Convexité

Marc SAGE (collab. Michel WIGNERON)

19 septembre 2017

Table des matières

1 Segments, parties étoilées, parties convexes, barycentres

Soit E un $\mathbb{R}$-espace vectoriel, soit A une partie de E.

Définitions (segment, partie étoilée, partie convexe)

*Soient $a, b \in E$. On appelle **segment** d'extrémités a et b la partie*[1]

$$[a,b] := \left\{a + \mu \ \overrightarrow{ab}\right\}_{\mu\in[0,1]} = \{(1-\mu)\,a + \mu b\}_{\mu\in[0,1]}\,.$$

*Soit $a \in A$. La partie A est dite **étoilée en** a lorsqu'elle inclut chaque segment d'extrémités a et un élément de A.*

FIG

*La partie A est dite **étoilée** lorsque qu'elle est étoilée en l'un de ses points :*

$$A \text{ étoilée} \overset{déf.}{\iff} \exists a \in A,\ \forall \alpha \in A,\ [a,\alpha] \subset A.$$

*La partie A est dite **convexe**[2] lorsqu'elle est étoilée en chacun de ses points :*

$$A \text{ convexe} \overset{déf.}{\iff} \forall a, \alpha \in A,\ [a,\alpha] \subset A.$$

FIG

Remarques – Soient $a, b \in E$.

- Le segment $[a,b]$ contient ses extrémités a et b (correspondant resp. à $\mu = 0$ et $\mu = 1$) :

$$\{a,b\} \subset [a,b]\,.$$

- Quand $a = b$, on a pour chaque réel μ les égalités $a + \mu\overrightarrow{ab} = a + \mu 0 = a + 0 = a$, d'où l'inclusion $[a,a] \subset \{a\}$, ce qui montre avec la remarque précédente que le segment $[a,b]$ se réduit à un singleton :

$$[a,a] = \{a\}\,.$$

- Reparamétrer $[a,b]$ à l'aide de la bijection $\left\{\begin{array}{ccc}[0,1] & \xrightarrow{\sim} & [0,1] \\ t & \longmapsto & 1-t\end{array}\right.$ livre l'égalité des segments

$$[a,b] = \{\lambda a + (1-\lambda)\,b\}_{\lambda\in[0,1]} = [b,a]\,.$$

- Lorsque $E = \mathbb{R}$, il pourrait y avoir confusion entre[3]

le segment ordonné $\{t \in \mathbb{R}\ ;\ a \leq t \leq b\}$ et le segment convexe $\{\lambda a + (1-\lambda)\,b\}_{\lambda\in[0,1]}$.

Montrons leur égalité lorsque $a \leq b$.

[1] *Rappel* : on a noté $\overrightarrow{ab} := b - a$.

[2] On dit aussi d'une telle partie qu'elle est *stable par passage au segment*.

[3] Quand $a > b$, seul le contexte permettra de savoir si le segment $[a,b]$ est vide (car ordonné) ou non (car convexe et valant alors $[b,a]$).

La fonction $\lambda \overset{\pi}{\mapsto} \lambda a + (1-\lambda)\, b$ décroissant continûment de b à a, elle atteint (par le théorème des valeurs intermédiaires) chaque réel t tel que $a \leq t \leq b$, d'où l'inclusion $\subset$.

Soit par ailleurs $\lambda \in [0,1]$: l'encadrement $a \leq \lambda a + (1-\lambda)\, b \leq b$ équivaut alors aux comparaisons $(1-\lambda)\, a \leq (1-\lambda)\, b$ et $\lambda a \leq \lambda b$, lesquelles s'obtiennent en multipliant celle $a \leq b$ par les réels positifs $1-\lambda$ et λ resp., d'où l'inclusion $\supset$.

- Chaque convexe est étoilé ou vide (passage de $\forall$ à $\exists$).
- En termes de lois "parties", la partie A est convexe ssi

$$\forall \lambda \in [0,1]\,,\ \lambda A + (1-\lambda)\, A \subset A.$$

On peut d'ailleurs remplacer l'inclusion $\subset$ par une égalité vu que l'inclusion réciproque $\supset$ est toujours vérifiée, chaque point $a \in A$ s'écrivant $a = \lambda a + (1-\lambda)\, a$ pour chaque $\lambda \in [0,1]$.

- Chaque énoncé de la forme[4] $P(\lambda, 1-\lambda)$ quantifié sur $\lambda \in [0,1]$ peut se réécrire $P(\lambda, \mu)$ où l'on quantifie sur les couples $\binom{\lambda}{\mu} \in [0,1]$ de somme 1. Nous préférerons dans ce cours cette réécriture où λ et μ sont souvent interchangeables. Par exemple,

$$\begin{gathered}\text{on a l'égalité } [a,b] = \{\lambda a + \mu b\}_{\lambda,\mu \in [0,1]}^{\lambda+\mu=1} \text{ et} \\ A \text{ est convexe ssi } \forall \lambda, \mu \in [0,1]\,,\ \lambda + \mu = 1 \Longrightarrow \lambda A + \mu A = A.\end{gathered}$$

- *Culture hors programme* : lorsqu'un musée est étoilé, il est théoriquement possible pour un visiteur doué d'une vue perçante de contempler toutes les œuvres du musée depuis un même point (où il pourra s'asseoir). Un théorème de Mark Alexandrovich Krasnosselsky (publié en 1946 dans le *Recueil Mathématique* de la Société mathématique de Moscou) affirme qu'il suffit pour cela, lorsque le musée est plan, que chaque *trio* d'œuvres soit visible depuis un même point (pas forcément le même selon le trio). La lectrice et le lecteur pourront à ce sujet consulter avec profit les ouvrages de géométrie de Michel Berger.

Exemples

1. Les convexes de $\mathbb{R}$ sont ses intervalles (théorème de première année).
2. Est convexe chaque singleton, chaque segment, chaque triangle/carré/disque plein, chaque tétraèdre/cube/boule plein(e), chaque demi-plan/demi-espace.
3. Dans ces exemples précédents, retirer le "bord[5]" de la partie convexe considérée conserve sa convexité (par exemple, est convexe chaque intervalle ouvert, chaque disque ouvert, chaque demi-plan ouvert).
4. Sont convexes l'espace total E et la partie vide $\emptyset$.
5. Les endomorphismes nilpotents de E forment une partie de $L(E)$ étoilée en l'endomorphisme nul.
6. N'est convexe aucun cercle de $\mathbb{R}^2$ (à moins d'être de rayon nul), aucune sphère de $\mathbb{R}^3$ (sauf à être réduite à son centre).

[4] La lettre P désigne ici un ***prédicat binaire***, *i. e.* une expression à deux symboles "libres". Pour chaques symboles $\heartsuit$ et $\clubsuit$, la notation $P(\heartsuit, \clubsuit)$ abrège l'expression symbolique obtenue en remplaçant dans P le premier symbole libre par $\heartsuit$ et le second par $\clubsuit$.

[5] Cette notion de "bord" sera formalisée au chapitre ? ? ?evn1.

7. Pour chaque naturel n, la matrice identité et les matrices des transvections forment une partie de $M_n(\mathbb{R})$ étoilée en I_n.
8. La réunion des arêtes de chaque carré (de côté non nul) n'est pas convexe, de même pour[6] la réunion des arêtes de chaque triangle (à moins d'être aplati) et pour la réunion des faces de chaque cube (sauf cas d'un singleton).

Montrons à titre d'exercice que *la convexité est stable par intersection* (non nécessairement finie).

Soit $(C_i)_{i\in I}$ une famille de convexes de E, soient $c, d \in \bigcap_{i\in I} C_i$ et soient $\lambda, \mu \in [0,1]$ de somme 1. Puisque le point $\lambda c + \mu d$ reste pour chaque $i \in I$ dans C_i par convexité de ce dernier, il tombe dans l'intersection des C_i, *c. q. f. d.*

En revanche, *la convexité n'est pas toujours préservée par union* : par exemple, la réunion de deux singletons/disques disjoints n'est jamais convexe.

Montrons également qu'*est convexe la réunion de chaque suite croissante de parties convexes.*

Soit $(C_n)_{n\in\mathbb{N}}$ une telle suite, soient $c, d \in \bigcup_{n\in\mathbb{N}} C_n$, soient $p, q \in \mathbb{N}$ tels que $(c, d) \in C_p \times C_q$ et soient $\lambda, \mu \in [0,1]$ de somme 1. En notant $m := \max\{p, q\}$, les comparaisons $p, q \leq m$ et la croissance de la suite évoquée livrent les inclusions $C_p, C_q \subset C_m$, donc ce dernier (qui contient les points c et d) contient le barycentre $\lambda c + \mu d$, lequel tombe par conséquent dans $\bigcup_{n\in\mathbb{N}} C_n$.

Définition (barycentre)

On appelle ***barycentre***[7] *de A toute combinaison linéaire de points de A dont la somme des coefficients vaut* 1*.*

Remarques

• Chaque segment est formé des barycentres à coefficients *positifs* de ses deux extrémités.

• Lorsque $E = \mathbb{R}$, on peut expliciter ces coefficients à l'aide, d'une part, des extrémités du segment considéré et, d'autre part, du barycentre considéré : on a en effet pour chaques réels[8] m, a, M l'implication

$$m < M \Longrightarrow a = \frac{M-a}{M-m}m + \frac{a-m}{M-m}M.$$

(ces coefficients ont été intuités en résolvant l'équation $a = \lambda m + (1-\lambda) M$ d'inconnue réelle λ).

• Lorsque les poids d'un barycentre sont égaux, on parle d'***isobarycentre*** ou de ***centre de gravité***. Par exemple, l'isobarycentre de chaque couple de points est le milieu du segment les joignant et (*sanity check* justifiant la terminologie) le centre de gravité de chaque triplet de points est le centre de gravité du triangle qu'ils forment.

• Étymologiquement, *barycentre* signifie « le point où les charges sont équilibrées » : *barus* = chargé/pesant, *kentron* = aiguille/pointe (l'équilibration est sous-entendue).

[6] Retenir qu'une figure "en fil de fer" n'est (presque) jamais convexe.

[7] *Culture* : on parle aussi de ***combinaison affine***. Lorsque les coefficients sont en outre positifs, on parle de ***combinaison convexe***.

[8] m comme « **min**imum », M comme « **max**imum »

• En mécanique, une tige de masse négligeable devant celles de ses extrémités tenue par le barycentre de ces dernières sera en équilibre :

FIG

De même, une tige de masse totale 1 aura pour centre de gravité le "barycentre continu"

$$\int_{P\in \text{ tige}} \delta m_P \ P \qquad \text{où } \delta m_P \text{ désigne la "masse infinitésimale" du point } P.$$

Plus généralement, si λ dénote la densité linéique de la tige, on aura en tout point P de la tige l'égalité $\delta m_P = \lambda_P \ \delta\ell_P$, d'où le centre de gravité[9]

$$\frac{\int_{P\in \text{ tige}} P \ \lambda_P \ \delta\ell_P}{\int_{P\in \text{ tige}} \lambda_P \ \delta\ell_P} \qquad \begin{array}{c}\text{où } \delta\ell_P \text{ désigne la "longueur infinitésimale"} \\ \text{du point } P \text{ pensé comme "bout de tige".}\end{array}$$

On aurait des formules analogues en dimension 2 et 3 en remplaçant « linéique » par « surfacique » ou « volumique ».

• La barycentration possède une très utile propriété d'*associativité*. Soient $p, q \in \mathbb{N}^*$, soient $a \in E^p$ et $b \in E^q$, soient $\lambda \in \mathbb{R}^p$ et $\mu \in \mathbb{R}^{q+1}$ chacun de somme non nulle. L'égalité

$$\mu_0 \left(\sum_{i=1}^{p} \lambda_i a_i \right) + \sum_{j=1}^{q} \mu_j b_j = \sum_{i=1}^{p} (\mu_0 \lambda_i)\, a_i + \sum_{j=1}^{q} \mu_j b_j$$

énonce alors une égalité entre un barycentre à $q+1$ points (le premier point étant lui-même un barycentre à p points) et un barycentre à $p+q$ points. Cette propriété peut par exemple servir pour démontrer que les trois médianes d'un triangle se rencontrent en son centre de gravité. Elle permet également de raisonner par récurrence (comme dans la proposition suivante).

Proposition (convexité & barycentres)

La partie A est convexe ssi elle est stable par barycentration à coefficients positifs.

Démonstration

Le sens réciproque est immédiat vu que chaque point d'un segment de A est un barycentre de deux points de A à coefficients positifs (λ et $1-\lambda$ pour un certain $\lambda \in [0,1]$).

Imposons maintenant A stable par passage au segment. Pour chaque naturel $n \geq 1$, notons B_n l'énoncé[10]

$$\begin{array}{l}\forall a \in A^n \\ \forall \lambda \in \mathbb{R}_+^n\end{array}, \ \sum_{i=1}^{n} \lambda_i = 1 \Longrightarrow \sum_{i=1}^{n} \lambda_i a_i \in A$$

et montrons $\forall n \in [2, \infty[\![\,,\ B_n$ par récurrence (B_1 est trivial).

L'énoncé B_2 équivaut à notre imposition.

[9] Lorsque λ est somme de deux "Dirac" en chacune des extrémités de la tige, on retrouve un barycentre discret.

[10] L'énoncé B_n énonce la stabilité de A par **b**arycentration de n points à coefficients positifs.

Soit $n \in [2, \infty[\![$ tel que B_n. Soient $a \in A^{n+1}$ et $\lambda \in \mathbb{R}_+^{n+1}$ tel que $\lambda_0 + \sum_{i=1}^n \lambda_i = 1$.

Si $\lambda_0 = 1$, la somme $\Lambda := \sum_{i=1}^n \lambda_i$ est alors nulle et la positivité des λ_i impose $\forall i \in [1, n]$, $\lambda_i = 0$, de sorte que le barycentre $\sum_{i=0}^n \lambda_i a_i$ se réduit à $\lambda_0 a_0 = a_0$ et reste bien dans A.

Dans le cas contraire, la somme Λ est non nulle, donc le barycentre $\beta := \sum_{i=1}^n \frac{\lambda_i}{\Lambda} a_i$ fait sens et reste dans A d'après B_n, donc le barycentre $\lambda_0 a_0 + \Lambda\beta$ reste dans A d'après B_2 ; or ce dernier barycentre vaut

$$\lambda_0 a_0 + \Lambda \left(\sum_{i=1}^n \frac{\lambda_i}{\Lambda} a_i \right) = \sum_{i=0}^n \lambda_i a_i, \text{ ce qui conclut à } B_{n+1}.$$

Exercice d'application

On appelle ***enveloppe convexe*** *de A l'ensemble de ses barycentres à coefficients positifs. On la note*

$$\operatorname{Conv} A := \left\{ \sum_{i=1}^n \lambda_i a_i \ ; \ n \in \mathbb{N}^*, \quad \begin{matrix} a \in A^n \\ \lambda \in \mathbb{R}_+^n \end{matrix} \ , \ \sum_{i=1}^n \lambda_i = 1 \right\}.$$

a. *Montrer l'inclusion*

$$A \subset \operatorname{Conv} A \text{ avec égalité ssi } A \text{ est convexe.}$$

b. *Soit $n \in \mathbb{N}$. Trouver une partie de $\mathbb{R}^{n+1}$ de cardinal $n+1$ dont l'enveloppe convexe vaut la partie*

$$\left\{ \lambda \in \mathbb{R}_+^{n+1} \ ; \ \sum_{i=0}^n \lambda_i = 1 \right\}.$$

Que peut-on en déduire ? Faire le lien avec des exemples du cours.

---

a. Chaque point de A étant barycentre de lui-même affecté du poids 1, on a l'inclusion $A \subset \operatorname{Conv} A$.

La partie A étant convexe ssi elle est stable par barycentration à coefficients positifs, sa convexité revient par définition à l'inclusion réciproque $\operatorname{Conv} A \subset A$, d'où le cas d'égalité quand A est convexe.

Montrons que $\operatorname{Conv} A$ est convexe : on en déduira la convexité de A lorsqu'il y a égalité $A = \operatorname{Conv} A$. Soient donc $\lambda, \mu \in [0, 1]$ de somme 1, soient $c, d \in (\operatorname{Conv} A)$, soient $p, q \in \mathbb{N}^*$, soient $\begin{matrix} \overrightarrow{\lambda} \in \mathbb{R}_+^p \\ \overrightarrow{\mu} \in \mathbb{R}_+^q \end{matrix}$ chacun de somme 1 et soient $\begin{matrix} a \in A^p \\ b \in A^q \end{matrix}$ tels que $\begin{cases} c = \sum_{i=1}^p \lambda_i a_i \\ d = \sum_{j=1}^q \mu_j b_j \end{cases}$. Le barycentre $\lambda c + \mu d$ s'écrit alors

$$\lambda c + \mu d = \lambda \left(\sum_{i=1}^p \lambda_i a_i \right) + \mu \left(\sum_{j=1}^q \mu_j b_j \right) = \sum_{i=1}^p (\lambda \lambda_i) a_i + \sum_{j=1}^q (\mu \mu_j) b_j = \sum_{k=1}^n \nu_k e_k$$

où l'on a défini $n := p + q$,

$$\nu := \begin{cases} [1, n|] & \longrightarrow & \mathbb{R}_+ \\ k & \longmapsto & \begin{cases} \lambda\lambda_k \text{ si } k \leq p \\ \mu\mu_{k-p} \text{ si } k > p \end{cases} \end{cases} \quad \text{et}$$

$$e := \begin{cases} [1, n|] & \longrightarrow & A \\ k & \longmapsto & \begin{cases} a_k \text{ si } k \leq p \\ b_{k-p} \text{ si } k > p \end{cases} \end{cases}.$$

Puisque chaque point e_k tombe dans A et que la famille ν a pour somme

$$\sum_{k=1}^{n} \nu_k = \sum_{k=1}^{p} \nu_k + \sum_{k=p+1}^{n} \nu_k = \sum_{i=1}^{p} \nu_i + \sum_{j=1}^{q} \nu_j$$

$$= \sum_{i=1}^{p} \lambda\lambda_i + \sum_{j=1}^{q} \mu\mu_j = \underbrace{\lambda \overbrace{\sum_{i=1}^{p} \lambda_i}^{=1} + \mu \overbrace{\sum_{j=1}^{q} \mu_j}^{=1}}_{=\lambda+\mu} = 1,$$

le barycentre $\lambda c + \mu d = \sum_{k=1}^{n} \nu_k e_k$ reste bien dans $\operatorname{Conv} A$, *c. q. f. d.*

Remarque – On montrerait aisément que l'enveloppe convexe de A est la plus petite partie convexe le contenant (description "externe" : intersection des convexes de E incluant A). C'est donc l'analogue en convexité du sous-espace vectoriel engendré en théorie vectorielle. On pourra ainsi montrer qu'une partie donnée est convexe en en exhibant un ensemble dont elle est l'enveloppe convexe.

b. Notons[11] Δ la partie considérée et notons A la partie de $\mathbb{R}^{n+1}$ formée par les vecteurs de sa base canonique.

FIG

Soit $\delta \in \Delta$. En décomposant δ selon la base canonique de $\mathbb{R}^{n+1}$, on voit que δ est combinaison linéaire de points de A à coefficients positifs et de somme 1 (car les coordonnées de δ sont positives et de somme 1), d'où l'appartenance $\delta \in \operatorname{Conv} A$ et l'inclusion $\Delta \subset \operatorname{Conv} A$.

Soit $c \in \operatorname{Conv} A$, soit $p \in \mathbb{N}^*$ et soient $\begin{matrix} \lambda \in \mathbb{R}_+^p \\ a \in A^p \end{matrix}$ tels que $\begin{cases} \sum_{i=1}^{p} \lambda_i = 1 \\ c = \sum_{i=1}^{p} \lambda_i a_i \end{cases}$. Chaque $\lambda_i a_i$ étant à coordonnées positives, la somme $\sum_{i=1}^{p} \lambda_i a_i = c$ est à coordonnées positives. L'application "somme des coordonnées" étant par ailleurs linéaire et valant 1 en chaque point de A, l'appliquer en $c = \sum_{i=1}^{p} \lambda_i a_i$ montre que la somme des coordonnées de c vaut $\sum_{i=1}^{p} \lambda_i 1 = 1$, ce qui conclut à l'appartenance $c \in \Delta$ et à l'inclusion $\operatorname{Conv} A \subset \Delta$.

Nous déduisons de ce qui précède la convexité de[12] Δ. Lorsque $n \leq 3$, on retrouve quatre des exemples du cours : point, segment, triangle plein, tétraèdre plein.

Culture hors programme : lorsque $E = \mathbb{R}^n$, un théorème de Constantin Carathéodory (publié en 1907 dans les *Mathematische Annalen*) affirme que

[11] Les dimensions 2 et 3 motivent la notation Δ en forme de triangle.

[12] *Culture* : ce convexe s'appelle un ***simplexe*** de dimension n.

chaque point de $\operatorname{Conv} A$ est barycentre d'$\underline{\underline{\text{au plus } n+1}}$ points de A à coefficients positifs (regarder par exemple le cas où A est un carré). Un théorème de Mark Krein et David Milman (publié en 1940 dans la revue polonaise *Studia Mathematica*) généralise par ailleurs le fait "intuitif" que chaque polygone convexe est enveloppe convexe de l'ensemble de ses sommets.

2 Fonctions convexes

Le cours de première année étudiait les fonctions préservant les barycentres, *i. e.* les combinaisons affines (fonctions justement dites *affines*). Concernant la convexité, préserver les barycentres *à coefficients positifs* ne s'avère guère plus intéressant. En revanche, lorsque l'espace but est *ordonné*, par exemple réel, il se révèle pertinent d'étudier les fonctions par lesquelles l'image de chaque barycentre à coefficients positifs est *plus petit que* le barycentre des images associées : le programme qualifie ces fonctions de "convexes" et se restreint au cas où l'espace but est $\mathbb{R}$.

On évoque pour toute cette partie un convexe de $\mathbb{R}$, *i. e.* un intervalle réel, noté I, et une fonction $f : I \longrightarrow \mathbb{R}$.

2.1 Convexité de l'épigraphe, graphe et cordes

Proposition – Définition (épigraphe, fonction convexe/concave, comparaisons de Jensen)

Les conditions suivantes sont équivalentes :

1. *Est une partie convexe*[13] *de* $\mathbb{R}^2$ *l'**épigraphe** de* f *défini par l'ensemble des couples* $(i,j) \in I \times \mathbb{R}$ *tels que* $j \geq f(i)$ *;*

   FIG : épigraphe d'une fonction

2. *pour chaques points* $a, b \in I$ *et pour chaques réels* $\lambda, \mu \in [0,1]$ *de somme* 1*, on a la comparaison*[14]
$$f(\lambda a + \mu b) \leq \lambda f(a) + \mu f(b).$$

3. *Pour chaque naturel* $n \in \mathbb{N}^*$*, pour chaque* n*-uplet* $\overrightarrow{a}$ *de points de* I *et pour chaque* n*-uplet* $\overrightarrow{\lambda}$ *de réels de* $[0,1]$ *de somme* 1*, on a la comparaison*[15]
$$f(\lambda_1 a_1 + \lambda_2 a_2 + \cdots + \lambda_n a_n) \leq \lambda_1 f(a_1) + \lambda_2 f(a_2) + \cdots + \lambda_n f(a_n).$$

[13] C'est ce point qui motive le qualificatif *convexe* attribué à un tel f.

[14] Définition officielle de la convexité de f.

[15] Ces comparaisons éponymes furent publiées en 1906 par Johan Jensen (prononcer *yène-çène*) dans la revue *Acta Mathematica*.

Lorsqu'une des conditions ci-dessus est vérifiée, la fonction f est dite ***convexe.***

L'application f est dite **concave** *si $-f$ est convexe,* i. e. *si l'on a les comparaisons ci-dessus dans l'autre sens.*

Pour chaque intervalle $\mathcal{I} \subset I$, la fonction f est dite ***convexe sur*** $\mathcal{I}$ *(resp.* ***concave sur*** $\mathcal{I}$*) lorsque la restriction $f_{|\mathcal{I}}$ est convexe (resp. concave).*

Démonstration

$\boxed{3 \Longrightarrow 2}$ Cas particulier $n = 2$ (remplacer $\binom{a_1}{a_2}$ par $\binom{a}{b}$ et $\binom{\lambda_1}{\lambda_2}$ par $\binom{\lambda}{\mu}$).

$\boxed{2 \Longrightarrow 1}$ Soient $\binom{a}{a'}$ et $\binom{b}{b'}$ dans l'épigraphe de f, soient $\lambda, \mu \in [0,1]$ de somme 1. On a alors les comparaisons

$$f(\lambda a + \mu b) \leq \lambda f(a) + \mu f(b) \overset{\lambda, \mu \geq 0}{\leq} \lambda a' + \mu b',$$

ce qui montre que le couple $\binom{\lambda a + \mu b}{\lambda a' + \mu b'} = \lambda \binom{a}{a'} + \mu \binom{b}{b'}$ est dans l'épigraphe de f, *c. q. f. d.*

$\boxed{1 \Longrightarrow 3}$ Soit $n \in \mathbb{N}^*$, soit $a \in I^n$ et soit $\lambda \in [0,1]^n$ de somme 1. Le point $\binom{a_i}{f(a_i)}$ étant pour chaque $i \in [1, n|]$ sur l'épigraphe de f, cet épigraphe contient par convexité le barycentre

$$\sum_{i=1}^{n} \lambda_i \begin{pmatrix} a_i \\ f(a_i) \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{n} \lambda_i a_i \\ \sum_{i=1}^{n} \lambda_i f(a_i) \end{pmatrix},$$

donc l'ordonnée de ce dernier point majore (par définition de l'épigraphe de f) l'image par f de son abscisse, *c. q. f. d.*

Exemples

1. Chaque fonction affine est convexe et concave ; réciproquement, chaque fonction convexe et concave est affine. Cet exemple correspondant au cas d'égalité dans les comparaisons de Jensen, on observera en conséquence qu'

ajouter une fonction affine n'altère pas le caractère convexe (resp. concave).

