}$ ou $\complement A$ (s'il n'y a pas d'ambigüité sur l'ensemble E pour les deux dernières notations). C'est-à-dire
$$\complement_E A = \{x \in E \ / \ x \notin A\}$$
- On appelle *différence de A dans B*, notée $B\backslash A$, l'ensemble des éléments appartenant à B, mais pas à A. C'est-à-dire
$$B\backslash A = B \cap \overline{A}$$

Définition 2.9 (Différence symétrique). Soit E et F deux ensembles. On appelle *différence symétrique de E et F*, que l'on note $E\Delta F$, l'ensemble des éléments qui sont dans un et un seul des deux ensembles E et F. C'est-à-dire

$$E\Delta F = (E\cup F)\setminus(E\cap F)$$

Définition 2.10 (n-uplet et Produit cartésien). Soit $n\in\mathbb{N}^*$, $n\geq 2$ et $E_1,E_2,\dots,E_n$, n ensembles.

— Soit $1\leq k\leq n$ et $x_k\in E_k$. On appelle *n-uplet de composantes* x_1, x_2,…,x_n (dans cet ordre) l'objet noté $(x_1,x_2,\dots,x_n)$.

— On appelle *produit cartésien* de E_1, E_2, …, E_n, que l'on note $E_1\times E_2\times\cdots\times E_n$, l'ensemble des n-uplets $(x_1,x_2,\dots,x_n)$.
Si $E_1=E_2=\cdots=E_n=E$, on note simplement E^n le produit cartésien.

Exemple 2.2. Si E et F sont deux ensembles, alors $E\times F=\{(x;y)\ /\ x\in E \text{ ET } y\in F\}$

Remarque 2.10. [Attention !] Ne pas confondre un n-uplet dans lequel l'ordre des éléments compte et un ensemble $\{x_1,x_2,\dots,x_n\}$ dans lequel l'ordre des éléments importe peu :

$$(1;2)\neq(2;1)\qquad \{1,2\}=\{2,1\}$$

Définition 2.11 (Partition). Soit E un ensemble, I un ensemble d'indices et $(E_i)_{i\in I}$ une famille de parties de E. On dit que $(E_i)_{i\in I}$ est une *partition* de E si :

— Aucun des ensembles E_i n'est vide : $\forall i\in I,\ E_i\neq\emptyset$.

— Les E_i sont disjoints deux à deux : $\forall(i;j)\in I^2,\ i\neq j\Rightarrow E_i\cap E_j=\emptyset$.

— La réunion des ensembles E_i est égale à E tout entier : $\bigcup\limits_{i\in I}E_i=E$.

Dans ces conditions, tout élément x de E appartient à un sous-ensemble E_i unique.

2.3. Propriétés

Propriété 2.2 (Règles de calculs sur l'intersection et l'union). Soit E un ensemble et A, B et C trois sous-ensembles de E.

— $A\cap B=B\cap A$ (L'intersection est *commutative*)

— $A\cap(B\cap C)=(A\cap B)\cap C$ (L'intersection est *associative*)

— $A\cap E=E\cap A=A$ (E est l'élément neutre de l'intersection)

— $A\cup B=B\cup A$ (L'union est *commutative*)

— $A\cup(B\cup C)=(A\cup B)\cup C$ (L'union est *associative*)

— $A\cup\emptyset=\emptyset\cup A=A$ ($\emptyset$ est l'élément neutre de l'union)

Démonstration. Toutes ces propriétés découlent des définitions et des mêmes propriétés sur les opérateurs logiques ET et OU . Toutefois, nous allons détailler le troisième point de l'intersection (à vous d'adapter pour l'union) :

$\boxed{(A \cap E) \subset A}$: Évident par définition de l'intersection.

$\boxed{A \subset (A \cap E)}$: Soit $x \in A$. Puisque $A \subset E$, nous avons $x \in E$. Par conséquent, il appartient à A et à E, donc à $A \cap E$. D'où l'inclusion.

Nous avons bien démontré l'égalité par double inclusion. □

Propriété 2.3 (Distributivité de l'intersection sur l'union et réciproquement). Soit E, F, G trois ensembles. Alors

— $(E \cap F) \cup G = (E \cup G) \cap (F \cup G)$ (distributivité à droite de l'union sur l'intersection).

— $E \cup (F \cap G) = (E \cup F) \cap (E \cup G)$ (distributivité à gauche de l'union sur l'intersection).

— $(E \cup F) \cap G = (E \cap G) \cup (F \cap G)$ (distributivité à droite de l'intersection sur l'union).

— $E \cap (F \cup G) = (E \cap F) \cup (E \cap G)$ (distributivité à gauche de l'intersection sur l'union).

Démonstration. Cela provient tout simplement des propriétés démontrées dans l'exercice 1.1 du chapitre 1 « Logique et raisonnement ». □

Remarque 2.11. Afin de retenir plus facilement ces quatre formules, il suffit de remplacer « $\cap$ » par « $+$ » et « $\cup$ » par « $\times$ », et inversement. En effet, vous avez depuis longtemps (sans l'avoir nommé) que la multiplication est distributive à droite et à gauche sur l'addition.

Remarque 2.12. [Généralisation] Ces propriétés peuvent se généraliser à une famille (infinie) d'ensembles :

$$E \cup \left(\bigcap_{i \in I} F_i \right) = \bigcap_{i \in I} (E \cup F_i) \qquad E \cap \left(\bigcup_{i \in I} F_i \right) = \bigcup_{i \in I} (E \cap F_i)$$

Dans le cas d'une famille finie, il suffit de faire une récurrence. Montrons la seconde égalité pour une famille quelconque.

Démonstration.

$\boxed{\subset}$: Soit $x \in E \cap \left(\bigcup_{i \in I} F_i \right)$. Il appartient donc d'une part à E, d'autre part à l'union $\bigcup_{i \in I} F_i$. Donc il existe $j \in I$, tel que $x \in F_j$. Donc, puisque x appartient aussi à E, il appartient à $E \cap F_j$. Et puisqu'il existe un $j \in I$ tel que $x \in E \cap F_j$, il appartient à l'union : $\bigcup_{i \in I} (E \cap F_i)$.

$\boxed{\supset}$: Soit $x \in \bigcup_{i \in I} (E \cap F_i)$. Donc il existe un $j \in I$, tel que $x \in E \cap F_j$. Il appartient donc d'une part à E et d'autre part à F_j. Ainsi, il appartient aussi à l'union $\bigcup_{i \in I} F_i$. Mais comme il appartient aussi à E, il appartient à l'intersection :
$E \cap \left(\bigcup_{i \in I} F_i \right)$

Propriété 2.4 (Passage au complémentaire). Soit E un ensemble et, A et B deux sous-ensembles de E.
- $\overline{\overline{A}} = A$
- $\overline{A \cup B} = \overline{A} \cap \overline{B}$
- $\overline{A \cap B} = \overline{A} \cup \overline{B}$

Démonstration. Cela découle encore des propriétés de logique, notamment des lois de De Morgan vues dans la propriété 1.2 du chapitre 1 « Logique et raisonnement ». □

Remarque 2.13. [Généralisation] Là encore, cela se généralise à une famille (infinie) d'ensembles :

$$\overline{\bigcup_{i\in I} A_i} = \bigcap_{i\in I} \overline{A_i} \qquad \overline{\bigcap_{i\in I} A_i} = \bigcup_{i\in I} \overline{A_i}$$

La démonstration est simple (la deuxième vous est laissée en exercice) :

Démonstration.

$$x \in \overline{\bigcup_{i\in I} A_i} \Leftrightarrow x \notin \bigcup_{i\in I} A_i \Leftrightarrow \forall i \in I,\ x \notin A_i \Leftrightarrow \forall i \in I,\ x \in \overline{A_i} \Leftrightarrow x \in \bigcap_{i\in I} \overline{A_i}$$

□

3. Relations binaires

L'autorité, c'est moins la qualité d'un homme qu'une relation entre deux êtres.

(Maurice Barrès.)

3.1. Vocabulaire

Définition 3.1 (Relation). On appelle *relation*, notée $\mathcal{R}$, la donnée de :

- Un ensemble de départ E (non vide).
- Un ensemble d'arrivée F (non vide).
- Un graphe $G \subset E \times F$, c'est-à-dire une partie de $E \times F$.

Soit $x \in E$ et $y \in F$. Si $(x;y) \in G$, on dit que x est *en relation avec* y pour $\mathcal{R}$ et on notera $x\mathcal{R}y$. Dans ce cas, y est appelé *image* de x par $\mathcal{R}$ et x un antécédent de y par $\mathcal{R}$.

Définition 3.2 (Relation binaire). Soit E un ensemble (non vide). On appelle *relation binaire sur* E une relation dont E est à la fois l'ensemble de départ et l'ensemble d'arrivée.

Remarque 3.1. Une relation binaire sur E est donc tout simplement un sous-ensemble de E^2.

Définition 3.3. Soit $\mathcal{R}$ une relation binaire sur un ensemble E. On dit que $\mathcal{R}$ est

Réflexive lorsque :

$$\forall x \in E,\ x\mathcal{R}x$$

Symétrique lorsque :

$$\forall (x;y) \in E^2,\ x\mathcal{R}y \Rightarrow y\mathcal{R}x$$

Antisymétrique lorsque :

$$\forall (x;y) \in E^2,\ (x\mathcal{R}y \text{ ET } y\mathcal{R}x) \Rightarrow x = y$$

Transitive lorsque :

$$\forall (x;y;z) \in E^3,\ (x\mathcal{R}y \text{ ET } y\mathcal{R}z) \Rightarrow (x\mathcal{R}z)$$

3.2. Relations d'équivalence

Définition 3.4 (Relation d'équivalence). On dit qu'une relation binaire $\mathcal{R}$ sur un ensemble E est une *relation d'équivalence* lorsqu'elle est réflexive, symétrique, transitive (« RST »).

Définition 3.5 (Classe d'équivalence). Soit $\mathcal{R}$ une relation d'équivalence sur un ensemble E et $x \in E$. On appelle *classe d'équivalence de* x l'ensemble des éléments de E en relation avec x :

$$c\ell(x) = \{y \in E \ / \ x\mathcal{R}y\}$$

Propriété 3.1. Soit E un ensemble et $\mathcal{R}$ une relation d'équivalence sur E. Alors

1. $\forall x, y \in E, \ (c\ell(x) = c\ell(y) \Leftrightarrow x\mathcal{R}y)$
2. Les classes d'équivalence forment une partition de E.

Démonstration.

1. Soit $(x;y) \in E^2$.
   $\boxed{\Rightarrow}$ $y \in c\ell(y) = c\ell(x)$, donc $x\mathcal{R}y$.
   $\boxed{\Leftarrow}$ $\boxed{\subset}$ Soit $z \in c\ell(x)$. Nous avons donc $x\mathcal{R}z$. Or par hypothèse, $x\mathcal{R}y$. Donc par symétrie et transitivité, nous avons $y\mathcal{R}z$, c'est-à-dire $z \in c\ell(y)$.
   $\boxed{\supset}$ Même démonstration en intervertissant x et y.
2. a) Elles sont clairement non vides, puisque pour tout $x \in E$, $x \in c\ell(x)$ (réflexivité).
   b) Soit $z \in c\ell(y) \cap c\ell(x)$, par transitivité et symétrie nous avons :

   $$\forall t \in E, \ t \in c\ell(x) \Leftrightarrow x\mathcal{R}t \Leftrightarrow t\mathcal{R}z \Leftrightarrow t\mathcal{R}y \Leftrightarrow t \in c\ell(y)$$

   D'où $c\ell(x) = c\ell(y)$. Les classes sont donc deux à deux disjointes.
   c) $\boxed{\subset}$ Par définition d'une classe d'équivalence, il est évident que $\bigcup_{x \in E} c\ell(x) \subset E$
   $\boxed{\supset}$ Puisque pour tout $x \in E$, nous avons $x \in c\ell(x)$ (c'est-à-dire $\{x\} \subset c\ell(x)$), nous en déduisons que :

   $$E = \bigcup_{x \in E} \{x\} \subset \bigcup_{x \in E} c\ell(x)$$

   Par double inclusion, nous avons bien $E = \bigcup_{x \in E} c\ell(x)$.

Exemple 3.1. Vous connaissez déjà des relations d'équivalence :

— L'égalité « $=$ » (entre deux nombres complexes, deux ensembles, deux fonctions, deux ...)

— La relation d'équivalence (oui !, c'est pour cela que cela s'appelle ainsi. Comme quoi les maths, c'est bien fait, non ?) « $\Leftrightarrow$ ».

— La relation de congruence modulo un réel (rencontrée depuis la seconde en trigonométrie) : Soit $\theta \in \mathbb{R}^*$ et $x, y \in \mathbb{R}$. On dit que x est congru à y modulo θ, si et seulement si :

$$x = y \mod \theta \Leftrightarrow \exists k \in \mathbb{Z}, \ x = y + k\theta$$

Remarque 3.2. Ceux qui ont fait de l'arithmétique en spécialité maths ont aussi vu la congruence sur $\mathbb{Z}$. C'est exactement la même, mais avec des entiers.

3.3. Relations d'ordre

Définition 3.6 (Relation d'ordre). Soit E un ensemble et $\mathscr{R}$ une relation binaire sur E. On dit que $\mathscr{R}$ est une *relation d'ordre* lorsqu'elle est réflexive, antisymétrique, transitive.
On les note généralement $\leq$ ou $\preccurlyeq$. On dit alors que $(E\,;\mathscr{R})$ est *un ensemble ordonné*.

Exemple 3.2. Là encore vous connaissez déjà des relations d'ordre :

— $\leq$ sur $\mathbb{R}$, mais aussi sur l'ensemble des fonctions.

— $\subset$ sur $\mathscr{P}(E)$.

— Mais aussi « l'ordre alphabétique » sur l'ensemble des mots (d'où son nom !)

De plus si $\preccurlyeq$ est une relation d'ordre, on peut toujours définir une relation d'ordre « inverse », que l'on notera $\succcurlyeq$, par

$$\forall(x;y) \in E^2,\ x \succcurlyeq y \Leftrightarrow y \preccurlyeq x$$

Définition 3.7 (Relation d'ordre stricte). Soit E un ensemble et $\mathscr{R}$ une relation binaire sur E. On dit que $\mathscr{R}$ est une *relation d'ordre stricte* lorsqu'elle est transitive et pour tout $(x;y) \in E^2$, $(x\mathscr{R}y \Rightarrow x \neq y)$

Définition 3.8 (Relation stricte associée à une relation d'ordre). Soit E un ensemble et $\preccurlyeq$ une relation d'ordre sur E. La relation $\prec$ sur E définie pour tous $x,y \in E$ par :

$$x \prec y \Leftrightarrow x \preccurlyeq y \text{ ET } x \neq y$$

est transitive et est appelée la *relation stricte associée à* $\preccurlyeq$.

Démonstration. Montrons que la relation $\prec$ ainsi définie est bien transitive (et donc une relation d'ordre stricte, puisque la condition $(x\mathscr{R}y \Rightarrow x \neq y)$ est incluse dans la définition). Soit donc $x,y,z \in E$ tels que $x \prec y$ et $y \prec z$. Montrons que $x \prec z$. Nous avons donc $x \preccurlyeq y$ et $y \preccurlyeq z$, donc par transitivité de $\preccurlyeq$, $x \preccurlyeq z$. Il ne reste plus qu'à démontrer que $x \neq z$. Supposons que $x = z$, alors $x \preccurlyeq y$ et $y \preccurlyeq z = x$, donc par antisymétrie de $\preccurlyeq$, nous avons $y = x$. Or $x \neq y$ puisque $x \prec y$. Donc $x \neq z$ et $x \prec z$.
La relation ainsi définie est donc bien une relation d'ordre stricte. $\square$

Remarque 3.3. Comme l'indique [2] sur son excellent site, la relation est aussi antisymétrique. Cependant la raison de son antisymétrie est vraiment tordue, au point que la démonstration de [2] est fausse (ce qui ne dénature en rien l'excellente qualité de ses cours !) Voyons cela en détail :
Nous devons montrer que $\forall(x;y) \in E^2$, $(x \prec y \text{ ET } y \prec x) \Rightarrow x = y$. Soit donc $(x;y) \in E^2$.

Si $x = y$: Dans ce cas, la relation $x \prec y$ est fausse, donc $(x \prec y \text{ ET } y \prec x)$ est fausse aussi et l'implication est vraie. Oui ! Rappelez-vous, si P est fausse, alors $P \Rightarrow Q$ est vraie.

Si $x \neq y$: Supposons que l'on ait $x \prec y$ et $y \prec x$, alors dans ce cas, on a $x \preccurlyeq y$ et $y \preccurlyeq x$, donc pas antisymétrie de $\preccurlyeq$, $x = y$. Ceci est absurde, donc l'hypothèse de départ ne peut être vraie. Cela veut dire que soit $x \not\prec y$, soit $y \not\prec x$

(ou les deux à la fois, voir la définition d'un ordre total 3.9).
Là encore, puisque la proposition $(x \prec y \text{ ET } y \prec x)$ est forcément fausse, l'implication est vraie !

Vous constatez que l'implication n'est vraie uniquement parce que la propriété $(x \prec y \text{ ET } y \prec x)$ est toujours fausse.

Vous constatez du même coup l'importance d'avoir mis une implication dans la définition de l'antisymétrie. Si on se ramène à la première relation antisymétrique rencontrée dans votre scolarité, c'est-à-dire « $\leq$ » (sur les réels), on serait tenté de se dire que si $x = y$, alors oui, $x \leq y$ et $y \leq x$, et donc on peut mettre une équivalence à la place de l'implication. C'est là que se ramener à des exemples simples que l'on maitrise pour comprendre une définition peut vous faire tomber dans un piège. En effet, si on met une équivalence dans la définition de l'antisymétrie, la relation d'ordre stricte associée à une relation d'ordre n'est plus antisymétrique puisque la proposition $x = y \Rightarrow (x \prec y \text{ ET } y \prec x)$ est fausse.
Une remarque sur cette remarque : d'après la définition d'un ordre stricte, rien n'oblige la relation stricte associée à être antisymétrique. Donc finalement rien n'empêche de mettre une équivalence dans la définition de l'antisymétrie. C'est vrai. Cela vous montre qu'en mathématiques, tout est question de choix dans les définitions mais surtout de cohérence des les propriétés énoncées par la suite !

Définition 3.9 (Ordre total). Soit E un ensemble et $\preccurlyeq$ une relation d'ordre sur E. On dit que $\preccurlyeq$ est une relation *d'ordre total* et que $(E\,;\preccurlyeq)$ est un ensemble *totalement ordonné* lorsque

$$\forall (x;y) \in E^2,\ x \preccurlyeq y \text{ OU } y \preccurlyeq x$$

Dans le cas contraire, l'ordre est dit partiel et $(E\,;\preccurlyeq)$ est dit ensemble *partiellement ordonné*.

Remarque 3.4. En des termes plus simples, l'ordre est total si tous les éléments peuvent se comparer entre eux. La relation « $\leq$ » est un ordre total. Cependant la relation « $\subset$ » est un ordre partiel puisque tous les ensembles ne sont pas forcément inclus les uns dans les autres.

Définition 3.10 (Majorant/Minorant). Soit $(E\,;\preccurlyeq)$ un ensemble ordonné, $A \subset E$ et $m \in E$. On dit que :

— m est un *minorant* de A si $\forall x \in A,\ m \preccurlyeq x$. On dit alors que A est *minoré*.

— m est un *majorant* de A si $\forall x \in A,\ x \preccurlyeq m$. On dit alors que A est *majoré*.

— A est *borné* si il est minoré et majoré.

Définition 3.11 (Plus petit/grand élément, minimum/maximum). Soit $(E\,;\preccurlyeq)$ un ensemble ordonné, $A \subset E$ et $m \in E$. On dit que :

— m est un *plus petit élément* ou un *minimum* de A si m est un minorant de A et $m \in A$.

— m est un *plus grand élément* ou un *maximum* de A si m est un majorant de A et $m \in A$.

Remarque 3.5. D'après la définition, il n'est pas clair qu'un minimum (ou maximum) existe toujours, ni qu'il soit unique. Par exemple, dans l'ensemble $E = (\mathbb{R}\,;\leq)$, $A =]0;1]$ est bien minoré, majoré, donc borné, admet un maximum (1), mais pas de minimum ($0 \neq]0;1]$). Quant à l'unicité. . .

Propriété 3.2 (Unicité du minimum/maximum). Soit $(E\,;\preccurlyeq)$ un ensemble ordonné et $A \subset E$.

— Si A admet un minimum, alors il est unique et on le note $\min(A)$.

— Si A admet un maximum, alors il est unique et on le note $\max(A)$.

Démonstration. Supposons que A possède deux minima m_1 et m_2. Par définition d'un minimum, nous avons $m_1 \in A$ et $m_2 \in A$. De plus, par définition d'un minorant, nous avons aussi $m_1 \preccurlyeq m_2$ et $m_2 \preccurlyeq m_1$, d'où $m_1 = m_2$ et l'unicité (s'il existe) du minimum. □

Définition 3.12 (Borne inférieure/supérieure). Soit $(E\,;\preccurlyeq)$ un ensemble ordonné et $A \subset E$.

— Si l'ensemble B des minorants de A est non vide et si B a un plus grand élément, m, alors celui-ci est appelé *borne inférieure* de A. On le note $m = \inf(A)$.

— Si l'ensemble C des majorants de A est non vide et si C a un plus petit élément, M, alors celui-ci est appelé *borne supérieure* de A. On le note $M = \sup(A)$.

Remarque 3.6. Vous remarquez qu'une borne supérieure ou inférieure n'appartient pas forcément à l'ensemble A.

Exemple 3.3. — Dans $(\mathbb{R}\,;\leq)$, $A =]0\,;1]$. Comme nous l'avons vu, A ne possède pas de minimum, son maximum est 1. Sa borne supérieure est $\sup(A) = 1$ et sa borne inférieure est $\inf(A) = 0$.

— Soit E un ensemble et, A et B deux parties de E. On munit $\mathscr{P}(E)$ de la relation d'ordre (partiel) « $\subset$ ». L'ensemble $\{A,B\}$ admet un plus petit élément (et du coup un plus grand) si et seulement si $A \subset B$ ou $B \subset A$. Dans ce cas, la borne inférieure est A (ou B) et la borne supérieure est B (ou A).
Dans tous les cas, la borne inférieure est $A \cap B$ et la borne supérieure est $A \cup B$. En effet, $A \cap B \subset A$ et $A \cap B \subset B$, donc $A \cap B$ est un minorant de $\{A,B\}$. De plus, si C est un minorant de $\{A,B\}$, alors $C \subset A$ et $C \subset B$, donc $C \subset A \cap B$. En conclusion, $A \cap B$ est bien le pus grand des minorants, soit la borne inférieure. De même pour $A \cup B$.

Propriété 3.3. Soit $(E\,;\preccurlyeq)$ un ensemble ordonné et $A \subset E$.

— Si A admet un minimum, alors A admet une borne inférieure et $\min(A) = \inf(A)$.

— Si A admet un maximum, alors A admet une borne supérieure et $\max(A) = \sup(A)$.

Démonstration. Supposons que A admette un minimum m et montrons que m est le plus grand des minorants de A. Tout d'abord, par définition d'un minimum, m est bien un minorant de A. Ensuite, soit x un minorant de A, par définition et puisque $m \in A$ (c'est un minimum), nous avons $x \preccurlyeq m$. Ainsi, m est bel et bien le plus grand des minorants, soit la borne inférieure de A. □

4. Applications

4.1. Fondamentaux

Définition 4.1 (Application). Soit $\mathscr{R}$ une relation d'ensemble de départ E, d'arrivée F et de graphe G. La relation $\mathscr{R}$ est appelée *application* si et seulement si tout antécédent x de E possède une unique image y de F.
Dans ce cas, au lieu d'écrire $x\mathscr{R}y$, on écrira $y = \mathscr{R}(x)$ et y est l'image de x par $\mathscr{R}$.
L'ensemble des applications de E vers F est noté $\mathscr{F}(E, F)$ ou encore F^E.

Remarque 4.1.

- Les applications ressemblent énormément aux fonctions. La seule différence (qui n'est pas au programme, donc vous avez le droit de les confondre) est que pour une application, tout élément de E possède une et une seule image (l'application est définie partout), alors que pour une fonction, tout élément de E possède au plus une image (donc soit pas d'image, soit une seule, elle n'est pas forcément définie partout).
- Pour désigner une application (donc une fonction), on utilise en générale une lettre minuscule et on écrit :

$$f : \left\{ \begin{array}{ccc} E & \longrightarrow & F \\ x & \longmapsto & y = f(x) \end{array} \right.$$

- La nature des ensembles E et F n'est pas précisée. Ce ne sont donc pas forcément des ensembles de réels. On peut avoir des applications sur n'importe quels objets mathématiques (des vecteurs, des matrices, des applications...). De même pour l'ensemble d'arrivé. Ainsi, on peut avoir une fonction qui à un vecteur associe une fonction.

Définition 4.2 (Égalité entre applications). Soit E, F deux ensembles et f, g deux applications de E dans F. On dit que f et g sont *égales* si et seulement si

$$\forall x \in E, f(x) = g(x)$$

Dans ce cas, on note $f = g$.

Remarque 4.2. Oui, vous savez déjà ce que veut dire l'égalité entre deux fonctions, mais nous sommes là pour être rigoureux et définir correctement toutes les notations utilisées !

Définition 4.3 (Famille). Soit E un ensemble et I un ensemble dont les éléments sont appelés *indices*. On appelle *famille d'éléments de E indexée par I* toute application de I dans E.

Remarque 4.3.

- Vous remarquez que finalement une famille est une application et une application est une famille. C'est juste deux manières différentes de voir le même objet mathématique, mais aussi deux manières différentes de le noter : au lieu de noter une famille comme une fonction on notera plutôt $(x_i)_{i\in I} \in E^I$.
- Nous avons déjà utilisé, de manière purement intuitive, la notion de famille d'ensembles dans le chapitre précédent « Ensembles ». C'est finalement une application de I dans $\mathscr{P}(E)$.

Exemple 4.1. Vous avez déjà rencontré cette notation pour les suites réelles : $(U_n)_{n\in\mathbb{N}}$. Une suite réelle est donc tout simplement une fonction f de $\mathbb{N}$ dans $\mathbb{R}$, telle que $f(n) = U_n$.
Encore plus simplement, un famille $(x_1,\dots,x_n)$ d'éléments de E est aussi une application g de $[\![1;n]\!]$ dans E telle que $g(1) = x_1,\dots,g(n) = x_n$.

Définition 4.4 (Fonction identité). Soit E un ensemble. On appelle *identité de E* l'application de E dans E définie par

$$\forall x \in E,\ Id_E(x) = x$$

Remarque 4.4. Cette application peut sembler inutile puisqu'elle ne fait rien. Cependant, nous en avons besoin par exemple pour déterminer la dérivée d'une fonction réciproque. De plus, vous la connaissez déjà ! Dans $\mathbb{R}$, c'est la fonction dont la courbe représentative (le graphe, donc) est la droite d'équation $y = x$.

Définition 4.5 (Restriction, prolongement). Soit E, F deux ensembles, $A \subset E$ et $f \in F^E$.

- La *restriction de f à A* est l'application :

$$f_{|A} : \left\{ \begin{array}{ccc} A & \longrightarrow & F \\ x & \longmapsto & f(x) \end{array} \right.$$

- Si g est une application de A dans F, on dit que f est un *prolongement de g à E* si et seulement si $f_{|A} = g$.

4.2. Composition

Définition 4.6 (Composition). Soit E, F, G trois ensembles et f, g deux applications : $f \in F^E$ et $g \in G^F$. On appelle *composée* des applications f et g, l'application de E dans G, notée $g \circ f$, définie par :

$$\forall x \in E,\ g\circ f(x) = g(f(x))$$

Remarque 4.5. Afin de ne pas se perdre entre les ensembles d'arrivées, de départs et le sens de la composition (surtout si on cumule plusieurs compositions), on peut avoir recours à un *diagramme commutatif* :

$$E \xrightarrow{\ f\ } F \xrightarrow{\ g\ } G$$

$$x \longmapsto f(x) \longmapsto g(f(x))$$

Propriété 4.1 (Associativité). La composition des applications est *associative*, c'est-à-dire, soit E, F, G, H quatre ensembles et soit $f \in F^E$, $g \in G^F$, $h \in H^G$ alors

$$h \circ (g \circ f) = (h \circ g) \circ f$$

Démonstration. Soit $x \in E$, $y = f(x)$, $z = g(y)$, $t = h(z)$, $gf = g \circ f$ et $hg = h \circ g$. Nous avons $gf(x) = g \circ f(x) = g(y) = z$, d'où $h \circ (g \circ f)(x) = h \circ gf(x) = h(gf(x)) = h(z) = t$. De même, $hg(y) = h \circ g(y) = h(g(y)) = h(z) = t$ et $(h \circ g) \circ f(x) = hg \circ f(x) = hg(y) = t$. D'où l'égalité quelque soit $x \in E$, donc des deux applications. □

Remarque 4.6. Cette propriété 4.1, permet d'alléger l'écriture en retirant les parenthèses (puisqu'il n'y a pas d'ambigüité) : $h \circ (g \circ f) = h \circ g \circ f$.

Propriété 4.2 (Élément neutre de la composition). Soit E, F deux ensembles et f une application de E dans F. Alors

$$Id_F \circ f = f \circ Id_E = f$$

On dit que l'identité est un élément neutre pour la composition.

Démonstration. Évident. □

4.3. Fonction indicatrice

Définition 4.7 (Fonction indicatrice). Soit E un ensemble et $A \subset E$. On appelle *fonction indicatrice* ou *fonction caractéristique* de A, l'application définie sur E par :

$$\mathbb{1}_A : \left\{ \begin{array}{ccl} E & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \left\{ \begin{array}{ll} \mathbb{1}_A(x) = 1 & \text{si } x \in A \\ \mathbb{1}_A(x) = 0 & \text{si } x \notin A \end{array} \right. \end{array} \right.$$

Remarque 4.7. La fonction indicatrice nous permet simplement de savoir si un élément appartient à un ensemble ou pas.

Propriété 4.3. Soit E un ensemble et A, B deux sous-ensembles de E. Alors

$$\mathbb{1}_A = \mathbb{1}_B \Leftrightarrow A = B$$

Démonstration.

$\boxed{\Rightarrow}$ **:**

$\boxed{A \subset B}$ **:** Soit $x \in A$. Nous avons $\mathbb{1}_B(x) = \mathbb{1}_A(x) = 1$, donc $x \in B$.

$\boxed{B \subset A}$ **:** Même raisonnement.

$\boxed{\Leftarrow}$ **:** Évident.

Propriété 4.4. Soit E un ensemble et A, B deux sous-ensembles. Alors

1. $\mathbb{1}_A^2 = \mathbb{1}_A$
2. $\mathbb{1}_{\overline{A}} = 1 - \mathbb{1}_A$
3. $\mathbb{1}_{A\cap B} = \mathbb{1}_A \mathbb{1}_B$
4. $\mathbb{1}_{A\cup B} = \mathbb{1}_A + \mathbb{1}_B - \mathbb{1}_A \mathbb{1}_B$

Démonstration.
1. Évident du fait que $0^2 = 0$ et $1^2 = 1$.
2. Cela provient simplement du fait que $1-0=1$ et $1-1=0$ (C'est d'ailleurs une astuce souvent utilisée en informatique pour échanger 1 et 0).
3. Si $x \in A\cap B$, alors $\mathbb{1}_{A\cap B}(x) = 1$, mais on a aussi $x \in A$ et $x \in B$, d'où $\mathbb{1}_A(x)\mathbb{1}_B(x) = 1\times 1 = 1$. Si $x \notin A\cap B$, alors $\mathbb{1}_{A\cap B}(x) = 0$. Mais si $x \notin A\cap B$, alors soit $x \notin A$, soit $x \notin B$, dans les deux cas, le produit $\mathbb{1}_A(x)\mathbb{1}_B(x)$ est nul. D'où l'égalité des fonctions.
4. Trois cas de figure se présentent :
   — Soit $x \in A\cap B$, dans ce cas, $x \in A\cup B$ et $\mathbb{1}_{A\cup B}(x) = 1$, mais on a aussi $\mathbb{1}_A(x)+\mathbb{1}_B(x)-\mathbb{1}_{A\cap B}(x) = 1+1-1 = 1$. D'où l'égalité.
   — Soit x appartient à l'un et un seul des deux ensembles (et donc pas à l'intersection). Dans ce cas $x \in A\cup B$ et $\mathbb{1}_{A\cup B}(x) = 1$. Mais on a aussi $\mathbb{1}_A(x)+\mathbb{1}_B(x)-\mathbb{1}_{A\cap B}(x) = 1+0-0$. D'où l'égalité.
   — Soit x n'appartient à aucun des deux ensembles. Dans ce cas, $\mathbb{1}_{A\cup B}(x) = 0$ et $\mathbb{1}_A(x)+\mathbb{1}_B(x)-\mathbb{1}_{A\cap B}(x) = 0+0-0 = 0$. D'où l'égalité.

   D'où l'égalité des fonctions.

Exemple 4.2. En se basant sur la propriété 4.3, les fonctions indicatrices permettent de démontrer très simplement l'égalité entre deux ensembles. Par exemple : Si A, B et C sont trois parties d'un ensemble E, alors

$$\begin{aligned}\mathbb{1}_{A\cap(B\cup C)} &= \mathbb{1}_A\mathbb{1}_{B\cup C} \\ &= \mathbb{1}_A\left(\mathbb{1}_B+\mathbb{1}_C-\mathbb{1}_B\mathbb{1}_C\right) \\ &= \mathbb{1}_A\mathbb{1}_B+\mathbb{1}_A\mathbb{1}_C-\mathbb{1}_A\mathbb{1}_B\mathbb{1}_C\end{aligned} \qquad \begin{aligned}\mathbb{1}_{(A\cap B)\cup(A\cap C)} &= \mathbb{1}_{A\cap B}+\mathbb{1}_{A\cap C}-\mathbb{1}_{A\cap B}\mathbb{1}_{A\cap C} \\ &= \mathbb{1}_A\mathbb{1}_B+\mathbb{1}_A\mathbb{1}_C-\mathbb{1}_A\mathbb{1}_B\mathbb{1}_A\mathbb{1}_C \\ &= \mathbb{1}_A\mathbb{1}_B+\mathbb{1}_A\mathbb{1}_C-\mathbb{1}_A\mathbb{1}_B\mathbb{1}_C\end{aligned}$$

D'où l'égalité $A\cap(B\cup C) = (A\cap B)\cup(A\cap C)$.

4.4. Image directe et réciproque

Définition 4.8 (Image directe). Soit E, F deux ensembles, $f \in F^E$ et $X \subset E$. On appelle *image de X par f*, le sous-ensemble de F défini par :

$$f(X) = \{y \in F \ / \ \exists x \in X,\ f(x) = y\}$$

Remarque 4.8. L'image de X par f est tout simplement l'ensemble des images des éléments x de X par f.

Propriété 4.5. Soit E, F deux ensembles, $f \in F^E$ et A, B deux parties de E.

1. Si $A \subset B$, alors $f(A) \subset f(B)$.
2. $f(A\cap B) \subset f(A)\cap f(B)$.
3. $f(A\cup B) = f(A)\cup f(B)$.

Démonstration.

1. Supposons que $A \subset B$. Soit $y \in f(A)$, montrons que $y \in f(B)$. Puisque $y \in f(A)$, il existe $x \in A$ tel que $f(x) = y$. Or $x \in A \subset B$, donc il existe $x \in B$ tel que $f(x) = y$. D'où $y \in f(B)$.
2. Voir correction du TD.
3. Voir correction du TD.

Définition 4.9 (Image réciproque). Soit E, F deux ensembles, $f \in F^E$ et $Y \subset F$. On appelle *image réciproque de Y par f* le sous-ensemble de E défini par :

$$f^{-1}(Y) = \{x \in E \ / \ f(x) \in Y\}$$

Remarque 4.9. L'image réciproque de Y est tout simplement l'ensemble des antécédents qui ont une image dans Y. Donc ce que l'on note pour l'instant f^{-1} n'est PAS une application, $f^{-1}(Y)$ est un ensemble (contrairement à ce que vous avez peut-être déjà vu et que l'on verra plus tard...)

Propriété 4.6. Soit E, F deux ensembles, $f \in F^E$ et C, D deux parties de F.

1. Si $C \subset D$, alors $f^{-1}(C) \subset f^{-1}(D)$.
2. $f^{-1}(C \cap D) = f^{-1}(C) \cap f^{-1}(D)$.
3. $f^{-1}(C \cup D) = f^{-1}(C) \cup f^{-1}(D)$.

Démonstration.

1. Supposons que $C \subset D$. Soit $x \in f^{-1}(C)$, montrons que $x \in f^{-1}(D)$. Puisque $x \in f^{-1}(C)$, $f(x) \in C$. Mais comme $C \subset D$, $f(x) \in D$, ainsi, $x \in f^{-1}(D)$.
2. Voir correction du TD.
3. Voir correction du TD.

4.5. Injections, Surjections, Bijections

Définition 4.10 (Injection, Surjecction, Bijection). Soit E, F deux ensembles et $f \in F^E$. On dit que :

— f est *injective* si et seulement si $\forall x_1, x_2 \in E,\ x_1 \neq x_2 \Rightarrow f(x_1) \neq f(x_2)$
si et seulement si $\forall x_1, x_2 \in E,\ f(x_1) = f(x_2) \Rightarrow x_1 = x_2$

— f est *surjective* si et seulement si $\forall y \in F,\ \exists x \in E,\ f(x) = y$

— f est *bijective* si et seulement si f est injective et surjective
si et seulement si $\forall y \in F,\ \exists! x \in E,\ f(x) = y$

Remarque 4.10.

— La fonction f est injective signifie que chaque élément de F a au plus un antécédent.
— La fonction f est surjective signifie que chaque élément de F a au moins un antécédent.
— La fonction f est bijective signifie que chaque élément de F a exactement un antécédent.

Remarque 4.11. Une fonction est toujours surjective sur son ensemble image. Ainsi, une fonction injective est toujours bijective sur son ensemble image.

Exemple 4.3. — L'exemple le plus simple de fonction bijective est tout simplement la fonction identité Id_E.

— L'injectivité, la surjectivité et donc la bijectivité d'une fonction dépend grandement des ensembles E et F.

La fonction $f : \left\{ \begin{array}{ccc} \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & x^2 \end{array} \right.$ n'est pas injective car $f(-1) = f(1)$.

n'est pas surjective car $y < 0$ n'a pas d'antécédent.
n'est pas bijective car ni injective ni surjective.

$g = f_{\mathbb{R}_+} : \left\{ \begin{array}{ccc} \mathbb{R}_+ & \longrightarrow & \mathbb{R} \\ x & \longmapsto & x^2 \end{array} \right.$ est injective car nous avons restreint l'ensemble des antécédent à $\mathbb{R}_+$ (Nous aurions pu le restreindre à $\mathbb{R}_-$ ou à toute autre partie de $\mathbb{R}_-$ ou $\mathbb{R}_+$).
n'est pas surjective car $y < 0$ n'a pas d'antécédent.
n'est pas bijective car pas surjective.

$h : \left\{ \begin{array}{ccc} \mathbb{R} & \longrightarrow & \mathbb{R}_+ \\ x & \longmapsto & x^2 \end{array} \right.$ n'est pas injective.

est surjective car nous avons restreint l'ensemble des images à $\mathbb{R}_+$.
n'est pas bijective car pas injective.

$c : \left\{ \begin{array}{ccc} \mathbb{R}_+ & \longrightarrow & \mathbb{R}_+ \\ x & \longmapsto & x^2 \end{array} \right.$ est injective.

est surjective.
est bijective car injective et surjective.

Propriété 4.7. Soit E, F, G trois ensembles et $f \in F^E$, $g \in G^F$

1. Si $g \circ f$ est injective, alors f est injective.
2. si $g \circ f$ est surjective alors g est surjective.

Démonstration.

1. Soit x_1 et x_2 dans E tels que $f(x_1) = f(x_2)$. Montrons que $x_1 = x_2$. De l'égalité $f(x_1) = f(x_2)$ nous déduisons que $g(f(x_1)) = g(f(x_2))$, c'est-à-dire $g \circ f(x_1) = g \circ f(x_2)$. Par injectivité de $g \circ f$, nous déduisons que $x_1 = x_2$.
2. Soit $z \in G$, montrons qu'il possède un antécédent par g dans F. La surjectivité de $g \circ f$ nous donne l'existence de $x \in E$, tel que $g \circ f(x) = z$. En posant $y = f(x) \in F$, nous avons l'existence de $y \in F$ tel que $g(y) = z$, d'où la surjectivité de g.

Définition 4.11 (Bijection réciproque). Soit E, F deux ensembles, $f \in F^E$ une application bijective. On appelle *application réciproque de* f l'application de F dans E qui à tout élément de F associe son unique antécédent dans E.

$$f^{-1} : \left\{ \begin{array}{ccc} F & \longrightarrow & E \\ y & \longmapsto & x \text{ tel que } f(x) = y \end{array} \right.$$

Remarque 4.12. Soit f une fonction bijective de E dans F.

— Soit $x \in E$ et $y = f(x)$ son image. Donc par définition, $f^{-1}(y) = x$. Ainsi $f^{-1}(f(x)) = f^{-1}(y) = x$. Nous avons montré que $f^{-1} \circ f = Id_E$.

— Soit $y \in F$ et x son unique antécédent. Donc par définition $f(x) = y$ et $f^{-1}(y) = x$. Ainsi $f\left(f^{-1}(y)\right) = f(x) = y$. Nous avons montré que $f \circ f^{-1} = Id_F$.

Remarque 4.13. Si f est bijective, alors pour tous $x \in E$ et $y \in F$,

$$y = f(x) \Leftrightarrow x = f^{-1}(y)$$

Si f est une fonction de $\mathbb{R}$ dans $\mathbb{R}$, cette équivalence signifie que la courbe représentative de f et celle de f^{-1} sont symétriques l'une de l'autre par rapport à la droite d'équation $y = x$ (première bissectrice).

Propriété 4.8. Soit E, F deux ensembles et $f \in F^E$.
La fonction f est bijective si et seulement si il existe $g \in E^F$ telle que $f \circ g = Id_F$ et $g \circ f = Id_E$.
Dans ce cas, la fonction g est égale à la réciproque de f.

Démonstration. Condition nécessaire : Cela vient tout simplement de la définition de la réciproque f^{-1} qui vérifie $f \circ f^{-1} = Id_F$ et $f^{-1} \circ f = Id_E$. On pose alors $g = f^{-1}$.

Condition suffisante : Nous avons $f \circ g = Id_F$ or Id_F est surjective, donc d'après la propriété 4.7, f est surjective. De même, $g \circ f = Id_E$ est injective donc f est injective. En conclusion, f est bijective.

Puisque f est bijective, elle possède une réciproque f^{-1}. En composant à gauche par f^{-1} l'égalité $f \circ g = Id_F$, nous obtenons $g = Id_E \circ g = (f^{-1} \circ f) \circ g = f^{-1} \circ (f \circ g) = f^{-1} \circ Id_F = f^{-1}$. La réciproque est donc l'unique fonction vérifiant les deux égalités (et elle est unique, au cas où nous ayons pas assez insisté là-dessus). $\square$

Propriété 4.9. Soit E, F et G trois ensembles, et $f \in F^E$, $g \in G^F$ deux fonctions.

1. Si f et g sont injectives alors $g \circ f$ est injective.
2. Si f et g sont surjectives alors $g \circ f$ est surjective.
3. Si f et g sont bijectives alors $g \circ f$ est bijective. Dans ce cas, $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$

Démonstration.

1. Soit x_1 et x_2 deux éléments de E tels que $g \circ f(x_1) = g \circ f(x_2)$. Par injectivité de g, nous avons donc $f(x_1) = f(x_2)$. Par injectivité de f, nous en déduisons que $x_1 = x_2$. Donc $g \circ f$ est injective.
2. Soit $z \in G$. Par surjectivité de g, il existe $y \in F$ tel que $g(y) = z$. Par surjectivité de f, il existe $x \in E$ tel que $f(x) = y$. Finalement, nous avons $g \circ f(x) = z$, d'où la surjectivité de $g \circ f$.
3. La bijectivité de $g \circ f$ est immédiate à partir des points précédents. Pour le second point : $\left(f^{-1} \circ g^{-1}\right) \circ (g \circ f) = f^{-1} \circ g^{-1} \circ g \circ f = f^{-1} \circ f = Id_E$. De même, $(g \circ f) \circ \left(f^{-1} \circ g^{-1}\right) = g \circ f \circ f^{-1} \circ g^{-1} = g \circ g^{-1} = Id_F$. Donc d'après la propriété 4.8, nous avons bien $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$.

Propriété 4.10 (Réciproque et image réciproque). Soit E, F deux ensembles, $f \in F^E$ bijective et $B \subset F$. Alors :

$$f^{\leftarrow}(B) = f^{-1}(B)$$

où l'on note $f^{\leftarrow}(B)$ l'image *réciproque* de B par f et $f^{-1}(B)$ l'image *directe* de B par f^{-1}.

Remarque 4.14. Avant de démontrer cette propriété, arrêtons nous un instant afin de bien comprendre ce qui est écrit.

D'abord nous avons déjà vu que si A est un ensemble, alors on note $f(A)$ l'ensemble des images des éléments de A par f, c'est ce que l'on appelle l'image directe de A par f. Ainsi, il est normal de noter l'image directe de B par f^{-1}, $f^{-1}(B)$.

Ensuite, nous avons défini pour une partie B de F son image réciproque par f, comme étant l'ensemble des antécédents (s'ils existent) des éléments de B. Nous avons noté cet ensemble $f^{-1}(B)$.

Vous remarquez que nous avons la même notation pour deux choses bien distinctes (en général), ce qui pose problème. C'est pour cela que dans le cadre de cette propriété, le second ensemble est noté $f^{\leftarrow}(B)$. Ce que dit cette propriété, c'est que lorsque f est bijective, finalement ces deux ensembles sont égaux et il n'y a donc pas d'ambigüité dans la notation.

Sachant que l'application f^{-1} n'est définie que lorsque f est bijective, il n'y a donc jamais ambigüité :

— Si f est quelconque, alors $f^{-1}(B)$ représente l'image réciproque de B par f.

— Si f est bijective, alors $f^{-1}(B)$ représente l'image réciproque de B par f, qui est égale à l'image directe de B par f^{-1}.

Démonstration. Soit $x \in E$:

$$\begin{aligned} x \in f^{-1}(B) &\Leftrightarrow \exists b \in B \ / \ x = f^{-1}(b) \Leftrightarrow \exists b \in B \ / \ f(x) = b \\ &\Leftrightarrow f(x) \in B \Leftrightarrow x \in f^{\leftarrow}(B) \end{aligned}$$

Définition 4.12 (Involution). Soit E un ensemble et $f \in E^E$. On dit que f est *involutive* (ou que f est une *involution* de E) lorsque $f \circ f = Id_E$.
Dans ce cas, f est bijective et elle est sa propre réciproque : $f^{-1} = f$.

Exemple 4.4. La conjugaison dans $\mathbb{C}$ est une involution de $\mathbb{C}$, les symétries axiales et ponctuelles sont des involutions du plan et la fonction inverse : $f : \left\{ \begin{array}{ccc} \mathbb{R}^* & \longrightarrow & \mathbb{R}^* \\ x & \longmapsto & \frac{1}{x} \end{array} \right.$ est une involution de $\mathbb{R}^*$.

5. Nombres complexes

Dans ce chapitre sur les nombres complexes, nous allons partir du principe que vous les avez déjà vu et acquis en terminale. Nous nous contenterons de faire de brefs rappels dans la première partie.

5.1. Rappels de Terminale

Théorème 5.1 (Et définitions).

— L'ensemble des nombres complexes est noté $\mathbb{C}$ et contient l'ensemble des réels $\mathbb{R}$ ainsi qu'un certain élément i pour lequel $i^2 = -1$.

— Tout nombre complexe z s'écrit d'une et une seule manière sous la forme dite *algébrique* : $z = a + ib$ pour certains $a, b \in \mathbb{R}$. Le réel a est appelé *partie réelle* de z et noté $\operatorname{Re}(z)$, le réel b est appelé *partie imaginaire* de z et noté $\operatorname{Im}(z)$.

— Les réels sont exactement les nombres complexes de partie imaginaire nulle. Enfin, un nombre complexe de partie réelle nulle est appelé un *imaginaire pur*.

Remarque 5.1. L'unicité de la forme algébrique d'un nombre complexe est utilisée fréquemment pour faire des identifications. Nous avons $z = a + ib = z' = a' + ib' \Leftrightarrow a = a'$ et $b = b'$. Ainsi une égalité sur des nombres complexes est ramenée à deux égalités sur les réels.

