 etienne.miquey[at]ens-lyon.fr

Espaces vectoriels

1 Motivation

Dans la vraie vie, on a généralement une représentation assez figée de l'espace, en 3 dimensions. Quand on doit le dessiner, on triche et on passe à deux, et certains physiciens fumeux vous diront que le temps en est une quatrième légitime. Toujours est-il que notre esprit est relativement à l'aise avec ce type de représentation. Par exemple, en physique, le jour où l'on vous a raconté qu'une force, c'était essentiellement un vecteur, ça ne vous a ni choqué, ni posé de gros problème de visualisation du concept, peut-être même bien au contraire. Le but des Espaces Vectoriels est justement de redonné ce type de structure à des ensembles qui à première vue n'ont pas la chance d'en avoir une naturelle. Si l'on en a le temps, on verra qu'il y a tout un tas d'autres trucs que l'on essaie de même de formaliser et de donner comme structure à des espaces.

2 Définition

Quand vous considérez un espace vectoriel, il faut bien distinguer les vecteurs d'une part, et les coefficients par lesquels on peut multiplier les vecteurs, que l'on nomme scalaires. Les vecteurs seront pris dans l'espace E, et les coefficients dans un corps $\mathbb{K}$. On parle de $\mathbb{K}$-espace vectoriel, je vous donne les vraies définitions, mais si ça vous complique la vie, vous pouvez remplacer $\mathbb{K}$ par $\mathbb{R}, \mathbb{Q}$ ou $\mathbb{C}$.

Il y a deux manières de voir la chose. Soit on prends la définition, et l'on essaie de comprendre ce qu'elle signifie, soit on essaie de voir quelles propriétés ont nos espaces vectoriels usuels, et on regarde ce qu'il manque. À vous de choisir votre méthode.

Définition 1. $(E, +, \cdot)$ est un $\mathbb{K}$-*espace vectoriel* si :

- $\forall x, x' \in E, x + x' \in E$ *(loi de composition interne)*
- $\exists 0_E$ *tel que* $\forall x \in E, x + 0_E = 0_E + x = x$ *(neutre)*
- $\forall x, y, z \in E, x + (y + z) = (x + y) + z$ *(associativité)*
- $\forall x \in E, \exists y, x + y = y + x = 0_E$ *(inverse)*

- $\forall \lambda \in \mathbb{K}, x \in E, \lambda \cdot x \in E$
- $\forall x \in E, 1_{\mathbb{K}} \cdot x = x$
- $\forall \lambda \in \mathbb{K}, x, y \in E, \lambda \cdot (x + y) = \lambda \cdot x + \lambda \cdot y$
- $\forall \lambda, \mu \in \mathbb{K}, x \in E, (\lambda + \mu) \cdot x = \lambda \cdot x + \mu \cdot x$
- $\forall \lambda, \mu \in \mathbb{K}, x \in E, \lambda \cdot (\mu \cdot x) = (\lambda\mu) \cdot x$

Exemples : Vous pouvez ainsi vérifier que $\mathbb{R}, \mathbb{R}^2, \mathbb{R}^3, \mathbb{R}^n = \{(x_1, \ldots, x_n), \forall i, x_i \in \mathbb{R}\}$ sont bien des espaces vectoriels. Mais on peut aussi montrer que l'ensemble des polynômes (noté $\mathbb{K}[X]$), l'ensemble des fonctions de $\mathcal{I}$ dans $\mathbb{K}$ (noté $\mathcal{F}(\mathcal{I}, \mathbb{K})$) ou l'ensemble des suites d'un espace vectoriel E (noté $E^{\mathbb{N}}$), munis des règles intuitives d'addition et de multiplication par scalaire, sont bien des espaces vectoriels. En revanche, $\mathbb{Z}^2$ n'en est pas un, si l'on prends $\mathbb{K} = \mathbb{Q}$ ou $\mathbb{R}$ par exemple, on n'a pas la règle $\forall \lambda \in \mathbb{K}, x \in \mathbb{Z}^2, \lambda \cdot x \in \mathbb{Z}^2$. Et de façon générale, on ne peut construire de corps $\mathbb{K}$ tel que $\mathbb{Z}^n$ soit un espace vectoriel.