FIG

2. La fonction "norme" $t \mapsto |t|$ est convexe sur $\mathbb{R}$ vu pour chaques réels a, b et pour chaques réels λ, μ positifs les comparaisons

$$|\lambda a + \mu b| \underset{\text{triangulaire}}{\overset{\text{comparaison}}{\leq}} |\lambda a| + |\mu b| = |\lambda| |a| + |\mu| |b| \overset{\lambda, \mu \geq 0}{=} \lambda |a| + \mu |b| .$$

FIG

3. Soient $\alpha < \beta$ deux réels. Montrons que la fonction χ caractéristique de la paire $\{\alpha, \beta\}$ est convexe sur le segment $[\alpha, \beta]$:

FIG

Soient $a<b$ dans $[\alpha,\beta]$ et soient $\lambda,\mu>0$ de somme 1. Le réel $\lambda a+\mu b$ est alors un barycentre *non trivial* de a et b, donc tombe dans $]a,b[\subset]\alpha,\beta[$, d'où les égalités et comparaison

$$\chi(\lambda a+\mu b)=0=\lambda 0+\mu 0 \underset{\chi\geq 0}{\overset{\lambda,\mu\geq 0}{\leq}} \lambda\chi(a)+\mu\chi(b).$$

Il reste à établir la comparaison ci-dessus sans les hypothèses $a<b$ et $\lambda,\mu>0$: la remarque suivante montre que cela est en fait inutile.

Remarques

• Si $a=b$, la comparaison du point 2 s'écrit $f(a)\leq f(a)$, ce qu'on a. De même, si $\{\lambda,\mu\}=\{0,1\}$, cette comparaison s'écrit $f(a)\leq f(a)$ ou $f(b)\leq f(b)$, ce qu'on a. On pourra donc, pour montrer la convexité/concavité de f, se restreindre aux cas $\lambda,\mu\in]0,1[$ et $a\neq b$, voire $a<b$ quitte à échanger les couples $\binom{a}{\lambda}$ et $\binom{b}{\mu}$:

$$f \textit{ convexe ssi } \begin{array}{c}\forall a,b\in I\\ \forall\lambda,\mu>0\end{array}, \left\{\begin{array}{c}a<b\\ \lambda+\mu=1\end{array}\right. \Longrightarrow f(\lambda a+\mu b)\leq\lambda f(a)+\mu f(b).$$

• Imposons f convexe et soit S un segment inclus dans I. L'épigraphe de $f_{|S}$ est alors l'intersection de l'épigraphe de f (qui est une partie convexe) avec la partie convexe $S\times\mathbb{R}$, donc est convexe, *i. e.* $f_{|S}$ est convexe. Réciproquement, si f est convexe sur chacun des segments de I, en évoquant une suite croissante $(S_n)_{n\in\mathbb{N}}$ de tels segments dont la réunion vaut[16] I, l'épigraphe de f vaudra alors la réunion croissante des épigraphes (chacun convexe) de $f_{|S_n}$, donc sera convexe. Par conséquent :

f est convexe sur l'intervalle I ssi elle l'est sur chacun de ses segments.

• Soient $a<b$ dans I. Observer que l'application

$$t\mapsto(\lambda_t,\mu_t):=\left(\frac{b-t}{b-a},\frac{t-a}{b-a}\right)$$

réalise une bijection de $[a,b]$ sur le segment Σ de $\mathbb{R}^2$ formé des $\binom{\lambda}{\mu}\in[0,1]^2$ de somme 1 (de réciproque $\binom{\lambda}{\mu}\mapsto\lambda a+\mu b$). Ce reparamétrage permet d'affirmer l'équivalence

$$\left[\forall\tbinom{\lambda}{\mu}\in\Sigma,\ f(\lambda a+\mu b)\leq\lambda f(a)+\mu f(b)\right]\Longleftrightarrow\forall t\in[a,b],\ f(t)\leq\lambda_t f(a)+\mu_t f(b).$$

Or le membre à droite $\lambda_t f(a)+\mu_t f(b)$ est affine en t et coïncide en a et b avec[17] f, donc a pour graphe (lorsque t varie) la corde du graphe de f reliant les points $\binom{a}{f(a)}$ et $\binom{b}{f(b)}$. Par conséquent :

f est convexe sur I ssi, pour chaque segment $S\subset I$,
le graphe de $f_{|S}$ est en dessous de la corde reliant ses extrémités.

FIG

[16] Par exemple $]0,\infty[=\bigcup_{n\in\mathbf{N}^*}\left[\frac{1}{n},n\right]$.
[17] Observer les égalités $\binom{\lambda_a}{\mu_a}=\binom{1}{0}=\binom{\mu_b}{\lambda_b}$.

• **Mnémonique**. Imposons f convexe ou concave : comment trancher ? Elle sera con**V**exe si son graphe est en forme de **V**, conc**A**ve si[18] son graphe est en forme de **A** :

FIG

• Pour une fonction convexe, la comparaison de Jensen s'énonce (de façon très abusive)

l'image du barycentre est sous le barycentre des images.

FIG

À l'aide du graphe de la fonction considérée, de la mnémonique précédente et de l'énoncé (abusif) ci-dessus, il devient *impossible* de se tromper dans le sens des comparaisons définissant une fonction convexe ou concave.

Les exemples à suivre devront convaincre *graphiquement* en attente d'une preuve formelle à venir (*cf.* section 2.3).

Exemples

1. Pour chaque naturel n, la fonction $t \mapsto t^n$ est convexe sur $\mathbb{R}_+$, $\begin{array}{l} \text{convexe sur } \mathbb{R}_- \text{ si } n \text{ est pair} \\ \text{concave sur } \mathbb{R}_- \text{ si } n \text{ est impair} \end{array}$ et les deux (*i. e.* affine) ssi $n \in \{0, 1\}$:

   FIG

2. Pour chaque réel $\alpha \geq 0$, la fonction $t \mapsto t^\alpha$ (définie sur $\mathbb{R}_+$) est $\begin{array}{l} \text{convexe si } \alpha \geq 1 \\ \text{concave si } \alpha \in [0, 1] \end{array}$.
   On a de même la convexité sur $\mathbb{R}_+^*$ de la fonction $t \mapsto \frac{1}{t^\alpha}$ dès que $\alpha > 0$:

   FIG

3. *Attention* : le domaine de définition $\mathbb{R}^*$ de la fonction « inverse » n'étant pas convexe, ni la convexité ni la concavité de cette fonction ne font sens. En revanche, on peut tout à fait affirmer la convexité de $\left\{ \begin{array}{ccc} \mathbb{R}_+^* & \longrightarrow & \mathbb{R}_+^* \\ t & \longmapsto & \frac{1}{t} \end{array} \right.$ et la concavité de $\left\{ \begin{array}{ccc} \mathbb{R}_-^* & \longrightarrow & \mathbb{R}_-^* \\ t & \longmapsto & \frac{1}{t} \end{array} \right.$.

   FIG

4. L'exponentielle et le cosinus hyperbolique sont convexes sur $\mathbb{R}$, le logarithme est concave sur $\mathbb{R}_+^*$:

   FIG

5. La fonction[19] $t \mapsto t \ln t$ (prolongée par $0 \mapsto 0$) est convexe sur $\mathbb{R}_+$:

   FIG

[18] Attention, les réciproques sont généralement fausses : le graphe de la fonction convexe exp n'est pas en forme de V.
[19] *Culture hors programme* : cette fonction intervient dans les définitions usuelles de l'*entropie*.

6. La fonction cos est concave sur $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$, la fonction sin est convexe sur $[-\pi, 0]$ et concave sur $[0, \pi]$:

FIG

En particulier, le graphe de sin est au-dessus de sa corde sur $\left[0, \frac{\pi}{2}\right]$, ce qui s'écrit

$$\forall \theta \in \left[0, \frac{\pi}{2}\right], \ \frac{2\theta}{\pi} \leq \sin \theta.$$

7. Les fonctions th, arctan et argsh sont concaves sur $\mathbb{R}_+$ (et convexes sur $\mathbb{R}_-$), la fonction argch est concave sur $[1, \infty[$:

FIG

8. La fonction tan est convexe sur $\left[0, \frac{\pi}{2}\right[$, la fonction argth est convexe sur $[0, 1[$, les fonctions arcsin et arccos sont resp. convexe et concave sur $[0, 1]$:

FIG

9. Soit $z \in \mathbb{Z}$. La fonction "partie entière" est convexe sur $[z, z+1]$ et concave (car affine) sur $[z, z+1[$ mais n'est ni convexe ni concave sur aucun intervalle ouvert contenant z vu pour chaque $\varepsilon \in \left]0, 1\right[$ les comparaisons strictes

$$\frac{\lfloor z-\varepsilon \rfloor + \lfloor z+\varepsilon \rfloor}{2} = \frac{(z-1)+z}{2} = z - \frac{1}{2} < z = \lfloor z \rfloor = \left\lfloor \frac{(z-\varepsilon)+(z+\varepsilon)}{2} \right\rfloor \text{ et}$$
$$\frac{\lfloor z-\varepsilon \rfloor + \lfloor z \rfloor}{2} = \frac{(z-1)+z}{2} = z - \frac{1}{2} > z - 1 = \left\lfloor z - \frac{\varepsilon}{2} \right\rfloor = \left\lfloor \frac{(z-\varepsilon)+(z)}{2} \right\rfloor.$$

FIG

Exercices d'application

1. *Imposons f convexe et soient $a < b < c$ dans I dont les points associés sur le graphe de f sont alignés. Montrer alors que f est affine sur $[a, c]$.*
2. *Lorsque f est continue, montrer qu'elle est convexe sur I ssi elle est convexe sur l'intérieur*[20] $\mathring{I}$.
3. *Imposons f continue. Montrer alors l'équivalence*

$$f \text{ convexe} \iff \forall a, b \in I, \ f\left(\frac{a+b}{2}\right) \leq \frac{f(a)+f(b)}{2}.$$

[20] *Rappel* : l'intérieur $\mathring{I}$ de l'intervalle I est défini par ce dernier privé de ses éventuelles extrémités.

1. Notons δ la différence[21] entre f et la fonction affine qui coïncide avec en a et en c : il s'agit de montrer la nullité sur $[a,c]$ de la fonction δ. Pour cela, on montre qu'elle est nulle sur $[a,b[$ et nulle sur $]b,c]$ (elle est déjà nulle en b d'après l'hypothèse de colinéarité).

FIG

Soient donc $x \in [a,b[$ et $y \in]b,c]$. Le point b est alors un barycentre *non trival* de x et y, mettons $b = \lambda x + \mu y$ pour certains réels λ, μ *strictement* positifs de somme 1. On a alors les comparaisons

$$0 = \delta(b) = \delta(\lambda x + \mu y) \overset{\delta \text{ convexe}}{\leq} \lambda\delta(x) + \mu\delta(y) \overset{\lambda,\mu\geq 0}{\leq} \lambda 0 + \mu 0 = 0,$$

d'où égalité partout, ce qui impose les égalités $\lambda\delta(x) = 0 = \mu\delta(y)$, d'où (puisque $\lambda \neq 0$ et $\mu \neq 0$) les nullités de $\delta(x)$ et $\delta(y)$, *c. q. f. d.*

2. Le sens direct étant trivial (sans hypothèse), imposons f continue sur I et convexe sur $\mathring{I}$. Soient $a, b \in I$. On peut alors approcher a et b par des points *intérieurs* de I, mettons $\binom{a}{b} = \binom{\lim a_n}{\lim b_n}$ pour certaines suites $\overrightarrow{a}, \overrightarrow{b} \in \mathring{I}^{\mathbb{N}}$. Soient $\lambda, \mu \in [0,1]$ de somme 1. La convexité de f sur $\mathring{I}$ livre alors les comparaisons $\forall n \in \mathbb{N},\ f(\lambda a_n + \mu b_n) \leq \lambda f(a_n) + \mu f(b_n)$, lesquelles donnent après passage à la limite (et en utilisant la continuité de f) la comparaisons $f(\lambda a + \mu b) \leq \lambda f(a) + \mu f(b)$, *c. q. f. d.*

   Remarque – L'exercice 3 permettrait d'alléger un peu la rédaction en imposant $\lambda = \frac{1}{2} = \mu$.

3. Le sens direct étant immédiat en utilisant des isobarycentres (remplacer λ et μ par $\frac{1}{2}$), on peut imposer les comparaisons de gauche. Soient alors $a, b \in I$ et notons φ la fonction affine[22] coïncidant avec f sur $\{a, b\}$. Il s'agit de montrer la négativité de la fonction $\delta := f - \varphi$, laquelle vérifie les comparaisons de l'hypothèse d'après le sens direct (et car $-\varphi$ est convexe). Nous proposons deux solutions.

   (a) **par un maximum.** La fonction δ est continue sur le segment $[a,b]$, donc atteint son maximum en un certain $m \in [a,b]$ et l'on veut alors montrer $\delta(m) \leq 0$.

   FIG

   Vu les égalités $\frac{a+(2m-a)}{2} = m = \frac{(2m-b)+b}{2}$ où l'un des réels $2m-a$ ou $2m-b$ tombe dans $[a,b]$, on peut écrire $m = \frac{\alpha+\beta}{2}$ pour certains[23] $\alpha, \beta \in [a,b]$ tels que $\delta(\beta) = 0$, d'où les comparaisons

   $$\delta(m) = \delta\left(\frac{\alpha+\beta}{2}\right) \underset{\text{l'hypothèse}}{\overset{\delta \text{ vérifie}}{\leq}} \frac{\delta(\alpha) + \delta(\beta)}{2} \underset{\text{maximum}}{\overset{\delta(m) \text{ est}}{\leq}} \frac{\delta(m) + 0}{2},$$

   d'où $\delta(m) \leq 0$, *c. q. f. d.*

[21] La fonction δ (comme « **d**ifférence ») mesure l'écart entre le graphe de f et la corde de son graphe reliant les points d'abscisses a et b.

[22] φ comme « a**f**fine »

[23] β comme « **b**orne »

(b) **par dichotomie.** Soit $t \in [a, b]$: par dichotomie, on peut évoquer une suite $(t_n)_{n\in\mathbb{N}} \in I^{\mathbb{N}}$ tendant vers t telle que

$$\begin{pmatrix} t_0 \\ t_1 \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix} \text{ et } \forall n \in [2, \infty|[,\ \exists p, q \in [0, n|[,\ t_n = \frac{t_p + t_q}{2}.$$

Vu la continuité de δ et la tendance $t_n \longrightarrow t$, il suffit d'établir les comparaisons

$$\forall n \in \mathbb{N},\ \delta(t_n) \leq 0, \text{ ce que nous allons faire par récurrence forte.}$$

Puisque δ est nulle sur la paire $\{a, b\} = \{t_0, t_1\}$, on a bien les comparaisons ci-dessus aux rangs 0 et 1. Soit ensuite $n \in [2, \infty|[$ tel que $\forall i \in [0, n|[,\ \delta(t_i) \leq 0$ et soient $p, q \in [0, n|[$ tels que $t_n = \frac{t_p+t_q}{2}$ (permis par construction de la suite dichotomique) : on a alors les comparaisons

$$\delta(t_n) = \delta\left(\frac{t_p + t_q}{2}\right) \underset{\text{l'hypothèse}}{\overset{\delta \text{ vérifie}}{\leq}} \frac{\delta(t_p) + \delta(t_q)}{2} \underset{\text{de récurrence}}{\overset{\text{hypothèse}}{\leq}} \frac{0+0}{2} = 0, \ c.\ q.\ f.\ d.$$

Remarque – Cet exercice possède un analogue *topologique* (plus général) : lorsque E est normé (*cf.* exoEntr7 ? ?chap Evn1 ? ?), chaque partie A *fermée* de E est convexe ssi $\frac{A+A}{2} \subset A$.

2.2 Comparaisons de Jensen

Fixons pour cette section

1. un naturel $n \geq 1$;
2. un n-uplet $\overrightarrow{a}$ de points de I ;
3. un n-uplet $\overrightarrow{\lambda}$ de réels positifs de somme 1.

Rappelons lorsque f est convexe la comparaison de Jensen :

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

Signalons-en au passage une formulation *probabiliste* simple :

f est convexe ssi, pour chaque variable aléatoire[24] X finie sur I, on a la comparaison $f(\mathbb{E}(X)) \leq \mathbb{E}(f(X))$.

En pratique, ces comparaisons sont *utilisées* et l'on légitime leur utilisation en montrant la convexité/concavité par d'*autres moyens* (que nous verrons, typiquement la monotonie de la dérivée).

Exemples

[24] Les valeurs prises par X correspondent aux réels a_i et les probabilités $\mathbb{P}(X = a_i)$ correspondent aux poids λ_i.

1. Lorsque $I = \mathbb{R}_+^*$ et $f = \ln$ (qui est concave), la comparaison de Jensen donne

$$\ln\left(\sum_{i=1}^n \lambda_i a_i\right) \geq \sum_{i=1}^n \lambda_i \ln(a_i) = \ln\left(\prod_{i=1}^n a_i^{\lambda_i}\right), \text{ d'où } \prod_{i=1}^n a_i^{\lambda_i} \leq \sum_{i=1}^n \lambda_i a_i.$$

La somme $\sum_{i=1}^n \lambda_i a_i$ s'appelle la ***moyenne arithmétique*** des a_i (pondérée par les λ_i) et le produit $\prod_{i=1}^n a_i^{\lambda_i}$ leur ***moyenne géométrique***[25].

FIG : sens géom de moy géom de 2 réel>0

Quand les poids sont égaux (alors à $\frac{1}{n}$), cette comparaison s'écrit

$$\boxed{\sqrt[n]{a_1 a_2 \cdots a_n} \leq \frac{a_1+a_2+\cdots+a_n}{n}}. \text{ Très utile !}$$

Quand $n = 2$, abréger $(p, q) := \left(\frac{1}{\lambda_1}, \frac{1}{\lambda_2}\right)$ et $(A, B) := \left(\sqrt[p]{a_1}, \sqrt[q]{a_2}\right)$ donne la comparaison[26]

$$AB \leq \frac{A^p}{p} + \frac{B^q}{q}.$$

2. Imposons $I = \mathbb{R}_+^*$, soient $\alpha, \beta \in [0, 1]$ de somme 1 et soit $b \in \mathbb{R}^n$. Montrons alors la comparaison[27]

$$\sum_{i=1}^n a_i^\alpha b_i^\beta \leq \left(\sum_{i=1}^n a_i\right)^\alpha \left(\sum_{i=1}^n b_i\right)^\beta.$$

Observer que cette dernière est *homogène* en le n-uplet a, au sens où multiplier chaque a_i par un même réel strictement positif donne une comparaison équivalente. La somme des a_i se retrouvant au passage multipliée par le scalaire multiplicateur, on peut sans nuire à la généralité *imposer* (dans $\mathbb{R}_+^*$) la valeur de cette somme. Le même argument pour les b_i permet ainsi d'imposer les égalités $\sum_{i=1}^n a_i = 1 = \sum_{i=1}^n b_i$ et on en déduit alors[28] les comparaisons

$$\begin{aligned} \sum_{i=1}^n a_i^\alpha b_i^\beta \quad &\underset{\text{à } i \text{ fixé } (\alpha+\beta=1)}{\overset{\text{comp. arith.-géom.}}{\leq}} \quad \sum_{i=1}^n (\alpha a_i + \beta b_i) \;=\; \underbrace{\alpha \overbrace{\sum_{i=1}^n a_i}^{=1} + \beta \overbrace{\sum_{i=1}^n b_i}^{=1}}_{=\alpha+\beta=1} \\ &= \qquad 1^\alpha 1^\beta = \left(\sum_{i=1}^n a_i\right)^\alpha \left(\sum_{i=1}^n b_i\right)^\beta, \text{ c. q. f. d.} \end{aligned}$$

3. Déterminons l'aire maximale d'un n-gone inscrit dans le cercle unité.

Soit un tel n-gone, soit $(A_1, A_2, ..., A_n)$ une énumération *consécutive* de ses sommets et notons θ_i la mesure dans $[0, \pi]$ de l'angle $\widehat{A_i O A_{i+1}}$ pour chaque $i \in$

[25] Ainsi la comparaison ci-contre se nomme-t-elle ***comparaison arithmético-géométrique***.

[26] Publiée en 1912 par William Henry Young dans les *Proceedings of the Royal Society A*.

[27] Cette comparaison porte le nom d'Otto Ludwig Hölder et apparaît dans ses travaux en analyse fonctionnelle. Lorsque $\alpha = \beta$, on retrouve la comparaison de Cauchy-Schwarz.

[28] Ce raisonnement se généralise immédiatement à chaque nombre de n-uplets (autre que deux), *cf.* exercice d'entraînement 5 pour une application.

$[1, n|]$. L'aire du n-gone vaut alors la somme des aires des triangles A_iOA_{i+1}, à savoir

$$\sum_{i=1}^{n} \frac{1}{2} \, OA_i \cdot OA_{i+1} \cdot \sin \widehat{A_iOA_{i+1}} = \frac{n}{2} \sum_{i=1}^{n} \frac{\sin \theta_i}{n}.$$

Or la concavité de sin sur $[0, \pi]$ permet de majorer la somme ci-dessus par $\sin\left(\sum_{i=1}^{n} \frac{\theta_i}{n}\right)$ et l'égalité $\sum_{i=1}^{n} \theta_i = 2\pi$ permet finalement de majorer l'aire par $\frac{n}{2} \sin \frac{2\pi}{n}$. Le cas d'égalité étant en particulier atteint lorsque chaque θ_i vaut $\frac{2\pi}{n}$ (cas d'un n-gone régulier), on a trouvé notre aire maximale.

Sanity check : quand n tend vers ∞, un n-gone régulier "tend" vers le cercle où il s'inscrit, donc son aire tend vers l'aire π délimitée par ce disque, ce que confirme l'équivalent séquentiel

$$\frac{N}{2} \sin \frac{2\pi}{N} \underset{N \to \infty}{\sim} \frac{N}{2} \frac{2\pi}{N} = \pi.$$

4. Montrons que, dans chaque arbre binaire à n feuilles, la hauteur[29] moyenne des feuilles est minorée par $\lg_2 n$.

   On raisonne par récurrence sur n. Dans chaque arbre binaire à une feuille, cette dernière est la racine, donc l'unique hauteur d'une feuille est nulle et la hauteur moyenne vaut 0, laquelle est bien minorée par $\lg_2 1 = 0$. Dans chaque arbre binaire à deux feuilles, ces dernières sont filles de la racine, donc l'unique hauteur vaut 1 et la hauteur moyenne vaut 1, laquelle est bien minorée par $\lg_2 2 = 1$.

   Soit $b \geq 3$ un entier et soit B un arbre binaire à b feuilles. Appelons C et D les arbres binaires de racines respectives les fils de la racine de B, notons c et d les nombres respectifs de feuilles de C et D et abrégeons β, γ, δ les hauteurs moyennes d'une feuille dans B, C, D resp. Chaque feuille de B étant ou bien une feuille de C ou bien une feuille D (la racine de B est exclue car $b > 2$), d'une part le nombre de feuilles de B vaut $b = c + d$, d'autre part la somme des hauteurs[30] des feuilles de C vaut

$$\begin{aligned} b\beta &= \sum_{\spadesuit \text{ feuille de } B} \operatorname*{ht}_B(\spadesuit) = \sum_{\spadesuit \text{ feuille de } C} \operatorname*{ht}_B(\spadesuit) + \sum_{\spadesuit \text{ feuille de } D} \operatorname*{ht}_B(\spadesuit) \\ &= \sum_{\spadesuit \text{ feuille de } C} \left[1 + \operatorname*{ht}_C(\spadesuit)\right] + \sum_{\spadesuit \text{ feuille de } D} \left[1 + \operatorname*{ht}_D(\spadesuit)\right] \\ &= \sum_{\spadesuit \text{ feuille de } C} 1 + \sum_{\spadesuit \text{ feuille de } C} \operatorname*{ht}_C(\spadesuit) + \sum_{\spadesuit \text{ feuille de } D} 1 + \sum_{\spadesuit \text{ feuille de } D} \operatorname*{ht}_D(\spadesuit) \\ &= c + c\gamma + d + d\delta = b + c\gamma + d\delta. \end{aligned}$$

   Les entiers c et d valant au moins 1 et étant de somme b, ils sont chacun strictement inférieurs à b, ce qui permet d'appliquer l'hypothèse de récurrence. On utilisera par ailleurs la convexité de $t \mapsto \frac{t \ln t}{\ln 2}$ (laquelle découle de celle

[29] Dans un arbre binaire fixé, la ***hauteur*** d'un nœud donné est définie par sa distance à la racine de l'arbre.

