$\boxed{\text{Quelques inégalités mathématiques}}$

Christian Carimalo

1 $\boxed{a_1^n + a_2^n + \cdots + a_n^n - na_1a_2\cdots a_n \geq 0}$ (1)

1.1 Démonstration

Tous les nombres réels $a_1, \cdots, a_n$ sont supposés positifs. Commençons par le cas $n = 2$. On a $f(a_2) = a_2^2 + a_1^2 - 2a_2a_1 = (a_2 - a_1)^2$ et $f(a_2)$ est positif ou nul, le dernier cas étant réalisé *si et seulement si* $a_2 = a_1$. La propriété est donc vraie pour $n = 2$. Supposons-la vraie pour l'entier n et considérons $f_{n+1}(a_{n+1}) = a_{n+1}^{n+1} + a_n^n + \cdots + a_1^n - (n+1)a_{n+1}a_n \cdots a_1$. On a $f'(a_{n+1}) = (n+1)\left[a_{n+1}^n - a_n \cdots a_1\right]$ et cette dérivée est nulle lorsque $a_{n+1} = a_0 = \sqrt[n]{a_n \cdots a_1}$, auquel cas $f_{n+1}(a_{n+1})$ atteint son <u>seul</u> minimum $f_{n+1}(a_0) = a_n^{n+1} + \cdots a_1^{n+1} - n\,(a_n \cdots a_1)^{1+1/n}$. En posant $b_k = a_k^{1+1/n}$, cette valeur s'exprime comme $f(a_0) = b_n^n + \cdots b_1^n - nb_n \cdots b_1$, qui est une quantité positive par hypothèse, et nulle si et seulement si $b_1 = b_2 = \cdots = b_n$, soit encore $a_1 = \cdots = a_n$. On en déduit que $f(a_{n+1})$ est positif pour tout a_{n+1}. Si $f(a_{n+1})$ est nul, alors $a_{n+1} = a_0$ <u>et</u> $f(a_0) = 0$, ce qui, on vient de voir, n'est réalisé que si et seulement si $a_1 = \cdots = a_n$, auquel cas $a_0 = \sqrt[n]{a_n \cdots a_1} = a_1 = \cdots = a_n$. La propriété est donc vraie aussi pour $n + 1$ et est ainsi vraie pour tout n.

1.2 Application

On notera que dans (1) les nombres a_k peuvent être considérés comme les valeurs propres d'une matrice $n \times n$ A, positive. Comme $\mathrm{Trace}\, A^n = a_1^n + a_2^n + \cdots + a_n^n$ et $\det A = a_1a_2 \cdots a_n$, on a donc pour une telle matrice

$$\boxed{\mathrm{Trace}\, A^n \geq n \det A} \tag{1.1}$$

2 $\boxed{B(x) = x^n - 1 - n(x-1) \geq 0 \quad (n \geq 2)\ \textbf{(Bernoulli)}}$ (2)

$B(1) = 0$; $B'(x) = n(x^{n-1} - 1)$ est nul pour $x = 1$ seulement ; en ce point, $B(x)$ prend sa seule valeur minimum, qui est nulle. Par conséquent, $B(x) \geq 0$ pour tout x. On en déduit

$$\boxed{\sum_{k=1}^{m} x_k^n \geq \sum_{k=1}^{m} x_k + m(n-1)} \tag{2.1}$$

En particulier, si les x_i sont les m valeurs propres d'une matrice $m \times m$ A, cette inégalité devient

$$\boxed{\text{Trace}\, A^n \geq \text{Trace}\, A + m(n-1)} \tag{2.2}$$

En posant $x = 1 + y$, l'inégalité donne aussi

$$\boxed{(1+y)^n \geq 1 + ny} \tag{2.3}$$

inégalité qui se démontre aussi aisément par récurrence.

3 Inégalité de Gibbs

Considérons la fonction $G(x) = x \ln x + 1 - x$ (pour $x > 0$!). On a $G(1) = 0$ et $G'(x) = \ln x$. La dérivée s'annule uniquement pour $x = 1$ et cette valeur de x correspond au minimum nul de $G(x)$. On en conclut

$$G(x) = x \ln x + 1 - x \geq 0 \tag{3.1}$$

l'égalité étant réalisée uniquement pour $x = 1$. Changeons ensuite x en x/y avec $y > 0$; on obtient

$$x\,(\ln x - \ln y) \geq x - y \tag{3.2}$$

l'égalité se réalisant uniquement si $x = y$. Envisageons ensuite deux séries de n nombres chacune, l'une $\{x_i\}$, l'autre $\{y_i\}$. L'inégalité (3.2) étant valable pour chaque indice i, on a

$$\sum_{i=1}^{n} x_i\,[\ln x_i - \ln y_i] \geq \sum_{i=1}^{n} [x_i - y_1]$$

Supposons alors que $\sum_i x_i = \sum_i y_i$. Cette égalité est notamment réalisée si les nombres x_k et y_k sont des probabilités discrètes, la somme des premiers comme la somme des seconds étant égales à 1. On aboutit alors à l'inégalité de Gibbs [1] :

$$-\sum_{i=1}^{n} x_i\,\ln x_i \leq -\sum_{i=1}^{n} x_i\,\ln y_i \tag{3.3}$$

1. La grandeur $S = -\sum_{i=1}^{n} x_i\,\ln x_i$ représente alors l'Entropie, voir sec. 23.

qui a sa version "intégrale"

$$-\int_\Gamma f\ln f\,d\Gamma \le -\int_\Gamma f\ln g\,d\Gamma \quad \text{si} \quad \int_\Gamma f\,d\Gamma = \int_\Gamma g\,d\Gamma \tag{3.4}$$

4 $\boxed{\dfrac{\lambda^p a^p}{p} + \dfrac{\lambda^{-q}b^q}{q} \ge ab,\ a,b>0,\ \dfrac{1}{p}+\dfrac{1}{q}=1}$ (4)

On définit $f(\lambda) = \dfrac{\lambda^p a^p}{p} + \dfrac{\lambda^{-q}b^q}{q}$. On a $f'(\lambda) = \lambda^{p-1}a^p - \lambda^{-q-1}b^q = \dfrac{1}{\lambda^{q+1}}\left[\lambda^{p+q}a^p - b^q\right]$. Le minimum est atteint pour $\lambda = \lambda_0 = \left(\dfrac{b^q}{a^p}\right)^{1/(p+q)} = \left(\dfrac{b^q}{a^p}\right)^{1/(pq)}$; comme $\lambda_0^p a^p = \lambda_0^{-q}b^q$ et $1/p+1/q=1$, on a $f(\lambda_0) = \lambda_0^p a^p = ba^{p-p/q} = ab$. L'inégalité est bien vérifiée, l'égalité n'ayant lieu que si et seulement si $\lambda = \lambda_0$.

5 $\boxed{\dfrac{x^p}{p} + \dfrac{x^{-q}}{q} \ge 1,\ x>0,\ \dfrac{1}{p}+\dfrac{1}{q}=1}$ (5)

Dans l'inégalité (4) on pose $a = a'/\lambda$, $b = \lambda b'$, d'où

$$\frac{a'^p}{p} + \frac{b'^q}{q} \ge a'b' \tag{5.1}$$

l'égalité n'étant réalisée que si et seulement si $\lambda = \lambda_0 = b^{1/p}/a^{1/q} = \lambda b'^{1/p}/a'^{1/q}$, soit $a'^{1/q} = b'^{1/p}$. C'est l'inégalité de Young. Divisons-la par $a'b'$ et posons $x = a'^{1/q}/b'^{1/p}$. On obtient

$$\frac{a'^{p-1}}{b'p} + \frac{b'^{q-1}}{a'q} = \frac{a'^{p/q}}{b'p} + \frac{b'^{q/p}}{a'q} = \frac{x^p}{p} + \frac{x^{-q}}{q} \ge 1$$

l'inégalité n'étant réalisée que si et seulement si $x=1$. On peut bien sûr partir plus directement de $f(x) = \dfrac{x^p}{p} + \dfrac{x^{-q}}{q}$ et en chercher le minimum.

6 Inégalité de Hölder

$$\boxed{\left(\sum_i a_i^p\right)^{1/p}\left(\sum_i b_i^q\right)^{1/q} \geq \sum_i a_i\, b_i,\ \ a_i, b_i \geq 0,\ \ \frac{1}{p}+\frac{1}{q}=1} \quad \textbf{(6)}$$

D'après l'inégalité (4), $\dfrac{\lambda^p a_i^p}{p}+\dfrac{\lambda^{-q} b_i^q}{q} \geq a_i b_i$ et l'égalité est obtenue si et seulement s'il existe une valeur de λ telle que $b_i^q = \lambda^{p+q}\, a_i^p$, d'où l'on déduit

$$X = \frac{\lambda^p A}{p}+\frac{\lambda^{-q}B}{q} \geq C, \quad \text{où}$$

$$A = \sum_i a_i^p, \quad B = \sum_i b_i^q, \quad C = \sum_i a_i\, b_i$$

L'égalité étant réalisée si et seulement si $b_i^q = \lambda^{p+q}\, a_i^p$ pour tout i, c'est-à-dire, si et seulement si les deux suites $\{a_i^p\}$ et $\{b_i^q\}$ sont *proportionnelles*. En outre, A, B et C étant donnés, X est minimum lorsque $\lambda^{p+q} = B/A$, et a pour valeur $X_{\min} = A^{1/p}B^{1/q}$. Cette valeur étant supérieure ou égale à C, on en déduit l'inégalité de Hölder :

$$\left(\sum_i a_i^p\right)^{1/p}\left(\sum_i b_i^q\right)^{1/q} \geq \sum_i a_i\, b_i \tag{6.1}$$

l'égalité étant réalisée si et seulement si les suites $\{a_i^p\}$ et $\{b_i^q\}$ sont proportionnelles, ou bien si l'une des suites est nulle.

De (6) on déduit notamment

$$\left(\sum_i a_i^p\right)^{1/p}\left(\sum_i a_i^{-q}\right)^{1/q} \geq n \tag{6.2}$$

Prenant par exemple $a_k = k$, $p = q = 2$, et compte tenu de $\displaystyle\sum_1^n k^2 = \frac{n(n+1)(2n+1)}{6}$, il vient

$$\sum_1^n \frac{1}{k^2} \geq \frac{6n}{(n+1)(2n+1)} \tag{6.3}$$

7 Inégalité de Cauchy-Schwarz

Dans l'inégalité de Hölder (6.1) on prend $p = q = 2$ pour en déduire la célèbre inégalité de Cauchy-Schwarz :

$$\sqrt{\sum_i a_i^2}\sqrt{\sum_i b_i^2} \geq \sum_i a_i\, b_i \tag{7.1}$$

qui ne devient égalité que si et seulement si $b_k = \lambda a_k$ pour tout k. Remarquons ici que

$$\left(\sum_i a_i^2\right)\left(\sum_j b_j^2\right) - \left(\sum_i a_i\, b_i\right)^2 = \sum_{i<j}(a_i b_j - a_j b_i)^2$$

et que le second membre n'est effectivement nul que si et seulement si les suites sont proportionnelles.