On se rend rapidement compte que montrer qu'un ensemble est bien un espace vectoriel, avec cette définition, est une tâche pénible et peu enrichissante. On a en fait un moyen un petit peu plus simple de le faire, en se placant dans un espace vectoriel plus gros. On hérite alors naturellement des propriétés de bon comportement des opérations les unes relativement aux autres, il n'y a plus qu'à montrer les stabilités, ce qui est beaucoup plus rapide. C'est la notion de *sous-espace vectoriel*[1].

Définition 2. Soit E un $\mathbb{K}$-espace vectoriel, et $F \subset E$. On dit que F est un *sous-espace vectoriel* de E si :
- $0_E \in F$
- F est stable par $+$: $\forall x, x' \in F, x + x' \in F$
- F est stable par $\cdot$: $\forall x \in F, \lambda \in \mathbb{K}, \lambda \cdot x \in F$

Théorème 3. *Soit $F \subset E$. Alors F est un sous-espace vectoriel de E si et seulement si F est un espace vectoriel.*

Là, on commence à voir un petit peu mieux à quoi ça ressemble, et à s'imaginer quelle tête peut avoir un (sous-)espace vectoriel.

1. *Dès que l'on fait un peu d'algèbre, en réalité on passe son temps à montrer que l'on a des sous-machins, parce que ça va toujours plus vite.*

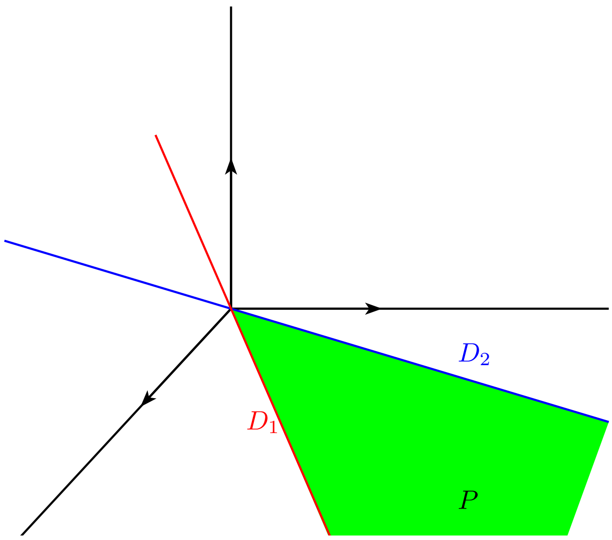


Figure 1 – Sous-espaces

Sur les exemples,
- D_1 et D_2 sont des sous-espaces vectoriels de $\mathbb{R}^3$
- $D_1 \cup D_2$ n'est pas un espace-vectoriel (non stable par +)
- P n'est pas non plus un sous-espace vectoriel (non stable par +)

Très facilement, on peut en fait montrer que tout sous-espace F contenant D_1 et D_2 contient nécessairement tout le plan horizontal[2].

Exemples : En se servant des gros espaces vectoriels englobant de tout à l'heure, on a par exemple :
- $\mathbb{K}_n[X] = \{P \in \mathbb{K}[X] \ / \ deg(P) \leq n\}$ qui est un sous-espace vectoriel de $\mathbb{K}[X]$[3]
- $\mathcal{C}^0(\mathcal{I}, \mathbb{K}) = \{f \in \mathcal{F}(\mathcal{I}, \mathbb{K}), f \ continue\}$ qui est un sous-espace vectoriel de $\mathcal{F}(\mathcal{I}, \mathbb{K})$
- l'ensemble des suites de E qui convergent ($\mathcal{L} = \{(u_n)_{n\in\mathbb{N}} \in E^{\mathbb{N}} \ / \ \exists l \in E, u_n \underset{n\to+\infty}{\to} l\}$) est un sous-espace de l'ensemble des suites de E

Pour vérifier que vous avez compris, vous pouvez essayer de montrer que :
- les fonctions paires[4] (resp. impaires) forment un espace-vectoriel.
- $\{f : \mathbb{R} \to \mathbb{R}, f(0) = 0\}$ et $\{f : \mathbb{R} \to \mathbb{R}, f(0) \geq 0\}$ sont des sous-espaces vectoriels de $\mathcal{F}(\mathbb{R}, \mathbb{R})$, mais que $\{f : \mathbb{R} \to \mathbb{R}, f(0) = 1\}$ et $\{f : \mathbb{R} \to \mathbb{R}, f(0) > 0\}$ ne le sont pas.
- l'ensemble des suites de E qui convergent vers 0 ($\{(u_n)_{n\in\mathbb{N}} \in E^{\mathbb{N}} \ / \ \in E, u_n \underset{n\to+\infty}{\to} 0\}$ forme un sous-espace vectoriel de $\mathcal{L}$, mais pas celles convergeant vers 1. Et même, que $\mathcal{L}_l = \{(u_n)_{n\in\mathbb{N}} \in E^{\mathbb{N}} \ / \ \in E, u_n \underset{n\to+\infty}{\to} l\}$ est un espace vectoriel si et seulement si $l = 0$.