Théorème 5.2 (Et définitions). L'ensemble $\mathbb{C}$ est muni de deux opérations d'addition et de multiplication qui généralisent celles que nous connaissons sur $\mathbb{R}$. Précisément, pour tout $z, z' \in \mathbb{C}$:

— $z + z' = (\operatorname{Re}(z) + \operatorname{Re}(z')) + i\,(\operatorname{Im}(z) + \operatorname{Im}(z'))$

— $zz' = (\operatorname{Re}(z)\operatorname{Re}(z') - \operatorname{Im}(z)\operatorname{Im}(z')) + i\,(\operatorname{Re}(z)\operatorname{Im}(z') + \operatorname{Im}(z)\operatorname{Re}(z'))$

D'où :

— $\operatorname{Re}(z+z') = \operatorname{Re}(z) + \operatorname{Re}(z')$

— $\operatorname{Im}(z+z') = \operatorname{Im}(z) + \operatorname{Im}(z')$

— $\operatorname{Re}(zz') = \operatorname{Re}(z)\operatorname{Re}(z') - \operatorname{Im}(z)\operatorname{Im}(z')$

— $\operatorname{Im}(zz') = \operatorname{Re}(z)\operatorname{Im}(z') + \operatorname{Im}(z)\operatorname{Re}(z')$

Enfin, pour tout $z = x + iy \in \mathbb{C}^*$ avec $x, y \in \mathbb{R}$: $\frac{1}{z} = \frac{1}{x+iy} = \frac{x-iy}{x^2+y^2}$.

Définition 5.1 (Affixe et image). On munit le plan d'un repère orthonormal direct $\left(O; \vec{i}\,; \vec{j}\right)$.

— Pour tout $z = x + iy \in \mathbb{C}$ avec $x, y \in \mathbb{R}$, le point M du plan de coordonnées $(x; y)$ est appelé l'*image* de z tandis que z est appelé l'*affixe* de M. On dit aussi que z est l'*affixe* du vecteur du plan de coordonnées $(x; y)$.

— **Règles de calcul sur les affixes**

  — Pour tout vecteur $\vec{u}$ et $\vec{v}$ du plan d'affixes respectifs u et v et pour tout $\lambda, \mu \in \mathbb{R}$, le vecteur $\lambda\vec{u} + \mu\vec{v}$ a pour affixe $\lambda u + \mu v$.

  — Pour tout point A et B du plan d'affixes respectifs a et b, le vecteur $\overrightarrow{AB}$ a pour affixe $b - a$.

Définition 5.2 (Conjugué, module). Soi $z \in \mathbb{C}$.

— On appelle *conjugué* de z le nombre complexe $\overline{z} = \mathrm{Re}(z) - i\,\mathrm{Im}(z)$.

— On appelle *module* de z le réel positif ou nul $|z| = \sqrt{\mathrm{Re}(z)^2 + \mathrm{Im}(z)^2}$.

Remarque 5.2.

— Module et valeur absolue coïncident sur $\mathbb{R}$ car pour tout $x \in \mathbb{R}$, $|x| = \sqrt{x^2}$.

— De par sa définition, le module $|z|$ s'interprète comme norme du vecteur d'affixe z. De même, pour tout $a, b \in \mathbb{C}$ d'images A, B, le module $|a - b|$ est autre que la distance AB. Il en découle que pour tout $R > 0$:

  — le cercle de centre A et de rayon R est l'ensemble $\{z \in \mathbb{C} \ / \ |z - a| = R\}$

  — le disque fermé de centre A et de rayon R est l'ensemble $\{z \in \mathbb{C} \ / \ |z - a| \leq R\}$

  — le disque ouvert de centre A et de rayon R est l'ensemble $\{z \in \mathbb{C} \ / \ |z - a| < R\}$

Théorème 5.3 (Propriété du conjugué). Pour tout $z, z' \in \mathbb{C}$:

$$\mathrm{Re}(z) = \frac{z + \overline{z}}{2} \qquad \mathrm{Im}(z) = \frac{z - \overline{z}}{2i} \qquad \overline{\overline{z}} = z \qquad \overline{z + z'} = \overline{z} + \overline{z'} \qquad \overline{zz'} = \overline{z} \times \overline{z'}$$

Théorème 5.4 (Propriétés du module). Soit $z, z' \in \mathbb{C}$:

Propriétés algébriques : $|\overline{z}| = |z|$, $z\overline{z} = |z|^2$, $|zz'| = |z| \times |z'|$ et si $z' \neq 0$, $\left|\frac{z}{z'}\right| = \frac{|z|}{|z'|}$. Enfin, $|z| = 0 \Leftrightarrow z = 0$.

Propriétés géométriques : $|\mathrm{Re}(z)| \leq |z|$ et $|\mathrm{Im}(z)| \leq |z|$.

Inégalités triangulaires : $|z + z'| \leq |z| + |z'|$ et $||z| - |z'|| \leq |z \pm z'| \leq |z| + |z'|$.

Théorème 5.5. Soit $z \in \mathbb{C}$.

— $z \in \mathbb{R} \Leftrightarrow \mathrm{Im}(z) = 0 \Leftrightarrow z = \overline{z}$.

— $z \in i\mathbb{R} \Leftrightarrow \mathrm{Re}(z) = 0 \Leftrightarrow z = -\overline{z}$.

5.2. Exponentielle complexe

5.2.1. Groupe $\mathbb{U}$ des nombres complexes de module 1

Définition 5.3 (Ensemble $\mathbb{U}$). On note $\mathbb{U}$ l'ensemble des nombres complexes de module 1, c'est-à-dire

$$\mathbb{U} = \{z \in \mathbb{C} \ / \ |z| = 1\}$$

L'ensemble $\mathbb{U}$ est représenté par le cercle trigonométrique de centre O et de rayon 1.

Propriété 5.1. L'ensemble $\mathbb{U}$ est stable par multiplication et par passage à l'inverse. C'est-à-dire que

$$\forall z,z' \in \mathbb{U},\ zz' \in \mathbb{U}$$

$$\forall z \in \mathbb{U},\ \frac{1}{z} \in \mathbb{U}$$

Démonstration. $|zz'| = |z|\,|z'| = 1 \times 1 = 1$ et $\left|\frac{1}{z}\right| = \frac{1}{|z|} = \frac{1}{1} = 1$ (0 n'appartient pas à $\mathbb{U}$). $\square$

Remarque 5.3. [Attention ! !] La propriété 5.1 parle du produit de deux complexes, mais pas de la somme. Il y a une bonne raison à cela. Un (contre-)exemple tout bête : $z = 1 \in \mathbb{U}$ et $z' = -1 \in \mathbb{U}$, pourtant $z + z' = 0 \notin \mathbb{U}$.

5.2.2. Exponentielle imaginaire

Lemme 5.1. Soit a et b deux réels tels que $a^2 + b^2 = 1$. Il existe un réel (non unique) θ tel que $a = \cos(\theta)$ et $b = \sin(\theta)$.

Démonstration. Puisque $a^2 + b^2 = 1$, nécessairement $-1 \leq a \leq 1$ et $-1 \leq b \leq 1$. En nous appuyant sur le cercle trigonométrique, on remarque qu'il existe un réel α tel que $\cos(\alpha) = a$. Comme $\cos^2(\alpha) + \sin^2(\alpha) = 1$, nous avons $b^2 = \sin^2\alpha$. Ainsi on a $\sin(\alpha) = b$ ou $\sin(\alpha) = -b$. Si la première égalité est vraie, on prend $\theta = \alpha$, sinon on prend $\theta = -\alpha$. Dans les deux cas, le réel θ construit vérifie les deux égalités mentionnées dans l'énoncé du lemme. $\square$

Propriété 5.2. Pour tout complexe $z \in \mathbb{U}$, il existe un réel θ tel que $z = \cos(\theta) + i\sin(\theta)$.

Démonstration. Soit $z = a + ib \in \mathbb{U}$ avec a et b réels. Par définition de $\mathbb{U}$, a et b vérifient $a^2 + b^2 = 1$. Par application du lemme 5.1, il existe $\theta \in \mathbb{R}$ tel que $a = \cos(\theta)$ et $b = \sin(\theta)$. Ainsi $z = \cos(\theta) + i\sin(\theta)$. $\square$

Propriété 5.3 (Réciproque de la propriété 5.2).
Tout nombre complexe s'écrivant sous la forme $z = \cos(\theta) + i\sin(\theta)$ avec $\theta \in \mathbb{R}$, appartient à $\mathbb{U}$.

Démonstration. C'est évident compte tenu du fait que $\cos^2(\theta) + \sin^2(\theta) = 1$. $\square$

Propriété 5.4. Soit la fonction f définie par

$$f : \begin{cases} \mathbb{R} \to \mathbb{C} \\ \theta \mapsto f(\theta) = \cos(\theta) + i\sin(\theta) \end{cases}$$

Alors

$$\forall \theta \in \mathbb{R},\ \overline{f(\theta)} = f(-\theta) = \frac{1}{f(\theta)}$$

$$\forall \alpha,\beta \in \mathbb{R},\ f(\alpha)f(\beta) = f(\alpha+\beta)$$

Démonstration.

- $\overline{f(\theta)} = \cos(\theta) - i\sin(\theta) = \cos(-\theta) + i\sin(-\theta) = f(-\theta)$.
- $\frac{1}{f(\theta)} = \frac{\overline{f(\theta)}}{f(\theta)\overline{f(\theta)}} = \frac{\overline{f(\theta)}}{|f(\theta)|^2} = \overline{f(\theta)}$.
- $$\begin{aligned} f(\alpha)f(\beta) &= (\cos(\alpha) + i\sin(\alpha))(\cos(\beta) + i\sin(\beta)) \\ &= (\cos(\alpha)\cos(\beta) - \sin(\alpha)\sin(\beta)) + i(\cos(\alpha)\sin(\beta) + \sin(\alpha)\cos(\beta)) \\ &= \cos(\alpha+\beta) + i\sin(\alpha+\beta) = f(\alpha+\beta) \end{aligned}$$

Définition 5.4 (Exponentielle imaginaire). Compte tenu de l'analogie de la propriété 5.4 avec les propriétés de l'exponentielle réelle (voire de sa définition, suivant comment l'exponentielle vous a été présentée au lycée), on définit l'*exponentielle imaginaire* par

$$\forall \theta \in \mathbb{R},\ e^{i\theta} = \cos(\theta) + i\sin(\theta)$$

Remarque 5.4. À votre niveau, il faut garder en tête que ceci n'est qu'une notation. Cela n'aurait pas de sens de dire que c'est « e multiplié $i\theta$ fois par lui même. » En réalité, ce n'est pas qu'une simple notation, mais la définition rigoureuse d'une exponentielle (réelle ou complexe ou autre) n'est pas encore à votre portée.

5.2.3. Argument d'un nombre complexe

Propriété 5.5. Soit $z \in \mathbb{C}$ non nul. Il existe un unique réel $\rho > 0$ et un unique réel θ défini à 2π près tel que

$$z = \rho e^{i\theta}$$

Démonstration. Puisque z est non nul, on peut poser $z' = \frac{z}{|z|}$. Nous avons $|z'| = 1$ donc $z' \in \mathbb{U}$ et d'après la propriété 5.2, il existe $\theta \in \mathbb{R}$ tel que $z' = e^{i\theta}$. Ainsi en posant $\rho = |z| > 0$ nous avons $z = \rho e^{i\theta}$. Nous avons montré l'existence de ρ et de θ, il ne reste plus qu'à démontrer l'unicité de ρ et l'unicité à 2π près de θ.

Soit $z = \rho_1 e^{i\theta_1} = \rho_2 e^{i\theta_2}$ avec $\rho_1 > 0$ et $\rho_2 > 0$. Nous avons $1 = \frac{z}{z} = \frac{\rho_1}{\rho_2} e^{i(\theta_1-\theta_2)}$. Ainsi en prenant le module : $1 = \left|\frac{\rho_1}{\rho_2} e^{i(\theta_1-\theta_2)}\right| = \frac{|\rho_1|}{|\rho_2|} = \frac{\rho_1}{\rho_2}$. Donc $\rho_1 = \rho_2$ et par conséquent $e^{i\theta_1} = e^{i\theta_2} = \cos(\theta_1) + i\sin(\theta_1) = \cos(\theta_2) + i\sin(\theta_2)$. Or on sait que deux complexes sont égaux si et seulement si leurs parties réelles et imaginaires respectives sont égales. Ainsi, $\cos(\theta_1) = \cos(\theta_2)$ et $\sin(\theta_1) = \sin(\theta_2)$. La première égalité nous donne

$$\theta_1 = \theta_2 \mod 2\pi \qquad \text{ou} \qquad \theta_1 = -\theta_2 \mod 2\pi$$

La seconde, quant à elle nous donne

$$\theta_1 = \theta_2 \mod 2\pi \qquad \text{ou} \qquad \theta_1 = \pi - \theta_2 \mod 2\pi$$

Les deux équations réunies nous donnent donc bien $\theta_1 = \theta_2 \mod 2\pi$ □

Remarque 5.5. On remarque que ρ est nécessairement égal à $|z|$. On peut donc tout simplement écrire $z = |z|\, e^{i\theta}$.

Définition 5.5 (Argument). Soit $z \in \mathbb{C}$ non nul, $\rho > 0$ et θ un réel tel que $z = \rho e^{i\theta}$.

— L'écriture $\rho e^{i\theta}$ s'appelle la *forme trigonométrique* ou encore *forme exponentielle* de z.

— Le couple $(\rho\,;\theta)$ est un *couple de coordonnées polaires* du point du plan complexe d'affixe z.

— Le réel θ est appelé un *argument* de z et est noté $\arg(z)$. Lorsque $\theta \in]-\pi\,;\pi]$ on parle d'argument principal.

Remarque 5.6. Puisque $(\rho\,;\theta)$ est appelé couple de coordonnées polaires, vous vous doutez bien que le calcul de ρ et de θ va être le même que celui vu dans le chapitre de géométrie du plan.

Soit $z = a + ib = \rho e^{i\theta}$. Alors

$$\begin{aligned}
a &= \rho\cos(\theta)\\
b &= \rho\sin(\theta)\\
\rho &= \sqrt{a^2+b^2}\\
\cos(\theta) &= \frac{a}{\rho} = \frac{a}{\sqrt{a^2+b^2}}\\
\sin(\theta) &= \frac{b}{\rho} = \frac{b}{\sqrt{a^2+b^2}}
\end{aligned}$$

Propriété 5.6. Soit z et z' deux nombres complexes non nuls.

$$z = z' \Leftrightarrow \left(|z| = |z'| \text{ ET } \arg(z) = \arg(z') \mod 2\pi\right)$$

Démonstration. La réciproque est évidente. Le sens direct est quant à lui une conséquence directe de la propriété 5.5. Plus qu'une conséquence, c'est en fait une autre manière d'écrire la propriété 5.5. □

Propriété 5.7. Soit $z = \rho e^{i\theta}$ et $z' = \rho' e^{i\theta'}$ deux nombres complexes non nuls. Alors

$$\begin{aligned}
zz' &= \rho\rho' e^{i(\theta+\theta')}\\
\frac{z}{z'} &= \frac{\rho}{\rho'} e^{i(\theta-\theta')}
\end{aligned}$$

Démonstration. Les calculs sont simples. Il faut tout de même tenir compte de la propriété 5.4. □

Propriété 5.8. Soit z et z' deux nombres complexes non nuls.

$$\arg(zz') = \arg(z) + \arg(z') \mod 2\pi$$
$$\arg\left(\frac{z}{z'}\right) = \arg(z) - \arg(z') \mod 2\pi$$
$$\arg(\overline{z}) = \arg\left(\frac{1}{z}\right) = -\arg(z) \mod 2\pi$$
$$\arg(z^n) = n\arg(z) \mod 2\pi$$

Démonstration. C'est une conséquence directe de la propriété 5.7 □

5.2.4. Exponentielle complexe

Définition 5.6 (Fonction exponentielle complexe). Soit $z = a + ib$ un nombre complexe avec a et b réels. On appelle *exponentielle complexe* de z le nombre complexe

$$\mathrm{e}^z = \mathrm{e}^{a+ib} = \mathrm{e}^a\mathrm{e}^{ib}$$

La fonction qui à tout nombre complexe z associe le nombre complexe e^z s'appelle *fonction exponentielle complexe*.

Propriété 5.9. Soit z et z' deux nombres complexes de formes algébriques $z = a + ib$ et $z' = a' + ib'$. Alors

$$\arg(\mathrm{e}^z) = b \tag{5.1}$$
$$|\mathrm{e}^z| = \mathrm{e}^a \tag{5.2}$$
$$\mathrm{e}^z \neq 0 \tag{5.3}$$
$$\mathrm{e}^{z+z'} = \mathrm{e}^z\mathrm{e}^{z'} \tag{5.4}$$
$$(\mathrm{e}^z)^{-1} = \mathrm{e}^{-z} \tag{5.5}$$
$$\forall n \in \mathbb{Z},\ (\mathrm{e}^z)^n = \mathrm{e}^{nz} \tag{5.6}$$
$$\overline{\mathrm{e}^z} = \mathrm{e}^{\overline{z}} \tag{5.7}$$

Démonstration.

1. $|\mathrm{e}^z| = |\mathrm{e}^a|\left|\mathrm{e}^{ib}\right| = |\mathrm{e}^a| = \mathrm{e}^a$ car l'exponentielle réelle est positive.
2. $|\mathrm{e}^z| = \mathrm{e}^a \neq 0$.
3. $\mathrm{e}^{z+z'} = \mathrm{e}^{(a+a')+i(b+b')} \overset{def}{=} \mathrm{e}^{a+a'}\mathrm{e}^{i(b+b')} = \mathrm{e}^a\mathrm{e}^{a'}\mathrm{e}^{ib}\mathrm{e}^{ib'} = \mathrm{e}^a\mathrm{e}^{ib}\mathrm{e}^{a'}\mathrm{e}^{ib'} = \mathrm{e}^{a+ib}\mathrm{e}^{a'+ib'} = \mathrm{e}^z\mathrm{e}^{z'}$.
4. $\mathrm{e}^z\mathrm{e}^{-z} = \mathrm{e}^{z-z} = \mathrm{e}^0 = 1$.
5. Pour $n \in \mathbb{N}$, c'est une conséquence (par récurrence immédiate) de l'équation 5.4. Pour $n < 0$, on applique simplement l'équation 5.5 pour avoir $-n \in \mathbb{N}$.
6. $\overline{\mathrm{e}^z} = \overline{\mathrm{e}^a\mathrm{e}^{ib}} = \mathrm{e}^a\overline{\mathrm{e}^{ib}} = \mathrm{e}^a\left(\overline{\cos(b) + i\sin(b)}\right) = \mathrm{e}^a(\cos(b) - i\sin(b)) = \mathrm{e}^a(\cos(-b) + i\sin(-b)) = \mathrm{e}^a\mathrm{e}^{-ib} = \mathrm{e}^{a-ib} = \mathrm{e}^{\overline{z}}$.

Propriété 5.10. Soit z et z' deux complexes. Alors

$$e^z = e^{z'} \Leftrightarrow z = z' + 2ik\pi \quad \text{avec } k \in \mathbb{Z} \Leftrightarrow z = z' \mod 2i\pi$$

Démonstration. Posons $z = a + ib$ et $z' = a' + ib'$. D'après la propriété 5.6, nous avons

$$\begin{aligned} e^z = e^{z'} &\Leftrightarrow \left(|e^z| = \left|e^{z'}\right| \text{ ET } \arg(e^z) = \arg\left(e^{z'}\right) \mod 2\pi\right) \\ &\Leftrightarrow \left(e^a = e^{a'} \text{ ET } b = b' \mod 2\pi\right) \\ &\Leftrightarrow \left(a = a' \text{ ET } b = b' + 2k\pi \quad \text{avec } k \in \mathbb{Z}\right) \\ &\Leftrightarrow z = z' + 2ik\pi \quad \text{avec } k \in \mathbb{Z} \end{aligned}$$

Remarque 5.7. Faites attention au « i » dans le modulo !

5.3. Applications

5.3.1. Trigonométrie

Propriété 5.11 (Formules d'Euler). Pour tout réel θ,

$$\cos(\theta) = \frac{e^{i\theta} + e^{-i\theta}}{2}$$
$$\sin(\theta) = \frac{e^{i\theta} - e^{-i\theta}}{2i}$$

Démonstration. Il suffit d'écrire $e^{i\theta}$ et $e^{-i\theta}$ puis de faire la somme et la différence pour obtenir les deux formules. □

Remarque 5.8. Faites bien attention au « i » qui est au dénominateur de la formule du sinus.

Propriété 5.12 (Formule de De Moivre). Pour tout réel θ et tout entier naturel n,

$$(\cos(\theta) + i\sin(\theta))^n = \cos(n\theta) + i\sin(n\theta)$$

Démonstration. La démonstration est évidente lorsque l'on a remarqué que l'égalité s'écrit avec les exponentielles complexes :

$$\left(e^{i\theta}\right)^n = e^{in\theta}$$

Il suffit ensuite d'utiliser la propriété 5.9 pour conclure. □

Remarque 5.9. Les formules d'Euler et de De Moivre servent principalement à *linéariser* ou *délinéariser* des formules de trigonométrie. C'est-à-dire par exemple calculer $\cos^3(\theta)$ en fonction de $\cos(\theta)$, $\cos(2\theta)$, $\cos(3\theta)$...et $\sin(\theta)$, $\sin(2\theta)$, $\sin(3\theta)$...

Propriété 5.13 (Factorisation par l'angle moitié). Pour tout réel $\theta \in \mathbb{R}$:

$$e^{i\theta}+1 = e^{i\frac{\theta}{2}}\left(e^{i\frac{\theta}{2}}+e^{-i\frac{\theta}{2}}\right) = 2e^{i\frac{\theta}{2}}\cos\left(\frac{\theta}{2}\right)$$
$$e^{i\theta}-1 = e^{i\frac{\theta}{2}}\left(e^{i\frac{\theta}{2}}-e^{-i\frac{\theta}{2}}\right) = 2ie^{i\frac{\theta}{2}}\sin\left(\frac{\theta}{2}\right)$$

Démonstration. Il suffit de factoriser $e^{i\theta}+1$ par $e^{i\frac{\theta}{2}}$! □

Remarque 5.10. Il existe en fait une formule plus générale :

$$e^{ix}+e^{iy} = 2e^{i\frac{x+y}{2}}\cos\left(\frac{x-y}{2}\right)$$

L'avantage de la forme exponentielle est qu'elle permet de faire très facilement des calculs de trigonométrie. Ainsi vous pouvez très facilement démontrer les formules suivantes avant de les apprendre par cœur :

Propriété 5.14 (Formulaire de trigonométrie).

$$\begin{aligned}
\cos(a+b) &= \cos(a)\cos(b)-\sin(a)\sin(b)\\
\sin(a+b) &= \sin(a)\cos(b)+\cos(a)\sin(b)\\
\tan(a+b) &= \frac{\tan(a)+\tan(b)}{1-\tan(a)\tan(b)}\\
\cos(2a) &= \cos^2(a)-\sin^2(a) = 2\cos^2(a)-1 = 1-2\sin^2(a)\\
\sin(2a) &= 2\sin(a)\cos(a)\\
\tan(2a) &= \frac{2\tan(a)}{1-\tan^2(a)}\\
\cos(a) &= \frac{1-\tan^2\left(\frac{a}{2}\right)}{1+\tan^2\left(\frac{a}{2}\right)}\\
\sin(a) &= \frac{2\tan\left(\frac{a}{2}\right)}{1+\tan^2\left(\frac{a}{2}\right)}\\
\tan(a) &= \frac{2\tan\left(\frac{a}{2}\right)}{1-\tan^2\left(\frac{a}{2}\right)}\\
\cos(a)+\cos(b) &= 2\cos\left(\frac{a+b}{2}\right)\cos\left(\frac{a-b}{2}\right)\\
\cos(a)-\cos(b) &= -2\sin\left(\frac{a+b}{2}\right)\sin\left(\frac{a-b}{2}\right)\\
\sin(a)+\sin(b) &= 2\sin\left(\frac{a+b}{2}\right)\cos\left(\frac{a-b}{2}\right)\\
\sin(a)-\sin(b) &= 2\cos\left(\frac{a+b}{2}\right)\sin\left(\frac{a-b}{2}\right)
\end{aligned}$$

Démonstration. À vous de la faire. □

Propriété 5.15. Soit a et b deux réels non tous les deux nuls. Il existe un réel strictement positif A et un réel φ tel que pour tout $x \in \mathbb{R}$,

$$a\cos(x) + b\sin(x) = A\cos(x - \varphi)$$

Démonstration. On pose $z = a + ib$. Puisque a et b ne sont pas tous les deux nuls, z est non nul et il peut s'écrire sous forme trigonométrique : $z = Ae^{i\varphi}(= a + ib)$. Par identification des parties réelles et imaginaires, on en déduit que $a = A\cos(\varphi)$ et $b = A\sin(\varphi)$. En remplaçant dans l'expression de départ, on obtient $A(\cos(x)\cos(\varphi) + \sin(x)\sin(\varphi))$, qui est égal à $A\cos(x - \varphi)$. □

Remarque 5.11. La propriété 5.15 est souvent utilisée en physique lorsque l'on étudie par exemple un signal (électrique ou autre) qui peut se décomposer comme somme de deux signaux. Dans ce contexte, A est l'amplitude du signal et φ le décalage de phase. Pour déterminer A et φ il suffit de suivre la démonstration précédente : on met $z = a + ib$ sous forme trigonométrique : $z = Ae^{i\varphi}$.

5.3.2. Équations du second degré dans $\mathbb{C}$

Définition 5.7 (Racine carrée). Soit $z \in \mathbb{C}$ un nombre complexe. On appelle *racine carrée de z* un complexe Z vérifiant $Z^2 = z$.

Propriété 5.16. Tout nombre complexe non nul possède exactement deux racines carrées. De plus ces deux racines carrées sont opposées l'une de l'autre.

Démonstration. Soit $z = \rho e^{i\theta} \neq 0$. Posons $Z_0 = \sqrt{\rho}e^{i\frac{\theta}{2}}$. Nous avons bien $Z_0^2 = z$. Ainsi z possède au moins une racine carrée. De plus $(-Z_0)^2 = Z_0^2 = z$. Donc z possède au moins 2 racines carrées opposées. Montrons que ce sont les deux seules.
Prenons deux racines carrées quelconques de z, Z_1 et Z_2. Nous avons donc $Z_1^2 = Z_2^2$ c'est-à-dire $(Z_1 - Z_2)(Z_1 + Z_2) = 0$. Or dans $\mathbb{C}$, un produit de facteurs est nul si et seulement si l'un des facteurs est nul (on pourrait le démontrer rigoureusement en utilisant le fait que $|zz'| = |z||z'|$ et qu'un complexe est nul si et seulement si son module est nul). Ainsi $Z_1 = Z_2$ ou $Z_1 = -Z_2$.

Les deux seules racines possibles sont donc Z_0 et $-Z_0$. □

Remarque 5.12. Il est strictement INTERDIT d'utiliser le symbole $\sqrt{\ }$ pour un autre nombre qu'un réel positif.
En effet, pour un réel (strictement) positif, il existe deux racines, l'une positive, l'autre négative. Il est donc aisé de décider que l'on note $\sqrt{\ }$ celle qui est positive (l'autre étant alors $-\sqrt{\ }$).
Par contre, pour un nombre complexe, il existe deux racines que l'on ne peut pas distinguer aussi facilement (elles sont certes opposées, mais la notion de positivité n'a pas de sens dans $\mathbb{C}$.)

Un exemple simple qui devrait vous faire passer l'envie d'utiliser le symbole $\sqrt{\ }$ sur un nombre complexe est le suivant :

$$-1 \overset{def}{=} \left(\sqrt{-1}\right)^2 = \sqrt{-1} \times \sqrt{-1} = \sqrt{(-1)\times(-1)} = \sqrt{1} = 1$$

Remarque 5.13.

— Le complexe nul (0) ne possède qu'une seule racine carrée : 0.
— Si $z \in \mathbb{R}_+^*$, ses deux racines carrées sont données par $\sqrt{z}$ et $-\sqrt{z}$.
— Si $z \in \mathbb{R}_-^*$, ses deux racines carrées sont données par $i\sqrt{-z}$ et $-i\sqrt{-z}$. Cela se vérifie aisément en élevant au carré : $\left(i\sqrt{-z}\right)^2 = i^2(-z) = -(-z) = z$.

En pratique, comment calculer les deux racines carrées d'un complexe quelconque ?

1. Si z peut se mettre sous la forme trigonométrique $z = \rho e^{i\theta}$, alors l'une des racines est donnée par
$$Z = \sqrt{\rho} e^{i\frac{\theta}{2}}$$
2. Si z ne peut pas se mettre sous la forme trigonométrique, on est obligé de travailler avec la forme algébrique : $z = a + ib$. Posons $Z = A + iB$ une racine. On a alors
$$\begin{cases} |Z|^2 = |z| \\ \mathrm{Re}\left(Z^2\right) = \mathrm{Re}(z) \\ \mathrm{Im}\left(Z^2\right) = \mathrm{Im}(z) \end{cases}$$
Ce qui veut dire que
$$\begin{cases} A^2 + B^2 = \sqrt{a^2 + b^2} \\ A^2 - B^2 = a \\ 2AB = b \end{cases}$$
En pratique on utilise
$$\begin{cases} A^2 + B^2 = \sqrt{a^2 + b^2} \\ A^2 - B^2 = a \\ AB \text{ est du signe de } b \end{cases}$$

En additionnant et soustrayant les deux premières lignes, on trouve A^2 et B^2, c'est-à-dire A et B au signe près. La troisième ligne nous permet de savoir si A et B sont de même signe ou pas.

Plutôt que d'apprendre (mal) cette formule que vous oublierez trop vite, comprenez et retenez surtout la méthode qui est aisée à retenir puisque très logique.

Théorème 5.6 (Résolution d'une équation du second degré). Soit a, b et c trois nombres complexes avec $a \neq 0$. Considérons l'équation d'inconnue $z \in \mathbb{C}$
$$az^2 + bz + c = 0 \tag{$\star$}$$
Notons le discriminant de l'équation ($\star$)
$$\Delta = b^2 - 4ac$$
On a :
— Si $\Delta = 0$, l'équation ($\star$) admet une racine double z_0 donnée par
$$z_0 = \frac{-b}{2a}$$
— Si $\Delta \neq 0$ et si δ désigne une des deux racines carrées de Δ, alors l'équation ($\star$) admet deux racines distinctes z_1 et z_2 données par :
$$z_1 = \frac{-b - \delta}{2a}$$
$$z_2 = \frac{-b + \delta}{2a}$$

Démonstration. Soit $z \in \mathbb{C}$ une solution de l'équation ($\star$). Puisque $a \neq 0$, nous pouvons écrire le trinôme sous forme canonique

$$0 = a\left(z^2 + \frac{b}{a}z + \frac{c}{a}\right) = a\left[\left(z + \frac{b}{2a}\right)^2 - \frac{b^2-4ac}{4a^2}\right]$$

En notant $Z = z + \frac{b}{2a}$, on doit avoir $Z^2 = \frac{\Delta}{4a^2}$

— Si $\Delta = 0$, alors $Z = 0$, c'est-à-dire $z = -\frac{b}{2a}$.

— Si $\Delta \neq 0$, en notant δ une racine carrée complexe de Δ, $\left(Z - \frac{\delta}{2a}\right)\left(Z + \frac{\delta}{2a}\right) = 0$, c'est-à-dire $Z = \pm\frac{\delta}{2a}$ ou encore $z = \frac{-b-\delta}{2a}$ ou $z = \frac{-b+\delta}{2a}$.

On vérifie dans chacun des cas précédents que z est effectivement solution de l'équation. □

5.3.3. Racines n-ièmes

Définition 5.8 (Racine n-ième de l'unité).
Soit $n \in \mathbb{N}^*$.

— On appelle *racine n-ième de l'unité* un complexe Z tel que $Z^n = 1$.

— On note $\mathbb{U}_n$ l'ensemble des racines n-ièmes de l'unité :

$$\mathbb{U}_n = \{z \in \mathbb{C} \;/\; z^n = 1\}$$

Propriété 5.17. Soit $n \in \mathbb{N}^*$ et $\omega_1 = e^{\frac{2i\pi}{n}}$. Il y a exactement n racines n-ièmes de l'unité. Elles sont données par les puissances de ω_1 : ω_1^k où $k \in [\![0;n-1]\!]$.

$$\mathbb{U}_n = \left\{\omega_1^k \;/\; k \in [\![0;n-1]\!]\right\} = \left\{e^{\frac{2ik\pi}{n}} \;/\; k \in [\![0;n-1]\!]\right\}$$

Démonstration. Soit $Z \in \mathbb{C}$ tel que $Z^n = 1$. En prenant le module, on en déduit que $|Z| = 1$. Il existe donc un unique réel $\theta \in [0;2\pi[$ tel que $Z = e^{i\theta}$. Mais puisque $e^0 = 1 = Z^n = e^{in\theta}$, on doit avoir $n\theta = 0 \mod 2\pi$, c'est-à-dire $n\theta = 2k\pi$ avec $k \in \mathbb{Z}$, ou encore $\theta = \frac{2k\pi}{n}$ avec $k \in \mathbb{Z}$. Comme on veut que θ soit dans l'intervalle $[0;2\pi[$, on doit avoir $0 \leq k < n$, d'où $k \in [\![0;n-1]\!]$. Ainsi $Z = e^{\frac{2ik\pi}{n}} = \omega_1^k$.

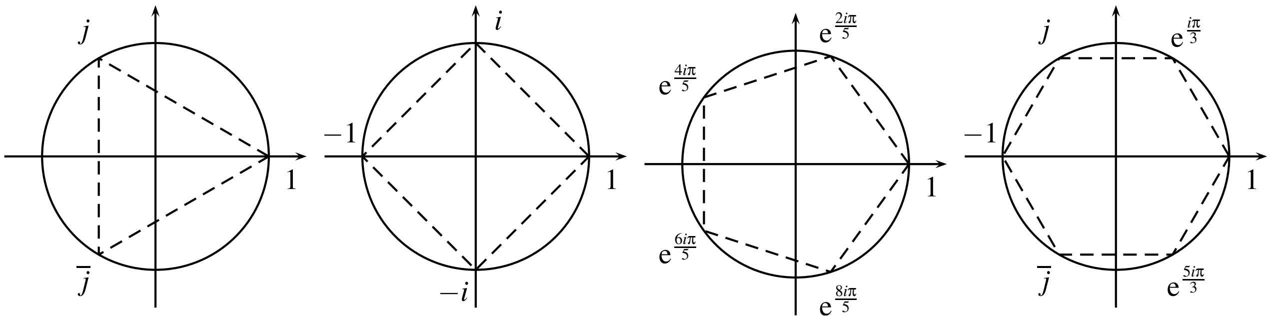
Réciproquement, tout complexe de cette forme vérifie bien $Z^n = 1$.

Étant donné que les exponentielles obtenues sont toutes distinctes, nous avons bien exactement n racines n-ièmes de l'unité. □

Remarque 5.14. D'un point de vue géométrique, on peut constater que les n racines n-ièmes sont les affixes des sommets du polygone régulier à n côtés, centré en O et dont l'un des sommets est le point d'affixe 1.

Propriété 5.18. Soit $n \geq 2$. La somme des racines n-ièmes de l'unité est nulle. C'est-à-dire, si $\omega_1 = e^{\frac{2i\pi}{n}}$ alors

$$1 + \omega_1 + \omega_1^2 + \cdots + \omega_1^{n-1} = \sum_{k=0}^{n-1} \omega_1^k = \sum_{z \in \mathbb{U}_n} z = 0$$



(a) $n=3$ (b) $n=4$ (c) $n=5$ (d) $n=6$

FIGURE 5.1. – Racines n-ièmes de l'unité.

Démonstration. C'est la somme des n premiers termes d'une suite géométrique de raison $\omega_1 \neq 1$ avec $\omega_1^n = 1$. □

Définition 5.9 (Racine n-ième).
Soit $n \in \mathbb{N}^*$ et $z \in \mathbb{C}$ non nul. On appelle *racine n-ième de z* un complexe Z tel que

$$Z^n = z$$

Propriété 5.19. Soit $n \in \mathbb{N}^*$ et $z \in \mathbb{C}$ non nul avec $z = \rho e^{i\theta}$ et $\rho > 0$. Le complexe z admet n racines n-ièmes données par

$$Z_k = \rho^{\frac{1}{n}} e^{i\left(\frac{\theta+2k\pi}{n}\right)} = \rho^{\frac{1}{n}} e^{i\frac{\theta}{n}} \omega_1^k \qquad k \in [\![0\,;n-1]\!]$$

où $\omega_1 = e^{\frac{2i\pi}{n}}$.

Démonstration. Notons $Z_0 = \rho^{\frac{1}{n}} e^{\frac{i\theta}{n}}$. On a bien $Z_0^n = z$. Par conséquent, $Z^n = z$ si et seulement si $\left(\frac{Z}{Z_0}\right)^n = 1$, c'est-à-dire si et seulement si $\frac{Z}{Z_0}$ est une racine n-ième de l'unité. □

Remarque 5.15. Il faut retenir de la propriété 5.19, d'abord qu'il y a exactement n racines n-ièmes d'un complexe non nul, mais aussi qu'elles s'obtiennent en multipliant une solution particulière par les racines n-ièmes de l'unité.

Remarque 5.16. [Résolution de $z^n = \lambda$] Pour résoudre l'équation $z^n = \lambda$ d'inconnue z, avec $\lambda \in \mathbb{C}$, il faut écrire λ sous forme trigonométrique. Il n'existe malheureusement pas de méthode générale permettant de la résoudre dans le cas où λ est sous forme algébrique, comme nous avons pu le voir à la section 5.3.2 dans le cas $n = 2$.

5.3.4. Géométrie

5.3.4.1. Translations

Remarque 5.17. Soit $u \in \mathbb{C}$ et $\vec{u}$ le vecteur d'affixe u. L'application $t_u : \left\{ \begin{array}{ccc} \mathbb{C} & \longrightarrow & \mathbb{C} \\ z & \longmapsto & z+u \end{array} \right.$ correspond, géométriquement à la translation de vecteur $\vec{u}$.

5.3.4.2. Symétries centrales

Remarque 5.18. L'application $S_O : \left\{ \begin{array}{ccc} \mathbb{C} & \longrightarrow & \mathbb{C} \\ z & \longmapsto & -z \end{array} \right.$ correspond, géométriquement à la symétrie de centre O.

Propriété 5.20. Soit $\omega \in \mathbb{C}$ et Ω le point d'affixe ω. L'application $S_\omega : \left\{ \begin{array}{ccc} \mathbb{C} & \longrightarrow & \mathbb{C} \\ z & \longmapsto & 2\omega - z \end{array} \right.$ correspond, géométriquement à la symétrie de centre Ω.

Démonstration. Soit M le point d'affixe z. Si N est l'image de M par la symétrie de centre Ω, alors $\overrightarrow{\Omega M} = \overrightarrow{N\Omega}$. D'où, en notant x l'affixe de N, $z - \omega = \omega - x$, ce qui équivaut à $x = 2\omega - z = S_\omega(z)$. □

5.3.4.3. Symétries Axiales

Remarque 5.19.

— L'application $S_x : \left\{ \begin{array}{ccc} \mathbb{C} & \longrightarrow & \mathbb{C} \\ z & \longmapsto & \overline{z} \end{array} \right.$ correspond, géométriquement à la symétrie d'axe (Ox).

— L'application $S_x : \left\{ \begin{array}{ccc} \mathbb{C} & \longrightarrow & \mathbb{C} \\ z & \longmapsto & -\overline{z} \end{array} \right.$ correspond, géométriquement à la symétrie d'axe (Oy).

5.3.4.4. Homothéties et Rotations

Remarque 5.20.

— Soit $\lambda \in \mathbb{R}$. L'application $H_\lambda : \left\{ \begin{array}{ccc} \mathbb{C} & \longrightarrow & \mathbb{C} \\ z & \longmapsto & \lambda z \end{array} \right.$ correspond, géométriquement, à l'homothétie de centre O et de rapport λ.

— Soit $\theta \in \mathbb{R}$. L'application $R_\theta : \left\{ \begin{array}{ccc} \mathbb{C} & \longrightarrow & \mathbb{C} \\ z & \longmapsto & e^{i\theta} z \end{array} \right.$ correspond, géométriquement, à la rotation centre O et d'angle θ.

— Soit $\omega = \rho e^{i\theta} \in \mathbb{C}$. L'application $HR_{\rho\theta} : \left\{ \begin{array}{ccc} \mathbb{C} & \longrightarrow & \mathbb{C} \\ z & \longmapsto & \omega z \end{array} \right.$ correspond, géométriquement, à la composée de la rotation centre O et d'angle θ et de l'homothétie de rapport ρ (quelque soit l'ordre de la composition).

Propriété 5.21. Soit $\omega \in \mathbb{C}, \lambda \in \mathbb{R}, \theta \in \mathbb{R}$ et Ω le point d'affixe ω.

— L'application $H_{\omega,\lambda} : \left\{ \begin{array}{ccc} \mathbb{C} & \longrightarrow & \mathbb{C} \\ z & \longmapsto & \omega + \lambda(z - \omega) \end{array} \right.$ correspond, géométriquement, à l'homothétie de centre Ω et de rapport λ.

— L'application $H_{\omega,\lambda} : \left\{ \begin{array}{ccc} \mathbb{C} & \longrightarrow & \mathbb{C} \\ z & \longmapsto & \omega + e^{i\theta}(z - \omega) \end{array} \right.$ correspond, géométriquement, à la rotation de centre Ω et d'angle θ.

Démonstration. Avec les même notations que précédemment, si N est l'image de M par l'homothétie de centre Ω et de rapport λ, alors $\Omega N = \lambda \Omega M$.
Pour la rotation, l'écriture est moins simple, mais un schéma nous permet de voir les choses. □

5.3.4.5. Cas général

Nous voyons que dans tous les cas, les applications considérées sont de la forme $f : z \mapsto az+b$ ou $f : z \mapsto a\bar{z}+b$, avec $a, b \in \mathbb{C}$ et $a \neq 0$. Réciproquement, est-ce que les applications de cette forme sont des symétries/rotations/translations ? Puisque l'application conjuguée est une symétrie axiale, nous pouvons nous contenter d'étudier la première forme : Soit $a \neq 0$, b deux complexes, $\theta = \arg(a)$ et $f : z \mapsto az+b$.

— Si $a = 1$, alors f est la translation de vecteur b.

— Si $a \neq 1$, alors f possède un unique point fixe : $f(\omega) = \omega \Leftrightarrow \omega = \frac{b}{1-a}$. On peut alors écrire f sous une forme plus simple : $f(z) - \omega = (az+b) - (a\omega + b) = a(z-\omega) = |a|\,e^{i\theta}(z-\omega) = e^{i\theta}\,|a|\,(z-\omega)$.
On voit alors que f est la composée de la rotation de centre ω et d'angle θ et de l'homothétie de centre ω et de rapport $|a|$.

Théorème 5.7 (Interprétation de $\frac{z-b}{z-a}$). Soit $a, b, z \in \mathbb{C}$ avec $z \neq a$ et $z \neq b$. On note A, B, M, les points d'affixes a, b, z. Alors

$$\left|\frac{z-b}{z-a}\right| = \frac{MB}{MA} \quad ; \quad \arg\left(\frac{z-b}{z-a}\right) \equiv \left(\overrightarrow{MA}, \overrightarrow{MB}\right) \mod 2\pi$$

Démonstration. Nous avons $\arg\left(\frac{z-b}{z-a}\right) \equiv \arg\left(\frac{b-z}{a-z}\right) \equiv \arg(b-z) - \arg(a-z) \equiv \left(\overrightarrow{i}, \overrightarrow{MB}\right) - \left(\overrightarrow{i}, \overrightarrow{MA}\right) \equiv \left(\overrightarrow{MA}, \overrightarrow{MB}\right) \mod 2\pi$. Donc $\frac{z-b}{z-a} = \frac{|b-z|e^{i\beta}}{|a-z|e^{i\alpha}} = \frac{MB}{MA}e^{i(\beta-\alpha)} = \frac{MB}{MA}e^{i\left(\overrightarrow{MA}, \overrightarrow{MB}\right)}$. □

Corollaire 5.1. Soit A, B, M trois points du plan d'affixes a, b, m, avec $A \neq M$. Alors

— A, B, M sont alignés si et seulement si $\left(\overrightarrow{MA}, \overrightarrow{MB}\right) \equiv 0 \mod \pi$, si et seulement si $\frac{z-b}{z-a} \in \mathbb{R}$.

— Les droites (AM) et (BM) sont orthogonales si et seulement si $\left(\overrightarrow{AM}, \overrightarrow{BM}\right) \equiv \frac{\pi}{2} \mod \pi$, si et seulement si $\frac{z-b}{z-a} \in i\mathbb{R}$.

Démonstration. Il suffit de reprendre le théorème et de remarquer que les réels sont les complexes d'arguments 0 ou π et que les imaginaires purs sont les complexes d'arguments $\frac{\pi}{2}$ ou $-\frac{\pi}{2}$. □

6. Polynômes

Note : Dans ce chapitre nous travaillerons à la fois sur $\mathbb{R}$ et sur $\mathbb{C}$. Afin d'alléger l'écriture, nous utiliserons la lettre $\mathbb{K}$ pour désigner $\mathbb{R}$ ou $\mathbb{C}$. Ainsi une propriété ou une définition qui est valable à la fois sur $\mathbb{R}$ et sur $\mathbb{C}$ sera énoncée sur $\mathbb{K}$.

6.1. Polynômes à une indéterminée

6.1.1. Définitions

Définition 6.1 (Scalaire). On appelle *scalaire* un élément de $\mathbb{K}$.

Définition 6.2 (Polynôme). On appelle *polynôme d'indéterminée X* à coefficients dans $\mathbb{K}$ une expression de la forme :

$$P = P(X) = a_0 + a_1 X + \cdots + a_n X^n = \sum_{k=0}^{n} a_k X^k$$

où $a_0, a_1, \ldots, a_n$ sont des éléments de $\mathbb{K}$ et sont appelés *coefficients* de $P(X)$.

On note $\mathbb{K}[X]$ l'ensemble des polynômes à coefficients dans $\mathbb{K}$.

Remarque 6.1. [Attention !] L'objet X est un objet mathématique bien précis que l'on appelle indéterminée. Ce n'est ni une valeur ni une variable.

Définition 6.3 (Vocabulaire). On appelle :

- *polynôme constant* un polynôme de la forme $P(X) = a_0$ avec $a_0 \in \mathbb{K}$.
- *polynôme unité* le polynôme $P(X) = 1$.
- *polynôme nul* le polynôme $P(X) = 0$.
- *monôme* un polynôme de la forme $P(X) = a_k X^k$ avec $a_k \in \mathbb{K}$.

Remarque 6.2. Tous les coefficients du polynôme nul sont nuls. Réciproquement, tout polynôme ayant tous ses coefficients nuls est le polynôme nul.

Définition 6.4 (Opérations). Soit $P(X) = \sum_{k=0}^{n} a_k X^k$ et $Q(X) = \sum_{k=0}^{m} b_k X^k$ deux polynômes de $\mathbb{K}[X]$ et $\lambda \in \mathbb{K}$ un scalaire.
On définit les polynômes :

— $(\lambda P)(X) = \sum_{k=0}^{n} \lambda a_k X^k$

— $(P+Q)(X) = \sum_{k=0}^{\max(n;m)} c_k X^k$ où $c_k = a_k + b_k$, en convenant que $a_k = 0$ si $k > n$ et $b_k = 0$ si $k > m$.

— $(P \times Q)(X) = \sum_{k=0}^{n+m} c_k X^k$ où $c_k = \sum_{i+j=k} a_i b_j = \sum_{i=0}^{k} a_i b_{k-i}$, en convenant que $a_k = 0$ si $k > n$ et $b_k = 0$ si $k > m$.

— $(P \circ Q)(X) = \sum_{k=0}^{n} a_k Q^k(X)$.

Remarque 6.3. C'est tout à fait cohérent avec ce que vous connaissez depuis le lycée sur les fonctions polynômiales.

Propriété 6.1. Soit $P(X)$, $Q(X)$ et $R(X)$, trois polynômes de $\mathbb{K}[X]$. On a alors :

$$(P+Q)(X) = (Q+P)(X) \quad \text{(la somme est commutative)}$$
$$(P \times Q)(X) = (Q \times P)(X) \quad \text{(le produit est commutatif)}$$
$$((P+Q)+R)(X) = (P+(Q+R))(X) \quad \text{(la somme est associative)}$$
$$((P \times Q) \times R)(X) = (P \times (Q \times R))(X) \quad \text{(le produit est associatif)}$$
$$(P \times (Q+R))(X) = (P \times Q)(X) + (P \times R)(X) \quad \text{(le produit est distributif par rapport à l'addition)}$$

Démonstration. En exercice. □

6.1.2. Degré d'un polynôme

Définition 6.5 (Degré). Soit $P(X) = \sum_{k=0}^{n} a_k X^k$ un polynôme non nul de $\mathbb{K}[X]$. On appelle *degré* de $P(X)$ et on note $\deg(P)$ le plus grand indice $k \in [\![0;n]\!]$ tel que $a_k \neq 0$.
Par convention on pose $\deg(0) = -\infty$.
On note $\mathbb{K}_n[X]$ l'ensemble des polynômes de $\mathbb{K}[X]$ de degré inférieur ou égal à n.

Définition 6.6 (Vocabulaire). On appelle :

— *Coefficient dominant* le coefficient du monôme de plus haut degré de $P(X)$.

— *Polynôme unitaire (ou normalisé)* un polynôme de coefficient dominant égal à 1.

Propriété 6.2. Soit $P(X)$ et $Q(X)$ deux polynômes de $\mathbb{K}[X]$. On a alors :

1. $\deg(P+Q) \leq \max(\deg(P);\deg(Q))$
2. $\deg(P\times Q) = \deg(P)+\deg(Q)$
3. $\deg(P\circ Q) = \deg(P)\times\deg(Q)$

Démonstration.