8 Inégalité de Minkowski

$$\boxed{\left(\sum_i a_i^p\right)^{1/p} + \left(\sum_i b_i^p\right)^{1/p} \geq \left[\sum_i (a_i + b_i)^p\right]^{1/p} \quad a_i, b_i \geq 0} \text{ (8)}$$

On écrit

$$\sum_i (a_i + b_i)^p = \sum_i (a_i + b_i)^{p-1} a_i + \sum_i (a_i + b_i)^{p-1} b_i$$

$$\leq \left(\sum_i a_i^p\right)^{1/p}\left(\sum_i (a_i + b_i)^{q(p-1)}\right)^{1/q} + \left(\sum_i b_i^p\right)^{1/p}\left(\sum_i (a_i + b_i)^{q(p-1)}\right)^{1/q}$$

et comme $q(p-1) = p$, $1 - 1/q = p$:

$$\left(\sum_i a_i^p\right)^{1/p} + \left(\sum_i b_i^p\right)^{1/p} \geq \left[\sum_i (a_i + b_i)^p\right]^{1/p} \tag{8.1}$$

l'égalité étant réalisée si et seulement si les suites $(a_i + b_i)^{q(p-1)} = (a_i + b_i)^p$ et a_i^p sont proportionnelles, c'est-à-dire finalement si et seulement si les suites $\{a_i\}$ et $\{b_i\}$ sont proportionnelles, ou bien si l'une des suites est nulle.

9 Inégalité de Hölder généralisée

$$\boxed{\left(\sum_i a_{1i}\right)^{\alpha_1}\cdots\left(\sum_i a_{ni}\right)^{\alpha_n} \geq \sum_i a_{1i}^{\alpha_1}\cdots a_{ni}^{\alpha_n},\ a_{ki}\geq 0,\ \sum_k \alpha_k = 1} \quad (9)$$

A partir de l'inégalité de Hölder "simple" (6), on fait le changement $a_i^p = a_i'$, $\alpha = 1/p$, $\beta = 1/q$ pour obtenir

$$\sum_i a_i'^{\alpha}\, b_i'^{\beta} \leq \left(\sum_i a_i'\right)^{\alpha}\left(\sum_i b_i'\right)^{\beta}, \quad \text{avec} \quad \alpha+\beta = 1$$

l'égalité étant réalisée si et seulement si les deux suites $b_i^q = b_i'$ et $a_i^p = a_i'$ sont proportionnelles. Supprimons les signes "primes" et redéfinissons $b_i^{\beta} = b'^{\beta_1} c_i^{\beta_2}$, avec $\beta_1+\beta_2 = \beta$. Il vient d'une part

$$\sum_i a_i^{\alpha}\, b_i'^{\beta_1} c_i^{\beta_2} \leq \left(\sum_i a_i'\right)^{\alpha}\left(\sum_i b_i'^{\beta_1/\beta} c^{\beta_2/\beta}\right)^{\beta}$$

et comme d'autre part

$$\sum_i b_i'^{\beta_1/\beta} c^{\beta_2/\beta} \leq \left(\sum_i b_i'\right)^{\beta_1/\beta}\left(\sum_i c_i\right)^{\beta_2/\beta}$$

on obtient finalement

$$\sum_i a_i^{\alpha}\, b_i'^{\beta_1} c_i^{\beta_2} \leq \left(\sum_i a_i\right)^{\alpha}\left(\sum_i b_i'\right)^{\beta_1}\left(\sum_i c_i\right)^{\beta_2}, \quad \text{avec} \quad \alpha+\beta_1+\beta_2 = 1$$

En réitérant cette opération, on peut ainsi établir l'inégalité annoncée, qui devient égalité si et seulement les suites sont toutes proportionnelles, ou bien si l'une des suites est nulle.

10 Inégalité de Hölder avec des intégrales

$$\boxed{\int f\, g\, d\Gamma \leq \left(\int f^p d\Gamma\right)^{1/p}\left(\int g^q d\Gamma\right)^{1/q}} \quad (10)$$

Les deux fonctions numériques positives f et g sont définies sur un ensemble E de $\mathbb{R}^n$ et intégrables sur cet ensemble. Pour u et v positifs et $1/p+1/q = 1$, on a (voir 3) $u^p/p + v^q/q \geq uv$. On pose

$$u = f/\left(\int f^p d\Gamma\right)^{1/p}, \quad v = g/\left(\int g^q d\Gamma\right)^{1/q}$$, d'où

$$\frac{1}{p}\frac{f^p}{\int f^p d\Gamma} + \frac{1}{q}\frac{g^q}{\int g^q d\Gamma} \geq \frac{fg}{\left(\int f^p d\Gamma\right)^{1/p}\left(\int g^q d\Gamma\right)^{1/q}}$$

et l'inégalité annoncée est obtenue par intégration, l'égalité étant réalisée si et seulement si les deux fonctions sont proportionnelles ou si l'une d'elles est nulle.

11 Inégalité de Cauchy-Schwarz avec des intégrales

$$\boxed{\left|\int f\,g\,d\Gamma\right| \leq \int |f|\,|g|\,d\Gamma \leq \sqrt{\int |f|^2\,d\Gamma \int |g|^2\,d\Gamma}} \tag{11}$$

12 $$\boxed{\frac{1}{p}\int f^p d\Gamma + \frac{1}{q}\int g^q d\Gamma \geq \int f\,g\,d\Gamma} \quad \mathbf{(12)}$$

Prendre $u = f$, $v = g$ et intégrer.

13 $$\boxed{\left(\int (f+g)^p\,d\Gamma\right)^{1/p} \leq \left(\int f^p d\Gamma\right)^{1/p} + \left(\int g^p d\Gamma\right)^{1/p}} \quad \mathbf{(13)}$$

On procède comme en (8).

14 Inégalité de réarrangement

Considérons deux suites $\{x_k\}$ et $\{y_k\}$ de n réels chacune et telles que $x_1 < x_2 < \cdots < x_n$ et $y_1 < y_2 < \cdots < y_n$. Pour toute permutation σ de $[1, \cdots, n]$, on a

$$x_1 y_n + x_2 y_{n-1} + \cdots x_n y_1 \leq x_1\,y_{\sigma(1)} + \cdots x_n\,y_{\sigma(n)} \leq x_1 y_1 + \cdots x_n\,y_n \tag{14}$$

c'est-à-dire, la somme des produits $x_k\,y_k$ est maximum lorsque les deux suites sont dans le même ordre et minimum lorsque les ordres sont complètement inverses.

Pour démontrer l'inégalité de droite, il suffit de montrer qu'une permutation de deux indices sur la suite des y diminue cette valeur maximum. Supposons $y_j \geq y_i$ et $x_j \geq x_i$. On a

$$(x_j - x_i)(y_j - y_i) = x_j y_j + x_i y_i - x_i y_j - x_j y_i \geq 0, \quad \text{soit}$$
$$x_j y_j + x_i y_i \geq x_i y_j + x_j y_i$$

l'égalité n'ayant lieu que si $x_j = x_i$ ou $y_j = y_i$. Pour démontrer l'inégalité de gauche, il suffit d'appliquer le résultat précédent à la suite $-x_n < \cdots < -x_1$.

15 Inégalité de Tchebychev

Etant donné deux suites réelles décroissantes $x_1 \geq x_2 \geq \cdots \geq x_n$ et $y_1 \geq y_2 \geq \cdots \geq y_n$, on a

$$\frac{1}{n}\sum_{k=1}^{n} x_k\, y_{n-k+1} \leq \left(\frac{1}{n}\sum_{k=1}^{n} x_k\right)\left(\frac{1}{n}\sum_{\ell=1}^{n} y_\ell\right) \leq \frac{1}{n}\sum_{k=1}^{n} x_k y_k \tag{15}$$

De l'inégalité $(x_k - x_\ell)(y_k - y_\ell) \geq 0$, valable pour tout k et tout ℓ puisque les deux suites sont conjointement décroissantes, on tire, en sommant d'abord sur k,

$$\sum_k x_k y_k - y_\ell \sum_k x_k - x_\ell \sum_k y_k + n x_\ell\, y_\ell \geq 0$$

puis, en sommant sur ℓ,

$$2n\sum_k x_k\, y_k - 2\sum_k x_k \sum_\ell y_\ell \geq 0$$

d'où l'inégalité de droite annoncée. Lorsque les deux suites ont des sens de monotonie opposés, le sens des inégalités change, d'où l'inégalité de gauche annoncée.

16 Inégalité de Schur

Etant donné trois nombres réels positifs x, y, z et un réel strictement positif r, on a

$$x^r(x-y)(x-z) + y^r(y-z)(y-x) + z^r(z-x)(z-y) \geq 0 \tag{16}$$

l'égalité n'étant réalisée que si et seulement si $x = y = z$ ou si deux des nombres parmi (x, y, z) sont égaux et le troisième nul.

L'inégalité étant complètement symétrique vis-à-vis des trois nombres (x, y, z), on peut supposer que $x \geq y \geq z$. On récrit alors le membre de gauche de l'ínégalité comme

$$(x-y)\left[x^r(x-z)-y^r(y-z)\right]+z^r(x-z)(y-z)$$

et cette expression est manifestement positive ; d'où l'inégalité annoncée.

17 Inégalité de Muirhead [2]

Il s'agit d'une généralisation de l'inégalité arithmético-géométrique. Etant donnés n réels strictement positifs $\{x_k\}$ et n autres réels quelconques $p_1, \cdots, p_n$, on appelle la p-moyenne des x_k la grandeur

$$[p]_x = \frac{1}{n!}\sum_{\sigma} x_{\sigma(1)}^{p_1}\, x_{\sigma(2)}^{p_2} \cdots x_{\sigma(n)}^{p_n}$$

σ étant une permutation quelconque de $(1, \cdots, n)$. Les moyennes arithmétiques et géométriques sont des cas particuliers :

$$[(1, 0, \cdots, 0)]_x = \frac{1}{n!}\sum_{\sigma} x_{\sigma(1)} \times 1 \times \cdots \times 1 = \frac{1}{n}\sum_{k} x_k$$

$$[(\frac{1}{n}, \cdots, \frac{1}{n})]_x = \frac{1}{n!}\sum_{\sigma} x_{\sigma(1)}^{1/n}\, x_{\sigma(2)}^{1/n} \cdots x_{\sigma(n)}^{1/n} = (x_1 \cdots x_n)^{1/n}$$

Etant donné deux suites $p = (p_1, \cdots, p_n)$ et $q = (q_1, \cdots, q_n)$ *décroissantes* : $p_1 \geq p_2 \cdots \geq p_n$ et $q_1 \geq q_2 \geq \cdots \geq q_n$. Si elles satisfont

$$p_1 + \cdots + p_k \leq q_1 + \cdots + q_k$$

avec égalité pour $k = n$, on dit que la suite q *majore* la suite p.

L'inégalité de Muirhead s'énonce ainsi : si les suites p et q sont décroissantes, on a $[p]_x \leq [q]_x$ pour toute suite $\{x_k\}$ si et seulement si q majore p.

2. Voir : G. Hardy, J.E. Littlewood, G. Pólya, "Inequalites", Cambridge University Press (digital printing 2001).

18 Inégalité de Bienaymé-Tchebychev

Soit X une variable aléatoire de densité de probabilité $f(x)$, possédant des moments d'ordre 1 et 2. La probabilité pour que l'on ait $|X| > a$ est

$$P(|X| > a) = \int_{|X|>a} f(x)dx$$

Dans le domaine où $|x| > a$, on a $x^2 > a^2$, et

$$P(|X| > a) \leq \int_{|X|>a} \frac{x^2}{a^2} f(x)dx \leq \frac{1}{a^2} \int_{-\infty}^{+\infty} x^2 f(x)dx, \quad \text{soit}$$

$$P(|X| > a) \leq \frac{\mathbb{E}(X^2)}{a^2} \quad \text{et aussi}$$

$$P(|X - \mathbb{E}(X)| > a) \leq \frac{\sigma^2}{a^2} \quad \text{où} \quad \sigma^2 = \int_{-\infty}^{+\infty} (x - \mathbb{E}(X))^2 f(x)dx \tag{18}$$

19 Fonctions convexes, fonctions concaves

On rappelle qu'une fonction numérique Ψ définie sur un espace vectoriel E est *convexe* si, quels que soient les éléments x et y de E et pour tout réel α compris entre 0 et 1, on a

$$\Psi(\alpha x + (1-\alpha)y) \leq \alpha\Psi(x) + (1-\alpha)\Psi(y) \tag{19.1}$$

La fonction Ψ est dite *concave* si $-\Psi$ est convexe (l'inégalité est inversée).