[30] Pour chaque arbre binaire T, on a noté ht_T la fonction "hauteur dans l'arbre T". On a alors pour chaque sous-arbre binaire S de T l'égalité $\mathrm{ht}_T = \mathrm{ht}_T(\text{racine de } S) + \mathrm{ht}_S$.

de $t \mapsto t \ln t$ puisque $\ln 2$ est positif)[31] :

$$\begin{aligned} b(\beta-1) &= c\gamma + d\delta \overset{\text{hypothèse}}{\underset{\text{de récurrence}}{\geq}} c \lg_2 c + d \lg_2 d \overset{\text{Jensen pour}}{\underset{t \mapsto t \lg_2 t}{\geq}} 2 \, \frac{c+d}{2} \lg_2 \frac{c+d}{2} \\ &= b(\lg_2 b - 1), \text{ d'où } \beta \geq \lg_2 b, \ c.\ q.\ f.\ d. \end{aligned}$$

Remarque – **Coût minimal d'un algorithme de tri**. Lorsque $n = N!$ pour un certain nombre N d'objets à trier, ce résultat et la minoration $\ln N! \geq \frac{N}{2} \ln N$ (valide dès que $N \geq 2$) livrent l'interprétation suivante : *chaque algorithme de tri nécessite en moyenne au moins*[32] $C \cdot N \cdot \ln N$ *comparaisons avec* $C := \frac{1}{2 \ln 2} \simeq 0,72$.

5. *(plus difficile)* Soit $N \geq 2$ un entier et déterminons la borne supérieure[33]

$$\sup_{u \in \mathbb{U}^N} \sqrt[N(N-1)]{\prod_{i \neq j}^{i,j \in [1,N|]} |u_i - u_j|}.$$

Soit $u \in \mathbb{U}^N$. Le produit $\prod_{i \neq j}^{i,j \in [1,N|]} |u_i - u_j|$ étant invariant par permutation des u_i, on peut imposer les arguments principaux $\theta_i := \operatorname{Arg} u_i$ formant une suite strictement croissante. En notant $\binom{u_{N+i}}{\theta_{N+i}} := \binom{u_i}{\theta_i + 2\pi}$ pour chaque $i \in [1, N|]$, d'une part on peut réécrire[34]

$$\begin{aligned} \prod_{j=i+1}^{N} |u_i - u_j| \prod_{j=1}^{i-1} |u_i - u_{j+N}| &\overset{\text{reparamétrages}}{\underset{\substack{D := j-i \\ D' := j+N-i}}{=}} \prod_{D=1}^{N-i} |u_i - u_{i+D}| \prod_{D'=N-i+1}^{N-1} |u_i - u_{i+D'}| \\ &= \prod_{D=1}^{N-1} |u_i - u_{i+D}|, \end{aligned}$$

d'autre part on a l'égalité $u_s = e^{i\theta_s}$ pour chaque $s \in [1, 2N|]$ et la suite $(\theta_i)_{i \in [1,2N|]}$ croît, ce qui permet d'exprimer[35]

$$\begin{aligned} \prod_{i \neq j}^{i,j \in [1,N|]} |u_i - u_j| &= \prod_{i=1}^{N} \prod_{j \neq i}^{j \in [1,N|]} |u_i - u_j| = \prod_{i=1}^{N} \prod_{D=1}^{N-1} |u_i - u_{i+D}| \\ &= \prod_{D=1}^{N-1} \prod_{i=1}^{N} 2 \sin \frac{|\theta_i - \theta_{i+D}|}{2} \overset{(\theta_i) \text{ croît}}{=} \prod_{D=1}^{N-1} 2^N \prod_{i=1}^{N} \sin \frac{\theta_{i+D} - \theta_i}{2}. \end{aligned}$$

Or la fonction $\ln \circ \sin$ est dérivable sur $]0, \pi[$, où tombe chaque $\frac{\theta_{i+D} - \theta_i}{2}$ (distinguer soigneusement les deux cas $i + D \lesseqgtr N$), et y a pour dérivée $\frac{\sin'}{\sin} = \cot$ qui

[31] Le réel $b \geq 3$ est simplifiable car non nul.

[32] La lectrice et le lecteur retrouveront cette minoration à l'aide de celle $\ln N! \geq \int_1^N \ln$.

[33] *Culture* : étant donnée une partie $P \subset \mathbb{C}$ (non vide), la borne supérieure (lorsque A décrit les parties de P de cardinal N) des moyennes géométriques des distances entre les points distincts de A s'appelle le ***diamètre d'ordre*** N de P (quand $N = 2$, on retrouve le diamètre usuel).

[34] D comme « **d**istance » (entre numéros des sommets).

[35] *Rappel* : pour chaques réels α, β, on a l'égalité $e^{i\alpha} - e^{i\beta} = 2e^{i\frac{\alpha+\beta}{2}} \sin \frac{\alpha-\beta}{2}$.

décroît, donc est concave, d'où à $D \in [0, N|[$ fixé les comparaisons[36]

$$\begin{aligned}\prod_{i=1}^{N} \sin \frac{\theta_{i+D} - \theta_i}{2} &= \exp \sum_{i=1}^{N} \ln \sin \frac{\theta_{i+D} - \theta_i}{2} \\ \underset{\exp \text{ croît}}{\overset{\text{Jensen et}}{\leq}} \exp \left(N \ \ln \sin \sum_{i=1}^{N} \frac{1}{N} \frac{\theta_{i+D} - \theta_i}{2} \right) &= \sin^N \frac{\sum_{i=1}^{N} (\theta_{i+D} - \theta_i)}{2N}.\end{aligned}$$

Vu par ailleurs le calcul de la somme (toujours à $D \in [0, N|[$ fixé)

$$\begin{aligned}\sum_{i=1}^{N} \theta_{i+D} &= \sum_{i=1}^{N-D} \theta_{i+D} + \sum_{i>N-D}^{N} (\theta_{i+D-N} + 2\pi) \\ &\underset{\substack{j:=i+D \\ k:=i+D-N}}{\overset{\text{reparamétrage}}{=}} \sum_{j=D+1}^{N} \theta_j + \left(\sum_{k=1}^{D} \theta_k \right) + 2\pi D = 2\pi D + \sum_{i=1}^{N} \theta_i,\end{aligned}$$

mettre tout bout à bout permet de majorer

$$\begin{aligned}\sqrt[N]{\prod_{i \neq j}^{i,j \in [1,N|]} |u_i - u_j|} &\leq \sqrt[N]{\prod_{D=1}^{N-1} 2^N \sin^N \frac{2\pi D}{2N}} = \prod_{D=1}^{N-1} 2 \sin \frac{\pi D}{N} = \prod_{D=1}^{N-1} \left| 1 - e^{2\pi i \frac{D}{N}} \right| \\ &= |P(1)| \text{ où } P = \prod_{D=1}^{N-1} \left(X - e^{2\pi i \frac{D}{N}} \right) = \frac{X^N - 1}{X - 1} = \sum_{i=0}^{N-1} X^i,\end{aligned}$$

d'où la majoration[37] $\sqrt[N(N-1)]{\prod_{i \neq j}^{i,j \in [1,N|]} |u_i - u_j|} \leq \sqrt[N-1]{N}$. On a égalité lorsque pour chaque $D \in [0, N|[$ la suite $(\theta_{i+D} - \theta_i)_{i \in [1,N|]}$ est constante, ce qui est le cas lorsque la suite $(\theta_i)_{i \in [1,N|]}$ est arithmétique de raison $\frac{2\pi}{N}$, *i. e.* quand les u_i sont les sommets d'un N-gone régulier.

Conclusion : la borne supérieure cherchée vaut[38] $\sqrt[N-1]{N}$.

Exercices d'application

1. *Soient $C : \mathbb{R} \longrightarrow \mathbb{R}$ convexe continue et $s > 0$ un réel tel que $I = [0, s]$. Imposons de plus $f : I \longrightarrow \mathbb{R}$ continue. Montrer alors la comparaison*

$$C \left(\frac{1}{s} \int_0^s f \right) \leq \frac{1}{s} \int_0^s C \circ f.$$

2. *On impose $\lambda_i > 0$ pour chaque $i \in [1, n|]$. Montrer alors (en notant $\lambda_{n+1} := \lambda_1$) la comparaison*

$$\sum_{i=1}^{n} \frac{\lambda_i^2}{\lambda_i + \lambda_{i+1}} \geq \frac{1}{2} \qquad \textit{(on rappelle l'égalité } \sum_{i=1}^{n} \lambda_i = 1\textit{).}$$

[36] *Culture* : on dit que la fonction sin est ***logarithmiquement concave*** sur $]0, \pi[$.

[37] La lectrice et le lecteur connaisseurs de la comparaison d'Hadamard pourrons retrouver cette majoration bien plus rapidement en écrivant $\prod_{i \neq j}^{i,j \in [1,N|]} |a_i - a_j|$ comme un déterminant de Vandermonde.

[38] *Sanity check* : quand $N = 2$, on retrouve le diamètre du cercle unité $\sqrt[1]{2} = 2$.

3. *Pour chaque réel t, appelons* ***moyenne d'ordre*** *t (des a_i pondérés par les λ_i) le réel*

$$M_t := \sqrt[t]{\sum_{i=1}^{n} \lambda_i a_i^t} \qquad \text{si } t \neq 0\,;\ \text{sinon on définit} \qquad M_0 := \prod_{i=1}^{n} a_i^{\lambda_i}.$$

Montrer la croissance de la fonction $t \mapsto M_t$. (Rappel : les λ_i sont positifs et de somme 1.)

---

1. Utilisons des sommes d'Eudoxe-Archimède (aussi appelées sommes de Riemann). Pour chaque naturel $n > 0$, la convexité de C permet de majorer

$$\begin{aligned} C\left(\sum_{i=1}^{n} \frac{1}{n} f\left(\frac{si}{n}\right)\right) &\leq \sum_{i=1}^{n} \frac{1}{n} C\left(f\left(\frac{si}{n}\right)\right),\ i.\ e. \\ C\left(\frac{1}{s}\sum_{i=1}^{n} \frac{s}{n} f\left(\frac{si}{n}\right)\right) &\leq \frac{1}{s}\sum_{i=1}^{n} \frac{s}{n} [C \circ f]\left(\frac{si}{n}\right) \end{aligned}$$

et un passage à la limite (avec la continuité[39] de C) donne la comparaison voulue.

Remarque – Il s'agit d'une version continue des comparaisons de Jensen. La borne 0 ne joue aucun rôle (sinon de simplifier l'écriture des sommes de Riemann). Nous en verrons une autre preuve section 2.3 suivante (premier exemple).

2. Chaque quotient de gauche s'écrit $\frac{\lambda_i^2}{\lambda_i+\lambda_{i+1}} = \frac{\lambda_i}{1+\frac{\lambda_{i+1}}{\lambda_i}} = \lambda_i F(a_i)$ avec $a_i := \frac{\lambda_{i+1}}{\lambda_i}$ et $F := \left[t \mapsto \frac{1}{t+1}\right]$. Cette dernière fonction étant convexe sur $\mathbb{R}_+^*$ (car $t \mapsto \frac{1}{t}$ est convexe sur $]1, \infty[$), on peut minorer $\sum_{i=1}^{n} \lambda_i F(a_i) \geq F\left(\sum_{i=1}^{n} \lambda_i a_i\right)$. Or la somme à droite vaut

$$\sum_{i=1}^{n} \lambda_i \frac{\lambda_{i+1}}{\lambda_i} = \sum_{i=1}^{n} \lambda_{i+1} = \lambda_{n+1} + \sum_{i=1}^{n-1} \lambda_{i+1} \overset{\text{reparamétrage}}{\underset{j:=i+1}{=}} \lambda_1 + \sum_{j=2}^{n} \lambda_j = \sum_{j=1}^{n} \lambda_j = 1,$$

ce qui permet de conclure

$$\sum_{i=1}^{n} \frac{\lambda_i^2}{\lambda_i + \lambda_{i+1}} = \sum_{i=1}^{n} \lambda_i F(a_i) \geq F\left(\sum_{i=1}^{n} \lambda_i a_i\right) = F(1) = \frac{1}{1+1} = \frac{1}{2}.$$

3. Convenons de sous-entendre le domaine de sommation/production $[1, n|]$ afin d'alléger les écritures.

Soient $u < v$ deux réels strictement positifs. On a alors les équivalences

$$\begin{aligned} &M_u \leq M_v \iff \sqrt[u]{\sum \lambda_i a_i^u} \leq \sqrt[v]{\sum \lambda_i a_i^v} \\ &\iff \left(\sum \lambda_i a_i^u\right)^{\frac{v}{u}} \leq \sum \lambda_i a_i^v \iff \left(\sum \lambda_i \left[a_i^u\right]\right)^{\frac{v}{u}} \leq \sum \lambda_i \left[a_i^u\right]^{\frac{v}{u}}, \end{aligned}$$

[39] La continuité de C découle en fait de sa convexité (*cf.* section 2.4.2).

ce qu'on a par convexité de la fonction $t \mapsto t^{\frac{v}{u}}$ vu la comparaison $\frac{v}{u} > 1$.

Soient $u < v$ deux réels strictement *négatifs*. On a alors les équivalences

$$\begin{aligned} M_u \leq M_v \iff & \sqrt[u]{\sum \lambda_i a_i^u} \leq \sqrt[v]{\sum \lambda_i a_i^v} \\ \iff & \sqrt[-v]{\sum \lambda_i \left[\frac{1}{a_i}\right]^{-v}} \leq \sqrt[-u]{\sum \lambda_i \left[\frac{1}{a_i}\right]^{-u}} \iff M^*_{-v} \leq M^*_{-u} \end{aligned}$$

(où l'étoile signifie que chaque a_i a été remplacé par son inverse), ce qui découle du paragraphe précédent vu les comparaisons $0 < -v < -u$.

Il reste à comparer M_0 aux autres moyennes. On a pour chaque $t > 0$ les équivalences

$$\begin{aligned} M_0 \leq M_t \iff & \prod a_i^{\lambda_i} \leq \sqrt[t]{\sum \lambda_i a_i^t} \\ \iff & \sum \lambda_i \ln a_i \leq \frac{1}{t} \ln \left(\sum \lambda_i a_i^t\right) \iff \sum \lambda_i \ln \left[a_i^t\right] \leq \ln \left(\sum \lambda_i \left[a_i^t\right]\right), \end{aligned}$$

ce qu'on a par concavité de ln. De même, on a pour chaque $t < 0$ les équivalences

$$\begin{aligned} M_t \leq M_0 \iff & \sqrt[t]{\sum \lambda_i a_i^t} \leq \prod a_i^{\lambda_i} \\ \iff & \prod \left[\frac{1}{a_i}\right]^{\lambda_i} \leq \sqrt[-t]{\sum \lambda_i \left[\frac{1}{a_i}\right]^{-t}} \iff M^*_0 \leq M^*_{-t}, \end{aligned}$$

ce qui découle du paragraphe précédent vu la comparaison $0 < -t$.

Remarque – Lorsque les a_i sont égaux, la fonction M est constante (et vaut la valeur commune[40] des a_i), sinon des arguments de *stricte* convexité (*cf.* section 2.4.6) montreraient la *stricte* croissance de M lorsque chaque poids λ_i est non nul. Par ailleurs, de l'analyse élémentaire établirait la continuité de la fonction M en 0 ainsi que les tendances $\begin{cases} M \xrightarrow[\infty]{} \max a_i \\ M \xrightarrow[-\infty]{} \min a_i \end{cases}$. On retiendra les encadrements suivants dans le cas où les λ_i sont égaux[41], ainsi que le nom des moyennes correspondantes :

$$\begin{aligned} \min_{i \in [1,n|]} a_i \quad \leq \quad & \underbrace{\frac{n}{\frac{1}{a_1} + \frac{1}{a_2} + \cdots \frac{1}{a_n}}}_{\textit{harmonique}} \leq \underbrace{\sqrt[n]{a_1 a_2 \cdots a_n}}_{\textit{géométrique}} \\ \leq \quad & \underbrace{\frac{a_1 + a_2 + \cdots + a_n}{n}}_{\textit{arithmétique}} \leq \underbrace{\sqrt{\frac{a_1^2 + a_2^2 + \cdots + a_n^2}{n}}}_{\textit{quadratique}} \leq \max_{i \in [1,n|]} a_i \end{aligned}$$

avec égalité (dans n'importe quelle comparaison) ssi les a_i sont égaux.

[40] Chaque moyenne vaut alors la valeur commune, la moindre des attentes pour une *moyenne*.
[41] *Sanity check* : on retrouve la comparaison arithmético-géométrique.

2.3 Croissance des pentes

Exprimons à présent la convexité en termes de pentes (de cordes du graphe de f). Pour chaque $a \in I$, nous abrégerons pour toute la suite de ce cours[42]

$$\tau_a^b := \frac{f(b) - f(a)}{b - a} \text{ (pour chaque } b \neq a \text{ dans } I) \qquad \text{et} \qquad \tau_a := \left\{ \begin{array}{ccc} I \setminus \{a\} & \longrightarrow & \mathbb{R} \\ t & \longmapsto & \tau_a^t \end{array} \right. .$$

On observera dès à présent l'égalité $\tau_i^j = \tau_j^i$ pour chaques points $i \neq j$ de I.

Lemme (hors programme)

Soient $a < b < c$ dans I. On a alors les équivalences[43]

$$\tau_a^b \leq \tau_a^c \iff \tau_a^c \leq \tau_b^c \iff \tau_a^b \leq \tau_b^c \iff f(b) \leq \frac{c-b}{c-a} f(a) + \frac{b-a}{c-a} f(c).$$

FIG

Démonstration

Abrégeons $i' := f(i)$ pour chaque $i \in I$. Observer que les réels $(\lambda, \mu) := \left(\frac{c-b}{c-a}, \frac{b-a}{c-a}\right)$ font sens, sont strictement positifs, de somme 1 et vérifient l'égalité $b = \lambda a + \mu c$. À l'aide de ces considérations, on a premièrement les équivalences

$$\tau_a^b \leq \tau_a^c \iff \frac{b' - a'}{\mu(c-a)} \leq \frac{c' - a'}{c - a} \underset{c-a>0}{\overset{\mu>0}{\iff}} b' - a' \leq \mu c' - \mu a' \iff b' \leq (1-\mu) a' + \mu c',$$

deuxièmement les équivalences

$$\tau_a^c \leq \tau_b^c \iff \frac{c' - a'}{c - a} \leq \frac{c' - b'}{\lambda(c-a)} \underset{c-a>0}{\overset{\lambda>0}{\iff}} \lambda c' - \lambda a' \leq c' - b' \iff b' \leq \lambda a' + (1-\lambda) c'$$

et troisièmement les équivalences

$$\tau_a^b \leq \tau_b^c \iff \frac{b' - a'}{\mu(c-a)} \leq \frac{c' - b'}{\lambda(c-a)} \underset{c-a>0}{\overset{\lambda,\mu>0}{\iff}} \lambda b' - \lambda a' \leq \mu c' - \mu b' \iff (\lambda + \mu) b' \leq \lambda a' + \mu c'.$$

En utilisant les égalités $\left\{ \begin{array}{l} \lambda + \mu = 1 \\ b = \lambda a + \mu c \end{array} \right.$, on voit que chacune des comparaisons ci-dessus équivaut à $b' \leq \lambda a' + \mu c'$, ce qui conclut.

Corollaire (convexité & croissance des pentes des cordes)

Les assertions suivantes sont équivalentes :

1. *la fonction f est convexe ;*
2. *pour chaques points a, b, c de I, on a l'implication*

$$a < b < c \Longrightarrow \tau_a^b \leq \tau_a^c \leq \tau_b^c ;$$

FIG

[42] τ comme « **tau**x d'accroissement »

[43] Un dessin doit convaincre de l'équivalence de la quatrième comparaison avec chacune des trois premières.

3. *pour chaque point $a \in I$, la fonction τ_a croît*[44] *;*

FIG

Démonstration

Nous abrégerons $i' := f(i)$ pour chaque $i \in I$.

$\boxed{1 \Longrightarrow 2}$ Soient $a < b < c$ dans I. D'après le lemme, les comparaisons désirées équivalent chacune à $b' \leq \lambda a' + \mu c'$ avec $(\lambda, \mu) := \left(\frac{c-b}{c-a}, \frac{b-a}{c-a}\right)$. En se souvenant que les réels λ et μ tombent dans $]0,1[$ et vérifient les égalités $\begin{cases} \lambda + \mu = 1 \\ b = \lambda a + \mu c \end{cases}$, la convexité de f permet de conclure.

$\boxed{2 \Longrightarrow 3}$ Soit $a \in I$ et soient $u < v$ deux points de $I \setminus \{a\}$. Il suffit pour conclure d'établir la comparaison $\tau_a^u \leq \tau_a^v$. Notons pour cela $A < B < C$ les réels a, u, v rangés dans l'ordre *croissant* et discutons selon la position de a par rapport à u et v :

a. si $a < u$, *i. e.* si $(A, B, C) = (a, u, v)$, l'hypothèse $\tau_A^B \leq \tau_A^C$ s'écrit alors $\tau_a^u \leq \tau_a^v$, *c. q. f. d.*;
b. si $u < a < v$, *i. e.* si $(A, B, C) = (u, a, v)$, l'hypothèse $\tau_A^B \leq \tau_B^C$ donne alors $\tau_u^a \leq \tau_a^v$, *i. e.* $\tau_a^u \leq \tau_a^v$, *c. q. f. d.*;
c. si enfin $a > v$, *i. e.* si $(A, B, C) = (u, v, a)$, l'hypothèse $\tau_A^C \leq \tau_B^C$ s'écrit alors $\tau_u^a \leq \tau_v^a$, *i. e.* $\tau_a^u \leq \tau_a^v$, *c. q. f. d.*

$\boxed{3 \Longrightarrow 1}$ Soient $a < b$ dans I, soient $\lambda, \mu \in]0,1[$ de somme 1 et abrégeons $c := \lambda a + \mu b$. Vu les égalités

$$\begin{cases} c - a = (\lambda a + \mu b) - a = \mu b - (1 - \lambda) a = \mu (b - a) \\ c - b = (\lambda a + \mu b) - b = \lambda a - (1 - \mu) b = \lambda (a - b) \end{cases}, \ i.\ e.\ (\lambda, \mu) = \left(\frac{b-c}{b-a}, \frac{c-a}{b-a}\right),$$

on peut utiliser le lemme et la comparaison désirée $c' \leq \lambda a' + \mu b'$ équivaut à $\tau_a^c \leq \tau_a^b$, ce qu'on a par croissance de τ_a vu la comparaison $c < b$.

Corollaire (convexité & croissance des pentes des tangentes, position relative du graphe et des tangentes)

Imposons f dérivable. Alors f est convexe ssi f' croît et l'on a dans ce cas les comparaisons[45]

$$\forall a, t \in I, \ f(t) \geq (t - a) f'(a) + f(a).$$

FIG

Démonstration

La dérivabilité de f permet pour chaque $a \in I$ de prolonger par continuité la fonction τ_a en imposant $\tau_a^a := f'(a)$.

Supposons f convexe et soient $a < b$ dans I. Observer que les croissances de τ_a et τ_b se propagent par continuité à tout l'intervalle I. On a alors les comparaisons

$$f'(a) = \tau_a(a) \overset{\tau_a \text{ croît}}{\leq} \tau_a(b) = \tau_b(a) \overset{\tau_b \text{ croît}}{\leq} \tau_b(b) = f'(b), \text{ d'où la croissance de } f'.$$

[44] En termes "pentus" : les pentes des cordes de f croissent.
[45] En d'autres termes : le graphe de f est au-dessus de chacune de ses tangentes.

Supposons f' croissante et soient $a < b < c$ dans I. D'après le lemme et la proposition précédente, il suffit pour conclure à la convexité de f de montrer la comparaison $\tau_a^b \leq \tau_b^c$. Or la formule des accroissements finis nous livre deux réels $\begin{array}{r} u \in \left]a, b\right[\\ v \in \left]b, c\right[\end{array}$ tels que $\left\{ \begin{array}{l} \tau_a^b = f'(u) \\ \tau_b^c = f'(v) \end{array} \right.$ et la croissance de f' conclut vu la comparaison $u < v$.

Soient $a \neq t$ dans I. La croissance de τ_a donne alors la positivité du taux de variation $\frac{\tau_a(t) - \tau_a(a)}{t-a}$, laquelle après multiplication par le positif $(t-a)^2$ s'écrit $0 \leq [f(t) - f(a)] - f'(a)(t-a)$. Cette comparaison restant valide quand $t = a$, on a terminé.

Remarques

• Coincer le graphe d'une fonction convexe entre une corde et une tangente permet d'éclairer géométriquement de nombreuses comparaisons. À exploiter sans retenue !

• *Sanity checks* : remplacer a par 0 et f par l'une des trois fonction exp, $u \mapsto \ln(1+u)$ ou sin redonne les comparaisons classiques

$$\forall t \in \mathbb{R},\ e^t \geq 1 + t, \qquad \forall u \in \left]-1, \infty\right[,\ \ln(1+u) \leq u \qquad \text{et} \qquad \forall \theta \in [0, \pi],\ \sin\theta \leq \theta.$$

FIG

• *Mnémonique* : pour retenir sans effort le signe du premier terme non linéaire des développements limités des fonctions convexes/concaves usuelles, il suffit de tracer leur graphe au voisinage du point considéré ainsi que la tangente associée. Par exemple, on retrouvera les égalités locales

$$e^t \underset{t\sim 0}{=} 1 + t + \frac{t^2}{2} + o\left(t^2\right), \quad \ln(1+u) \underset{u\sim 0}{=} u - \frac{u^2}{2} + o\left(u^2\right) \quad \text{et} \quad \sin\theta \underset{\theta\sim 0}{=} \theta - \frac{\theta^3}{6} + o\left(\theta^3\right).$$

• Le sens réciproque du corollaire peut être montré sans le lemme. Soient $a < b$ dans I et notons φ la fonction affine coïncidant avec f en a et en b : il s'agit de montrer la négativité sur $[a, b]$ de la différence $\delta := f - \varphi$. La fonction δ s'annulant en a et en b, sa dérivée s'annule par Rolle en un certain $m \in \left]a, b\right[$. Or cette dérivée $\delta' = f' - \tau_a^b$ croît, donc est négative sur $\left]a, m\right[$ et positive sur $\left]m, b\right[$, donc δ décroît sur $\left]a, m\right[$ et croît sur $\left]m, b\right[$; puisque δ s'annule en a et en b, elle reste négative sur $[a, b]$, *c. q. f. d.*

• En pratique, il arrive souvent que la fonction considérée ne soit pas dérivable à une extrémité de son intervalle de définition. Cela n'est en fait pas un problème, l'exercice d'application 2 section 2.1 permettant d'affirmer que

si f est continue sur I et dérivable sur $\mathring{I}$, alors
f est convexe sur I ssi f' croît sur l'intérieur $\mathring{I}$.