3 Fabrication

L'objectif de cette partie est de regarder comment la structure d'espace vectoriel se comporte vis-à-vis des opérateurs ensemblistes, et de voir après comment construire de tels espaces.

Tout d'abord, on a remarqué tout à l'heure (Figure 1) que si E et F étaient des espaces vectoriels, il n'en était pas forcément de même pour $E \cup F$. En revanche, on a facilement le résultat suivant :

Propriété 4. *Si E et F sont des espaces-vectoriels, alors $E \cap F$ est un espace-vectoriel.*

Démonstration. On a $E \cap F \in E$ et $E \cap F \in F$. Toutes les propriétés de sous-espace vectoriel sont vraies dans E et F, elles le sont donc a fortiori dans $E \cap F$. $\square$

Pour pouvoir quand même travailler avec l'idée intuitive d'union, on va définir la notion d'espace vectoriel engendré par des parties $X_1, \ldots X_n$ de E. Une première façon de le faire, en s'inspirant de l'exemple de la section précédente est de dire que $Vect(X_1, X_2 \ldots X_n)$ est le plus petit espace vectoriel contenant $X_1 \cup \ldots \cup X_n$: $Vect(X_1, X_2 \ldots X_n) = \bigcap_{F \ ss-ev, \ F \supset X_1 \cup \ldots \cup X_n} F$. En tant qu'intersection d'espace vectoriel, c'est bien lui même

2. *Car on peut en construire les vecteurs de bases à partir de ceux de D_1 et D_2*

3. *Ce n'est pas vrai pour $\{P \in \mathbb{K}[X] \ / \ deg(P) \geq n\}$, puisque par exemple si $P = X^2 + 1$ et $Q = X^2 - 3$, $deg(P) = deg(Q) \geq 2$, mais $deg(P + Q) < 2$*

4. $\forall x, f(x) = f(-x)$

un espace-vectoriel. Cependant, cette définition n'est pas très commode pour travailler ensuite, on ne connaît pas forcément de tels F, et certainement pas tous. On a la définition équivalente[5] suivante :

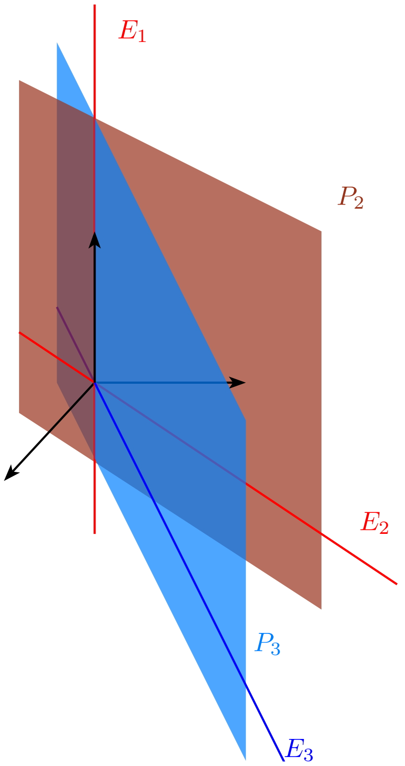
Définition 5. $Vect(X) = \{\sum_i \lambda_i x_i,\ x_i \in X, \lambda_i \in \mathbb{K})\}$

Avec des mots, cela revient juste à dire que c'est l'ensemble des combinaisons linéaires d'éléments de X. Par exemple, si $x_1, x_2 \in \mathbb{R}^2$, alors :

$$Vect(x_1, x_2) = \begin{cases} D \text{ la droite engendrée par } x_1 \text{ et } x_2 \text{ si les deux vecteurs sont colinéaires} \\ \mathbb{R}^2 \text{ sinon} \end{cases}$$

En s'appuyant sur cette idée, on définit la notion de somme[6] d'espace vectoriel :

Définition 6. Si E et F sont deux espaces vectoriels, on pose $E + F = Vect(E \cup F) = \{e + f, e \in E, f \in F\}$. On parle de somme directe, et on note $E \oplus F$, si $E \cap F = \{0\}$.



$$\text{On a : } \begin{cases} E_1 \oplus E_2 = P_2 \\ E_1 \oplus E_3 = P_3 \\ P_1 + P_2 = \mathbb{R}^3 \\ P_1 \oplus E_3 = \mathbb{R}^3 \end{cases}$$

FIGURE 2 – Somme d'espaces vectoriels

Pendant qu'on y est, profitons-en pour introduire une petite notion supplémentaire.