L'ensemble E lui-même est dit convexe si $z = \alpha x + (1-\alpha)y$ est encvore un élément de E.

19.1 Inégalité de Jensen

$$\boxed{\Psi(\sum_{i=1}^{m}\alpha_i x_i) \leq \sum_{i=1}^{m}\alpha_i\Psi(x_i),\ \sum_{i=1}^{m}\alpha_i = 1} \tag{19.2}$$

L'inégalité est vraie pour $m = 1$ et vraie aussi pour $m = 2$ car, Ψ étant convexe, on a par définition $\Psi(\alpha_1 x_1 + \alpha_2 x_2) \leq \alpha_1\Psi(x_1) + \alpha_2\Psi(x_2)$, avec $0 \leq \alpha_i \leq 1$ et $\alpha_1 + \alpha_2 = 1$. Posons $z = \sum_{i=1}^{m}\alpha_i x_i = \sum_{k=1}^{m-1}\alpha_k x_k + \alpha_m x_m$, puis $\alpha = \sum_{k=1}^{m-1}\alpha_k = 1 - \alpha_m$, et $\beta_k = \alpha_k/\alpha$ pour $1 \leq k \leq m-1$, de sorte que $\sum_{k=1}^{m-1}\beta_k = 1$, puis enfin $u = \sum_{k=1}^{m-1}\beta_k x_k$. On a $z = \alpha u + (1-\alpha)x_m$ et donc $\Psi(z) \leq \alpha\Psi(u) + (1-\alpha)\Psi(x_m)$. Supposons maintenant l'ínégalité vraie pour $m-1$. Alors $\Psi(u) \leq \sum_{k=1}^{m-1}\beta_k\Psi(x_k)$, et l'inégalité plus haut donne $\Psi(z) \leq \sum_{k=1}^{m-1}\alpha_k\Psi(x_k) + \alpha_m\Psi(x_m)$, c'est-à-dire, l'inégalité annoncée.

En prenant $m = n$, $\alpha_i = 1/n$, l'inégalité (19.2) donne

$$\Psi(\frac{1}{n}\sum_{i=1}^{n} x_i) \leq \frac{1}{n}\sum_{i=1}^{n}\Psi(x_i) \tag{19.3}$$

19.2 Fonction convexe sur l'ensemble des réels

On considère trois réels x, y et z vérifiant $x < y$, $x \leq z \leq y$ et $\alpha = \dfrac{y-z}{y-x}$. On a $0 \leq \alpha \leq 1$, $1 - \alpha = \dfrac{z-x}{y-x}$, et $z = \alpha x + (1-\alpha)y$. On remplace dans (19.1) pour obtenir

$$\frac{\Psi(x) - \Psi(z)}{x-z} \leq \frac{\Psi(y) - \Psi(z)}{y-z} \tag{19.4}$$

Posons $P(t) = \dfrac{\Psi(t) - \Psi(z)}{t-z}$. L'inégalité (19.4) se récrit comme $P(x) < P(y)$, et montre que lorsque $y \to z+0$, la borne inférieure de $P(y)$ est finie et que lorsque $x \to z-0$, la borne supérieure de $P(x)$ est finie. On en conclut qu'une fonction convexe définie sur les réels possède en tout point z une dérivée à gauche $\Psi'_g(z)$ et une dérivée à droite $\Psi'_d(z)$ et que pour $x < y$ on a

$$P(x) \leq \Psi'_g(z) \leq \Psi'_d(z) \leq P(y) \tag{19.5}$$

Supposons Ψ dérivable en tout point. En faisant tendre z vers x, la même inégalité (19.4) donne $\Psi'(x) \leq \dfrac{\Psi(y) - \Psi(x)}{y-x}$ tandis que si l'on fait tendre z vers y, elle donne $\dfrac{\Psi(y) - \Psi(x)}{y-x} \leq \Psi'(y)$. On en déduit que

$$\Psi'(x) \leq \Psi'(y) \quad \text{pour} \quad x \leq y \tag{19.6}$$

c'est-à-dire, la dérivée de Ψ est nécessairement croissante. Si cette dérivée est elle-même dérivable, la dérivée seconde de Ψ est donc nécessairement positive. On montre que la positivité de la dérivée seconde d'une fonction deux fois dérivable est une condition nécessaire et suffisante de sa convexité.

Les inégalités (19.5) apportent aussi les conclusions suivantes. Pour $y \geq z$ on a $\Psi(y) - \Psi(z) \geq \Psi'_g(z)(y-z)$, tandis que pour $x \leq z$ on a $\Psi(x) - \Psi(z) \geq (x-z)\Psi'_d(z)$. Par conséquent, quel que soit les réels y et z ($y \leq z$ ou $y \geq z$), il est toujours possible de trouver un nombre M tel que

$$\Psi(y) - \Psi(z) \geq (y-z)M \tag{19.7}$$

19.3 Le logarithme

La fonction $\ln x$ pour $x > 0$ est deux fois dérivable et sa dérivée seconde est négative. Elle est donc *concave*. On en déduit par exemple

$$\ln(\sum_{i=1}^{m}\alpha_i x_i) \geq \sum_{i=1}^{m}\alpha_i \ln(x_i) \quad \text{et} \quad \sum_{i=1}^{m}\alpha_i x_i \geq \prod_i x_i^{\alpha_i} \tag{19.8}$$

les réels x_i étant supposés positifs et $\sum_{i=1}^{m}\alpha_i = 1$. Une application immédiate est la comparaison entre la moyenne *géométrique* $x_g = (x_1 \cdots x_n)^{1/n}$ et la moyenne *arithmétique* $x_a = (x_1 + \cdots + x_n)/n$ de n nombres positifs : en prenant $\alpha_i = 1/n$, l'inégalité précédente donne $\ln x_a \geq \ln x_g$, soit

$$\boxed{\sqrt[n]{x_1 \cdots x_n} \leq \frac{1}{n}\left[x_1 + \cdots + x_n\right]} \tag{19.9}$$

En remplaçant dans (19.9) les nombres par leurs inverses, on obtient une inégalité entre la moyenne géométrique de n nombres x_k et leur moyenne dite *harmonique*

$$x_h = \frac{n}{\dfrac{1}{x_1} + \cdots + \dfrac{1}{x_n}} \quad : \quad x_g \geq x_h \tag{19.10}$$

Les moyennes ci-dessus s'ordonnent selon

$$x_h \leq x_g \leq x_a \tag{19.11}$$

les égalités se réalisant si et seulement si les nombres impliqués sont tous égaux. Le cas dégalité $x_a = x_h$ peut être retrouvé comme suit. Le produit

$$P = \left[x_1 + \cdots + x_n\right]\left[\frac{1}{x_1} + \cdots + \frac{1}{x_n}\right] = n + \sum_{i<j}\left[\frac{x_i}{x_j} + \frac{x_j}{x_i}\right]$$

comporte $n(n-1)/2$ termes de la forme $x/y + y/x$, qui prennent leur seule valeur *minimum*, égale à 2, pour $x = y$. Donc $P \geq n^2$, et $P = n^2$ (soit $x_a = x_h$) lorsque tous les x_i sont égaux.

On peut aussi utiliser la concavité du logarithme pour retrouver l'inégalité de Young (5.1). Définissons $\alpha_1 = 1/p$, $\alpha_2 = 1/q$ avec $1/p + 1/q = 1$, $x_1 = a^p$, $x_2 = b^q$ et appliquons (19.8) :

$$\ln(a^p/p + b^q/q) \geq \frac{1}{p}\ln a^p + \frac{1}{q}\ln b^q = \ln(ab), \quad \text{soit} \quad a^p/p + b^q/q \geq ab$$

19.4 La fonction puissance

La fonction $f(x) = x^p$ avec $x \geq 0$, $p \geq 2$ a pour dérivée seconde $f''(x) = p(p-1)x^{p-2}$. Celle-ci est positive et la fonction puissance est donc convexe. On en déduit

$$\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right)^p \leq \frac{1}{n}\sum_{i=1}^{n} x_i^p, \quad \text{soit}$$

$$\frac{1}{n}\sum_{i=1}^{n} x_i \leq \left[\frac{1}{n}\sum_{i=1}^{n} x_i^p\right]^{1/p} \tag{19.12}$$

On peut alors comparer la moyenne arithmétique d̀'autres types de moyennes. Ainsi, en prenant $p = 2$, on trouve que la moyenne *quadratique* de n nombres positifs est toujours plus grande que leur moyenne arithmetique :

$$x_q = \left[\frac{1}{n}\sum_{i=1}^{n} x_i^2\right]^{1/2} \geq x_a \tag{19.13}$$

19.5 Autre application

Considérons la fonction $F(x) = \ln[1 + e^x]$. Sa dérivée seconde $F''(x) = \dfrac{e^x}{(1+e^x)^2}$ est toujours positive et F est donc convexe sur l'ensemble des réels. En lui appliquant (19.3), on trouve

$$\ln\left[1 + \exp(\frac{x_1 + \cdots + x_n}{n})\right] \leq \frac{1}{n}\sum_{k=1}^{n} \ln(1 + e^{x_k}) \tag{19.14}$$

ou, en posant $y_k = e^{x_k}$,

$$\boxed{\left[\prod_k (1 + y_k)\right]^{1/n} \geq 1 + (y_1 \cdots y_n)^{1/n}} \tag{19.15}$$

Cette dernière inégalité peut être exploitée comme suit. Les nombres y_k, positifs mais quelconques, peuvent être envisagés comme les n valeurs propres d'une matrice A *définie positive* $n \times n$, ce qui donne

$$[\det(1 + A)]^{1/n} \leq 1 + (\det A)^{1/n} \tag{19.16}$$

Et, pour aller plus loin, on peut exprimer A sous la forme $A = MN^{-1}$ où M et N sont deux matrices $n \times n$, définies positives. On en déduit aisément l'inégalité

$$\boxed{[\det(M+N)]^{1/n} \geq (\det M)^{1/n} + (\det N)^{1/n}} \tag{19.17}$$

qui ne devient égalité que si et seulement si les deux matrices sont proportionnelles. Enfin, si dans (19.17) on fait le remplacement $M = \alpha B$, $N = (1-\alpha)C$, avec $0 \leq \alpha \leq 1$, on obtient

$$[\det(\alpha B + (1-\alpha)C)]^{1/n} \geq \alpha (\det B)^{1/n} + (1-\alpha)(\det C)^{1/n} \tag{19.18}$$

ce qui montre que l'application $\Phi(M) = (\det M)^{1/n}$ est concave.