1. Si $P=0$ ou $Q=0$, la propriété est évidente. Si $P\neq 0$ et $Q\neq 0$, on pose $P(X)=\sum_{k=0}^{n} a_kX^k$ et $Q(X)=\sum_{k=0}^{m} b_kX^k$. Si $k>\max(n;m)$ alors $c_k=a_k+b_k=0$. Ainsi le plus grand indice k tel que $c_k=a_k+b_k$ est non nul, est forcément inférieur ou égal à $\max(n;m)$. Ainsi $\deg(P+Q)\leq\max(n;m)$.
2. Si $P=0$ ou $Q=0$, la propriété est évidente. Si $P\neq 0$ et $Q\neq 0$, on pose $P(X)=\sum_{k=0}^{n} a_kX^k$ et $Q(X)=\sum_{k=0}^{m} b_kX^k$ (avec $a_n\neq 0$ et $b_m\neq 0$), et on pose $c_k=\sum_{i=0}^{k} a_ib_{k-i}$.
   Soit $k>n+m$. Puisque pour tout $i>n$, $a_i=0$ et pour tout $i\leq n$, $k-i>n+m-i\geq m$ donc $b_{k-i}=0$, on en déduit que $c_k=0$.
   Ainsi, le degré de $(P\times Q)(X)$ est forcément inférieur ou égal à $n+m$. Or
   $$c_{n+m}=a_0\underbrace{b_{n+m}}_{=0}+a_1\underbrace{b_{n+m-1}}_{=0}+\cdots+a_{n-1}\underbrace{b_{m+1}}_{=0}+a_nb_m+\underbrace{a_{n+1}}_{=0}b_{m-1}+\cdots+\underbrace{a_{m+n}}_{=0}b_0$$
   $$c_{n+m}=a_nb_m\neq 0$$
   Le degré est donc exactement égal à $n+m$.
3. $P\circ Q(X)=\sum_{k=0}^{n} a_kQ^k(X)$. Or $\deg\left(Q(X)^k\right)=k\deg(Q)$ (en appliquant $(k-1)$ fois le point précédent) d'où la conclusion.

Remarque 6.4.

— Si $\deg(P)\neq\deg(Q)$ alors $\deg(P+Q)=\max(\deg(P);\deg(Q))$. En effet les termes dominants ne peuvent pas s'éliminer entre eux.

— Si $\deg(P+Q)<\max(\deg(P);\deg(Q))$ alors $\deg(P)=\deg(Q)$ et les coefficients dominants sont opposés.

Propriété 6.3. Deux polynômes sont égaux si et seulement si tous leurs coefficients correspondants aux monômes de mêmes degrés sont égaux. Dans ce cas les deux polynômes sont de même degré.

Démonstration. Soit $P(X)=\sum_{k=0}^{n} a_kX^k$ et $Q(X)=\sum_{k=0}^{m} b_kX^k$. Alors $P(X)=Q(X)\Leftrightarrow(P-Q)(X)=0\Leftrightarrow$ tous les coefficients de $(P-Q)(X)$ sont nuls $\Leftrightarrow\forall k\in[\![0;\max(n;m)]\!]$, $a_k-b_k=0\Leftrightarrow\forall k\in[\![0;\max(n;m)]\!]$, $a_k=b_k$.

Pour $k\in[\![\min(n;m)+1;\max(n;m)]\!]$, on a soit $a_k=0$, soit $b_k=0$, donc $a_k=b_k=0$. Les deux polynômes ont donc bien le même degré. $\square$

6.2. Dérivation

6.2.1. Polynôme dérivé

Définition 6.7 (Polynôme dérivé). Soit $P(X)=\sum_{k=0}^{n} a_k X^k$ un polynôme. On appelle *polynôme dérivé* de $P(X)$, noté $P'(X)$, le polynôme défini par :

$$P'(X)=\begin{cases} 0 & \text{si } \deg(P)\leq 0 \\ \sum_{k=1}^{n} k a_k X^{k-1} = a_1+2a_2X+\cdots+na_nX^{n-1} & \text{si } \deg(P)\geq 1 \end{cases}$$

On définit par récurrence les polynômes dérivés successifs par $\begin{cases} P^{(0)}(X)=P(X) \\ \forall k\in\mathbb{N}^*, P^{(k)}(X)=\left(P^{(k-1)}\right)'(X) \end{cases}$

Propriété 6.4. Soit $P(X)\in\mathbb{K}[X]$ un polynôme, alors

$$\deg(P')=\begin{cases} \deg(P)-1 & \text{si } \deg(P)\geq 1 \\ -\infty & \text{si} \deg(P)\leq 0 \end{cases}$$

Démonstration. Cela vient tout simplement de la définition de la dérivée. □

Propriété 6.5. Soit $P(X)$ et $Q(X)$ deux polynômes de $\mathbb{K}[X]$ et λ un scalaire de $\mathbb{K}$. Alors :

1. $(P+Q)'(X)=(P'+Q')(X)$
2. $(\lambda P)'(X)=(\lambda P')(X)$
3. $(P\times Q)'(X)=(P'\times Q+P\times Q')(X)$

Démonstration. Cela vient là aussi tout simplement des définitions des sommes et produits de polynômes. □

6.2.2. Dérivées d'ordre supérieur

Propriété 6.6 (Formule de Leibniz). Soit $P(X)$ et $Q(X)$ deux polynômes de $\mathbb{K}[X]$ et $n\in\mathbb{N}$. Alors

$$(P\times Q)^{(n)}(X)=\sum_{k=0}^{n}\binom{n}{k}\left(P^{(k)}\times Q^{(n-k)}\right)(X)$$

Démonstration. La démonstration se fait par récurrence : La formule est évidente lorsque $n=0$; Pour l'hérédité, c'est un jeu de changement d'indice et d'application de la formule de Pascal (que vous connaissez bien évidemment par cœur). □

Théorème 6.1 (Formule de Taylor en 0). Soit $P(X)$ un polynôme de degré n, alors :

$$P(X)=P(0)+P'(0)X+\frac{P''(0)}{2!}X^2+\cdots+\frac{P^{(n)}(0)}{n!}X^n=\sum_{k=0}^{n}\frac{P^{(k)}(0)}{k!}X^k$$

Remarque 6.5. Rien ne vous choque dans cette formule ? Nous n'avons pas défini ce que voulait dire $P(0)$! En effet, on a bien précisé dans la remarque 6.1 que X n'était ni une variable ni une valeur. Alors que représente ce $P(0)$? En fait, nous pouvons démontrer qu'il existe un lien bijectif entre le polynôme $P(X)=\sum_{k=0}^{n}a_kX^k$ et la *fonction polynômiale* $\tilde{P}: x\mapsto\sum_{k=0}^{n}a_kx^k$. Ainsi, on va pouvoir confondre (et on ne va pas s'en priver !) les polynômes et les fonctions polynômiales en notant abusivement P au lieu de $\tilde{P}$. Donc $P(0)$ est en fait égal à $\tilde{P}(0)=a_0$.

Démonstration. Par récurrence, vous pouvez montrer aisément que $P^{(k)}(0)=k!a_k$. □

Corollaire 6.1 (Formule de Taylor en α). Soit $P(X)$ un polynôme de degré n et $\alpha\in\mathbb{K}$, alors :

$$P(X)=P(\alpha)+P'(\alpha)(X-\alpha)+\frac{P''(\alpha)}{2!}(X-\alpha)^2+\cdots+\frac{P^{(n)}(\alpha)}{n!}(X-\alpha)^n=\sum_{k=0}^{n}\frac{P^{(k)}(\alpha)}{k!}(X-\alpha)^k$$

Démonstration. On applique la formule de Taylor en 0 (théorème 6.1) au polynôme $Q(X)=P(X+\alpha)$. □

6.3. Arithmétique dans $\mathbb{K}[X]$

6.3.1. Divisibilité

Définition 6.8 (Divisibilité). Soit $A(X)$ et $B(X)$ deux polynômes. On dit que $A(X)$ *divise* $B(X)$ (dans $\mathbb{K}[X]$), et on note $A|B$ si et seulement si il existe $C(X)\in\mathbb{K}[X]$ tel que $B(X)=(A\times C)(X)$.
Le polynôme $A(X)$ est appelé *diviseur* de $B(X)$ et $B(X)$ un *multiple* de $A(X)$.

Remarque 6.6.

1. Un polynôme $P(X)$ non nul est divisible par les polynômes λ et $\lambda P(X)$ avec $\lambda\in\mathbb{K}^*$.
2. Réciproquement, un polynôme de degré 0 (polynôme constant et non nul) divise tous les polynômes.
3. Tout polynôme divise le polynôme nul.

Définition 6.9 (Polynômes associés). Soit $P(X)$ et $Q(X)$ deux polynômes de $\mathbb{K}[X]$. On dit que $P(X)$ et $Q(X)$ sont *associés* si et seulement si il existe un scalaire non nul $\lambda\in\mathbb{K}^*$ tel que $P(X)=\lambda Q(X)$.

Propriété 6.7. Soit $A(X), B(X), C(X), D(X)$ quatre polynômes de $\mathbb{K}[X]$ et $n \in \mathbb{N}^*$. On a :

1. $A|A$
2. $\left\{\begin{array}{l} A|B \\ B|C \end{array}\right. \Rightarrow A|C$
3. $A|B \Rightarrow A|(B \times C)$
4. $\left\{\begin{array}{l} A|B \\ A|C \end{array}\right. \Rightarrow A|(B+C)$
5. $\left\{\begin{array}{l} A|C \\ B|D \end{array}\right. \Rightarrow (A \times B)|(C \times D)$
6. $A|B \Rightarrow A^n|B^n$

Démonstration. Laissée en exercice. □

Propriété 6.8. Soit $P(X)$ et $Q(X)$ deux polynômes non nuls. Alors

$$(P|Q \text{ et } Q|P) \Leftrightarrow \exists \lambda \in \mathbb{K}^*,\ P(X) = \lambda Q(X) \Leftrightarrow P(X) \text{ et } Q(X) \text{ sont associés}$$

Démonstration. $\boxed{\Rightarrow}$ Puisque $P(X)$ divise $Q(X)$, il existe un polynôme non nul $R(X)$ tel que $Q(X) = P(X)R(X)$. D'où $\deg(Q) \geq \deg(P)$. De même, puisque $Q(X)$ divise $P(X)$, on a $\deg(P) \geq \deg(Q)$. Finalement, $\deg(P) = \deg(Q)$ et $\deg(R) = 0$. D'où $R(X)$ est un polynôme constant $R(X) = \lambda$ et $Q(X) = \lambda P(X)$.

$\boxed{\Leftarrow}$ Évident.

Propriété 6.9. Soit A et B deux polynômes. Alors $AB = 1$ si et seulement si A et B sont constants et inverse l'un de l'autre.

Démonstration. $\boxed{\Leftarrow}$ Évident.

$\boxed{\Rightarrow}$ Nécessairement A et B sont non nuls. Ainsi $\deg(A) + \deg(B) = 0$ donc $\deg(A) = \deg(B) = 0$. Ce sont donc des polynômes constants inverses l'un de l'autre.

Remarque 6.7. Cela veut dire qu'un polynôme non nul n'est pas forcément inversible. Les seuls polynômes inversibles sont les constantes non nulles.

6.3.2. Division euclidienne

Théorème 6.2 (Division euclidienne). Soit $A(X)$ et $B(X)$ deux polynômes de $\mathbb{K}[X]$ tels que $B(X) \neq 0$. Alors il existe un unique couple de polynômes $(Q(X); R(X)) \in \mathbb{K}[X] \times \mathbb{K}[X]$ tel que

$$\left\{\begin{array}{l} A(X) = (B \times Q)(X) + R(X) \\ \deg(R) < \deg(B) \end{array}\right.$$

Les polynômes $Q(X)$ et $R(X)$ sont respectivement appelés *quotient* et *reste* de la *division euclidienne* de $A(X)$ par $B(X)$.

Démonstration. Les détails de la démonstration sont laissés en exercice.

Unicité si existence : On pose $A(X) = (B \times Q)(X) + R(X) = (B \times Q')(X) + R'(X)$. On démontre alors que $\deg(B) + \deg(Q - Q') < \deg(B)$ ce qui permet de conclure.

Existence : Si $\deg(A) < \deg(B)$ alors $Q(X) = 0$ et $R(X) = A(X)$ conviennent. Si $\deg(A) \geq \deg(B)$ on fait une récurrence forte sur le degré de $A(X)$. Pour l'hérédité, en notant $a_{n+1}X^{n+1}$ et $b_m X^m$ les termes dominants de $A(X)$ et $B(X)$, il faut s'appuyer sur le polynôme $A(X) - B(X) \times \frac{a_{n+1}}{b_m} X^{n+1-m}$ qui est de degré inférieur ou égal à n.

Corollaire 6.2. Soit $P(X)$ et $Q(X)$ deux polynômes de $\mathbb{K}[X]$. Alors $P|Q$ si et seulement si le reste de la division euclidienne de $Q(X)$ par $P(X)$ est nul.

Exemple 6.1. Voyons concrètement comment on fait une division euclidienne entre deux polynômes. Il faut pour cela se rappeler comment vous faisiez les divisions euclidiennes en primaire.

Nous allons faire la division euclidienne de $2X^4 + X^3 - X^2 + X + 1$ par $2X^2 - X - 2$, en utilisant la technique dite « division selon les puissances décroissantes », c'est-à-dire en éliminant au fur et à mesure les monômes de plus haut degré.

$$\begin{aligned}
2X^4 + X^3 - X^2 + X + 1 &= X^2(2X^2 - X - 2) + 2X^3 + X^2 + X + 1 \\
&= X^2(2X^2 - X - 2) + X(2X^2 - X - 2) + 2X^2 + 3X + 1 \\
&= X^2(2X^2 - X - 2) + X(2X^2 - X - 2) + 1(2X^2 - X - 2) + 4X + 3 \\
&= (2X^2 - X - 2)\underbrace{(X^2 + X + 1)}_{\text{Quotient}} + \underbrace{4X + 3}_{\text{Reste}}
\end{aligned}$$

On représente cela ainsi :

$$\begin{array}{rrrrrrrrrr|l}
2X^4 & + & X^3 & - & X^2 & + & X & + & 1 & & 2X^2 - X - 2 \\
\cline{11-11}
-(2X^4 & - & X^3 & - & 2X^2) & & & & & & X^2 + X + 1 \\
\cline{1-5}
 & & 2X^3 & + & X^2 & + & X & & & & \\
 & & -(2X^3 & - & X^2 & - & 2X) & & & & \\
\cline{3-7}
 & & & & 2X^2 & + & 3X & + & 1 & & \\
 & & & & -(2X^2 & - & X & - & 2) & & \\
\cline{5-9}
 & & & & & & 4X & + & 3 & &
\end{array}$$

6.4. Racines d'un polynôme

6.4.1. Racines

Définition 6.10 (Racine). Soit $P(X) \in \mathbb{K}[X]$. On dit que $\alpha \in \mathbb{K}$ est une *racine* ou un *zéro* de $P(X)$ si $P(\alpha) = 0$.

Théorème 6.3. Soit $P(X) \in \mathbb{K}[X]$ et $\alpha \in \mathbb{K}$. Alors

$$\alpha \text{ est un zéro de } P(X) \Leftrightarrow P(X) \text{ est divisible par } (X - \alpha)$$

Démonstration. $\boxed{\Rightarrow}$ La division euclidienne de $P(X)$ par $(X-\alpha)$ donne $P(X) = Q(X)(X-\alpha)+R(X)$ avec $\deg(R) < 1$, donc $R(X) = \lambda$ avec $\lambda \in \mathbb{K}$. D'où $P(\alpha) = 0 = \lambda$ et $(X-\alpha)|P$.

$\boxed{\Leftarrow}$ Nous avons $P(X) = (X-\alpha)Q(X)$ donc $P(\alpha) = (\alpha-\alpha)Q(\alpha) = 0$, et α est une racine de $P(X)$.

□

Définition 6.11 (Ordre de multiplicité). Soit $P(X) \in \mathbb{K}[X]$ un polynôme non nul et α une racine de $P(X)$. On définit *l'ordre de multiplicité* de α comme étant le plus grand entier m tel que $(X-\alpha)^m$ divise $P(X)$.

Remarque 6.8. Grâce au théorème 6.3 nous pouvons affirmer que $m \geq 1$. De plus, puisque $(X-\alpha)^m$ divise $P(X)$ nous avons aussi $m \leq \deg(P)$. Donc finalement $1 \leq m \leq \deg(P)$.

Propriété 6.10. Soit $P(X) \in \mathbb{K}[X]$, $\alpha \in \mathbb{K}$ et $k \geq 1$ un entier. Alors

$$(X-\alpha)^k \text{ divise } P(X) \Leftrightarrow P(\alpha) = P'(\alpha) = \cdots = P^{(k-1)}(\alpha) = 0$$

Démonstration. $\boxed{\Leftarrow}$ D'après la formule de Taylor en α (corollaire 6.1), nous avons

$$P(X) = (X-\alpha)^k \left[\sum_{i=k}^{n} \frac{P^{(i)}(\alpha)}{i!}(X-\alpha)^{i-k}\right]$$

$\boxed{\Rightarrow}$ Par contraposée, toujours en utilisant la formule de Taylor en α. Si $P^{(j)}(\alpha) \neq 0$ pour un certain $0 \leq j \leq k-1$, alors $Q(X) = \sum_{i=0}^{k-1} \frac{P^{(i)}(\alpha)}{i!}(X-\alpha)^i$ est un polynôme non nul de degré strictement inférieur à k, il ne peut donc pas être divisible par $(X-\alpha)^k$. Or le polynôme $(P-Q)(X)$ est divisible par $(X-\alpha)^k$, donc $P(X)$ ne peut pas être divisible par $(X-\alpha)^k$.

Théorème 6.4. Soit $P(X) \in \mathbb{K}[X]$, $\alpha \in \mathbb{K}$ et $m \geq 1$ un entier. Alors α est une racine de multiplicité m de $P(X)$ si et seulement si $P(\alpha) = P'(\alpha) = \cdots = P^{(m-1)}(\alpha) = 0$ et $P^{(m)}(\alpha) \neq 0$.

Démonstration. Cela découle tout simplement de la définition de la multiplicité d'une racine et de la propriété 6.10. □

6.4.2. Nombre de racines

Propriété 6.11. Soit $P(X) \in \mathbb{K}[X]$ et $\alpha_1, \ldots, \alpha_p$, p racines distinctes de $P(X)$. Alors P est divisible par $(X-\alpha_1)(X-\alpha_2)\ldots(X-\alpha_p)$.

Démonstration. Par récurrence :

— La propriété est vraie pour $p = 1$.

— Supposons qu'elle soit vraie pour $p \geq 1$ et montrons la au rang $p+1$. Par hypothèse de récurrence, $P(X)$ est divisible par $(X-\alpha_1)\dots(X-\alpha_p)$, donc $P(X) = (X-\alpha_1)\dots(X-\alpha_p)Q(X)$.
Or $P(\alpha_{p+1}) = 0 = \underbrace{(\alpha_{p+1}-\alpha_1)}_{\neq 0}\dots\underbrace{(\alpha_{p+1}-\alpha_p)}_{\neq 0}Q(\alpha_{p+1})$. On en déduit que α_{p+1} est une racine de $Q(X)$ et donc $Q(X)$ est divisible par $(X-\alpha_{p+1})$ et par conséquent $P(X)$ est divisible par $(X-\alpha_1)\dots(X-\alpha_{p+1})$.

Propriété 6.12. Soit $P(X) \in \mathbb{K}[X]$ et $\alpha_1, \dots, \alpha_p$, p racines distinctes de $P(X)$ de multipicité respectives $m_1, \dots m_p$. Alors $P(X)$ est divisible par $(X-\alpha_1)^{m_1}(X-\alpha_2)^{m_2}\dots(X-\alpha_p)^{m_p}$.

Démonstration. Admis. □

Corollaire 6.3. Un polynôme non nul de degré n possède au plus n racines comptées avec leur ordre de multiplicité.

Démonstration. D'après la propriété 6.12, $P(X)$ est divisible par $Q(X) = (X-\alpha_1)^{m_1}\dots(X-\alpha_p)^{m_p}$ de degré $\deg(Q) = m_1 + \dots + m_p$. Donc si $P(X)$ est non nul, $m_1 + \dots + m_p \leq \deg(P)$. □

Corollaire 6.4. Un polynôme de degré au plus n admettant plus de n racines distinctes est le polynôme nul.

Démonstration. C'est la contraposée du corollaire 6.3. □

Corollaire 6.5. Un polynôme admettant une infinité de racines distinctes est nul.

Démonstration. La démonstration du corollaire du corollaire du corollaire devrait être à votre niveau...

6.5. Décomposition en facteurs irréductibles

6.5.1. Polynôme irréductible

Définition 6.12 (Polynôme irréductible). Soit $P(X) \in \mathbb{K}[X]$. On dit que $P(X)$ est *irréductible* s'il est de degré supérieur ou égal à 1 et si ses seuls diviseurs sont les polynômes λ et $\lambda P(X)$ avec $\lambda \in \mathbb{K}^*$ (c'est-à-dire les polynômes constants non nuls et les polynômes associés à $P(X)$).

Remarque 6.9.

— Un polynôme de degré 1 est forcément irréductible.

— Un polynôme de degré supérieur ou égal à 2 qui possède une racine est forcément réductible.

Théorème 6.5. Tout polynôme de $\mathbb{K}[X]$ de degré supérieur ou égal à 1 se décompose de manière unique en produit d'une constante non nulle et de polynômes irréductibles unitaires à l'ordre des facteurs près.

Démonstration. Hors programme. □

Définition 6.13 (Polynôme scindé). Soit $P(X) \in \mathbb{K}[X]$. On dit que $P(X)$ est scindé dans $\mathbb{K}$ si et seulement si c'est un produit de polynômes de degré 1 de $\mathbb{K}[X]$.

6.5.2. Décomposition dans $\mathbb{C}[X]$

La décomposition d'un polynôme dépend de $\mathbb{K}$. Nous allons donc distinguer les décompositions sur $\mathbb{C}[X]$ et sur $\mathbb{R}[X]$.

Théorème 6.6 (de D'Alembert-Gauss ou théorème fondamental de l'algèbre). Tout polynôme de $\mathbb{C}[X]$ de degré supérieur ou égal à 1 admet au moins une racine dans $\mathbb{C}$.

Démonstration. Hors programme. □

Remarque 6.10. Vous remarquerez que le théorème 6.6 n'est pas énoncé dans $\mathbb{K}$ mais dans $\mathbb{C}$. En effet, vous savez que dans $\mathbb{R}$ il est faux : $1+X^2$ ne possède pas de racines dans $\mathbb{R}$.

Corollaire 6.6. Les polynômes irréductibles de $\mathbb{C}[X]$ sont les polynômes de degré 1.

Démonstration. Nous avons déjà vu que les polynômes de degré 1 étaient forcément irréductibles. Ensuite, par contraposée du théorème de D'Alembert-Gauss (théorème 6.6) un polynôme irréductible est forcément de degré 1. □

Corollaire 6.7. Tout polynôme $P(X) \in \mathbb{C}[X]$ peut s'écrire sous la forme

$$P(X) = \lambda(X-\alpha_1)^{m_1}(X-\alpha_2)^{m_2}\dots(X-\alpha_p)^{m_p}$$

avec $\lambda \in \mathbb{C}$ et $\alpha_1,\dots,\alpha_p$ des complexes distincts.

Démonstration. Nous avons vu dans le théorème 6.5 que tout polynôme se décomposait en produit de polynômes irréductibles. Or d'après le corollaire 6.6 les polynômes irréductibles de $\mathbb{C}[X]$ sont les polynômes de degré 1. □

Remarque 6.11. Le corollaire 6.7 revient à dire que tous les polynômes non nuls de $\mathbb{C}[X]$ sont scindés.
De plus la somme des multiplicités est égale au degré de $P(X)$ (si $P(X)$ est non nul).

6.5.3. Décomposition dans $\mathbb{R}[X]$

Passons maintenant à la décomposition dans $\mathbb{R}[X]$.

Lemme 6.1. Soit $P(X) \in \mathbb{R}[X]$. Si on considère $P(X)$ comme un polynôme de $\mathbb{C}[X]$ et que $\alpha \in \mathbb{C} \setminus \mathbb{R}$ est une racine de $P(X)$ alors $\overline{\alpha}$ est aussi une racine complexe de $P(X)$ avec même multiplicité.

Démonstration. Puisque les coefficients de $P(X)$ sont réels, on remarque que $\overline{P(\alpha)} = P(\overline{\alpha})$ donc $\overline{\alpha}$ est une racine. □

Propriété 6.13. Les polynômes irréductibles de $\mathbb{R}[X]$ sont les polynômes de degré 1 et les polynômes de degré 2 n'admettant pas de racine réelle (discriminant strictement négatif).

Démonstration. Il est clair que les polynômes de degré 1 et de degré 2 à discriminant strictement négatif sont irréductibles dans $\mathbb{R}[X]$. Reste à montrer que les polynômes de degré 2 à discriminant positif ou nul et de degré supérieur ou égal à 3 sont réductibles.
Nous savons déjà qu'un polynôme de degré 2 à discriminant positif ou nul possède 2 racines réelles, il est donc réductible dans $\mathbb{R}[X]$.
Si $P(X)$ est de degré supérieur ou égal à 3, il possède au moins une racine α dans $\mathbb{C}$. Si $\alpha \in \mathbb{R}$, alors $P(X)$ est réductible, sinon d'après le lemme 6.1, $\overline{\alpha} \neq \alpha$ est aussi une racine. Le polynôme $P(X)$ est donc divisible par

$$\begin{aligned}(X-\alpha)(X-\overline{\alpha}) &= X^2 - (\alpha+\overline{\alpha})X + \alpha\overline{\alpha} \\ &= X^2 - 2\,\mathrm{Re}(\alpha)X + |\alpha|^2 \in \mathbb{R}[X]\end{aligned}$$

Le polynôme $P(X)$ n'est donc pas irréductible. □

Propriété 6.14. Tout polynôme $P(X) \in \mathbb{R}[X]$ peut s'écrire sous la forme

$$P(X) = \lambda(X-\alpha_1)^{m_1}(X-\alpha_2)^{m_2}\dots(X-\alpha_p)^{m_p}(X^2+\beta_1X+\gamma_1)^{n_1}(X^2+\beta_2X+\gamma_2)^{n_2}\dots(X^2+\beta_qX+\gamma_q)^{n_q}$$

Avec $\lambda \in \mathbb{R}$, $\alpha_1,\dots,\alpha_p$ des réels distincts et $\beta_1,\dots,\beta_q$, $\gamma_1,\dots,\gamma_q$ des réels tels que pour tout $k \in [\![1\,;q]\!]$, $\beta_k^2 - 4\gamma_k < 0$.

Démonstration. Là encore cela découle directement du théorème 6.5 et de la propriété 6.13. □

6.6. Somme et produit des racines d'un polynôme

Propriété 6.15 (Somme et produit des racines d'un polynôme scindé). Soit $P(X) \in \mathbb{K}[X]$ un polynôme scindé sur $\mathbb{K}$ de degré $n \geq 1$. On note $P(X) = a_0 + a_1X + \cdots + a_nX^n = \lambda(X-\alpha_1)(X-\alpha_2)\dots(X-\alpha_n)$. Alors

$$\begin{aligned}\alpha_1+\alpha_2+\cdots+\alpha_n &= -\frac{a_{n-1}}{a_n} \\ \alpha_1\alpha_2\dots\alpha_n &= (-1)^n\frac{a_0}{a_n}\end{aligned}$$

Démonstration. Il suffit de développer le produit et d'identifier les coefficients. Le terme constant (a_0) est égal à $(-1)^n\lambda$, le terme dominant (a_n) est égal à λ et le terme de degré $n-1$ (a_{n-1}) est égal à $-\lambda(\alpha_1+\cdots+\alpha_n)$. □

Remarque 6.12. Pour $n=2$ on retrouve la formule (censée être) connue : $x_1+x_2=S=-\frac{b}{a}$, $x_1x_2=P=\frac{c}{a}$. Ainsi, résoudre le système

$$\begin{cases} \alpha+\beta &= S \\ \alpha\times\beta &= P \end{cases}$$

revient à déterminer les solutions de

$$X^2-SX+P=0$$

7. Fractions Rationnelles

Note : Dans ce chapitre nous travaillerons à la fois sur $\mathbb{R}$ et sur $\mathbb{C}$. Afin d'alléger l'écriture, nous utiliserons la lettre $\mathbb{K}$ pour désigner $\mathbb{R}$ ou $\mathbb{C}$. Ainsi une propriété ou une définition qui est valable à la fois sur $\mathbb{R}$ et sur $\mathbb{C}$ sera énoncée sur $\mathbb{K}$.

Le début de ce chapitre contient des choses qui peuvent sembler « évidentes » car vous les utilisez depuis tout petit avec les fractions. En effet, les fractions rationnelles sont aux polynômes ce que les nombres rationnels ($\mathbb{Q}$) sont aux entiers relatifs ($\mathbb{Z}$). Mais vous l'avez déjà remarqué, en maths on aime être rigoureux et poser clairement les choses, mêmes si elles semblent simples.

7.1. Ensemble des fractions rationnelles

7.1.1. Fractions rationnelles

Définition 7.1 (Fraction rationnelle). On appelle *fraction rationnelle* d'indéterminée X un « quotient » de deux polynômes P et Q de $\mathbb{K}[X]$.

On note $\mathbb{K}(X)$ l'ensemble des fractions rationnelles et $F = \frac{P}{Q}$ un élément de $\mathbb{K}(X)$.

Définition 7.2 (Égalité de deux fractions). On dit que deux fractions rationnelles $F_1 = \frac{P_1}{Q_1}$ et $F_2 = \frac{P_2}{Q_2}$ sont égales si et seulement si $P_1Q_2 = P_2Q_1$.

Remarque 7.1. Tout comme un entier peut être considéré comme un rationnel en écrivant $n = \frac{n}{1}$, un polynôme (de $\mathbb{K}[X]$) peut être considéré comme une fraction rationnelle (de $\mathbb{K}(X)$) en l'écrivant $P = \frac{P}{1}$.

Définition 7.3 (Forme irréductible). Soit $F \in \mathbb{K}(X)$ une fraction rationnelle. On appelle *forme irréductible* de F toute écriture de la forme $F = \frac{A}{B}$ avec A et B deux polynômes premiers entres eux. Une telle écriture est toujours possible et unique à multiplications près par des scalaires non nuls.

7.1.2. Opérations

Définition 7.4 (Opérations). Soit $F_1 = \frac{P_1}{Q_1}$ et $F_2 = \frac{P_2}{Q_2}$ deux fractions rationnelles et λ un réel. On définit :

$$\lambda F_1 = \frac{\lambda P_1}{Q_1} \qquad F_1 + F_2 = \frac{P_1Q_2 + P_2Q_1}{Q_1Q_2}$$

$$F_1 \times F_2 = \frac{P_1P_2}{Q_1Q_2} \qquad \frac{F_1}{F_2} = \frac{P_1Q_2}{Q_1P_2}$$

Remarque 7.2.

- Lorsque l'on définit quelque chose il faut vérifier que c'est bien défini (!!!). Nous avons défini la somme et le produit à partir de P_1 et Q_1, mais est-ce que la somme ou le produit changent si on on utilise une autre représentation de $F_1 = \frac{P_1}{Q_1} = \frac{P_1'}{Q_1'}$? Je vous laisse le vérifier par vous même.
- Comme dit plus haut, les définitions de somme et produit de fractions rationnelles sont exactement celles des nombres rationnels.

7.1.3. Dérivées

Définition 7.5 (Dérivée). Soit $F = \frac{P}{Q}$ une fraction rationnelle. On appelle *dérivée* de F la fraction rationnelle égale à $\frac{P'Q-PQ'}{Q^2}$.

Remarque 7.3. Bien entendu, vous avez vérifié que cette définition ne dépendait pas du choix de P et Q...

Propriété 7.1. Toutes les propriétés « classiques » de dérivées restent vraies :

$$\begin{aligned}(\lambda F_1)' &= \lambda F_1' \\ (F_1+F_2)' &= F_1'+F_2' \\ (F_1F_2)' &= F_1'F_2+F_1F_2' \\ \left(\frac{F_1}{F_2}\right)' &= \frac{F_1'F_2-F_1F_2'}{F_2^2}\end{aligned}$$

Démonstration. Ce n'est que du calcul...

7.1.4. Degré

Définition 7.6 (Degré d'une fraction rationnelle). Soit une fraction rationnelles $F(X) = \frac{P}{Q} \in \mathbb{K}(X)$. On appelle degré de F l'entier relatif :

$$\deg(F) = \deg(P) - \deg(Q)$$

Remarque 7.4.

- Là encore je vous laisse le soin de vérifier que le degré d'une fraction rationnelle ne dépend pas du choix de P et Q.
- Le degré des fractions rationnelles prolonge le degré des polynômes. C'est-à-dire que

$$\underbrace{\deg(P)}_{P\in\mathbb{K}[X]} = \underbrace{\deg\left(\frac{P}{1}\right)}_{P\in\mathbb{K}(X)}$$

Propriété 7.2. Soit $F_1(X)$ et $F_2(X)$ deux fractions rationnelles de $\mathbb{K}(X)$. On a alors :

1. $\deg(F_1 \times F_2) = \deg(F_1) + \deg(F_2)$
2. $\deg(F_1 + F_2) \leq \max(\deg(F_1); \deg(F_2))$

Démonstration.

1. Posons $F_1 = \frac{P_1}{Q_1}$ et $F_2 = \frac{P_2}{Q_2}$, alors $F_1 \times F_2 = \frac{P_1 \times P_2}{Q_1 \times Q_2}$, donc $\deg(F_1 \times F_2) = \deg(P_1 \times P_2) - \deg(Q_1 \times Q_2) = \deg(P_1) - \deg(Q_1) + \deg(P_2) - \deg(Q_2) = \deg(F_1) + \deg(F_2)$.
2. De même, $\deg(F_1 + F_2) = \deg(F_1 Q_2 + F_2 Q_1) - \deg(Q_1 Q_2)$, or $\deg(F_1 Q_2 + F_2 Q_1) \leq \max(\deg(F_1 Q_2); \deg(F_2 Q_1))$, donc on a $\deg(F_1 + F_2) \leq \deg(F_1 Q_2) - \deg(Q_1 Q_2)$ ou $\deg(F_1 + F_2) \leq \deg(F_2 Q_1) - \deg(Q_1 Q_2)$, c'est-à-dire $\deg(F_1 + F_2) \leq \deg(F_1)$ ou $\deg(F_1 + F_2) \leq \deg(F_2)$, finalement, $\deg(F_1 + F_2) \leq \max(\deg(F_1); \deg(F_2))$.

Remarque 7.5. Encore une fois, à vous de vérifier que les calculs précédents ne dépendent pas de la représentation choisie pour F_1 et F_2.

7.1.5. Zéros et pôles

Définition 7.7 (Zéro et pôle). Soit $F = \frac{P}{Q} \in \mathbb{K}(X)$ irréductible et $\lambda \in \mathbb{K}$.

— λ est un *zéro* de F si et seulement si λ est une racine de P. La *multiplicité* de λ en tant que zéro de F est égale à la multiplicité de λ en tant que racine de P.

— λ est un *pôle* de F si et seulement si λ est une racine de Q. La *multiplicité* de λ en tant que pôle de F est égale à la multiplicité de λ en tant que racine de Q.

Remarque 7.6. Cette fois nous avons imposé à la fraction rationnelle d'être sous forme irréductible. Cela implique qu'un réel ne peut pas être à la fois pôle et zéro : un réel est soit un pôle, soit un zéro, soit rien du tout.

Exemple 7.1. Dans $\mathbb{R}[X]$, la fraction $F(X) = \frac{(1+X+X^2)(X-1)^3(X-2)(X+3)^2(X+4)}{X^5(X-2)(X-3)^3(X^2+1)^2(X-5)(X-4)^2}$ a pour zéros :

— 1 de multiplicité 3.

— -3 de multiplicité 2.

— -4 de multiplicité 1.

et pour pôles :

— 0 de multiplicité 5.

— 3 de multiplicité 3 (pôle triple).

— 5 de multiplicité 1 (pôle simple).

— 4 de multiplicité 2 (pôle double).

2 n'est ni un pôle ni un zéro.

Théorème 7.1 ((et définition) Partie entière). Soit $F = \frac{P}{Q} \in \mathbb{K}(X)$. Il existe un unique polynôme $E \in \mathbb{K}[X]$ et une unique fraction rationnelle $R \in \mathbb{K}(X)$ tels que

$$F = E + R \qquad \text{et } \deg(R) < 0$$

Le polynôme E est appelé *partie entière* de F.

Démonstration.

Existence : Effectuons la division euclidienne de P par Q : il existe A et B éléments de $K[X]$ tels que $P = AQ + B$ et $\deg(B) < \deg(Q)$. Nous avons alors $F = \frac{AQ+B}{Q} = A + \frac{B}{Q}$. Posons alors $E = A$ et $R = \frac{B}{Q}$. Nous avons $\deg(R) = \deg(B) - \deg(Q) < 0$ et $F = E + R$.

Unicité : Soit E_1, E_2, R_1, R_2 des polynômes et fractions rationnelles satisfaisant l'énoncé du théorème. Alors $E_1 + R_1 = E_2 + R_2$, d'où $E_1 - E_2 = R_2 - R_1$. Ainsi, $\deg(E_1 - E_2) = \deg(R_2 - R_1) \leq \max(\deg(R_1); \deg(R_2)) < 0$. Le polynôme $E_1 - E_2$ est donc nul et $R_1 = R_2$.

Remarque 7.7.

— La démonstration précédente vous permet de constater que pour calculer la partie entière d'une fraction rationnelle il suffit de faire la division euclidienne du numérateur par le dénominateur.

— Cela veut aussi dire que si le degré de la fraction rationnelle est négatif, alors la partie entière est nulle.

7.2. Décomposition en éléments simples

7.2.1. Théorie

Théorème 7.2 (Décomposition sur $\mathbb{C}$). Soit $F \in \mathbb{C}(X)$ de partie entière E et de pôles distincts $\lambda_1, \ldots, \lambda_r$ de multiplicités respectives $m_1, \ldots m_r$. Il existe une unique famille $(a_{ik})_{\substack{1\leq i\leq r\\ 1\leq k\leq m_i}}$ de nombres complexes telle que :

$$F = E + \sum_{i=1}^{r} \underbrace{\sum_{k=1}^{m_i} \frac{a_{ik}}{(X-\lambda_i)^k}}_{\text{Partie polaire associée au pôle } \lambda_i}$$

Cette décomposition de F est appelée décomposition en éléments simples sur $\mathbb{C}$.

Démonstration. Hors programme. $\square$

Théorème 7.3 (Décomposition sur $\mathbb{R}$). Soit $F = \frac{P}{Q} \in \mathbb{R}(X)$ <u>irréductible</u> de partie entière E. Soit la décomposition de Q en facteurs irréductibles :

$$Q = \alpha \prod_{i=1}^{r} (X-\lambda_i)^{m_i} \prod_{j=1}^{s} (X^2 + b_j X + c_j)^{n_j}$$

Il existe des familles uniques $(a_{ik})_{\substack{1\leq i\leq r\\ 1\leq k\leq m_i}}$, $(u_{jk})_{\substack{1\leq j\leq s\\ 1\leq k\leq n_j}}$, $(v_{jk})_{\substack{1\leq j\leq s\\ 1\leq k\leq n_j}}$ telles que :

$$F = E + \sum_{i=1}^{r} \underbrace{\sum_{k=1}^{m_i} \frac{a_{ik}}{(X-\lambda_i)^k}}_{\text{Partie polaire associée au pôle } \lambda_i} + \sum_{j=1}^{s} \underbrace{\sum_{k=1}^{n_j} \frac{u_{jk}X + v_{jk}}{(X^2 + b_j X + c_j))^k}}_{\text{Élément simple de seconde espèce}}$$

Cette décomposition de F est appelée décomposition en éléments simples sur $\mathbb{R}$.

Démonstration. Hors programme. □

Remarque 7.8. Bon...
Ok, ces deux théorèmes vous semblent un peu abscons. Passons donc tout de suite à la pratique pour voir ce qu'il se passe !

7.2.2. Pratique

Nous allons calculer la décomposition en éléments simples de

$$F(X) = \frac{(X-1)(X+3)}{X^3(X-3)^4(X^2+1)^2(X-5)(X-4)^2}$$

Le degré du numérateur étant inférieur à celui du dénominateur, la partie entière de F est nulle. D'après le théorème 7.3, il existe des réels tels que F se décompose ainsi :

$$F(X) = \frac{a}{X} + \frac{b}{X^2} + \frac{c}{X^3} + \frac{d}{X-3} + \frac{e}{(X-3)^2} + \frac{f}{(X-3)^3} + \frac{g}{(X-3)^4} + \frac{h}{X-5} + \frac{j}{X-4} + \frac{k}{(X-4)^2} + \frac{lX+m}{X^2+1} + \frac{nX+o}{(X^2+1)^2}$$

Bien sûr on pourrait tout développer, identifier tous les coefficients et, après 42 pages de calculs et 37 relectures pour traquer les erreurs de calculs, déterminer les 14 inconnues.

Nous allons voir des méthodes systématiques et efficaces adaptées à chaque cas.

7.2.2.1. Éléments simples de première espèce

7.2.2.1.1. Pôle simple Soit λ un pôle simple de $F = \frac{P}{Q}$, alors $Q = (X-\lambda)Q_1$ avec $Q_1(\lambda) \neq 0$ (racine simple).

7.2.2.1.1.1. Multiplication-évaluation Nous avons $F = \frac{P}{Q} = \frac{a}{X-\lambda} + R$. En multipliant l'égalité par $X-\lambda$ et en simplifiant la fraction, nous obtenons $\frac{P}{Q_1} = a + (X-\lambda)R$. Nous avons donc

$$a = \frac{P(\lambda)}{Q_1(\lambda)}$$

Appliquons cela à notre exemple pour calculer h (5 étant le seul pôle simple) :

$$h = \frac{(5-1)(5+3)}{5^3(5-3)^4(5^2+1)^2(5-4)^2} = \frac{1}{5^3 \times 2 \times 13^2} = \frac{1}{42250}$$

7.2.2.1.1.2. Dérivation Remarquons que $Q' = (X-\lambda)Q_1' + Q_1$ donc $Q'(\lambda) = Q_1(\lambda)$. Ainsi, en reprenant le résultat de la méthode précédente :

$$a = \frac{P(\lambda)}{Q_1(\lambda)} = \frac{P(\lambda)}{Q'(\lambda)}$$

Cette méthode fonctionne exclusivement sur les pôles simples et est surtout pratique lorsque l'on connait la forme développée du dénominateur et que l'on ne veut pas le factoriser (par exemple si le dénominateur vaut $X^n - 1$).

7.2.2.1.2. Pôles doubles Soit λ un pôle double de $F = \frac{P}{Q}$, alors $Q = (X-\lambda)^2 Q_1$ et $Q_1(\lambda) \neq 0$.

7.2.2.1.2.1. Premier coefficient Nous pouvons utiliser le principe de « multiplication-évaluation » du paragraphe 7.2.2.1.1.1 : Nous avons $F = \frac{P}{Q} = \frac{a}{(X-\lambda)^2} + \frac{b}{(X-\lambda)} + R$. En multipliant l'égalité pas $(X-\lambda)^2$, nous obtenons

$$\frac{P}{Q_1} = a + b(X-\lambda) + (X-\lambda)^2 R \tag{7.1}$$

Nous avons donc

$$a = \frac{P(\lambda)}{Q_1(\lambda)}$$

Appliquons cela à notre exemple pour calculer k.

$$k = \frac{3\times 7}{4^3 \times 17^2 \times (-1)}$$
$$k = \frac{-21}{18496}$$

7.2.2.1.2.2. Second coefficient Maintenant que l'on connaît le coefficient a, il existe deux méthodes pour calculer b.

7.2.2.1.2.2.1. Soustraction Nous pouvons calculer $F_1 = F - \frac{a}{(X-\lambda)^2}$. Après mise au même dénominateur, le numérateur et le dénominateur se simplifient forcément (sinon il y a une erreur de calcul) et λ devient une racine simple de F_1. Le coefficient b peut alors se calculer avec les méthodes vues au paragraphe 7.2.2.1.1.

Appliquons cela au calcul de j.

$$F + \frac{21}{18496(X-4)^2} = \frac{18496(X-1)(X+3) + 21X^3(X-3)^4(X^2+1)^2(X-5)}{18496X^3(X-3)^4(X^2+1)^2(X-5)(X-4)^2}$$
$$= \frac{\begin{gathered}21X^{11} - 273X^{10} + 1344X^9 - 3276X^8 + 4746X^7 - 5754X^6 + 5460X^5 \\ - 3108X^4 + 609X^3 - 6069X^2 - 5780X + 13872\end{gathered}}{18496X^3(X-3)^4(X^2+1)^2(X-5)(X-4)}$$

Bien sûr, en exercice et en examen vous n'aurez pas de tels calculs à faire à la main...

On trouve alors

$$j = \frac{\begin{gathered}21\times 4^{11} - 273\times 4^{10} + 1344\times 4^9 - 3276\times 4^8 + 4746\times 4^7 - 5754\times 4^6 + 5460\times 4^5 \\ - 3108\times 4^4 + 609\times 4^3 - 6069\times 16 - 5780\times 4 + 13872\end{gathered}}{4^3 \times 17^2(-1)}$$
$$j = \frac{6019}{1257728}$$

7.2.2.1.2.2.2. Dérivation En dérivant l'équation (7.1), nous obtenons $\left(\frac{P}{Q_1}\right)' = b + (X-\lambda)(2R + (X-\lambda)R')$. Ainsi, en évaluant en λ nous avons :

$$b = \left(\frac{P}{Q_1}\right)'(\lambda)$$

Cette méthode demande d'être capable de calculer simplement la dérivée de la fraction rationnelle $\frac{P}{Q_1}$, c'est-à-dire que P et Q_1 devront de préférence être sous forme développé.

7.2.2.1.3. Pôles d'ordre supérieur ou égal à 3

7.2.2.1.3.1. Premier coefficient Les méthodes de « multiplication-évaluation » vues aux paragraphes 7.2.2.1.1.1 et 7.2.2.1.2.1 peuvent se généraliser à toutes les multiplicités. Ainsi nous pouvons calculer simplement les coefficients

c et g :

$$
\begin{aligned}
c &= \frac{-3}{(-3)^4(-5)(-4)^2} \\
c &= \frac{1}{27 \times 5 \times 16} \\
c &= \frac{1}{2160}
\end{aligned}
$$

$$
\begin{aligned}
g &= \frac{2 \times 6}{3^3 \times 10^2 \times (-2)} \\
g &= \frac{-1}{450}
\end{aligned}
$$

7.2.2.1.3.2. Division selon les puissances croissantes Pour calculer les autres coefficients, on peut toujours appliquer la méthode de soustraction vue au paragraphe 7.2.2.1.2.2.1. Au fur et à mesure des soustractions, les fractions se simplifient, mais les calculs restent lourds. La méthode de la « division selon les puissances croissantes » permet de calculer tous les coefficients d'un pôle d'un coup.

Pour calculer les coefficients du pôle λ de multiplicité n, on effectue le changement de variable $Y = X - \lambda$, et on ré-écrit F en fonction de Y. Le nouveau pôle est alors 0, de multiplicité n. Nous avons donc $F \times Y^n = \frac{P \times Y^n}{Q} = \frac{P}{Q_1} = a_0 + a_1 Y + \cdots + a_{n-1} Y^{n-1} + Y^n R$. Les n coefficients se trouvent alors en faisant une division selon les puissances croissantes de P par Q_1, que l'on arrête dès que l'on obtient le n^{e} coefficient (correspondant au monôme de degré $n-1$).

Le principe est le même que la division euclidienne, mis à part qu'au lieu d'ordonner les polynômes selon les puissances décroissantes, on les ordonne selon les puissances croissantes. Bien évidement, pour se faire, les polynômes doivent être sous forme développée.

Calculons les coefficients d, e, f et g correspondant au pôle 3 de multiplicité 4. Effectuons le changement de variable $Y = X - 3$, c'est-à-dire $X = Y + 3$. Ainsi

$$
\begin{aligned}
F(Y) &= \frac{(Y+2)(Y+6)}{(Y+3)^3 Y^4 (Y^2+6Y+10)^2 (Y-2)(Y-1)^2} \\
F(Y) &= \frac{Y^2+8Y+12}{Y^4\left(Y^{10}+17Y^9+112Y^8+314Y^7+29Y^6-1919Y^5-3538Y^4+1168Y^3+7596Y^2+1620Y-5400\right)}
\end{aligned}
$$

Lors de la division, il faut garder en tête que l'on ne s'intéresse qu'aux degrés inférieurs ou égaux à 3 (pôle de multiplicité 4).

$$
\begin{array}{rrrrrrrr|l}
12 & + & 8Y & + & Y^2 & & & & -5400+1620Y+7596Y^2+1168Y^3\dots \\
\cline{9-9}
-(12 & - & \frac{18}{5}Y & - & \frac{422}{25}Y^2 & - & \frac{584}{225}Y^3 & +\dots) & -\frac{1}{450}-\frac{29}{13500}Y-\frac{89}{22500}Y^2-\frac{14243}{3037500}Y^3 \\
\cline{1-8}
 & & \frac{58}{5}Y & + & \frac{447}{25}Y^2 & + & \frac{584}{225}Y^3 & +\dots & \\
 & - & (\frac{58}{5}Y & - & \frac{87}{25}Y^2 & - & \frac{6119}{375}Y^3 & +\dots) & \\
\cline{2-8}
 & & & + & \frac{534}{25}Y^2 & + & \frac{21277}{1125}Y^3 & +\dots & \\
 & & & - & (\frac{534}{25}Y^2 & - & \frac{801}{125}Y^3 & +\dots) & \\
\cline{4-8}
 & & & & & + & \frac{28486}{1125} & +\dots & \\
 & & & & & - & \frac{28486}{1125} & &
\end{array}
$$

Nous avons donc $g = -\frac{1}{450}$, $f = -\frac{29}{13500}$, $e = -\frac{89}{22500}$ et $d = -\frac{14243}{3037500}$. On remarque que l'on trouve le même g que précédemment (ouf !).

