

# Convexité

Extrait du programme officiel :

L'objectif de ce chapitre est double :

- introduire brièvement la notion de partie convexe d'un espace vectoriel réel ;
- étudier les fonctions convexes d'une variable réelle.

Le cours gagne à être illustré par de nombreuses figures.

La notion de barycentre est introduite exclusivement en vue de l'étude de la convexité.

CONTENUS

CAPACITÉS & COMMENTAIRES

## a) Parties convexes d'un espace vectoriel réel

Barycentre.

⇔ PC et SI : centre de masse (ou centre de gravité).

Partie convexe. Caractérisation à l'aide de barycentres à coefficients positifs.

## b) Fonctions convexes d'une variable réelle

Une fonction  $f$  est convexe sur l'intervalle  $I$  de  $\mathbb{R}$  si pour tout  $(x, y)$  de  $I^2$  et tout  $\lambda$  de  $[0, 1]$  :

$$f((1 - \lambda)x + \lambda y) \leq (1 - \lambda)f(x) + \lambda f(y).$$

Pour  $f$  convexe, les étudiants doivent connaître l'inégalité

$$f\left(\sum_{i=1}^n \lambda_i x_i\right) \leq \sum_{i=1}^n \lambda_i f(x_i)$$

où  $x_1, \dots, x_n$  sont des points de  $I$  et  $\lambda_1, \dots, \lambda_n$  des réels positifs de somme 1.

Caractérisations : convexité de l'épigraph, inégalité des pentes.

Position relative du graphe et de ses cordes.

Fonction concave.

## c) Fonctions convexes dérivables, deux fois dérivables

Caractérisation des fonctions convexes dérivables sur  $I$ , des fonctions convexes deux fois dérivables sur  $I$ .

Exemples d'inégalités de convexité.

Position relative du graphe d'une fonction convexe dérivable et de ses tangentes.

# TABLE DES MATIÈRES

## XII CONVEXITÉ

<b>I Barycentres et parties convexes d'un espace vectoriel</b>	<b>1</b>
1 Barycentre . . . . .	1
2 Parties convexes . . . . .	2
<b>II Fonctions convexes d'une variable réelle</b>	<b>3</b>
1 Définitions . . . . .	3
2 Caractérisations . . . . .	4
3 Cas des fonctions dérivables . . . . .	4
<b>III Inégalités de convexité</b>	<b>5</b>
<b>IV Que faut-il savoir faire ?</b>	<b>6</b>

Dans ce chapitre, le plus important, ce sont les dessins !

# I BARYCENTRES ET PARTIES CONVEXES D'UN ESPACE VECTORIEL

Soit  $E$  un espace vectoriel réel.



# 1 Barycentre

## Définition : Barycentre

Soit  $n \in \mathbb{N}^*$ ,  $(x_1, \dots, x_n) \in E^n$ ,  $(\lambda_1, \dots, \lambda_n) \in \mathbb{R}^n$  tel que  $\lambda = \sum_{i=1}^n \lambda_i \neq 0$ .

On appelle **barycentre** de  $((x_1, \lambda_1), \dots, (x_n, \lambda_n))$ , l'élément  $g$  de  $E$  défini par

$$g = \frac{1}{\lambda} \sum_{i=1}^n \lambda_i x_i.$$

On note  $g = \text{Bar}((x_1, \lambda_1), \dots, (x_n, \lambda_n))$ .

## Remarques

R1 – Cela s'interprète de façon affine : les vecteurs de  $E$  sont vus comme des points, et la relation se réécrit

$$\overrightarrow{OG} = \frac{1}{\lambda} \sum_{i=1}^n \lambda_i \overrightarrow{OA_i}$$

relation indépendante du point  $O$  : en utilisant la relation de Chasles  $\overrightarrow{O'G} = \overrightarrow{O'O} + \overrightarrow{OG}$  on vérifie qu'on a encore  $\overrightarrow{O'G} = \frac{1}{\lambda} \sum_{i=1}^n \lambda_i \overrightarrow{O'A_i}$ .

Il s'agit du point appelé **centre de masse** en sciences physiques, lorsque les poids  $\lambda_i$  sont des... masses.

On a donc une interprétation affine d'une combinaison linéaire de vecteurs dont la somme des coefficients est égale à 1.