Exemple : soit $C : \mathbb{R} \longrightarrow \mathbb{R}$ convexe, soient $a < b$ dans I et abrégeons $\binom{S}{\ell} := \binom{[a,b]}{b-a}$. Le graphe de C est alors au-dessus de sa tangente en $m := \frac{1}{\ell}\int_S f$, d'où (en composant à droite par f) la minoration $C \circ f \geq (f - m)\, C'(m) + C(m)$. La fonction $f - m$ étant de moyenne nulle, appliquer $\frac{1}{\ell}\int_S$ livre la comparaison[46]

$$\frac{1}{\ell}\int_S C \circ f \geq C\left(\frac{1}{\ell}\int_S f\right).$$

[46] *Sanity check* : lorsque $S = [0, s]$, on retrouve la version continue de Jensen établie à l'exercice d'application 1 section 2.2.

Exemples (ceux de la section 2.1)

1. L'exponentielle et le cosinus hyperbolique sont dérivables sur $\mathbb{R}$ où leur dérivées respectives exp et sh croissent, donc sont convexes sur $\mathbb{R}$. De même, le logarithme est dérivable sur $\mathbb{R}_+^*$ où sa dérivée $t \mapsto \frac{1}{t}$ décroît, donc est concave sur $\mathbb{R}_+^*$.
2. La fonction sin est dérivable sur $[0, \pi]$ où sa dérivée cos décroît, donc est concave sur $[0, \pi]$. On montrerait de même la concavité de cos sur $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$.
3. Soit $\alpha \geq 1$ et imposons $f = [t \mapsto t^\alpha]$, laquelle est définie et dérivable sur $\mathbb{R}_+$. Sa dérivée $f' = \left[t \mapsto \alpha t^{\alpha-1}\right]$ est alors croissante (car l'exposant $\alpha - 1$ est positif et le coefficient α également), donc f est convexe sur $\mathbb{R}_+$. On montrerait de même la convexité sur $\mathbb{R}_+^*$ de $t \mapsto t^\beta$ pour chaque réel $\beta < 0$.
4. Sur $\mathbb{R}_+$, les fonctions th, arctan et argsh sont dérivables et dérivées respectives $1 - \mathrm{th}^2$, $t \mapsto \frac{1}{1+t^2}$ et $s \mapsto \frac{1}{\sqrt{1+s^2}}$, chacune décroissante[47], d'où la concavité de ces trois fonctions.
5. Sur $\left[0, \frac{\pi}{2}\right[$, la fonction tan est dérivable de dérivée $1 + \tan^2$ croissante, donc est convexe. De même sur $[0, 1[$ pour argth (de dérivée $t \mapsto \frac{1}{1-t^2}$ croissante).
6. Soit $n \in \mathbb{N}$ et imposons $f = [t \mapsto t^n]$, laquelle est définie et dérivable sur $\mathbb{R}$. La dérivée $f' = \left[t \mapsto nt^{n-1}\right]$ est alors croissante sur $\mathbb{R}_+$ et $\begin{cases} \text{croissante sur } \mathbb{R}_- \text{ si } n-1 \text{ impair} \\ \text{décroissante sur } \mathbb{R}_- \text{ si } n-1 \text{ pair} \end{cases}$, donc la fonction f est convexe sur $\mathbb{R}_+$ et $\begin{cases} \text{convexe sur } \mathbb{R}_- \text{ si } n \text{ pair} \\ \text{concave sur } \mathbb{R}_- \text{ si } n \text{ impair} \end{cases}$.
7. La fonction argch est continue sur $[1, \infty[$ et dérivable sur l'intérieur $]1, \infty[$ de dérivée $c \mapsto \frac{1}{\sqrt{c^2-1}}$ décroissante, donc est concave sur $[1, \infty[$.
8. La fonction arcsin est continue sur $[0, 1]$ et dérivable sur l'intérieur $[0, 1[$ de dérivée $s \mapsto \frac{1}{\sqrt{1-s^2}}$ croissante, donc est convexe sur $[0, 1]$. De même, la fonction $\arccos = \frac{\pi}{2} - \arcsin$ est concave sur $[0, 1]$.
9. Soit $\gamma \in [0, 1]$ et imposons $f = [t \mapsto t^\gamma]$, laquelle est continue sur $\mathbb{R}_+$ et dérivable sur $\mathbb{R}_+^*$. Sa dérivée $f' = \left[t \mapsto \gamma t^{\gamma-1}\right]$ est alors décroissante (car l'exposant $\gamma - 1$ est négatif et le coefficient γ positif), donc f est concave sur $\mathbb{R}_+$.
10. Imposons $f = [t \mapsto t \ln t]$, laquelle est définie et dérivable sur $\mathbb{R}_+^*$, prolongée continûment par $0 \mapsto 0$. Pour chaque $t > 0$, on a les égalités
$$f'(t) = \frac{\partial}{\partial t} f(t) = \frac{\partial}{\partial t}(t \ln t) = \ln t + t\frac{1}{t} = \ln t + 1,$$
donc la dérivée $f' = \ln + 1$ croît sur $\mathbb{R}_+^*$, d'où la convexité sur $\mathbb{R}_+$ de la fonction f.

Corollaire (convexité & signe de la dérivée seconde)

Imposons f deux fois dérivable. Alors f est convexe ssi $f'' \geq 0$ (et concave ssi $f'' \leq 0$).

[47] Ces monotonies s'établissent par composition de fonctions monotones, *surtout pas* en dérivant à nouveau ! Par exemple, on pourra écrire $\mathrm{argsh}' = i \circ r \circ q$ où la fonction **q**uadratique $q := \left[t \mapsto 1 + t^2\right]$ croît de $\mathbb{R}_+$ vers $[1, \infty[$, la fonction "**r**acine" $r := \left[t \mapsto \sqrt{t}\right]$ croît de $[1, \infty[$ dans $[1, \infty[$ et la fonction "**i**nverse" $i := \left[t \mapsto \frac{1}{t}\right]$ décroît sur $[1, \infty[$.

Démonstration : la fonction f'' faisant sens, il est sensé d'utiliser la caractérisation de la croissance de f' en fonction du signe de sa dérivée f''.

Remarque – Ne pas abuser de ce corollaire : il est souvent plus simple d'écrire f' comme composée de fonctions monotones "simples", comme nous l'avons fait dans les exemples précédents.

Exercices d'application

1. *Étudier la convexité des fonctions* $t \mapsto \ln(1+e^t)$ *et* $t \mapsto \sqrt{1+t^2}$.
2. *Imposons* f *convexe. Montrer alors que* f *décroît sur un certain intervalle (éventuellement vide) puis croît sur un certain intervalle (éventuellement vide). On pourra considérer l'ensemble des* $a \in I$ *tel que* $\tau_a \geq 0$ *à droite de* a.
3. *Soit* ABC *un triangle non aplati. On note* α, β, γ *les mesures de ses angles dans* $]0, \pi[$. *Montrer la comparaison*

$$\frac{1}{\sin\alpha} + \frac{1}{\sin\beta} \geq \frac{8}{3 + 2\cos\gamma}.$$

---

1. Notons ℓ et r les fonctions respectives de l'énoncé[48], chacune définie et infiniment dérivable sur $\mathbb{R}$.

   Vu à $a \in \mathbb{R}$ fixé les égalités

$$\ell'(a) = \frac{\partial}{\partial a} \ln(1+e^a) = \frac{e^a}{1+e^a} = \frac{1}{e^{-a}+1},$$

   la dérivée $\ell' = \left[t \mapsto \frac{1}{t}\right] \circ \left[t \mapsto 1 + e^{-t}\right]$ est composée de deux fonctions décroissantes, donc croît, d'où la convexité de ℓ.

   Vu pour chaque réel $a > 0$ les égalités

$$r'(a) = \frac{\partial}{\partial a}\sqrt{1+a^2} = \frac{2a}{2\sqrt{1+a^2}} = \frac{|a|}{\sqrt{1+a^2}} = \sqrt{\frac{a^2}{1+a^2}} = \sqrt{\frac{1}{a^{-2}+1}},$$

   la dérivée $r'_{|\mathbb{R}_+^*} = \underbrace{\left[q \mapsto \sqrt{q}\right]}_{\text{croît}} \circ \underbrace{\left[c \mapsto \frac{1}{1+c}\right]}_{\text{décroît}} \circ \underbrace{\left[t \mapsto t^{-2}\right]}_{\text{décroît sur } \mathbb{R}_+^*}$ croît sur sur $\mathbb{R}_+^*$.

   De même, pour chaque réel $a < 0$, l'égalité $a = -|a|$ permet en reprenant le calcul ci-dessus de décomposer

$$r'_{|\mathbb{R}_-^*} = \underbrace{\left[q \mapsto -\sqrt{q}\right]}_{\text{décroît}} \circ \underbrace{\left[c \mapsto \frac{1}{1+c}\right]}_{\text{décroît}} \circ \underbrace{\left[t \mapsto t^{-2}\right]}_{\text{croît sur } \mathbb{R}_-^*},$$

   d'où la croissance sur $\mathbb{R}_-^*$ de r'. Cette dérivée étant par ailleurs continue en 0, elle croît sur tout $\mathbb{R}$, d'où la convexité de r.

[48] ℓ et r comme « **l**ogarithme » et « **r**acine »

Remarque – On aurait également pu, afin d'éviter la discussion de signe, dériver r' et obtenir

$$\begin{aligned} r''(a) &= \frac{\partial}{\partial a} r'(a) = \frac{\partial}{\partial a}\frac{a}{\sqrt{1+a^2}} = \frac{\sqrt{1+a^2} - a\frac{a}{\sqrt{1+a^2}}}{\sqrt{1+a^2}^2} \\ &= \frac{\left(1+a^2\right)-a^2}{\sqrt{1+a^2}^3} = \left(1+a^2\right)^{-\frac{3}{2}} \geq 0. \end{aligned}$$

2. Notons comme suggéré $I^+ := \{a \in I \ ; \ \tau_a \geq 0 \text{ sur } I \cap]a,\infty[\}$ et soient $u < v$ dans I^+.

En notant $m := \frac{u+v}{2}$, on a alors d'une part $v > m$, d'où la minoration $\tau_u^v \geq \tau_u^m$ (car τ_u croît), d'autre part $m > u$, d'où la minoration $\tau_u^m \geq 0$ (car $u \in I^+$). Il en résulte la croissance de f sur I^+.

Soit par ailleurs $i \in [u,v]$ et soit $t > i$ dans I. On a alors (comme ci-dessus) les minorations $\tau_i^t \geq \tau_u^t \geq 0$, d'où l'appartenance $i \in I^+$. Il en résulte la convexité de I^+.

On montrerait de même que la partie $I^- := \{a \in I \ ; \ \tau_a \leq 0 \text{ sur } I \cap]-\infty,a[\}$ est un intervalle sur lequel f décroît.

Montrons enfin que I^+ et I^- recouvrent I. Soit $a \in I \backslash I^-$ et soit $i < a$ dans I tel que $\tau_a^i > 0$. On a alors pour chaque $t > a$ dans I les majorations $\tau_a^t \geq \tau_a^i \geq 0$, d'où l'appartenance $a \in I^+$.

Remarque – On montrerait sans peine que les intervalles $I^\pm$ sont fermés et que leur intersection est un intervalle fermé sur lequel f est constante.

3. Regardons la fonction $C := \frac{1}{\sin}$ intervenant dans la comparaison à établir[49].

FIG

Le tracé de son graphe nous incite à croire qu'elle est convexe : montrons-le. La fonction C est infiniment dérivable sur $]0,\pi[$ et a pour dérivée

$$C' = \left[\frac{1}{\sin}\right]' = \frac{-\sin'}{\sin^2} = -\frac{\cos}{1-\cos^2} = \underbrace{[-\mathrm{Id}]}_{\text{décroît}} \circ \left[c \mapsto \frac{c}{1-c^2}\right] \circ \underbrace{\cos}_{\text{décroît}},$$

donc C' varie sur $\mathbb{R}$ comme $c \mapsto \frac{c}{1-c^2}$ varie sur $]-1,1[$. Or d'une part le quotient $\frac{c}{1-c^2}$ se réécrit $\frac{-1}{c-\frac{1}{c}}$ lorsque $c \neq 0$, d'autre part la fonction $\left[d \mapsto \frac{-1}{d}\right] \circ \left[t \mapsto t - \frac{1}{t}\right]$ croît sur $]-1,0[$ et sur $]0,1[$ comme composée de fonctions croissantes[50] ; étant par ailleurs continue en 0, la fonction $c \mapsto \frac{c}{1-c^2}$ croît sur tout $]-1,1[$, d'où la croissance de C' et la convexité de C.

On peut par conséquent utiliser une comparaison de Jensen :

$$\begin{aligned} \frac{1}{2}\frac{1}{\sin\alpha} + \frac{1}{2}\frac{1}{\sin\beta} &= \frac{C(\alpha)+C(\beta)}{2} \\ &\overset{C \text{ convexe}}{\geq} C\left(\frac{\alpha+\beta}{2}\right) \overset{\alpha+\beta+\gamma=\pi}{=} \frac{1}{\sin\left(\frac{\pi-\gamma}{2}\right)} = \frac{1}{\cos\frac{\gamma}{2}}. \end{aligned}$$

[49] Il serait possible de résoudre cet exercice entièrement de manière géométrique à l'aide de formules trigonométriques pertinentes.

[50] La fonction $t \mapsto t - \frac{1}{t}$ croît sur $\mathbb{R}_-^*$ et sur $\mathbb{R}_+^*$ comme somme de deux fonctions croissantes.

Il suffit pour conclure d'établir la comparaison $\frac{1}{\cos\frac{\gamma}{2}} \geq \frac{4}{3+2\cos\gamma}$, ce qui s'écrit aussi $\frac{1}{c} \geq \frac{4}{3+2(2c^2-1)}$ en abrégeant $c := \cos\frac{\gamma}{2}$. Or cette comparaison (à un seul paramètre dans $]0,1[$) découle des équivalences

$$\frac{1}{c} \geq \frac{4}{3+2\left(2c^2-1\right)} \iff 3+2\left(2c^2-1\right) \geq 4c$$
$$\iff 4c^2-4c+1 \geq 0 \iff (2c-1)^2 \geq 0$$

et de la positivité du carré $(2c-1)^2$.

Remarques – La croissance de C' aurait pu s'obtenir en dérivant directement (à $c \in]-1,1[$ fixé)

$$\frac{\partial}{\partial c}\left(\frac{c}{1-c^2}\right) = \frac{\left(1-c^2\right)-c\left(-2c\right)}{\left(\text{dénominateur}\right)^2} = \frac{1+2c^2}{\left(\text{dén.}\right)^2} \underset{\text{est positif}}{\overset{\text{chaque carré}}{\geq}} 0.$$

On aurait également pu utiliser la convexité de la fonction "inverse" $i := t \mapsto \frac{1}{t}$: il serait venu

$$\frac{1}{\sin\alpha}+\frac{1}{\sin\beta} = 2\frac{i\left(\sin\alpha\right)+i\left(\sin\beta\right)}{2} \geq 2i\left(\frac{\sin\alpha+\sin\beta}{2}\right) = \frac{2}{\sin\frac{\alpha+\beta}{2}\cos\frac{\alpha-\beta}{2}}$$

et, en minorant le $\frac{1}{\cos}$ par 1, on serait retombé sur le $\frac{2}{\cos\frac{\gamma}{2}}$ de la solution ci-dessus.

2.4 Compléments hors programme

2.4.1 Stabilité de la convexité par diverses opérations

Propriétés (convexité, combinaisons linéaires, bornes supérieures, composées et réciproques)

1. *L'opposé de chaque fonction convexe est concave.*
2. *Chaque combinaison linéaire à coefficients <u>positifs</u> de fonctions convexes reste convexe*[51].
3. *Lorsqu'elle fait sens, la borne <u>supérieure</u> de chaque ensemble de fonctions convexes reste convexe.*
4. *Pour chaque bijection φ <u>affine</u>, la fonction f est convexe sur I ssi*[52] *la fonction $f\circ\varphi$ est convexe sur l'intervalle $\varphi^{-1}(I)$.*
5. *Lorsqu'elle fait sens, chaque composée de fonctions convexes <u>croissantes</u> reste convexe.*
6. *Si f est convexe et bijective, alors f^{-1} est convexe si f <u>dé</u>croît (et concave si f croît).*

FIG

Exemples typiques : $t\mapsto t^2$ et $t\mapsto\sqrt{t}$ sur $[0,1]$, $t\mapsto\frac{1}{t}$, $t\mapsto\sqrt{1-t^2}$.

[51] *Culture* : on dit que les fonctions convexes sur I forment un ***cône*** de $\mathbb{R}^I$.
[52] La convexité est dite *invariante par reparamétrage affine.*

Démonstration Soient $a,b\in I$ et soient $\lambda,\mu\in[0,1]$ de somme 1.

1. C'est une définition.
2. Une récurrence immédiate permet de traiter seulement le cas de *deux* fonctions. Imposons donc f convexe, soit $g:I\longrightarrow\mathbb{R}$ convexe et soit $\rho\in\mathbb{R}_+$. Puisque f est convexe, on a la comparaison $f(\lambda a+\mu b)\leq\lambda f(a)+\mu f(b)$; puisque ρ est positif, on peut multiplier cette dernière par ρ ; ajouter enfin la comparaison $g(\lambda a+\mu b)\leq\lambda g(a)+\mu g(b)$ découlant de la convexité de g justifie la comparaison

$$\begin{aligned}[\rho f+g](\lambda a+\mu b) &= \begin{matrix}\rho f(\lambda a+\mu b)\\ +g(\lambda a+\mu b)\end{matrix} \leq \begin{matrix}\rho(\lambda f(a)+\mu f(b))\\ +(\lambda g(a)+\mu g(b))\end{matrix}\\ &= \begin{matrix}\lambda\rho f(a) & +\mu\rho f(b)\\ +\lambda g(a) & +\mu g(b)\end{matrix} = \lambda\begin{pmatrix}\rho f(a)\\ +g(a)\end{pmatrix}+\mu\begin{pmatrix}\rho f(b)\\ +g(b)\end{pmatrix}\\ &= \lambda\ [\rho f+g](a)+\mu\ [\rho f+g](b),\ c.\ q.\ f.\ d.\end{aligned}$$

3. Soit $\mathcal{F}$ un ensemble de fonctions convexes sur I. On impose que la fonction $F:=\sup_{f\in\mathcal{F}}f$ fasse sens, *i. e.* est une fonction réelle définie sur I (cela impose $\mathcal{F}\neq\emptyset$) On a alors pour chaque $f\in\mathcal{F}$ les comparaisons

$$f(\lambda a+\mu b)\underset{\text{convexe}}{\overset{f\text{ est}}{\leq}}\lambda f(a)+\mu f(b)\underset{\lambda,\mu\geq0}{\overset{f\leq F}{\leq}}\lambda F(a)+\mu F(b),$$

d'où (en faisant varier f) la majoration $\sup_{f\in\mathcal{F}}f(\lambda a+\mu b)\leq\lambda F(a)+\mu F(b)$, *i. e.* $F(\lambda a+\mu b)\leq\lambda F(a)+\mu F(b)$, *c. q. f. d.*

Remarque – Cela est faux pour les bornes *inf*érieures : si $I=[0,2]$, alors $\inf\{1,\mathrm{Id}\}$ est concave :

FIG

4. Soit φ comme dans l'énoncé et imposons $f\circ\varphi$ convexe. En notant $\binom{\alpha}{\beta}:=\binom{\varphi^{-1}(a)}{\varphi^{-1}(b)}$ (réels qui tombent dans $\varphi^{-1}(I)$), on a alors les comparaisons

$$\begin{aligned}&f(\lambda a+\mu b)=f(\ \lambda\varphi(\alpha)+\mu\varphi(\beta)\ )\underset{\text{affine}}{\overset{\varphi\text{ est}}{=}}f(\ \varphi(\lambda\alpha+\mu\beta)\ )\\ =\ &[f\circ\varphi](\lambda\alpha+\mu\beta)\underset{\text{convexe}}{\overset{f\circ\varphi\text{ est}}{\leq}}\lambda[f\circ\varphi](\alpha)+\mu[f\circ\varphi](\beta)\\ =\ &\lambda f(\varphi(\alpha))+\mu f(\varphi(\beta))=\lambda f(a)+\mu f(b),\end{aligned}$$

d'où la convexité de f. Pour le sens réciproque, appliquer le sens direct que l'on vient d'établir en remplaçant (I,f,φ) par $\left(\varphi^{-1}(I),f\circ\varphi,\varphi^{-1}\right)$.

Remarque – Le résultat est faux si φ n'est pas affine : la fonction $t\mapsto t^2$ convexe sur $[0,1]$ devient concave après composition à droite par la bijection $t\mapsto\sqrt[4]{t}$.

5. Une récurrence immédiate permet de traiter juste le cas de *deux* fonctions. Imposons donc f convexe et soit g convexe croissante sur $\operatorname{Im}f$. Puisque f est convexe, on a la comparaison $f(\lambda a+\mu b)\leq\lambda f(a)+\mu f(b)$; puisque g croît, on en déduit la comparaison $g(f(\lambda a+\mu b))\leq g(\lambda f(a)+\mu f(b))$; enfin,

la convexité de g sur $f(I)$ permet de majorer ce dernier terme par $\lambda g(f(a)) + \mu g(f(b))$. Finalement, on obtient la comparaison

$$[g \circ f](\lambda a + \mu b) \leq \lambda \ [g \circ f](a) + \mu \ [g \circ f](b), \ c.\ q.\ f.\ d.$$

Remarque – La croissance du composé tout à droite est inutile. En revanche, celle des composés de gauche est nécessaire : dès que f est convexe et *négative*, la composée $|f| = -f$ par la fonction convexe "module" est *concave* (donc non convexe si f n'est pas affine, par exemple quand $f = \mathrm{th}_{|\mathbb{R}_-}$).

6. Imposons f convexe, bijective et décroissante (le cas croissant se traiterait de façon analogue). Soient α, β dans $f(I)$ et notons $\binom{u}{v} := \binom{f^{-1}(\alpha)}{f^{-1}(\beta)}$. On a alors les comparaisons

$$\lambda\alpha + \mu\beta = \lambda f(u) + \mu f(v) \underset{\text{convexe}}{\overset{f \text{ est}}{\geq}} f(\lambda u + \mu v),$$

d'où (par décroissance de f^{-1}) les majorations

$$f^{-1}(\lambda\alpha + \mu\beta) \leq f^{-1}(f(\lambda u + \mu v)) = \lambda u + \mu v = \lambda f^{-1}(\alpha) + \mu f^{-1}(\beta), \ c.\ q.\ f.\ d.$$

Remarques

- La convexité et la concavité étant des notions "duales" (d'après le point 1), la lectrice et le lecteur énonceront et établiront aisément les analogues concaves des propriétés précédentes.
- D'après le point 2, la convexité est préservée par addition. Elle n'est toutefois pas préservée par multiplication (même positive croissante) : la fonction $t \mapsto t^{\frac{2}{3}}$ est concave sur $[0, 1]$ (et positive croissante) mais son carré $t \mapsto t^{\frac{4}{3}}$ est convexe (non affine) :

FIG

- Le point 4 exprime que la convexité est une propriété qui ne dépend *ni de la hauteur du graphe* de l'application considérée (terme constant de φ) *ni du sens dans lequel on parcourt ce graphe*[53] (signe de la pente de φ) *ni de la vitesse constante avec laquelle on le parcourt* (module de la pente φ). Attention, la convexité peut se perdre si la vitesse n'est plus constante (*cf.* contre-exemple en remarque dans la preuve ci-dessus).
- *Sanity check* : la fonction $-\sin$ est convexe sur $]0, \pi[$, elle prend ses valeurs dans $\mathbb{R}_-^*$ où croît et est convexe la fonction $t \mapsto -\frac{1}{t}$, donc la composée $\frac{1}{\sin}$ est bien convexe sur $]0, \pi[$ (comme nous l'avions montré directement à l'exercice d'application 3 section 2.2).
- La convexité n'est pas préservée par recollement : recoller deux constantes distinctes l'une sur $\mathbb{R}_-$ et l'autre sur $\mathbb{R}_+$ ne donne une fonction ni convexe ni concave :

FIG

Et même par recollement *continu* : prolonger par parité n'importe quelle fonction convexe strictement croissante sur $\mathbb{R}_-$ (par exemple exp) :

FIG

[53] *Cas particulier* : quand $I = \mathbb{R}$, la fonction f est convexe ssi la fonction $t \mapsto f(-t)$ est convexe.