19.6 Inégalité de Jensen pour les intégrales

Soit f une fonction numérique définie sur un ensemble D et intégrable sur cet ensemble, selon une mesure $\mu(u)$ telle que $\int_D d\mu(u) = 1$. Soit Ψ une fonction convexe sur les réels. Appliquons l'inégalité (19.7) en choisissant $z = \int_D f(u)d\mu(u)$ et $y = f(u)$. Il vient

$$M\left[f(u) - \int_D f(u)d\mu(u)\right] \leq \Psi(f(u)) - \Psi(\int_D f(u)d\mu(u))$$

En intégrant, le membre de gauche disparaît (compte tenu de la normalisation de la mesure) et l'on obtient

$$\boxed{\Psi(\int_D f(u)d\mu(u)) \leq \int_D \Psi(f(u))d\mu(u)} \tag{19.19}$$

20 Inégalité de Prékopa-Leindler [3]

Etant donné trois fonctions f, g, h définies et positives dur $\mathbb{R}$ et un nombre réel α compris entre 0 et 1, si pour tout x et tout y dans $\mathbb{R}$ on a $h(\alpha x + (1-\alpha)y) \geq f^\alpha(x) g^{1-\alpha}(y)$, alors

$$\boxed{\int_{\mathbb{R}^n} h\, d\Gamma \geq \left(\int_{\mathbb{R}^n} f\, d\Gamma \right)^\alpha \left(\int_{\mathbb{R}^n} g\, d\Gamma \right)^{1-\alpha}} \tag{20.1}$$

La démonstration procède par récurrence.

20.1 Démonstration pour $n = 1$

Posons $I_f = \displaystyle\int_{\mathbb{R}} f(x)\, dx$, $I_g = \displaystyle\int_{\mathbb{R}} g(y)\, dy$, $J = I_f^\alpha\, I_g^{1-\alpha}$, et $u(x) = f(x)/I_f$, $v(y) = g(y)/I_g$, $w(z) = h(z)/J$.

Définissons aussi $x(t)$ et $y(t)$ par

$$\int_{-\infty}^{x(t)} u(s)\, ds = t, \qquad \int_{-\infty}^{y(t)} v(s)\, ds = t$$

de telle sorte que $X(t)$ et $y(t)$ tendent vers $-\infty$ pour $t \to 0$, et tendent vers $+\infty$ lorsque $t \to 1$. On a $x'(t)u(x(t)) = 1$ et $y'(t)v(y(t)) = 1$, et par conséquent u et v étant strictement positifs, $x'(t)$ et $y'(t)$ sont strictement positifs. $x(t)$ et $y(t)$ sont strictement croissantes et il y a une correspondance biunivoque entre t et $x(t)$ d'une part, t et $y(t)$ d'autre part. Définissons aussi $z(t) = \alpha x(t) + (1-\alpha)y(t)$. On a $z'(t) = \alpha x'(t) + (1-\alpha)y'(t)$ et d'après (19.8)

$$z'(t) \geq {x'}^\alpha(t)\, {y'}^{1-\alpha}(t)$$

De façon évidente, on a $\displaystyle\int_{\mathbb{R}} w(s)\, ds \geq \int_0^1 w(z(t))\, z'(t)\, dt$, et comme par hypothèse, $w(z) \geq f^\alpha(x) g^{1-\alpha}(y)/J = u^\alpha(x)\, v^{1-\alpha}(y)$, il vient

$$\int_{\mathbb{R}} w(s)\, ds \geq \int_0^1 dt\, \left[\, x'(t)\, u(x(t))\,\right]^\alpha \left[\, y'(t)\, v(y(t))\,\right]^{1-\alpha} = 1$$

puisque chacune des grandeurs entre crochet vaut 1. on en déduit

3. L. Leindler, *On a certain converse Hölder inequality*, Acta Sci. Math. **33** (1972), 217-223. ; A. Prékopa, *On logarithmic concave mesasures and functions*, Acta Sci. Math. **34** 1973, 335-343.

$$\boxed{\int_{\mathbb{R}} h\,ds \geq \left(\int_{\mathbb{R}} f\,dx\right)^{\alpha} \left(\int_{\mathbb{R}} g\,dy\right)^{1-\alpha}} \tag{20.2}$$

La propriété annoncée est donc vraie pour $n = 1$.

20.2 Démonstration pour $\mathbb{R}^n$

Supposons que (20.1) soit vraie pour $\mathbb{R}^n$, et prenons pour hypothèse la relation

$$h(\alpha r + (1-\alpha)s, \alpha x + (1-\alpha)y) \geq f(r,x)^{\alpha}\, g^{1-\alpha}(s,y)$$

où (r,x) et (s,y) sont des éléments de $\mathbb{R}^{n+1}$.

Posons $H(\alpha r + (1-\alpha)s) = \int_{\mathbb{R}^n} h\,d\Gamma_n$, $F(r) = \int_{\mathbb{R}^n} f\,d\Gamma_n$, $G(s) = \int_{\mathbb{R}^n} g\,d\Gamma_n$.

En conséquence de l'hypothèse, on a $H(\alpha r + (1-\alpha)s) \geq F^{\alpha}(r)\,G^{1-\alpha}(s)$, et donc, d'après ce qui précède,

$$\int_{\mathbb{R}} H\,ds \geq \left(\int_{\mathbb{R}} F\,dx\right)^{\alpha} \left(\int_{\mathbb{R}} G\,dy\right)^{1-\alpha}$$

Comme

$$\int_{\mathbb{R}} \{H,F,G\}\,ds = \int_{\mathbb{R}^{n+1}} \{h,f,g\}\,d\Gamma_{n+1}$$

on en déduit que la propriété est aussi vraie pour $n+1$. Elle est donc vraie pour tout n.

20.3 Application : inégalité de Brunn-Minkowski-Lusternik [4]

En posant dans (20.1), $f = \mathrm{id}_M$, $g = \mathrm{id}_N$, $h = \mathrm{id}_{\alpha M + (1-\alpha)N}$, on trouve l'inégalité

$$\boxed{\mathrm{Vol}(\alpha M + (1-\alpha)N) \geq [\,\mathrm{Vol}M\,]^{\alpha}\;[\,\mathrm{Vol}N\,]^{1-\alpha}} \tag{20.3}$$

où Vol signifie "volume". L'inégalité de Brunn-Minkowski proprement dite est obtenue en posant

4. L. Lusternik, *Die Brunn-Minkowskische Ungleichung für beliebige messbare Mengen*, Doklady Akad, SSSR, 1935, No 3, 55-58.

$$M = A/\,[\,\mathrm{Vol}A\,]^{1/n}\;,\; N = A/\,[\,\mathrm{Vol}B\,]^{1/n}\;,\; \alpha = [\,\mathrm{Vol}A\,]^{1/n}\,/\,\left[[\,\mathrm{Vol}A\,]^{1/n} + [\,\mathrm{Vol}B\,]^{1/n}\right]$$

ce qui donne

$$\boxed{[\,\mathrm{Vol}(A+B)\,]^{1/n} \geq [\,\mathrm{Vol}A\,]^{1/n} + [\,\mathrm{Vol}B\,]^{1/n}} \tag{20.4}$$

Précisons ici que l'ensemble $A+B$ est constitué d'éléments de la forme $x+y$ où x est un élément de A et y est un élément de B (somme de Minkowski).

Il est intéressant de rapprocher (20.3) de la deuxième inégalité de (19.8) appliquée à deux variables I_1 et I_2 positives. On peut envisager l'une et l'autre comme les intervalles de variation de deux variables x et y sur la même droite réelle et l'on a, avec $0 \leq \alpha \leq 1$,

$$\alpha I_1 + (1-\alpha) I_2 \geq I_1^{\alpha}\, I_2^{1-\alpha}$$

conformément à (20.3), puisque I_1 et I_2 représentent bien des “volumes” pour la dimension 1.

21 Inégalité de Wirtinger

On considère une fonction f de la variable réelle t, à valeurs possiblement complexes, définie et dérivable dans l'intervalle $0 \leq t \leq 1$, et ayant les propriétés suivantes : $f(0) = f(1)$, $\displaystyle\int_0^1 f(t)dt = 0$.

On prolonge la fonction $f(t)$ sur l'ensemble des réels en une fonction périodique $F(t)$ telle que $F(t) = f(t)$ pour $0 \leq t \leq 1$ et $F(t+n) = F(t)$. La fonction F est continue et dérivable par morceaux sur $\mathbb{R}$; F ainsi que sa dérivée F' sont développables en série de Fourier ayant pour coefficients respectifs

$$f_n = \int_0^1 (F \text{ ou } f)(t) e^{-2i\pi nt}\, dt \quad \text{ et } \quad d_n = \int_0^1 (F' \text{ ou } f')(t) e^{-2i\pi nt}\, dt$$

On a, en intégrant par parties,

$$d_n = \left[f(t)e^{-2i\pi nt}\right]_0^1 + 2i\pi n f_n = 2i\pi n f_n$$

puisque $f(0) = f(1)$. Appliquant le théorème de Parseval, on trouve :

$$\int_0^1 |f'(t)|^2\, dt = \sum_n |d_n|^2 = 4\pi^2 \sum_n n^2 |f_n|^2$$

$$\geq\ 4\pi^2 \sum_n |f_n|^2 = 4\pi^2 \int_0^1 |f(t)|^2\, dt \tag{21.1}$$

en tenant compte de $f_0 = \int_0^1 f(t)dt = 0$. C'est l'inégalité de Wirtinger.

22 Inégalité isopérimétrique

Le *problème isopérimétrique* du plan consiste à trouver, parmi toutes les courbes fermées, continues et rectifiables [5] et dont on se donne le périmètre, celle qui enferme la plus grande aire.

En préalable, rappelons le théorème de Stokes :

$$\oint_\Gamma \overrightarrow{A} \cdot \overrightarrow{d\ell} = \iint_S \overrightarrow{\mathrm{rot}}\, \overrightarrow{A} \cdot \overrightarrow{dS} \tag{22.1}$$

lequel, appliqué au plan, devient la formule de Green-Riemann :

$$\oint_\Gamma A_x dx + A_y dy = \iint_S \left[\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y}\right] dxdy \tag{22.2}$$

En prenant $A_x = -y/2$, $A_y = x/2$, on obtient la relation

$$A = \iint_S dxdy = \frac{1}{2} \oint_\Gamma [-ydx + xdy] \tag{22.3}$$

qui permet d'exprimer l'aire A délimitée par la courbe fermée Γ comme une intégrale d'une forme différentielle le long de Γ.