On peut refaire cette méthode pour calculer a, b et c (pôle triple). Les calculs nous donnent $c = \frac{1}{2160}$, $b = \frac{41}{64800}$ et $a = -\frac{4741}{7776000}$.

À ce stade, nous pouvons écrire :

$$F = -\frac{4741}{7776000X} + \frac{41}{64800X^2} + \frac{1}{2160X^3} - \frac{14243}{3037500(X-3)} - \frac{89}{22500(X-3)^2} - \frac{29}{13500(X-3)^3} - \frac{1}{450(X-3)^4} + \frac{1}{42250(X-5)} + \frac{6019}{1257728(X-4)} - \frac{21}{18496(X-4)^2} + \frac{lX+m}{(X^2+1)} + \frac{nX+o}{(X^2+1)^2}$$

Il nous reste les quatre coefficients correspondants au polynôme de degré 2 irréductible.

7.2.2.2. Éléments simples de seconde espèce

On appelle ainsi les éléments simples faisant intervenir des polynômes du second degré irréductibles (dans $\mathbb{R}$).

7.2.2.2.1. Passage dans $\mathbb{C}$ (pour les pôles simples) Pour calculer les coefficients correspondant, on peut faire la décomposition dans $\mathbb{C}$.

Nous avons vu dans le chapitre 6 « Polynômes » que si un nombre complexe est racine d'un polynôme de $\mathbb{R}[X]$ alors son conjugué aussi. Un élément simple de seconde espèce (dans $\mathbb{R}(X)$) se met donc (dans $\mathbb{C}(X)$) sous la forme

$$\frac{a}{X-\lambda} + \frac{b}{X-\overline{\lambda}}$$

avec a et b éléments de $\mathbb{C}$. Or, puisque la fraction rationnelle (F) appartient à $\mathbb{R}(X)$, on a $F(X) = \overline{F(X)}$, d'où nécessairement $\frac{a}{X-\lambda} + \frac{b}{X-\overline{\lambda}} = \frac{\overline{a}}{X-\overline{\lambda}} + \frac{\overline{b}}{X-\lambda}$. Par unicité de la décomposition, on a

$$b = \overline{a}$$

Finalement, pour décomposer dans $\mathbb{C}(X)$, il suffit de trouver l'un des coefficients, le second étant son conjugué. Pour revenir à $\mathbb{R}(X)$ on regroupe entre eux les éléments simples correspondant à des racines conjuguées.

> **Exemple 7.2.** Prenons un exemple particulier pour voir ce qu'il se passe. Soit $F = \frac{X^2+X-1}{X(X^2+1)} = \frac{a}{X} + \frac{bX+c}{X^2+X+1}$. En décomposant F dans $\mathbb{C}(X)$, on obtient $F = -\frac{1}{X} + \frac{1-\frac{i}{2}}{X-i} + \frac{1+\frac{i}{2}}{X+i}$. En mettant les deux dernières fractions au même dénominateur, nous obtenons $F = -\frac{1}{X} + \frac{\left(1-\frac{i}{2}\right)(X+i)+\left(1+\frac{i}{2}\right)(X-i)}{X^2+1} = -\frac{1}{X} + \frac{2X+1}{X^2+1}$.

Remarque 7.9. Attention, cette méthode ne fonctionne que si le pôle est simple. En effet, si le pôle est multiple, lors de la mise sous même dénominateur, le polynôme du numérateur deviendra alors de degré supérieur ou égal à 2, ce qui n'est donc pas un élément simple. Il faudra alors le redécomposer en éléments simples.

7.2.2.2.2. Multiplication-évaluation dans $\mathbb{C}$ Une autre méthode consiste à utiliser la méthode de multiplication-évaluation, mais dans $\mathbb{C}$. Il faut garder en tête que les coefficients sont forcément réels.

Reprenons notre exemple initial et calculons n et o : Multiplions F par $(X^2+1)^2$ et évaluons le résultat en $X = i$. Nous avons alors

$$ni + o = \frac{(i-1)(i+3)}{i^3(i-3)^4(i-5)(i-4)^2}$$
$$ni + o = \frac{291}{9392500}i - \frac{4837}{9392500}$$

En identifiant parties réelles et imaginaires, nous obtenons $n = \frac{291}{9392500}$ et $o = -\frac{4837}{9392500}$.

7.2.2.3. Autres méthodes

Il nous reste encore deux coefficients à calculer. Voyons différentes méthodes moins efficaces que les précédentes mais qui peuvent parfois permettre de finir rapidement les calculs.

7.2.2.3.1. Méthode de la limite Après avoir multiplié F par X, on calcule la limite lorsque X tend vers $+\infty$ des deux écritures de XF. Nous obtenons ainsi des relations entre les coefficients. Cette méthode n'est donc efficace que lorsqu'il reste peu de coefficients à calculer. Notez surtout qu'elle est totalement inutile si vous n'avez pas supprimé la partie entière. En effet, dans ce cas, $XF(X)$ tendra vers $+\infty$.

Sur notre exemple on a clairement $\lim_{X\to+\infty} XF = 0$, ce qui nous donne

$$0 = a+d+h+j+l$$

D'où $l = -a-d-h-j = \frac{4741}{7776000} + \frac{14243}{3037500} - \frac{1}{42250} - \frac{6019}{1257728} = \frac{5080047}{10378712500}$

7.2.2.3.2. Évaluation On peut évaluer F en certaines valeurs précises afin d'obtenir d'autres relations entre les coefficients et calculer les derniers coefficients restants. D'une part,

$$F(1) = \frac{0\times 4}{1^3(-2)^4(2)^2(-3)(-3)^2} = 0$$

D'autre part

$$F(1) = a+b+c-\frac{d}{2}+\frac{e}{4}-\frac{f}{8}+\frac{g}{16}-\frac{h}{4}-\frac{j}{3}+\frac{k}{9}+\frac{l+m}{2}+\frac{n+o}{4}$$

Ainsi,

$$\begin{aligned}
m &= -2a-2b-2c+d-\frac{e}{2}+\frac{f}{4}-\frac{g}{8}+\frac{h}{2}+\frac{2j}{3}-\frac{2k}{9}-l-\frac{n+o}{2}\\
m &= 2\times\frac{4741}{7776000}-2\times\frac{41}{64800}-\frac{2}{2160}-2\times\frac{14243}{3037500\times 2}+2\times\frac{89}{22500\times 4}-2\times\frac{29}{13500\times 8}+\frac{2}{450\times 16}\\
&\quad+\frac{2}{42250\times 4}+2\times\frac{6019}{1257728\times 3}+2\times\frac{21}{18496\times 9}-\frac{291-4837}{2\times 9392500}-\frac{5080047}{10378712500}\\
m &= -\frac{3816337}{5189356250}
\end{aligned}$$

7.2.2.3.3. Parité/imparité En exploitant la parité ou l'imparité de F nous pouvons faire apparaitre des relations simples entre les coefficients.

Exemple 7.3. $F = \frac{X}{X^2-1} = \frac{a}{X-1} + \frac{b}{X+1}$. Puisque F est impaire ($F(X) = -F(-X)$) nous avons $\frac{a}{-X-1} + \frac{b}{-X+1} = -\left(\frac{a}{X-1} + \frac{b}{X+1}\right)$, c'est-à-dire $\frac{-b}{X-1} + \frac{-a}{X+1} = \frac{-a}{X-1} + \frac{-b}{X+1}$. Par unicité des coefficients (oui, rappelez-vous des théorèmes 7.2 et 7.3), On en déduit que $a = b$.
Il ne reste plus qu'à calculer l'un des deux coefficients par « multiplication-évaluation ». On obtient

$$F = \frac{1}{2(X-1)} + \frac{1}{2(X+1)}$$

Au final, nous avons :

$$\begin{aligned}
F = &-\frac{4741}{7776000X}+\frac{41}{64800X^2}+\frac{1}{2160X^3}-\frac{14243}{3037500(X-3)}-\frac{89}{22500(X-3)^2}-\frac{29}{13500(X-3)^3}-\frac{1}{450(X-3)^4}\\
&+\frac{1}{42250(X-5)}+\frac{6019}{1257728(X-4)}-\frac{21}{18496(X-4)^2}+\frac{7(725721X-1090382)}{10378712500(X^2+1)}+\frac{291X-4837}{9392500(X^2+1)^2}
\end{aligned}$$

Deuxième partie

S2

8. Groupes

8.1. Lois de composition interne

8.1.1. Définitions

Définition 8.1 (Loi de composition interne). Soit E un ensemble.

— On appelle *loi de composition interne* (ou *opération*) sur E une application de $E \times E$ dans E.

— On appelle *magma* tout couple $(E\,;\star)$ constitué d'un ensemble E et d'une loi de composition interne $\star$ sur E.

Remarque 8.1. Vous connaissez déjà un grand nombre de lci :

— Sur $\mathbb{R}$ ou $\mathbb{C}$: $+$, $-$, $\times$, $\div$.

— Sur les fonctions : $+$, $-$, $\circ$, mais aussi $\times$ et $\div$ lorsque c'est définit.

— Sur $\mathbb{Z}$: $+$, $-$, $\times$, $\wedge$, $\vee$.

— Sur l'ensemble des parties d'un ensemble : $\cap$, $\cup$, Δ.

Lorsque l'on a à faire à une loi pas habituelle, on note ces lois $\top$ (truc), $\perp$ (machin), $\star$.

Définition 8.2 (Stabilité). Soit $(E\,;\star)$ un magma. Une partie A de E est dite *stable* par « $\star$ » lorsque $\forall (x;y) \in A^2$, $x \star y \in A$. L'application de $A \times A$ dans A ainsi définie est appelée l'*opération induite* sur A.

Définition 8.3 (Loi produit). Soit $(E_1\,;\top)$ et $(E_2\,;\perp)$ deux magmas. On définit sur $E_1 \times E_2$ la *loi produit* par :

$$(x_1\,;x_2) \star (y_1\,;y_2) = (x_1 \top y_1\,;x_2 \perp y_2)$$

La loi $\star$ ainsi définie est une loi de composition interne sur $E_1 \times E_2$.

Exemple 8.1. Soit $z = \rho e^{i\theta}$ et $z' = \rho' e^{i\theta'}$ deux nombres complexes, alors $zz' = \rho\rho' e^{i(\theta+\theta')}$. Si on munit $\mathbb{R}_+^*$ de la loi de multiplication usuelle et $\mathbb{R}$ de la loi d'addition, alors la loi produit sur $\mathbb{R}_+^* \times \mathbb{R}$ est définie par $(\rho\,;\theta) \star (\rho'\,;\theta') = (\rho\rho'\,;\theta + \theta')$.

Définition 8.4 (Associativité, commutativité). Soit $(E\,;\star)$ un magma. La loi $\star$ est dite :

— *Associative* si et seulement si $\forall (x;y;z) \in E^3$, $x \star (y \star z) = (x \star y) \star z$ (dans ce cas on dit que $(E\,;\star)$ est un magma associatif).

— *Commutative* si et seulement si $\forall (x;y) \in E^2$, $x \star y = y \star x$ (dans ce cas on dit que $(E\,;\star)$ est un magma commutatif).

Exemple 8.2. Vous connaissez déjà plein de lois associatives ou commutatives.

Remarque 8.2. Comme nous l'avons vu dans un chapitre antérieur, l'associativité permet de se passer des parenthèses dans l'écriture des calculs : $(x \star y) \star z = x \star y \star z$.

Définition 8.5 (Distributivité). Soit E un ensemble et, $\star$ et $\top$ deux lois de composition interne sur E. On dit que $\star$ est :

— *Distributive à droite* par rapport à $\top$ si et seulement si

$$\forall (x;y;z) \in E^3,\ (x \top y) \star z = (x \star z) \top (y \star z)$$

— *Distributive à gauche* par rapport à $\top$ si et seulement si

$$\forall (x;y;z) \in E^3,\ x \star (y \top z) = (x \star y) \top (x \star z)$$

— *Distributive* par rapport à $\top$ si et seulement si elle l'est à droite et à gauche.

Exemple 8.3. Là encore vous connaissez pleins de lois distributives les unes par rapport aux autres.

8.1.2. Éléments particuliers

Définition 8.6 (Élément neutre). Soit $(E;\star)$ un magma et $e \in E$. On dit que e est *un élément neutre pour* $\star$ (ou de $(E;\star)$) si et seulement si

$$\forall x \in E,\ x \star e = e \star x = x$$

Exemple 8.4. Vous en connaissez plusieurs.

Théorème 8.1 (Unicité de l'élément neutre). Soit $(E;\star)$ un magma. Alors E possède au plus un élément neutre.

Démonstration. Si e et e' sont deux éléments neutre de E pour $\star$, alors $e = e \star e' = e'$. □

Remarque 8.3. S'il existe un élément neutre de $(E;\star)$ alors il est unique et dans ce cas on parle de <u>l'élément</u> neutre de $(E;\star)$. On le note généralement 1_E ou 1 en notation multiplicative et 0_E ou 0 en notation additive.

Définition 8.7 (Itérés). Soit $(E;\star)$ un magma dont la loi $\star$ est <u>associative</u> et possédant un élément neutre e. Pour tout $n \in \mathbb{N}$ et pour tout $x \in E$, on appelle *itéré n^e de x*, noté x^n (en notation multiplicative) ou nx (en notation additive) l'élément $\overbrace{x \star \cdots \star x}^{n \text{ fois}}$ défini par récurrence par

$x^0 = e$ et $x^{n+1} = x \star x^n$ si $n \in \mathbb{N}$	En notation multiplicative.
$0x = e$ et $(n+1)x = x \star (nx)$ si $n \in \mathbb{N}$	En notation additive.

Exemple 8.5. Dans l'ensemble des fonctions, muni de la loi de composition des fonctions, $f^2(x)$ désigne $f \circ f(x)$ et non $f(x)^2$.

Définition 8.8 (Symétrique, Inverse). Soit $(E\,;\star)$ un magma possédant un élément neutre e et $x \in E$. On dit que x

- Admet un *symétrique à droite* (ou *un inverse à droite*) si et seulement si il existe $x' \in E$ tel que $x \star x' = e$.
- Admet un *symétrique à gauche* (ou *un inverse à gauche*) si et seulement si il existe $x' \in E$ tel que $x' \star x = e$.
- Admet un *symétrique* (ou *inverse*) si et seulement si il existe $x' \in E$ qui est à la fois symétrique à droite et à gauche.
- Est *inversible* (ou *symétrisable*) si et seulement si il possède un inverse, c'est-à-dire il existe $x' \in E$ tel que

$$x \star x' = x' \star x = e$$

Exemple 8.6. Dans l'ensemble des fonctions, muni de la composition de fonctions, f est inversible (c'est-à-dire bijective) si et seulement si il existe g tel que $g \circ f = f \circ g = Id$.

Propriété 8.1 (Unicité de l'inverse (dans un magma associatif)). Soit $(E\,;\star)$ un magma associatif possédant un élément neutre e et $x \in E$.
Si x possède un symétrique à droite (x') et un symétrique à gauche (x''), alors $x' = x''$.
Si x possède un symétrique, alors il est unique.

Démonstration. On a $x \star x' = e = x'' \star x$. Par associativité de la loi $\star$, nous avons

$$x' = e \star x' = (x'' \star x) \star x' = x'' \star (x \star x') = x'' \star e = x''$$

Le second point découle du premier : si x_1 et x_2 sont deux symétriques de x, alors x_1 est un symétrique à droite et x_2 un symétrique à gauche. □

Remarque 8.4. Puisque un inverse, lorsqu'il en existe un, est unique, on peut parler du symétrique ou de l'inverse de x. On le note x^{-1} en notation multiplicative et $-x$ en notation additive (dans ce cas, on parle aussi d'*opposé*).

Propriété 8.2 (Simplification par un élément inversible). Soit $(E\,;\star)$ un magma associatif possédant un élément neutre et $(x;y;z) \in E^3$. Si x est inversible alors

$$x \star y = x \star z \Rightarrow y = z$$
$$y \star x = z \star x \Rightarrow y = z$$

Remarque 8.5.

- Attention ! Si x n'est pas inversible, les implications sont fausses.
- Il suffit que x soit inversible à gauche pour la première implication et à droite pour la seconde.

Démonstration. Il suffit de multiplier à gauche ou à droite par l'inverse de x et utiliser l'associativité de la loi $\star$. □

Définition 8.9 (Élément régulier). Soit $(E;\star)$ un magma et $a \in E$. On dit que a est :

— *Régulier à droite* si et seulement si $\forall (x;y) \in E^2$, $x \star a = y \star a \Rightarrow x = y$.

— *Régulier à gauche* si et seulement si $\forall (x;y) \in E^2$, $a \star x = a \star y \Rightarrow x = y$.

— *Régulier* s'il est régulier à droite et à gauche.

Remarque 8.6. On en déduit qu'un élément inversible (resp. à droite, resp. à gauche) est régulier (resp. à droite, resp. à gauche), mais la réciproque est fausse. En effet sur $(\mathbb{N};+)$, tout élément est régulier, cependant, seul 0 possède un inverse (un opposé en l'occurrence) car si $n \in \mathbb{N}^*$ alors $-n \notin \mathbb{N}$.

Propriété 8.3 (Inversibilité d'un produit). Soit $(E;\star)$ un magma associatif possédant un élément neutre e et $(x;y) \in E^2$.
Si x et y sont inversibles, alors $x \star y$ aussi et $(x \star y)^{-1} = y^{-1} \star x^{-1}$

Démonstration. Là encore, il suffit de faire les multiplications en utilisant l'associativité de la loi $\star$. □

Remarque 8.7. Attention à l'ordre des inverses qui est important si la loi n'est pas commutative.

Propriété 8.4 (Puissances négatives). Soit $(E;\star)$ un magma associatif possédant un élément neutre e et $x \in E$. Si x est inversible alors pour tout $n \in \mathbb{N}$, x^n aussi et

$$(x^n)^{-1} = \left(x^{-1}\right)^n$$

On note alors cet élément x^{-n}, définissant ainsi x^k pour tout $k \in \mathbb{Z}$.

Démonstration. Facile. □

Propriété 8.5 (Inversibilité de l'inverse). Soit $(E;\star)$ un magma associatif possédant un élément neutre e et $x \in E$.
Si x est inversible alors son inverse x^{-1} aussi

$$\left(x^{-1}\right)^{-1} = x$$

Démonstration. Facile. □

Exemple 8.7. Soit E et F deux ensembles non vides et $\top$ une lci sur F. On peut définir sur F^E la loi de composition interne $\star$ par :

$$\forall f,g \in F^E, f\star g : \begin{cases} E & \longrightarrow & F \\ x & \longmapsto & f(x)\top g(x) \end{cases}$$

— Si $\top$ est commutative ou associative, alors $\star$ aussi.

— Si e est l'élément neutre de $(F\,;\top)$, alors l'application constante $\varphi : \begin{cases} E & \longrightarrow & F \\ x & \longmapsto & e \end{cases}$ est l'élément neutre de $\left(F^E\,;\star\right)$.

C'est en particulier ainsi que l'on définit la somme et le produit de fonctions réelles.

8.2. Groupes

8.2.1. Groupes

Définition 8.10 (Groupe). Soit $(G;\star)$ un magma. On dit que $(G;\star)$ est un *groupe* si et seulement si :

1. La loi $\star$ est associative.
2. La loi $\star$ admet un élément neutre.
3. Tout élément est inversible (dans G).

Un groupe $(G;\star)$ est dit *commutatif* ou *abélien* lorsque la loi $\star$ est commutative.

Remarque 8.8.

— Un groupe possédant un élément neutre il est nécessairement non vide.

— Un groupe est un magma et est donc noté $(G;\star)$, cependant, afin d'alléger les notations, on omet souvent la loi $\star$. On dit par exemple simplement « Soit G un groupe ».

— Par convention, la loi d'un groupe est notée multiplicativement ($\star$ ou $\cdot$), dans ce cas il arrive souvent que l'on omette carrément la loi en notant simplement xx' au lieu de $x\star x'$.

— Par convention encore, si le groupe est commutatif, on note la loi additivement, c'est-à-dire $(G;+)$.

Exemple 8.8.

— $(\mathbb{Z};+)$, $(\mathbb{Q};+)$, $(\mathbb{R};+)$, $(\mathbb{C};+)$, $(\mathbb{Q}^*;\times)$, $(\mathbb{R}^*;\times)$ et $(\mathbb{C}^*;\times)$ sont des groupes abéliens (notez bien que l'on a retiré le zéro pour les trois derniers).

— $(\mathbb{N};+)$, $(\mathbb{Z}^*;\times)$, $(\mathbb{R};\times)$ ne sont pas des groupes (un ou plusieurs éléments ne sont pas inversibles).

Propriété 8.6. Dans un groupe tout élément est régulier. C'est-à-dire, si $(G;\star)$ est un groupe alors

$$\forall(a;b;c)\in G^3,\ a\star c = b\star c \Rightarrow a=b \quad \text{et} \quad c\star a = c\star b \Rightarrow a=b$$

Démonstration. Nous avons déjà remarqué que « inversible »$\Rightarrow$« régulier ». Or tout élément d'un groupe est inversible. $\square$

Propriété 8.7. Soit $(G_1\,;\star)$, $(G_2\,;\top)$ deux groupes et e_1, e_2 leurs éléments neutres.

- L'ensemble $G_1 \times G_2$ est un groupe pour la loi produit.
- L'élément neutre de $G_1 \times G_2$ est $(e_1\,;e_2)$.
- Pour $(x_1\,;x_2) \in G_1 \times G_2$, $(x_1\,;x_2)^{-1} = \left(x_1^{-1}\,;x_2^{-1}\right)$.

Démonstration. Les deux derniers points permettent de montrer le premier. Commençons par ceux-là. Notons $\perp$ la loi produit sur $G_1 \times G_2$.

- Élément neutre : Soit $(x_1\,;x_2) \in G_1 \times G_2$, alors

$$(x_1\,;x_2)\perp(e_1\,;e_2) = (x_1 \star e_1\,;x_2 \top e_2) = (x_1\,;x_2) = (e_1 \star x_1\,;e_2 \star x_2) = (e_1\,;e_2)\perp(x_1\,;x_2)$$

- Inverse : Soit $(x_1\,;x_2) \in G_1 \times G_2$, alors

$$(x_1\,;x_2)\perp\left(x_1^{-1}\,;x_2^{-1}\right) = \left(x_1 \star x_1^{-1}\,;x_2 \top x_2^{-1}\right) = (e_1\,;e_2)$$

- Les lois $\star$ et $\top$ étant associatives, la loi produit aussi.

D'où la propriété. □

Définition 8.11 (Permutation, Groupe symétrique). Soit E un ensemble non vide.

- On appelle *permutation* de E toute bijection de E dans E.
- On appelle *groupe symétrique* de E l'ensemble des permutations de E, noté S_E. Le magma $(S_E\,;\circ)$ est un groupe d'élément neutre Id_E.

Démonstration.

1. La loi de composition $\circ$ est associative.
2. L'identité de E, Id_E est bien une bijection de E dans E. De plus pour toute application f (en particulier une bijection) de E dans E, nous avons $f \circ Id_E = Id_E \circ f = f$. Donc Id_E est bien l'élément neutre de $\circ$.
3. Si f est une bijection de E sur E alors sa bijection réciproque f^{-1} est aussi une bijection de E sur E.

8.2.2. Sous-groupes

Définition 8.12 (Sous-groupe). Soit $(G\,;\star)$ un groupe et H une partie de G. On dit que H est *un sous groupe de* G si et seulement si $(H\,;\star)$ est un groupe.

Remarque 8.9. Un sous groupe est donc tout simplement un groupe dans un autre. Cependant, est-ce qu'ils partagent le même élément neutre ? Si $x \in H$, l'inverse de x dans H est-il le même que sont inverse dans G ?

Théorème 8.2 (Élément neutre et inverse dans un sous-groupe). Soit G un groupe d'élément neutre 1_G et H un sous-groupe de G.

1. $1_G \in H$
2. H est stable par passage à l'inverse : $\forall h \in H, h^{-1} \in H$.

Démonstration. 1. Puisque H est un groupe il possède un élément neutre $1_H \in H \subset G$, qui vérifie $1_H 1_H = 1_H$. Or G possède aussi un élément neutre 1_G qui vérifie $1_H 1_G = 1_H$. Ainsi, $1_H 1_H = 1_H 1_G$. Puisque G est un groupe on peut simplifier à gauche par 1_H, pour obtenir $1_H = 1_G \in H$.

2. Soit $h \in H$, notons h' son inverse dans H et h^{-1} son inverse dans G. Nous avons alors $h' = 1_G h' = (h^{-1}h)h' = h^{-1}(hh') = h^{-1}1_G = h^{-1}$. Donc $h^{-1} \in H$.

Théorème 8.3 (Caractérisation des sous-groupes). Soit G un groupe d'élément neutre 1_G et H une partie de G. Alors H est un sous-groupe de G si et seulement si :

1. $1_G \in H$
2. H est stable par produit et passage à l'inverse, c'est-à-dire $\forall x, y \in H,\ xy^{-1} \in H$.

Démonstration. $\boxed{\Rightarrow}$ Si H est un sous-groupe, alors il est stable par produit, nous avons vu que $1_G \in H$ et que H est stable par passage à l'inverse. D'où le résultat.

$\boxed{\Leftarrow}$ — $1_G \in H$ et est élément neutre.

— Comme $1_G \in H$ et d'après le deuxième point de l'hypothèse, si $h \in H$, alors $h^{-1} = 1_G h^{-1} \in H$, d'où H stable par passage à l'inverse.

— Ensuite si h et h' sont deux éléments de H, alors $h'h = h'(h^{-1})^{-1} \in H$, donc H est stable par produit (le produit est bien une loi de composition interne de H).

— L'associativité de la loi sur G est transmise sur H.

La loi de G est donc bien une lci sur H, associative, possédant un élément neutre (1_G) dans H et tout élément de H est inversible (dans H), conclusion, H est bien un groupe pour la loi de G, donc un sous-groupe de G.

Remarque 8.10. Pour montrer qu'une partie est un sous-groupe à partir de la seule définition, nous avons de nombreuses choses à montrer : associativité de la loi, la stabilité par la loi et la stabilité par l'inverse. Or si on utilise le théorème précédent, on a juste à montrer que l'élément neutre est dans H (souvent simple) et la stabilité par le produit xy^{-1}. En pratique on utilise donc le théorème de caractérisation.

De même, pour montrer qu'un magma est un groupe, on cherche plutôt à montrer que c'est un sous-groupe d'un groupe connu.

Exemple 8.9. Pour tout groupe G, G et $\{1_G\}$ sont des sous-groupes de G.
Un sous groupe H de G différent de G et $\{1_G\}$ est appelé *sous-groupe propre* de G.

Exemple 8.10. $(\mathbb{Z};+)$ est un sous groupe de $(\mathbb{Q};+)$ qui est un sous-groupe de $(\mathbb{R};+)$ qui est un sous-groupe de $(\mathbb{C};+)$. De même pour $(\mathbb{Q}^*;\times)$, $(\mathbb{R}^*;\times)$, $(\mathbb{C}^*;\times)$.

Exemple 8.11. — $(\mathbb{U};\times)$ est un groupe.

Démonstration. Il suffit de montrer que c'est un sous-groupe de $(\mathbb{C}^*;\times)$.

— Nous avons bien $\mathbb{U} \subset \mathbb{C}^*$.

— L'élément neutre de $(\mathbb{C}^*;\times)$ est 1 et $|1| = 1$, donc $1 \in \mathbb{U}$.

— Pour tous $u, u' \in \mathbb{U}$, $\left|u^{-1}u'\right| = \left|\frac{u'}{u}\right| = \frac{|u'|}{|u|} = \frac{1}{1} = 1$, donc $u^{-1}u' \in \mathbb{U}$.

— $(\mathbb{U}_n;\times)$ est ou groupe.

Démonstration. C'est un sous-groupe de $(\mathbb{U};\times)$.

— Nous avons bien $\mathbb{U}_n \subset \mathbb{U}$.

— L'élément neutre de $(\mathbb{U};\times)$ est 1 et $1^n=1$, donc $1\in\mathbb{U}_n$.

— Pour tous $u,u'\in\mathbb{U}_n$, $\left(u^{-1}u'\right)^n=\left(\frac{u'}{u}\right)^n=\frac{u'^n}{u^n}=\frac{1}{1}=1$, donc $u^{-1}u'\in\mathbb{U}_n$.

8.3. Morphismes

8.3.1. Définitions

Définition 8.13 (Morphisme). Soit $(E\,;\star)$, $(F\,;\top)$ deux magmas et f une application de E dans F. On dit que f est un *morphisme* de $(E\,;\star)$ dans $(F\,;\top)$ si et seulement si

$$\forall(x;y)\in E^2,\ f(x\star y)=f(x)\top f(y)$$

Remarque 8.11. Un morphisme f est dit injectif, surjectif, bijectif lorsque l'application f est injective, surjective, bijective.

Exemple 8.12. L'application

$$\varphi:\left\{\begin{array}{ccc}\mathbb{R}_+^*\times\mathbb{R} & \longrightarrow & \mathbb{C}^*\\ (\rho\,;\theta) & \longmapsto & \rho e^{i\theta}\end{array}\right.$$

est un morphisme surjectif de $\left(\mathbb{R}_+^*\,;\times\right)\times(\mathbb{R}\,;+)$ (muni de la loi produit) dans $(\mathbb{C}^*\,;\times)$.
En effet, la loi produit est définie par $(\rho\,;\theta)\star(\rho'\,;\theta')=(\rho\rho'\,;\theta+\theta')$ et

$$\varphi\left((\rho\,;\theta)\star(\rho'\,;\theta')\right)=\varphi(\rho\rho'\,;\theta+\theta')=\rho\rho' e^{i(\theta+\theta')}=\left(\rho e^{i\theta}\right)\times\left(\rho' e^{i\theta'}\right)=\varphi(\rho\,;\theta)\times\varphi(\rho'\,;\theta')$$

Définition 8.14 (Homomorphisme, Endomorphisme, Isomorphisme, Automorphisme). Un morphisme est aussi appelé un *homomorphisme*.
Soit $(E\,;\star)$ un magma. On appelle *endomorphisme* de $(E\,;\star)$ un morphisme de $(E\,;\star)$ dans lui même.
On appelle *isomorphisme* un morphisme bijectif.
On appelle *automorphisme* un endomorphisme bijectif.

Définition 8.15 (Morphisme de groupes). On appelle *morphisme de groupes* un morphisme (de magmas) entre deux groupes.

8.3.2. Propriétés

Propriété 8.8.

— La composée de deux morphismes est un morphisme.

— La réciproque d'un isomorphisme est un isomorphisme.

Démonstration. Soit f un morphisme de $(E\,;\star)$ dans $(F\,;\top)$ et g un morphisme de $(F\,;\top)$ dans $(G;\Delta)$.

— Pour x et y dans E, on a

$$g\circ f(x\star y)=g\left(f(x)\top f(y)\right)=g\circ f(x)\Delta g\circ f(y)$$

Cela montre que $g\circ f$ est un morphisme.

— Si f est un isomorphisme, pour tout x et y de F, on a

$$f\left(f^{-1}(x)\star f^{-1}(y)\right)=f\left(f^{-1}(x)\right)\top f\left(f^{-1}(y)\right)=x\top y=f\left(f^{-1}(x\top y)\right)$$

Donc par injectivité de f, $f^{-1}(x\top y)=f^{-1}(x)\star f^{-1}(y)$ et f^{-1} est un morphisme.

Propriété 8.9. Soit $(G_1\,;\star)$ et $(G_2\,;\top)$ deux groupes d'éléments neutres e_1 et e_2 et f un morphisme de G_1 dans G_2. Alors

1. $f(e_1)=e_2$
2. pour tout $x\in G_1$, $f\left(x^{-1}\right)=(f(x))^{-1}$ et pour tout $n\in\mathbb{Z}$, $f(x^n)=f(x)^n$.

Démonstration.
1. Nous avons $e_1=e_1\star e_1$, donc $f(e_1)=f(e_1\star e_1)=f(e_1)\top f(e_1)$ d'où $f(e_1)\top e_2=f(e_1)\top f(e_1)$, puis en simplifiant par $f(e_1)$ (nous sommes dans un groupe), $e_2=f(e_1)$.
2. Nous avons $x\star x^{-1}=x^{-1}\star x=e_1$, donc $f(x)\top f(x^{-1})=f(x^{-1})\top f(x)=f(e_1)=e_2$. Donc $(f(x)^{-1})=f(x^{-1})$. Le dernier point est évident (récurrence par exemple).

Propriété 8.10 (Image d'un sous-groupe). Soit $(G_1\,;\star)$ et $(G_2\,;\top)$ deux groupes et f un morphisme de G_1 dans G_2. Alors si H_1 est un sous groupe de $(G_1\,;\star)$, alors $f(H_1)$ est un sous groupe de $(G_2\,;\top)$

Démonstration. Puisque H_1 est non vide, $f(H_1)$ non plus. Soit y_1 et y_2 dans $f(H_1)$, il existe donc x_1 et x_2 dans H_1 tels que $y_1=f(x_1)$ et $y_2=f(x_2)$. Nous avons alors $y_1\top y_2^{-1}=f(x_1)\top(f(x_2))^{-1}=f(x_1)\top f(x_2^{-1})=f(x_1\star x_2^{-1})$. Or H_1 est un sous-groupe donc $x_1\star x_2^{-1}\in H_1$ d'où $y_1\top y_2^{-1}\in f(H_1)$. Donc $f(H_1)$ est un sous-groupe. □

Propriété 8.11 (Image réciproque d'un sous-groupe). Soit $(G_1\,;\star)$ et $(G_2\,;\top)$ deux groupes et f un morphisme de G_1 dans G_2. Alors si H_2 est un sous groupe de $(G_2\,;\top)$, alors $f^{-1}(H_2)$ est un sous groupe de $(G_1\,;\star)$

Démonstration. Puisque $f(e_1)=e_2\in H_2$, nous avons $e_1\in f^{-1}(H_2)$ et donc $f^{-1}(H_2)$ est non vide. Soit x_1 et x_2 dans $f^{-1}(H_2)$, alors $f(x_1)$ et $f(x_2)$ sont dans H_2, d'où $f(x_1)\top(f(x_2))^{-1}\in H_2$, donc $f(x_1\star x_2^{-1})\in H_2$ et $x_1\star x_2^{-1}\in f^{-1}(H_2)$. D'où la propriété. □

8.3.3. Noyau et image

Définition 8.16 (Noyau, Image). Soit $(G_1\,;\star)$ et $(G_2\,;\top)$ deux groupes et f un morphisme de G_1 dans G_2.

— $f(G_1)$ est appelé *image* du morphisme f et est noté $\operatorname{Im} f$

— $f^{-1}(\{e_2\})$ est appelé *noyau* de f et est noté $\operatorname{Ker} f$.

Exemple 8.13. — Soit $n \in \mathbb{N}^*$. L'application $f : \left\{ \begin{array}{ccc} \mathbb{C}^* & \longrightarrow & \mathbb{C}^* \\ z & \longmapsto & z^n \end{array} \right.$ est un endomorphisme surjectif du groupe $(\mathbb{C}^* ; \times)$. Son noyau est le groupe $\mathbb{U}_n$.

— L'application $f : \left\{ \begin{array}{ccc} \mathbb{R} & \longrightarrow & \mathbb{U} \\ \theta & \longmapsto & e^{i\theta} \end{array} \right.$ est un morphisme surjectif du groupe $(\mathbb{R} ; +)$ dans le groupe $(\mathbb{U} ; \times)$. Son noyau est le groupe $2\pi\mathbb{Z}$.

— L'application $f : \left\{ \begin{array}{ccc} \mathbb{R}_+^* \times \mathbb{R} & \longrightarrow & \mathbb{C}^* \\ (\rho ; \theta) & \longmapsto & \rho e^{i\theta} \end{array} \right.$ est un morphisme surjectif du groupe $\left(\mathbb{R}_+^* \times ; \times\right)(\mathbb{R} ; +)$ dans le groupe $(\mathbb{C}^* ; \times)$. Son noyau est le groupe $\{1\} \times 2\pi\mathbb{Z}$.

— L'application $f : \left\{ \begin{array}{ccc} \mathbb{C} & \longrightarrow & \mathbb{C}^* \\ z & \longmapsto & e^z \end{array} \right.$ est un morphisme surjectif du groupe $(\mathbb{C} ; +)$ dans le groupe $(\mathbb{C}^* ; \times)$. Son noyau est le groupe $2i\pi\mathbb{Z}$.

Théorème 8.4. Soit G un groupe d'élément neutre e et f un morphisme de G dans un autre groupe. Alors f est un morphisme injectif si et seulement $\operatorname{Ker} f = \{e\}$

Démonstration. Notons $(G ; \star)$ le premier groupe et $(G_2 ; \top)$ le second groupe.

— Supposons f injectif et soit $x \in \operatorname{Ker} f$. Alors par définition, $f(x) = e_2 = f(e)$ et par injectivité de f, $x = e$. Ainsi $\{e\} \subset \operatorname{Ker} f \subset \{e\}$, d'où l'égalité.

— Supposons que $\operatorname{Ker} f = \{e\}$ et soit x_1 et x_2 dans G tels que $f(x_1) = f(x_2)$. Nous avons alors $e_2 = f(x_1) \top (f(x_2))^{-1} = f(x_1 \star x_2^{-1})$. Ainsi $x_1 \star x_2^{-1} \in \operatorname{Ker} f = \{e\}$. Donc $x_1 x_2^{-1} = e$ et $x_1 = x_2$. Ainsi, f est injective.

9. Systèmes linéaires

Vous avez rencontré pour la première fois les systèmes linéaires en 3ᵉ avec des systèmes « 2×2 ». Vous avez appris à les résoudre avec la méthode dite de « substitution ». Cette méthode semble vous avoir suffi tout au long du lycée, mais elle n'est plus viable lorsque les systèmes deviennent gros. Ce chapitre va introduire une méthode efficace et systématique (programmable sur machine) de résolution de systèmes linéaires.
Les matrices, bien que pas encore étudiées vont être utilisées comme « tableaux de nombres » nous permettant de simplifier les notations lors des résolutions de systèmes linéaires.

L'ensemble $\mathbb{K}$ désigne soit $\mathbb{R}$ soit $\mathbb{C}$, n et p sont des entiers naturels non nuls.

9.1. Définitions

Définition 9.1 (Équation linéaire). Soit $p \in \mathbb{N}^*$. On appelle *équation linéaire* à p inconnues x_1, …, x_p, une équation de la forme $a_1x_1+\dots+a_px_p = b$ avec $(a_1,\dots,a_p,b) \in \mathbb{K}^{p+1}$.
Une solution d'une telle équation est notée sous forme d'un vecteur : $(s_1,\dots,s_p) \in \mathbb{K}^p$.
On appelle *équation homogène* associée à l'équation linéaire, l'équation obtenue en remplaçant b par 0.

Remarque 9.1.

- L'équation cartésienne d'une droite du plan est une équation linéaire à deux inconnues.
- L'équation cartésienne d'un plan de l'espace est une équation linéaire à trois inconnues.
- Le vecteur nul $(0,\dots,0)$ est toujours solution d'une équation homogène.
- Si $(s_1,\dots,s_p)$ est solution d'une équation homogène alors pour tout $\lambda \in \mathbb{K}$, $(\lambda s_1,\dots,\lambda s_p)$ est aussi solution.
- Si $(s_1,\dots,s_p)$ et $(s'_1,\dots,s'_p)$ sont solutions d'une équation homogène alors $(s_1+s'_1,\dots,s_p+s'_p)$ est aussi solution. Donc l'ensemble des solutions homogènes est un espace vectoriel.

Définition 9.2 (Système linéaire). Soit n et p appartenant à $\mathbb{N}^*$. On appelle *système linéaire* de n équations à p inconnues x_1, …, x_p un système de la forme :

$$\left\{\begin{array}{rcl} a_{1,1}x_1+\dots+a_{1,p}x_p & = & b_1 \\ \vdots & & \vdots \\ a_{n,1}x_1+\dots+a_{n,p}x_p & = & b_n \end{array}\right.$$

où les $a_{i,j}$ et b_i ($1 \le i \le n$, $1 \le j \le p$) sont des éléments de $\mathbb{K}$.
Un vecteur $(s_1,\dots,s_p) \in \mathbb{K}^p$ est solution du système si et seulement si il est solution de toutes les équations linéaires composant le système.
On appelle *système d'équations homogènes* le système obtenu en remplaçant tous les b_i par 0.

Remarque 9.2.

- Le système d'équations cartésiennes d'une droite de l'espace est un système linéaire.
- Si $(s_1,\dots,s_p)$ est solution d'un système d'équations homogènes alors pour tout $\lambda\in\mathbb{K}$, $(\lambda s_1,\dots,\lambda s_p)$ est aussi solution.
- Si $(s_1,\dots,s_p)$ et $(s'_1,\dots,s'_p)$ sont solutions d'un système d'équations homogènes alors $(s_1+s'_1,\dots,s_p+s'_p)$ est aussi solution (espace vectoriel).

Définition 9.3 (Matrice associée à un système linéaire). Soit S le système linéaire

$$\left\{\begin{array}{rcl} a_{1,1}x_1+\cdots+a_{1,p}x_p & = & b_1 \\ \vdots & & \vdots \\ a_{n,1}x_1+\cdots+a_{n,p}x_p & = & b_n \end{array}\right.$$

On appelle *matrice* associée au système S le « tableau de nombre » noté

$$\begin{pmatrix} a_{1,1} & \dots & a_{1,p} \\ \vdots & \ddots & \vdots \\ a_{n,1} & \dots & a_{n,p} \end{pmatrix}$$

Définition 9.4 (Matrice augmentée). Soit S le système linéaire :

$$S=\left\{\begin{array}{rcl} a_{1,1}x_1+\cdots+a_{1,p}x_p & = & b_1 \\ \vdots & & \vdots \\ a_{n,1}x_1+\cdots+a_{n,p}x_p & = & b_n \end{array}\right.$$

la matrice A, la matrice associée :

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

et B le vecteur colonne :

$$B=\begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$$

On appelle *matrice augmentée* du système S la matrice obtenue en juxtaposant A et B :

$$(A|B)=\begin{pmatrix} a_{1,1} & \dots & a_{1,p} & b_1 \\ \vdots & \ddots & \vdots & \vdots \\ a_{n,1} & \dots & a_{n,p} & b_n \end{pmatrix}$$

Exemple 9.1. Soit le système à cinq équations et 3 inconnues :

$$S = \left\{ \begin{array}{ccccccc} 3x_1 & + & 2x_2 & + & 5x_3 & = & 7 \\ 2x_1 & + & & & 2x_3 & = & -4 \\ -x_1 & + & x_2 & - & 3x_3 & = & 2 \\ 4x_1 & - & 2x_2 & & & = & \frac{3}{2} \\ & & 3x_2 & + & 5x_3 & = & 4 \end{array} \right.$$

Alors la matrice associée est

$$\begin{pmatrix} 3 & 2 & 5 \\ 2 & 0 & 2 \\ -1 & 1 & -3 \\ 4 & -2 & 0 \\ 0 & 3 & 5 \end{pmatrix}$$

Et la matrice augmentée est

$$\begin{pmatrix} 3 & 2 & 5 & 7 \\ 2 & 0 & 2 & -4 \\ -1 & 1 & -3 & 2 \\ 4 & -2 & 0 & \frac{3}{2} \\ 0 & 3 & 5 & 4 \end{pmatrix}$$

9.2. Systèmes équivalents

Définition 9.5 (Opérations élémentaires). On appelle *opération élémentaire sur les lignes* d'un système (ou d'une matrice) l'une des opérations suivantes :

1. Échange des lignes L_i et L_j.
2. Multiplication de la ligne L_i par un scalaire $\lambda \in \mathbb{K}^*$ ($\lambda \neq 0$).
3. Ajouter $\lambda \cdot L_j$ à la ligne L_i.

Les opérations élémentaires sur les lignes sont notées ainsi :

1. $L_i \leftrightarrow L_j$
2. $L_i \leftarrow \lambda L_i$ ($\lambda \neq 0$)
3. $L_i \leftarrow L_i + \lambda L_j$

Définition 9.6 (Systèmes équivalents). Deux systèmes linéaires sont dits *équivalents* si et seulement si on passe de l'un à l'autre par une suite finie d'opérations élémentaires sur les lignes.

Définition 9.7 (Matrices équivalentes en lignes). Deux matrices sont dites *équivalentes en lignes* si et seulement si on passe de l'une à l'autre par une suite finie d'opérations élémentaires sur les lignes. Deux matrices M et M' équivalentes en lignes sont notées :

$$M \underset{L}{\sim} M'$$

Remarque 9.3. — Si on passe du système S_1 au système S_2 par une suite d'opérations élémentaires sur les lignes $O_1, \ldots, O_m$, alors la suite d'opérations $O_m^{-1}, \ldots, O_1^{-1}$ permet de passer de S_2 à S_1. Tout d'abord il faut se convaincre que les O_i^{-1} sont bien des opérations élémentaires :

1. Si $O = L_i \leftrightarrow L_j$ alors $O^{-1} = L_j \leftrightarrow L_i$ est aussi une opération élémentaire sur les lignes.
2. Si $O = L_i \leftarrow \lambda L_i$ $(\lambda \neq 0)$ alors $O^{-1} = L_i \leftarrow \frac{1}{\lambda} L_i$ est une opération élémentaire sur les lignes.
3. Si $O = L_i \leftarrow L_i + \lambda L_j$ alors $O^{-1} = L_i \leftarrow L_i - \lambda L_j$ est une opération élémentaire sur les lignes.

Ainsi si S_1 est équivalent à S_2 alors S_2 est équivalent à S_1, ce qui justifie le fait de parler de *deux systèmes équivalents*.

Même remarque avec les matrices équivalentes en lignes.

— C'est une relation d'équivalence.

Propriété 9.1. Deux systèmes équivalents ont le même ensemble de solutions.

Démonstration. Soit $(s_1, \ldots, s_p)$ une solution d'un système linéaire S_1. Soit S_2 un système équivalent à S_1 et O l'une des opérations élémentaires sur les lignes permettant de passer de S_1 à S_2.

1. Si $O = L_i \leftrightarrow L_j$ alors il est évident que les solutions ne changent pas.
2. Si $O = L_i \leftarrow \lambda L_i$ $(\lambda \neq 0)$ alors nous avons simplement multiplié la iᵉ équation par λ non nul, donc les solutions ne changent pas.
3. Si $O = L_i \leftarrow L_i + \lambda L_j$ alors, puisque $(s_1, \ldots, s_p)$ est à la fois solution de L_j et de L_i, il est aussi solution de λL_j et de $L_i + \lambda L_j$.

Donc les solutions de S_1 sont bien solutions de S_2. Puisque l'on passe de S_2 à S_1 par des opérations élémentaires sur les lignes, les solutions de S_2 sont aussi des solutions de S_1. Finalement, S_1 et S_2 ont les mêmes solutions. □

Propriété 9.2. Si on passe d'un système S_1 à un système S_2 par une suite finie d'opérations élémentaires sur les lignes, alors la matrice augmentée de S_2 s'obtient en effectuant la même suite d'opérations élémentaires sur les lignes sur la matrice augmentée de S_1.

Démonstration. C'est assez évident. Soit M_1 et M_2 les matrices augmentées de S_1 et S_2.

1. Si on échange les lignes L_i et L_j du système S_1 alors M_2 s'obtient en échangeant les lignes L_i et L_j de M_1.

2. Si on multiplie la ligne L_i du système S_1 par $\lambda \neq 0$ alors M_2 s'obtient en multipliant la ligne L_i de M_1 par $\lambda \neq 0$.
3. Si on remplace la ligne L_i du système S_1 par la ligne $L_i + \lambda L_j$ alors M_2 s'obtient en remplaçant la ligne L_i de M_1 par la ligne $L_i + \lambda L_j$.

Remarque 9.4. Cette propriété justifie le fait d'utiliser la notation matricielle pour résoudre les systèmes.

Exemple 9.2. Reprenons l'exemple 9.1 précédent et effectuons la suite d'opérations élémentaires sur les lignes suivante :

$$L_1 \leftrightarrow L_3 \qquad L_2 \leftarrow L_2 + 2L_1 \qquad L_3 \leftarrow L_3 + 3L_1 \qquad L_4 \leftarrow L_4 + 4L_1$$

Le système obtenu après la première opération (échange des lignes 1 et 3) est :

$$S_1 = \begin{cases} -x_1 + x_2 - 3x_3 = 2 \\ 2x_1 + 2x_3 = -4 \\ 3x_1 + 2x_2 + 5x_3 = 7 \\ 4x_1 - 2x_2 = \frac{3}{2} \\ 3x_2 + 5x_3 = 4 \end{cases}$$

Puis les suivantes donnent :

$$S_2 = \begin{cases} -x_1 + x_2 - 3x_3 = 2 \\ 2x_2 - 4x_3 = 0 \\ 5x_2 - 4x_3 = 13 \\ 2x_2 - 12x_3 = \frac{19}{2} \\ 3x_2 + 5x_3 = 4 \end{cases}$$

Les systèmes S_1 et S_2 sont équivalents et ont donc exactement les mêmes solutions.