Exercices d'application

Une fonction F est dite ***logarithmiquement convexe*** *(en abrégé* ***log-convexe****) si* $\ln \circ F$ *fait sens et est convexe.*

1. (a) *Montrer que f est log-convexe ssi $f > 0$ et si l'on a les comparaisons*

$$f(\lambda a + \mu b) \leq f(a)^{\lambda} f(b)^{\mu} \quad \text{pour chaques } a < b \text{ dans } I \text{ et pour chaques } \lambda, \mu \in [0,1] \text{ de somme } 1.$$

   (b) *Montrer que la fonction* ch *est log-convexe sur $\mathbb{R}$ et que, à $\alpha \in \mathbb{R}$ fixé, la fonction $t \mapsto t^{\alpha}$ est log-convexe sur $\mathbb{R}_+^*$ ssi $\alpha \leq 0$.*

   (c) *Montrer que chaque fonction log-convexe est convexe.*

   (d) *Montrer que f est log-convexe ssi $f > 0$ et si $t \mapsto f(t) C^t$ est convexe pour chaque réel $C > 0$.*

2. *Soit F log-convexe sur $\mathbb{R}_+$ vérifiant $F(0) = 1$ et $\forall t \geq 1,\ F(t) = tF(t-1)$.*

   (a) *Montrer que F coïncide sur $\mathbb{N}$ avec la fonction "factorielle".*

   (b) *Soient $n \in \mathbb{N}^*$ et $t \in \,]0,1]$. En exploitant les comparaisons des pentes $\tau_{n-1}^{n} \leq \tau_n^{n+t} \leq \tau_n^{n+1}$, montrer la tendance*

$$\frac{N!\ N^t}{(t+1)(t+2)\cdots(t+N)} \underset{N\to\infty}{\longrightarrow} F(t).$$

   (c) *En déduire l'unicité d'un F comme évoqué.*

---

1. (a) La composée $\ln \circ f$ fait sens ssi f prend ses valeurs dans le domaine de définition de ln, *i. e.* ssi $f > 0$. Par ailleurs, ayant évoqué a, b, λ, μ comme dans l'énoncé, la comparaison de Jensen s'écrit

$$\ln(f(\lambda a + \mu b)) \leq \lambda \ln(f(a)) + \mu \ln(f(b)) = \ln\left(f(a)^{\lambda} f(b)^{\mu}\right),$$

   donc équivaut à celle voulue en appliquant l'exponentielle (laquelle croît).

   Remarque – **Culture hors programme**. Ces comparaisons sont parfois prises comme définition de la log-convexité afin d'inclure des fonctions s'annulant.

   (b) La fonction ch est strictement positive et dérivable sur $\mathbb{R}$, donc $\ln \circ \mathrm{ch}$ fait sens. Sa dérivée vaut $\frac{\mathrm{ch}'}{\mathrm{ch}} = \frac{\mathrm{sh}}{\mathrm{ch}} = \mathrm{th}$ qui croît, d'où la log-convexité désirée.

   Soit $\alpha \in \mathbb{R}$. La fonction $t \mapsto t^{\alpha}$ est strictement positive sur $\mathbb{R}_+^*$ et son logarithme vaut $\alpha \ln$ qui est concave pour $\alpha \geq 0$ et convexe sinon, *c. q. f. d.*

   (c) Supposons $\ln \circ f$ convexe. Composer à gauche par la fonction convexe croissante exp préserve la convexité d'après notre preuve du point (5), d'où la convexité de $\exp \circ \ln \circ f = f$.

   Sanity checks : la fonction ch est convexe sur $\mathbb{R}$, la fonction $t \mapsto t^{\alpha}$ est convexe pour chaque réel $\alpha < 0$.

(d) $\boxed{\Longrightarrow}$ Soit $C > 0$ un réel. Ajouter une fonction affine préservant la convexité, on a la convexité de la fonction $t \mapsto \ln(f(t)) + t \ln C = \ln(f(t) C^t)$, donc $t \mapsto f(t) C^t$ est log-convexe, *a fortiori* convexe d'après le point (1c).

$\boxed{\Longleftarrow}$ Soient $a < b$ dans I et soient $\lambda, \mu > 0$ de somme 1. Soit $C > 0$. On a alors la comparaison

$$f(\lambda a + \mu b) C^{\lambda a + \mu b} \leq \lambda f(a) C^a + \mu f(b) C^b,$$

d'où (puisque C est strictement positif) les majorations (en notant $c := C^{a-b}$)

$$f(\lambda a + \mu b) \leq \lambda f(a) C^{a - \lambda a - \mu b} + \mu f(b) C^{b - \mu b - \lambda a} = \lambda f(a) c^{\mu} + \mu f(b) c^{-\lambda}.$$

Vu l'hypothèse $f > 0$, il suffit pour conclure de majorer le dernier terme ci-dessus par $f(a)^{\lambda} f(b)^{\mu}$. Or la comparaison arithmético-géométrique nous permet de *min*orer

$$\lambda\ f(a) c^{\mu} + \mu\ f(b) c^{\lambda} \geq [f(a) c^{\mu}]^{\lambda}\ \left[f(b) c^{-\lambda}\right]^{\mu} = f(a)^{\lambda} f(b)^{\mu}.$$

Pour avoir la *maj*oration voulue, il suffit de réaliser le cas d'*égalité* $f(a) c^{\mu} = f(b) c^{-\lambda}$, lequel équivaut à $c = \frac{f(b)}{f(a)}$ (ce quotient fait sens car $f > 0$). On peut donc conclure en imposant $C := \left(\frac{f(b)}{f(a)}\right)^{\frac{1}{a-b}}$ (cette puissance fait sens car $a \neq b$ et est non nulle car $f(b) > 0$).

Sanity check : en remplaçant C par 1, on retrouve le point (1c).

2.

(a) Soit $N \in \mathbb{N}$. Les deuxième et troisième hypothèses livrent les égalités

$$\begin{aligned} F(N) &= N\ F(N-1) = N(N-1)\ F(N-2) \\ &= \underset{\text{immédiate}}{\overset{\text{récurrence}}{\cdots}} = N(N-1)(N-2)\cdots 1\ F(0) = N!. \end{aligned}$$

(b) Vu d'une part les comparaisons $n - 1 < n < n + t \leq n + 1$ et d'autre part la convexité de $\ln F$, les comparaisons de l'énoncé font sens et sont valides. Par ailleurs, on peut réécrire comme le point précédent

$$F(n+t) \underset{\text{immédiate}}{\overset{\text{récurrence}}{=}} F(t)\ \underbrace{(t+1)(t+2)\cdots(t+n)}_{\text{abrégé en } t^{\blacktriangle n} \text{ (il y a } n \text{ facteurs)}} = F(t)\ t^{\blacktriangle n}.$$

Les comparaisons de l'énoncé se réécrivent maintenant

$$\frac{\ln[F(n)] - \ln[F(n-1)]}{(n) - (n-1)} \leq \frac{\ln[F(n+t)] - \ln[F(n!)]}{(n+t) - (n)} \leq \frac{\ln[F(n+1)] - \ln[F(n)]}{(n+1) - (n)},$$

$$i.\ e.\ \ln \frac{n!}{(n-1)!} \leq \frac{1}{t} \ln \frac{F(n+t)}{n!} \leq \ln \frac{(n+1)!}{n!},$$

$$i.\ e.\ t \ln n \leq \ln \frac{F(t)\ t^{\blacktriangle n}}{n!} \leq t \ln(n+1),\ i.\ e.\ n^t \leq \frac{F(t)\ t^{\blacktriangle n}}{n!} \leq (n+1)^t$$

$$\text{ou encore } 1 \leq \frac{F(t)\ t^{\blacktriangle n}}{n!\ n^t} \leq \left(1 + \frac{1}{n}\right)^t.$$

Lorsque n tend vers ∞, les deux membres de l'encadrement précédent tendent vers 1, d'où l'équivalent $\frac{F(t)\ t^{\blacktriangle N}}{N!\ N^t} \underset{N \to \infty}{\sim} 1$ et la tendance désirée.

(c) Étant donné un réel $a \geq 0$, on peut écrire

$$\begin{array}{rcl} a & = & n+t \text{ où } \left\{ \begin{array}{l} n := \lfloor a \rfloor \text{ est entier} \\ t := a - \lfloor a \rfloor \in [0,1[\end{array} \right. , \\ \text{d'où } F(a) & = & F(n+t) \underset{\text{précédent}}{\overset{\text{point}}{=}} F(t)\ t^{\blacktriangle n}, \end{array}$$

ce qui montre que F est déterminée par sa restriction à $[0,1[$. Or d'une part l'hypothèse $F(0) = 1$ détermine F en 0, d'autre part à $t \in]0,1[$ fixé la tendance établie au point précédent montre que $F(t)$ est limite d'une suite indépendante de F, donc ne saurait dépendre de F. Il en découle l'unicité désirée.

Remarque – **Culture hors programme**. La fonction "factorielle" définie sur $\mathbb{N}$ peut se prolonger sur $\mathbb{R}_+$ d'une infinité de manières, même en imposant ou bien les égalités $F(0) = 1$ et $\forall t \geq 1,\ F(t) = t\ F(t-1)$ ou bien la log-convexité de F. L'exercice montre que la *conjonction* de ces deux conditions détermine un *unique* prolongement. Ce résultat apparaît dans un manuel d'analyse de Harald Bohr et Johannes Mollerup publié en 1922. La démonstration précédente s'inspire de celle d'Emil Artin présente dans l'ouvrage *die Theorie der Gammafunktion* publié en 1931 – de fait, nous verrons avec la fonction Γ d'Euler (*cf.* chapitre ? ? ?) l'*existence* d'un tel prolongement.

2.4.2 Convexité, dérivabilité et continuité

Propriétés (convexité & dérivabilité à droite/gauche, position relative du graphe et des demi-tangentes)

Imposons f convexe. Alors

1. *f est dérivable à droite et à gauche en chaque point de l'intérieur $\mathring{I}$;*
2. *f est continue sur $\mathring{I}$.*
3. *le graphe de f est au-dessus de chacune de ses demi-tangentes*[54].

Démonstration Soit $a \in \mathring{I}$.

1. Soit $\varepsilon > 0$ tel que $[a-\varepsilon, a+\varepsilon] \subset I$. Puisque l'application τ_a est définie et croît (par convexité de f) sur l'intervalle $]a-\varepsilon, a+\varepsilon[$, elle admet une limite finie à gauche et à droite en a, *c. q. f. d.*
2. L'application f est continue des deux côtés en a (car elle y est dérivable des deux côtés), donc est continue en a.
3. Soit $t \in I$. Si $t < a$ (resp. $>$), on a alors la comparaison $\tau_a(t) \leq \lim_{a^+} \tau_a$ (resp. $\geq$), d'où après multiplication par $t-a$ qui est positif (resp. négatif) la comparaison $f(t) - f(a) \leq (t-a) f'_d(a)$, laquelle reste valide quand $t = a$, ce qui conclut. Raisonnement *identique* en remplaçant « dérivée à droite » par « dérivée à gauche ».

[54] Les demi-tangentes *verticales* (correspondant aux éventuelles "dérivées" infinies au bord de I) n'ont pas de sens à être « au-dessus de » quelque autre graphe et ne sont donc pas concernées.

Remarque – La restriction à l'*intérieur* de I est indispensable vu les exemples 2 et 3 ci-après.

Exemples

1. L'application convexe $t \mapsto |t|$ est dérivable sur $\mathbb{R}^*$, est dérivable à droite et à gauche en 0 et est continue. Son graphe est au-dessus des première et seconde bissectrice.
2. La fonction caractéristique de $\{0,1\}$ est dérivable (et continue) sur $]0,1[$ mais n'est continue ni en 0 ni en 1. Son graphe est au-dessus de l'axe des abscisses.
3. L'application continue $t \mapsto \sqrt{|t|}$ est concave sur $\mathbb{R}_+$ et sur $\mathbb{R}_-$, est dérivable sur $\mathbb{R}^*$ mais pas en 0. Son graphe est sur $\mathbb{R}_+$ en dessous de chacune de ses tangentes, la "tangente" en l'origine faisant exception faute de sens. *Idem* sur $\mathbb{R}_-$ (mais pas sur $\mathbb{R}^*$!).

Propriété (convexité & caractère C^1)

Imposons f convexe et dérivable. Alors f est de classe C^1.

Démonstration

Puisque f' croît, elle sera continue si elle vérifie le théorème des valeurs intermédiaires. Montrons que c'est le cas, de deux manières différentes et ce *indépendamment de l'hypothèse de convexité*[55].

Idée 1 : soient $a, b \in I$ tels que $f'(a) < f'(b)$ et montrons l'inclusion $]f'(a), f'(b)[\subset \operatorname{Im} f'$. Pour chaque $\gamma \in]f'(a), f'(b)[$, la dérivée $f' - \gamma$ de la fonction $f - \gamma \operatorname{Id}$ n'est pas de signe constant sur $[a, b]$ (puisque $f'(a) < \gamma < f'(b)$), donc la primitive $f - \gamma \operatorname{Id}$ n'est pas monotone sur $[a, b]$, *a fortiori* (puisqu'elle est continue) n'est pas injective sur $[a, b]$, donc prend sur $[a, b]$ une même valeur en deux points distincts, d'où par Rolle (vu que $f - \gamma \operatorname{Id}$ est dérivable sur $[a, b]$) un point $c \in]a, b[$ annulant sa dérivée, ce qui s'écrit $f'(c) = \gamma$ et conclut à l'inclusion désirée.

Idée 2 : soient $a, b \in I$. Les fonctions τ_a et τ_b se prolongent continûment sur $[a, b]$, donc les images de $[a, b]$ par ces fonctions sont deux intervalles se rencontrant (en $\tau_a(b) = \tau_b(a)$), donc leur réunion est un intervalle ; comme ce convexe contient $f'(a) = \tau_a(a)$ et $f'(b) = \tau_b(b)$, il inclut tout le segment $[f'(a), f'(b)]$. Enfin, chaque image par τ_a ou τ_b est (d'après la formule des accroissements finis) une image par f', d'où les inclusions

$$[f'(a), f'(b)] \subset \operatorname{Im} \tau_a \cup \operatorname{Im} \tau_b \subset \operatorname{Im} f' \cup \operatorname{Im} f' = \operatorname{Im} f', \text{ c. q. f. d.}$$

Exercice d'application

[55] *Culture* : le fait que chaque dérivée vérifie le théorème des valeurs intermédiaires (la convexité n'intervient pas) est attribué à Gaston Darboux et apparaît dans son « Mémoire sur les fonctions discontinues » de 1875 publié aux *Annales scientifiques de l'École Normale Supérieure*. On pourrait montrer (très hors programme) que chaque fonction $I \longrightarrow \mathbb{R}$ est continue ssi elle vérifie le théorème des valeurs intermédiaires et si l'image réciproque de chaque singleton est *fermée*.

Montrer que f est convexe ssi f est continue sur $\mathring{I}$ et si pour chaques points $a < b$ dans I il y a deux réels $\lambda, \mu \in \left]0, 1\right[$ de somme 1 tels que

$$f(\lambda a + \mu b) \leq \lambda f(a) + \mu f(b).$$

$\boxed{\Longrightarrow}$ La continuité sur $\mathring{I}$ découle de la proposition précédente et la convexité de f valide la comparaison ci-dessus pour $\lambda := \frac{1}{2} =: \mu$.

$\boxed{\Longleftarrow}$ Soient $u < v < w$ dans I. La comparaison $f(v) \leq \frac{w-v}{w-u} f(u) + \frac{v-u}{w-u} f(w)$ à montrer et la conclusion étant inchangées par ajout à f d'une fonction affine, on peut afin d'alléger imposer $f(u) = 0 = f(w)$ (quitte à soustraire à f la fonction affine coïncidant avec sur $\{u, w\}$). Cela d'une part transforme la majoration voulue en $f(v) \leq 0$, d'autre part donne sens aux bornes

$$a := \sup\{t \in [u, v] \ ; \ f(t) = 0\} \qquad \text{et} \qquad b := \inf\{t \in [v, w] \ ; \ f(t) = 0\}.$$

Observer alors l'encadrement $a \leq v \leq b$ et montrons les nullités $f(a) = 0 = f(b)$: si $a = u$, on a directement $f(a) = f(u) = 0$, sinon a tombe dans l'intervalle $\left]u, v\right] \subset \left]u, w\right[\subset \mathring{I}$ où f est continue et l'on a encore la nullité de $f(a)$ en approchant a par une suite de zéros de f (permis par définition de a) (on procéderait de même pour b).

FIG

Si $v = a$, on aura alors $f(v) = f(a) = 0$, d'où la conclusion (*idem* si $v = b$), ce qui permet d'imposer $a < v < b$. L'hypothèse nous donne alors deux réels $\lambda, \mu \in \left]0, 1\right[$ de somme 1 tels que $f(\lambda a + \mu b) \leq \lambda f(a) + \mu f(b) = 0$, d'où un point de $\left]a, b\right[$ où f est négative. La fonction f ne s'annule par ailleurs ni sur $\left]a, v\right]$ (par maximalité de a) ni sur $\left[v, b\right[$ (par minimalité de b), *i. e.* ne s'annule pas sur la réunion $\left]a, v\right] \cup \left[v, b\right[= \left]a, b\right[$, donc garde (par continuité sur $\mathring{I}$) un signe constant strict sur $\left]a, b\right[$. Des deux phrases précédentes résulte la comparaison $f < 0$ sur $\left]a, b\right[$, en particulier en v, ce qui conclut.

Sanity check : on retrouve le résultat de l'exercice d'application 3 section 2.1 dans le cas singulier où $\lambda = \frac{1}{2} = \mu$.

2.4.3 Convexité et cordes

Propriété (positions relatives du graphe et des cordes d'une fonction convexe)

Imposons f convexe et soit $S \subset I$ un segment infini. On note φ la fonction affine coïncidant avec f aux bornes de S. On a alors les comparaisons $\begin{cases} f \leq \varphi \text{ sur } S \\ f \geq \varphi \text{ sur } I \setminus S \end{cases}$:

FIG

Démonstration

Nous avons déjà vu la comparaison $f \leq \varphi$ sur S. Notons $a < b$ les extrémités de[56] S et soit $t \in I \setminus S$. Rappelons l'égalité $\varphi(t) = (t - a)\, \tau_a^b + f(a)$ et discutons selon la position de t par rapport à $S = [a, b]$:

[56] La comparaison $a < b$ est stricte car S est infini.

1. si $t < a$, on a alors la comparaison $\tau_t^a \leq \tau_a^b$, *i. e.* $f(a) - f(t) \leq (a-t)\,\tau_a^b$, ou encore $\varphi(t) \leq f(t)$;
2. si $t > b$, on a alors la comparaison $\tau_a^b \leq \tau_a^t$, *i. e.* $(t-a)\,\tau_a^b \leq f(t) - f(a)$, ou encore $\varphi(t) \leq f(t)$.

Exercices d'application

1. *Montrer que chaque fonction convexe sur un intervalle borné y est minorée.*
2. *Montrer que chaque fonction convexe majorée sur $\mathbb{R}$ est constante.*
3. *Montrer que chaque fonction convexe majorée sur $\mathbb{R}_+$ décroît.*

---

1. Imposons I borné et f convexe. Si I est fini, l'image $f(I)$ est alors finie et est minorée. Imposons donc I infini, ce qui permet d'évoquer un segment $S \subset I$ infini. La fonction f est alors d'une part minorée sur S (car continue sur le segment S), d'autre part minorée en dehors de S par une fonction affine (d'après la propriété ci-dessus), laquelle est minorée sur I car ce dernier est borné. Total : f est minorée.

   FIG

   Remarque – Quand $I =]0,1[$, l'application $t \mapsto \frac{1}{t}$ est convexe (bien sûr minorée) mais non majorée.
2. Imposons $I = \mathbb{R}$ et f convexe non constante. Soient $a < b$ deux réels tels que $\tau_a^b \neq 0$. Selon le signe de cette pente, la fonction affine coïncidant avec f sur $\{a, b\}$ tend alors vers ∞ en ∞ ou en $-\infty$; or cette fonction minore f hors de $[a,b]$ (d'après la propriété ci-dessus et par convexité de f), donc cette dernière tend aussi vers ∞ (en ∞ ou en $-\infty$) et n'est par conséquent pas majorée.

   FIG
3. Imposons $I = \mathbb{R}_+$ et f convexe majorée. Soient $a < b$ dans I : si $\tau_a^b > 0$, le même raisonnement qu'au point précédent montrerait alors la tendance $f \underset{\infty}{\longrightarrow} \infty$ et contredirait le caractère majoré de f ; on en déduit $\tau_a^b \leq 0$, d'où la décroissance de f.

   FIG

Remarque – Au lieu de cordes, on aurait également pu utiliser des demi-tangentes afin de minorer f.

2.4.4 Convexité et extrema

Propriétés (*maxima* et *minima* d'une fonction convexe)

1. *Chaque fonction convexe sur un segment y est majorée et atteint sa borne supérieure en l'une des extrémités de ce segment.*

2.

(a) *Si f est convexe, alors chaque minimum local de f est un minimum global.*

(b) *Si de plus f est dérivable, alors chaque point annulant f' réalise le minimum de f.*

Démonstration

1. Imposons que $S := I$ soit un segment et que f y soit convexe. Notons φ la fonction affine coïncidant avec f en les bornes de S. On a alors les comparaisons $f \leq \varphi \leq \max \varphi$: la première (qui découle de la convexité de f) est une égalité en chacune des bornes de S (par définition de φ), la seconde est une égalité en l'une des bornes de S (car φ est affine), ce qui montre que $\sup f = \max \varphi$ est atteint en un point du bord de S, *c. q. f. d.*

2.

(a) Imposons f convexe et soit $m \in I$ un point réalisant un mimum local pour f. On imposera I infini (il n'y a sinon rien à faire). Distinguons alors selon que m est une borne de I ou bien est à l'intérieur de I.

Supposons $m = \inf I$. Il y a alors un $i > m$ dans I tel que $f \geq f(m)$ sur $[m, i]$, donc la fonction τ_m est positive en i, donc (par croissance) positive sur $]i, \infty[\cap I$, d'où la minoration $f \geq f(m)$ à droite de m, çed sur tout I, ce qui conclut. Le même raisonnement tiendrait si $m = \sup I$ en regardant à gauche de m. Enfin, reprendre les deux arguments précédents lorsque $m \in \mathring{I}$ conclurait de même en regardant des deux côtés de m.

(b) Imposons f convexe et dérivable. Soit $m \in I$ tel que $f'(m) = 0$: le graphe de f est alors au-dessus de sa tangente en le point $\binom{m}{f(m)}$, comparaison qui s'écrit $f \geq f(m)$, *c. q. f. d.*

Rappel. La réciproque n'est valide que pour les points *intérieurs* de I : en effet, la fonction convexe dérivable Id admet sur $[0, 1]$ un minimum global en 0 mais sa dérivée ne s'y annule pas.

Contre-exemples : si $f = \left[t \mapsto \left\{ \begin{array}{c} 1 \text{ si } t = 0 \\ \frac{1}{2-t} \text{ si } t > 0 \end{array} \right. \right]$, alors

FIG

1. f est convexe sur l'intervalle $[1, 2[$ mais n'y est pas majorée,
2. f est convexe sur le segment $[0, 1]$ mais n'y atteint pas sa borne inférieure.

Exercices d'application

1. *Montrer que f est convexe ssi pour chaque segment $S \subset I$ et pour chaque réel ρ la fonction $t \mapsto f(t) + \rho t$ restreinte à S atteint sa borne supérieure en l'une des extrémités de S.*

2. *Soient* $a, b, c \in [0,1]$*. Montrer alors la majoration*

$$\frac{a}{b+c+1} + \frac{b}{c+a+1} + \frac{c}{a+b+1} + (1-a)(1-b)(1-c) \leq 1.$$

On pourra voir le membre de gauche comme une fonction convexe dont on cherche le maximum.

---

1. $\boxed{\Longrightarrow}$ Soit $\rho \in \mathbb{R}$ et soit $S \subset I$ un segment. Alors la fonction $t \mapsto f(t) + \rho t$ est convexe comme somme de deux fonctions convexes, donc est majorée sur S et atteint sa borne supérieure en l'une des extrémités de S.

   $\boxed{\Longleftarrow}$ Observer que la propriété « atteindre sa borne supérieure en l'une des extrémités de S » est inchangée par ajout d'une constante : on pourra par conséquent appliquer cette propriété à $f + A$ pour chaque fonction A affine.

   Soit $a < b$ dans I, notons $S := [a, b]$ et φ la fonction affine coïncidant avec f sur $\{a, b\}$. Il suffit pour conclure d'établir la majoration $f \leq \varphi$. Or, d'une part (d'après le paragraphe précédent) la fonction $f - \varphi$ atteint sa borne supérieure en l'une des extrémités de S, d'autre part (par définition de φ) en chacune de ces extrémités la différence $f - \varphi$ est nulle : on en déduit la majoration $f - \varphi \leq \max(f - \varphi) = 0$, *c. q. f. d.*

2. Le membre de gauche de la majoration désirée peut s'écrire $F(a, b, c)$ pour une certaine fonction $F : [0,1]^3 \longrightarrow \mathbb{R}$ somme de quatre fonctions. Montrons comme suggéré que F est convexe en chacun de ses trois arguments.

   Soient $A, B \in [0,1]$. La fonction $t \mapsto \frac{t}{A+B+1}$ est affine, donc convexe. *Idem* pour $t \mapsto (1-A)(1-B)(1-t)$. La fonction $t \mapsto \frac{A}{B+t+1}$ est dérivable et sa dérivée $t \mapsto \frac{-A}{(1+B+t)^2}$ croît sur $]-1-B, \infty[$ car $A \geq 0$, donc est convexe sur $[0,1]$ car $B > -1$. De même pour $t \mapsto \frac{B}{t+A+1}$. Finalement, la fonction $t \mapsto F(A, B, t)$ est convexe sur $[0,1]$. Vu par ailleurs à $t \in [0,1]$ fixé les égalités[57]

   $$F(A, B, t) = F(t, A, B) = F(B, t, A),$$

   la fonction F est bien convexe en chacun de ses arguments.