Notons L le périmètre de Γ. Il s'exprime comme

$$L = \oint ds$$

où s est lélément de longueur sur Γ, que l'on prend pour abcisse curviligne du point M courant sur Γ, à partir d'un point de réf'erence M_0 sur Γ. Les coordonnées x et y de M sont fonctions de s et l'on a $ds^2 = (dx)^2 + (dy)^2$. Pour repérer la position de M sur Γ, on peut plutôt utiliser le paramètre $t = s/L$, On notera

5. C'est-à-dire, dont peut calculer la longueur, au sens usuel du terme.

$x'(t) = \dfrac{dx}{dt} = L\dfrac{dx}{ds}$, $y'(t) = \dfrac{dy}{dt} = L\dfrac{dy}{ds}$, et l'on a $x'^2 + y'^2 = L^2$. Considérons ensuite la fonction $F(x,y) = x + iy$. En changeant l'origine des coordonnées, on peut toujours faire en sorte que

$$\int_0^1 F(x(t),y(t))dt = 0$$

En outre, la courbe Γ étant fermée, on a certainement $F(x(0),y(0)) = F(x(1),y(1))$. Pour simplifier, posons $f(t) = F(x(t),y(t))$. On a

$$\int_0^1 f^\star(t)f'(t)dt = \int_0^1 (x-iy)(x'+iy')dt = \int_0^1 [xx'+yy']\,dt$$

$$+i\int_0^1 [xy'-yx']\,dt = \frac{1}{2}\left[x^2+y^2\right]_0^1 + 2iA, \quad \text{soit}$$

$$I = \int_0^1 f^\star(t)f'(t)dt = 2iA$$

puisque $[x^2+y^2]_0^1 = 0$. La fonction $f(t)$ satisfait aux conditions du théorème de Wirtinger. En appliquant alors l'inégalité de Cauchy-Schwarz (11) à l'intégrale I puis l'inégalité de Wirtinger, on trouve, en tenant compte de $|f'| = L$,

$$|I| = 2A \le \int_0^1 |f|\ |f'|\,dt \le \left[\int_0^1 |f|^2 dt \int_0^1 |f'|^2 dt\right]^{1/2} \le \frac{1}{2\pi}\int_0^1 |f'|^2 dt = \frac{L^2}{2\pi}$$

donc

$$\boxed{L^2 \ge 4\pi A} \tag{22.4}$$

C'est *l'inégalité isopérimétrique*. Si l'on a $L^2 = 4\pi A$, toutes les inégalités deviennent des égalités et ceci n'est réalisé que si et seulement si f et f' sont proportionnels. Comme $|f'| = L = \text{constante}$, on a alors $|f| = \sqrt{x^2+y^2} = \text{constante}$: la courbe fermée Γ est donc un cercle. Réciproquement, si Γ est un cercle de rayon R, on a évidemment $L^2 = 4\pi^2R^2 = 4\pi A$.

Autre démonstration [6]

Le problème isopérimétrique consiste à trouver les fonctions $x(t)$ et $y(t)$ qui rendent extremum l'intégrale $A = \dfrac{1}{2}\displaystyle\int_0^1 [xy'-yx']\,dt$ et qui prennent des valeurs données

6. Voir par exemple : J. Bass, "Cours de Mathématiques", Tome II, Chap. XLII, Masson et C$^{\text{ie}}$, Ed. 1964.

pour $t = 0$ et $t = 1$, de telle sorte que l'intégrale $L = \displaystyle\int_0^1 \sqrt{x'^2 + y'^2}dt$ ait une valeur fixée. Pour ce faire, on peut appliquer les équations d'Euler-Lagrange à la fonction

$$G(x, x', y, y') = xy' - yx' + K\sqrt{x'^2 + y'^2}$$

où K est une constante. On obtient ainsi les deux équations

$$\frac{\partial G}{\partial x} - \frac{d}{dt}\left(\frac{\partial G}{\partial x'}\right) = y' - K\left[-y' + \frac{d}{dt}\left(\frac{x'}{\sqrt{x'^2 + y'^2}}\right)\right] = 0$$

$$\frac{\partial G}{\partial y} - \frac{d}{dt}\left(\frac{\partial G}{\partial y'}\right) = -x' - K\left[x' + \frac{d}{dt}\left(\frac{y'}{\sqrt{x'^2 + y'^2}}\right)\right] = 0, \quad \text{soit}$$

$$(1+K)y' = K\frac{d}{dt}\left(\frac{x'}{\sqrt{x'^2 + y'^2}}\right), \quad (1+K)x' = -K\frac{d}{dt}\left(\frac{y'}{\sqrt{x'^2 + y'^2}}\right)$$

L'intégration des deux dernières équations conduit à

$$(1+K)y = K\frac{x'}{\sqrt{x'^2 + y'^2}} + C_2, \quad (1+K)x = -K\frac{y'}{\sqrt{x'^2 + y'^2}} + C_1$$

C_1 et C_2 étant deux constantes. Multipliant la première équation par y', la seconde par x', ajoutant les deux expressions obtenues et intégrant une nouvelle fois, on aboutit à

$$(1+K)(x^2 + y^2) = C_2 y + C_1 x + C$$

c'est-à-dire, à l'équation d'un cercle :

$$(x - x_0)^2 + (y - y_0)^2 = R^2$$

Les constantes x_0, y_0 et R étant ensuite déterminées à l'aide des données du problème.

23 L'Entropie

23.1 Maximum d'Entropie

En Physique, la théorie quantique est la seule acceptable pour décrire les phénomènes relevant du comportement microscopique des systèmes. Dans ce cadre, il apparaît qu'un système donné ne peut généralement occuper que certains micro-états correspondant à une quantification de l'énergie. Soit alors un système pouvant occuper des micro-états repérés par un indice α.

Les théories statistiques quantiques attribuent à chacun de ces états une probabilité a priori de réalisation notée p_α et définissent, à un facteur près, l'entropie du macro-état du système par

$$\boxed{S = -\sum_\alpha p_\alpha \ \ln p_\alpha} \tag{23.1}$$

Cette définition de l'entropie rejoint celle de Shannon dans sa théorie de l'information élaborée vers 1947. Bien entendu, la loi de probabilité p_α dépend crucialement des contraintes extérieures imposées au système. Il est intéressant de noter ici deux choses. D'une part, l'entropie ainsi définie est toujours positive ($p_\alpha \leq 1$). D'autre part, si le système ne peut se trouver que dans un seul micro-état, la probabilité de l'y trouver est bien sûr égale à 1, et son entropie est alors égale à 0. Cette dernière remarque suggère que l'expression adoptée ci-dessus pour l'entropie est bien adaptée pour donner une mesure du désordre qui peut régner à l'échelle microscopique au sein du système : à mesure que le nombre d'états accessibles est important, plus il y a de termes positifs dans l'expression de l'entropie, plus grande est donc celle-ci, et plus grand est alors le *manque d'information* sur l'état microscopique du système.

Supposons le système isolé. Selon le principe général qu'un système a tendance à occuper la totalité des états auxquels il peut accéder, son état d'équilibre macroscopique sera tout naturellement défini comme celui correspondant au *maximum d'entropie*. Soit p_α^e la loi de probabilité correspondante et $p_\alpha = p_\alpha^e + \delta p_\alpha$ une loi de probabilité infiniment voisine ($|\delta p_\alpha| \ll p_\alpha^e$) correspondant à un macro-état du système infiniment voisin de l'équilibre. L'écart, infinitésimal, entre les entropies correspondant à ces deux états est

$$\delta S = S(\{p_\alpha\}) - S(\{p_\alpha^e\})$$

Or, effectuant un développement limité jusqu'au second ordre, on a

$$(p^e_\alpha + \delta p_\alpha)\ln(p^e_\alpha + \delta p_\alpha) \approx p^e_\alpha \ln p^e_\alpha + \delta p_\alpha\,(1 + \ln p^e_\alpha) + \frac{(\delta p_\alpha)^2}{2p^e_\alpha}$$

d'où

$$\delta S \approx -\sum_\alpha \left[\delta p_\alpha\,(1 + \ln p^e_\alpha) + \frac{(\delta p_\alpha)^2}{2p^e_\alpha}\right] \qquad (23.2)$$

Comme on a affaire à un extremum de l'entropie, le premier terme de ce développement, qui est linéaire suivant les variations δp_α, doit être nul, soit

$$\sum_\alpha \delta p_\alpha\,(1 + \ln p^e_\alpha) = 0 \qquad (23.3)$$

Cependant, les variations δp_α ne sont pas toutes indépendantes, puisque, la somme des probabilités étant toujours égale à 1, on a

$$\sum_\alpha p_\alpha \; - \; \sum_\alpha p^e_\alpha \; = \; \sum_\alpha \delta p_\alpha \; = \; 1 - 1 = 0 \qquad (23.4)$$

Pour simplifier, supposons que le nombre M de micro-états accessibles soit fini. La contrainte précédente nous permet de réexprimer par exemple δp_M comme

$$\delta p_M = -\sum_{\alpha=1}^{M-1} \delta p_\alpha \qquad (23.5)$$

et par suite la condition d'extremum prend la forme

$$\sum_{\alpha=1}^{M-1} \delta p_\alpha\,(\ln p^e_\alpha \; - \; \ln p^e_M) \; = \; 0 \qquad (23.6)$$

Dans cette nouvelle relation, les $M - 1$ variations δp_α doivent être considérées comme *indépendantes*. L'égalité n'est donc réalisée que si et seulement si on a

$$\ln p^e_\alpha \; = \; \ln p^e_M \;\; \text{pour} \;\; \alpha = 1, \ldots, M - 1 \qquad (23.7)$$

soit

$$p^e_1 \; = \; p^e_2 \; = \; \cdots \; = \; p^e_M \qquad (23.8)$$

On retrouve ainsi le fait qu'à l'équilibre les micro-états se réalisent tous avec la même probabilité, qui, si $M = \Omega$ est leur nombre, est égale à

$$p^e \;=\; \frac{1}{\Omega} \tag{23.9}$$

L'entropie est alors

$$S^e = -\sum_{\alpha=1}^{\Omega} \frac{1}{\Omega} \ln \frac{1}{\Omega} \;=\; \ln \Omega \tag{23.10}$$

Cette expression correspond à un *maximum* d'entropie. En effet, les termes du second ordre dans la variation d'entropie sont

$$\delta S \approx -\sum_{\alpha} \frac{(\delta p_\alpha)^2}{2p^e_\alpha} = -\frac{\Omega}{2} \sum_{\alpha} \,(\delta p_\alpha)^2 \tag{23.11}$$

et donc tous *négatifs*. S^e représente donc bien une valeur *maximum* et il s'agit d'un maximum absolu.

23.2 Corrélations et Entropie

L'entropie d'un système permet de quantifier l'étendue du désordre qui y règne. Voici un aspect du sujet.

Considérons deux systèmes Σ_1 et Σ_2 pouvant exercer l'un sur l'autre une interaction. Cette interaction peut rendre compte d'une tendance à un certain ordre au sein du système global Σ constitué de Σ_1 et Σ_2, surtout si elle est de nature attractive. Nous supposerons l'ensemble Σ isolé du monde extérieur.