Lorsque la somme des coefficients n'est pas égale à 1, la combinaison linéaire n'a pas d'interprétation affine (c'est-à-dire qu'elle ne s'interprète pas comme un point remarquable).

R2 – En posant pour tout  $i$ ,  $\mu_i = \frac{\lambda_i}{\lambda}$ , on obtient  $g = \sum_{i=1}^n \mu_i x_i$  avec  $\sum_{i=1}^n \mu_i = 1$ , et donc  $g = \text{Bar}((x_1, \mu_1), \dots, (x_n, \mu_n))$ .

Plus généralement, multiplier ou diviser tous les coefficients par un même réel non nul ne change pas le barycentre.

R3 – Lorsque tous les poids sont égaux, on parle d'**isobarycentre** :  $\text{Isobar}(x_1, \dots, x_n) = \frac{1}{n} \sum_{i=1}^n x_i = \text{Bar}((x_1, 1), \dots, (x_n, 1))$ .

R4 – Soit  $X$  une variable aléatoire finie prenant des valeurs  $x_1, \dots, x_n$  sur un univers  $\Omega$  muni d'une probabilité  $\mathbb{P}$  (programme de MPSI). Alors l'espérance de  $X$  est  $\mathbb{E}(X) = \sum_{i=1}^n \mathbb{P}(X = x_i) x_i$  avec  $\sum_{i=1}^n \mathbb{P}(X = x_i) = 1$  : il s'agit donc de  $\text{Bar}((x_i, \mathbb{P}(X = x_i)))_{1 \leq i \leq n}$ .

## Propriété : Associativité du barycentre

Soit  $n, p \in \mathbb{N}^*$ ,  $(x_1, \dots, x_n) \in E^n$ ,  $(\lambda_1, \dots, \lambda_n) \in \mathbb{R}^n$  tel que  $\lambda = \sum_{i=1}^n \lambda_i \neq 0$  et  $(y_1, \dots, y_p) \in E^p$ ,  $(\mu_1, \dots, \mu_p) \in \mathbb{R}^p$  tel que

$\mu = \sum_{i=1}^p \mu_i \neq 0$ . On suppose en outre que  $\lambda + \mu \neq 0$  et on pose

$$\bullet \quad x = \text{Bar}((x_1, \lambda_1), \dots, (x_n, \lambda_n)),$$

$$\bullet \quad y = \text{Bar}((y_1, \mu_1), \dots, (y_p, \mu_p)),$$

$$\bullet \quad g = \text{Bar}((x_1, \lambda_1), \dots, (x_n, \lambda_n), (y_1, \mu_1), \dots, (y_p, \mu_p)).$$

Alors

$$g = \text{Bar}((x, \lambda), (y, \mu)).$$

## Démonstration

Il suffit de l'écrire...



## 2 Parties convexes

### Définition : Segment

Si  $a$  et  $b$  sont deux éléments de  $E$ , on définit le segment  $[a, b]$

$$[a, b] = \{(1-t)a + tb ; t \in [0, 1]\}$$

### Remarques

- R1 – Segment dans  $E$  défini à partir d'un... segment dans  $\mathbb{R}$  ! Mais dans  $\mathbb{R}$ , le segment se définit directement par des inégalités.
- R2 –  $[a, b] = [b, a]$ .
- R3 – Il s'agit donc de l'ensemble des barycentres de  $a$  et  $b$  à coefficients positifs.

### Exercice

Soit  $A, B, C$  trois points du plan  $\mathbb{R}^2$ . Déterminer la position de  $G = \text{Bar}((A, \alpha), (B, \beta), (C, \gamma))$  où  $\alpha + \beta + \gamma = 1$  en fonction du signe de  $\alpha$ ,  $\beta$  et  $\gamma$ .

### Définition : Partie convexe

On dit qu'une partie  $C$  de  $E$  est convexe lorsque, pour tout couple  $(a, b)$  d'éléments de  $C$ , le segment  $[a, b]$  est inclus dans  $C$ , c'est-à-dire lorsque

$$\forall (a, b) \in C^2, \forall t \in [0, 1], (1-t)a + tb \in C$$

### Remarques

- R1 – On peut se contenter de  $t \in ]0, 1[$ ...
- R2 – Vu en MPSI : les intervalles sont les convexes de  $\mathbb{R}$ .