   Concluons. La fonction $t \mapsto F(t, b, c)$ est convexe, donc maximale sur $[0,1]$ en l'une des bornes de $[0,1]$, mettons en un certain $P \in \{0,1\}$. Ensuite, la fonction $t \mapsto F(P, t, c)$ est de même maximale sur $[0,1]$ en un certain $Q \in \{0,1\}$, puis la fonction $t \mapsto F(P, Q, t)$ est maximale sur $[0,1]$ en un certain $R \in \{0,1\}$. On obtient ainsi les majorations[58]

   $$F(a, b, c) \leq F(P, b, c) \leq F(P, Q, c) \leq F(P, Q, R) \ ;$$

   or, d'une part le point (P, Q, R) appartient à $\{0,1\}^3$, d'autre part en chaque point de $\{0,1\}^3$ la fonction F prend la valeur 1 (quatre calculs à effectuer selon le nombre de coordonnées valant 1), ce qui conclut.

   Remarque – Cette majoration se généralise immédiatement à n'importe quel nombre de lettres.

---

[57] On dit que F est *invariante par permutation cyclique* de ses arguments.

[58] Les points P, Q, R dépendent *a priori* de a, b, c mais cette dépendance n'est pas pertinente pour notre preuve.

2.4.5 Concavité et sous-additivité

Propriété (sous-addititivité des fonctions concaves)

Chaque fonction concave sur $\mathbb{R}_+$ *et fixant* 0 *est sous-additive :*

$$I = \mathbb{R}_+ \text{ et } f \text{ concave et } f(0) = 0 \Longrightarrow \forall a, b \geq 0,\ f(a+b) \leq f(a) + f(b).$$

Démonstration

Supposons $I = \mathbb{R}_+$ et f concave fixant 0. Soient $a, b \geq 0$. Voici deux
Preuve algébrique. Si $a = 0 = b$, la comparaison voulue s'écrit $0 \leq 0 + 0$ et est vérifiée. Dans le cas contraire, les réels $(\lambda, \mu) := \left(\frac{a}{a+b}, \frac{b}{a+b}\right)$ font sens, tombent dans $[0,1]$ et sont de somme 1, d'où les comparaisons

$$f(a) = f(\lambda(a+b) + \mu 0) \underset{\text{concave}}{\overset{f \text{ est}}{\geq}} \lambda f(a+b) + \mu f(0) = \lambda f(a+b)\ ;$$

on montrerait de même la comparaison $f(b) \geq \mu f(a+b)$, d'où en additionnant la comparaison

$$f(a) + f(b) \geq (\lambda + \mu) f(a+b) = f(a+b),\ c.\ q.\ f.\ d.$$

Preuve géométrique. La comparaison voulue étant symétrique en a et b, on peut imposer $a \leq b$. Appelons alors A, B, S les points du graphe de f d'abscisses respectives $a, b, a+b$. Sur la bande $[a,b] \times \mathbb{R}$, la corde (AB) est alors au-dessus de la corde (OB), laquelle est au-dessus de la corde (OS) : les milieux de $[AB]$ et $[OS]$ ayant par ailleurs même abscisse $\frac{a+b}{2}$, l'ordonnée du premier majore donc celle du second, ce qui s'écrit $\frac{f(a)+f(b)}{2} \leq \frac{f(0)+f(a+b)}{2}$, *c. q. f. d.*

FIG

Remarque – La réciproque est fausse : la fonction $t \mapsto \lceil t \rceil$ fixe 0 et est sous-additive (à détailler) mais n'est pas concave sur $\mathbb{R}_+$. Une réciproque "faible" est cependant présentée à l'exercice d'entraînement[59] 9b.

Exercice d'application

Soient $a, b, c \geq 1$ *trois réels. Montrer alors la comparaison*

$$a + b + c + \frac{1}{abc} \leq \frac{1}{a} + \frac{1}{b} + \frac{1}{c} + abc.$$

On pourra reparamétrer à l'aide de la fonction exp *et utiliser la convexité de* sh.

---

[59] Cet exercice pourra être contextualisé par la lectrice et le lecteur intéressés à la lumière d'un article de Philippe Turpin publié en 1973 dans la revue *Studia Mathematica* (archives disponibles en ligne).

Notons comme suggéré α, β, γ les antécédents respectifs de a, b, c par exp. La comparaison voulue se réécrit alors

$$\begin{array}{rcl} e^{\alpha}+e^{\beta}+e^{\gamma}+e^{-\alpha-\beta-\gamma} & \leq & e^{-\alpha}+e^{-\beta}+e^{-\gamma}+e^{\alpha+\beta+\gamma}, \\ i.\ e.\ \dfrac{e^{-(\alpha+\beta+\gamma)}-e^{\alpha+\beta+\gamma}}{2} & \leq & \dfrac{e^{-\alpha}-e^{\alpha}}{2}+\dfrac{e^{-\beta}-e^{\beta}}{2}+\dfrac{e^{-\gamma}-e^{\gamma}}{2}, \\ \text{ou encore } -\operatorname{sh}(\alpha+\beta+\gamma) & \leq & -\operatorname{sh}\alpha-\operatorname{sh}\beta-\operatorname{sh}\gamma. \end{array}$$

Or la fonction $-\operatorname{sh}$ est concave sur $\mathbb{R}_+$ (où tombent α, β, γ puisque $a, b, c \geq 1$) et fixe 0, donc est sous-additive, ce qui conclut à la comparaison ci-dessus.

Remarque – Généralisation immédiate à un nombre quelconque (au lieu de trois) de réels dans $[1, \infty[$.

2.4.6 Convexité stricte

Proposition (admise) – Définition (stricte convexité)

Les assertions suivantes sont équivalentes :

1. *l'épigraphe de f privé de $\{P\}$ est une partie convexe*[60] *pour chaque point P du graphe de f ;*
2. *f est convexe sur I et n'est affine sur aucun segment (infini) de I ;*
3. *pour chaques points $a < b$ de I, pour chaques réels $\lambda, \mu > 0$ de somme 1, on a la comparaison stricte*

$$f(\lambda a + \mu b) < \lambda f(a) + \mu f(b) \text{ ;}$$

4. *pour chaques points $a < b < c$ de I, on a les comparaisons strictes*

$$\tau_a^b < \tau_a^c < \tau_b^c \text{ ;}$$

5. *pour chaque point $a \in I$, l'application τ_a croît strictement ;*
6. *(si de plus f est dérivable) la dérivée f' croît strictement.*
   Lorsque l'une des conditions ci-dessus est vérifiée[61]*, la fonction f est dite* ***strictement convexe****.*

L'unique intérêt (dans ce cours) de la *stricte* convexité est de préciser le cas d'égalité dans l'égalité de Jensen : pour chaque naturel $n \geq 1$, pour chaque n-uplet λ *strictement* positif de somme 1, pour chaque $a \in I^n$, on a lorsque f est *strictement* convexe la comparaison

$$f\left(\sum_{i=1}^{n} \lambda_i a_i\right) \leq \sum_{i=1}^{n} \lambda_i f(a_i) \text{ avec égalité ssi les } a_i \text{ sont égaux,}$$

[60] En d'autres termes : l'épigraphe de f est convexe et ne contient aucun segment infini sur son "bord du bas".

[61] *Attention* : lorsque f est deux fois dérivable, ces assertions n'équivalent pas à « $f'' > 0$ » ! mais à ce que f'' reste positive et ne s'annule sur aucun intervalle infini de I.

la majoration s'écrivant alors $f(a) \leq f(a)$ où a est la valeur commune des a_i. En d'autres termes, le cas d'égalité correspond au cas vraiment simple où le barycentre $\sum_{i=n}^{n} \lambda_i a_i$ se confond avec chacun de ses points pondérés (et ce indépendamment des poids).

Exercices d'application

1. *On reprend les notations de l'exercice d'application 2 section 2.2. Montrer l'implication*
$$\sum_{i=1}^{n} \frac{\lambda_i^2}{\lambda_i + \lambda_{i+1}} = \frac{1}{2} \Longrightarrow \forall i \in [1, n|] \,,\ \lambda_i = \frac{1}{n}.$$
2. *On reprend les notations de l'exercice d'application 3 section 2.2 et on impose chaque poids λ_i non nul. Soient $u < v$ deux réels tels que $M_u = M_v$. Montrer alors que les a_i sont égaux.*
3. *On reprend les notations de l'exercice d'application 3 section 2.3. Montrer l'implication*
$$\frac{1}{\sin \alpha} + \frac{1}{\sin \beta} = \frac{8}{3 + 2\cos \gamma} \Longrightarrow ABC \text{ isocèle et } \gamma = 120°.$$

---

1. On avait utilisé dans la preuve la convexité de la fonction $t \mapsto \frac{1}{t+1}$ appliquée aux réels $\frac{\lambda_{i+1}}{\lambda_i}$ et aux poids λ_i. Puisqu'aucun poids n'est nul et que la fonction convexe utilisée l'est *strictement*, le cas d'égalité force l'égalité des réels $\frac{\lambda_{i+1}}{\lambda_i}$. Soit donc $C > 0$ tel que $\frac{\lambda_{i+1}}{\lambda_i} = C$ pour chaque $i \in [1, n|]$. Multiplier ces n égalités donne alors (utiliser un produit télescopique) l'égalité $1 = C^n$, d'où $C = 1$ (puisque $C > 0$) et les égalités $\lambda_{i+1} = \lambda_i$, *c. q. f. d.*.
2. Si $u = 0$, nous avions utilisé la concavité de ln appliquée aux a_i^v, laquelle fonction est *strictement* concave, d'où l'égalité des a_i^v (aucun poids n'est nul) et celle des a_i (appliquer $t \mapsto t^{\frac{1}{v}}$ fait sens car $v > u = 0$). De même, si $v = 0$, l'hypothèse $M_u = M_0$ équivalait à $M_0^* = M_{-u}^*$, d'où l'égalité des $\left(\frac{1}{a_i}\right)^{-u}$ et celle des a_i.

   Si $0 \leq u < v$, c'est la convexité de $t \mapsto t^{\frac{v}{u}}$ qui avait été appliquée aux a_i^u, convexité *stricte* vu la comparaison $\frac{v}{u} > 1$. De même, si $u < v \leq 0$, l'hypothèse $M_u = M_v$ équivaut à $M_{-v}^* = M_{-u}^*$, d'où l'égalité des $\left(\frac{1}{a_i}\right)^{-v}$ et celle des a_i.

   Enfin, si $u < 0 < v$, la croissance de M force alors les égalités $M_u = M_0 = M_v$ et l'on peut utiliser le premier paragraphe.
3. Nous avions utilisé la convexité de $\frac{1}{\sin}$ appliquée à α et β avec les poids valant $\frac{1}{2}$. Cette convexité étant stricte (reprendre les décompositions en fonctions *strictement* monotones ou bien utiliser la *stricte* positivité de la dérivée seconde), le cas d'égalité force l'égalité $\alpha = \beta$, d'où le caractère isocèle de ABC. Par ailleurs, nous avions utilisé la positivité du carré $(2c-1)^2$: le cas d'égalité force donc la nullité de $2c - 1$, d'où $\cos \frac{\gamma}{2} = \frac{1}{2}$, *i. e.* $\frac{\gamma}{2} = \frac{\pi}{3}$ (car $0 < \gamma < \pi$), ou encore $\gamma = \frac{2\pi}{3} = 120°$.

Remarque – Chacune des trois implications montrées est en fait une équivalence, comme la lectrice et le lecteur avisés sauront le vérifier par eux-mêmes.

3 Le point des compétences

Formulaire

1. Parties convexes d'un espace vectoriel réel

Soit E un $\mathbb{R}$-espace vectoriel et soit A une partie de E.

• On appelle ***barycentre*** de A toute combinaison linéaire de points de A dont la somme des coefficients vaut 1.

• Soient $a, b \in E$. On appelle ***segment*** d'extrémités a et b l'ensemble des barycentres de $\{a, b\}$ à coefficients positifs. Ce segment est noté

$$[a, b] := \{\lambda a + (1 - \lambda) b \ ; \ \lambda \in [0, 1]\} = \{\lambda a + \mu b\}_{\lambda+\mu=1}^{\lambda,\mu\in[0,1]} = [b, a] .$$

• La partie A est dite ***étoilée*** si elle est stable par passage au segment dont une extrémité est fixée :

$$A \ \textit{\textbf{étoilée}} \overset{\text{déf.}}{\iff} \exists a \in A, \ \forall \alpha \in A, \ [a, \alpha] \in A.$$

• La partie A est dite ***convexe*** si elle est stable par passage au segment :

$$A \ \textit{\textbf{convexe}} \overset{\text{déf.}}{\iff} \begin{array}{c} \forall a, \alpha \in A \\ \forall \lambda, \mu \in [0, 1] \end{array}, \lambda + \mu = 1 \Longrightarrow \lambda a + \mu \alpha \in A.$$

• La partie A est convexe ssi elle est stable par barycentration à coefficients positifs :

$$A \text{ convexe} \iff \forall n \in \mathbb{N}^*, \quad \begin{array}{c} \forall \overrightarrow{a} \in A^n \\ \forall \overrightarrow{\lambda} \in [0, 1]^n \end{array}, \ \sum_{i=1}^{n} \lambda_i = 1 \Longrightarrow \sum_{i=1}^{n} \lambda_i a_i \in A.$$

2. Fonctions convexes réelles définies sur un convexe de $\mathbb{R}$

Soit I un intervalle de $\mathbb{R}$, soit $f : I \longrightarrow \mathbb{R}$.

• On appelle ***épigraphe*** de f l'ensemble des couples dont l'ordonnée majore l'image par f de l'abscisse :

$$\textit{\textbf{épigraphe}} \text{ de } f \ := \ \{(i, j) \in I \times \mathbb{R} \ ; \ j \geq f(i)\} .$$

• La fonction f est dite ***convexe*** si

$$\begin{array}{c} \forall a, b \in I \\ \forall \lambda, \mu \in [0, 1] \end{array}, \ \lambda + \mu = 1 \Longrightarrow f(\lambda a + \mu b) \leq \lambda f(a) + \mu f(b)$$

(retenir : « *l'image du barycentre est sous le barycentre des images* »).

• La fonction f est dite ***concave*** si $-f$ est convexe.

• **Convexité et épigraphe** : chaque fonction $I \longrightarrow \mathbb{R}$ est convexe ssi son épigraphe est une partie convexe de $\mathbb{R}^2$:

$$f \text{ convexe} \iff \text{l'épigraphe de } f \text{ est convexe.}$$

• **Convexité et croissance des pentes** : la fonction f est convexe ssi pour chaques points a, b, c de I, on a l'implication

$$a < b < c \Longrightarrow \frac{f(b) - f(a)}{b - a} \leq \frac{f(c) - f(a)}{c - a} \leq \frac{f(c) - f(b)}{c - b}$$

ou encore ssi, pour chaque point $a \in I$, la fonction $\left\{ \begin{array}{ccc} I \setminus \{a\} & \longrightarrow & \mathbb{R} \\ t & \longmapsto & \frac{f(t)-f(a)}{t-a} \end{array} \right.$ croît.

• **Convexité et positions relatives du graphe avec ses cordes** : la fonction f est convexe ssi, pour chaque segment $S \subset I$, le graphe de la restriction $f_{|S}$ est en dessous de la corde reliant les extrémités de ce graphe :

$$f \text{ convexe} \iff \forall a < b \text{ dans } I, \, \forall t \in [a, b], \ f(t) \leq \frac{f(b) - f(a)}{b - a}(t - a) + f(a).$$

• Chaque fonction $I \longrightarrow \mathbb{R}$ dérivable est convexe ssi sa dérivée croît :

$$f \text{ convexe} \underset{\text{dérivable}}{\overset{\text{si } f}{\iff}} f' \text{ croît.}$$

• Chaque fonction $I \longrightarrow \mathbb{R}$ deux fois dérivable est convexe ssi sa dérivée seconde est positive :

$$f \text{ convexe} \underset{\text{dérivable}}{\overset{\text{si } f \text{ deux fois}}{\iff}} f'' \geq 0.$$

• Le graphe de chaque fonction convexe dérivable est au-dessus de chacune de ses tangentes :

$$f \text{ convexe dérivable} \Longrightarrow \forall a, t \in I, \ f(t) \geq (t - a) f'(a) + f(a).$$

• **Comparaisons de Jensen** : si f est convexe, on a alors les comparaisons

$$\forall n \in \mathbb{N}^*, \quad \begin{array}{c} \forall \overrightarrow{a} \in I^n \\ \forall \overrightarrow{\lambda} \in [0,1]^n \end{array}, \ \sum_{i=1}^n \lambda_i = 1 \Longrightarrow f\left(\sum_{i=1}^n \lambda_i a_i\right) \leq \sum_{i=1}^n \lambda_i f(a_i).$$

• **Exemples de comparaisons de Jensen** :

1. Soit $n \in \mathbb{N}^*$, soit $\overrightarrow{\lambda} \in [0,1]^n$ de somme 1 et soit $\overrightarrow{a} \in \mathbb{R}_+^n$. La concavité de ln livre alors la comparaison (dite ***arithmético-géométrique***)

$$\underset{\substack{\textbf{\textit{moyenne géométrique}} \\ \text{des } a_i \text{ pondérés par les } \lambda_i}}{a_1^{\lambda_1} a_2^{\lambda_2} \cdots a_n^{\lambda_n}} \leq \underset{\substack{\textbf{\textit{moyenne arithmétique}} \\ \text{des } a_i \text{ pondérés par les } \lambda_i}}{\lambda_1 a_1 + \lambda_2 a_2 + \cdots + \lambda_n a_n}.$$

En particulier, quand $\lambda_i = \frac{1}{n}$ pour chaque $i \in [1, n|]$, on obtient la comparaison

$$\boxed{\sqrt[n]{a_1 a_2 \cdots a_n} \leq \frac{a_1 + a_2 + \cdots + a_n}{n}}.$$

2. Soit n un naturel, soient $a, b \in \mathbb{R}_+^n$ et soient $\alpha, \beta \geq 0$ des réels de somme 1. On a alors la comparaison (dite de Hölder)

$$\sum_{i=1}^{n} a_i^\alpha b_i^\beta \leq \left(\sum_{i=1}^{n} a_i\right)^\alpha \left(\sum_{i=1}^{n} b_i\right)^\beta .$$

Le nombre de couples de lettres $\binom{a}{\alpha}, \binom{b}{\beta}$ (ici deux) peut en fait être imposé quelconque dans $\mathbb{N}$ (preuve inchangée).

Exercices d'entraînement

1. ★*Soit E un $\mathbb{R}$-espace vectoriel. Montrer que chaque somme de parties convexes de E reste convexe.*

2. ★*Soit $f : \mathbb{R}_+ \longrightarrow \mathbb{R}$ convexe dont le graphe admet une droite asymptote au voisinage de ∞. Déterminer la position relative de ce graphe et de cette asymptote.*

3. ★★*Soit E un ensemble fini non vide et notons P l'ensemble formé par les fonctions $f : E \longrightarrow \mathbb{R}_+^*$ telles que $\sum_{e\in E} f(e) = 1$. On définit alors deux applications*

$$H := \left\{ \begin{array}{ccc} P & \longrightarrow & \mathbb{R} \\ p & \longmapsto & -\sum_{e\in E} p_e \ln p_e \end{array} \right. \quad \textit{et } H_2 := \left\{ \begin{array}{ccc} P^2 & \longrightarrow & \mathbb{R} \\ \binom{p}{q} & \longmapsto & -\sum_{e\in E} p_e \ln \frac{p_e}{q_e} \end{array} \right. .$$

   (a) *En utilisant la stricte concavité de $\ln$, montrer que H_2 est négative et s'annule exactement sur la diagonale de P. (Culture : la **diagonale** d'un ensemble A est formée des couples (a, a) où a décrit A.)*

   (b) *En déduire pour chaque $p \in P$ la comparaison*

$$H(p) \leq \ln \operatorname{Card} E \textit{ avec égalité ssi } p \textit{ vaut constamment } \frac{1}{\operatorname{Card} E}.$$

4. ★★*(On pourra tout au long de cet exercice penser aux comparaisons arithmético-géométriques.)*

   (a) *Soit $n \in \mathbb{N}^*$ et notons $R := \{\overrightarrow{r} \in \mathbb{R}^n \ ; \ \forall i \in [1, n|] , \ r_i > 0\}$. Pour chaque $\overrightarrow{a} \in R$, donner sens à et montrer l'égalité*

$$\min \left\{ \frac{\varepsilon_1 a_1 + \varepsilon_2 a_2 + \cdots + \varepsilon_n a_n}{n} \ ; \ \overrightarrow{\varepsilon} \in R \textit{ et } \varepsilon_1 \varepsilon_2 \cdots \varepsilon_n = 1 \right\} = \sqrt[n]{a_1 a_2 \cdots a_n}.$$

   (b) *Soit $n \in \mathbb{N}^*$ et soit $\overrightarrow{a} \in \mathbb{R}_+^n$ de produit 1. Montrer alors la comparaison*

$$(2 + a_1)(2 + a_2) \cdots (2 + a_n) \geq 3^n.$$

   (c) *Soit $n \geq 2$ un entier et soient $a_2, a_3 ..., a_n$ des réels positifs de produit 1. Montrer alors la comparaison*

$$(1 + a_2)^2 (1 + a_3)^3 \cdots (1 + a_n)^n \geq n^n.$$

   (d) *Soit $n \geq 2$ un entier et soit $\overrightarrow{a} \in \,]0, 1[^n$ de somme 1. Montrer alors la comparaison*

$$\sum_{i=1}^{n} \frac{a_i}{\sqrt{1 - a_i}} \geq \sum_{i=1}^{n} \sqrt{\frac{a_i}{n-1}}.$$

5. ★★*Soit $n \in \mathbb{N}$ et soient $a, b \in \mathbb{R}_+^n$. Comparer*

$$\sqrt[n]{a_1 a_2 \cdots a_n} + \sqrt[n]{b_1 b_2 \cdots b_n} \textit{ et } \sqrt[n]{(a_1 + b_1)(a_2 + b_n) \cdots (a_n + b_n)}.$$

   On pourra utiliser une comparaison de Hölder ou bien la convexité de[62] $t \mapsto \ln(1 + e^t)$.

[62] *cf.* exercice d'application 1 section 2.3

6. ★★

   (a) *Soient $a < b$ deux réels et soit $f : [a, b] \longrightarrow \mathbb{R}$ concave dérivable. Montrer alors l'encadrement*

   $$\frac{f(a) + f(b)}{2} \leq \frac{1}{b-a} \int_a^b f \leq f\left(\frac{a+b}{2}\right).$$

   (b) *Soit $f : [0, 1] \longrightarrow \mathbb{R}$ concave telle que $f(0) = 1$. Montrer alors les comparaisons*

   $$\forall u \in [0, 1], \ u f(u) \leq 2\left(\int_0^u f\right) - u$$

   et en déduire la majoration

   $$\int_0^1 t f(t) \, \mathrm{d}t \leq \frac{2}{3}\left(\int_0^1 f\right)^2.$$

   On pourra utiliser l'égalité $\int_{u=0}^1 \int_{t=0}^u \varphi(t) \ \mathrm{d}t \ \mathrm{d}u = \int_{t=0}^1 \int_{u=t}^1 \varphi(t) \ \mathrm{d}u \ \mathrm{d}t$ valide pour chaque fonction $\varphi : [0, 1] \longrightarrow \mathbb{R}$ continue.

7. ★★*Soient $a < b$ deux réels. On note E l'ensemble des fonctions concaves de $C^2([a, b], \mathbb{R})$ s'annulant sur $\{a, b\}$. Montrer que*

   $$\text{la fonction } L \text{ "longueur du graphe" } f \mapsto \int_a^b \sqrt{1 + f'^2} \text{ croît sur } E.$$

   On pourra utiliser la convexité de l'application[63] $t \mapsto \sqrt{1 + t^2}$.

8. ★★★*Soit $f : \mathbb{R} \longrightarrow \mathbb{R}$ convexe dérivable et soit $n \in \mathbb{N}^*$. On veut montrer l'encadrement*

   $$0 \leq \frac{f(0)}{2} + f(1) + f(2) + \cdots + f(n-1) + \frac{f(n)}{2} - \int_0^n f \leq \frac{f'(n) - f'(0)}{8}.$$

   (a) *Justifier pourquoi on peut se ramener au cas $n = 1$.*

   (b) *Commenter géométriquement l'encadrement voulu et en déduire l'une des comparaisons souhaitée.*

   (c) *Justifier pourquoi l'on peut imposer $f(0) = 0$, $f'(0)$ et $f(1) = 1$. On pourra observer comment évolue l'encadrement à montrer lorsque l'on ajoute à f une fonction affine.*

   (d) *Conclure.*

9. ★★★*Soit I un intervalle réel. Pour chaque $a \in I$ on note*

   $$\Phi_a := \left\{\varphi \in C_{p.\ m.}([0, 1], I) \ ; \ \int_0^1 \varphi = a\right\}.$$

   Soit $f : I \longrightarrow \mathbb{R}$ telle que la fonction $C := \left\{ \begin{array}{ccc} I & \longrightarrow & \mathbb{R} \\ t & \longmapsto & \inf_{\varphi \in \Phi_t} \int_0^1 f \circ \varphi \end{array} \right.$ fait sens.