Les deux sous-systèmes peuvent chacun occuper des micro-états que nous repèrerons par l'indice α pour ceux de Σ_1 et par β pour ceux de Σ_2. La probabilité de trouver le système global Σ dans un micro-état où Σ_1 est dans l'état α *et* Σ_2 dans l'état β est notée $p_{\alpha\beta}$. Nous noterons $p^{(1)}_\alpha$ la probabilité de trouver Σ_1 dans le micro-état α, quel que soit l'état de Σ_2. De même, nous noterons $p^{(2)}_\beta$ la probabilité de trouver Σ_2 dans le micro-état β, quel que soit l'état de Σ_1. Faisant appel aux probabilités conditionnelles, la probabilité de trouver Σ_1 dans l'état α sachant que Σ_2 est dans l'état β est

$$p_{\alpha|\beta} = \frac{p_{\alpha\beta}}{p^{(2)}_\beta} \tag{23.12}$$

et l'on a

$$\sum_{\alpha} p_{\alpha|\beta} \;=\; 1 \;, \quad \text{soit} \;\; \sum_{\alpha} p_{\alpha\beta} \;=\; p^{(2)}_\beta \tag{23.13}$$

De même, la probabilité de trouver Σ_2 dans l'état β sachant que Σ_1 est dans l'état α est

$$p_{\beta|\alpha} = \frac{p_{\alpha\beta}}{p_\alpha^{(1)}} \tag{23.14}$$

et l'on a

$$\sum_\beta p_{\beta|\alpha} \;=\; 1\;, \quad \text{soit} \;\; \sum_\beta p_{\alpha\beta} \;=\; p_\alpha^{(1)} \tag{23.15}$$

On sait que si les deux systèmes sont *indépendants* on a alors $p_{\beta|\alpha} = p_\beta^{(2)}$, $p_{\alpha|\beta} = p_\alpha^{(1)}$, et

$$p_{\alpha\beta} = p_\alpha^{(1)} p_\beta^{(2)} \tag{23.16}$$

Le rapport

$$C_{\alpha\beta} = \frac{p_\alpha^{(1)} p_\beta^{(2)}}{p_{\alpha\beta}} \tag{23.17}$$

donne une mesure de la *corrélation* entre les deux systèmes Σ_1 et Σ_2 en interaction. Il vaut 1 s'ils ne sont pas corrélés, auquel cas correspond le plus grand désordre dans Σ.
L'entropie de Σ est

$$S = -\sum_{\alpha,\beta} p_{\alpha\beta}\;\ln p_{\alpha\beta} \tag{23.18}$$

Introduisons les entropies respectives S_1 et S_2 de Σ_1 et Σ_2 comme si ces systèmes étaient indépendants. On a

$$S_1 = -\sum_\alpha p_\alpha^{(1)}\;\ln p_\alpha^{(1)} \;\; \text{et} \;\; S_2 = -\sum_\beta p_\beta^{(2)}\;\ln p_\beta^{(2)} \tag{23.19}$$

Tenant compte des relations

$$\sum_\beta p_{\alpha\beta} \;=\; p_\alpha^{(1)}, \quad \text{et} \quad \sum_\alpha p_{\alpha\beta} \;=\; p_\beta^{(2)}$$

on a

$$S_1 + S_2 - S = -\sum_{\alpha\beta} p_{\alpha\beta}\;\ln\left(\frac{p_\alpha^{(1)} p_\beta^{(2)}}{p_{\alpha\beta}}\right)$$

On connait l'inégalité $x - 1 - \ln x \geq 0$ pour tout $x > 0$, qui ne devient égalité que pour $x = 1$. En prenant $x = \dfrac{p_\alpha^{(1)} p_\beta^{(2)}}{p_{\alpha\beta}}$, on obtient

$$- \sum_{\alpha\beta} p_{\alpha\beta} \; \ln \left(\frac{p_\alpha^{(1)} p_\beta^{(2)}}{p_{\alpha\beta}} \right) \geq \; \sum_{\alpha\beta} p_{\alpha\beta} \; \left[1 - \frac{p_\alpha^{(1)} p_\beta^{(2)}}{p_{\alpha\beta}} \right] = \sum_{\alpha\beta} \; \left[p_{\alpha\beta} \; - \; p_\alpha^{(1)} p_\beta^{(2)} \right]$$

Mais

$$\sum_{\alpha\beta} \left[p_{\alpha\beta} \; - \; p_\alpha^{(1)} p_\beta^{(2)} \right] = \sum_{\alpha\beta} p_{\alpha\beta} \; - \; \sum_{\alpha} p_\alpha^{(1)} \sum_{\beta} p_\beta^{(2)} \; = \; 1 - 1 \times 1 = 0$$

On en déduit l'inégalité

$$\boxed{\boxed{S_1 + S_2 \geq S}} \tag{23.20}$$

l'égalité n'ayant lieu que si et seulement si

$$p_{\alpha\beta} = p_\alpha^{(1)} p_\beta^{(2)} \tag{23.21}$$

pour *tous* les micro-états, c'est-à-dire si les deux sous-systèmes Σ_1 et Σ_2 sont *complè-tement décorrélés*. Une telle situation correspond au désordre le plus complet dans Σ où son entropie prend une valeur maximum, égale à la somme des entropies des sous-systèmes.

Du point de vue de la Physique, l'inégalité précédente peut être considérée sous deux aspects.

Imaginons par exemple un processus où les deux systèmes Σ_1 et Σ_2, initialement isolés séparément, viennent à interagir pour former un système composite [7]. A l'état initial les deux systèmes étaient complètement décorrélés et l'entropie du système global Σ était alors égale à la somme des entropies de Σ_1 et Σ_2 pris séparément. L'interaction des deux systèmes introduit évidemment une corrélation entre eux et à l'état d'équilibre final l'entropie de Σ est devenue plus petite. Ainsi, lorsqu'on passe d'une situation de désordre à une situation plus ordonnée, l'entropie *décroît*. A l'inverse, lorsqu'on passe d'un état ordonné à un état désordonné, on observe une *croissance* d'entropie.

D'un autre côté, considérant le postulat selon lequel l'état macroscopique de Σ usuellement observé est celui pour lequel son entropie est maximum, l'inégalité

7. Un alliage par exemple.

mathématique établie plus haut peut vouloir signifier que, si les circonstances le permettent [8], cet état doit correspondre à une situation où la corrélation statistique entre les deux sous-systèmes en interaction est en fait *minimale*, voire inexistante. Pour se convaincre de la réalité de cette situation, il suffit de considérer l'équilibre thermique de deux corps séparables. Bien qu'au niveau microscopique il existe des échanges de chaleur continuels entre les deux corps, leurs énergies respectives ont des valeurs pratiquement bien définies, et, à l'échelle macroscopique, chacun d'eux semble se comporter comme s'ils étaient statistiquement indépendants. Ceci est d'ailleurs conforté par le fait que si, à la suite de l'équilibre, on vient interposer entre les deux corps des parois adiabatiques, chacun garde la température acquise dans l'équilibre et aucune évolution macroscopique n'est observée, ni pour l'un ni pour l'autre [9] .

24 Inégalités de stabilité - Principe de modération [10].

Dans de nombreux domaines, la stabilité d'un système est étudiée au moyen d'une fonction $F(x, y, \cdots)$ dépendant d'un certain nombre de paramètres $x, y, \cdots$ caractérisant un état donné de ce système. La situation d'équilibre de celui-ci est censée correspondre à un extremum de cette fonction, et nous supposerons ici que l'équilibre trouvé est *stable* s'il correspond à un *minimum* de F.

Pour étudier la forme générale des conditions de stabilité, on imagine des transformations virtuelles à partir d'une situation d'équilibre. Pour simplifier le raisonnement, supposons que la fonction F ne dépend que de deux variables x et y. Soit F_0 sa valeur à l'équilibre considéré, et $\Delta' F = F - F_0$ son écart consécutif à une transformation faisant varier x et y de δx et δy, respectivement. Nous supposerons que cet écart peut être développé en série de Taylor à tous les ordres :

$$\Delta' F = \delta F + \frac{1}{2!}\,\delta^2 F + \frac{1}{3!}\,\delta^3 F + \cdots \tag{24.1}$$

8. Cela dépend bien sûr de la nature des systèmes et de leur interaction.

9. A noter à ce propos l'hypothèse émise par J. W. Gibbs, selon laquelle *deux systèmes mécaniques en équilibre statistique restent en équilibre dans le cas limite d'un couplage nul, même s'ils sont séparés*. Voir à ce sujet : "Physique Statistique", B. Diu, C. Guthmann, D. Lederer, B. Roulet, Hermann ed., 1989, Chap. II, Complément II.A ; "Thermodynamics and Statistical Mechanics", A. Sommerfeld, Academic Press Inc. Pub., 1956, & 36 A ; "Elementary Principles in Statistical Mechanics", J. W. Gibbs, Yale University Press, New Haven, 1902.

10. Consulter : L. Landau et E. Lifchitz, "Physique Statistique", §22, Ed. Mir, Moscou, 1967 ; H. B. Callen, "Thermodynamics and an introduction to Thermostatistics", §8-4, 2nde édition, John Wiley & Sons, 1985 ; P. Papon, J. Leblond, "Thermodynamique des états de la matière", §3.2, 3.4, Hermann, 1990.

où ici δF est une variation virtuelle au premier ordre, $\delta^2 F$ une variation virtuelle au second ordre, etc. Explicitement :

$$\delta F = \frac{\partial F}{\partial x}\ \delta x + \frac{\partial F}{\partial y}\ \delta y \tag{24.2}$$

$$\delta^2 F = \frac{\partial^2 F}{\partial x^2}\ (\delta x)^2 + 2\ \frac{\partial^2 F}{\partial x \partial y}\ \delta x \delta y + \frac{\partial^2 F}{\partial y^2}\ (\delta y)^2 \tag{24.3}$$

Au sens le plus large, un équilibre correspond à un extremum de F, ce qui se traduit par l'équation $\delta F = 0$. On a alors les possibilités suivantes.

• Les conditions $\delta F = 0$ **et** $\Delta' F \geq 0$ sont vérifiées quelles que soient les (petits) écarts δx et δy. On a alors $\delta^n F > 0$ pour $n \geq 2$. L'équilibre est bien stable : à la suite d'écarts suffisamment petits, le système, s'il est libre d'évoluer, reviendra tôt ou tard à son état d'équilibre initial.

• On a bien $\delta F = 0$ **et** $\delta^2 F \geq 0$, mais $\Delta' F \geq 0$ n'est pas vérifié pour certaines perturbations. Cela signifie que certains termes $\delta^n F$ avec $n \geq 2$, sinon tous, sont négatifs et l'emportent sur le terme du second ordre. On a alors affaire à un équilibre de départ *métastable*.

• On a $\delta F = 0$, mais certaines perturbations sont telles que $\delta^2 F < 0$. L'équilibre est instable.

Cherchons la condition générale de positivité de $\delta^2 F$. Les variations δx et δy sont a priori indépendantes. En prenant par exemple $\delta y = 0$, $\delta x \neq 0$ ou $\delta x = 0$, $\delta y \neq 0$, la condition de positivité de $\delta^2 F$ impose que l'on ait

$$\frac{\partial^2 F}{\partial x^2} > 0, \quad \underline{\text{et}} \quad \frac{\partial^2 F}{\partial y^2} > 0 \tag{24.4}$$

Dans le cas le plus général, posons $u = (\delta x)/(\delta y)$. Il vient

$$\delta^2 F = (\delta y)^2\ \left[\ \frac{\partial^2 F}{\partial x^2}\ u^2 + 2\ u\ \frac{\partial^2 F}{\partial x \partial y}\ + \frac{\partial^2 F}{\partial y^2}\ \right] \tag{24.5}$$

Le terme entre crochets doit donc rester positif. Or, ce terme se présente comme un trinôme selon la variable u, dont le coefficient de u^2, $\partial^2 F/\partial x^2$, est positif comme trouvé plus haut. Il suffit donc d'imposer que ce trinôme n'ait pas de racine réelle, c'est-à-dire que le discriminant des racines soit négatif. On obtient ainsi l'inégalité

$$\boxed{\left[\frac{\partial^2 F}{\partial x \partial y}\right]^2 \leq \frac{\partial^2 F}{\partial x^2}\ \frac{\partial^2 F}{\partial y^2}} \tag{24.6}$$

laquelle, en posant $A = \dfrac{\partial F}{\partial x}$, $B = \dfrac{\partial F}{\partial y}$, prend la forme

$$\boxed{\left(\frac{\partial A}{\partial x}\right)_y \left(\frac{\partial B}{\partial y}\right)_x \;\geq\; \left[\left(\frac{\partial A}{\partial y}\right)_x\right]^2 \;=\; \left[\left(\frac{\partial B}{\partial x}\right)_y\right]^2} \tag{24.7}$$