9.3. Algorithme de Gauss

Vous remarquez sur l'exemple 9.2 précédent que les opérations élémentaires n'ont pas été choisies totalement par hasard. En effet, le but était de supprimer la variable x_1 de toutes les lignes sauf de la première. En réitérant le procédé on peut ainsi réduire le nombre de variables et d'équations, et résoudre ainsi facilement le système. Voyons cela plus en détail.

Afin de simplifier les notations, nous allons dorénavant travailler avec les matrices augmentées.

Définition 9.8 (Matrice échelonnée). Une matrice est dite échelonnée en lignes si elle vérifie les deux propriétés suivantes :

1. Si une ligne est entièrement nulle, toutes les lignes suivantes le sont aussi.
2. À partir de la deuxième ligne, dans chaque ligne non entièrement nulle, le premier coefficient non nul à partir de la gauche est situé à droite (strictement) du premier coefficient non nul de la ligne précédente.

Cette définition ne semble *a priori* pas très claire, pourtant elle décrit une chose très simple.

Exemple 9.3. La matrice

$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & 0 & 0 \\ 2 & 3 & 1 \end{pmatrix}$$

n'est pas échelonnée en lignes puisque la première condition n'est pas vérifiée : la deuxième ligne est entièrement nulle alors que la troisième ne l'est pas.

La matrice

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & 0 & 4 & 5 & 6 \\ 0 & 1 & 3 & 4 & 5 \\ 0 & 0 & 0 & 1 & 2 \end{pmatrix}$$

n'est pas échelonnée en lignes puisque la seconde condition n'est pas vérifiée : le premier coefficient non nul de la deuxième ligne est dans la troisième colonne alors que le premier coefficient non nul de la troisième ligne est dans la deuxième colonne.

La matrice

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 2 & 3 & 4 & 1 & 2 & 5 & 6 & 0 & 2 \\ 0 & 1 & 2 & 3 & 4 & 5 & 2 & 5 & 6 \\ 0 & 1 & 2 & 3 & 0 & 5 & 4 & 2 & 5 \\ 0 & 0 & 3 & 3 & 4 & 5 & 2 & 4 & 5 \\ 0 & 0 & 2 & 3 & 4 & 5 & 2 & 0 & 5 \\ 0 & 0 & 4 & 2 & 4 & 0 & 4 & 8 & 9 \\ 0 & 0 & 1 & 4 & 2 & 5 & 4 & 2 & 5 \\ 0 & 0 & 0 & 0 & 0 & 4 & 2 & 0 & 5 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 4 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 7 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

n'est pas une matrice échelonnée en lignes puisque la seconde condition n'est toujours pas vérifiée : le premier coefficient non nul de la deuxième ligne n'est pas à droite mais en dessous de celui de la première ligne.

La matrice

$$\begin{pmatrix}
1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\
0 & 3 & 4 & 1 & 2 & 5 & 6 & 0 & 2 \\
0 & 0 & 3 & 3 & 4 & 5 & 2 & 4 & 5 \\
0 & 0 & 0 & 0 & 0 & 4 & 2 & 0 & 5 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 4 & 2 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{pmatrix}$$

est une matrice échelonnée en lignes. Chaque « marche » de l'échelonnement fait obligatoirement une hauteur de 1. Toutes les lignes nulles se retrouvent en dernier.

Définition 9.9 (Pivot). Dans une matrice échelonnée en lignes, on appelle *pivot* le premier coefficient non nul de chaque ligne non entièrement nulle.

Exemple 9.4. Pour la matrice échelonnée en lignes de l'exemple 9.3, les pivots sont dans l'ordre : 1, 3, 3, 4, 4.

Propriété 9.3. Toute matrice non nulle est équivalente en ligne à une matrice échelonnée en lignes.

Démonstration. Admis. □

L'algorithme de Gauss est un algorithme efficace, facilement programmable sur machine et permettant de calculer les solutions de n'importe quel système linéaire. Il consiste principalement à transformer un système en un système équivalent et échelonné en lignes. Un tel système est alors facile à résoudre en « remontant les équations ». Voyons son principe sur un exemple.

Exemple 9.5. Soit le système à quatre inconnues et trois équations,

$$S = \left\{ \begin{array}{ccccccccc}
2x_1 & + & 7x_2 & + & 11x_3 & + & 15x_4 & = & 19 \\
3x_1 & + & 5x_2 & + & 8x_3 & + & 11x_4 & = & 14 \\
x_1 & + & x_2 & + & 2x_3 & + & 3x_4 & = & 4
\end{array} \right.$$

La matrice augmentée est

$$M = \left(\begin{array}{cccc|c}
2 & 7 & 11 & 15 & 19 \\
3 & 5 & 8 & 11 & 14 \\
1 & 1 & 2 & 3 & 4
\end{array} \right)$$

1. Première règle, il vaut toujours mieux avoir un pivot qui vaut 1. Nous allons donc échanger la première et

la dernière ligne : $L_1 \leftrightarrow L_3$.

$$M_1 = \left(\begin{array}{cccc|c} 1 & 1 & 2 & 3 & 4 \\ 3 & 5 & 8 & 11 & 14 \\ 2 & 7 & 11 & 15 & 19 \end{array}\right)$$

2. Le pivot de la première ligne vaut 1, nous pouvons donc remplir la première colonne avec des 0 en faisant : $L_2 \leftarrow L_2 - 3L_1$ et $L_3 \leftarrow L_3 - 2L_1$.

$$M_2 = \left(\begin{array}{cccc|c} 1 & 1 & 2 & 3 & 4 \\ 0 & 2 & 2 & 2 & 2 \\ 0 & 5 & 7 & 9 & 11 \end{array}\right)$$

3. Nous pouvons facilement diviser la deuxième ligne par 2 : $L_2 \leftarrow \frac{1}{2}L_2$.

$$M_3 = \left(\begin{array}{cccc|c} 1 & 1 & 2 & 3 & 4 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 5 & 7 & 9 & 11 \end{array}\right)$$

4. Maintenant que le pivot de la deuxième ligne vaut 1, nous pouvons recommencer le processus pour mettre des 0 dans la deuxième colonne : $L_3 \leftarrow L_3 - 5L_2$.

$$M_4 = \left(\begin{array}{cccc|c} 1 & 1 & 2 & 3 & 4 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 2 & 4 & 6 \end{array}\right)$$

5. Nous pouvons encore diviser la troisième ligne par deux ($L_3 \leftarrow \frac{1}{2}L_3$) et obtenir la matrice échelonnée en lignes :

$$M_5 = \left(\begin{array}{cccc|c} 1 & 1 & 2 & 3 & 4 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 2 & 3 \end{array}\right)$$

Le système S est donc équivalent au système

$$S_5 = \left\{\begin{array}{ccccccccc} x_1 & + & x_2 & + & 2x_3 & + & 3x_4 & = & 4 \\ & & x_2 & + & x_3 & + & x_4 & = & 1 \\ & & & & x_3 & + & 2x_4 & = & 3 \end{array}\right.$$

De la dernière ligne nous tirons que $x_3 = 3-2x_4$, que nous pouvons réinjecter dans les deux première lignes :

$$S_5 = \left\{\begin{array}{ccccccccc} x_1 & + & x_2 & + & 2(3-2x_4) & + & 3x_4 & = & 4 \\ & & x_2 & + & 3-2x_4 & + & x_4 & = & 1 \\ & & & & x_3 & & & = & 3-2x_4 \end{array}\right.$$

Soit

$$S_5 = \left\{\begin{array}{ccccccccc} x_1 & + & x_2 & & & - & x_4 & = & -2 \\ & & x_2 & & & - & x_4 & = & -2 \\ & & & & x_3 & & & = & 3-2x_4 \end{array}\right.$$

De la seconde ligne nous tirons maintenant que $x_2 = x_4 - 2$, que nous réinjectons dans la première ligne :

$$S_5 = \left\{\begin{array}{ccccccccc} x_1 & + & x_4-2 & & & - & x_4 & = & -2 \\ & & x_2 & & & & & = & x_4-2 \\ & & & & x_3 & & & = & 3-2x_4 \end{array}\right.$$

Soit

$$S_5 = \left\{\begin{array}{ccccc} x_1 & & & = & 0 \\ & x_2 & & = & x_4-2 \\ & & x_3 & = & 3-2x_4 \end{array}\right.$$

Les solutions sont donc :

$$\mathscr{S} = \left\{\left(0;x_4-2;3-2x_4;x_4\right) \mid x_4 \in \mathbb{K}\right\}$$

Puisque nous avons travaillé exclusivement avec des opérations élémentaires sur les lignes, nous sommes sûrs que nous avons trouvé toutes les solutions du système d'origine.

Remarque 9.5. Soit S un système linéaire et M sa matrice augmentée associée.

— Les lignes nulles de la réduction en lignes de M correspondent à des équations du type « $0 = 0$ ». Elles peuvent donc être supprimées du système sans changer les solutions.

— Si le pivot de la dernière ligne de la réduction en lignes de M (après avoir enlevé les lignes nulles) est dans la dernière colonne, alors cela correspond à une équation du type « $0 = b \neq 0$ ». Cette équation n'ayant pas de solution, le système d'origine n'a pas de solution.

9.4. Résolution d'un système linéaire

Nous avons vu sur l'exemple 9.5 la méthode permettant de résoudre un système linéaire. Vous savez déjà qu'un système linéaire peut avoir 0, 1 ou une infinité de solutions (comme dans l'exemple 9.5). Peut-on prédire à l'avance le nombre de solutions ? Peut-il y a avoir, disons 7 solutions ? Dans le cas où il y a une infinité de solutions, certaines des variables jouent un rôle particulier (comme x_4 sur l'exemple 9.5). Combien de variables jouent un tel rôle ? Nous allons répondre à cela dans cette section.

Propriété 9.4. Soit M une matrice et M_1 et M_2 deux matrices échelonnées en lignes équivalentes en lignes à M. Alors le nombre de pivots de M_1 est égal au nombre de pivots de M_2.

Démonstration. Admis. □

Remarque 9.6. La propriété 9.4 signifie que quelle que soit la manière dont on réduit M pour la rendre échelonnée en lignes, le nombre de pivots est toujours le même.

Définition 9.10 (Rang d'une matrice / d'un système linéaire). On appelle rang d'une matrice M, noté $\mathrm{rg}(M)$ le nombre de pivots obtenus après réduction de M en matrice échelonnée en lignes.
On appelle rang d'un système linéaire S, noté $\mathrm{rg}(S)$ le rang de sa matrice associée.

Remarque 9.7. Il est assez évident que le rang d'une matrice est inférieur à la fois au nombre de lignes et au nombre de colonnes de la matrice.

Définition 9.11 (Inconnues principales/secondaires). Soit S un système linéaire à p inconnues de rang r dont la matrice associée est échelonnée en lignes.
On appelle *inconnues principales* les r inconnues correspondant aux colonnes contenant les pivots.
On appelle *inconnues secondaires* les $p-r$ inconnues restantes.

Définition 9.12 (Système compatible/incompatible). Un système linéaire est dit *incompatible* s'il n'admet aucune solution. Il est dit *compatible* s'il admet au moins une solution.

Exemple 9.6. Soit $S_{\alpha,\beta}$ le système

$$\left\{\begin{array}{ccccccc} x_1 & - & 3x_2 & + & 5x_3 & = & -4 \\ 2x_1 & + & x_2 & - & 4x_3 & = & -1 \\ 3x_1 & - & 2x_2 & + & x_3 & = & \alpha \\ 4x_1 & - & 5x_2 & + & 6x_3 & = & \beta \end{array}\right.$$

La matrice augmentée associée est

$$\begin{pmatrix} 1 & -3 & 5 & -4 \\ 2 & 1 & -4 & -1 \\ 3 & -2 & 1 & \alpha \\ 4 & -5 & 6 & \beta \end{pmatrix}$$

Sa réduction en lignes nous donne

$$\begin{pmatrix} 1 & -3 & 5 & -4 \\ 0 & 1 & -2 & 1 \\ 0 & 0 & 0 & \alpha+5 \\ 0 & 0 & 0 & \beta+9 \end{pmatrix}$$

On en déduit que le système admet des solutions si et seulement si $\alpha = -5$ et $\beta = -9$. Dans ce cas, il est de rang 2, les inconnues principales sont x_1 et x_2, l'inconnue secondaire est x_3.
Les solutions sont $\mathscr{S} = \left\{ \left(x_3 - 1; 2x_3 + 1; x_3\right) \mid x_3 \in \mathbb{R} \right\}$

On voit sur l'exemple 9.6 que nous pouvons décrire toutes les solutions à l'aide des inconnues secondaires.

Propriété 9.5. Soit S un système linéaire homogène à p inconnues, de rang r. Alors l'ensemble des solutions est un espace vectoriel de dimension $p - r$.

Démonstration. Admis. □

Corollaire 9.1. Soit un système à $n \in \mathbb{N}^*$ équations et $p \in \mathbb{N}^*$ inconnues de matrice augmentée $[A, B]$ tel que $\mathrm{rg}(S) = r$.

1. Si $r = 0$, alors A est la matrice nulle. Si B n'est pas le vecteur nul, il n'y a pas de solutions, sinon tout vecteur de $\mathbb{R}^p$ est solution.
2. Si $r = n = p$, alors A se réduit en une matrice triangulaire supérieure. Quelque soit B, il y a une unique solution.
3. Si $r = n < p$, alors A se réduit en une matrice rectangulaire telle que les r premières colonnes forment une matrice triangulaire supérieure. Il y a alors une infinité de solutions dépendants de $p - r$ paramètres (les $p - r$ dernières inconnues).
4. Si $r = n < p$, alors A se réduit en une matrice rectangulaire telle que les r premières colonnes forment une matrice triangulaire supérieure. Il y a alors une infinité de solutions dépendants de $p - r$ paramètres (les $p - r$ dernières inconnues).
5. Si $r = p > n$, alors A se réduit en une matrice rectangulaire telle que les r premières lignes forment une matrice triangulaire supérieure et les lignes suivantes sont nulles.
6. Si $r < \min(p, n)$, alors A se réduit en une matrice telle que les r premières lignes et colonnes forment une matrice triangulaire supérieure et les lignes suivantes sont nulles.
   Dans les deux derniers cas, si l'un des $n - r$ seconds membres $b_{r+1}, \ldots, b_n$ est non nul, il n'y a pas de solution. Sinon, on supprime les $n - r$ dernières lignes et on retombe dans le cas 2 ou 3 (solution unique dans le cas 4 et une infinité dans le cas 5).

Démonstration. Intuitif. □

Remarque 9.8. Notez qu'un système homogène possède obligatoirement au moins une solution : le vecteur nul.

Exemple 9.7. Soit un système linéaire homogène à cinq inconnues et de rang trois. Supposons qu'après résolution, nous obtenions le système (après suppression des lignes « 0=0 ») :

$$\begin{cases} x_2 = 2x_1 + x_3 \\ x_4 = x_1 + 2x_3 \\ x_5 = 2x_1 - x_3 \end{cases}$$

Nous avons bien 3 inconnues principales : x_2, x_4 et x_5 et 2 inconnues secondaires : x_1 et x_3. De plus un vecteur solution s'écrit :

$$\overrightarrow{v} = (x_1;2x_1+x_3;x_3;x_1+2x_3;2x_1-x_3) \quad \text{avec } (x_1\,;x_3) \in \mathbb{R}^2$$

Ainsi $\overrightarrow{v}$ peut s'écrire

$$\overrightarrow{v} = x_1\,(1;2;0;1;2) + x_3\,(0;1;1;2;-1) \quad \text{avec } (x_1\,;x_3) \in \mathbb{R}^2$$

et les deux vecteurs $(1;2;0;1;2)$ et $(0;1;1;2;-1)$ suffisent à décrire toutes les solutions du système.

Propriété 9.6. Les solutions d'un système linéaire compatible s'obtiennent en additionnant une solution particulière à toutes les solutions du système homogène associé.

Démonstration. Admis. □

Exemple 9.8. Si on reprend le système de l'exemple 9.6 alors une solution particulière est $(-1\,;1\,;0)$ et les solutions du système homogène s'écrivent $\lambda(1\,;2\,;1)$ avec $\lambda \in \mathbb{R}$. Les solutions du système sont bien $(-1\,;1\,;0)+\lambda(1\,;2\,;1) = (\lambda-1\,;1+2\lambda\,;\lambda)$ avec $\lambda \in \mathbb{R}$.

Exemple 9.9. Résoudre le système suivant (on donnera le rang du système ainsi que toutes les solutions).

$$S = \begin{cases} 2x_1 + 3x_2 + 4x_3 = 1 \\ x_1 + 2x_2 + 3x_3 = 2 \\ 4x_1 + 2x_2 + x_3 = 3 \end{cases}$$

Appliquons le pivot de Gauss à la matrice augmentée.

$$\left(\begin{array}{ccc|c} 2 & 3 & 4 & 1 \\ 1 & 2 & 3 & 2 \\ 4 & 2 & 1 & 3 \end{array}\right)$$

$L1 \leftrightarrow L_2$, $L_2 \leftarrow L_2 - 2L_1$, $L_3 \leftarrow L_3 - 4L_1$

$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & 2 \\ 0 & -1 & -2 & -3 \\ 0 & -6 & -11 & -5 \end{array}\right)$$

$L2 \leftarrow -L_2$, $L_3 \leftarrow L_3 + 6L_2$

$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & 2 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 13 \end{array}\right)$$

Le rang du système est donc 3. Ainsi il y a 3 inconnues principales et 0 inconnues secondaires. Le système linéaire homogène associé n'est donc décrit par aucun vecteur, sa seule solution est donc le vecteur nul. Le système S possède donc au plus une solution. Or après réduction par le pivot de Gauss, nous n'avons obtenu aucune équation incompatible (de type $0 = b$ avec $b \neq 0$). Ainsi le système S possède une unique solution (c'est en fait une propriété générale que l'on verra plus tard [corollaire 12.1]).

La solution est donnée en remontant de la dernière ligne du système à la première : $x_3 = 13$, $x_2 = 3 - 2x_3 = 3 - 26 = -23$ et $x_1 = 2 - 2x_2 - 3x_3 = 2 + 46 - 39 = 9$. La solution du système est donc $\mathscr{S} = \{(9; -23; 13)\}$.

10. Espaces vectoriels

Dans tout ce chapitre, $\mathbb{K}$ désigne soit $\mathbb{R}$ soit $\mathbb{C}$ (en fait, dans la plupart des cas, $\mathbb{K}$ peut-être simplement un corps quelconque).

10.1. Espaces vectoriels

10.1.1. Espaces vectoriels

Définition 10.1 (Scalaire). Les éléments de $\mathbb{K}$ sont appelés *scalaires*.

Définition 10.2 (Espace vectoriel). On appelle $\mathbb{K}$-*espace vectoriel* (on note souvent $\mathbb{K}ev$) un ensemble E muni d'une addition (notée $\oplus$ et appelée *loi de composition interne*) de $E \times E$ dans E, et d'une multiplication par un scalaire (notée $\cdot$ et appelée *loi de composition externe*) de $\mathbb{K} \times E$ dans E, telles que :

1. $(E\,;\oplus)$ est un groupe commutatif (l'élément neutre est noté $\overrightarrow{0_E}$). Pour rappel :
   - E est stable par addition : pour tout $(u\,;v) \in E^2$, $u \oplus v \in E$
   - l'addition est associative : pour tout $(u\,;v\,;w) \in E^3$, $u \oplus (v \oplus w) = (u \oplus v) \oplus w$
   - l'addition est commutative : pour tout $(u\,;v) \in E^2$, $u \oplus v = v \oplus u$
   - l'addition admet un élément neutre : il existe un élément $\overrightarrow{0}_E \in E$ tel que pour tout $u \in E$, $u \oplus \overrightarrow{0}_E = u$.
   - tout élément de E admet un opposé : pour tout $u \in E$, il existe $u' \in E$ tel que $u \oplus u' = \overrightarrow{0}_E$. (On note $-u$ l'opposé de u).
2. E est stable par multiplication par un scalaire : pour tout $u \in E$, pour tout $\lambda \in \mathbb{K}$, $\lambda \cdot u \in E$.
3. la multiplication par un scalaire est associative : pour tout $(\lambda\,;\mu) \in \mathbb{K}^2$, pour tout $u \in E$, $\lambda \cdot (\mu \cdot u) = (\lambda\mu) \cdot u$.
4. la multiplication par un scalaire est distributive sur l'addition (à droite et à gauche) : pour tout $(\lambda\,;\mu) \in \mathbb{K}^2$, pour tout $(u\,;v) \in E^2$, $\lambda \cdot (u \oplus v) = (\lambda \cdot u) \oplus (\lambda \cdot v)$ et $(\lambda + \mu) \cdot u = (\lambda \cdot u) \oplus (\mu \cdot u)$.
5. pour tout $u \in E$, $1 \cdot u = u$

Lorsqu'il n'y a pas d'ambiguïté sur les lois utilisées, on peut noter simplement E l'espace vectoriel, sinon on le note $(E\,,\oplus\,,\cdot)$.

Remarque 10.1. La loi $\oplus$ est généralement notée simplement « + ».

Voici quelques contre-exemples d'espaces vectoriels :

1.
   - $E = [0\,;3]$ n'est pas un espace vectoriel car $1 \in E$ et $3 \in E$ et pourtant $1 + 3 = 4 \notin E$: l'ensemble n'est pas stable pour l'addition.
   - $E = [1\,;3]$ n'est pas un espace vectoriel car le seul élément neutre pour l'addition est 0 et $0 \notin E$.
   - $E = [-1\,;2]$ n'est pas un espace vectoriel car 2 n'a pas d'opposé.
2. $E = \left\{(x\,;y) \in \mathbb{R}^2 \;/\; y < 1\right\}$ n'est pas une espace vectoriel car $\overrightarrow{u} = \left(1\,;\frac{1}{2}\right) \in E$ mais $4 \cdot \overrightarrow{u} = (4\,;2) \notin E$.

Exemple 10.1.

— $\{0\}$ est un $\mathbb{R}ev$.

— $\mathbb{R}$ est un $\mathbb{R}ev$.

— $\mathbb{C}$ est un $\mathbb{C}ev$ et un $\mathbb{R}ev$.

— $\mathbb{R}^2$, $\mathbb{R}^3$ et plus généralement $\mathbb{K}^n$ sont des $\mathbb{R}ev$ et des $\mathbb{K}ev$.

— L'ensemble des fonctions réelles continues muni de l'addition de fonctions et du produit par un réel est un $\mathbb{R}ev$.

Définition 10.3 (Vecteur). Soit E un $\mathbb{K}$-espace vectoriel, les éléments de E sont appelés *vecteurs*.

Remarque 10.2. Puisque les éléments d'un espace vectoriel sont des vecteurs, ils pourront être notés avec des flèches. Cependant, parfois, pour alléger l'écriture nous pourrons omettre la flèche.

Propriété 10.1. Soit E un $\mathbb{K}ev$, alors pour tout $u \in E$:

— $0 \cdot u = \overrightarrow{0}_E$

— $(-1) \cdot u = -u$

Démonstration. — $0 \cdot u = 0 \cdot u + \overrightarrow{0}_E = 0 \cdot u + (u + (-u)) = (0 \cdot u + u) + (-u) = (0 \cdot u + 1 \cdot u) + (-u) = (1 \cdot u) + (-u) = u + (-u) = \overrightarrow{0}_E$

— $u + (-1) \cdot u = 1 \cdot u + (-1) \cdot u = (1 + (-1)) \cdot u = 0 \cdot u = \overrightarrow{0}_E$. Donc $(-1) \cdot u$ est bien l'opposé de u.

Définition 10.4 (Famille de vecteurs). Soit E un $\mathbb{K}$-ev. On appelle *famille de vecteurs* de E tout n-uplet $(u_1; \dots; u_n) \in E^n$.

Remarque 10.3.

— L'ordre des vecteurs dans la famille est important.

— Une famille peut comporter plusieurs fois le même vecteur.

Définition 10.5 (Combinaison linéaire d'une famille de vecteurs). Soit E un $\mathbb{K}$-espace vectoriel. On dit que $u \in E$ est *combinaison linéaire* de la famille $(u_1; \dots; u_n)$ s'il existe des scalaires $\lambda_1, \dots, \lambda_n$ de $\mathbb{K}$ tels que :

$$u = \sum_{i=1}^{n} \lambda_i u_i = \lambda_1 u_1 + \cdots + \lambda_n u_n$$

Cette écriture est également appelée *décomposition* du vecteur u suivant la famille de vecteurs $(u_1; \cdots; u_n)$

Exemple 10.2. Dans l'espace vectoriel des fonctions réelles, la fonction $f : x \mapsto \sin(x+1)$ est combinaison linéaire de la famille $(\sin;\cos)$.

Remarque 10.4. La décomposition sur une famille n'est en générale pas unique. Prenons par exemple $\overrightarrow{e}_1 = (1;2;0)$, $\overrightarrow{e}_2 = (1;0;1)$ et $\overrightarrow{e}_3 = (-1;2;-2)$. Alors le vecteur $\overrightarrow{u} = (1;4;-1)$ se décompose en $\overrightarrow{u} = \overrightarrow{e}_1 + \overrightarrow{e}_2 + \overrightarrow{e}_3 = 2\overrightarrow{e}_1 - \overrightarrow{e}_2 + 0\overrightarrow{e}_3$.

Propriété 10.2. Soit E un $\mathbb{K}$-ev et $v_1, \dots, v_p$, p vecteurs de E tous combinaisons linéaires d'une famille $(u_1;\dots;u_n)$. Alors toute combinaison linéaire u des vecteurs $v_1, \dots, v_p$ est une combinaison linéaire des vecteurs $u_1, \dots, u_n$.

Démonstration. Nous avons par hypothèse, pour tout $i \in [\![1;p]\!]$, $v_i = \sum_{k=1}^{n} \lambda_{i,k} u_k$ et $u = \sum_{j=1}^{p} \alpha_j v_j$. On en déduit :

$$u = \sum_{j=1}^{p} \alpha_j \left(\sum_{k=1}^{n} \lambda_{j,k} u_k \right) = \sum_{j=1}^{p} \sum_{k=1}^{n} \alpha_j \lambda_{j,k} u_k$$
$$u = \sum_{k=1}^{n} \sum_{j=1}^{p} \alpha_j \lambda_{j,k} u_k = \sum_{k=1}^{n} \left(\sum_{j=1}^{p} \alpha_j \lambda_{j,k} \right) u_k$$
$$u = \sum_{k=1}^{n} \beta_k u_k \qquad \text{en posant } \beta_k = \sum_{j=1}^{p} \alpha_j \lambda_{j,k}$$

10.1.2. Produit d'une famille finie d'espaces vectoriels

Théorème 10.1. Soit $n \geq 2$ un entier naturel, $E_1, \dots, E_n$ des $\mathbb{K}$-espaces vectoriels et $E = E_1 \times \cdots \times E_n$. On définit sur E :

1. L'addition par $(u_1;\dots;u_n) \oplus_E (v_1;\dots;v_n) = (u_1 \oplus_{E_1} v_1;\dots;u_n \oplus_{E_n} v_n)$
2. Le produit externe par $\lambda \cdot_E (u_1;\dots;u_n) = (\lambda \cdot_{E_1} u_1;\dots;\lambda \cdot_{E_n} u_n)$

Où $(u_1;\dots;u_n)$ et $(v_1;\dots;v_n)$ sont des éléments quelconques de E et λ un élément de $\mathbb{K}$.
Alors $(E;\oplus_E;\cdot_E)$ est un $\mathbb{K}$-espace vectoriel.

Démonstration. Il suffit d'écrire les choses. Nous avons déjà vu que le produit de groupes était un groupe pour la loi produit. Il ne reste plus qu'à vérifier les propriétés pour le produit externe. □

Corollaire 10.1. Pour tout $\mathbb{K}$-espace vectoriel E et tout entier $n \in \mathbb{N}^*$, E^n est un $\mathbb{K}$-espace vectoriel.

Démonstration. C'est le théorème précédent avec $E_1 = E_2 = \cdots = E_n$. □

10.1.3. Sous-espaces vectoriels

Définition 10.6 (Sous-espace vectoriel). Soit $(E, \oplus, \cdot)$ un $\mathbb{K}$-espace vectoriel. On dit que $(F, \oplus, \cdot)$ est un sous espace vectoriel (noté sev) de $(E, \oplus, \cdot)$ si $(F, \oplus, \cdot)$ est un $\mathbb{K}$-espace vectoriel et $F \subset E$.

Exemple 10.3. $\{\overrightarrow{0}_E\}$ et E sont des sous-espaces vectoriels de E.

A priori pour montrer qu'un ensemble est un sous-espace vectoriel d'un autre, il faut redémontrer la pléthore de propriétés de la définition 10.2. Heureusement la propriété suivante est là pour vous simplifier la vie :

Propriété 10.3 (Caractérisation d'un sous-espace vectoriel). Soit $(E, \oplus, \cdot)$ un $\mathbb{K}$-espace vectoriel. Alors $(F, \oplus, \cdot)$ est un sous espace vectoriel de $(E, \oplus, \cdot)$ si et seulement si :

1. $F \subset E$
2. $\overrightarrow{0}_E \in F$
3. F est stable par combinaison linéaire : pour tout $(\lambda;\mu) \in \mathbb{K}^2$ et pour tout $(u;v) \in F^2$, $\lambda u \oplus \mu v \in F$

Démonstration. $\boxed{\Rightarrow}$ Si F est un sev de E, alors c'est un $\mathbb{K}$ev, donc par définition d'un $\mathbb{K}$ev, il contient un élément neutre pour $\oplus$. Or il est unique. Puisque $F \subset E$, s'il y a un élément neutre dans F, c'est le même que celui de E. Donc $\overrightarrow{0}_E \in F$.

$\boxed{\Leftarrow}$

1. Si F est stable par combinaison linéaire, alors $u \oplus v \in F$.
2. L'addition étant associative dans E elle l'est aussi dans $F \subset E$.
3. L'addition étant commutative dans E elle l'est aussi dans $F \subset E$.
4. Par hypothèse, $\overrightarrow{0}_E \in F$. Puisque $\overrightarrow{0}_E$ est un élément neutre dans E, il l'est aussi dans F.
5. Soit $u \in F \subset E$. E étant un ev, u admet un opposé dans E. Reste à montrer que cet opposé est dans F : On a montré que dans E, $-u = (-1) \cdot u$. Or F est stable par combinaison linéaire, donc $(-1) \cdot u \in F$. Donc u admet bien un opposé dans F.
6. F est stable par multiplication par un scalaire, par hypothèse.
7. Le multiplication par un scalaire est associative dans E, donc elles l'est aussi dans F.
8. Le multiplication par un scalaire est distributive sur l'addition dans E, donc elles l'est aussi dans F.
9. $1 \cdot u = u$ dans E, donc aussi dans F.

Remarque 10.5.

— Le premier point est souvent écrit explicitement dans la définition de F.

— Le deuxième point est souvent très simple à montrer.

— Il est souvent plus aisé de décomposer le troisième point en deux étapes :

a) Montrer que F est stable par addition : prendre deux vecteurs u et v quelconques de F et montrer que $w = u \oplus v$ appartient bien à F.

b) Montrer que F est stable par multiplication par un scalaire : prendre un vecteur u quelconque de F et un scalaire λ quelconque de $\mathbb{K}$ et montrer que $\lambda \cdot u$ appartient F.

— Le troisième point est équivalent à $\forall \lambda \in \mathbb{K}$, $\forall (u;v) \in F^2$, $\lambda u \oplus v \in F$. En effet :

$\boxed{\Rightarrow}$ Il suffit de prendre $\mu = 1$.

$\boxed{\Leftarrow}$ Avec $v = \overrightarrow{0}_E$, nous avons que pour tout $\lambda \in \mathbb{K}$ et pour tout $u \in F$, $\lambda u \in F$. Soit alors $(\lambda;\mu) \in \mathbb{K}^2$ et $(u;v) \in F^2$. Si $\mu = 0$, alors $\lambda u + \mu v = \lambda u \in F$. Et si $\mu \neq 0$, $\lambda u + \mu v = \mu\left(\frac{\lambda}{\mu}u + v\right) = \mu w \in F$, avec $w = \frac{\lambda}{\mu}u + v \in F$ par hypothèse.

Exemple 10.4. Montrer que $D = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \ / \ x+y=0 \right\}$ est un sous-espace vectoriel de $\mathbb{R}^2$.

1. $D \subset \mathbb{R}^2$ par définition.
2. $0+0=0$ donc $\overrightarrow{0} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \in D$.
3. — Soit $\overrightarrow{u} \begin{pmatrix} x \\ y \end{pmatrix}$ et $\overrightarrow{v} \begin{pmatrix} x' \\ y' \end{pmatrix}$ deux éléments de D. Alors x, y, x', y' vérifient : $x+y=0$ et $x'+y'=0$.

   Posons $\overrightarrow{w} = \overrightarrow{u} \oplus \overrightarrow{v}$ et montrons que $\overrightarrow{w} \in D$. Les coordonnées de $\overrightarrow{w}$ sont $\overrightarrow{w} \begin{pmatrix} x+x' \\ y+y' \end{pmatrix}$. Calculons alors $(x+x')+(y+y') = (x+y)+(x'+y') = 0+0 = 0$. Donc $\overrightarrow{w} \in D$.

   — Soit $\overrightarrow{u} \begin{pmatrix} x \\ y \end{pmatrix} \in D$ et $\lambda \in \mathbb{R}$. Alors x et y vérifient $x+y=0$. Posons $\overrightarrow{w} = \lambda \overrightarrow{u}$ et montrons que $\overrightarrow{w} \in D$. Les coordonnées de $\overrightarrow{w}$ sont $\overrightarrow{w} \begin{pmatrix} \lambda x \\ \lambda y \end{pmatrix}$. Calculons alors $(\lambda x)+(\lambda y) = \lambda(x+y) = \lambda \times 0 = 0$. Donc $\overrightarrow{w} \in D$.

Nous pouvons en déduire que D est un sous-espace vectoriel de $\mathbb{R}^2$.

Propriété 10.4. L'intersection de deux sous-espaces vectoriels d'un $\mathbb{K}$-espace vectoriel E est un sous-espace vectoriel de E.

Démonstration. Très bon exercice d'entraînement ! Il n'y a rien de compliqué, c'est juste long à faire... □

Remarque 10.6. Attention. La réunion de deux (sous-)espaces vectoriels n'est PAS un espace vectoriel (sauf vraiment dans des cas très précis que vous pouvez trouver en exercices). En général le problème vient de la stabilité par l'addition qui n'est pas vérifiée.

Propriété 10.5. Toute intersection (finie ou infinie) de sous-espaces vectoriels d'un $\mathbb{K}$-espace vectoriel E est un sous-espace vectoriel de E.

Démonstration. C'est une généralisation de la propriété précédente et la démonstration est à votre portée. □

Propriété 10.6 (Sous-espace vectoriel engendré par une famille). Soit E un $\mathbb{K}$-ev et $(u_1\,;\dots;u_n)$ une famille de E. L'ensemble des combinaisons linéaires de cette famille est un sous-espace vectoriel de E, appelé *sous-espace vectoriel engendré* par la famille $(u_1\,;\dots;u_n)$. On note cet ensemble $\text{Vect}(u_1\,,\dots,u_n)$

Démonstration. Soit $F = \text{Vect}\,(u_1,\dots,u_n)$

— On a bien évidemment $F \subset E$.

— $\overrightarrow{0}_E = 0\cdot u_1 + \cdots + 0\cdot u_n \in F$

— Soit u et v deux éléments de F. Alors pour tout $(\lambda;\mu) \in \mathbb{K}^2$, $\lambda u + \mu v \in F$ en raison de la propriété 10.2

□

Propriété 10.7. Soit E un $\mathbb{K}$-ev et $(u_1,\dots,u_n)$ une famille de vecteurs de E. Alors $\text{Vect}\,(u_1,\dots,u_n)$ est le plus petit espace vectoriel (au sens de l'inclusion) contenant les vecteurs $u_1,\dots,u_n$. On a alors

$$\text{Vect}\,(u_1,\dots,u_n) = \bigcap_{\substack{F_i \text{ sev de } E\\ u_1,\dots,u_n \in F_i}} F_i$$

Démonstration. Dire que $\text{Vect}\,(u_1,\dots,u_n)$ est le plus petit espace vectoriel contenant les vecteurs $u_1,\dots,u_n$ signifie que si un sous-espace vectoriel H contient les vecteurs $u_1,\dots,u_n$, alors nécessairement $\text{Vect}\,(u_1,\dots,u_n) \subset H$. Soit donc H un sev de E contenant tous les vecteurs $u_1,\dots,u_n$.
Montrons que $\text{Vect}\,(u_1,\dots,u_n) \subset H$. Soit $u \in \text{Vect}\,(u_1,\dots,u_n)$. On a alors $u = \sum_{i=1}^{n} \lambda_i u_i$. Or les u_i appartiennent à H et H est stable par combinaison linéaire, donc $u \in H$ et $\text{Vect}\,(u_1,\dots,u_n) \subset H$.

L'égalité est simplement une autre manière de dire que c'est le plus petit espace vectoriel contenant les vecteurs. Démontrons la tout de même par double inclusion :

$\boxed{\subseteq}$ De la même manière que pour H, si F_i est un sev contenant les vecteurs u_i, alors nécessairement $\text{Vect}\,(u_1,\dots,u_n) \subset F_i$. Donc on a $\text{Vect}\,(u_1,\dots,u_n) \subset \bigcap_{\substack{F_i \text{ sev de } E\\ u_1,\dots,u_n \in F_i}} F_i$

$\boxed{\supseteq}$ $\text{Vect}\,(u_1,\dots,u_n)$ est un sev de E qui contient les u_i, donc $\bigcap_{\substack{F_i \text{ sev de } E\\ u_1,\dots,u_n \in F_i}} F_i \subset \text{Vect}\,(u_1,\dots,u_n)$ (de la même manière que pour tout ensemble A et B, $A\cap B \subset A$ et $A \cap B \subset B$).

Définition 10.7 (Sous-espace engendré par une partie). Soit E un espace vectoriel et $A \subset E$ une partie de E. L'intersection de tous les sous-espaces de E qui contiennent A est un sous-espace vectoriel de E.
On l'appelle le *sous-espace vectoriel engendré par* A. Il est noté $\text{Vect}(A)$.

Remarque 10.7. Si A est de cardinal fini, alors $\text{Vect}(A)$ est l'espace vectoriel engendré par la famille constituée de tous les vecteurs de A.

Propriété 10.8. Soit E un espace vectoriel et $A \subset E$ une partie de E. Alors $\text{Vect}(A)$ est le plus petit sous-espace vectoriel de E contenant A (au sens de la relation d'ordre « inclusion »).

Démonstration. Si F est un sous-espace vectoriel contenant A, alors par définition $\text{Vect}(A) \subset F$ puisque $\text{Vect}(A)$ est l'intersection de F et d'autres sous-espaces. □

Propriété 10.9. Soit E un espace vectoriel, A et B deux parties de E. Si $A \subset B$ alors $\text{Vect}(A) \subset \text{Vect}(B)$.

Démonstration. Nous avons $A \subset B \subset \text{Vect}(B)$. Donc $\text{Vect}(B)$ est un espace vectoriel contenant A, d'où $\text{Vect}(A) \subset \text{Vect}(B)$. □

10.1.4. Somme de sous-espaces vectoriels

Propriété 10.10 (Somme de sous-espaces vectoriels). Soit E un espace vectoriel, et F et G deux sous-espaces vectoriels de E. Alors $F+G = \{u+v \mathbin{/} u \in F \text{ et } v \in G\}$ est le plus petit sous-espace vectoriel de E contenant $F \cup G$.

Démonstration. Notons $S = F+G$.

— Vérifions d'abord que c'est bien un sous-espace vectoriel de E.

a) Soit $u \in S$, par définition de S, il existe $x \in F$ et $y \in G$ tel que $u = x+y$. Puisque F et G sont inclus dans E, on a x et y qui appartiennent à E. Par stabilité de E, $x+y$ appartient à E et donc $u \in E$. On a montré que $S \subset E$.

b) F et G étant des sev de E, ils contiennent $\overrightarrow{0}_E$. Or $\overrightarrow{0}_E = \overrightarrow{0}_E + \overrightarrow{0}_E$. Donc $\overrightarrow{0}_E \in S$

c) — Soit u et u' deux éléments de S. $u = x+y$ avec $x \in F$ et $y \in G$. $u' = x'+y'$ avec $x' \in F$ et $y' \in G$. Posons $v = u+u'$ et montrons que $v \in S$. On a $v = u+u' = (x+y)+(x'+y') = (x+x')+(y+y')$ (par associativité de l'addition dans E). Or par stabilité de F et G on a $(x+x') \in F$ et $(y+y') \in G$. D'où $u+u' \in S$. D'où la stabilité de S par l'addition.

— Soit $u = x+x' \in S$ avec $x \in F$ et $x' \in G$, et $\lambda \in \mathbb{K}$. Posons $v = \lambda \cdot u$ et montrons que $v \in S$. On a $v = \lambda \cdot (x+x') = \lambda \cdot x + \lambda \cdot x'$ (par distributivité de $\cdot$ sur $+$ dans E). Or par stabilité de F et G on a $\lambda \cdot x \in F$ et $\lambda \cdot x' \in G$. D'où $\lambda \cdot u \in S$. D'où la stabilité de S par la multiplication par un scalaire.

Donc S est bien un sous-espace vectoriel de E.

— Reste à montrer que c'est le plus petit. Dire que S est le plus petit espace vectoriel contenant $F \cup G$ signifie que si un sous-espace vectoriel H contient $F \cup G$, alors nécessairement $S \subset H$. Supposons donc que nous avons un sous espace H contenant $F \cup G$. Tout d'abord remarquons que si $F \cup G \subset H$ alors $F \subset H$ et $G \subset H$. Soit maintenant $u = x+x' \in S$, montrons que $u \in H$. Nous avons $x \in F \subset H$ et $x' \in G \subset H$. Par stabilité de H, nous avons $x+x' \in H$. Nous avons donc montré que $u \in H$, c'est-à-dire $S \subset H$.

S est donc bien le plus petit sous-espace vectoriel contenant $F \cup G$. □

Remarque 10.8. Nous avons donc $F+G = \text{Vect}(F \cup G)$.

Définition 10.8 (Somme directe). Soit E un $\mathbb{K}$-ev, et F et G deux sev de E. On dit que F et G sont en *somme directe* si tout élément de $F+G$ se décompose de **manière unique** en somme d'un élément de F et d'un élément de G. Dans ce cas on note la somme de F et G :

$$F+G = F \oplus G$$

Exemple 10.5. [Contre-exemple] Dans $\mathbb{R}^3$, soit $F = \left\{(x;y;z) \in \mathbb{R}^3 \;/\; x=y\right\}$ et $G = \left\{(x;y;z) \in \mathbb{R}^3 \;/\; y=z\right\}$. Le vecteur $(1;3;2)$ se décompose de deux manières différentes sur $F+G$: $(1;3;2) = \underbrace{(1;1;0)}_{\in F} + \underbrace{(0;2;2)}_{\in G} = \underbrace{(2;2;1)}_{\in F} + \underbrace{(-1;1;1)}_{\in G}$

Propriété 10.11. Soit E un $\mathbb{K}$-ev, et F et G deux sev de E. Alors F et G sont en somme directe si et seulement si $F \cap G = \left\{\overrightarrow{0}_E\right\}$.

Démonstration. $\boxed{\Leftarrow}$: Supposons que $F \cap G = \left\{\overrightarrow{0}_E\right\}$. Soit $u \in F+G$ et montrons que u se décompose de manière unique en $u = u_1 + u_2$ avec $u_1 \in F$ et $u_2 \in G$. Supposons que

$$u = \underbrace{u_1}_{\in F} + \underbrace{u_2}_{\in G} = \underbrace{v_1}_{\in F} + \underbrace{v_2}_{\in G}$$

On a alors $\underbrace{u_1 - v_1}_{\in F} = \underbrace{v_2 - u_2}_{\in G} = w \in F \cap G$. Or $F \cap G$ étant réduit au vecteur nul, on a $w = \overrightarrow{0}_E$, d'où $u_1 = v_1$ et $u_2 = v_2$ et l'écriture de u est unique.

$\boxed{\Rightarrow}$: Supposons maintenant que F et G sont en somme directe et montrons que leur intersection est réduite au vecteur nul. Soit $\overrightarrow{u} \in F \cap G$. On a tout simplement

$$u = \underbrace{u}_{\in F} + \underbrace{\overrightarrow{0}_E}_{\in G} = \underbrace{\overrightarrow{0}_E}_{\in F} + \underbrace{u}_{\in G}$$

Par unicité de la décomposition, on a $u = \overrightarrow{0}_E$.

Définition 10.9 (Sous-espaces supplémentaires). Soit E un $\mathbb{K}$-ev, et F et G deux sev de E. On dit que F et G sont *supplémentaires dans E* si tout élément de E se décompose de **manière unique** en somme d'un élément de F et d'un élément de G.

Propriété 10.12. Soit E un $\mathbb{K}$-ev, et F et G deux sev de E. Alors F et G sont supplémentaires dans E si et seulement si F et G sont en somme directe et $F+G=E$. Dans ce cas on note :

$$F \oplus G = E$$

Démonstration. $\boxed{\Rightarrow}$ Soit $x \in E$, il se décompose de manière unique en somme d'un élement x_f de f et x_G de G : $x = x_F + x_G$. Donc $x \in F+G$ et $E \subset F+G$. Or $F+G \subset E$. Donc $F+G = E$. De plus cette décomposition est unique, donc tout élément de $F+G$ se décompose de manière unique sur F et G, donc F et G sont en somme directe.

$\boxed{\Leftarrow}$ Puisque $F+G=E$, pour tout $x \in E$, il existe $x_F \in F$ et $x_G \in G$ tels que $x = x_F + x_G$. Or F et G étant en somme directe, cette décomposition est unique. Finalement tout élément de E se décompose de manière unique en somme d'un élément de F et d'un élément de G.

Exemple 10.6. Dans $\mathbb{R}^3$, un plan passant par $(0;0;0)$ et une droite passant par $(0;0;0)$ non incluse dans le plan sont supplémentaires.

10.2. Familles de vecteurs

10.2.1. Familles génératrices

Définition 10.10 (Famille génératrice). Soit E un $\mathbb{K}$-ev et $(u_1,\dots,u_n)$ une famille de vecteurs de E. On dit que $(u_1,\dots,u_n)$ est une *famille génératrice* de E (ou que les vecteurs $u_1,\dots,u_n$ *engendrent* E) si tout vecteur u de E est combinaison linéaire des vecteurs $u_1,\dots,u_n$.

Autrement dit $E = \mathrm{Vect}\,(u_1,\dots,u_n)$.

Exemple 10.7. Les vecteurs $(1;0;0)$, $(0;1;0)$ et $(0;0;1)$ engendrent $\mathbb{R}^3$. Les vecteurs $(1;1;1)$, $(0;1;1)$ et $(1;1;0)$ engendrent aussi $\mathbb{R}^3$.

Propriété 10.13. Dans un espace vectoriel E, toute sur-famille d'une famille génératrice de E est génératrice de E.

Démonstration. C'est assez évident. Si $(u_1,\dots,u_n)$ est une famille génératrice d'un espace vectoriel E, alors tout vecteur de E s'écrit comme combinaison linéaire de vecteur u_i. On peut toujours rajouter des vecteurs à la famille, les vecteurs de E continuerons à être des combinaisons linéaires de la famille (il suffit de prendre 0 pour les coefficients correspondants aux vecteurs ajoutés). $\square$

Propriété 10.14. Soit E un $\mathbb{K}$-ev et $\mathscr{F} = (u_1,\dots,u_n)$ une famille génératrice. Alors $\mathscr{F}\setminus\{u_p\}$ est une famille génératrice si et seulement si $u_p \in \mathrm{Vect}\,(\mathscr{F}\setminus\{u_p\})$.

Démonstration. $\boxed{\Rightarrow}$ **:** Supposons que $\mathscr{F}\setminus\{u_p\}$ est une famille génératrice. Puisque $u_p \in E$, par définition d'une famille génératrice, u_p s'écrit comme combinaison linéaire d'éléments de $\mathscr{F}\setminus\{u_p\}$, d'où $u_p \in \mathrm{Vect}\,(\mathscr{F}\setminus\{u_p\})$.

$\boxed{\Leftarrow}$ **:** Puisque $(u_1,\dots,u_n)$ est une famille génératrice, tout vecteur de E est une combinaison linéaire des vecteurs de cette famille. Or par hypothèse, u_p est une combinaison linéaire des autres vecteurs. Donc on peut décomposer la première combinaison linéaire sur les autres vecteurs. Démontrons-le de manière plus rigoureuse :

Soit $u \in E$. On peut écrire $u = \sum_{i=1}^{n} \lambda_i u_i = \sum_{\substack{i=1\\ i\neq p}}^{n} \lambda_i u_i + \lambda_p u_p$. Or par hypothèse, $u_p = \sum_{\substack{k=1\\ k\neq p}}^{n} \alpha_k u_k$. Ainsi :

$$u = \sum_{\substack{i=1\\ i\neq p}}^{n} \lambda_i u_i + \lambda_p u_p = \sum_{\substack{i=1\\ i\neq p}}^{n} \lambda_i u_i + \lambda_p \sum_{\substack{k=1\\ k\neq p}}^{n} \alpha_k u_k = \sum_{\substack{i=1\\ i\neq p}}^{n} (\lambda_i + \lambda_p \alpha_i)\, u_i$$

Donc $\mathscr{F}\setminus\{u_p\}$ est bien une famille génératrice de E.