[63] *cf.* exercice d'application 1 section 2.3

(a) *Montrer que C est la plus grande fonction convexe minorant f. Pour établir sa convexité, étant donnés un $\lambda \in]0,1[$, deux $a < b$ dans I et un $(\alpha, \beta) \in \Phi_a \times \Phi_b$, on pourra considérer la fonction*

$$\left\{ \begin{array}{rcl} [0,\lambda[\ni t & \longmapsto & \alpha\left(\frac{t}{\lambda}\right) \\ [\lambda,1] \ni t & \longmapsto & \beta\left(\frac{t-\lambda}{1-\lambda}\right) \end{array} \right. .$$

(b) *Soit $F : \mathbb{R}_+ \longrightarrow \mathbb{R}$ croissante sous-additive et fixant 0. Montrer l'encadrement*

$\frac{c}{2} \leq F \leq c$ pour une certain fonction concave c croissante et fixant 0.

On rappelle au besoin l'encadrement $a \leq \lceil a \rceil \leq a+1$ pour chaque réel a.

10. ★★★

(a) *Soit S un segment réel, soient $f, g : S \longrightarrow \mathbb{R}$ deux fonctions convexes dérivables telles que $\max\{f,g\} \geq 0$. On veut montrer que le segment $[f,g]$ de $\mathbb{R}^S$ contient une fonction positive.*

i. *Conclure si f ou g est positive. On exclut ces cas par la suite.*

ii. *Montrer l'existence d'un point de S où $\max\{f,g\}$ est minimale, où f et g sont positives et où $f'g'$ est négative.*

iii. *Conclure.*

(b) *Reprendre la question (10a) sans l'hypothèse de dérivabilité. On pourra se ramener au cas où f et g sont continues.*

(c) *Reprendre la question (10a) en remplaçant le segment S par n'importe quel intervalle réel. On pourra chercher à établir que f ou g tend (en l'une des bornes de cet intervalle) vers la borne inférieure de $\max\{f,g\}$.*

(d) *Soit I un intervalle réel, soit $n \in \mathbb{N}^*$ et soit $\overrightarrow{f}$ un n-uplet de fonctions $I \longrightarrow \mathbb{R}$ convexes. Montrer qu'il y a un $\overrightarrow{\lambda} \in [0,1]^n$ de somme 1 tel que $\sum_{i=1}^n \lambda_i f_i \geq 0$ ssi $\max_{i \in [1,n|]} f_i \geq 0$.*

Solutions des exercices d'entraînement

1. Par une récurrence immédiate, il suffit de montrer le résultat pour *deux* convexes. Soient donc A et B deux convexes de E et soit $\lambda \in [0,1]$. On a alors (en notant $\mu := 1-\lambda$) les égalités

$$\begin{aligned} & \lambda (A+B) + \mu (A+B) = (\lambda A + \lambda B) + (\mu A + \mu B) \\ = \; & \lambda A + (\lambda B + \mu A) + \mu B = \lambda A + (\mu A + \lambda B) + \mu B \\ = \; & \underbrace{(\lambda A + \mu A)}_{=A \text{ car } A \text{ convexe}} + \underbrace{(\lambda B + \mu B)}_{=B \text{ car } B \text{ convexe}} = A + B, \text{ ce qui conclut.} \end{aligned}$$

2. Un dessin suggère de montrer que le graphe de f reste toujours au-dessus de chaque droite asymptote. Soit φ une fonction affine telle que $f - \varphi \underset{\infty}{\longrightarrow} 0$. La fonction $F := f - \varphi$ est alors convexe, tend vers 0 en ∞ et l'on aimerait montrer la comparaison $F \geq 0$. Montrons pour cela que F décroît : elle tendra alors en ∞ vers son *infimum* (dans $\overline{\mathbb{R}}$), lequel sera nul (par unicité de la limite), d'où la conclusion $F \geq \inf F = 0$.

   Soit par l'absurde $a < b$ dans $\mathbb{R}_+$ tels que $\tau_a^b > 0$ (taux d'accroissement pour F). La fonction affine coïncidant avec F sur $\{a,b\}$ tend alors vers ∞ en ∞ ; or cette fonction "corde" *min*ore F hors du segment $[a,b]$ (*cf.* section 2.4.3), en particulier sur $[b,\infty[$, d'où la tendance $F \underset{\infty}{\longrightarrow} \infty$, contredisant l'hypothèse $F \underset{\infty}{\longrightarrow} 0$.

   FIG ? ?

   Remarque – Par décroissance de F, la comparaison $F \geq 0$ ou bien est stricte ou bien est une égalité au voisinage de ∞.

3.
   (a) Soient $p, q \in P$. On a alors les comparaisons

$$\begin{aligned} H_2 \binom{p}{q} &= -\sum_{e \in E} p_e \ln \frac{p_e}{q_e} = \sum_{e \in E} p_e \ln \frac{q_e}{p_e} \\ &\underset{\text{pour } \ln}{\overset{\text{Jensen}}{\leq}} \ln \sum_{e \in E} p_e \frac{q_e}{p_e} = \ln \sum_{e \in E} q_e = \ln 1 = 0, \end{aligned}$$

   d'où la négativité de H_2. Par stricte concavité de ln, on a égalité ssi la famille $\left(\frac{q_e}{p_e}\right)_{e \in E}$ est constante ; or les familles p et q ont même somme non nulle (à savoir 1), donc cette constance équivaut à l'égalité $p = q$.

   Remarque – Il serait élémentaire de donner sens à la concavité de H_2 sur P^2 – et un peu fastidieux de l'établir.

   (b) Soit $p \in P$ et abrégeons $C := \operatorname{Card} E$ (dont l'inverse fait sens vu que E est par hypothèse non vide). Notons u la famille constante valant $\frac{1}{C}$ (loi **u**niforme sur E). Le point 3a montre alors la comparaison

$$\begin{aligned} 0 \;\geq\; H_2 \binom{p}{u} &= -\sum_{e \in E} p_e \ln \frac{p_e}{\frac{1}{C}} = -\sum_{e \in E} p_e (\ln p_e + \ln C) \\ &= -\sum_{e \in E} p_e \ln p_e - \left(\sum_{e \in E} p_e\right) \ln C = H(p) - \ln C, \end{aligned}$$

d'où la comparaison voulue avec (toujours d'après le point 3a) égalité ssi $p = u$. (Ce point se généraliserait à peu de frais à *toutes* les distributions de probabilités sur E – y compris celles s'annulant.)

REMARQUE – **Culture entropique**. La fonction H est appelée ***entropie***[64]. Le résultat montré ici s'interprète comme suit : l'entropie d'un système fini est *maximale* lorsque ce système est décrit par une loi de probabilité *uniforme* (et ne peut plus alors "évoluer"). La lectrice et le lecteur souhaitant approfondir ces notions d'entropie gagneront à consulter l'œuvre de Ludwig BOLTZMANN en physique statistique.

4.

(a) Soient $a, \varepsilon \in R$. Abrégeons $g := \sqrt[n]{\prod_{i=1}^n a_i}$. Une inégalité arithmético-géométrique donne de suite

$$\frac{1}{n}\sum_{i=1}^n \varepsilon_i a_i \geq \sqrt[n]{\prod_{i=1}^n \varepsilon_i a_i} = \underbrace{\sqrt[n]{\prod_{i=1}^n \varepsilon_i}}_{=1}\underbrace{\sqrt[n]{\prod_{i=1}^n a_i}}_{=g} = g.$$

On a égalité si les $\varepsilon_i a_i$ sont égaux, cette valeur commune valant alors la moyenne g. On a donc égalité si $\varepsilon_i = \frac{g}{a_i}$ pour chaque $i \in [1, n|]$, ce qui conclut.

(b) Il suffit pour chaque $i \in [1, n|]$ de minorer

$$2 + a_i = 1 + 1 + a_i \overset{\text{comp. arith.-}}{\underset{\text{géométrique}}{\geq}} 3\ \sqrt[3]{1 \cdot 1 \cdot a_i}, \text{ d'où la comparaison}$$

$$\prod_{i=1}^n (2 + a_i) \geq \prod_{i=1}^n (3\ \sqrt[3]{a_i}) = \prod_{i=1}^n 3\ \sqrt[3]{\prod_{i=1}^n a_i} = 3^n\ \sqrt[3]{1} = 3^n,\ c.\ q.\ f.\ d.$$

Bonus : les comparaisons multipliées ne mettant en jeu que des réels strictement positifs, le cas d'égalité final équivaut à celui dans chacun des comparaisons multipliées, *i. e.* à $1 = 1 = a_i$ pour chaque $i \in [1, n|]$, *i. e.* à l'égalité des a_i (à 1).

(c) On adapte la même stratégie en réécrivant à $i \in [2, n|]$ fixé chaque produit $1 + a_i$ comme une somme de i termes afin de neutraliser l'exposant i :

$$(1 + a_i)^i = \left((i-1)\frac{1}{i-1} + a_i\right)^i$$
$$\overset{\text{comp. arith.-}}{\underset{\text{géométrique}}{\geq}} \left(i\sqrt[i]{\frac{1}{i-1}\frac{1}{i-1}\cdots\frac{1}{i-1}a_i}\right)^i = \frac{i^i}{(i-1)^{i-1}}a_i.$$

Multiplier ces comparaisons permet alors de conclure, vu d'une part le télescopage du produit $\prod_{i=2}^n \frac{i^i}{(i-1)^{i-1}} = \frac{n^n}{1^1}$, d'autre part l'hypothèse $\prod_{i=2}^n a_i = 1$.

[64]Elle apparaît dans l'article fondateur « A mathematical theory of communication » de Claude SHANNON publié en 1948 dans le *Bell System Technical Journal*. Le nom *entropie* semble apparaître pour la première fois avec Rudolf CLAUSIUS en 1865 dans son livre *Théorie Mécanique de la Chaleur* (contexte thermodynamique où elle était notée S).

Bonus : comme au point précédent, on a égalité à la fin ssi on a égalité partout, *i. e.* ssi $a_i = \frac{1}{i-1}$ pour chaque $i \in [1, n|]$, ce que bloque l'hypothèse $\prod_{i=2}^{n} a_i = 1$ (sauf si $n \leq 2$). Par conséquent, la comparaison est stricte dès que $n \geq 3$.

(d) Nous sous-entendrons le domaine de sommation $[1, n|]$ afin d'alléger les écritures.

Une condition $\sum a_i = 1$ *doit* nous faire penser à une comparaison de Jensen. Dans notre problème, la fonction $t \mapsto \frac{1}{\sqrt{1-t}}$ est convexe sur $]0, 1[$ en tant que composée de la fonction convexe $t \mapsto t^{-\frac{1}{2}}$ à droite par la fonction affine $t \mapsto 1 - t$, d'où la minoration

$$\sum \frac{a_i}{\sqrt{1-a_i}} \overset{\text{Jensen}}{\geq} \frac{1}{\sqrt{1 - \sum a_i \times a_i}} = \frac{1}{\sqrt{1 - \sum a_i^2}}.$$

Par ailleurs, la concavité de $t \mapsto \sqrt{t}$ et l'hypothèse $\sum a_i = 1$ nous permettent de majorer

$$\sum \sqrt{\frac{a_i}{n-1}} = n\sqrt{\frac{1}{n-1}} \sum \frac{1}{n}\sqrt{a_i} \overset{\text{Jensen}}{\leq} n\sqrt{\frac{1}{n-1}}\sqrt{\sum \frac{a_i}{n}} = \sqrt{\frac{n}{n-1}}.$$

Il suffit donc pour conclure de minorer $\frac{1}{\sqrt{1-\sum a_i^2}} \overset{?}{\geq} \sqrt{\frac{n}{n-1}}$. Or cette comparaison équivaut à $\frac{n-1}{n} \geq 1 - \sum a_i^2$, *i. e.* à $n \sum a_i^2 \geq 1$, ce qui est une comparaison de Cauchy-Schwarz $\sum 1^2 \sum a_i^2 \geq \left(\sum 1 a_i\right)^2$.

Bonus : le cas d'égalité impose équivaut à l'égalité dans toutes nos comparaisons utilisées, *i. e.* à l'égalité des a_i (à $\frac{1}{n}$ vu par ailleurs l'hypothèse $\sum a_i = 1$).

5. La première chose est d'intuiter la comparaison à montrer en regardant les cas simples, à savoir ici les petites valeurs de n. Quand $n = 1$, les deux membres à comparer sont égaux et il n'y a rien à en tirer. Regardons le cas $n = 2$: vu à $A, B, U, V \geq 0$ fixé les comparaisons

$$\begin{aligned} & \left(\sqrt{A+B}\sqrt{U+V}\right)^2 - \left(\sqrt{AU} + \sqrt{BV}\right)^2 \\ = \;& AU + AV + BU + BV - \left(AU + BV + 2\sqrt{ABUV}\right) \\ = \;& AV - 2\sqrt{ABUV} + BU = \left(\sqrt{AV} - \sqrt{BU}\right)^2 \geq 0, \end{aligned}$$

il est raisonnable d'espérer avoir la comparaison

$$\sqrt[n]{a_1 a_2 \cdots a_n} + \sqrt[n]{b_1 b_2 \cdots b_n} \leq \sqrt[n]{(a_1 + b_1)(a_2 + b_2) \cdots (a_n + b_n)}.$$

Nous en proposons *quatre* démonstrations. Le domaine de sommation/production $[1, n|]$ sera sous-entendu afin d'alléger les écritures.

(a) **par Hölder.** Suivant l'indication, énonçons formellement et utilisons la comparaison de Hölder sous sa forme générale[65] : pour chaques na-

[65] *cf.* exemple 2 section 2.2

turels p et q, pour chaque $\lambda \in [0,1]^q$ de somme 1 et pour chaque matrice $M \in M_{p,q}(\mathbb{R})$ à coefficients positifs, on a la majoration

$$\sum_{i=1}^{p} \prod_{j=1}^{q} m_{i,j}^{\lambda_j} \leq \prod_{j=1}^{q} \left(\sum_{i=1}^{p} m_{i,j} \right)^{\lambda_j}.$$

En particulier, remplacer $\binom{p}{q}$ par $\binom{2}{n}$ puis remplacer $\left(\begin{smallmatrix} m_{1,j} \\ m_{2,j} \end{smallmatrix}, \lambda_j \right)$ par $\left(\begin{smallmatrix} a_j \\ b_j \end{smallmatrix}, \frac{1}{n} \right)$ pour chaque $j \in [1,n|]$ donne

$$\prod_{j=1}^{n} \sqrt[n]{a_j} + \prod_{j=1}^{n} \sqrt[n]{b_j} \leq \prod_{j=1}^{n} \sqrt[n]{a_j + b_j}, \ c.\ q.\ f.\ d.$$

(b) **par Jensen.** Si l'un des a_i est nul, la somme $\sqrt[n]{a_1 a_2 \cdots a_n} + \sqrt[n]{b_1 b_2 \cdots b_n}$ se réduit alors à $\sqrt[n]{b_1 b_2 \cdots b_n}$ qui est clairement majoré par $\sqrt[n]{(a_1+b_1)(a_2+b_n)\cdots(a_n+b_n)}$ vu la positivité des a_j.

Dans le cas contraire, le réel $c_i := \frac{b_i}{a_i}$ fait sens pour chaque $i \in [1,n|]$ et la comparaison voulue se réécrit

$$\begin{aligned} 1 + \sqrt[n]{c_1 c_2 \cdots c_n} &\leq \sqrt[n]{(1+c_1)(1+c_2)\cdots(1+c_n)}, \\ i.\ e.\ \ln\left(1 + \sqrt[n]{\prod c_i}\right) &\leq \sum \frac{\ln(1+c_i)}{n}. \end{aligned}$$

Or la fonction $t \mapsto \ln(1+e^t)$ est dérivable de dérivée $t \mapsto \frac{1}{1+e^{-t}}$ croissante, donc est convexe, d'où les comparaisons

$$\begin{aligned} \ln\left(1 + \sqrt[n]{\prod c_i}\right) &= \ln\left(1 + \exp\left(\sum \frac{\ln c_i}{n}\right)\right) \\ &\overset{\text{Jensen}}{\leq} \sum \frac{1}{n} \underbrace{\ln(1+\exp(\ln c_i))}_{=\ln(1+c_i)}, \ c.\ q.\ f.\ d. \end{aligned}$$

(c) **inclusions et *minima*.** Pour chaque $r \in \mathbb{R}_+^n$, notons

$$M_r := \left\{ \frac{1}{n} \sum \lambda_i r_i \ ; \ \lambda \in \mathbb{R}_+^{*n} \text{ et } \prod \lambda_i = 1 \right\}.$$

D'une part, on vérifiera aisément les inclusions $M_{a+b} \subset M_a + M_b$, d'autre part, l'exercice 4a donne sens à et montre l'égalité $\min M_r = \sqrt[n]{\prod r_i}$ pour chaque $r \in \mathbb{R}_+^{*n}$. On en déduit les comparaisons

$$\begin{aligned} \prod \sqrt[n]{a_i + b_i} &= \min M_{a+b} \geq \min(M_a + M_b) = \min M_a + \min M_b \\ &= \sqrt[n]{\prod a_i} + \sqrt[n]{\prod b_i}, \ c.\ q.\ f.\ d. \end{aligned}$$

(d) **combinatoire.** Développons et utilisons des comparaison arithmético-géométriques :

$$\prod_{i=1}^{n}(a_i+b_i)=\sum_{I\amalg J=[1,n|]}\prod_{i\in I}a_i\prod_{j\in J}b_j=\sum_{p+q=n}^{p,q\in\mathbb{N}}\quad\sum_{\substack{\text{Card } I\,=\,p\\ \text{Card } J\,=\,q}}^{I\amalg J=[1,n|]}\prod_{i\in I}a_i\prod_{j\in J}b_j$$

$$\underset{\substack{\text{géométrique}\\ \text{à } (p,q) \text{ fixé}}}{\overset{\text{comp. arith.-}}{\geq}}\sum_{p+q=n}^{p,q\in\mathbb{N}}\binom{n}{p}\left[\prod_{\substack{\text{Card } I\,=\,p\\ \text{Card } J\,=\,q}}^{I\amalg J=[1,n|]}\left(\prod_{i\in I}a_i\right)\left(\prod_{j\in J}b_j\right)\right]^{\frac{1}{\binom{n}{p}}}$$

Le produit dans la parenthèse comporte $\binom{n}{p}n$ facteurs, dont $\binom{n}{p}p$ facteurs a_i et $\binom{n}{p}q$ facteurs b_j. Par symétrie, chacun des n réels a_i apparaît exactement $\binom{n}{p}\frac{p}{n}$ fois et chaque réel b_j apparaît $\binom{n}{p}\frac{q}{n}$ fois. En prenant la racine $\binom{n}{p}$-ième, il reste $\prod a_i^{\frac{p}{n}}\prod b_j^{\frac{q}{n}}$. En notant α et β les moyennes géométriques des a_i et b_j, la somme ci-dessus devient finalement

$\sum_{p+q=n}^{p,q\in\mathbb{N}}\binom{n}{p}\alpha^p\beta^q=(\alpha+\beta)^n$, d'où le résultat en prenant la racine n-ième.

(e) **par Mac Laurin (hors programme).** En notant $e_1, e_2, ..., e_n$ les fonctions symétriques élémentaires des c_i (*cf.* preuve 5b), les comparaisons de Mac Laurin[66] affirment la décroissance de la suite $\left(\sqrt[k]{\frac{e_k}{\binom{n}{k}}}\right)_{k\in[1,n|]}$, d'où l'on tire pour chaque $i\in[0,n|]$ la comparaison $\binom{n}{i}e_n^{\frac{i}{n}}\leq e_i$. Il en résulte les comparaisons

$$(1+\sqrt[n]{e_n})^n=\sum_{i=0}^{n}\binom{n}{i}e_n^{\frac{i}{n}}\leq\sum_{i=0}^{n}e_i=\prod_{i=1}^{n}(1+c_i)\text{ , } c.\ q.\ f.\ d.$$

6.

(a) Notons A, B, M les points du graphe de f d'abscisses respectives $a, b, \frac{a+b}{2}$. Réécrivons alors l'encadrement désiré sous la forme

$$(b-a)\,\frac{f(a)+f(b)}{2}\leq\int_a^b f\leq(b-a)\,f\left(\frac{a+b}{2}\right)$$

et interprétons géométriquement chacun de ses termes.

FIG

Le terme de gauche est l'aire d'un rectangle dont un des côtés horizontaux passe en le milieu de $[AB]$, *i. e.* vaut l'aire d'un trapèze rectangle de côté incliné dirigé par la corde $[AB]$.

[66] publiées en 1729 dans les *Philosophical Transactions of the Royal Society*

Le terme du milieu est l'aire sous le graphe de f.

Le terme de droite est l'aire d'un rectangle dont un des côtés horizontal passe en M, *i. e.* l'aire d'un trapèze rectangle de côté incliné dirigé par la tangente en M au graphe de f.

Le résultat est alors immédiat car la tangente en M est au-dessus du graphe de f qui lui-même est au-dessus de la corde $[AB]$ (tout cela en vertu de la concavité de f).

Proprement, il suffit d'intégrer sur $[a,b]$ les comparaisons traduisant ces positionnements géométriques, à savoir

$$\frac{f(b)-f(a)}{b-a}(\mathrm{Id}-a)+f(a)\leq f\leq f'(m)(\mathrm{Id}-m)+f(m) \text{ où } m:=\frac{a+b}{2}.$$

Remarque – La dérivabilité ne sert que pour la comparaison de droite et est en fait superflue : en effet, d'une part on pourrait utiliser une *demi*-tangente (au lieu d'une tangente), d'autre part (sans faire appel aux demi-tangentes hors programme) intégrer lorsque t décrit $[a,b]$ les majorations $\frac{f(t)+f(2m-t)}{2}\leq f\left(\frac{t+(2m-t)}{2}\right)$ (lesquelles découlent de la concavité de f) donne directement $\int_a^b f\leq (b-a)f(m)$.

(b) Soit $u\in\,]0,1]$. D'après le point 6a où l'on remplacé $\binom{a}{b}$ par $\binom{0}{u}$, on a la comparaison $\frac{f(0)+f(u)}{2}\leq\frac{1}{u-0}\int_0^u f$, *i. e.* $uf(u)\leq 2\left(\int_0^u f\right)-u$, laquelle reste valide lorsque $u=0$.

Abrégeons $I:=\int_0^1 f$ et $J:=\int_0^1 tf(t)\mathrm{d}t$. Intégrer la dernière comparaison lorsque u décrit $[0,1]$ donne alors

$$J\leq 2\int_0^1\left(\int_0^u f\right)\mathrm{d}u-\int_0^1 u\ \mathrm{d}u.$$

La dernière intégrale vaut $\left[\frac{u^2}{2}\right]_{u=0}^1=\frac{1}{2}$ et l'intégrale "double" s'évalue en utilisant l'égalité donnée (conséquence d'un théorème de Fubini hors-programme : intuitivement, reparamétrer le triangle $\left\{(t,u)\in[0,1]^2\ ;\ t\leq u\right\}$ où l'on intègre en droites horizontales ou verticales) :

$$\int_{u=0}^1\int_{t=0}^u f(t)\ \mathrm{d}t\ \mathrm{d}u\overset{\text{indication}}{=}\int_{t=0}^1\int_{u=t}^1 f(t)\ \mathrm{d}u\ \mathrm{d}t$$
$$=\int_{t=0}^1(1-t)f(t)\ \mathrm{d}t=I-J.$$

La dernière comparaison s'écrit donc $J\leq 2I-2J-\frac{1}{2}$, *i. e.* $\frac{3}{2}J\leq I-\frac{1}{4}$. Il suffirait alors pour conclure d'avoir la majoration $I-\frac{1}{4}\leq I^2$, laquelle s'écrit encore $0\leq\left(I-\frac{1}{2}\right)^2$, ce qu'on a.

7. Tout d'abord, pour chaque $e\in E$ la "longueur" $L(e)$ fait sens car l'intégrande $\sqrt{1+e'^2}$ est de classe C^1 (donc intégrable). Soient à présent $f\leq g$ dans E. La comparaison alors voulue $L(f)\leq L(g)$ se réécrit $\int_a^b C\circ f'\leq\int_a^b C\circ g'$ où l'on a noté C la fonction $t\mapsto\sqrt{1+t^2}$ suggérée. Cette dernière étant convexe

sur $\mathbb{R}$, son graphe est au-dessus de chacune de ses tangentes, d'où les comparaisons

$$\forall u,v\in[a,b]\,,\ C(v)-C(u)\geq C'(u)\cdot(v-u)\,.$$

Y remplacer à $t\in[a,b]$ fixé $\binom{u}{v}$ par $\binom{f'(t)}{g'(t)}$ puis intégrer lorsque t décrit $[a,b]$ permet alors de minorer

$$L(g)-L(f)\geq\int_a^b(C'\circ f')\times(g'-f')=\int_a^b\Gamma\delta' \text{ où } \begin{pmatrix}\delta\\\Gamma\end{pmatrix}:=\begin{pmatrix}g-f\\C'\circ f'\end{pmatrix}.$$

Or d'une part δ est positive vu l'hypothèse $f\leq g$, d'autre part Γ décroît comme composée d'une fonction croissante (dérivée de la fonction convexe C) par une décroissante (dérivée de la fonction concave f). Puisque Γ est de classe C^1 (on utilise le fait que f est de classe C^2), on peut d'une part donc affirmer $\Gamma'\leq 0$, d'autre part intégrer par parties puis minorer

$$\int_a^b\Gamma\delta'=[\Gamma\ \delta]_a^b+\int_a^b\underbrace{-\Gamma'}_{\geq0}\underbrace{\delta}_{\geq0}\geq\Gamma(b)\underbrace{\delta(b)}_{=0}-\Gamma(a)\underbrace{\delta(a)}_{=0}=0, \text{ ce qui conclut}$$

FIG

Remarque – La convexité de g a servi *uniquement* pour donner sens à $L(f)=\int_a^b C\circ f'$ (la fonction monotone f' étant en effet intégrable). De même, on peut supposer f seulement de classe C^1 : la fonction Γ décroissant et restant positive, le second théorème de la moyenne (hors programme) nous donne alors un $t\in[a,b]$ tel que $\int_a^b\Gamma\delta'=\Gamma(a)\int_a^t\delta'\overset{\delta(a)=0}{=}\Gamma(a)\,\delta(t)\geq 0$, ce qui conclut.