Une conséquence importante des inégalités (24.4) et (24.6) ou (24.7) est la suivante. Le système étudié, se trouve initialement en équilibre stable avec le milieu ambiant. On a alors $A = A_e = 0$, $B = B_e = 0$. Supposons que le milieu ambiant impose au système une petite variation δx, n'affectant pas la gandeur B. Sous cette action, l'équilibre est rompu et la grandeur A, initialement nulle, subit la variation $\Delta^a A = \left(\dfrac{\partial A}{\partial x}\right)_y \delta x$. Cependant, une réaction va s'opérer dans le système, lequel, laissé libre d'évoluer, va revenir de lui-même vers un nouvel équilibre avec le milieu ambiant, qui tient éventuellement compte d'une modification de la variable x. Lors de ce retour vers un équilibre, la grandeur A aura subit la variation $\Delta^r A = \left(\dfrac{\partial A}{\partial x}\right)_{B=0} \delta x$, puisque B reste inchangé. Or, des différentielles $dA = \left(\dfrac{\partial A}{\partial x}\right)_y dx + \left(\dfrac{\partial A}{\partial y}\right)_x dy$ et $dB = \left(\dfrac{\partial B}{\partial x}\right)_y dx + \left(\dfrac{\partial B}{\partial y}\right)_x dy$, on tire

$$\left(\frac{\partial y}{\partial x}\right)_B = -\left(\frac{\partial B}{\partial x}\right)_y \Big/ \left(\frac{\partial B}{\partial y}\right)_x \qquad \text{et}$$

$$\left(\frac{\partial A}{\partial x}\right)_{B=0} = \left(\frac{\partial A}{\partial x}\right)_y + \left(\frac{\partial A}{\partial y}\right)_x \left(\frac{\partial y}{\partial x}\right)_B, \quad \text{soit}$$

$$\left(\frac{\partial A}{\partial x}\right)_{B=0} = \left(\frac{\partial A}{\partial x}\right)_y - \left[\left(\frac{\partial A}{\partial y}\right)_x\right]^2 \Big/ \left(\frac{\partial B}{\partial y}\right)_x$$

compte tenu de $\left(\dfrac{\partial A}{\partial y}\right)_x = \left(\dfrac{\partial B}{\partial x}\right)_y$. Ainsi

$$\Delta^r A = (\Delta^a A)(1-K), \quad \text{avec} \quad K = \left[\left(\frac{\partial A}{\partial y}\right)_x\right]^2 \Big/ \left[\left(\frac{\partial A}{\partial x}\right)_y \left(\frac{\partial B}{\partial y}\right)_x\right] \tag{24.8}$$

D'après (24.7), on a $0 \leq K \leq 1$. Par conséquent

$$\boxed{|\Delta^r A| \leq |\Delta^a A|} \tag{24.9}$$

L'action du milieu ambiant mesurée par δx, a pour effet une variation $\Delta^a A$. L'inégalité (24.9) montre que lors du rétablissement d'un équilibre, des processus internes au système tendent à diminuer le résultat de cette action. Ce résultat est conforme à un principe général de *modération* (principe de Le Chatelier-Braun), selon lequel

Une action extérieure mettant hors équilibre un système provoque en celui-ci des processus tendant à diminuer les résultats de cette action

Ce principe, dont la première formulation à propos des réactions chimiques fut donnée par Le Châtelier en 1888 [11], est de portée générale et rend compte de l'évolution de nombreux phénomènes, comme par exemple les phénomènes d'induction en électromagnétisme (loi de Lenz) ou encore de phénomènes irréversibles (conduction électrique, transferts de chaleur, etc...).

Comme présenté ci-dessous, on peut donner à ce principe une formulation mathématique très simple, faisant référence à la stabilité de l'équilibre.

Soit x une variable servant à décrire l'état d'un certain système Σ. A un instant initial de date t_0, ce système est supposé se trouver dans un état d'équilibre où x prend la valeur x_0. A partir de cet instant, l'évolution ultérieure du système sera décrite par une certaine loi horaire

$$x(t) \;=\; f(t, x_0, t_0) \tag{24.10}$$

Que se passe-t-il si l'état initial du système est perturbé ? Cette perturbation sera apportée sous la forme d'un petit déplacement ϵ_0 de la valeur initiale de x : $x'_0 = x_0 + \epsilon_0$. La nouvelle loi horaire est alors

$$x'(t) \;=\; f(t, x_0 + \epsilon_0, t_0) \tag{24.11}$$

La différence

$$x'(t) \;-\; x(t) \;=\; \epsilon(t) \;=\; f(t, x_0 + \epsilon_0, t_0) \;-\; f(t, x_0, t_0) \tag{24.12}$$

représente l'évolution de la perturbation au cours du temps. L'état initial de Σ défini par (x_0, t_0) sera reconnu comme stable si l'amplitude de la perturbation *décroît* au cours du temps, ce que, selon le *critère de stabilité de Lyapunov*, on exprime par l'inégalité

11. "Si une réaction chimique à l'équilibre est sujette à une modification de certains paramètres qui fait que celle-ci est déplacée par rapport à sa position d'équilibre, il s'ensuit que ladite réaction cherche à se réajuster à un nouvel état d'équilibre. La réaction évolue dans la direction qui, au moins en partie, contrecarre la modification imposée".

$$\frac{d\epsilon^2}{dt} \ \leq \ 0 \tag{24.13}$$

Revenons au système binaire auquel le milieu ambiant impose la variation δx . Toute perturbation par rapport à l'équilibre provoque une variation positive de la fonction F. Comme l'équilibre doit être stable, le fluide doit évoluer de telle sorte que

$$\frac{d(\Delta F)^2}{dt} \ = \ 2\Delta F \frac{d\Delta F}{dt} \leq \ 0 \tag{24.14}$$

Comme $\Delta F > 0$, on en déduit que l'on doit avoir

$$\frac{d\Delta F}{dt} \leq \ 0 \tag{24.15}$$

On a

$$\Delta F = \delta^2 F = \frac{1}{2} \left(\frac{\partial A}{\partial y} \right)_x^e (\delta x)^2 = \frac{1}{2} (\Delta A)^2 / \left(\frac{\partial A}{\partial y} \right)_x^e$$

où la notation "e" signifie que la grandeur est considérée à l'équilibre. Or, $\left(\frac{\partial A}{\partial y} \right)_x^e > 0$, d'où l'on déduit

$$\frac{d(\Delta A)^2}{dt} \ = \ 2\Delta A \ \frac{d\Delta A}{dt} \leq \ 0 \tag{24.16}$$

Si ΔA est positif (resp. négatif), alors ΔA décroît (resp. croît) au cours du temps. Autrement dit, le système réagit bien de façon à minimiser l'effet de la perturbation.

25 Inégalités diverses

25.1 $$\left(1+\frac{1}{n}\right)^{n} \leq \left(1+\frac{1}{n+1}\right)^{n+1}$$

Cette inégalité se démontre aisément en considérant la fonction $f(x) = \ln\left[(1+\frac{1}{x})^x\right] = x\left[\ln(1+x) - \ln x\right]$. On a $f'(x) = \ln\frac{1+x}{x} + \frac{1}{1+x} > 0$ et $f(x)$ est strictement croissant. On a donc $f(n+1) > f(n)$ et l'inégalité s'en déduit.

25.2 $$n\left(\sqrt[n]{n+1}-1\right) \leq 1+\frac{1}{2}+\cdots+\frac{1}{n} \leq n - \frac{n-1}{\sqrt[(n-1)]{n}}$$

Posons $x_k = 1+\frac{1}{k}$. D'après (19.9), on a

$$\sum_{k=1}^{n} x_k = n + \sum_{k=1}^{n}\frac{1}{k} \geq n\left(\prod_{k}(1+\frac{1}{k})\right)^{1/n} = n(n+1)^{1/n},$$

d'où l'inégalité à gauche annoncée. Posons ensuite $y_k = 1-\frac{1}{k}$ avec $k \geq 2$. D'après (19.9) encore, on a

$$\sum_{k=2}^{n} y_k = n-1-\sum_{k=2}^{n}\frac{1}{k} = n-\sum_{k=1}^{n}\frac{1}{k} \geq (n-1)\left(\prod_{k}(1-\frac{1}{k})\right)^{1/(n-1)} = (n-1)/n^{1/(n-1)}$$

d'où l'inégalité à droite annoncée.

25.3 $$(n)^{n/2} \leq n! \leq \left(\frac{n+1}{2}\right)^{n}$$

Pour établir ces inégalités, une très belle astuce consiste à exprimer le carré de la factorielle sous la forme

$$(n!)^2 = \prod_{k=1}^{n} k(n+1-k) \tag{25.3.1}$$

et à étudier les variations de la fonction $f(x) = x(n+1-x)$. Celle-ci présente un maximum égal à $(\frac{n+1}{2})^2$ pour $x = x_M = (n+1)/2$. Elle est croissante dans l'intervalle $1 \leq x \leq x_M$ et son minimum dans cet intervalle est donc égal à $f(1) = n$. Chacun des n termes de $(n!)^2$ étant de la forme $f(x)$, on en déduit

$$\sqrt{(f(1))^n} \leq n! \leq \sqrt{(f(x_M))^n}$$

ce qui donne bien

$$(n)^{n/2} \leq n! \leq \left(\frac{n+1}{2}\right)^n \tag{25.3.2}$$

Faisons les remarques suivantes.

• L'inégalité de droite peut être aussi obtenue en appliquant (19.9) au cas $x_k = k$, qui donne $(n!)^{1/n} \leq \dfrac{1}{n}\displaystyle\sum_{k=1}^{n} k = \dfrac{n+1}{2}$.

• Cette même inégalité donne

$$\frac{(n!)^{1/n}}{n} \leq \frac{1}{2}\left(1 + \frac{1}{n}\right), \quad \text{soit} \quad \xi = \lim_{n\to+\infty}\left[\frac{(n!)^{1/n}}{n}\right] \leq \frac{1}{2}$$

Or, $\xi = 1/e$ d'après la formule de Stirling.

25.4 $X = \dfrac{a}{b+c} + \dfrac{b}{c+a} + \dfrac{c}{a+b} \geq \dfrac{3}{2}$ (Nesbitt)

Tous les nombres considérés sont positifs. Voici une première démonstration, assez directe.

• Posons $x = a + b$, $y = b + c$, $z = c + a$. On exprime a, b et c en fonction de x, y et z : $a = (x + z - y)/2$, $b = (x + y - z)/2$, $c = (y + z - x)/2$ et on reporte

$$X = \frac{a}{b+c} + \frac{b}{c+a} + \frac{c}{a+b} = \frac{1}{2}\left[f(x,y) + f(y,z) + f(z,x)\right] - 3/2, \quad \text{où}$$
$$\text{par exemple} \quad f(x,y) = x/y + y/x = f(y,x)$$

Or, la fonction $g(u) = u + 1/u$ est minimum pour $u = 1$ et $g(1) = 2$. On a donc

$$X \geq 3 - 3/2 = 3/2$$

• Une seconde démonstration consiste à comparer la moyenne arithmétique et la moyenne harmonique des trois nombres $(a+b)$, $(b+c)$ et $(c+a)$ (formules (19.9) et (19.10)) :

$$\frac{(a+b)+(b+c)+(c+a)}{3} \geq \frac{3}{\dfrac{1}{a+b}+\dfrac{1}{b+c}+\dfrac{1}{c+a}}, \quad \text{soit}$$

$$[(a+b)+(b+c)+(c+a)]\left[\frac{1}{a+b}+\frac{1}{b+c}+\frac{1}{c+a}\right] \geq 9, \quad \text{et}$$

$$2(a+b+c)\left[\frac{1}{a+b}+\frac{1}{b+c}+\frac{1}{c+a}\right] = 6+2\left[\frac{a}{b+c}+\frac{b}{c+a}+\frac{c}{a+b}\right] \geq 9$$

d'où l'inégalité cherchée.