10.2.2. Familles liées, familles libres

Propriété 10.15 ((Et définition) Famille liée). Soit E un $\mathbb{K}$-ev et $(u_1,\dots,u_n)$ une famille de vecteurs de E. Les conditions suivantes sont équivalentes :

1. L'un des vecteurs u_i est combinaison linéaire des autres.
2. Il existe $\lambda_1,\dots,\lambda_n \in \mathbb{K}$ non tous nuls tels que :

$$\lambda_1 u_1 + \cdots + \lambda_n u_n = \overrightarrow{0}_E$$

Dans ce cas, on dit que la famille $(u_1,\dots,u_n)$ est *liée*

Démonstration. $\boxed{1 \Rightarrow 2}$ **:** Supposons que l'un des vecteurs u_{i_0} est combinaison linéaire des autres. Alors on peut écrire

$$u_{i_0} = \sum_{\substack{i=1\\ i\neq i_0}}^{n} \alpha_i u_i$$

Ainsi, $\alpha_1 u_1 + \cdots + \alpha_{i_0-1}u_{i_0-1} + (-1)u_{i_0} + \alpha_{i_0+1}u_{i_0+1} + \cdots + \alpha_n u_n = \overrightarrow{0}_E$.

$\boxed{2 \Rightarrow 1}$ **:** Puisque les λ_i ne sont pas tous nuls, il existe i_0 tel que $\lambda_{i_0} \neq 0$. On a alors

$$\lambda_{i_0} u_{i_0} = \sum_{\substack{i=1\\ i\neq i_0}}^{n} (-\lambda_i) u_i$$

ou encore

$$u_{i_0} = \sum_{\substack{i=1\\ i\neq i_0}}^{n} \left(-\frac{\lambda_i}{\lambda_{i_0}}\right) u_i$$

D'où l'équivalence. □

Définition 10.11 (Famille libre). Soit E un $\mathbb{K}$-ev et $(u_1,\dots,u_n)$ une famille. On dit que la famille $(u_1,\dots,u_n)$ est *libre* si elle n'est pas liée. On dit aussi que les vecteurs $u_1,\dots u_n$ sont *linéairement indépendants*.

De manière équivalente, une famille est libre :

- si et seulement si pour tous $\lambda_1,\dots,\lambda_n \in \mathbb{K}$ non tous nuls, on a $\lambda_1 u_1 + \cdots + \lambda_n u_n \neq \overrightarrow{0}_E$
- si et seulement si

$$\forall \lambda_1,\dots,\lambda_n \in \mathbb{K},\ \left(\lambda_1 u_1 + \cdots + \lambda_n u_n = \overrightarrow{0}_E \Rightarrow \lambda_1 = \cdots = \lambda_n = 0\right)$$

- si et seulement si aucun des vecteurs n'est combinaison linéaire des autres.

Remarque 10.9. Une famille qui contient le vecteur nul est forcément liée.

Propriété 10.16. Toute sous famille (ou famille extraite) d'une famille libre est libre.

Démonstration. C'est assez évident si l'on prend n'importe laquelle des formulations équivalentes de la définition.□

Propriété 10.17. Soit E un $\mathbb{K}$-ev, $\in E$ et $(u_1,\dots,u_n)$ une famille libre de E. Alors $(u_1,\dots,u_n,u)$ est une famille libre si et seulement si $u \notin \operatorname{Vect}(u_1,\dots,u_n)$.

Démonstration. $\boxed{\Rightarrow}$ **:** Si $u \in \operatorname{Vect}(u_1,\dots,u_n)$ alors la famille $(u_1,\dots,u_n,u)$ est liée. Donc par contraposée, si la famille $(u_1,\dots,u_n,u)$ est libre alors $u \notin \operatorname{Vect}(u_1,\dots,u_n)$.

$\boxed{\Leftarrow}$ **:** Toujours par contraposée, supposons que la famille $(u_1,\dots,u_n,u)$ est liée. Il existe $n+1$ éléments de $\mathbb{K}$, $\lambda_1,\dots,\lambda_n,\lambda$ non tous nuls tels que $\lambda_1 u_1+\dots+\lambda_n u_n+\lambda u=\overrightarrow{0}_E$. Si $\lambda=0$ alors nous avons une combinaison linéaire de la famille $(u_1,\dots,u_n)$ nulle dont les coefficients λ_i sont non tous nuls, ce qui contredit le fait que la famille est libre. Donc nécessairement λ est non nul et

$$u=-\sum_{i=1}^{n}\frac{\lambda_i}{\lambda}u_i$$

Donc $u\in\operatorname{Vect}(u_1,\dots,u_n)$.

10.2.3. Bases

Définition 10.12 (Base). Soit E un $\mathbb{K}$-ev et $(u_1,\dots,u_n)$ une famille de vecteurs de E. On dit que la famille $(u_1,\dots,u_n)$ est une *base* de E si la famille est à la fois libre et génératrice.

Exemple 10.8.

— Soit $n\in\mathbb{N}^*$, la famille $(e_i)_{1\le i\le n}$ où pour tout $i\in[\![1\,;n]\!]$, e_i est défini par

$$e_i=(0;\dots;1;0;\dots;0)\quad \text{où le 1 est en } n^{\mathrm{e}} \text{ position}$$

est une une base de $\mathbb{K}^n$ appelée *base canonique de* $\mathbb{K}^n$

— La famille $(X^k)_{0\le k\le n}$ est une base de $\mathbb{K}_n[X]$, appelée base canonique de $\mathbb{K}_n[X]$. Plus généralement, la famille $(X^n)_{n\in\mathbb{N}}$ est une base (canonique) de $\mathbb{K}[X]$.

— Si $a\in\mathbb{K}$ est fixé, alors les familles $\left((X-a)^k\right)_{0\le k\le n}$ et $((X-a)^n)_{n\in\mathbb{N}}$ sont des bases respectivement de $\mathbb{K}_n[X]$ et $\mathbb{K}[X]$.

Théorème 10.2. Soit E un $\mathbb{K}$-ev et $(u_1,\dots,u_n)$ une base de E. Alors tout vecteur u de E s'écrit de manière unique comme combinaison linéaire des vecteurs $u_1,\dots,u_n$. C'est-à-dire :

$$\forall u\in E,\ \exists!\,(\lambda_1,\lambda_2,\dots,\lambda_n)\in\mathbb{K}^n,\ u=\lambda_1u_1+\lambda_2u_2+\dots+\lambda_nu_n$$

Démonstration. Par définition d'une base, la famille est génératrice de E, donc il est clair que tout vecteur de E s'écrit comme combinaison linéaire de vecteurs de la famille.

Puisque nous avons utilisé le côté génératrice de la famille pour montrer l'existence de la décomposition, cela sera très certainement le côté libre de la famille qui va nous permettre de montrer l'unicité de la décomposition. Supposons qu'un vecteur u se décompose de deux manières et montrons qu'en fait se sont les mêmes. Supposons donc

$$u=\lambda_1u_1+\dots+\lambda_nu_n=\mu_1u_1+\dots+\mu_nu_n$$

Or $(\lambda_1-\mu_1)u_1+\cdots+(\lambda_n-\mu_n)u_n=\overrightarrow{0}_E$. La famille étant libre, cela signifie que chaque coefficient est nul, $\lambda_1=\mu_1$, ..., $\lambda_n=\mu_n$. D'où l'unicité. □

Définition 10.13 (Coordonnées). Soit E un $\mathbb{K}$-ev et $(u_1,\dots,u_n)$ une base de E. On appelle *coordonnées* d'un vecteur u de E dans la base $(u_1,\dots,u_n)$ l'unique $n-$uplet $(\lambda_1,\dots,\lambda_n)$ de $\mathbb{K}^n$ tel que :

$$u=\sum_{k=1}^{n}\lambda_k u_k$$

10.3. Espaces vectoriels de dimension finie

10.3.1. Dimension d'un espace vectoriel

Définition 10.14 (Espace vectoriel de dimension finie). Soit E un $\mathbb{K}$-ev. On dit que E est de *dimension finie* s'il existe une famille génératrice finie de E. Dans le cas contraire E est de *dimension infinie*

Théorème 10.3. Soit E un $\mathbb{K}$-ev de dimension finie.

1. Si E admet une famille libre $\left(\overrightarrow{l}_1,\dots,\overrightarrow{l}_p\right)$ de p vecteurs et une famille génératrice $(\overrightarrow{g}_1,\dots,\overrightarrow{g}_m)$ de m vecteurs, alors $p\leq m$ (toute famille libre est plus petite que toute famille génératrice).
2. Si de plus E admet une base $(\overrightarrow{e}_1,\dots,\overrightarrow{e}_n)$ de n vecteurs, alors $p\leq n\leq m$.
3. Enfin sous les conditions précédentes,
   — Si $p=n$, alors $\left(\overrightarrow{l}_1,\dots,\overrightarrow{l}_p\right)$ est une base de E.
   — Si $m=n$, alors $(\overrightarrow{g}_1,\dots,\overrightarrow{g}_m)$ est une base de E.

Démonstration. 1. Admis.

2. Cela découle du point précédent.
3. — Supposons que $p=n$ et montrons que la famille $\left(\overrightarrow{l}_1,\dots,\overrightarrow{l}_n\right)$ est génératrice. Soit $\overrightarrow{u}\in E$. D'après le point précédent, la famille $\left(\overrightarrow{l}_1,\dots,\overrightarrow{l}_n,\overrightarrow{u}\right)$ ne peut pas être libre, elle est donc liée. Il existe donc une combinaison linéaire à coefficients non tous nuls égale au vecteur nul. Par un raisonnement analogue que celui fait dans la démonstration de la propriété 10.17, le coefficient de $\overrightarrow{u}$ est non nul. On en déduit donc que $\overrightarrow{u}$ est une combinaison linéaire des $\overrightarrow{u}_1,\dots,\overrightarrow{u}_n$.
   — Supposons que $n=m$ et montrons que la famille $(\overrightarrow{g}_1,\dots,\overrightarrow{g}_n)$ est libre. Si cette famille n'était pas libre, l'un des vecteurs (supposons g_n pour simplifier) serait combinaison linéaire des autres et donc tout vecteur de E serait combinaison linéaire de $(\overrightarrow{g}_1,\dots,\overrightarrow{g}_{n-1})$. Cela contredirait donc le point précédent. Donc la famille est libre.

Théorème 10.4 (Fondamental). Soit E un $\mathbb{K}$-ev (pas forcément de dimension finie). Si $\left(\overrightarrow{l}_1, \ldots, \overrightarrow{l}_p\right)$ est une famille libre de E et $(\overrightarrow{g}_1, \ldots, \overrightarrow{g}_m)$ une famille génératrice de E, alors il existe une base $\mathscr{B}$ de E de la forme

$$\mathscr{B} = \left(\overrightarrow{l}_1, \ldots, \overrightarrow{l}_p, \overrightarrow{l}_{p+1}, \ldots, \overrightarrow{l}_n\right)$$

où les vecteurs $\overrightarrow{l}_{p+1}, \ldots, \overrightarrow{l}_n$ sont pris dans la famille $(\overrightarrow{g}_1, \ldots, \overrightarrow{g}_m)$.

Démonstration. Admis. On part de la famille libre et on construit par récurrence, à l'aide de la propriété 10.17, des familles libres, jusqu'à obtenir une famille génératrice. □

Corollaire 10.2. Tout espace vectoriel de dimension finie non réduit à $\left\{\overrightarrow{0}_E\right\}$ possède une base.

Démonstration. Par définition, un tel espace possède une famille génératrice $\mathscr{G}$. Soit $\overrightarrow{u} \neq 0_E \in \mathscr{G}$. La famille $(\overrightarrow{u})$ est libre. Par application du théorème 10.4, on peut la compléter en une base de E. □

Corollaire 10.3 (Théorème de la base incomplète). Soit E un $\mathbb{K}$-ev de dimension finie et $\left(\overrightarrow{l}_1, \ldots, \overrightarrow{l}_p\right)$ une famille libre de E. Alors on peut compléter cette famille en une base $\left(\overrightarrow{l}_1, \ldots, \overrightarrow{l}_p, \overrightarrow{l}_{p+1}, \ldots, \overrightarrow{l}_n\right)$ de E.

Démonstration. Par définition, E possède une famille génératrice finie $\mathscr{G}$. Donc en utilisant le théorème 10.4, on peut compléter la famille libre en une base. □

Corollaire 10.4. Soit E un $\mathbb{K}$-ev et $(\overrightarrow{g}_1, \ldots, \overrightarrow{g}_m)$ une famille génératrice de E. Alors on peut extraire de cette famille une base, c'est-à-dire, qu'on peut choisir des éléments parmi $(\overrightarrow{g}_1, \ldots, \overrightarrow{g}_m)$ de façon à former une base de E.

Démonstration. Ce n'est qu'un cas particulier du théorème fondamental 10.4. □

Théorème 10.5. Soit E un $\mathbb{K}$-ev de dimension finie non nulle. Alors toutes les bases de E ont même cardinal.

Démonstration. Puisque E est de dimension finie il possède au moins une base. Soit $\mathscr{B}_1$ et $\mathscr{B}_2$ deux bases de E. Comme $\mathscr{B}_1$ est libre et $\mathscr{B}_2$ génératrice, on a $\operatorname{Card}(\mathscr{B}_1) \leq \operatorname{Card}(\mathscr{B}_2)$. De la même manière en intervertissant les bases, on a $\operatorname{Card}(\mathscr{B}_2) \leq \operatorname{Card}(\mathscr{B}_1)$. D'où l'égalité. □

Ce théorème et le corollaire 10.2 nous permettent de justifier la définition suivante :

Définition 10.15 (Dimension d'un espace vectoriel).

— Si $E = \left\{\overrightarrow{0}_E\right\}$ on dit que E est de *dimension* 0 et on note $\dim(E) = 0$.

— Si E est un espace vectoriel de dimension finie non réduit à $\left\{\overrightarrow{0}_E\right\}$, on appelle *dimension* de E le cardinal d'une base de E et on le note $\dim(E)$.

Propriété 10.18. Soit E un $\mathbb{K}$-ev de dimension n et $\mathscr{F}$ une famille de p vecteurs de E.

1. Si $\mathscr{F}$ est libre alors $p \leq n$ et on a égalité si et seulement si $\mathscr{F}$ est une base de E. (Toute famille libre de n vecteurs de E est une base de E.)
2. Si $\mathscr{F}$ est génératrice alors $p \geq n$ et on a égalité si et seulement si $\mathscr{F}$ est une base de E. (Toute famille génératrice de n vecteurs de E est une base de E.)

Démonstration. À vous de jouer. □

Remarque 10.10. En résumé, si $\dim(E) = n$ et X est une famille de n vecteurs de E, alors on a

$$X \text{ libre } \Leftrightarrow X \text{ génératrice } \Leftrightarrow X \text{ base}$$

10.3.2. Dimension d'un sous-espace vectoriel

Propriété 10.19. Soit E un $\mathbb{K}$-ev de dimension finie et F un sev de E. Alors F est de dimension finie et $\dim(F) \leq \dim(E)$. De plus $\dim(F) = \dim(E)$ si et seulement si $F = E$.

Démonstration. Admis. □

Propriété 10.20 (Formule de Grassmann). Soit F et G deux sev d'un $\mathbb{K}$-ev tous les deux dimension finie. Alors $F+G$ et $F \cap G$ sont de dimension finie et :

$$\dim(F+G) = \dim(F) + \dim(G) - \dim(F \cap G)$$

Démonstration. Admis. On part d'une base de $F \cap G$ que l'on complète en une base de F et en une base de G. On montre alors que la famille contenant tous les vecteurs ainsi obtenus est une base de $F+G$. □

Définition 10.16 (Rang d'une famille de vecteurs). Soit E un $\mathbb{K}$-espace vectoriel (pas forcément de dimension finie) et $\mathscr{F}$ une famille finie de vecteurs de E. On appelle *rang de la famille* $\mathscr{F}$, noté $\operatorname{rg}(\mathscr{F})$, la dimension (finie) de $\operatorname{Vect}(\mathscr{F})$.

10.3.3. Sous-espace vectoriels supplémentaires en dimension finie

Propriété 10.21. Soit E un $\mathbb{K}$-ev de dimension finie, et F et G deux sev de E. Les trois propriétés suivantes sont équivalentes :

(i) $E = F \oplus G$

(ii) $\dim(E) = \dim(F) + \dim(G)$ et $F \cap G = \left\{\overrightarrow{0}_E\right\}$

(iii) $\dim(E) = \dim(F) + \dim(G)$ et $E = F + G$

Démonstration. On montre $(i) \Rightarrow (ii) \Rightarrow (iii) \Rightarrow (i)$

$(i) \Rightarrow (ii)$ **:** Puisque $E = F \oplus G$, on a $E = F + G$ et $F \cap G = \left\{\overrightarrow{0}_E\right\}$. Donc $\dim(E) = \dim(F+G) = \dim(F) + \dim(G) - \underbrace{\dim(F \cap G)}_{=0} = \dim(F) + \dim(G)$.

$(ii) \Rightarrow (iii)$ **:** Puisque $\dim(E) = \dim(F) + \dim(G)$ et $F \cap G = \left\{\overrightarrow{0}_E\right\}$, on a $\dim(F+G) = \dim(F) + \dim(G) - \underbrace{\dim(F \cap G)}_{=0} = \dim(F) + \dim(G) = \dim(E)$. Donc $F + G = E$.

$(iii) \Rightarrow (i)$ **:** Puisque $\dim(E) = \dim(F) + \dim(G)$ et $E = F + G$, on a $\dim(F \cap G) = \dim(F) + \dim(G) - \dim(F+G) = \dim(F) + \dim(G) - \dim(E) = 0$. Donc $F \cap G = \left\{\overrightarrow{0}_E\right\}$.

Propriété 10.22. Soit E un $\mathbb{K}$-ev de dimension finie. Alors tout sev F de E admet un sev supplémentaire G. De plus :

$$\dim(G) = \dim(E) - \dim(F)$$

Démonstration. On pose $n = \dim(E)$. Soit F un sev de E, et $(\overrightarrow{e}_1, \dots, \overrightarrow{e}_p)$ une base de F. On la complète en une base $(\overrightarrow{e}_1, \dots, \overrightarrow{e}_p, \overrightarrow{e}_{p+1}, \dots, \overrightarrow{e}_n)$ de E.
Soit $G = \mathrm{Vect}(\overrightarrow{e}_{p+1}, \dots, \overrightarrow{e}_n)$. On vérifie facilement grâce à la propriété 10.21 que G convient. En effet, on a $\dim(E) = \dim(F) + \dim(G)$, et de plus si $u \in E$, on a :

$$u = \underbrace{\lambda_1 e_1 + \cdots + \lambda_p e_p}_{\in F} + \underbrace{\lambda_{p+1} e_{p+1} + \cdots + \lambda_n e_n}_{\in G}$$

Donc $u \in F + G$ $\square$

11. Applications linéaires

Dans tout ce chapitre $\mathbb{K}$ désigne le corps $\mathbb{R}$ ou le corps $\mathbb{C}$. Sauf mention contraire, les $\mathbb{K}$-espaces vectoriels auront pour loi $+$ et $\cdot$. Les ensembles $(E\,;+\,;\cdot)$ et $(F\,;+\,;\cdot)$ désigneront des $\mathbb{K}$-espaces vectoriels. Afin d'alléger les notations, le produit externe sera omis.

11.1. Définitions et propriétés

Définition 11.1 (Application linéaire). On appelle *application linéaire* ou *morphisme de $\mathbb{K}$-espaces vectoriels* une application f de E dans F vérifiant :

1. $\forall (x;y) \in E^2,\ f(x+y) = f(x) + f(y)$
2. $\forall x \in E,\ \forall \lambda \in \mathbb{K},\ f(\lambda x) = \lambda f(x)$.

On note $\mathscr{L}(E\,,F)$ l'ensemble des applications linéaires de E dans F.

Définition 11.2 (Isomorphisme, Endomorphisme, Automorphisme). On appelle *isomorphisme* une application linéaire bijective.
On appelle *endomorphisme de E* une application linéaire de E dans E.
On appelle *automorphisme de E* un endomorphisme bijectif.
On note $\mathrm{Isom}(E\,,F)$, $\mathscr{L}(E)$ et $\mathrm{Aut}(E)$ l'ensemble, respectivement, des isomorphimes de E dans F, des endomorphismes de E et des automorphismes de E.

Propriété 11.1. Soit $f \in \mathscr{L}(E\,,F)$. On a :

1. $f(0_E) = 0_F$.
2. Pour tout $x \in E$, $f(-x) = -f(x)$.

Démonstration.
1. Nous avons pour tout $y \in E$, $f(y) = f(0_E + y) = f(0_E) + f(y)$. D'où le résultat.
2. Cela provient directement de la définition, avec $\lambda = -1$ et des propriétés vues au chapitre précédent : $f(-x) = f((-1)x) = (-1)f(x) = -f(x)$.

Propriété 11.2 (Caractérisation d'une application linéaire). Une application f de E dans F est linéaire si et seulement si

$$\forall x,y \in E,\ \forall \lambda \in \mathbb{K},\ f(x+\lambda y) = f(x) + \lambda f(y)$$

Démonstration. $\boxed{\Rightarrow}$ L'application linéaire conserve la somme donc $f(x+\lambda y) = f(x) + f(\lambda y)$. Elle conserve aussi le produit externe, donc $f(\lambda y) = \lambda f(y)$. Ainsi $f(x+\lambda y) = f(x) + \lambda f(y)$.

$\boxed{\Leftarrow}$ En prenant $\lambda = 1$, l'application conserve la somme. En prenant $x = 0_E$, nous avons $f(\lambda y) = f(0_E) + \lambda f(y)$. Mais d'après la propriété précédente, $f(0_E) = 0_F$, donc l'application conserve le produit externe.

11.2. Image, Noyau

Propriété 11.3 (Image d'un sous-espace). Soit $f \in \mathscr{L}(E,F)$. Si A est un sous-espace vectoriel de E, alors $f(A)$ est un sous-espace vectoriel de F.

Démonstration. Puisque A contient 0_E, l'ensemble $f(A)$ contient $f(0_E) = 0_F$. Ainsi $f(A)$ est non vide.
Soit $(x,y) \in f(A)^2$. Alors il existe $(a;b) \in A^2$, tel que $x = f(a)$ et $y = f(b)$. Ainsi, pour tout $\lambda \in \mathbb{K}$, $x + \lambda y = f(a) + \lambda f(b)$. Nous avons alors, par linéarité de f, $x + \lambda y = f(a) + f(\lambda b) = f(a + \lambda b)$. Mais comme A est un sev de E, il s'ensuit que $a + \lambda b \in A$ et par conséquent, $x + \lambda y \in f(A)$. □

Définition 11.3 (Image d'une application linéaire). Soit $f \in \mathscr{L}(E,F)$. L'ensemble $f(E)$ est un sous-espace vectoriel de F, appelé *image* de f et noté $\operatorname{Im} f$.

Remarque 11.1. Une application linéaire $f \in \mathscr{L}(E,F)$ est surjective si et seulement si $\operatorname{Im} f = F$.

Propriété 11.4 (Image réciproque d'un sous-espace). Soit $f \in \mathscr{L}(E,F)$. Si B est un sous-espace vectoriel de F, alors $f^{-1}(B)$ est un sous-espace vectoriel de E.

Démonstration. $f(0_E) = 0_F \in B$, donc $0_E \in f^{-1}(B)$. Soit $(x;y) \in \left(f^{-1}(B)\right)^2$ et $\lambda \in \mathbb{K}$. Par linéarité de f, on a $f(x + \lambda y) = f(x) + \lambda f(y) \in B$ (car $f(x)$ et $f(y)$ sont dans B). Ainsi, $x + \lambda y \in f^{-1}(B)$. □

Définition 11.4 (Noyau). Soit $f \in \mathscr{L}(E,F)$. On appelle *noyau de* f, noté $\operatorname{Ker} f$, le sous-espace vectoriel $f^{-1}(\{0_F\}) \subset E$.

Théorème 11.1 (Caractérisation de l'injectivité). Une application linéaire f de E dans F est injective si et seulement si son noyau est le sous-espace nul de E :

$$f \text{ injective } \Leftrightarrow \operatorname{Ker} f = \{0_E\}$$

Démonstration. $\boxed{\Rightarrow}$ Soit $x \in \operatorname{Ker} f$. Nous avons donc $f(0_E) = 0_F = f(x)$. Par injectivité de f, $x = 0_E$.

$\boxed{\Leftarrow}$ Soit $\operatorname{Ker} f = \{0_E\}$, et x et y tels que $f(x) = f(y)$. Par linéarité de f, nous avons $f(x - y) = f(x) - f(y) = 0_F$. Donc $x - y \in \operatorname{Ker} f$, c'est-à-dire $x - y = 0_E$, soit $x = y$. D'où l'injectivité de f.

Propriété 11.5. Soit $f \in \mathscr{L}(E, F)$ et $(u_i)_{i\in I}$ une famille génératrice de E. Alors $(f(u_i))_{i\in I}$ est une famille génératrice de $\operatorname{Im} f$.

Démonstration. Soit $y \in \operatorname{Im} f$. Il existe $x \in E$ tel que $f(x) = y$. Or il existe une famille $(\lambda_i)_{i\in I}$ de scalaire, de support fini I, telle que

$$x = \sum_{i\in I} \lambda_i u_i$$

D'où

$$y = \sum_{i\in I} \lambda_i f(u_i) \in \operatorname{Vect}\left((f(u_i))_{i\in I}\right)$$

De plus, tout élément de $\operatorname{Vect}\left((f(u_i))_{i\in I}\right)$ appartient à $\operatorname{Im} f$. Donc $\operatorname{Im} f = \operatorname{Vect}\left((f(u_i))_{i\in I}\right)$. □

Propriété 11.6. Soit $f \in \mathscr{L}(E, F)$.

1. Si $(u_i)_{i\in I}$ est une famille liée de E, alors $(f(u_i))_{i\in I}$ est une famille liée de F.
2. Si f est injective et $(u_i)_{i\in I}$ est une famille libre de E, alors $(f(u_i))_{i\in I}$ est libre dans F.

Démonstration.

1. Il existe une famille de scalaire $(\lambda_i)_{i\in I}$ de support fini non vide telle que

$$\sum_{i\in I} \lambda_i u_i = 0_E \quad \text{d'où} \quad \sum_{i\in I} \lambda_i f(u_i) = 0_F$$

D'où le résultat.

2. Soit $(u_i)_{i\in I}$ une famille libre d'éléments de E. Soit une famille de scalaire $(\lambda_i)_{i\in I}$ de support fini, telle que $\sum_{i\in I} \lambda_i f(u_i) = 0_F$. Cette égalité s'écrit $f\left(\sum_{i\in I} \lambda_i u_i\right) = 0_F = f(0_E)$. Ainsi, par injectivité de f, $\sum_{i\in I} \lambda_i u_i = 0_E$. La famille $(u_i)_{i\in I}$ étant libre, pour tout $i \in I$, $\lambda_i = 0$ et la famille $(f(u_i))_{i\in I}$ est libre.

Théorème 11.2 (Détermination d'une application linéaire). Soit $B = (b_i)_{i\in I}$ une base de E et $C = (c_i)_{i\in I}$ une famille de vecteurs de F.

1. Il existe une et une seule application linéaire f de E dans F telle que

$$\forall i \in I,\ f(b_i) = c_i$$

Cette application linéaire f vérifie :

2. f est injective si et seulement si C est une famille libre.
3. f est surjective si et seulement si C est une famille génératrice de F.
4. f est un isomorphisme si et seulement si C est une base de F.

Démonstration.

1. **Unicité :** Soit $x \in E$ et $(x_i)_{i\in I}$ la famille de ses coordonnées dans la base B. Par linéarité de f, on a $f(x) = f\left(\sum_{i\in I} x_i b_i\right) = \sum_{i\in I} x_i f(b_i) = \sum_{i\in I} x_i c_i$. L'image $f(x)$ est donc entièrement déterminée par les coordonnées de x, d'où l'unicité de f.

Existence Soit f l'application définie par $x \mapsto \sum_{x\in I} x_i c_i$ où $(x_i)_{i\in I}$ est la famille des coordonnées de x dans la base B.
Étant donné $\lambda \in \mathbb{K}$, on sait que λx a pour coordonnées $(\lambda x_i)_{i\in I}$ donc :

$$f(\lambda x) = \sum_{i\in I} \lambda x_i c_i = \lambda \left(\sum_{i\in I} x_i c_i \right) = \lambda f(x)$$

Étant donné x et y de coordonnées $(x_i)_{i\in I}$ et $(y_i)_{i\in I}$, celles de $x+y$ sont $(x_i+y_i)_{i\in I}$. On a donc $f(x+y) = \sum_{i\in I}(x_i+y_i)c_i = \sum_{i\in I} x_i c_i + \sum_{i\in I} y_i c_i = f(x)+f(y)$, ce qui montre la linéarité de f. Pour vérifier que $f(b_k) = c_k$ pour $k \in I$, il suffit de remarquer que les coordonnées de b_k sont $(\beta_i)_{i\in I}$ avec $\beta_k = 1$ et $\beta_i = 0$ pour tout $i \neq k$.

2. Supposons $(c_i)_{i\in I}$ libre. Pour $x = \sum_{i\in I} x_i b_i$, on a $f(x) = \sum_{i\in I} x_i c_i$. Ainsi, $f(x) = 0$ donne $\forall i \in I, x_i = 0$ et donc $x = 0_E$.
Il en résulte $\operatorname{Ker} f = \{0_E\}$: f est injective.
La réciproque a été vue au deuxième point de la propriété 11.6.
3. D'après la propriété 11.5, la fonction f est surjective si et seulement si C est une famille génératrice de F.
4. Cela découle simplement des deux points précédents.

Remarque 11.2. Le premier point signifie qu'une application linéaire est entièrement caractérisée par les images des vecteurs d'une base.

11.3. Structures

11.3.1. $\mathscr{L}(E;F)$

Théorème 11.3 (Espace vectoriel $\mathscr{L}(E;F)$). L'ensemble $\mathscr{L}(E;F)$ est un $\mathbb{K}$-espace vectoriel.

Démonstration. Montrons que c'est un sous-espace vectoriel de $\left(F^E;+;\cdot\right)$ qui est un $\mathbb{K}$-espace vectoriel de vecteur nul θ_{EF} : $\begin{cases} E & \longrightarrow & F \\ x & \longmapsto & 0_F \end{cases}$.
Soit f et g deux éléments de $\mathscr{L}(E;F)$ et $\lambda \in \mathbb{K}$. Pour tout $(x;y) \in E^2$ et tout $\alpha \in \mathbb{K}$,

$$\begin{aligned}
(f+\lambda g)(x+\alpha y) &= f(x+\alpha y) + (\lambda \cdot g)(x+\alpha y) \\
&= f(x+\alpha y) + \lambda g(x+\alpha y) \\
&= f(x) + \alpha f(y) + \lambda(g(x)+\alpha g(y)) \\
&= (f(x)+\lambda g(x)) + \alpha(f(y)+\lambda g(y)) \\
&= (f+\lambda g)(x) + \alpha((f+\lambda g)(y))
\end{aligned}$$

D'où $f+\lambda g$ est linéaire. □

Théorème 11.4 (Composition). Soit E, F et G des $\mathbb{K}$-espaces vectoriels, $f \in \mathscr{L}(E;F)$ et $g \in \mathscr{L}(F;G)$. Alors $g \circ f \in \mathscr{L}(E;G)$.

Démonstration. Nous avons bien évidemment que $g\circ f$ est une application de E dans G. Il ne reste donc qu'à démontrer la linéarité. Soit $\lambda\in\mathbb{K}$ et $(x;y)\in E^2$. Alors

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

D'où la linéarité. □

Propriété 11.7. Soit E, F et G des $\mathbb{K}$-espaces vectoriels, et $h\in\mathscr{L}(E;F)$ et $k\in\mathscr{L}(F;G)$. Alors les applications :

$$L_d : \left\{\begin{array}{ccc}\mathscr{L}(E;F) & \longrightarrow & \mathscr{L}(E;G)\\ f & \longmapsto & k\circ f\end{array}\right. \quad \text{et} \quad L_g : \left\{\begin{array}{ccc}\mathscr{L}(F;G) & \longrightarrow & \mathscr{L}(E;G)\\ g & \longmapsto & g\circ h\end{array}\right.$$

sont linéaires.

On dira que l'application $L : \left\{\begin{array}{ccc}\mathscr{L}(E;F)\times\mathscr{L}(F;G) & \longrightarrow & \mathscr{L}(E;G)\\ (f;g) & \longmapsto & g\circ f\end{array}\right.$ est bilinéaire (linéaire par rapport à chacune des deux variables).

Démonstration. Cela veut simplement dire que

$$\forall\alpha\in\mathbb{K};\ \forall(f_1;f_2)\in(\mathscr{L}(E;F))^2,\ k\circ(f_1+\alpha\cdot f_2)=k\circ f_1+\alpha\cdot(k\circ f_2)$$
$$\forall\alpha\in\mathbb{K};\ \forall(g_1;g_2)\in(\mathscr{L}(F;G))^2,\ (g_1+\alpha\cdot g_2)\circ h=g_1\circ h+\alpha\cdot(g_2\circ h)$$

Ce qui est évident. □

Propriété 11.8. Soit E, F et G des $\mathbb{K}$-espaces vectoriels, et $f\in\mathscr{L}(E;F)$ et $g\in\mathscr{L}(F;G)$. Alors

1. $\operatorname{Im} g\circ f\subset\operatorname{Im} g$ et $\operatorname{Ker} f\subset\operatorname{Ker} g\circ f$
2. $g\circ f=\theta_{EG}\Leftrightarrow\operatorname{Im} f\subset\operatorname{Ker} g$ (la fonction θ_{EG} étant la fonction nulle qui à tout vecteur de E associe le vecteur nul de G, 0_G).

Démonstration.
1. Puisque $f(E)\subset F$, on déduit que $\operatorname{Im} g\circ f=(g\circ f)(E)=g(f(E))\subset g(F)=\operatorname{Im} g$.
   Soit $x\in\operatorname{Ker} f$. Alors $f(x)=0_F$ et $g(f(x))=g(O_F)=O_G$, donc $x\in\operatorname{Ker} g\circ f$.
2. Soit $y in \operatorname{Im}(f)$. Il existe $x\in E$ tel que $y=f(x)$, d'où $g(y)=g(f(x))=0$.
   Réciproquement, soit $x\in E$, alors $f(x)\in\operatorname{Im}(f)\subset\operatorname{Ker}(g)$, donc $g\circ f(x)=0$.

11.3.2. $\mathscr{L}(E)$

Théorème 11.5 (Anneau). Le triplet $(\mathscr{L}(E);+;\circ)$ est un anneau.

Démonstration. $\mathscr{L}(E) = \mathscr{L}(E;E)$, donc $(\mathscr{L}(E);+;\cdot)$ est un $\mathbb{K}$-espace vectoriel, d'où $(\mathscr{L}(E);+)$ est un groupe abélien.
De plus, la loi $\circ$ est associative, admet Id_E pour élément neutre, et pour tous f, g et h dans $\mathscr{L}(E)$, on a $f\circ(g+h) = f\circ g + f\circ h$ et $(g+h)\circ f = g\circ f + h\circ f$. □

Remarque 11.3. [Attention] Cet anneau n'est pas commutatif. Par exemple, soit $p : (x;y) \mapsto (x;0)$ et $s : (x;y) \mapsto (y;x)$. Ces deux applications sont linéaires et pour tout $(x;y) \in \mathbb{R}^2$,

$$(p\circ s)(x,y) = p(y,x) = (y;0) \quad ; \quad (s\circ p)(x,y) = s(x,0) = (0;x)$$

Définition 11.5 (Itérés, nilpotent). Soit $f \in \mathscr{L}(E)$. On pose $f^0 = Id_E$, $f^1 = f$, et pour tout $n \in \mathbb{N}$, $f^{n+1} = f^n \circ f$.
Pour tout $n \in \mathbb{N}$, les endomorphismes f^n sont appelés *itérés* de f.
On dit que $f \in \mathscr{L}(E)$ est *nilpotent* lorsque l'un de ses itérés est l'endomorphisme nul sur E.

Remarque 11.4. Si f et g sont à valeurs dans un espace E qui n'est ni $\mathbb{R}$ ni $\mathbb{C}$, cela n'a pas de sens de parler de produit $f\times g$. Il ne faut donc pas confondre $f^2 = f\circ f$ avec $f\times f$ qui n'a pas de sens en général.

Exemple 11.1. [Calculs dans $\mathscr{L}(E)$] Soit f et g deux endomorphismes de E.

1. $(Id_E - 3f)^2 = Id_E^2 - 3f\circ Id_E - 3Id_E\circ f + 9f^2 = Id_E - 6f + 9f^2$
2. $(f+g)^2 = f^2 + f\circ g + g\circ f + g^2$
3. $(f-g)\circ(f+g) = f^2 - g^2 + f\circ g - g\circ f$
4. $f^2 - 5f + 6Id_E = (f - 3Id_E)\circ(f - 2Id_E) = (f - 2Id_E)\circ(f - 3Id_E)$

Remarque 11.5. Les identités remarquables ne sont pas les mêmes car nous avons vu que l'anneau n'est pas commutatif. Pour le dernier, l'idée est de factoriser le polynôme $X^2 - 5X + 6 = (X-2)(X-3)$.

11.3.3. Aut(E) et $GL(E)$

Théorème 11.6 (Isomorphisme réciproque). Soit f un isomorphisme de E dans F. Alors l'application f^{-1} de F dans E est linéaire (et est donc un isomorphisme de F dans E).

Démonstration. Soit $(y_1;y_2) \in F$ et $\lambda \in \mathbb{K}$. L'application f étant bijective, il existe des antécédents par f, uniques, x_1 et x_2 dans E de y_1 et y_2. Nous avons alors

$$y_1 + \lambda y_2 = f(x_1) + \lambda f(x_2) = f(x_1 + \lambda x_2)$$

Ainsi, $x_1 + \lambda x_2 = f^{-1}(y_1 + \lambda y_2)$, ce qui montre que

$$f^{-1}(y_1 + \lambda y_2) = f^{-1}(y_1) + \lambda f^{-1}(y_2)$$

D'où la linéarité de f^{-1}. □

Théorème 11.7 (Groupe linéaire de E). Le couple $(\mathrm{Aut}(E)\,;\circ)$ est le groupe des éléments inversibles de l'anneau $(\mathscr{L}(E)\,;+\,;\circ)$ appelé *groupe linéaire de* E et noté $GL(E)$.

Démonstration.
1. Montrons que les éléments inversibles de $(\mathscr{L}(E)\,;+\,;\circ)$ sont dans $(\mathrm{Aut}(E)\,;\circ)$. Si $f \in \mathscr{L}(E)$ est inversible alors il existe $g \in \mathscr{L}(E)$ tel que $f\circ g = g\circ f = Id_E$, donc nécessairement f est injective et surjective. C'est donc bien un automorphisme de E.
2. Montrons maintenant que les automorphismes de E sont inversibles dans $\mathscr{L}(E)$. Si f est un automorphisme de E, alors c'est un isomorphisme de E dans E, donc d'après le théorème 11.6 précédent f^{-1} est un isomorphisme de E dans E. Donc f admet un symétrique (pour la loi $\circ$) dans $\mathscr{L}(E)$, c'est donc bien un élément inversible de $\mathscr{L}(E)$.

Définition 11.6 (Espaces isomorphes). Deux espaces vectoriels E et F sont dit *isomorphes* si et seulement si il existe un isomorphisme de E dans F.

Remarque 11.6. D'après le théorème 11.6, s'il existe un isomorphisme f de E dans F, alors f^{-1} est un isomorphisme de F dans E. La définition d'espaces isomorphes ne dépend donc pas de l'ordre dans lequel on considère E et F.

Lemme 11.1. Soit $f \in \mathscr{L}(E\,;F)$ et H un sous-espace vectoriel de E et $g = f_{|H}$. Alors $\mathrm{Ker}\, g = H \cap \mathrm{Ker}\, f$.

Démonstration. Un vecteur x appartient à $\mathrm{Ker}\, g$ si et seulement si il appartient à H et si $g(x) = 0_E$, c'est à dire $x \in H$ et $f(x) = 0_E$. $\square$

Théorème 11.8 (Théorème noyau-image). Soit $f \in \mathscr{L}(E\,;F)$. Alors $\mathrm{Im}\, f$ est isomorphe à tout supplémentaire de $\mathrm{Ker}\, f$.

Démonstration.
1. Soit H un supplémentaire de $\mathrm{Ker}\, f$ et $g = f_{|H}$. Le noyau de g est $H \cap \mathrm{Ker}\, f = \{0_E\}$. Il s'ensuit que g est injective. Or g est surjective sur son image, on veut donc montrer que $\mathrm{Im}\, g = \mathrm{Im}\, f$.
2. Pour tout $y \in \mathrm{Im}\, f$, il existe $x \in E$ tel que $f(x) = y$. Puisque $H + \mathrm{Ker}\, f = E$, il existe $(h;k) \in H \times \mathrm{Ker}\, f$ tel que $x = h + k$. Il s'ensuit que $y = f(h) + f(k) = f(h) = g(h)$ et $\mathrm{Im}\, f \subset \mathrm{Im}\, g$.
   Comme, bien entendu, $\mathrm{Im}\, g \subset \mathrm{Im}\, f$, il vient $\mathrm{Im}\, g = \mathrm{Im}\, f$.
3. L'application $\overline{g} : H \to \mathrm{Im}\, f$ induite par g, est donc un isomorphisme de H vers $\mathrm{Im}\, f$.

11.4. Applications linéaires en dimension finie

11.4.1. Dimensions

Propriété 11.9. Des $\mathbb{K}$-espaces vectoriels de dimension finies sont isomorphes si et seulement si ils ont la même dimension.

Démonstration. $\boxed{\Leftarrow}$ Soit E et F sont deux espaces vectoriels de dimension finie n, et $B=(b_i)_{1\leq i\leq n}$ et $C=(c_i)_{1\leq i\leq n}$ des bases de E et F. L'application linéaire φ telle que $\forall i,\ \varphi(b_i)=c_i$ est bijective d'après le théorème 11.2.

$\boxed{\Rightarrow}$ S'il existe un isomorphisme $\varphi\in\mathscr{L}(E;F)$, alors l'image $\varphi(B)=(\varphi(b_i))_{1\leq i\leq n}$ est une base de F (toujours le théorème 11.2), et on a ainsi $\dim(F)=n=\dim(E)$.

Théorème 11.9 (Dimension de $\mathscr{L}(E;F)$). Si E et F sont deux $\mathbb{K}$ espaces vectoriels de dimensions finies, l'espace vectoriel $\mathscr{L}(E;F)$ est de dimension finie et

$$\dim(\mathscr{L}(E;F))=\dim(E)\times\dim(F)$$

Démonstration. Soit $(e_j)_{1\leq j\leq p}$ et $(f_i)_{1\leq i\leq n}$ des bases de E et F respectivement.
On considère $\varphi_{i,j}\in\mathscr{L}(E;F)$ caractérisée par $\forall k\in[\![1;p]\!],\ \varphi_{i,j}(e_k)=\delta_{j,k}f_i$, où $\delta_{i,j}=0$ si $i\neq j$ et $\delta_{i,i}=1$ ($\delta_{i,j}$ est le symbole de Kronecker). Vérifions que $(\varphi_{i,j})_{(i;j)\in[\![1;n]\!]\times[\![1;p]\!]}$ est une base de $\mathscr{L}(E;F)$.

Soit $u\in\mathscr{L}(E;F)$. On sait que u est caractérisée par la donnée des $u(e_j)$ pour $1\leq j\leq p$. Posons donc, pour tout $j\in[\![1;p]\!]$, $u(e_j)=\sum_{i=1}^{n}a_{i,j}f_i$. Ainsi, u est caractérisée par la famille de scalaires $(a_{i,j})_{\substack{1\leq i\leq n\\1\leq j\leq p}}$. Posons alors $v=\sum_{i=1}^{n}\sum_{j=1}^{p}a_{i,j}\varphi_{i,j}$.

Pour tout $k\in[\![1;p]\!]$, on a $v(e_k)=\sum_{i=1}^{n}\sum_{j=1}^{p}a_{i,j}\varphi_{i,j}(e_k)$ et, comme $\varphi_{i,j}(e_k)=\delta_{j,k}f_i$, nous avons $\varphi_{i,k}(e_k)=f_i$ et les autres termes sont nuls. Donc il reste $v(e_k)=\sum_{i=1}^{n}a_{i,k}\varphi_{i,k}(e_k)=\sum_{i=1}^{n}a_{i,k}f_i=u(e_k)$. Il en résulte que $u=v$, donc la famille $(\varphi_{i,j})_{(i;j)\in[\![1;n]\!]\times[\![1;p]\!]}$ est génératrice de $\mathscr{L}(E;F)$.

Si $(a_{i,j})_{\substack{1\leq i\leq n\\1\leq j\leq p}}$ est une famille de scalaires telle que $\sum_{i=1}^{n}\sum_{j=1}^{p}a_{i,j}\varphi_{i,j}=0_{\mathscr{L}(E;F)}$, le calcul précédent montre que $\forall k\in[\![1;p]\!],\ \sum_{i=1}^{n}a_{i,k}f_i=0_F$.

Or $(f_i)_{1\leq i\leq n}$ est une base, donc $\forall k\in[\![1;p]\!],\ \forall i\in[\![1;n]\!],\ a_{i,k}=0$. Il en résulte que la famille $(\varphi_{i,j})_{(i;j)\in[\![1;n]\!]\times[\![1;p]\!]}$ est libre et finalement c'est une base de $\mathscr{L}(E;F)$.

Par conséquent $\dim(\mathscr{L}(E;F))=\mathrm{Card}(\varphi_{i,j})_{(i;j)\in[\![1;n]\!]\times[\![1;p]\!]}=np$. □

11.4.2. Rang

Définition 11.7 (Rang). Soit $f\in\mathscr{L}(E;F)$. Si $\mathrm{Im}\,f$ est de dimension finie, sa dimension est appelée *rang de f*. On le note $\mathrm{rg}(f)$.

Théorème 11.10 (Théorème du rang). Soit $f\in\mathscr{L}(E;F)$. Si E est de dimension finie, alors $\mathrm{Im}\,f$ est de dimension finie, et on a :

$$\dim(\mathrm{Im}\,f)+\dim(\mathrm{Ker}\,f)=\dim(E)$$

Ou sous forme équivalente :

$$\mathrm{rg}(f)=\dim(E)-\dim(\mathrm{Ker}\,f)$$

Démonstration. Soit H un supplémentaire de $\operatorname{Ker} f$ dans E. D'après le théorème noyau-image 11.8, $\operatorname{Im} f$ est isomorphe à H. On en déduit que $\dim(\operatorname{Im} f) = \dim(H)$, donc que

$$\dim(\operatorname{Im} f) = \dim(E) - \dim(\operatorname{Ker} f)$$

D'où le résultat. □

Corollaire 11.1. Soit E et F deux $\mathbb{K}$-espaces vectoriels et $f \in \mathscr{L}(E;F)$.

1. Si F est de dimension finie, alors f est de rang fini et $\operatorname{rg}(f) \leq \dim(F)$, avec égalité si et seulement si f est surjective.
2. Si E est de dimension finie, alors f est de rang fini et $\operatorname{rg}(f) \leq \dim(E)$, avec égalité si et seulement si f est injective.

Démonstration.
1. Comme $\operatorname{Im} f \subset F$, alors $\operatorname{Im} f$ est de dimension finie et $\operatorname{rg}(f) = \dim(\operatorname{Im} f) \leq \dim(F)$, avec égalité si et seulement si $\operatorname{Im} f = F$, c'est-à-dire si et seulement si f est surjective.
2. Cela découle directement du théorème du rang 11.10. Avec égalité si et seulement $\dim(\operatorname{Ker} f) = 0$.

Corollaire 11.2. Soit E et F deux $\mathbb{K}$-espaces vectoriels de dimensions finies et $f \in \mathscr{L}(E;F)$.

1. Si f est injective alors $\dim(E) \leq \dim(F)$, avec égalité si et seulement si f est bijective.
2. Si f est surjective alors $\dim(F) \leq \dim(E)$, avec égalité si et seulement si f est bijective.

Démonstration. Cela découle directement du corollaire précédent (11.1).

1. Si f est injective alors $\dim(\operatorname{Ker} f) = 0$ et $\dim(E) = \operatorname{rg}(f) \leq \dim(F)$. Il y a égalité si et seulement si $\operatorname{rg}(f) = \dim(F)$, c'est-à-dire si et seulement si f est surjective.
2. Si f est surjective alors $\dim(F) = \operatorname{rg}(f) \leq \dim(E)$. Il y a égalité si et seulement si $\operatorname{rg}(f) = \dim(E)$, c'est-à-dire si et seulement f est injective.

Corollaire 11.3 (Application linéaire entre espaces vectoriels de mêmes dimensions finies). Soit E et F deux $\mathbb{K}$-espaces vectoriels de dimensions finies et $f \in \mathscr{L}(E;F)$. Si $\dim(E) = \dim(F)$, alors :

$$f \text{ bijective } \Leftrightarrow f \text{ injective } \Leftrightarrow f \text{ surjective}$$

Démonstration. Cela découle des corollaires précédents (une démonstration directe avec le théorème du rang a été vue en T.D.). □

12. Matrices

Le possible est une matrice formidable.