Définition 7. Soit E un espace vectoriel, et $F \subset E$ un sous-espace vectoriel. On dit que F' est un *supplémentaire* de F si $F \oplus F' = E$.

Un exemple intéressant est celui des fonctions paires et impaires. En notant $\mathcal{P}$ les fonctions paires de $\mathbb{R}$, et $\mathcal{I}$ les fonctions impaires, on a :

$$\mathcal{F}(\mathbb{R}, \mathbb{R}) = \mathcal{P} \oplus \mathcal{I}$$

En effet, $\forall f \in \mathcal{F}(\mathbb{R}, \mathbb{R}), f_p : x \mapsto \frac{f(x)+f(-x)}{2} \in \mathcal{P}, f_i : x \mapsto \frac{f(x)-f(-x)}{2} \in \mathcal{I}$ et $f = f_p + f_i$. D'où $\mathcal{P} + \mathcal{I} = \mathcal{F}(\mathbb{R}, \mathbb{R})$. De plus, si $f \in \mathcal{P} \cap \mathcal{I}$, alors $\forall x, f(x) = f(-x) = -f(-x) = 0$. D'où $\mathcal{P} \cap \mathcal{I} = \{0\}$.

À titre culturel, je vous donne le résultat suivant, qui confirme un petit peu l'intuition qu'on pourrait avoir. Vous pourrez très facilement le démontrer en dimension finie dès que l'on aura vu ce qu'est une base.

Théorème 8. *Tout sous-espace vectoriel admet un supplémentaire.*

Fin de l'interlude.

La but initial de cette partie était de s'intéresser à la construction d'espaces vectoriels. On peut désormais le faire à l'aide de sommes de sous-espaces, mais aussi en prenant l'espace engendré par une famille idoine.

5. *On montre facilement que l'on définit bien ainsi un espace-vectoriel, et que celui-ci contient la grosse intersection précédente et réciproquement.*

6. *Qui a donc plus de sens que celle d'union*

Définition 9. Une famille $(x_1, \ldots, x_n)$ d'un espace E est dite *génératrice* si $Vect(x_1, \ldots, x_n) = E$.

À l'inverse, on aimerait (en tant que matheux, donc flemmards), pouvoir le faire avec une famille aussi petite que possible. Par exemple, avec $x_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, x_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, x_3 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, on a $Vect(x_1, x_2, x_3) = Vect(x_1, x_2) = Vect(x_1, x_3) = Vect(x_2, x_3)$. En effet, on ne gagne rien avec un troisième de ces vecteurs si l'on dispose de ces deux autres, puisque l'on peut déjà construire tout l'espace avec ces deux vecteurs. On dit ainsi qu'une famille est liée, si tous ses éléments ne sont pas indispensables. Sinon, à l'inverse, on dira qu'elle est libre.

Définition 10. Une famille $(x_1, \ldots, x_n)$ est dite *liée* si $\exists \lambda_1, \ldots, \lambda_n$, tels que $\sum\limits_{i=1}^{n} \lambda_i x_i = 0$ et $\exists i / \lambda_i = 0$. Une famille $(x_1, \ldots, x_n)$ est dite *libre* si $\forall \lambda_1, \ldots, \lambda_n, \sum\limits_{i=1}^{n} \lambda_i x_i = 0 \Rightarrow \forall i, \lambda_i = 0$

Vous pouvez maintenant commencer à pressentir que l'on a tout ce qu'il faut pour définir une base, c'est-à-dire une famille minimale qui engendre tout l'espace.

Définition 11. Une famille $(x_1, \ldots, x_n)$ est une *base* si elle est libre et génératrice

Et disposant de cette notion de base, on peut du même coup définir la notion de dimension.

Définition 12. On appelle *dimension* de E le cardinal d'une base de E (si tant est qu'il en existe une).