### Propriété : Caractérisation par les barycentres

$C$  est convexe si et seulement si elle est stable par barycentre à coefficients positifs, autrement dit si et seulement si pour tout  $n \geq 1$ , pour tout  $n$ -uplet  $(\lambda_1, \dots, \lambda_n)$  de réels vérifiant  $\lambda_1 \geq 0, \lambda_2 \geq 0, \dots, \lambda_n \geq 0$  et  $\lambda_1 + \lambda_2 + \dots + \lambda_n = 1$ , et pour tout  $n$ -uplet  $(x_i)_{1 \leq i \leq n}$  d'éléments de  $C$ ,  $\sum_{i=1}^n \lambda_i x_i \in C$ .

### Remarque

Géométriquement, cela signifie que si un polygone a ses sommets dans  $C$ , il est tout entier (arêtes, intérieur) inclus dans  $C$ .

## II FONCTIONS CONVEXES D'UNE VARIABLE RÉELLE

$I$  désigne un intervalle réel contenant au moins deux points.

### 1 Définitions

#### Définition : Fonction convexe, concave

- $f : I \rightarrow \mathbb{R}$  est dite **convexe** sur  $I$  lorsque

$$\forall x, y \in I, \forall t \in [0, 1], f((1-t)x + ty) \leq (1-t)f(x) + tf(y)$$

- $f$  est dite **concave** sur  $I$  lorsque  $-f$  est convexe.



### Remarques

- R1** – On peut se contenter de  $t \in ]0, 1[$  de nouveau.  
**R2** – Concave n'est pas le contraire de convexe !  
 On peut être les deux à la fois (quand ?) et on n'est en général ni l'un ni l'autre.

La définition donne immédiatement l'interprétation géométrique suivante :

#### Propriété : Caractérisation par interprétation géométrique

$f$  est convexe (respectivement concave) si et seulement si tout sous-arc de sa courbe est situé en dessous (respectivement au-dessus) de la corde correspondante.

### Remarque

Un point en lequel il y a un changement de concavité est appelé **point d'inflexion**.

#### Définition : Épigraphe

La partie  $\mathcal{E}(f)$  du plan située au dessus (au sens large) de la courbe de  $f$  s'appelle son **épigraphe**.

$$(x, y) \in \mathcal{E}(f) \iff y \geq f(x)$$

## 2 Caractérisations

Ces caractérisations sont à visualiser sur des dessins !

#### Propriété : Caractérisation avec l'épigraphe

$f$  est convexe sur  $I$  si et seulement si son épigraphe  $\mathcal{E}(f)$  est une partie convexe de  $\mathbb{R}^2$ .

#### Propriété : Caractérisations avec les pentes des cordes

Soit  $f : I \rightarrow \mathbb{R}$ . Les assertions suivantes sont équivalentes :

- (i)  $f$  est convexe sur  $I$ .  
 (ii) **Inégalité des trois cordes** : si  $x, y, z \in I$ ,  $x < y < z$ ,

$$\frac{f(y) - f(x)}{y - x} \leq \frac{f(z) - f(x)}{z - x} \leq \frac{f(z) - f(y)}{z - y}$$

- (iii) **Croissance des pentes** : pour tout  $a \in I$ ,

$$\tau_a : \begin{cases} I \setminus \{a\} & \longrightarrow \mathbb{R} \\ x & \longmapsto \frac{f(x) - f(a)}{x - a} \end{cases} \text{ est croissante}$$

### Démonstration

- (i)  $\Rightarrow$  (ii)  $y = (1-t)x + tz \dots$   
 (ii)  $\Rightarrow$  (iii) Séparer trois cas suivant la position par rapport à  $a$ .  
 (iii)  $\Rightarrow$  (i) Utiliser  $\tau_x$ .

□

### Exercice

Montrer qu'une fonction convexe sur  $I$  admet en chaque point de  $\overset{\circ}{I}$  une dérivée à gauche et une dérivée à droite et en déduire que  $f$  est continue sur  $\overset{\circ}{I}$ .

Est-ce le cas sur  $I$  ?