• Une troisième démonstration est la suivante.

$$\begin{aligned} X &= \frac{a+b+c}{b+c}+\frac{b+c+a}{c+a}+\frac{c+a+b}{a+b}-3 \\ &= (a+b+c)\left[\frac{1}{b+c}+\frac{1}{c+a}+\frac{1}{a+b}\right]-3 \\ &= \frac{1}{2}[(b+c)+(c+a)+(a+b)]\left[\frac{1}{b+c}+\frac{1}{c+a}+\frac{1}{a+b}\right]-3 \end{aligned}$$

Posant ensuite $x_1^2 = b+c$, $x_2^2 = c+a$, $x_3^2 = a+b$, $y_1 = 1/x_1$, $y_2 = 1/x_2$, $y_3 = 1/x_3$, le premier terme de X prend la forme $(x_1^2+x_2^2+x_3^2)(y_1^2+y_2^2+y_3^2)$ pour laquelle l'inégalité de Cauchy-Schwarz (7.1) donne pour borme inférieure $(x_1y_1+x_2y_2)^2 = 9$. D'où $X \geq 9/2-3 = 3/2$, c'est-à-dire, l'inégalité annoncée.

• Voici une quatrième démonstration, plus simple !

Comme $X(a,b,c)$ est complètement symétrique sous une permutation quelconque de (a,b,c), on peut supposer que $a \geq b \geq c$ et l'on a $x_3 = 1/(a+b) \leq x_2 = 1/(a+c) \leq x_1 = 1/(c+b)$. D'après les inégalités de réarrangement, on a donc

$$X = ax_1+bx_2+cx_3 \geq bx_1+cx_3+ax_2 \quad \text{et} \quad X \geq cx_1+ax_2+bx_3, \quad \text{d'où}$$

$$2X \geq (b+c)x_1+(a+c)x_2+(a+b)x_3 = 3 \quad \text{et} \quad X \geq 3/2$$

25.5 $\displaystyle X = \sum_{k=1}^{n}\frac{a_k^2}{a_{k+1}} \geq \sum_{k=1}^{n} a_k, \quad \text{avec} \quad a_{n+1} = a_1$

On applique l'inégalité de Cauchy-Schwarz de façon astucieuse :

$$\left(\sum_{k=1}^{n} a_k\right)^2 = \left(\sum_{k=1}^{n} \sqrt{a_{k+1}}\frac{a_k}{a_{k+1}}\right)^2 \leq \left(\sum_{k=1}^{n} a_{k+1}\right)\left(\sum_{k=1}^{n} \frac{a_k^2}{a_{k+1}}\right)$$
$$= \left(\sum_{k=1}^{n} a_k\right)\left(\sum_{k=1}^{n} \frac{a_k^2}{a_{k+1}}\right)$$

d'où le résultat.

25.6 $\dfrac{x_1}{x_2+\cdots+x_n}+\dfrac{x_2}{x_3+\cdots+x_n+x_1}+\cdots+\dfrac{x_n}{x_1+\cdots+x_{n-1}} \geq \dfrac{n}{n-1}$

Posons $s = x_1 + \cdots + x_n$. Alors

$$X = \frac{x_1}{x_2+\cdots+x_n}+\frac{x_2}{x_3+\cdots+x_n+x_1}+\cdots+\frac{x_n}{x_1+\cdots+x_{n-1}}$$
$$= \frac{x_1}{s-x_1}+\cdots\frac{x_n}{s-x_n} = s\left[\frac{1}{s-x_1}+\cdots+\frac{1}{s-x_n}\right]-n$$
$$= \frac{1}{n-1}\left[(s-x_1)+\cdots+(s-x_n)\right]\left[\frac{1}{s-x_1}+\cdots+\frac{1}{s-x_n}\right]-n$$
$$\geq \frac{1}{n-1}\times n^2 - n = \frac{n}{n-1}$$

où l'inégalité de Cauchy-Schwarz a été utilisée dans la dernière étape.

25.7 $X = (1+\dfrac{1}{x})(1+\dfrac{1}{y})(1+\dfrac{1}{z}) \geq 64 \quad \text{avec} \quad x+y+z=1$

$$X = 1+\frac{1}{x}+\frac{1}{y}+\frac{1}{z}+\frac{1}{xy}+\frac{1}{xz}+\frac{1}{yz}+\frac{1}{xyz}$$

Posant $g = 1/(xyz)^{1/3}$, et utilisant les inégalités entre moyennes arithmétiques et moyennes géométriques, on obtient

$$\frac{1}{x}+\frac{1}{y}+\frac{1}{z} \geq 3g,\ \frac{1}{xy}+\frac{1}{xz}+\frac{1}{yz} \geq 3g^2$$

et donc $X \geq 1+3g+3g^2+g^3 = (1+g)^3$, et comme $3/g \leq x+y+z=1$, soit $g \geq 3$, il vient finalement $X \geq 4^3 = 64$.

25.8 $\quad X = \dfrac{c^3}{a} + \dfrac{d^3}{b} \geq 1 \quad \text{avec} \quad a^2 + b^2 = (c^2 + d^2)^3$

La particularité intéressante de cette inégalité est qu'elle-même et la contrainte qui l'accompagne sont invariantes sous la transformation

$$a' = \alpha^3\, a,\ b' = \alpha^3\, b,\ c' = \alpha\, c,\ d' = \alpha\, d$$

qui permet, première astuce, de supposer d'emblée que $a^2 + b^2 = 1 = c^2 + d^2$, sans altérer sa généralité. La seconde astuce consiste à écrire

$$c^2 + d^2 = 1 = \sqrt{ac}\sqrt{\frac{c^3}{a}} + \sqrt{db}\sqrt{\frac{d^3}{b}}$$

et à appliquer l'inégalité de Cauchy-Schwarz pour obtenir

$$1 \leq \sqrt{(ac+bd)X} \quad \text{soit} \quad X \geq \frac{1}{ac+bd}$$

Mais, en appliquant à nouveau l'inégalité de Cauchy-Schwarz, on a $ac + bd \leq \sqrt{(a^2+b^2)(c^2+d^2)} = 1$, donc $1/(ac+bd) \geq 1$ et finalement $X \geq 1$.

25.9 $\quad X = \sqrt{\dfrac{2a}{a+b}} + \sqrt{\dfrac{2b}{b+c}} + \sqrt{\dfrac{2c}{c+a}} \leq 3$

Sauf existence d'une astuce démoniaque, il semble que cette inégalité puisse difficilement être prouvée à l'aide des inégalités classiques, telles celle de Cauchy-Schwarz, car celles-ci donnent des majorations trop larges de X. Par exemple, pour éliminer les racines carrées, on peut se servir du fait que la fonction "$\sqrt{}$" est concave et écrire

$$\frac{1}{3}X \leq \sqrt{H/3}, \quad \text{où} \quad H = \frac{2a}{a+b} + \frac{2b}{b+c} + \frac{2c}{c+a}, \quad \text{soit} \quad X \leq \sqrt{3H} \qquad (25.9.3)$$

On montre facilement que H est minimum et égal à 3 pour $a = b = c$, mais cela ne prouve en rien que X reste toujours inférieur à 3. Il reste donc la méthode directe consistant à rechercher les extrema de X en annulant ses dérivées premières. Pour simplifier cette recherche, posons $u = b/a$, $v = b/c$, $w = a/c$. Ces trois nouvelles variables sont liées par $u\,v\,w = 1$, ce qui montre qu'en fait X n'est fonction que de deux variables que nous choisirons comme étant u et v, Ainsi

$$X = X(u,v) = \sqrt{\frac{2}{1+u}} + \sqrt{\frac{2}{1+v}} + \sqrt{\frac{2uv}{1+uv}} \qquad (25.9.4)$$

Deux remarques concernant cette expression.

1) Lorsque $u > 1$ et $v > 1$, les deux premiers termes sont inférieurs à 1. Cependant, comme $w = 1/(uv) < 1$, le dernier terme est supérieur à 1. A l'opposé, lorque $u < 1$ et $v < 1$, les deux premiers termes sont supérieurs à 1, mais le dernier terme est alors inférieur à 1. Ainsi, le dernier terme vient contre-balancer les deux premiers, et cet effet s'observe aussi bien si l'on choisit les autres couples de variables (u, w) ou (v, w), car X est complètement symétrique vis-à-vis des $u,\ v, w$. C'est ce qui rend plausible a priori l'inégalité annoncée.

2) On observe aussi que $X(u, 0) = X(u, +\infty) = \sqrt{\dfrac{2}{1+u}} + \sqrt{2}$, d'où l'on conclut que, pour u fixé fini, X doit présenter des extrema dans l'intervalle $0 \leq v < +\infty$. Comme $X(u, v) = X(v, u)$, on a la même conclusion pour les variations vis-à-vis de u à v fixé fini.

Les extrema de $X(u, v)$ sont obtenus en annulant les dérivées partielles premières de X :

$$\begin{aligned}
\frac{\partial X}{\partial u} &= \frac{1}{\sqrt{2}}\left[\sqrt{\frac{v}{u}}\frac{1}{(1+uv)^{3/2}} - \frac{1}{(1+u)^{3/2}}\right] = 0 \\
\frac{\partial X}{\partial v} &= \frac{1}{\sqrt{2}}\left[\sqrt{\frac{u}{v}}\frac{1}{(1+uv)^{3/2}} - \frac{1}{(1+v)^{3/2}}\right] = 0
\end{aligned} \tag{25.9.5}$$

ce qui conduit aux équations

$$\frac{1}{(1+uv)^{3/2}} = \sqrt{\frac{u}{v}}\frac{1}{(1+u)^{3/2}} = \sqrt{\frac{v}{u}}\frac{1}{(1+v)^{3/2}} \tag{25.9.6}$$

La dernière donne $\dfrac{v}{(1+v)^{3/2}} = \dfrac{u}{(1+u)^{3/2}}$ dont une solution évidente est $u = v$. En fait, en posant $\alpha = 1/3$, cette équation se récrit comme

$$(v^\alpha - u^\alpha)(u^\alpha + v^\alpha - u^{2\alpha}v^{2\alpha}) = 0 \tag{25.9.7}$$

d'où l'on tire la solution $v = u$, mais aussi la solution

$$v^\alpha = \frac{1}{2u^{2\alpha}}\left[1 + \sqrt{1+4u}\right] \tag{25.9.8}$$

en excluant la solution négative pour v^α.

En reportant $v = u$ dans la première équation de (25.9.6), on obtient la seule solution $u = 1$, et donc finalement $u = v = w = 1$, soit encore $a = b = c$ et

$X = 3$. En calculant les dérivées partielles secondes de X pour cette solution, on vérifie qu'elle correspond bien à un maximum car

$$\frac{\partial^2 X}{\partial u^2} = \frac{\partial^2 X}{\partial v^2} = -\frac{1}{8} \quad \text{et} \quad \frac{\partial^2 X}{\partial u \partial v} = -\frac{1}{16} \tag{25.9.9}$$

Par contre, la solution (25.9.8) reportée dans la première équation de (25.9.6) conduit au résultat $u = 0$ qui ne correspond à aucun extremum.

En conclusion, X n'a qu'un seul maximum, égal à 3, obtenu si et seulement si $a = b = c$. L'inégalité annoncée est donc bien vérifiée.

• • • •

Et bien d'autres inégalités à prouver, ou à découvrir....