Il convient de noter ici qu'il est clair que si un espace admet deux bases, celles-ci ont même cardinal. Si celle-ci sont de cardinal infini, c'est évident, sinon, on montre facilement que la plus petit n'est pas génératrice ou que la plus grande est liée.

4 Applications linéaires

La dernière partie de ce chapitre concernent les fonctions entre espaces vectoriels. Il en existe tout un tas, mais il est assez clair que celle qui vont nous intéresser sont celles qui respectent la structure d'espace vectoriel. En effet, l'intérêt principal de cette structure est de pouvoir définir tout un espace avec seulement une base de celui-ci, il serait donc appréciable de pouvoir également définir une fonction par son image sur les éléments d'une base. Les fonctions possédant cette propriété sont appellées *applications linéaires.*

Définition 13. Soient E, F deux espaces vectoriels. On dit que $f : E \to F$ est une application linéaire si

$$\forall x, y \in E, \lambda, \mu \in \mathbb{K}, f(\lambda x + \mu y) = \lambda f(x) + \mu f(y)$$

Si l'on regarde ce que ça donne avec $E = F = \mathbb{R}$ (cf. feuille d'introduction), on retrouve bien les fonctions linéaires que l'on connait usuellement, dont le graphe est une droite passant par 0. Plus généralement, on a les propriétés immédiates suivantes :

Propriété 14. *Soient $f, g : E \to F$ des applications linéaires. Alors*
- $f(0_E) = 0_F$
- $f \circ g : x \mapsto f(g(x))$ *est une application linéaire*
- $f + g :: x \mapsto f(x) + g(x)$ *est une application linéaire*

Les exemples suivants montrent bien l'intérêt de toutes ces définitions, puisque la dérivation et l'intégration, par exemple, sont des exemples d'applications linéaires (on laissera au lecteur le soin de vérifier que ce sont effectivement des applications linéaires) : $D : \left\{ \begin{array}{l} \mathcal{C}^\infty \to \mathcal{C}^\infty \\ f \mapsto f' \end{array} \right.$ $\Bigg|$ $I_{ab} : \left\{ \begin{array}{l} \mathcal{C}^0 \to \mathcal{C}^0 \\ f \mapsto \int_a^b f(t)\, \mathrm{d}t \end{array} \right.$

Définition 15. Soit $f : E \to F$ une application linéaire. On appelle *noyau* de f l'ensemble $Ker(f) = \{x \in E \;/\; f(x) = 0\}$ et *image* de f l'ensemble $Im(f) = \{y \in F \;/\; \exists x, f(x) = y\}$.

La propriété suivante met en évidence le caractère fondamental des deux définitions précédentes, vous pouvez faire la preuve en exercice :

Propriété 16. *$Im(f)$ et $Ker(f)$ sont des sous-espaces vectoriels.*

Petite digression dont se délecteront les amateurs de physique. Si l'on s'intéresse à un système masse-ressort horizontal sans frottement, que l'on note $x(t)$ sa position horizontal, m sa masse, k la constante de raideur du ressort et l_0 sa longueur à l'origine, on obtient par application du 2^{nd} principe de Newton l'équation suivante :

$$m\ddot{x} + k(x - l_0) = 0$$

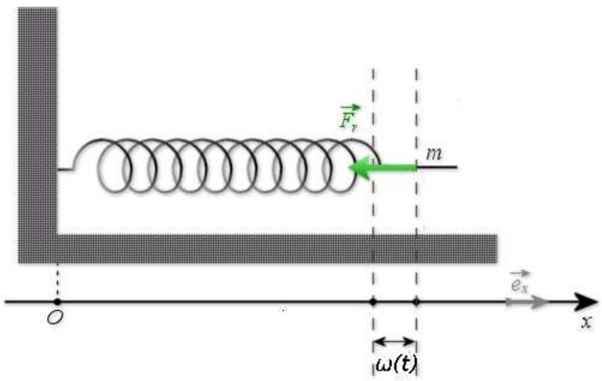


FIGURE 3 – Système masse-ressort horizontal

En posant, $\omega = x - l_0$ et $\alpha = \frac{k}{m}$, cette équation devient

$$\ddot{\omega} + \alpha\omega = 0$$

Revenons au point de vue du matheux. En posant, avec les notations précédentes, $R : x \mapsto D(D(x)) + \alpha x$, qui est une application linéaire, l'équation précédente devient donc :

$$R(\omega) = 0 \Leftrightarrow \omega \in Ker(R)$$

L'espace des solutions est donc un espace vectoriel, et c'est pour cela que vous ne cherchez qu'une base des solutions quand vous avez ce genre d'exercice. Si l'on rajoute une force constante, cela reviendra juste à considérer un espace affine, c'est-à-dire un espace vectoriel dont on a translaté l'origine. C'est pour cela que vous chercherez dans ce cas une solution particulière (la translation) et les solution de l'équation homogène. Maintenant vous savez ce que vous cachent vos profs de physique.