## 3 Cas des fonctions dérivables

**Propriété : Caractérisation avec la dérivée**

Soit  $f : I \rightarrow \mathbb{R}$  dérivable sur  $I$ .

$f$  est convexe si et seulement si  $f'$  est croissante sur  $I$ .

**Démonstration**

Si  $f$  est convexe, dans l'inégalité des trois cordes, faire tendre  $z$  vers  $x$  à gauche puis vers  $y$  à droite.

Si  $f'$  est croissante, dériver  $\varphi : z \mapsto f((1-t)z + ty) - (1-t)f(z) - tf(y)$  sur  $[x, y]$ . □

**Corollaire : Caractérisation avec la dérivée seconde**

Soit  $f : I \rightarrow \mathbb{R}$  deux fois dérivable sur  $I$ .

$f$  est convexe (respectivement concave) si et seulement si  $f'' \geq 0$  (respectivement  $f'' \leq 0$ ) sur  $I$ .

**Propriété : Caractérisation avec les tangentes**

Soit  $f : I \rightarrow \mathbb{R}$  dérivable sur  $I$ .

$f$  est convexe (respectivement concave) si et seulement si son graphe est au-dessus (respectivement au dessous) de toutes ses tangentes.

**Démonstration**

- Si la courbe de  $f$  est au dessus de ses tangentes, posons  $x < y$ . Soit  $t \in ]0, 1[$ ,  $z = (1-t)x + ty$  et donc  $t = \frac{z-x}{y-x}$ , le fait que la courbe de  $f$  est au dessus de la tangente en  $z$ , d'équation  $Y = f'(z)(X-z) + f(z)$  se traduit par  $f(x) \geq f'(z)(x-z) + f(z)$  et  $f(y) \geq f'(z)(y-z) + f(z)$  ce qui donne

$$(y-z)(f(z) - f(x)) \leq (z-x)(y-z)f'(z) \leq (z-x)(f(y) - f(z))$$

(avec  $z-x < 0$  et  $y-z > 0$ ) et donc  $(y-x)f(z) \leq (y-z)f(x) + (z-x)f(y)$  et enfin  $f(z) \leq (1-t)f(x) + tf(y)$ .

- Si  $f$  est convexe et dérivable, pour  $a \in I$ , on veut montrer que pour tout  $x \in I$ ,  $f(x) \geq f'(a)(x-a) + f(a)$ . Pour cela, on pose  $\varphi_a : x \in I \mapsto f(x) - f'(a)(x-a)$ , dérivable sur  $I$ , de dérivée  $\varphi'_a : x \mapsto f'(x) - f'(a)$ . Par croissance de  $f'$ ,  $\varphi_a$  est décroissante avant  $a$  et croissante après, donc admet un minimum en  $a$ , avec  $\varphi_a(a) = f(a)$ , ce qui permet bien de trouver l'inégalité voulue. □

**Exercice**

Étudier la convexité de  $\exp$ ,  $\ln$ ,  $t \mapsto t^\alpha$  sur  $\mathbb{R}_+^*$ ,  $\sin$ ,  $\tan$ .

**III INÉGALITÉS DE CONVEXITÉ****Propriété : Inégalité de Jensen**

Soit  $f : I \rightarrow \mathbb{R}$ ,  $n \geq 1$ ,  $(x_1, \dots, x_n) \in I^n$ ,  $(\lambda_1, \dots, \lambda_n) \in [0, 1]^n$  tel que  $\sum_{i=1}^n \lambda_i = 1$ .

- Si  $f$  est convexe,  $f\left(\sum_{i=1}^n \lambda_i x_i\right) \leq \sum_{i=1}^n \lambda_i f(x_i)$ .
- Si  $f$  est concave,  $f\left(\sum_{i=1}^n \lambda_i x_i\right) \geq \sum_{i=1}^n \lambda_i f(x_i)$ .