(Victor Hugo)

L'ensemble $\mathbb{K}$ désigne soit $\mathbb{R}$ soit $\mathbb{C}$, n et p sont des entiers naturels non nuls.

12.1. Définitions

12.1.1. Types de matrices

Définition 12.1 (Matrice). Soit $n, p \in \mathbb{N}^*$. On appelle *matrice* à n lignes et p colonnes à coefficients dans $\mathbb{K}$ un tableau de nombres que l'on écrit sous la forme :

$$M = \begin{pmatrix} a_{1,1} & \dots & a_{1,p} \\ \vdots & \ddots & \vdots \\ a_{n,1} & \dots & a_{n,p} \end{pmatrix}$$

On dit que M est de *taille* $n \times p$.

Afin de simplifier la notation on note parfois $M = [a_{i,j}]_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}}$.

On note $\mathscr{M}_{n,p}(\mathbb{K})$ l'ensemble des matrices à n lignes et p colonnes à coefficients dans $\mathbb{K}$.

Définition 12.2 (Vecteur ligne / Vecteur colonne). Une matrice de taille $n \times 1$ est appelée *vecteur colonne*.

Une matrice de taille $1 \times p$ est appelée *vecteur ligne*.

Remarque 12.1. Une matrice n'ayant qu'une colonne est assimilée à un vecteur que l'on représente en colonne et une matrice ayant une seule ligne est assimilée à un vecteur que l'on représente en lignes.

Définition 12.3 (Matrice nulle). On appelle *matrice nulle* de taille $n \times p$ la matrice de taille $n \times p$ dont tous les coefficients sont nuls. On la note $0_{n,p}$.

Définition 12.4 (Matrice carrée). On appelle *matrice carrée* de taille n une matrice à n lignes et n colonnes.

On note $\mathscr{M}_n(\mathbb{K})$ l'ensemble des matrices carrées de taille n à coefficients dans $\mathbb{K}$.

Définition 12.5 (Matrice diagonale). On appelle *matrice diagonale* une matrice carrée dont tous les éléments en dehors de la diagonale sont nuls.

$$D = \operatorname{diag}(a_{1,1}, \dots, a_{n,n}) = \begin{pmatrix} a_{1,1} & 0 & \dots & 0 \\ 0 & a_{2,2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & a_{n,n} \end{pmatrix}$$

Exemple 12.1. La matrice

$$M = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & -3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 75 & 0 \\ 0 & 0 & 0 & 0 & \frac{\sqrt{2}}{\pi} \end{pmatrix}$$

est une matrice diagonale de taille 5.

Définition 12.6 (Matrice identité). On appelle *matrice identité* de taille n la matrice diagonale de taille n dont tous les coefficients valent 1 :

$$I_n = \operatorname{diag}(1, \dots, 1) = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & 1 \end{pmatrix}$$

Définition 12.7 (Matrice triangulaire). On appelle *matrice triangulaire supérieure* une matrice carrée dont tous les éléments en dessous de la diagonales sont nuls.

$$\begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ 0 & a_{22} & \ddots & \vdots \\ \vdots & \ddots & \ddots & a_{n-1,n} \\ 0 & \dots & 0 & a_{n,n} \end{pmatrix}$$

On appelle *matrice triangulaire inférieure* une matrice carrée dont tous les éléments au dessus de la diagonales sont nuls.

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

On appelle *matrice triangulaire* une matrice qui est soit triangulaire inférieure soit triangulaire supérieure.

Exemple 12.2. La matrice

$$M_s = \begin{pmatrix} 1 & 5 & 0 & 2 & 1 \\ 0 & -3 & 4 & 7 & -1 \\ 0 & 0 & 0 & 3 & -15 \\ 0 & 0 & 0 & 75 & 2 \\ 0 & 0 & 0 & 0 & \frac{\sqrt{2}}{\pi} \end{pmatrix}$$

est triangulaire supérieure.

La matrice

$$M_i = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 \\ 3 & 4 & 35 & 0 & 0 \\ -1 & -2 & 1 & 75 & 0 \\ 47 & 0 & 3 & 5 & \frac{\sqrt{2}}{\pi} \end{pmatrix}$$

est triangulaire inférieure.

Les matrices M_s et M_i sont toutes les deux des matrices triangulaires.

12.1.2. Opérations

Définition 12.8 (Somme de matrices). Soit $A = [a_{i,j}]_{\substack{1\leq i\leq n\\ 1\leq j\leq p}}$ et $B = [b_{i,j}]_{\substack{1\leq i\leq n\\ 1\leq j\leq p}}$ deux matrices de tailles $n \times p$. On note $A + B = C = [c_{i,j}]_{\substack{1\leq i\leq n\\ 1\leq j\leq p}}$ la matrice de taille $n \times p$ définie par

$$\forall\ 1 \leq i \leq n,\ \forall\ 1 \leq j \leq n,\ c_{i,j} = a_{i,j} + b_{i,j}$$

Exemple 12.3.

$$\begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} + \begin{pmatrix} 2 & 8 & 1 \\ -1 & 2 & -3 \end{pmatrix} = \begin{pmatrix} 3 & 10 & 4 \\ 2 & 4 & -2 \end{pmatrix}$$

Remarque 12.2. Cela n'a strictement <u>AUCUN SENS</u> d'additionner deux matrices n'ayant pas la même taille.

Définition 12.9 (Multiplication par un scalaire). Soit $A = [a_{i,j}]_{\substack{1\leq i\leq n\\ 1\leq j\leq p}}$ une matrice de taille $n \times p$ et $\lambda \in \mathbb{K}$. On note $\lambda A = C = [c_{i,j}]_{\substack{1\leq i\leq n\\ 1\leq j\leq p}}$ la matrice de taille $n \times p$ définie par

$$\forall 1 \leq i \leq n,\ \forall 1 \leq j \leq n,\ c_{i,j} = \lambda a_{i,j}$$

Exemple 12.4.

$$3\begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 3 & 6 & 9 \\ 9 & 6 & 3 \end{pmatrix}$$

Définition 12.10 (Produit d'une matrice par un vecteur colonne). Soit $A = [a_{i,j}]_{\substack{1\leq i\leq n\\ 1\leq j\leq p}}$ une matrice de taille $n \times p$ et $B = [b_l]_{1\leq l\leq p}$ un vecteur colonne de taille p. On note $AB = C = [c_i]_{1\leq i\leq n}$ le vecteur colonne de taille n défini par

$$\forall\ 1 \leq i \leq n,\ c_i = \sum_{k=1}^{p} a_{i,k} b_k$$

Remarque 12.3. On remarque que l'élément à la ligne i du produit est tout simplement le produit scalaire de la i-ième ligne de A par le vecteur B. Autrement dit, le produit scalaire entre deux vecteurs est égal au produit (matriciel) d'un vecteur ligne par un vecteur colonne.

Exemple 12.5.

$$\begin{array}{cc} & \begin{pmatrix} 2 \\ -4 \\ 3 \end{pmatrix} \\ \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} & \begin{pmatrix} 2\times 1+(-4)\times 2+3\times 3 \\ 2\times 3+(-4)\times 2+3\times 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \end{pmatrix} \end{array}$$

Remarque 12.4.

— Remarquez que le nombre de lignes de B doit absolument être égal au nombre de colonnes de A.

— En fait, le vecteur produit est la combinaison linéaire des colonnes de A, les coefficients étant ceux de B. En effet, sur l'exemple précédent

$$\begin{pmatrix} 3 \\ 1 \end{pmatrix} = 2\begin{pmatrix} 1 \\ 3 \end{pmatrix} - 4\begin{pmatrix} 2 \\ 2 \end{pmatrix} + 3\begin{pmatrix} 3 \\ 1 \end{pmatrix}$$

— Soit $A = [a_{i,j}]_{\substack{1\leq i\leq n \\ 1\leq j\leq p}}$ une matrice de taille $n\times p$ et $X = (x_i)_{1\leq i\leq p}$ un vecteur colonne de taille p. Alors le produit AX donne un vecteur colonne dont la ligne i est égale à $a_{i1}x_1 + \dots a_{i,p}x_p$. C'est la raison pour laquelle on peut écrire un système sous la forme $AX = B$ où A est la matrice associée au système, X est le vecteur contenant les inconnues et B le vecteur du second membre.

Définition 12.11 (Produit de deux matrices). Soit $A = [a_{i,j}]_{\substack{1\leq i\leq n \\ 1\leq j\leq p}}$ une matrice de taille $n\times p$ et $B = [b_{i,j}]_{\substack{1\leq i\leq p \\ 1\leq j\leq m}}$ une matrice de taille $p\times m$. On note $AB = C = [c_{i,j}]_{\substack{1\leq i\leq n \\ 1\leq j\leq m}}$ la matrice de taille $n\times m$ définie par

$$\forall\ 1\leq i\leq n,\ \forall\ 1\leq j\leq m,\ c_{ij} = \sum_{k=1}^{p} a_{ik}b_{kj}$$

Exemple 12.6.

$$\begin{array}{cc} & \begin{pmatrix} 2 & 1 & 4 & 2 \\ -4 & 2 & 3 & 1 \\ 3 & 1 & -1 & 2 \end{pmatrix} \\ \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} & \begin{pmatrix} 3 & 8 & 7 & 10 \\ 1 & 8 & 17 & 10 \end{pmatrix} \end{array}$$

Remarque 12.5.

— Remarquez qu'encore une fois le nombre de lignes de B doit absolument être égal au nombre de colonnes de A.

— On remarque que la j-ième colonne du produit AB est tout simplement le produit de A par la j-ième colonne de B.

— De même, la i-ième ligne du produit AB est tout simplement le produit de la i-ième ligne de A par B.

— Le nombre de lignes de AB est égal au nombre de lignes de A et le nombre de colonnes de AB est égal au nombre de colonnes de B.

Remarque 12.6. [Attention !] Le produit de matrices n'a pas vraiment les mêmes propriétés que le produit de nombres.

1. Le produit de deux matrices n'est pas commutatif, c'est-à-dire que $AB \neq BA$. Tout d'abord si les matrices ne sont pas carrées les deux produits n'ont en général pas de sens en même temps. Si par contre les deux matrices sont carrées, les deux produits ont un sens, mais ne sont en général tout de même pas égaux.

Exemple 12.7.

$$\begin{array}{cc} & \begin{pmatrix}1 & 3\\ 2 & -1\end{pmatrix} \\ \begin{pmatrix}3 & 5\\ -1 & 4\end{pmatrix} & \begin{pmatrix}13 & 4\\ 7 & -7\end{pmatrix}\end{array}$$

et

$$\begin{array}{cc} & \begin{pmatrix}3 & 5\\ -1 & 4\end{pmatrix} \\ \begin{pmatrix}1 & 3\\ 2 & -1\end{pmatrix} & \begin{pmatrix}0 & 17\\ 7 & 6\end{pmatrix}\end{array}$$

De même,

$$\begin{array}{cc} & \begin{pmatrix}1 & 2 & 3\\ 2 & 3 & 1\end{pmatrix} \\ \begin{pmatrix}-1 & 4\\ 2 & -3\\ 1 & -1\end{pmatrix} & \begin{pmatrix}7 & 10 & 1\\ -4 & -5 & 3\\ -1 & -1 & 2\end{pmatrix}\end{array}$$

et

$$\begin{array}{cc} & \begin{pmatrix}-1 & 4\\ 2 & -3\\ 1 & -1\end{pmatrix} \\ \begin{pmatrix}1 & 2 & 3\\ 2 & 3 & 1\end{pmatrix} & \begin{pmatrix}6 & -5\\ 5 & -2\end{pmatrix}\end{array}$$

2. Nous pouvons avoir un produit nul ($AB = 0_{n,p}$) sans pour autant que A ou B soit nulle.

Exemple 12.8. Soit $A = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$ et $B = \begin{pmatrix} 1 & 2 \\ -1 & -2 \end{pmatrix}$. Alors

$$\begin{array}{cc} & \begin{pmatrix} 1 & 2 \\ -1 & -2 \end{pmatrix} \\ \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} & \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \end{array}$$

Définition 12.12 (Transposée). Soit $A = [a_{i,j}]_{\substack{1\leq i\leq n \\ 1\leq j\leq p}}$ une matrice de taille $n \times p$. On appelle *transposée* de A la matrice notée ${}^tA$ de taille $p \times n$ dont l'élément à la i-ième ligne et j-ième colonne est égal à l'élément à la j-ième ligne et i-ième colonne de A.

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} & \cdots & \cdots & \cdots & a_{1,p-1} & a_{1,p} \\ a_{2,1} & a_{2,2} & a_{2,3} & \cdots & \cdots & \cdots & a_{2,p-1} & a_{2,p} \\ a_{3,1} & a_{3,2} & a_{3,3} & \cdots & \cdots & \cdots & a_{3,p-1} & a_{3,p} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n-1,1} & a_{n-1,2} & a_{n-1,3} & \cdots & \cdots & \cdots & a_{n-1,p-1} & a_{n-1,p} \\ a_{n,1} & a_{n,2} & a_{n,3} & \cdots & \cdots & \cdots & a_{n,p-1} & a_{n,p} \end{pmatrix}$$

$${}^tA = \begin{pmatrix} a_{1,1} & a_{2,1} & a_{3,1} & \cdots & a_{n-1,1} & a_{n,1} \\ a_{1,2} & a_{2,2} & a_{3,2} & \cdots & a_{n-1,2} & a_{n,2} \\ a_{1,3} & a_{2,3} & a_{3,3} & \cdots & a_{n-1,3} & a_{n,3} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{1,p-1} & a_{2,p-1} & a_{3,p-1} & \cdots & a_{n-1,p-1} & a_{n,p-1} \\ a_{1,p} & a_{2,p} & a_{3,p} & \cdots & a_{n-1,p} & a_{n,p} \end{pmatrix}$$

Exemple 12.9. La transposée de $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$ est ${}^tA = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}$

Définition 12.13 (Matrices symétriques / anti-symétriques). Soit A une matrice carrée.
On dit que A est *symétrique* si et seulement si elle est égale à sa transposée ($A = {}^tA$).
On dit que A est *anti-symétrique* si et seulement si elle est égale à l'opposé de sa transposée ($A = -{}^tA$).

Remarque 12.7.
- Une matrice $A = [a_{i,j}]$ est symétrique si et seulement si, pour tout i et pour tout j, $a_{i,j} = a_{j,i}$.
- Une matrice $A = [a_{i,j}]$ est anti-symétrique si et seulement si, pour tout i et pour tout j, $a_{i,j} = -a_{j,i}$. En particulier, puisque $a_{i,i} = -a_{i,i}$, la diagonale d'une matrice anti-symétrique est forcément nulle.

Exemple 12.10. La matrice $A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 5 & 6 & 7 \\ 3 & 6 & 8 & 9 \\ 4 & 7 & 9 & 10 \end{pmatrix}$ est symétrique.
La matrice $B = \begin{pmatrix} 0 & 1 & -2 \\ -1 & 0 & 3 \\ 2 & -3 & 0 \end{pmatrix}$ est anti-symétrique.

Définition 12.14 (Trace). Soit A une matrice carrée. On appelle *trace* de A, le nombre noté $\mathrm{tr}(A)$ égal à la somme des éléments diagonaux de A :

$$\mathrm{tr}(A) = a_{1,1} + a_{2,2} + \cdots + a_{n,n} = \sum_{i=1}^{n} a_{i,i}$$

12.2. Propriétés

Prenons deux matrices carrées A et B de même taille et calculons le produit $(A+B)(A+B)$. Nous obtenons $A^2 + AB + BA + B^2$. Si $AB = BA$ alors nous retrouvons l'identité remarquable bien connue. Par contre si $AB \neq BA$ alors l'identité remarquable est fausse. La propriété 12.1 suivante permet de généraliser cela.

Propriété 12.1 (Formule du binôme). Soit A et B deux matrices carrées telles que $AB = BA$ (on dit que A et B commutent) alors

$$\forall n \in \mathbb{N},\ (A+B)^n = \sum_{k=0}^{n} \binom{n}{k} A^k B^{n-k}$$

Démonstration. La démonstration est la même que pour la formule du binôme de Newton dans $\mathbb{R}$. Il faut simplement montrer d'abord que si A et B commutent alors $A^{n_1} B^{n_2} = B^{n_2} A^{n_1}$. □

Propriété 12.2. Soit A et B deux matrices et $\lambda \in \mathbb{K}$. Alors

1. ${}^t({}^tA) = A$.
2. ${}^t(A+B) = {}^tA + {}^tB$.
3. ${}^t(\lambda A) = \lambda\, {}^tA$.
4. Si le produit AB est défini, alors ${}^t(AB) = {}^tB\,{}^tA$

Démonstration. Les trois premiers points sont évidents. Montrons le dernier.
Puisque le produit AB est défini, le nombre de colonnes de A est égal au nombre de lignes de B, donc le nombre de colonnes de ${}^tB$ est égal au nombre de lignes de ${}^tA$. Le produit ${}^tB\,{}^tA$ a donc un sens.
De plus, posons $AB = C$ et ${}^tB\,{}^tA = D$. Supposons $A \in \mathscr{M}_{n,p}(K)$ et $B \in \mathscr{M}_{p,m}(\mathbb{K})$. On a pour tout $1 \leq i \leq n$ et pour tout $1 \leq j \leq m$, $c_{i,j} = \sum_{k=1}^{p} a_{i,k}b_{k,j}$ et $d_{j,i} = \sum_{k=1}^{p} b_{k,j}a_{i,k} = c_{i,j}$. On a donc bien $D = {}^tC$. □

Propriété 12.3. Soit A et B deux matrices carrées de même taille. Alors

1. $\mathrm{tr}(A+B) = \mathrm{tr}(A) + \mathrm{tr}(B)$.
2. $\mathrm{tr}(AB) = \mathrm{tr}(BA)$.
3. Pour tout $\lambda \in \mathbb{K}$, $\mathrm{tr}(\lambda A) = \lambda \mathrm{tr}(A)$.

Démonstration. Alors

1. Les éléments diagonaux de $A+B$ sont la somme des éléments diagonaux de A et de B.
2. Soit $C = AB$ et $D = BA$. On a pour tout i, $c_{i,i} = \sum_{k=1}^{n} a_{i,k}b_{k,i}$ et pour tout k, $d_{k,k} = \sum_{i=1}^{n} b_{k,i}a_{i,k}$. Ainsi, $\mathrm{tr}(BA) = \mathrm{tr}(D) = \sum_{k=1}^{n} d_{kk} = \sum_{k=1}^{n}\sum_{i=1}^{n} b_{k,i}a_{i,k} = \sum_{k=1}^{n}\sum_{i=1}^{n} a_{i,k}b_{k,i} = \sum_{i=1}^{n}\sum_{k=1}^{n} a_{i,k}b_{k,i} = \sum_{i=1}^{n} c_{i,i} = \mathrm{tr}(C) = \mathrm{tr}(AB)$.
3. Évident.

Et voila. □

12.3. Matrices inversibles

Nous avons vu plus haut que nous pouvions écrire un système linéaire sous la forme $AX = B$ où A est la matrice associée au système linéaire, X le vecteur inconnu et B la matrice du second membre. Nous aimerions donc pouvoir écrire naturellement les solutions sous la forme $X = A^{-1}B$. Pour cela il nous faut d'abord définir ce que veux dire ce « A^{-1} ».

12.3.1. Théorie

Propriété 12.4. Soit A une matrice carrée de taille n, alors $AI_n = I_nA = A$.

Démonstration. Il suffit d'écrire le produit pour s'en rendre compte, mais faisons tout de même la démonstration. Notons $B = I_n$ et $C = AB$. Par définition de la matrice identité, nous avons $b_{i,j} = 0$ pour $i \neq j$ et $b_{i,i} = 1$. Par définition du produit, nous avons $c_{i,j} = \sum_{k=1}^{n} a_{i,k}b_{k,j}$ d'où $c_{i,j} = a_{i,j}$ et $C = A$. □

Nous voyons que la matrice identité joue le rôle du « 1 » pour les nombres ou de la fonction identité pour les fonctions. Nous pouvons encore pousser la comparaison plus loin en définissant l'inverse d'une matrice carrée.

Définition 12.15 (Inverse d'une matrice carrée). Soit A une matrice carrée de taille n. On appelle *inverse* de A que l'on note A^{-1} une matrice (si elle existe) telle que $AA^{-1} = A^{-1}A = I_n$.

Remarque 12.8. Puisque nous devons avoir $AA^{-1} = A^{-1}A$, la définition n'a de sens que pour les matrices carrées.

Propriété 12.5. Si une matrice est inversible alors elle admet un unique inverse.

Démonstration. Le magma $(\mathscr{M}\,;+)$ est associatif, donc voir la démonstration dans le chapitre strucure. □

Exemple 12.11. Soit $A = \begin{pmatrix} 4 & 1 \\ 7 & 2 \end{pmatrix}$ alors

$$\begin{array}{cc} & \begin{pmatrix} 4 & 1 \\ 7 & 2 \end{pmatrix} \\ \begin{pmatrix} 2 & -1 \\ -7 & 4 \end{pmatrix} & \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \end{array}$$

et

$$\begin{array}{cc} & \begin{pmatrix} 2 & -1 \\ -7 & 4 \end{pmatrix} \\ \begin{pmatrix} 4 & 1 \\ 7 & 2 \end{pmatrix} & \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \end{array}$$

Ainsi, $\begin{pmatrix} 2 & -1 \\ -7 & 4 \end{pmatrix}$ est l'inverse de A et réciproquement.

Propriété 12.6. Soit A et B deux matrices carrées de taille n telles que $AB = I_n$. Alors $BA = I_n$, et A et B sont inverses l'une de l'autre.

Démonstration. Admis. □

Propriété 12.7. Soit A une matrice carrée de taille n. Alors A est inversible si et seulement si tout système linéaire de matrice associée A (c'est-à-dire tout système de la forme $AX = B$) possède une unique solution.

Démonstration. Encore

- $\boxed{\Rightarrow}$ Si A est inversible, alors de l'égalité $AX = B$, on peut déduire que $A^{-1}AX = A^{-1}B$, c'est-à-dire $X = A^{-1}B$. Ainsi, un tel système possède au moins une solution. Montrons l'unicité : supposons que nous avons deux solutions X_1 et X_2. Nous avons donc $AX_1 = B$ et $AX_2 = B$. En soustrayant ces deux égalités nous obtenons $A(X_1 - X_2) = 0_{n,1}$. Puisque A est inversible, nous pouvons multiplier à gauche par A^{-1} : $A^{-1}A(X_1 - X_2) = A^{-1}0_{n,1}$, c'est-à-dire $X_1 - X_2 = 0_{n,1}$, d'où $X_1 = X_2$ et l'unicité de la solution.
- $\boxed{\Leftarrow}$ Notons B_i la matrice colonne formée de la i-ième colonne de I_n. Puisque tout système de la forme $AX = B$ admet une unique solution, nous pouvons noter X_i la solution de $AX = B_i$. Notons C la matrice carrée dont la i-ième colonne est X_i. Nous avons alors par construction $AC = I_n$. D'où A est inversible d'inverse C.

Remarque 12.9. Ainsi, pour résoudre un système linéaire de la forme $AX = B$ dont on connait déjà A^{-1}, il suffit de calculer

$$X = A^{-1}B$$

Propriété 12.8. Une matrice carrée de taille n est inversible si et seulement si elle est de rang n.

Démonstration. Admis. □

Corollaire 12.1. Un système linéaire à n équations et n inconnues possède une unique solution si et seulement si il est de rang n.

Démonstration. Cela découle des deux propriétés 12.8 et 12.7 précédentes. □

Propriété 12.9. Soit A une matrice triangulaire supérieure (resp. inférieure). Alors A est inversible si et seulement si tous ses coefficients diagonaux sont non nuls (c'est-à-dire si et seulement si le produit des coefficients diagonaux est non nul). Dans ce cas, A^{-1} est aussi une matrice triangulaire supérieure (resp. inférieure).

Démonstration. Si la matrice est triangulaire, alors son rang est égal au nombre de coefficients non nuls de sa diagonale. L'équivalence découle donc de la propriété 12.8. Le fait que la matrice inverse soit triangulaire est admis.□

Propriété 12.10. Soit A et B deux matrices de tailles n inversibles. Alors

1. le produit AB est inversible et $(AB)^{-1} = B^{-1}A^{-1}$.
2. A^m est inversible et $(A^m)^{-1} = \left(A^{-1}\right)^m$.
3. La transposée de A est inversible et $({}^tA)^{-1} = {}^t\left(A^{-1}\right)$.

Démonstration. Voili

1. $(B^{-1}A^{-1})(AB) = B^{-1}(A^{-1}A)B = B^{-1}I_nB = B^{-1}B = I_n$.
2. $(A^{-1})^m A^m = (A^{-1}A)^m = (I_n)^m = I_n$.
3. ${}^t(A^{-1})\,{}^tA = {}^t(AA^{-1}) = {}^tI_n = I_n$.

Définition 12.16. On appelle *groupe linéaire d'ordre n* l'ensemble des matrices carrées de tailles n inversibles et à coefficients dans $\mathbb{K}$. On note cet ensemble $\mathrm{GL}_n(\mathbb{K})$.

Propriété 12.11. Pour tout $n \in \mathbb{N}^*$, le groupe linéaire d'ordre n est un groupe pour la multiplication des matrices.

Démonstration. Il est inclus dans l'ensemble des matrices qui est un groupe. Il contient la matrice identité (l'élément neutre) et si deux matrices sont inversibles, alors leur produit aussi. □

12.3.2. Pratique

Voyons maintenant concrètement comment calculer l'inverse d'une matrice.

D'après la démonstration de la propriété 12.7, les colonnes de l'inverse d'une matrice sont déterminées en résolvant des systèmes linéaires dont les seconds membres sont les colonnes de la matrice identité. Mis à part les seconds membres, tous les systèmes à résoudre sont identiques, donc l'algorithme du pivot de Gauss nous mènera exactement aux mêmes calculs. Le principe de la méthode consiste à résoudre tous les systèmes en même temps.

Propriété 12.12 (Algorithme de Gauss-Jordan). Soit A une matrice carrée de taille n inversible. L'inverse A^{-1} de A se calcule de la manière suivante :

1. On forme la matrice augmentée $n \times 2n$ suivante $B = (A|I_n)$.
2. On applique l'algorithme de Gauss à B.
3. Une fois le système échelonné obtenu, on continue sur le même principe que l'algorithme de Gauss afin d'obtenir une matrice de la forme $(I_n|C)$
4. L'inverse est $A^{-1} = C$.

Démonstration. Admis. □

Exemple 12.12. Inverser la matrice $A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 6 \end{pmatrix}$.

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 2 & 3 & 4 & 0 & 1 & 0 \\ 3 & 4 & 6 & 0 & 0 & 1 \end{array}\right)$$

$L_2 \leftarrow L_2 - 2L_1$, $L_3 \leftarrow L_3 - 3L_1$

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & -1 & -2 & -2 & 1 & 0 \\ 0 & -2 & -3 & -3 & 0 & 1 \end{array}\right)$$

$L_2 \leftarrow -L_2$, $L_3 \leftarrow L_3 + 2L_2$

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & 2 & 2 & -1 & 0 \\ 0 & 0 & 1 & 1 & -2 & 1 \end{array}\right)$$

$L_2 \leftarrow L_2 - 2L_3$, $L_1 \leftarrow L_1 - 3L_3$

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 0 & -2 & 6 & -3 \\ 0 & 1 & 0 & 0 & 3 & -2 \\ 0 & 0 & 1 & 1 & -2 & 1 \end{array}\right)$$

$L_1 \leftarrow L_1 - 2L_2$

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -2 & 0 & 1 \\ 0 & 1 & 0 & 0 & 3 & -2 \\ 0 & 0 & 1 & 1 & -2 & 1 \end{array}\right)$$

L'inverse de A est donc

$$A^{-1} = \begin{pmatrix} -2 & 0 & 1 \\ 0 & 3 & -2 \\ 1 & -2 & 1 \end{pmatrix}$$

13. Déterminants

> There's al–gebra. That's like sums with letters. For... for people whose brains aren't clever enough for numbers, see ?
>
> *(Sergeant Colon, Jingo)*

Présentation

Nous avons déjà rencontré deux déterminants en première année : celui de deux vecteurs du plan ($\mathbb{R}^2$) et celui de trois vecteurs de l'espace ($\mathbb{R}^3$). Nous n'allons pas rappeler la définition géométrique qui avait été donnée, simplement rappeler deux propriétés :

Propriété 13.1 (déterminant de deux vecteurs de $\mathbb{R}^2$). Soit $\vec{u}\begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix}$ et $\vec{v}\begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix}$ deux vecteurs de $\mathbb{R}^2$ muni d'une base orthonormée directe. Le déterminant de $\vec{u}$ et $\vec{v}$ est donné par :

$$\det(\vec{u}, \vec{v}) = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{21}a_{12}$$

Propriété 13.2 (déterminant de trois vecteurs de $\mathbb{R}^3$). Soit $\vec{u}\begin{pmatrix} a_{11} \\ a_{21} \\ a_{31} \end{pmatrix}$, $\vec{v}\begin{pmatrix} a_{12} \\ a_{22} \\ a_{32} \end{pmatrix}$ et $\vec{w}\begin{pmatrix} a_{13} \\ a_{23} \\ a_{33} \end{pmatrix}$ trois vecteurs de $\mathbb{R}^3$ muni d'une base orthonormée directe. Le déterminant de $\vec{u}$, $\vec{v}$ et $\vec{w}$ est donné par :

$$\det(\vec{u}, \vec{v}, \vec{w}) = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{21}\begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + a_{31}\begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix}$$

Sur cette dernière propriété, vous voyez que l'on peut calculer le déterminant de trois vecteurs de $\mathbb{R}^3$ à partir de celui de deux vecteurs de $\mathbb{R}^2$. C'est en généralisant cette formule que nous allons définir le déterminant de n vecteurs de $\mathbb{R}^n$, en utilisant ce que l'on appelle une définition par récurrence.

13.1. Déterminant d'une matrice carré

13.1.1. Définitions

Définition 13.1 (Mineure d'une matrice).
Soit $A = (a_{ij})$ une matrice carrée de taille $n \times n$. Pour tout couple $(i, j) \in [\![1\,;n]\!]^2$, on appelle *mineure d'indice* i, j la matrice de taille $(n-1) \times (n-1)$ obtenue en supprimant la i^{e} ligne et la j^{e} colonne de A.

Exemple 13.1. Soit $A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 16 \end{pmatrix}$. La mineure d'indice $2,3$ est :

$$A_{2,3} = \begin{pmatrix} 1 & 2 & \cancel{3} & 4 \\ \cancel{5} & \cancel{6} & \cancel{7} & \cancel{8} \\ 9 & 10 & \cancel{11} & 12 \\ 13 & 14 & \cancel{15} & 16 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 4 \\ 9 & 10 & 12 \\ 13 & 14 & 16 \end{pmatrix}$$

Définition 13.2 (Déterminant d'une matrice carrée).
Soit $A = (a_{ij})$ une matrice carrée de taille $n \times n$ et pour tout $(i, j) \in [\![1\,;n]\!]^2$, A_{ij} la matrice obtenue en supprimant la i^{e} ligne et la j^{e} colonne de A. On appelle *déterminant de A* le nombre défini par :

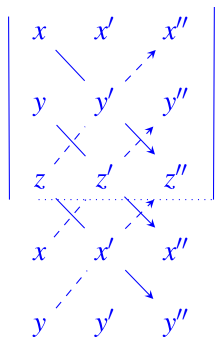
— Si $n = 2$, $\det(A) = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{21}a_{12}$

— Si $n > 2$, $\det(A) = \sum_{i=1}^{n} (-1)^{i+1} a_{i1} \det(A_{i1})$

Remarque 13.1.

— Notez bien le $(-1)^{1+i}$ dans la somme.

— Si A est une matrice 3×3, on peut utiliser la règle de Sarrus :



Exemple 13.2. Calculons le déterminant de la matrice de l'exemple 13.1. Nous avons

$$\det(A)=\begin{vmatrix}1&2&3&4\\5&6&7&8\\9&10&11&12\\13&14&15&16\end{vmatrix}$$

$$=1\times\begin{vmatrix}1&2&3&4\\5&6&7&8\\9&10&11&12\\13&14&15&16\end{vmatrix}-5\times\begin{vmatrix}1&2&3&4\\5&6&7&8\\9&10&11&12\\13&14&15&16\end{vmatrix}+9\times\begin{vmatrix}1&2&3&4\\5&6&7&8\\9&10&11&12\\13&14&15&16\end{vmatrix}-13\times\begin{vmatrix}1&2&3&4\\5&6&7&8\\9&10&11&12\\13&14&15&16\end{vmatrix}$$

$$=\begin{vmatrix}6&7&8\\10&11&12\\14&15&16\end{vmatrix}-5\begin{vmatrix}2&3&4\\10&11&12\\14&15&16\end{vmatrix}+9\begin{vmatrix}2&3&4\\6&7&8\\14&15&16\end{vmatrix}-13\begin{vmatrix}2&3&4\\6&7&8\\10&11&12\end{vmatrix}$$

Soit on continue à développer les déterminants, soit on utilise la règle de Sarrus (qui revient en réalité à faire un développement, mais à ne pas l'écrire) :

$$\begin{aligned}\det(A)=&\,[(6\times11\times16+10\times15\times8+14\times7\times12)-(14\times11\times8+15\times12\times6+16\times10\times7)]\\&-5\,[(2\times11\times16+10\times15\times4+14\times3\times12)-(14\times11\times4+15\times12\times2+16\times10\times3)]\\&+9\,[(2\times7\times16+6\times15\times4+14\times3\times8)-(14\times7\times4+15\times8\times2+16\times6\times3)]\\&-13\,[(2\times7\times12+6\times11\times4+10\times3\times8)-(10\times7\times4+11\times8\times2+12\times6\times3)]\\=&\,0\end{aligned}$$

13.1.2. Propriétés

Propriété 13.3. Soit $A=(a_{ij})$ une matrice de taille $n\times n$. Alors

$$\det\left({}^tA\right)=\det(A)$$

Démonstration. C'est évident pour une matrices 2×2, on peut le vérifier aisément pour une matrice 3×3 avec la règle de Sarrus. Pour les matrices plus grosse, la démonstration est hors programme. □

Remarque 13.2. Cette propriété 13.3 permet d'affirmer que toute les propriétés sur les lignes sont valables sur les colonnes (et inversement). Dans la suite nous énoncerons les propriétés à la fois sur les lignes et sur les colonnes, mais les démonstrations ne seront faites que sur les colonnes.

Propriété 13.4. Soit $A=(a_{ij})$ une matrice de taille $n\times n$, $j\in[\![1;n]\!]$ et B_j la matrice obtenue en intervertissant la première et la j^{e} colonne (ou ligne) de A. Alors

$$\det(B_j)=-\det(A)$$

Démonstration. Là encore c'est évident pour les matrices 2×2 et 3×3. On l'admet pour les matrices de tailles supérieures. □

Propriété 13.5. Échanger deux colonnes (ou deux lignes) de A a pour effet de multiplier le déterminant par (-1).

Démonstration. Supposons que l'on échange les colonnes i et j, avec $i \neq j$. Si $i = 1$ ou $j = 1$, on est dans le cadre de la propriété précédente. Supposons donc que $i \neq 1$ et $j \neq 1$. Alors l'inversion de la colonne i et de la colonne j revient à échanger la colonne i et la première colonne (donc multiplication par (-1)), puis échanger la colonne j et la première colonne (donc la colonne j passe en première position et la colonne i passe en j^e position, et multiplication par (-1)), et enfin à échanger la colonne i (qui était la première) et la première colonne (qui était la j^e) (donc multiplication par (-1)). Au final nous avons fait trois multiplications par (-1), donc le déterminant est bien multiplié par (-1). □

Propriété 13.6. Le déterminant d'une matrice ayant deux colonnes (ou deux lignes) égales est nul.

Démonstration. Supposons que les colonnes i et j soient identiques. Alors en intervertissant ces deux colonnes, on multiplie le déterminant par (-1). Or la matrice reste la même, donc le déterminant ne change pas. On en déduit que le déterminant est égal à son opposé, il est donc nul. □

Propriété 13.7. Soit λ un réel. Multiplier une colonne (ou une ligne) d'une matrice par λ, multiplie son déterminant par λ.

Démonstration. C'est évident pour la première colonne en revenant à la définition récursive du déterminant. Les propriétés précédentes permettent alors de conclure. □

Propriété 13.8. Soit A une matrice de taille $n \times n$ et $\lambda \in \mathbb{R}$. Alors

$$\det(\lambda A) = \lambda^n \det(A)$$

Démonstration. C'est évident si on remarque que toutes les colonnes (ou toutes les lignes) de A sont multipliées par λ. □

Propriété 13.9. Ajouter à une colonne (ou une ligne) une combinaison linéaire des autres colonnes (ou ligne) ne modifie par le déterminant.

Démonstration. Admis. □

Corollaire 13.1. Le déterminant d'une matrice ayant une colonne (ou une ligne) nulle est nul.

Démonstration. En ajoutant à la colonne nulle n'importe qu'elle autre colonne, on ne change pas le déterminant. Or on obtient une matrice ayant deux colonnes identiques. Le déterminant est donc nul. □

Propriété 13.10. Une matrice carrée est inversible si et seulement si son déterminant est non nul.

Démonstration. Admis. Nous avons déjà vu ce résultat lorsque nous avons inversé la matrice 2×2 $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ en exercice. □

Propriété 13.11. Le déterminant du produit de deux matrices est égal au produit des déterminants : Pour A et B de taille $n \times n$,

$$\det(AB) = \det(A) \times \det(B)$$

Démonstration. Hors programme. □

13.1.3. Calculs pratiques de déterminant

Toutes les propriétés de la section précédente vont nous permettre d'écrire d'autres propriétés qui vont faciliter les calculs de déterminants.

Tout d'abord remarquons que le calcul d'un déterminant $n \times n$ à partir de la définition, revient à calculer n déterminants de tailles $(n-1) \times (n-1)$. Si la première colonne a beaucoup de zéros (voire un seul coefficient non nul) alors cela fait moins de déterminants à calculer.
Les propriétés 13.9 et 13.7 vont nous permettre d'appliquer l'algorithme de Gauss sur la matrice pour tenter de mettre le plus de zéros possible en première colonne. Cependant, l'algorithme de Gauss étant lui aussi couteux en calculs, il faut réussir à faire un compromis et on ne le mène pas forcément jusqu'au bout.

Propriété 13.12 (Développement selon une ligne et une colonne). Soit $A = (a_{ij})$ une matrice de taille $n \times n$ et $(i, j) \in [\![1\,;n]\!]^2$. Alors

$$\det(A) = \sum_{k=1}^{n} (-1)^{k+j} a_{kj} \det\left(A_{kj}\right) \qquad \text{(développement selon la colonne } j\text{)}$$

$$\det(A) = \sum_{k=1}^{n} (-1)^{i+k} a_{ik} \det\left(A_{ik}\right) \qquad \text{(développement selon la ligne } i\text{)}$$

Démonstration. En intervertissant la colonne j avec la colonne $(j-1)$, puis la colonne $(j-1)$ avec la colonne $(j-2)$, *etc*, jusqu'à la colonne 2 et la colonne 1, on a déplacé la j^{e} colonne en première position et décalé les colonnes 1 à $(j-1)$ d'un cran vers la droite. On a de plus fait $(j-1)$ inversions de colonnes, donc multiplié le déterminant par $(-1)^{j-1}$. Maintenant que la colonne j est en première position, on peut appliquer la définition du déterminant et retrouver la formule annoncée $\left((-1)^{1+i} \times (-1)^{j-1} = (-1)^{i+j}\right)$. □

Cette propriété nous permet de choisir la ligne ou la colonne ayant le plus de zéros possibles pour calculer plus efficacement le déterminant. Attention cependant de ne pas oublier le $(-1)^{i+j}$!

Exemple 13.3. Soit $A = \begin{pmatrix} 1 & 2 & 0 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 0 & 12 \\ 14 & 15 & 2 & 18 \end{pmatrix}$. Calculons ce déterminant. On remarque d'abord que sur la 3^{e} colonne il y a la moitié des coefficients qui sont nuls. Nous allons donc développer selon la troisième colonne :

$$\begin{aligned}
\det(A) &= (-1)^{3+2} \times 7 \begin{vmatrix} 1 & 2 & 4 \\ 9 & 10 & 12 \\ 14 & 15 & 18 \end{vmatrix} + (-1)^{3+4} \times 2 \begin{vmatrix} 1 & 2 & 4 \\ 5 & 6 & 8 \\ 9 & 10 & 12 \end{vmatrix} \\
&= -7\,(180 + 15 \times 9 \times 4 + 336 - 560 - 15 \times 12 - 18 \times 18) - 2\,(12 \times 6 + 200 + 144 - 9 \times 24 - 80 - 120) \\
&= 7 \times 8 - 2 \times 0 \\
&= 56
\end{aligned}$$

Propriété 13.13 (Matrice triangulaire). Le déterminant d'une matrice triangulaire est égal au produit des coefficients de la diagonale.

Démonstration. Supposons qu'elle soit triangulaire supérieure. On développe selon la première colonne qui ne possède que des zéros sauf sur la première ligne. Donc le déterminant est égal à a_{11} multiplié par le déterminant de la matrice obtenue en supprimant la première ligne et la première colonne. On retombe sur une matrice triangulaire supérieure. Par récurrence on montre donc le résultat. □

Corollaire 13.2 (Matrice diagonale). Le déterminant d'une matrice diagonale est égal au produit des coefficients (de la diagonale).

Démonstration. Une matrice diagonale est une matrice triangulaire bien particulière. □

Corollaire 13.3 (Matrice identité). Le déterminant de la matrice identité est égal à 1.

Démonstration. Vraiment ? ? □

Propriété 13.14 (Matrice inverse). Le déterminant de l'inverse d'une matrice (inversible) est égal à l'inverse du déterminant :

$$\det\left(A^{-1}\right) = \frac{1}{\det(A)}$$

Démonstration. Tout d'abord remarquons que la formule a du sens puisqu'une matrice est inversible si et seulement si son déterminant est non nul (propriété 13.10). Soit A une matrice inversible. On a $AA^{-1} = I_n$ d'où $1 = \det(I_n) = \det(AA^{-1}) = \det(A)\det(A^{-1})$. D'où le résultat. □

Propriété 13.15 (Matrice triangulaire par bloc). Soit $n \in \mathbb{N}^*$, B une matrice de taille $p \times p$ avec $p \leq n$, C une matrice de taille $p \times (n-p)$, D une matrice de taille $(n-p) \times (n-p)$ et A la matrice définie par

$$A = \left(\begin{array}{c|c} B & C \\ \hline 0 & D \end{array}\right)$$

Alors :

$$\det(A) = \det(B) \times \det(D)$$

Démonstration. Admis. □

Remarque 13.3. La propriété a été énoncée pour deux blocs diagonaux, mais vous pouvez montrer facilement par récurrence que l'on peut la généraliser à plusieurs blocs diagonaux.

Exemple 13.4. Soit $A = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 6 & 7 & 8 & 9 & 10 \\ 0 & 0 & a & b & c \\ 0 & 0 & d & e & f \\ 0 & 0 & 0 & 0 & g \end{pmatrix}$. Alors $\det(A) = (1 \times 7 - 6 \times 2) \times (ae - bd) \times g = -5g(ae - bd)$.

13.2. Déterminant d'une famille de vecteurs

Nous avons introduit ce chapitre avec le déterminants de vecteurs de $\mathbb{R}^2$ et $\mathbb{R}^3$ pour justifier la définition du déterminant d'une matrice. Mais qu'en est-il de la définition du déterminant de n vecteurs de $\mathbb{R}^n$?

Définition 13.3 (Déterminant de n vecteurs).
Soit $\mathscr{B} = (e_1, \dots, e_n)$ une base de $\mathbb{R}^n$. Soit $u_1 \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{pmatrix}$, $u_2 \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{n2} \end{pmatrix}$, ..., $u_n \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{nn} \end{pmatrix}$, n vecteurs de $\mathbb{R}^n$ dont les coordonnées sont données dans la base $\mathscr{B}$. Soit $A = (a_{ij})$ la matrice $n \times n$ construites à partir des vecteurs u_i. On définit le déterminant de la famille $(u_1, \dots, u_n)$ dans la base $\mathscr{B}$ par

$$\det_{\mathscr{B}}(u_1, \dots, u_n) = \det(A)$$

Remarque 13.4.
Puisque les coordonnées d'un vecteur changent suivant la base choisie, le déterminant d'une famille de vecteurs change selon la base dans lequel on le calcul. Il existe une formule de « changement de bases » qui n'est pas au programme qui permet de calculer le déterminant d'une famille de vecteurs dans une base $\mathscr{B}'$ connaissant son déterminant dans la base $\mathscr{B}$: $\det_{\mathscr{B}'}(u_1, \dots, u_n) = \det_{\mathscr{B}}(u_1, \dots, u_n) \times \det_{\mathscr{B}'}(\mathscr{B})$.

On pourrait montrer que le déterminant d'une base orthonormée dans une autre base orthonormée vaut soit 1 soit (-1). Cela permet de classer les bases orthonormées en deux catégories : les bases directes (qui ont un déterminant égal à 1 dans la base canonique) et les bases indirectes. Je vous invite à faire le parallèle avec ce qui a été vue dans le chapitre « Géométrie de l'espace » de première année.

Tout cela permet de voir que le déterminant d'une famille de vecteurs ne change pas lorsque l'on passe d'une base orthonormée à une autre de même direction, et qu'il est multiplié par (-1) si on change l'orientation de la base.

Remarque 13.5. Toutes les propriétés des sections précédentes peuvent être retranscrites en termes de vecteurs. Par exemple, la propriété 13.5 se réécrit : Échanger deux vecteurs de la famille a pour effet de multiplier le déterminant par (-1).

Propriété 13.16 (Caractérisation d'une famille liée). Soit $(u_1, \dots, u_n)$ une famille de n vecteurs de $\mathbb{R}^n$. Il y a équivalence entre les assertions suivantes :

1. La famille $(u_1, \dots, u_n)$ est liée.
2. Pour toutes bases $\mathscr{B}$ de $\mathbb{R}^n$, on a $\det_{\mathscr{B}}(u_1, \dots, u_n) = 0$
3. Il existe une base $\mathscr{B}$ de $\mathbb{R}^n$ telle que $\det_{\mathscr{B}}(u_1, \dots, u_n) = 0$

Démonstration.

$\boxed{1 \Rightarrow 2}$ Si la famille est liée alors il existe une combinaison linéaire, à coefficients non tous nuls, nulle. Donc l'un des vecteurs peut s'écrire comme combinaison linéaire des autres. En soustrayant à ce vecteur cette combinaison linéaire des autres, on ne change pas le déterminant et on transforme la colonne correspondante en une colonne nulle. Le déterminant, quelque soit la base, est nul.

2⇒3 Si la propriété est vraie pour toute les bases alors elle est vraie pour l'une d'elle...

3⇒1 Admis.

Propriété 13.17 (Caractérisation d'une base). Soit $(u_1,\dots,u_n)$ une famille de n vecteurs de $\mathbb{R}^n$. Il y a équivalence entre les assertions suivantes :

1. La famille $(u_1,\dots,u_n)$ est une base.
2. Pour toutes bases $\mathscr{B}$ de $\mathbb{R}^n$, on a $\det_{\mathscr{B}}(u_1,\dots,u_n)\neq 0$
3. Il existe une base $\mathscr{B}$ de $\mathbb{R}^n$ telle que $\det_{\mathscr{B}}(u_1,\dots,u_n)\neq 0$

Démonstration. Un déterminant est soit nul, soit non nul. Une famille libre de n vecteurs de $\mathbb{R}^n$ est une base, donc une famille de n vecteurs est soit une base soit liée. Donc cette propriété est simplement une réécriture de la propriété 13.16. □

13.3. Applications

13.3.1. Géométrie

— Nous avons déjà vu l'année dernière que le déterminant de deux vecteurs du plan est égal à la surface algébrique du parallélogramme construit à partir de ces deux vecteurs.

— Nous avons déjà vu l'année dernière que le déterminant de trois vecteurs de l'espace est égal au volume algébrique du parallélépipède construit à partir de ces trois vecteurs.

13.3.2. Résolution de systèmes linéaires

Propriété 13.18 (Formules de Cramer). Soit A une matrice de taille $n\times n$ de déterminant non nul. Soit B une matrice (colonne) de taille $n\times 1$. Alors le système linéaire $AX=B$ possède une unique solution dont les composantes sont données par :

$$\forall i\in [\![1\,;n]\!],\ x_i=\frac{\det(C_1,\dots,C_{i-1},B,C_{i+1},\dots,C_n)}{\det(A)}$$

où $(C_1,\dots,C_n)$ sont les colonnes de A.

Démonstration. Puisque $B=x_1C_1+\cdots+x_nC_n$, on a $\det(C_1,\dots,C_{i-1},B,C_{i+1},\dots,C_n)=x_i\det(A)$. □

Remarque 13.6. Résoudre un système avec cette formule demande de calculer $(n+1)$ déterminants, chaque calcul de déterminant pouvant se faire par pivot de Gauss. Donc suivant la taille du système à résoudre, ce n'est pas forcément une méthode très efficace.