Les définitions suivantes sont peut-être plus ou moins à vos programmes, dans le doute, vous y avez droit.

Définition 17. Une fonction $f : E \to F$ est dite *injective* si $\forall x, y \in E,\ x \neq y \Rightarrow f(x) \neq f(y)$, ce qui est équivalent à $\forall x, y, f(x) = f(y) \Rightarrow x = y$ (tout élément de l'espace d'arrivée a au plus un antécédent par f). On dit que f est *surjective* si $\forall y \in F, \exists x \in E / f(x) = y$ (tout le monde a un antécédent).

f est dite *bijective* si elle est injective et surjective (tout le monde possède exactement un antécédent, il y a correspondance exacte entre les ensembles de départ et d'arrivée de f).

Ces notions sont relativement centrales dès lors que l'on fait des maths avec des fonctions. Vous aurez tôt ou tard l'occasion de vous en rendre compte. L'existence d'une application linéaire bijective entre deux espaces E et F nous indique par exemple qu'ils sont structurellement les mêmes (on dit alors qu'ils sont *isomorphes*), et l'on notera dans ce cas-là $E \simeq F$. On en a une caractérisation très simple avec les espaces vectoriels :

Propriété 18. *Soit $f : E \to F$ une application linéaire. Alors,*

1. *f injective $\Leftrightarrow Ker(f) = \{0\}$*
2. *f surjective $\Leftrightarrow Im(f) = F$*

Démonstration.

1. $$\begin{aligned} f \text{ injective} &\Leftrightarrow \forall x, y \in E, x \neq y \Rightarrow f(x) \neq f(y) \\ &\Leftrightarrow \forall x, y \in E, y - x \neq 0 \Rightarrow f(y) - f(x) = f(y - x) \neq 0 \\ &\Leftrightarrow \forall x \in E, x \neq 0 \Rightarrow f(x) \neq 0 \\ &\Leftrightarrow Ker(f) = \{0\} \end{aligned}$$
2. $$\begin{aligned} f \text{ surjective} &\Leftrightarrow \forall y \in F, \exists x \in E / f(x) = y \\ &\Leftrightarrow \forall y \in F, y \in Im(f) \\ &\Leftrightarrow Im(f) = F \end{aligned}$$

□

Le théorème suivant, qui sera le dernier du jour, est discutablement l'un des plus importants en algèbre linéaire. Je ne peux que vous inviter à faire de beaux dessins pour vous représenter le résultat.

Théorème 19 (du rang)**.** *Si* $f : E \to F$ *et* $S \oplus Ker(f)$*, alors* $S \simeq Im(f)$*.*

La preuve, en dimension quelconque est quelque peu compliquée et nécessiterait un peu trop de temps pour être faite ici. En revanche, en dimension finie, c'est presque trivial : on montre dans un premier temps que $f_{|S}$ (f restreinte au départ à S) est injective (même preuve que pour le noyau), puis on en déduit le résultat voulu. On a gratuitement, par simple application de ce résultat, le corollaire suivant :

Corollaire 20. *Si* $dim(E) = dim(F) < +\infty$ *et* $f : E \to F$*, alors :*

$$f \text{ injective} \Leftrightarrow f \text{ surjective} \Leftrightarrow f \text{ bijective}$$

Démonstration. Il suffit de voir que $dim(Im(f)) = dim(S) = dim(E) - dim(Ker(f))$, dont on déduit facilement $dim(Ker(f)) + dim(Im(f)) = dim(E)$. Puis on utilise le Lemme 18 pour conclure. □

Supposé qu'Euclide et ses prédécesseurs aient considéré le triangle
comme une moitié de carré ou, mieux, d'un parallélogramme :
ils auraient été immédiatement conduits au vecteur,
c'est-à-dire à la structure de l'espace comme espace vectoriel.
[Michel Serres, *Hermès I*]