**Démonstration**

Cas  $f$  convexe. Récurrence sur  $n \in \mathbb{N}^*$ , on pose  $P_n$  : « pour tous  $(x_1, \dots, x_n) \in I^n$ ,  $(\lambda_1, \dots, \lambda_n) \in [0, 1]^n$  tel que  $\sum_{i=1}^n \lambda_i = 1$ ,

$$f\left(\sum_{i=1}^n \lambda_i x_i\right) \leq \sum_{i=1}^n \lambda_i f(x_i) ».$$

**Initialisation** : rien à faire pour  $n = 1$ . Pour  $n = 2$ , c'est la définition d'une fonction convexe.

**Hérédité** : Soit un  $n \in \mathbb{N}^*$  pour lequel  $P_n$  est vraie. et soient  $(x_1, \dots, x_{n+1}) \in I^{n+1}$  et  $(\lambda_1, \dots, \lambda_{n+1}) \in [0, 1]^{n+1}$  tel que  $\sum_{i=1}^{n+1} \lambda_i = 1$ . L'idée est de faire apparaître une associativité de barycentre (à savoir retrouver par le calcul, on vous le demandera sûrement en colle). Si  $\lambda_{n+1} = 1$ , alors tous les autres  $\lambda_i$  sont nuls et il n'y a rien à faire.



Sinon,  $\lambda = \sum_{i=1}^n \lambda_i = 1 - \lambda_{n+1} \neq 0$ , on pose  $x = \frac{1}{\lambda} \sum_{i=1}^n \lambda_i x_i = \text{Bar}((x_i, \lambda_i)_{1 \leq i \leq n})$ . Alors

$$\sum_{i=1}^{n+1} \lambda_i x_i = \text{Bar}((x_i, \lambda_i)_{1 \leq i \leq n+1}) = \text{Bar}((x, \lambda), (x_{n+1}, \lambda_{n+1})) = (1 - \lambda_{n+1})x + \lambda_{n+1}x_{n+1}$$

Alors, par définition de la convexité, puis par HR, comme les  $\mu_i = \frac{\lambda_i}{\lambda}$  sont positifs de somme 1,

$$f\left(\sum_{i=1}^{n+1} \lambda_i x_i\right) \leq (1 - \lambda_{n+1})f(x) + \lambda_{n+1}f(x_{n+1}) \leq \lambda \sum_{i=1}^n \mu_i f(x_i) + \lambda_{n+1}f(x_{n+1}) = \sum_{i=1}^{n+1} \lambda_i f(x_i).$$

ce qui établit la récurrence.

Détail du calcul (facile) pour l'associativité du barycentre (évitale, mais c'est l'idée sous-jacente à la preuve, à comprendre) :

$$\sum_{i=1}^{n+1} \lambda_i x_i = \sum_{i=1}^n \lambda_i x_i + \lambda_{n+1} x_{n+1} = \lambda x + \lambda_{n+1} x_{n+1} = (1 - \lambda_{n+1})x + \lambda_{n+1} x_{n+1}. \quad \square$$

### Remarque

Se rencontre souvent avec des poids tous égaux à  $\frac{1}{n}$ , ce qui donne  $f\left(\frac{1}{n} \sum_{i=1}^n x_i\right) \leq \frac{1}{n} \sum_{i=1}^n f(x_i)$ .

### Exemples : Très classiques

**E1 – Inégalité arithmético-géométrique** : pour tout  $x_1, \dots, x_n \in \mathbb{R}_*^+$ ,  $\sqrt[n]{x_1 \cdots x_n} \leq \frac{x_1 + \cdots + x_n}{n}$ . Qu'obtient-on en remplaçant  $x_i$  par  $\frac{1}{x_i}$  ?

**E2 –**  $\forall x > -1, \ln(1+x) \leq x.$   $\forall x \in \mathbb{R}, e^x \geq 1+x.$

**E3 –**  $\forall x \in \left[0, \frac{\pi}{2}\right], \frac{2}{\pi}x \leq \sin x \leq x.$

## IV QUE FAUT-IL SAVOIR FAIRE ?

- Savoir démontrer qu'une fonction dérivable ou deux fois dérivable est convexe.
- Reconnaître une inégalité de convexité, sous forme de somme ou de produit (et donc d'abord écrire l'inégalité « générique », avec les  $1/n$  (surtout) ou avec les  $\lambda_i$  (un peu quand même).)
- Savoir traduire par une inégalité le fait que le graphe d'une fonction convexe est en-dessus de ses tangentes.
- Savoir que  $\ln(1+x) \leq x$  si  $x \geq -1$ .
- Savoir passer de la « convexité pour 2 » à la « convexité pour n ».
- Savoir dessiner une fonction convexe...