Exemple 13.5. Résoudre $\begin{cases} 2x-y+z &= 0\\ x-2y-z &= 3\\ 3x+y+2z &= 1\end{cases}$

On a $\det(A) = \begin{vmatrix} 2 & -1 & 1 \\ 1 & -2 & -1 \\ 3 & 1 & 2 \end{vmatrix} = 6$. D'où

$$x = \frac{\begin{vmatrix} 0 & -1 & 1 \\ 3 & -2 & -1 \\ 1 & 1 & 2 \end{vmatrix}}{6} = \frac{12}{6} = 2 \quad ; \quad y = \frac{\begin{vmatrix} 2 & 0 & 1 \\ 1 & 3 & -1 \\ 3 & 1 & 2 \end{vmatrix}}{6} = \frac{6}{6} = 1 \quad ; \quad z = \frac{\begin{vmatrix} 2 & -1 & 0 \\ 1 & -2 & 3 \\ 3 & 1 & 1 \end{vmatrix}}{6} = \frac{-18}{6} = -3$$

14. Représentation matricielle des applications linéaires

Dans tout ce chapitre, $\mathbb{K}$ désigne l'un des corps $\mathbb{R}$ ou $\mathbb{C}$. Les lettres n, p et q sont des entiers naturels non nuls.

14.1. Matrice d'une application linéaire

14.1.1. Matrice d'une famille finie de vecteurs dans une base finie

Définition 14.1 (Matrice d'une famille finie de vecteurs dans une base finie). Soit E un $\mathbb{K}$-espace vectoriel de dimension finie n, $\mathscr{B} = (e_1, \dots, e_n)$ une base de E et $\mathscr{X} = (X_1, \dots, X_p)$ une famille finie de vecteurs de E.
Pour tout $j \in [\![1\,;p]\!]$ on note $(a_{1j}, \dots, a_{nj})$ les coordonnées de X_j dans la base $\mathscr{B}$. La matrice $(a_{ij})_{\substack{1\leq i\leq n\\ 1\leq j\leq p}}$, notée $\mathrm{Mat}_{\mathscr{B}}(\mathscr{X})$, est appelée *matrice de $\mathscr{X}$ dans la base $\mathscr{B}$*.

$$\begin{array}{cc}
 & \begin{array}{ccccc} X_1 & & X_j & & X_p \\ \downarrow & & \downarrow & & \downarrow \end{array} \\
\mathrm{Mat}_{\mathscr{B}}(\mathscr{X}) = & \begin{pmatrix} a_{11} & \dots & a_{1j} & \dots & a_{1p} \\ \vdots & & \vdots & & \vdots \\ a_{i1} & \dots & a_{ij} & \dots & a_{ip} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \dots & a_{nj} & \dots & a_{np} \end{pmatrix} \begin{array}{l} \leftarrow \; e_1 \\ \\ \leftarrow \; e_i \\ \\ \leftarrow \; e_n \end{array} \\
 & \uparrow \\
 & \text{Coordonnées de } X_j \text{ dans } \mathscr{B} \text{ écrites en colonnes}
\end{array}$$

Théorème 14.1 (Bases et matrices inversibles). Soit E un $\mathbb{K}$-espace vectoriel de dimension finie n, $\mathscr{B}$ une base de E et $\mathscr{F}$ une famille de n vecteurs de E. Alors $\mathscr{F}$ est une base de E si et seulement si $\mathrm{Mat}_{\mathscr{B}}(\mathscr{F})$ est inversible.

Démonstration. Il suffit d'utiliser la caractérisation d'une base (propriété 13.16) et le fait qu'une matrice est inversible si et seulement si son déterminant est non nul (propriété 13.10) □

Propriété 14.1. Soit $E \neq \{0_E\}$ un $\mathbb{K}$-espace vectoriel de dimension finie, $\mathscr{B}$ une base de E et $\mathscr{X}$ une famille de vecteurs de E. Alors

$$\mathrm{rg}(\mathscr{X}) = \mathrm{rg}\left(\mathrm{Mat}_{\mathscr{B}}(\mathscr{X})\right)$$

Démonstration. Admis. □

14.1.2. Matrice d'une application linéaire dans des bases

Définition 14.2 (Matrice d'une application linéaire dans des bases finies). Soit E et F deux $\mathbb{K}$-espace vectoriels de dimensions respectives p et n, $\mathscr{B} = (e_1, \ldots, e_p)$ une base de E, $\mathscr{C} = (f_1, \ldots, f_n)$ une base de F et $f \in \mathscr{L}(E;F)$. On appelle *matrice de f dans $\mathscr{B}$ et $\mathscr{C}$* et on note $\mathrm{Mat}_{\mathscr{B},\mathscr{C}}(f)$ la matrice de la famille $f(\mathscr{B}) = (f(e_1), \ldots, f(e_p))$ dans la base $\mathscr{C}$.
Si $E = F$ et $\mathscr{B} = \mathscr{C}$, la matrice $\mathrm{Mat}_{\mathscr{B},\mathscr{B}}(f)$ est simplement notée $\mathrm{Mat}_{\mathscr{B}}(f)$.

$$\mathrm{Mat}_{\mathscr{B},\mathscr{C}}(f) = \mathrm{Mat}_{\mathscr{C}}(f(\mathscr{B})) = \begin{array}{c} \begin{array}{ccccc} f(e_1) & & f(e_j) & & f(e_p) \\ \downarrow & & \downarrow & & \downarrow \end{array} \\ \begin{pmatrix} a_{11} & \ldots & a_{1j} & \ldots & a_{1p} \\ \vdots & & \vdots & & \vdots \\ a_{i1} & \ldots & a_{ij} & \ldots & a_{ip} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \ldots & a_{nj} & \ldots & a_{np} \end{pmatrix} \\ \uparrow \end{array} \begin{array}{l} \leftarrow f_1 \\ \\ \leftarrow f_i \\ \\ \leftarrow f_n \end{array}$$

Coordonnées de $f(e_j)$ dans $\mathscr{C}$ écrites en colonnes

Remarque 14.1. Nous savons que l'on connaît tout d'une application linéaire lorsque l'on connaît l'image d'une base. Ainsi une application linéaire est entièrement déterminée par sa matrice associée dans deux bases données.

Exemple 14.1.

— Si E est un $\mathbb{K}$-espace vectoriel de dimension finie n et si $\mathscr{B}$ est une base de E, alors $\mathrm{Mat}_{\mathscr{B}}(Id_E) = I_n$.

— Soit T l'endomorphisme $P \mapsto X^2P'' + P(1)$ de $\mathbb{R}_3[X]$ et $\mathscr{B}_3$ la base canonique de $\mathbb{R}_3[X]$. Alors

$$\mathrm{Mat}_{\mathscr{B}_3}(T) = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 6 \end{pmatrix}$$

Définition 14.3 (Application linéaire canoniquement associée à une matrice). Soit $A \in \mathscr{M}_{n,p}(\mathbb{K})$, $\mathscr{B}_p$ et $\mathscr{B}_n$ les bases canoniques respectives de $\mathbb{K}^p$ et $\mathbb{K}^n$. On appelle *application linéaire canoniquement associée à A* l'application linéaire f telle que

$$\mathrm{Mat}_{\mathscr{B}_p,\mathscr{B}_n}(f) = A$$

Remarque 14.2. Ainsi, donner une matrice revient à donner une application linéaire.

Propriété 14.2 (Rang d'une application linéaire). Soit E et F deux $\mathbb{K}$-espaces vectoriels de dimension finie, $\mathscr{B}$ une base de E, $\mathscr{C}$ une base de F et $f \in \mathscr{L}(E;F)$. Alors

$$\operatorname{rg}(f) = \operatorname{rg}\left(\operatorname{Mat}_{\mathscr{B},\mathscr{C}}(f)\right)$$

Démonstration.

$$\operatorname{rg}\left(\operatorname{Mat}_{\mathscr{B},\mathscr{C}}(f)\right) = \operatorname{rg}\left(\operatorname{Mat}_{\mathscr{C}}\left(f\left(\mathscr{B}\right)\right)\right) = \operatorname{rg}\left(f\left(\mathscr{B}\right)\right) = \dim \operatorname{Vect}(f(\mathscr{B})) = \dim \operatorname{Im} f = \operatorname{rg}(f)$$

D'où le résultat. □

Théorème 14.2 (Calcul matriciel de l'image d'un vecteur par une application linéaire). Soit $E \neq \{0_E\}$ et $F \neq \{0_F\}$ deux $\mathbb{K}$-espace vectoriels de dimension finie, $\mathscr{B}$ une base de E, $\mathscr{C}$ une base de F, $u \in \mathscr{L}(E;F)$ et $x \in E$. Alors

$$\operatorname{Mat}_{\mathscr{C}}\left(u(x)\right) = \operatorname{Mat}_{\mathscr{B},\mathscr{C}}(u) \times \operatorname{Mat}_{\mathscr{B}}(x)$$

Démonstration. Soit $\mathscr{B} = (e_1, \dots, e_p)$, $\mathscr{C} = (f_1, \dots, f_n)$. Posons $X = \operatorname{Mat}_{\mathscr{B}}(x)$ et $U = \operatorname{Mat}_{\mathscr{B},\mathscr{C}}(u)$.

$$u(x) = u\left(\sum_{j=1}^{p} x_j e_j\right) = \sum_{j=1}^{p} x_j u(e_j) = \sum_{j=1}^{p} x_j \sum_{i=1}^{n} u_{ij} f_i = \sum_{i=1}^{n}\left(\sum_{j=1}^{p} u_{ij} x_j\right) f_i$$

Donc les coordonnées de $u(x)$ dans $\mathscr{C}$ sont $\left(\sum_{j=1}^{p} u_{1j x_j};\dots;\sum_{j=1}^{p} u_{nj} x_j\right)$, c'est-à-dire le produit $U \times X$. □

Remarque 14.3. Le calcul de l'image d'un vecteur par une application linéaire se réduit donc à un simple produit matriciel. Attention cependant à ce que les matrices et les vecteurs soient écris dans les bonnes bases.

Exemple 14.2. Soit f l'endomorphisme de $\mathbb{R}_2[X]$ de matrice $\begin{pmatrix} 3 & 3 & 6 \\ 0 & 1 & 2 \\ 0 & 2 & 4 \end{pmatrix}$ dans la base canonique. Déterminer $\operatorname{Im} f$ et $\operatorname{Ker} f$.

Par définition

$Im f = \operatorname{Vect}\big(f(1); f(X); f(X^2)\big) = \operatorname{Vect}\big(3; 2X^2 + X + 3; 4X^2 + 2X + 6\big) = \operatorname{Vect}\big(3; 2X^2 + X + 3\big) = \operatorname{Vect}\big(3; 2X^2 + X\big)$

Puis, soit $P = aX^2 + bX + c \in \mathbb{R}_2[X]$.

$$P \in \operatorname{Ker} f \Leftrightarrow f(P) = 0 \Leftrightarrow \begin{pmatrix} 3 & 3 & 6 \\ 0 & 1 & 2 \\ 0 & 2 & 4 \end{pmatrix}\begin{pmatrix} c \\ b \\ a \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\Leftrightarrow \begin{cases} 3c+3b+6a &= 0 \\ b+2a &= 0 \\ 2b+4a &= 0 \end{cases} \Leftrightarrow \begin{cases} c &= 0 \\ b+2a &= 0 \end{cases} \Leftrightarrow P = aX^2-2aX = a(X^2-2X)$$

Donc $\operatorname{Ker} f = \operatorname{Vect}(X^2-2X)$.

14.1.3. Liens entre applications linéaires et matrices

Propriété 14.3 (Lien entre $\mathscr{L}(E;F)$ et $\mathscr{M}_{n,p}(\mathbb{K})$). Soit E et F deux $\mathbb{K}$-espaces vectoriels de dimensions finies respectives p et n, $\mathscr{B}$ une base de E et $\mathscr{C}$ une base de F. L'application $f \mapsto \operatorname{Mat}_{\mathscr{B},\mathscr{C}}(f)$ est un isomorphisme de $\mathscr{L}(E;F)$ sur $\mathscr{M}_{n,p}(\mathbb{K})$.

Démonstration. Très bon exercice. □

Remarque 14.4. Lorsque l'on parle d'isomorphisme il faut préciser de quelles lois on parle. Il est clair ici que l'on parle d'isomorphisme d'espaces vectoriels munis des lois classiques. Nous déduisons trois choses de cette propriété :

— Comme nous l'avons mentionné plus haut, connaître une application linéaire est équivalent à connaître sa matrice dans deux bases.

— $\operatorname{Mat}_{\mathscr{B},\mathscr{C}}(\lambda f+g) = \lambda \operatorname{Mat}_{\mathscr{B},\mathscr{C}}(f) + \operatorname{Mat}_{\mathscr{B},\mathscr{C}}(g)$.

— $\dim(\mathscr{M}_{n,p}) = np = \dim(E)\times\dim(F) = \dim(\mathscr{L}(E;F))$

Propriété 14.4 (Matrice d'une composée). Soit E, F, G trois $\mathbb{K}$-espaces vectoriels de dimensions finies non nulles de bases respectives $\mathscr{B}$, $\mathscr{C}$, $\mathscr{D}$ et $f \in \mathscr{L}(E;F)$ et $g \in \mathscr{L}(F;G)$. Alors,

$$\operatorname{Mat}_{\mathscr{B},\mathscr{D}}(g\circ f) = \operatorname{Mat}_{\mathscr{C},\mathscr{D}}(g)\times \operatorname{Mat}_{\mathscr{B},\mathscr{C}}(f)$$

Démonstration. Notons $\mathscr{B} = (e_1,\dots,e_n)$. Alors pour tout $j \in [\![1;n]\!]$, en utilisant le théorème de calcul matriciel de l'image d'un vecteur (théorème 14.2) :

$$\operatorname{Mat}_{\mathscr{B},\mathscr{D}}(g\circ f)\times\operatorname{Mat}_{\mathscr{B}}(e_j) = \operatorname{Mat}_{\mathscr{D}}(g\circ f(e_j)) = \operatorname{Mat}_{\mathscr{C},\mathscr{D}}(g)\times\operatorname{Mat}_{\mathscr{C}}(f(e_j)) = \operatorname{Mat}_{\mathscr{C},\mathscr{D}}(g)\times\operatorname{Mat}_{\mathscr{B},\mathscr{C}}(f)\times\operatorname{Mat}_{\mathscr{B}}(e_j)$$

Or $\operatorname{Mat}_{\mathscr{B}}(e_j)$ est le j^{e} vecteur de la base canonique de $\mathbb{K}^n$, c'est-à-dire le vecteur ayant des 0 partout sauf un 1 en j^{e} position. Nous venons donc de montrer que pour tout $j \in [\![1;n]\!]$, $\operatorname{Mat}_{\mathscr{B},\mathscr{D}}(g\circ f)$ et $\operatorname{Mat}_{\mathscr{C},\mathscr{D}}(g)\times\operatorname{Mat}_{\mathscr{B},\mathscr{C}}(f)$ ont la même j^{e} colonne. D'où l'égalité des matrices. □

Remarque 14.5. En conclusion, la matrice de la composée de deux applications linéaires est tout simplement le produit des matrices associées (attention cependant d'être dans les bonnes bases !)

Propriété 14.5 (Matrice d'un isomorphisme). Soit E et F deux $\mathbb{K}$-espaces vectoriels de mêmes dimensions finies non nulles de bases respectives $\mathscr{B}$, $\mathscr{C}$ et $f \in \mathscr{L}(E;F)$. Alors f est un isomorphisme de E sur F si et seulement si $\mathrm{Mat}_{\mathscr{B},\mathscr{C}}(f)$ est inversible. Dans ce cas,

$$\mathrm{Mat}_{\mathscr{C},\mathscr{B}}\left(f^{-1}\right) = \left(\mathrm{Mat}_{\mathscr{B},\mathscr{C}}(f)\right)^{-1}$$

Démonstration. Si f est bijective et si on pose $n = \dim(F)$, on a $\mathrm{Mat}_{\mathscr{B},\mathscr{C}}(f) \times \mathrm{Mat}_{\mathscr{C},\mathscr{B}}(f^{-1}) = \mathrm{Mat}_{\mathscr{C}}(Id_F) = I_n$. Donc $\mathrm{Mat}_{\mathscr{B},\mathscr{C}}(f)$ est inversible d'inverse $\mathrm{Mat}_{\mathscr{C},\mathscr{B}}(f^{-1})$
Réciproquement, si $A = \mathrm{Mat}_{\mathscr{B},\mathscr{C}}(f)$ est inversible, notons g l'unique application linéaire de F dans E pour laquelle $\mathrm{Mat}_{\mathscr{C},\mathscr{B}}(g) = A^{-1}$. Dans ces conditions, $\mathrm{Mat}_{\mathscr{B}}(g \circ f) = \mathrm{Mat}_{\mathscr{C},\mathscr{B}}(g) \times \mathrm{Mat}_{\mathscr{B}}\mathscr{C}(f) = A^{-1}A = I_n$. De même, $\mathrm{Mat}_{\mathscr{C},\mathscr{B}} = I_n$. Donc $g \circ f = Id_E$ et $f \circ g = Id_F$ et f est bijective de E sur F. □

14.2. Changements de bases, équivalence et similitude

14.2.1. Matrices de passage

14.2.1.1. Définition

Définition 14.4 (Matrice de passage). Soit E un $\mathbb{K}$-espace vectoriel de dimension n. On note $\mathscr{B} = (e_1, e_2, \ldots, e_n)$ et $\mathscr{B}' = (e'_1, e'_2, \ldots, e'_n)$ deux bases de E. On appelle *matrice de passage* de la base $\mathscr{B}$ à la base $\mathscr{B}'$ la matrice de l'application Id_E relativement aux bases $\mathscr{B}'$ et $\mathscr{B}$.
On la note :

$$P_{\mathscr{B}}\left(\mathscr{B}'\right) = \mathrm{Mat}_{\mathscr{B}}(\mathscr{B}') = \mathrm{Mat}_{\mathscr{B}',\mathscr{B}}(Id_E) = \begin{array}{ccc} \begin{pmatrix} p_{1,1} & p_{1,2} & \cdots & p_{1,n} \\ p_{2,1} & p_{2,2} & \cdots & p_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{n,1} & p_{n,2} & \cdots & p_{n,n} \end{pmatrix} & \begin{matrix} \leftarrow \\ \leftarrow \\ \\ \leftarrow \end{matrix} & \begin{matrix} \overrightarrow{e_1} \\ \overrightarrow{e_2} \\ \vdots \\ \overrightarrow{e_n} \end{matrix} \\ \begin{matrix} \uparrow & \uparrow & & \uparrow \\ \overrightarrow{e'_1} & \overrightarrow{e'_2} & \cdots & \overrightarrow{e'_n} \end{matrix} & & \end{array}$$

Cette matrice donne la décomposition des vecteurs de la base $\mathscr{B}'$ dans la base $\mathscr{B}$.

Remarque 14.6.
— Attention à l'ordre des bases $\mathscr{B}$ et $\mathscr{B}'$!
— On remarque que $P_{\mathscr{B}}(\mathscr{B}') = \mathrm{Mat}_{\mathscr{B}}(\mathscr{B}')$. Donc si les coordonnées des vecteurs de $\mathscr{B}'$ sont données dans la base $\mathscr{B}$ (lorsque $\mathscr{B}$ est la base canonique par exemple), alors la matrice de passage de $\mathscr{B}$ à $\mathscr{B}'$ est simplement donnée en recopiant les coordonnées des vecteurs de $\mathscr{B}'$.

Exemple 14.3. Soit $E = \mathbb{R}_2[X]$, $\mathscr{B} = \left(1, X, X^2\right)$ et $\mathscr{B}' = \left(1, X-1, (X-1)^2\right)$. On a alors $1 = 1$, donc $Id_E(1) = 1 + 0X + 0X^2$. $X - 1 = -1 + X$, donc $Id_E(X-1) = -1 + 1X + 0X^2$. Enfin, $(X-1)^2 = 1 - 2X + X^2$, d'où $Id_E((X -$

$1)^2) = 1 - 2X + 1X^2$. La matrice de passage de $\mathscr{B}$ à $\mathscr{B}'$ est donc :

$$P_{\mathscr{B}}(\mathscr{B}') = \begin{array}{cc} \begin{pmatrix} 1 & -1 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix} & \begin{array}{l} \leftarrow \; 1 \\ \leftarrow \; X \\ \leftarrow \; X^2 \end{array} \\ \begin{array}{ccc} \uparrow & \uparrow & \uparrow \\ 1 & X-1 & (X-1)^2 \end{array} & \end{array}$$

Remarque 14.7. [Astuce] Si les coordonnées des vecteurs des bases $\mathscr{B}$ et $\mathscr{B}'$ sont données dans une même base (canonique par exemple), nous pouvons calculer $P_{\mathscr{B}}(\mathscr{B}')$ de manière astucieuse.

Pour déterminer la première colonne de P, nous cherchons les coordonnées du vecteur e'_1 dans la base $\mathscr{B}$, c'est-à-dire que nous cherchons des coefficients $(x_{11}, \ldots, x_{1n})$ tels que $e'_1 = x_{11}e_1 + \cdots + x_{1n}e_n$. Pour cela on résout un système linéaire dont la matrice associée est tout simplement la matrice B dont les colonnes sont les coordonnées des vecteurs de $\mathscr{B}$ dans la base canonique. Ce qui veut dire que l'on cherche X_1 tel que $BX_1 = e'_1$. On recommence le procédé pour X_2 jusqu'à X_n. On remarque alors qu'au lieu de résoudre n systèmes linéaires parfaitement identiques, sauf en ce qui concerne le second membre, il vaut mieux tout résoudre d'un coup : On crée la matrice B' dont les colonnes sont les coordonnées des vecteurs de $\mathscr{B}'$ dans la base canonique et la matrice X dont les colonnes sont les X_i. Ainsi, nous avons l'équation matricielle $BX = B'$ dont la solution est

$$X = B^{-1}B'$$

Par définition de X et de $P_{\mathscr{B}}(\mathscr{B}')$, nous avons $X = P_{\mathscr{B}}(\mathscr{B}') = B^{-1}B'$.

14.2.1.2. Propriétés

Propriété 14.6. Soit E un $\mathbb{K}$-espace vectoriel de dimension finie, et $\mathscr{B}$ et $\mathscr{B}'$ deux bases de E. Si $P = P_{\mathscr{B}}(\mathscr{B}')$, alors P est inversible et $P^{-1} = P_{\mathscr{B}'}(\mathscr{B})$

Démonstration. C'est tout simplement une application de la propriété 14.5. □

Propriété 14.7. Soit E un espace vectoriel de dimension finie et $P \in \mathscr{M}_n(\mathbb{K})$. Alors P est inversible si et seulement si il existe deux bases $\mathscr{B}$ et $\mathscr{B}'$ de E telles que $P = P_{\mathscr{B}}(\mathscr{B}') = \mathrm{Mat}_{\mathscr{B}',\mathscr{B}}(Id_E)$.

Démonstration. $\boxed{\Leftarrow}$ C'est la propriété 14.6.

$\boxed{\Rightarrow}$ Soit (e_i) la base canonique de E. Soit, pour tout $j \in [\![1;n]\!]$, f_j le j-ième vecteur colonne de P : $f_j = \sum_{i=1}^{n} p_{ij}e_i$. Alors $\mathscr{B}' = (f_j)$ est une base de E car $\det(\mathscr{B}') = \det(P) \neq 0$.

Ainsi, pour tout $j \in [\![1;n]\!]$, $Id(f_j) = f_j = \sum_{i=1}^{n} p_{ij}e_i = \begin{pmatrix} p_{1j} \\ \vdots \\ p_{nj} \end{pmatrix}_{\mathscr{B}_c}$. D'où $\mathrm{Mat}_{\mathscr{B}',\mathscr{B}_c}(Id_E) = P$.

Remarque 14.8. Cette propriété nous dit que toute matrice inversible est une matrice de passage.

Propriété 14.8. Soit E un $\mathbb{K}$-espace vectoriel de dimension finie, et $\mathscr{B}$, $\mathscr{B}'$ et $\mathscr{B}''$ trois bases de E. Alors $P_{\mathscr{B}}(\mathscr{B}'') = P_{\mathscr{B}}(\mathscr{B}') \times P_{\mathscr{B}'}(\mathscr{B}'')$.

Démonstration. C'est tout simplement une application de la propriété 14.4. □

Propriété 14.9 (Matrice de passage et coordonnées). Soit E un $\mathbb{K}$-espace vectoriel de dimension finie n, $\overrightarrow{x} \in E$, $\mathscr{B}$ et $\mathscr{B}'$ deux bases de E, et $P = P_{\mathscr{B}}(\mathscr{B}')$. On note $X_{\mathscr{B}}$ la matrice colonne des coordonnées de $\overrightarrow{x}$ dans la base $\mathscr{B}$ et $X'_{\mathscr{B}'}$ la matrice colonne des coordonnées de $\overrightarrow{x}$ dans la base $\mathscr{B}'$. Alors :

$$X_{\mathscr{B}} = PX'_{\mathscr{B}'}$$

Démonstration. $\overrightarrow{x} = Id_E(\overrightarrow{x})$, donc le résultat découle simplement de la définition de la matrice de passage. □

Exemple 14.4. Soit $X = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ dans la base canonique $\mathscr{B}_c$, $\mathscr{B}' = (e_1 ; e_1 + e_2 ; e_3 - e_2 - 3e_1)$. Déterminer les coordonnées de X dans la base $\mathscr{B}'$.

Sans utiliser la formule : On écrit e_1, e_2, e_3 en fonction des vecteurs ε_1, ε_2, ε_3 de $\mathscr{B}'$:

$$\begin{cases} \varepsilon_1 & = e_1 \\ \varepsilon_2 & = e_1 + e_2 \\ \varepsilon_3 & = e_3 - e_2 - 3e_1 \end{cases} \Leftrightarrow \begin{cases} e_1 & = \varepsilon_1 \\ e_2 & = \varepsilon_2 - \varepsilon_1 \\ e_3 & = \varepsilon_3 + \varepsilon_2 + 2\varepsilon_1 \end{cases}$$

Ainsi, de $X = e_1 + 2e_2 + 3e_3$, nous arrivons à $X = 5\varepsilon_1 + 5\varepsilon_2 + 3\varepsilon_3$

En utilisant la formule de changement de bases : Nous avons $X_{\mathscr{B}_c} = PX_{\mathscr{B}'}$, d'où $X_{\mathscr{B}'} = P^{-1}X_{\mathscr{B}_c}$. Il faut donc calculer l'inverse de $P = \begin{pmatrix} 1 & 1 & -3 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & -1 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$. Et on retrouve $X_{\mathscr{B}'} = \begin{pmatrix} 5 \\ 5 \\ 3 \end{pmatrix}$.

Remarque 14.9. Attention à l'ordre des bases et des vecteurs. Il serait plus naturel d'écrire $X' = PX$, ce qui est une erreur. P est la matrice de passage de $\mathscr{B}$ à $\mathscr{B}'$ et pourtant on calcule les coordonnées de $\overrightarrow{x}$ dans la base $\mathscr{B}$ à partir de celles de $\overrightarrow{x}$ dans la base $\mathscr{B}'$.

14.2.2. Matrices équivalentes

Définition 14.5 (Matrices équivalentes). Soit $(n;p) \in (\mathbb{N}^*)^2$ et deux matrices A, B appartenant à $\mathscr{M}_{n,p}(\mathbb{K})$. On dit que B est équivalente à A s'il existe deux matrices carrées inversibles $Q \in GL_n(\mathbb{K})$, $P \in GL_p(\mathbb{K})$ telles que :

$$B = Q^{-1}AP$$

Propriété 14.10. La relation « est équivalente à » sur $\mathscr{M}_{n,p}(\mathbb{K})$ est une relation d'équivalence. On dit alors simplement « A et B sont équivalentes » plutôt que « B est équivalente à A ».

Démonstration. **Réflexivité :** Nous avons simplement $A = I_n^{-1}AI_p$.

Symétrie : Nous avons $B = Q^{-1}AP$, donc $A = QBP^{-1}$. En posant $Q' = Q^{-1}$ et $P' = P^{-1}$, on a bien $A = Q'^{-1}BP'$.

Transitivité : Nous avons $C = Q^{-1}BP$ et $B = Q'^{-1}AP'$. En injectant dans l'expression de C, on obtient $C = Q^{-1}\left(Q'^{-1}AP'\right)P = Q^{-1}Q'^{-1}AP'P = (Q'Q)^{-1}A(P'P)$. En posant $Q'' = Q'Q$ et $P'' = P'P$ nous obtenons $C = Q''^{-1}AP''$.

Propriété 14.11 (Matrices de passage et matrices équivalentes). Soit E et F deux espaces vectoriels de dimensions finies, f une application linéaire de E dans F, $\mathscr{B}_1$ et $\mathscr{B}_1'$ deux bases de E, et $\mathscr{B}_2$ et $\mathscr{B}_2'$ deux bases de F. On note :

$$P = P_{\mathscr{B}_1}(\mathscr{B}_1') \qquad \text{et} \qquad Q = Q_{\mathscr{B}_2}(\mathscr{B}_2')$$

Soit A la matrice de f relativement aux bases $\mathscr{B}_1$ et $\mathscr{B}_2$, et B la matrice de f relativement aux bases $\mathscr{B}_1'$ et $\mathscr{B}_2'$. On a alors :

$$B = Q^{-1}AP$$

Démonstration. Soit $\overrightarrow{x}$ un vecteur de E. Notons X et X' les matrices colonnes des coordonnées de $\overrightarrow{x}$ dans les bases $\mathscr{B}_1$ et $\mathscr{B}_1'$. Notons aussi Y et Y' les matrices colonnes des coordonnées de $f(\overrightarrow{x})$ dans les bases $\mathscr{B}_2$ et $\mathscr{B}_2'$.

D'après la propriété 14.9, on a :

$$X = PX' \qquad \text{et} \qquad Y = QY'$$

Ou encore, puisque P est inversible, $X' = P^{-1}X$.

De plus, par définition de A et B, on a

$$Y = AX \qquad \text{et} \qquad Y' = BX'$$

Ainsi, $AX = Y = QY' = Q(BX') = QB\left(P^{-1}X\right)$. Puisque cette égalité est vraie pour tout $\overrightarrow{x}$, on peut identifier : $A = QBP^{-1}$, ou encore $B = Q^{-1}AP$. □

Remarque 14.10.

— Cette propriété nous dit que les matrices d'une même application dans des bases différentes sont des matrices équivalentes.

— Pour ne pas se perdre, il est recommandé de faire ce que l'on appel un diagramme commutatif :

$$\begin{tikzcd} E,\mathscr{B}_1' \arrow[r, "Id_E,P"] \arrow[d, "f,B"'] & E,\mathscr{B}_1 \arrow[d, "f,A"] \\ F,\mathscr{B}_2' \arrow[r, "Id_F,Q"'] & F,\mathscr{B}_2 \end{tikzcd}$$

On voit alors que $f_B = Id_F^{-1} \circ f_A \circ Id_E$, d'où l'égalité sur les matrices.

Propriété 14.12. Soit $A, B \in \mathscr{M}_{n,p}(\mathbb{K})$. Alors A et B sont équivalentes si et seulement si A et B sont les deux matrices d'une même applications linéaires dans deux bases différentes.

Démonstration. $\boxed{\Leftarrow}$ C'est la propriété 14.11.

$\boxed{\Rightarrow}$ Les matrices A et B étant équivalentes, il existe P et Q inversibles telles que $B = Q^{-1}Ap$. D'après la propriété 14.6, P et Q sont des matrices de passage de $\mathscr{B}_1$ à $\mathscr{B}_1'$ dans E (espace de dimension n) et de $\mathscr{B}_2$ à $\mathscr{B}_2'$ dans F (espace de dimension p).

Soit f l'application telle que $\mathrm{Mat}_{\mathscr{B}_1,\mathscr{B}_2}(f) = A$ et g celle telle que $\mathrm{Mat}_{\mathscr{B}_1',\mathscr{B}_2'}(g) = B$. Soit $x \in E$, comme pour la propriété 14.11, on pose X les coordonnées de x dans $\mathscr{B}_1$ et X' celles dans $\mathscr{B}_1'$. D'où $X = PX'$. Soit $y = g(x)$ et posons Y et Y' les coordonnées de y dans les bases $\mathscr{B}_2$ et $\mathscr{B}_2'$. On a alors $Y' = BX' = Q^{-1}APX' = Q^{-1}AX$. D'où $g(x) = Y = QY' = AX = f(x)$. Ainsi $f = g$.

Théorème 14.3 (Matrice J_r). Soit E et F deux $\mathbb{K}$-espaces vectoriels de dimensions respectives p et n, et $f \in \mathscr{L}(E;F)$ de rang r. Alors il existe une base $\mathscr{B}$ de E et une base $\mathscr{C}$ de F telles que

$$\mathrm{Mat}_{\mathscr{B},\mathscr{C}}(f) = J_r = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$$

Où J_r est une matrice de taille $n \times p$.

Démonstration. Supposons que $\mathscr{B} = (e_j)_{1\le j\le p}$ et $\mathscr{C} = (f_i)_{1\le i\le n}$ soient deux bases qui conviennent et notons $M = \mathrm{Mat}_{\mathscr{B},\mathscr{C}}(f)$. Alors

— Pour que les $p-r$ dernières colonnes de M soient nulles, il faut et il suffit que les vecteurs $e_{r+1}, \dots, e_p$ soient éléments de $\operatorname{Ker} f$ et linéairement indépendants. Comme $\operatorname{Ker} f$ est de dimension $p-r$, d'après le théorème du rang, nous n'avons qu'à choisir pour famille $(e_j)_{r+1\le j\le p}$ une base de $\operatorname{Ker} f$ et la compléter simplement en une base $\mathscr{B}$ de E.

— Pour que les r première colonnes de M soient ce que l'on veut, on peut poser : $f_j = f(e_j)$ pour tout $j \in [\![1;r]\!]$, puis compléter en une base $\mathscr{C}$ de F. Or ceci n'est possible qu'à condition que la famille $(f_i)_{1\le j\le r}$ soit libre. Or la famille $(e_i)_{1\le j\le r}$ engendre par construction un supplémentaire I de $\operatorname{Ker} f$ dans E, donc $f_{|_I}$ est un isomorphisme de I sur $\operatorname{Im} f$ (noyau réduit au vecteur nul donc injective). Donc la famille $(f_i)_{1\le j\le r} = \left(f(e_j)\right)_{1\le j\le r}$ est libre.

Nous avons donc bien réussi à construire deux bases $\mathscr{B}$ et $\mathscr{C}$ telles que $M = J_r$. □

Propriété 14.13 (Invariance du rang). Soit A et B deux matrices de $\mathscr{M}_{n,p}(\mathbb{K})$. Les matrices A et B sont équivalentes si et seulement si $\mathrm{rg}(A) = \mathrm{rg}(B)$.
Cela implique que toute matrice de $\mathscr{M}_{n,p}(\mathbb{K})$ de rang r est équivalente à J_r.

Démonstration. Soit $A \in \mathscr{M}_{n,p}(\mathbb{K})$ de rang r. Notons f l'application linéaire canoniquement associée à A et $\mathscr{B}_n$ et $\mathscr{B}_p$ les bases canoniques respectives de $\mathbb{K}^n$ et $\mathbb{K}^p$. D'après le théorème 14.3, il existe deux bases $\mathscr{B}$ de $\mathbb{K}^p$ et $\mathscr{C}$ de $\mathbb{K}^n$ telles que $\mathrm{Mat}_{\mathscr{B},\mathscr{C}}(f) = J_r$. Ainsi, en posant $P = P_{\mathscr{B}_p}(\mathscr{B})$ et $Q = P_{\mathscr{B}_n}(\mathscr{C})$, alors $J_r = Q^{-1}AP$. Ainsi A et J_r sont équivalentes.
On a ainsi montré que toute matrice de $\mathscr{M}_{n,p}(\mathbb{K})$ de rang r est équivalente à J_r, on en déduit, par transitivité, que deux matrices de $\mathscr{M}_{n,p}(\mathbb{K})$ ayant le même rang sont équivalentes.

Réciproquement, nous avons vu (propriété 14.12) que deux matrices équivalentes étaient les matrices d'une même application linéaire dans des bases différentes. Le rang d'une application linéaire étant indépendant des bases dans lesquelles ont travaille, on en déduit, grâce à la propriété 14.2 que les deux matrices ont le même rang. □

Remarque 14.11. C'est surtout le sens indirect qui va nous servir : pour montrer que deux matrices sont équivalentes, il suffit de montrer qu'elles ont le même rang (donc pas besoin de chercher les bases et les matrices de passage correspondantes).

Remarque 14.12. Deux matrices carrées équivalentes ont le même rang, cependant elles n'ont pas forcément le même déterminant et la même trace. Nous allons maintenant voir une notion plus forte que l'équivalence qui conserve le déterminant et la trace.

14.2.3. Matrices semblables

Dans cette section nous allons étudier le cas $F = E$, donc on ne s'intéresse qu'aux endomorphismes f de E dans E. Les matrices sont donc toutes carrées. De plus les matrices de f seront toujours relatives à la même base « au départ » et « à l'arrivée ».

Définition 14.6 (Matrices semblables). Soit A et B deux matrices carrées de $\mathscr{M}_n(\mathbb{K})$. On dit que B est *semblable* à A s'il existe une matrice inversible $P \in GL_n(\mathbb{K})$ telle que

$$B = P^{-1}AP$$

Remarque 14.13. On remarque que les matrices semblables ne sont que des cas particuliers des matrices équivalentes où $Q = P$.

Propriété 14.14. La relation « est semblable à » sur $\mathscr{M}_n(\mathbb{K})$ est une relation d'équivalence. On dit alors simplement « A et B sont semblables » plutôt que « B est semblable à A ».

Démonstration. **Réflexivité :** Nous avons simplement $A = I_n^{-1}AI_n$.
Symétrie : Nous avons $B = P^{-1}AP$, donc $A = PBP^{-1}$. En posant $P' = P^{-1}$, on a bien $A = P'^{-1}BP'$.

Transitivité : Nous avons $C = P^{-1}BP$ et $B = P'^{-1}AP'$. En injectant dans l'expression de C, on obtient $C = P^{-1}\left(P'^{-1}AP'\right)P = P^{-1}P'^{-1}AP'P = (P'P)^{-1}A\,(P'P)$. En posant $P'' = P'P$ nous obtenons $C = P''^{-1}AP''$.

Exemple 14.5. I_n est la seule matrice semblable à I_n alors que toute matrice inversible est équivalente à I_n.

Propriété 14.15. Si deux matrices A et B sont semblables alors elles sont équivalentes (semblables $\Rightarrow$ équivalentes).

Démonstration. Comme nous l'avons remarqué plus haut, il suffit de poser $Q = P$. □

Remarque 14.14. Attention ! ! La réciproque est fausse. Comme indiqué dans l'exemple précédent, n'importe quelle matrice inversible est équivalente à l'identité. Or seule l'identité est semblable à l'identité.

Propriété 14.16 (Matrices de passage et matrices semblables). Soit E un espace vectoriel de dimension finie, f une application linéaire de E dans E, $\mathscr{B}$ et $\mathscr{B}'$ deux bases de E. On note :

$$P = P_{\mathscr{B}}\left(\mathscr{B}'\right)$$

Soit A la matrice de f relativement à la base $\mathscr{B}$ et B la matrice de f relativement à la base $\mathscr{B}'$. On a alors :

$$B = P^{-1}AP$$

Démonstration. Nous avons vu que cette propriété n'est qu'un cas particulier de la propriété 14.11. Il suffit donc de reprendre la démonstration en supprimant simplement les indices sur les bases. □

Remarque 14.15. Cette propriété nous dit que les matrices carrées d'une même application dans des bases différentes sont des matrices semblables.

Propriété 14.17. Soit $A, B \in \mathscr{M}_n(\mathbb{K})$. Alors A et B sont semblables si et seulement si A et B sont les deux matrices d'une même application linéaire dans deux bases différentes.

Démonstration. On reprend là encore la démonstration de la propriété 14.12 □

Propriété 14.18 (Invariance du déterminant et de la trace). Deux matrices carrées semblables ont le même déterminant et la même trace : Pour toute matrice carrée A et toute matrice carrée inversible P, on a :

$$\det\left(P^{-1}AP\right) = \det(A)$$
$$\operatorname{tr}\left(P^{-1}AP\right) = \operatorname{tr}(A)$$

Démonstration.

Déterminant : On sait que pour toute matrice carrée, $\det(AB) = \det(A)\det(B)$. De plus, puisque P est inversible, $\det\left(P^{-1}\right) = \frac{1}{\det(P)}$.

Trace : Nous allons utiliser la propriété 12.3 du chapitre « Matrices » : $\mathrm{tr}(AB) = \mathrm{tr}(BA)$.
Ainsi, $\mathrm{tr}\left(P^{-1}AP\right) = \mathrm{tr}\left(P^{-1}(AP)\right) = \mathrm{tr}\left((AP)P^{-1}\right) = \mathrm{tr}\left(APP^{-1}\right) = \mathrm{tr}(A)$.

Remarque 14.16. Attention ! Cette fois, la réciproque est fausse : Deux matrices ayant même déterminant et même trace ne sont pas forcément semblables.

Remarque 14.17. Les traces de deux matrices semblables étant égales, la trace est invariante par changement de base. Cela nous permet de justifier la définition suivante.

Définition 14.7 (Trace d'un endomorphisme). Soit $E \neq \{0_E\}$ un $\mathbb{K}$-espace vectoriel de dimension finie et $f \in \mathscr{L}(E)$. On appelle *trace de* f et l'on note $\mathrm{tr}(f)$ la trace d'une matrice de f dans une base de E.

Propriété 14.19 (trace). Pour tout $f, g \in \mathscr{L}(E)$,

— Pour tout $\lambda \in \mathbb{K}$, $\mathrm{tr}(\lambda f + g) = \lambda \mathrm{tr}(f) + \mathrm{tr}(g)$.

— $\mathrm{tr}(f \circ g) = \mathrm{tr}(g \circ f)$

Démonstration. Il suffit de revenir aux matrices et utiliser les propriétés précédentes. □

Remarque 14.18. Alors qu'il est simple de montrer que deux matrices sont équivalentes (en calculant leurs rangs), il est plus difficile de montrer que deux matrices sont semblables.

Exemple 14.6. Soit $A = \begin{pmatrix} 1 & -2 \\ 2 & 3 \end{pmatrix}$ et $B = \begin{pmatrix} 11 & 14 \\ -6 & -7 \end{pmatrix}$. Montrer que A et B sont semblables et déterminer toutes les matrices P inversibles telles que $B = P^{-1}AP$.

D'abord on vérifie qu'elles « peuvent » être semblables en calculent leur rang (elles doivent être équivalentes), leur déterminant et leur trace :

$$\begin{aligned} \mathrm{rg}(A) &= 2 \\ \det(A) &= 3+4 = 7 \\ \mathrm{tr}(A) &= 1+3 = 4 \end{aligned} \qquad \begin{aligned} \mathrm{rg}(B) &= 2 \\ \det(B) &= -77+84 = 7 \\ \mathrm{tr}(B) &= 11-7 = 4 \end{aligned}$$

Nous n'avons pas prouvé qu'elles étaient semblables, mais il n'est pas impossible qu'elles le soient.

Posons maintenant $\mathscr{B} = (e_1 ; e_2)$ une base de $\mathbb{R}^2$. Notons f l'application telle que A soit la matrice de f dans la base $\mathscr{B}$. Nous cherchons une base $\mathscr{B}'(\varepsilon_1 ; \varepsilon_2)$ telle que B soit la matrice de f dans cette dernière base. Nous cherchons alors la matrice de passage P de $\mathscr{B}$ vers $\mathscr{B}'$. Notons $P = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$.

Nous avons $\begin{cases} f(e_1) &= e_1 + 2e_2 \\ f(e_2) &= -2e_1 + 3e_2 \end{cases}$ et $\begin{cases} \varepsilon_1 &= ae_1 + ce_2 \\ \varepsilon_2 &= be_1 + de_2 \end{cases}$.

Ainsi, $f(\varepsilon_1) = 11\varepsilon_1 - 6\varepsilon_2 = 11(ae_1 + ce_2) - 6(be_1 + 2e_2) = (11a - 6b)e_1 + (11c - 6d)e_2$. Mais nous pouvons aussi le calculer en utilisant la linéarité de f : $f(\varepsilon_1) = af(e_1) + cf(e_2) = a(e_1 + 2e_2) + c(-2e_1 + 3e_2) = (a - 2c)e_1 + (2a + 3c)e_2$.

Puisque $(e_1\,; e_2)$ est une base, il y a unicité de la décomposition. Nous avons donc $11a - 6b = a - 2c$ et $11c - 6d = 2a + 3c$. En raisonnant de même avec $f(\varepsilon_2)$, nous avons le système :

$$\begin{cases} 11a - 6b = a - 2c \\ 11c - 6d = 2a + 3c \\ 14a - 7b = b - 2d \\ 14c - 7d = 2b + 3d \end{cases}$$

Je vous ai détaillé le raisonnement pour que vous compreniez les mécanismes qui entrent en jeu. Mais en fait ce système peut être trouvé de manière plus rapide et directe, il s'obtient en résolvant simplement l'équation

$$PB = AP$$

Lorsque l'une des matrices possède beaucoup de zéros, il est parfois plus simple de poser des équations du type $f(\varepsilon_i) = \sum_k a_k e_k$.

Après résolution (par pivot de Gauss par exemple) nous trouvons :

$$\mathscr{S} = \left\{(4c - 3d\,; 7c - 5d\,; c\,; d) \ / \ (c\,; d) \in \mathbb{R}^2\right\}$$

Le système possède une infinité de solutions. Une seule suffit pour conclure que A et B sont semblables, à condition que cela donne une matrice (de passage) inversible.

On peut prendre $c = 0$ et $d = 1$: $P = \begin{pmatrix} -3 & -5 \\ 0 & 1 \end{pmatrix}$ avec $\det(P) = -3 \neq 0$ et $P^{-1} = -\frac{1}{3}\begin{pmatrix} 1 & 5 \\ 0 & -3 \end{pmatrix}$. Le calcul de $P^{-1}AP$ (que je vous laisse le soin de faire) donne bien B.

L'énoncer nous demande toutes les matrices P inversibles. P est nécessairement de la forme :

$$P = \begin{pmatrix} 4c - 3d & 7c - 5d \\ c & d \end{pmatrix}$$

Avec $(c\,; d) \in \mathbb{R}^2$ tel que $\det(P) = d(4c - 3d) - c(7c - 5d) = -3d^2 + 9cd - 7c^2 = -\frac{1}{28}\left((14c - 9d)^2 + 3d^2\right) \neq 0$. Ce déterminant s'annule donc pour $d = 0$ et $14c = 9d$, c'est-à-dire, $c = d = 0$. Donc $(c\,; d) \in \mathbb{R}^2 \setminus \{(0\,; 0)\}$.

Bibliographie

[1] Patrick BERGEON. *Cours de TSI*. 2015. URL : http://patrick.bergeon.perso.neuf.fr/Sup_TSI/Cours.htm (visité le 01/06/2015).

[2] Christophe BERTAULT. *Cours de MPSI*. 2017. URL : http://christophebertault.fr (visité le 21/08/2017).

[3] Lycée CORBUSIER. *Cours de TSI*. 2015. URL : http://www.prepa-corbusier.fr (visité le 01/06/2015).

[4] Denis FELDMANN. *Mathématiques dans le supérieur*. 2015. URL : http://denisfeldmann.fr/maths.htm (visité le 01/06/2015).

[5] Jean-Michel FERRARD. *Cours de MPSI, PCSI et PSI**. 2017. URL : http://mathprepa.fr/ (visité le 21/08/2017).

[6] Patrick FRADIN. *Cours de MPSI*. 2017. URL : http://mpsi.tuxfamily.org/ (visité le 21/08/2017).

[7] Laurent GARCIN. *Cours de MPSI*. 2017. URL : http://lgarcin.github.io (visité le 21/08/2017).

[8] Daniel GUININ, François AUBONNET et Bernard JOPPIN. *Algèbre 1*. Précis de Mathématiques - Cours, Exercices résolus 1. Breal, 1994. ISBN : 978-2-7495-1234-1.

[9] Daniel GUININ, François AUBONNET et Bernard JOPPIN. *Analyse 1*. Précis de Mathématiques - Cours, Exercices résolus 3. Breal, 1994. ISBN : 978-2-7495-1234-1.

[10] Daniel GUININ, Elisabeth LADAME et Hervé VANDEVEN. *Tout-En-Un Mathématiques*. Les nouveaux Précis de Mathématiques. Breal, 2014. ISBN : 978-2-7495-3230-1.

[11] Éric LEHMAN. *2. analyse*. Mathématiques pour l'étudiant de première année. Belin, 1986. ISBN : 978-2-7495-1234-1.

[12] Emmanuel MORRAND. *Cours de TSI*. 2015. URL : http://www.emmanuelmorand.net/supTSI-1213/supTSI-1213.php (visité le 01/06/2015).

[13] Lycée Pierre-Paul RICQUET. *Cours de TSI*. 2015. URL : http://pedagogie.ac-toulouse.fr/lyc-riquet-saint-orens/TSI-Maths/1tsi/html/cours.php (visité le 01/06/2015).

[14] Jacques ROGNAUX. *Cours de TSI*. 2015. URL : http://www.tsisoa.com/spip/Cours-de-Jacques-Rogniaux.html (visité le 01/06/2015).

[15] Jean-Paul VINCENT. *Cours de MPSI*. 2017. URL : http://tourbillon.pagesperso-orange.fr (visité le 21/08/2017).

Index

Table des matières