8.

(a) Commençons par réécrire le membre à encadrer de manière plus homogène en répartissant le $1=\frac12+\frac12$ chez tout le monde :

$$\begin{aligned}&\frac{f(0)}{2}+f(1)+f(2)+\cdots+f(n-1)+\frac{f(n)}{2}\\=\ &\frac{f(0)}{2}+\left(\frac{f(1)}{2}+\frac{f(1)}{2}\right)+\cdots+\left(\frac{f(n-1)}{2}+\frac{f(n-1)}{2}\right)+\frac{f(n)}{2}\\=\ &\left(\frac{f(0)}{2}+\frac{f(1)}{2}\right)+\cdots+\left(\frac{f(n-1)}{2}+\frac{f(n)}{2}\right)=\sum_{i=1}^n\frac{f(i)+f(i-1)}{2}.\end{aligned}$$

On en déduit les équivalences

$$\begin{aligned}&0\leq\frac{f(0)}{2}+f(1)+f(2)+\cdots+f(n-1)+\frac{f(n)}{2}-\int_0^n f\leq\frac{f'(n)-f'(0)}{8}\\&\iff 0\leq\sum_{i=1}^n\frac{f(i-1)+f(i)}{2}-\sum_{i=1}^n\int_{i-1}^i f\leq\sum_{i=1}^n\frac{f'(i)-f'(i-1)}{8}\\&\iff 0\leq\sum_{i=1}^n\left(\frac{f(i-1)+f(i)}{2}-\int_{i-1}^i f\right)\leq\sum_{i=1}^n\frac{f'(i)-f'(i-1)}{8}\\&\underset{\text{pour chaque } i\in[1,n|]}{\overset{\text{définir } f_i:=t\mapsto f(t+i-1)}{\iff}} 0\leq\sum_{i=1}^n\left(\frac{f_i(0)+f_i(1)}{2}-\int_0^1 f_i\right)\leq\frac{f_i'(1)-f_i'(0)}{8}\end{aligned}$$

Par conséquent, si l'on montre l'encadrement désiré quand $n = 1$ pour *chaque* fonction f comme dans l'énoncé, on pourra alors à $i \in [1, n|]$ fixé y remplacer f par f_i, additionner le tout et obtenir l'encadrement ci-dessus.

(b) Puisque nous sommes ramenés au cas $n = 1$, l'encadrement à montrer s'écrit

$$0 \leq \frac{f(0) + f(1)}{2} - \int_0^1 f \leq \frac{f'(1) - f'(0)}{8}.$$

Notons A et B les points du graphe de f d'abscisses respectives a et b et appelons T le point d'intersection des tangentes en A et en B. Notons α et β les fonctions affines de graphes les tangentes resp. en A et en B au graphe de f et appelons φ la fonction affine coïncidant avec f sur $[0, 1]$. La convexité de f permet alors d'affirmer que son graphe est d'une part au-dessus de ceux de α et β et d'autre part en dessous de celui de φ (proprement, on a sur $[0, 1]$ les comparaisons $\max\{\alpha, \beta\} \leq f \leq \varphi$).

FIG

Le terme-différence du milieu représente l'aire (signée) de la lunule comprise entre la corde $[AB]$ et le graphe de f restreinte à $[0, 1]$: cette aire est positive par convexité de f, d'où la comparaison de gauche (proprement, intégrer la comparaison $f \leq \varphi$ sur $[0, 1]$).

Par ailleurs, vu que l'aire du triangle ABT majore celle de la lunule (proprement, intégrer sur $[0, 1]$ la comparaison $\varphi - \max\{\alpha, \beta\} \geq \varphi - f$), il suffit pour conclure de majorer $\mathcal{A}(ABT)$ par $\frac{f'(1) - f'(0)}{8}$. Cela est même nécessaire puisque la comparaison de droite doit être valable pour chaque fonction convexe ayant mêmes dérivées que f en 0 et en 1 et dont le graphe colle les tangentes en 0 et 1 *aussi près que l'on veut*.

(c) Regardons la comparaison restant à montrer

$$\frac{f(0) + f(1)}{2} - \int_0^1 f \leq \frac{f'(1) - f'(0)}{8}.$$

Elle est tout d'abord *homogène* en f, au sens où pour chaque réel $a > 0$ cette comparaison équivaut à celle où l'on a remplacé f par af. Montrons par ailleurs qu'elle est invariante par ajout à f de fonctions affines. Soit $\lambda, \mu \in \mathbb{R}$ et notons $F := f + \lambda \operatorname{Id} + \mu$. On a alors $F' = f' + \lambda$ et le λ rajouté disparaît dans la différence $F'(1) - F'(0)$, donc le membre de droite est préservé. Par ailleurs, le milieu se voit rajouter $\frac{(\lambda 0 + \mu) + (\lambda 1 + \mu)}{2} = \frac{\lambda}{2} + \mu$ et l'intégrale se voit rajouter $\int_0^1 (\lambda \operatorname{Id} + \mu) = \frac{\lambda}{2} + \mu$, donc le membre de gauche est inchangé.

Par conséquent, nous pouvons imposer $f(0) = 0$ par "translation verticale" (soustraire $f(0)$ à f), puis $f'(0) = 0$ par "ajout de pente" (soustraire $f'(0) \operatorname{Id}$ à f). La fonction f est alors positive (car son graphe est au-dessus de sa tangente en l'origine), d'où $f(1) \geq 0$: si cette image était nulle, la fonction f serait négative (car son graphe est en dessous de la corde reliant $\binom{0}{0}$ et $\binom{f(1)}{1}$)), donc nulle et la comparaison désirée serait triviale. On peut donc imposer $f(1) > 0$ et même $f(1) = 1$ en divisant f par $f(1)$.

FIG

(d) Notons $\alpha := f'(1)$ l'unique paramètre restant. Vu l'égalité $y_B - y_T = \alpha(x_B - x_T)$, *i. e.* $1 = \alpha(1 - x_T)$, on obtient l'abscisse $x_T = 1 - \frac{1}{\alpha}$ (*sanity check* : on a $\alpha \geq 1$ car la pente de la tangente en B majore celle de la corde $[AB]$, d'où $x_T \in [0,1]$). On en déduit les équivalences

$$\mathcal{A}(ABT) \leq \frac{f'(1) - f'(0)}{8} \iff \frac{1}{2}x_T \leq \frac{\alpha}{8} \iff 1 - \frac{1}{\alpha} \leq \frac{\alpha}{4}$$
$$\iff \alpha^2 - 4\alpha + 4 \geq 0 \iff (\alpha - 2)^2 \geq 0, \text{ ce qu'on a.}$$

Remarque – Lorsque f est de classe C^2, donnons une autre preuve de la comparaison de droite[67]. On définit pour cela une suite $(\alpha_n)_{n\in\mathbb{N}}$ de fonctions $[0,1[\longrightarrow \mathbb{R}$ par la récurrence[68] $\alpha_0 := 1$ et $\forall n \in \mathbb{N}^*$, $\begin{cases} \alpha_n' = \alpha_{n-1} \\ \int_0^1 \alpha_n = 0 \end{cases}$. Sur $[0,1[$, on obtient aisément $\alpha_1 = X - \frac{1}{2}$ et $\alpha_2 = \frac{X^2}{2} - \frac{X}{2} + \frac{1}{12} \geq \alpha_2\left(\frac{1}{2}\right) = -\frac{1}{24}$. Prolongeons chaque α_n sur tout $\mathbb{R}$ par 1-périodicité. Intégrer deux fois par parties livre alors l'égalité

$$\int_0^1 f = [f\alpha_1]_0^1 - [f'\alpha_2]_0^1 + \int_0^1 f''\alpha_2.$$

Or, d'une part les égalités $\alpha_2(1^-) = \frac{1}{12} = \alpha_2(0^+)$ donnent $[f'\alpha_2]_0^1 = \frac{1}{12}[f']_0^1$, d'autre part les égalités $\alpha_1(1^-) = \frac{1}{2}$ et $\alpha_1(0^+) = -\frac{1}{2}$ donnent $[f\alpha_1]_0^1 = \frac{f(0)+f(1)}{2}$. On conclut en majorant[69]

$$\frac{f(0) + f(1)}{2} - \int_0^1 f = [f'\alpha_2]_0^1 - \int_0^1 \alpha_2 f'' \underset{f''\geq 0}{\overset{\alpha_2 \geq -\frac{1}{24}}{\leq}} \frac{1}{12}[f']_0^1 + \frac{1}{24}[f']_0^1 = \frac{1}{8}[f']_0^1.$$

9.

(a) Pour chaque $t \in I$, la fonction constante égale à t tombe dans Φ_t, d'où les majorations

$$c(t) = \inf_{\varphi \in \Phi_t} \int_0^1 f \circ \varphi \leq \int_{u=0}^1 f(t)\,\mathrm{d}u = f(t). \text{ Il en résulte } c \leq f.$$

Soit $c \geq f$ concave, soit $t \in I$. Pour chaque $\varphi \in \Phi_t$, on peut majorer (*cf.* exercice d'application 1 section 2.2)

$$\begin{aligned} c(t) &= c\left(\int_0^1 \varphi\right) \overset{\text{Jensen}}{\leq} \int_0^1 c \circ \varphi, \text{ d'où (faisant varier } \varphi) \\ c(t) &\leq \inf_{\varphi \in \Phi_t} \int_0^1 c \circ \varphi = C(t). \text{ On en déduit } c \leq C. \end{aligned}$$

Soient $a, b \in I$, soient $\lambda, \mu > 0$ de somme 1, soit $(\alpha, \beta) \in \Phi_a \times \Phi_b$ et notons $\varphi := \begin{cases} [0,1] & & \mathbb{R}_+ \\ [0,\lambda[\ni t & \longmapsto & \alpha\left(\frac{t}{\lambda}\right) \\ [\lambda, 1] \ni t & \longmapsto & \beta\left(\frac{t-\lambda}{\mu}\right) \end{cases}$ (suggéré par l'énoncé). Alors φ

[67] La lectrice et le lecteur intéressés davantage pourront enquêter sur l'identité d'Euler-Mac Laurin.

[68] *Culture* : les $n!\alpha_n$ sont les polynômes de Bernoulli.

[69] Bien repérer où intervient la convexité de f.

fait bien sens (les arguments de $\alpha\left(\frac{t}{\lambda}\right)$ et $\beta\left(\frac{t-\lambda}{\mu}\right)$ tombent bien dans le domaine $[0,1]$ de α et β), est continue par morceaux (comme recollement de deux fonctions continues par morceaux) et d'intégrale

$$\int_0^1 \varphi = \int_0^\lambda \varphi + \int_\lambda^1 \varphi = \int_0^\lambda \alpha\left(\frac{t}{\lambda}\right) \mathrm{d}t + \int_\lambda^1 \beta\left(\frac{u-\lambda}{\mu}\right) \mathrm{d}u$$

$$\underset{\substack{U:=\frac{u-\lambda}{\mu} \text{ et } \mathrm{d}u=\mu \mathrm{d}U \\ T:=\frac{t}{\lambda} \text{ et } \mathrm{d}t=\lambda \mathrm{d}T}}{\overset{\text{reparamétrages}}{=}} \int_0^1 \alpha(T)\,\lambda \mathrm{d}T + \int_0^1 \beta(U)\,\mu \mathrm{d}U = \lambda \int_0^1 \alpha + \mu \int_0^1 \beta$$

$= \lambda a + \mu b$, d'où l'appartenance $\varphi \in \Phi_{\lambda a+\mu b}$.

On en déduit les majorations

$$C(\lambda a + \mu b) \leq \int_0^1 f \circ \varphi \underset{\text{analogue}}{\overset{\text{calcul}}{=}} \lambda \int_0^1 f \circ \alpha + \mu \int_0^1 f \circ \beta,$$

d'où (en faisant varier α puis β) la comparaison $C(\lambda a+\mu b) \leq \lambda C(a) + \mu C(b)$. Il en résulte la convexité de C.

(b) On applique ce qui précède en remplaçant (I, f) par $(\mathbb{R}_+, -F)$: la fonction $c := -C$ fait alors sens (chaque intégrande $F \circ \varphi$ est intégrable car F est monotone), est concave et majore F.

Explicitement, on a $c(t) = \sup_{\varphi \in \Phi_t} \int_0^1 F \circ \varphi$ pour chaque réel $t \geq 0$.

Soit $\varphi \in \Phi_0$: alors φ est positive (car à valeurs dans $I = \mathbb{R}_+$), continue sauf en un nombre fini de points (car continue par morceaux sur un segment) et d'intégrale nulle, donc est nulle sauf peut-être en ses points de discontinuité, donc $f \circ \varphi$ est nulle par morceaux, *a fortiori* d'intégrale nulle. Faire varier φ donne alors $c(0) = 0$.

Soit $t > 0$ et soit $\varphi \in \Phi_t$. Puisque F est sous-additive, une récurrence immédiate livre pour chaque $(n, a) \in \mathbb{N} \times \mathbb{R}_+$ la majoration $F(na) \leq nF(a)$. On peut alors majorer

$$F \circ \varphi = F \circ \left(\frac{\varphi}{t} t\right) \overset{F \text{ croît}}{\leq} F \circ \left(\left\lceil \frac{\varphi}{t} \right\rceil t\right)$$

$$\underset{\text{additive}}{\overset{F \text{ sous-}}{\leq}} \left\lceil \frac{\varphi}{t} \right\rceil F(t) \leq \left(\frac{\varphi}{t} + 1\right) F(t),$$

d'où les majorations

$$\int_0^1 F \circ \varphi \leq \int_0^1 \left(\frac{\varphi}{t} + 1\right) F(t) = \left(\frac{\int_0^1 \varphi}{t} + 1\right) F(t) \overset{\varphi \in \Phi_t}{=} 2F(t).$$

Faire varier φ donne alors $c(t) \leq 2F(t)$, comparaison encore valide si $t = 0$.

Enfin, à $t \in \mathbb{R}_+$ fixé, puisque F est positive, l'intégrale $\int_0^1 F \circ \varphi$ est positive pour chaque $\varphi \in \Phi_t$, d'où (faisant varier φ) la positivité de $c(t)$. La fonction convexe $-c$ est par conséquent majorée (par 0) sur $\mathbb{R}_+$, donc décroît[70], d'où la croissance de c.

[70] *cf.* exercice d'application 3 section 2.4.3

(a) i. Supposons par exemple $f \geq 0$. Le barycentre trivial $1f + 0g$ convient alors.

ii. La fonction $M := \max\{f, g\}$ est continue sur le segment S (comme *supremum* de fonctions continues car dérivables), donc y atteint son minimum : soit $s \in S$ tel que $M \geq M(s)$. Abrégeons $\binom{\alpha}{\beta} := \binom{f'(s)}{g'(s)}$.

Si $f(s) < 0$, on a alors $f < 0$ au voisinage de s, donc $\max\{f, g\}$ (qui est positive) ne peut valoir f sur ce voisinage, donc M y vaut g ; puisque M admet un minimum local en s, la fonction g également ; or la convexité de g force ce minimum à être global, d'où $g \geq g(s) = M(s) \geq 0$, ce qu'on a exclu. Il en résulte $f(s) \geq 0$ et échanger f et g conduirait de même à $g(s) \geq 0$.

Supposons $\alpha, \beta > 0$. On a alors $\begin{cases} f < f(s) \\ g < g(s) \end{cases}$ à gauche de s, d'où $M < M(s)$ à gauche de s, ce qui force $s = \min S$ (sinon s ne minimiserait pas M). Puisque f' croît (par convexité de f), on a les comparaisons $f' \geq f'(\min S) = \alpha > 0$, donc f croît, d'où les comparaisons $f \geq f(\min S) = f(s) \geq 0$, ce qui est exclu. On montrerait de même l'implication $\alpha, \beta < 0 \Longrightarrow g \geq 0$ dont le conséquent est exclu. Reste finalement le dernier cas $\alpha\beta \leq 0$, *c. q. f. d.*

iii. Puisque α et β sont de signes opposés, le réel 0 appartient au segment $[\alpha, \beta]$, d'où deux réels $\lambda, \mu \geq 0$ de somme 1 tels que $0 = \lambda\alpha + \mu\beta$. La fonction $\lambda f + \mu g$ est alors convexe (combinaison positive de fonctions convexes) et dérivable en s de dérivée nulle, donc est minimale en s. Vu la positivité des quatre réels $\lambda, \mu, f(s), g(s)$, on peut conclure $\lambda f + \mu g \geq 0$.

1 : cas affine,ciseaux

Remarque – La figure ci-dessus devrait éclairer les lignes qui précèdent dans le cas affine (fermer les "ciseaux" formés par les deux graphes) et donner un appui géométrique à la démarche suivante consistant à se ramener à ce cas : traiter le cas où f est de signe constant, évoquer a un zéro de f, se ramener à $\begin{cases} \alpha \neq 0 \\ g(a) \geq f(a) \end{cases}$, puis minorer $\begin{cases} f \geq \alpha\delta \\ g \geq \beta\delta \end{cases}$ où $\delta := \mathrm{Id} - a$.

FIG : minorer par le cas affine

Une troisième démarche consisterait (*modulo* cas triviaux) à évoquer deux réels $a \neq b$ tels que $\begin{cases} f \geq f(a) < 0 \\ g \geq g(b) < 0 \end{cases}$, à montrer $f'g' \geq 0$ sur $[a, b]$, puis à évoquer deux réels $\alpha < \beta$ tels que $f(\alpha) = 0 = g(\beta)$; chaque m dans $[\alpha, \beta]$ pourrait alors être utilisé comme au point 9(a)iii. Notre première preuve part directement d'un tel m minimisant $\max\{f, g\}$, ce qui nous a été suggéré par la figure :

FIG

(b) Quitte à remplacer f et g par les prolongements continus de leur restriction à $\mathring{S}$, ce qui ne change ni leur convexité ni l'hypothèse $\max\{f,g\} \geq 0$ ni la conclusion (à détailler), on peut les imposer continues. La fonction continue M réalise ainsi toujours son minimum en un certain $s \in S$.

Sans l'hypothèse de dérivabilité, nous allons reprendre le même raisonnement avec les dérivées généralisées à gauche $\binom{\overleftarrow{\alpha}}{\overleftarrow{\beta}} := \binom{f'_g(s)}{g'_g(s)}$ ou celles à droite $\binom{\overrightarrow{\alpha}}{\overrightarrow{\beta}} := \binom{f'_d(s)}{g'_d(s)}$ (l'un des deux couples fait sens dans $\overline{\mathbb{R}}$ par convexité).

Supposons $s = \min S$. Si $\overrightarrow{\alpha}$ et $\overrightarrow{\beta}$ sont finis, ils vérifient les implications $\begin{cases} \overrightarrow{\alpha} \geq 0 \Longrightarrow f \geq 0 \\ \overrightarrow{\beta} \geq 0 \Longrightarrow g \geq 0 \end{cases}$, d'où $\overrightarrow{\alpha}, \overrightarrow{\beta} < 0$ et la contradiction $M < M(s)$ à droite de s. Par conséquent, chaque fonction croissante τ_s tend en s vers $-\infty$, donc reste < 0 au voisinage de s, d'où la même contradiction. On exclurait de même le cas $s = \max S$ avec les dérivées à gauche.

Si enfin $s \in \mathring{S}$, les dérivées $\overleftarrow{\alpha}, \overleftarrow{\beta}, \overrightarrow{\alpha}, \overrightarrow{\beta}$ sont alors finies et l'on a d'une part $\overleftarrow{\alpha} > 0$ ou $\overrightarrow{\alpha} < 0$ (sinon $f \geq 0$), d'autre part l'implication $\begin{cases} \overleftarrow{\alpha} > 0 \Longrightarrow \overleftarrow{\beta} \leq 0 \\ \overrightarrow{\alpha} < 0 \Longrightarrow \overrightarrow{\beta} \geq 0 \end{cases}$ (sinon $M < M(s)$ quelque part), ce qui permet d'obtenir $\overrightarrow{\alpha}\overrightarrow{\beta} \leq 0$ ou $\overleftarrow{\alpha}\overleftarrow{\beta} \leq 0$ et de conclure comme en 9(a)iii.

(c) Remplaçons à présent S par un intervalle I. Si $i := \inf M$ est atteint, rien à changer. Supposons donc le contraire. La fonction convexe M ne peut alors décroître puis croître, donc est monotone et tend vers i en l'une des bornes de I. Si cette borne est finie, on y prolonge alors f, g, M par continuité et on est ramené au cas où $i = \min M$. On supposera donc $M \underset{\infty}{\longrightarrow} i$ (noter au passage que la positivité de M se propage à son *infimum* : $i \geq 0$).

Si l'on montre (comme suggéré) que f tend vers i en ∞, la fonction convexe f admettra alors une asymptote en ∞ et son graphe sera au-dessus de cette asymptote (d'après l'exercice d'entraînement 2), d'où $f \geq i \geq 0$ et l'on pourra conclure (idem pour g). Soient donc par l'absurde $\varepsilon > 0$ et $a, b \in I^{\mathbb{N}}$ croissant chacune strictement vers ∞ telles que $\forall n \in \mathbb{N}$, $\begin{cases} |f(a_n) - i| > \varepsilon \\ |g(b_n) - i| > \varepsilon \end{cases}$.

Soit $A \in I$ tel que $i < M < i + \varepsilon$ sur $[A, \infty[$ (permis car $M \underset{\infty}{\longrightarrow} i$). Quitte à rétrospectivement décaler les suites a et b, on peut imposer qu'elle prennent des valeurs au-delà de A. Pour chaque $n \in \mathbb{N}$, l'intervalle $]i, i+\varepsilon[$ contient alors $M(b_n)$ mais pas $g(b_n)$, donc la fonction M ne peut coïncider avec g en b_n, d'où $f(b_n) = M(b_n)$. Soient enfin $p, q, N \in \mathbb{N}$ tels que $b_p < a_N < b_q$ (permis car $a, b \longrightarrow \infty$) : selon la position de $f(a_N)$ par rapport à $i \pm \varepsilon$, on a alors l'une des comparaisons $f(b_p) < f(a_N)$ ou $f(a_N) < f(b_q)$, donc f admet une pente strictement positive, d'où (par convexité) l'absurde tendance $M \geq f \longrightarrow \infty$.

(d) $\boxed{\Longrightarrow}$ Soient $\lambda \in [0,1]^n$ tel que $\begin{cases} \sum_{i=1}^n \lambda_i = 1 \\ \sum_{i=1}^n \lambda_i f_i \geq 0 \end{cases}$. Soit $a \in I$. Soit $k \in [1, n|]$ tel que $f_k(a) \geq f_i(a)$ pour chaque $i \in [1, n|]$. On a alors les comparaisons

$$\max_{i\in[1,n|]} f_i(a) = f_k(a) = 1 f_k(a) = \sum_{i=1}^{n} \lambda_i f_k(a)$$
$$\underset{\text{à } i \text{ fixé}}{\overset{\lambda_i\geq 0}{\geq}} \sum_{i=1}^{n} \lambda_i f_i(a) \geq 0,\ c.\ q.\ f.\ d.$$

$\boxed{\Longleftarrow}$ On raisonne par récurrence, initiée au rang 2 par le point 9c (le rang 1 est trivial). Supposons $n \geq 3$ et abrégeons $\binom{a}{b} := \binom{f_1}{f_2}$ En réécrivant $\max_{i\in[1,n|]} f_i = \max\{\max\{a,b\}, f_3, f_4, ..., f_n\}$ où la fonction $\max\{a,b\}$ est convexe comme *supremum* de deux fonctions convexes, on peut (par hypothèse de récurrence) évoquer $n-1$ réels $\lambda, \lambda_3, \lambda_4, ..., \lambda_n \geq 0$ de somme 1 tels que $\lambda \max\{a,b\} + \sum_{i=3}^{n} \lambda_i f_i \geq 0$. Or le membre de gauche se réécrit $\max\left\{\begin{smallmatrix}\lambda a+\sum_{i=3}^{n}\lambda_i f_i \\ \lambda b+\sum_{i=3}^{n}\lambda_i f_i\end{smallmatrix}\right\}$ où les deux sommes entre les accolades sont convexes[71], d'où (d'après l'initialisation) deux réels $\alpha, \beta \geq 0$ de somme 1 tels que $\begin{smallmatrix}\alpha\left(\lambda a+\sum_{i=3}^{n}\lambda_i f_i\right) \\ +\beta\left(\lambda b+\sum_{i=3}^{n}\lambda_i f_i\right)\end{smallmatrix} \geq 0$. Le membre de gauche se réécrivant $(\lambda\alpha) f_1 + (\lambda\beta) f_2 + \sum_{i=3}^{n} \lambda_i f_i$ où les coefficients ont pour somme 1, on a terminé.

Remarque – Nous avons montré que *l'enveloppe convexe de chaque ensemble fini de fonctions convexes contient une fonction positive ssi le maximum de cet ensemble est positif.* Cela devient faux pour un ensemble *infini* : la suite $\left(-\frac{1}{n}\right)_{n\in\mathbb{N}^*}$ a en effet un *supremum* positif (car nul) mais aucune combinaison convexe de ses termes n'est positive (car chaque telle combinaison est une suite constante strictement négative).

[71] comme combinaisons linéaires positives de fonctions convexes