Théorème de Hermite(1873) établissant la transcendance du nombre de Neper e=2,71828...

Guy PHILIPPE

12 juin 2005

Preuve détaillée donc un peu longue mais qui n'utilise que des outils élémentaires de niveau mathématiques supérieures ou spéciales comme la formule de Leibniz de dérivation d'ordre n d'un produit ou la formule donnant l'équivalent de n ! (Stirling). Les grandes lignes sont empruntées à Jean-Paul Delahaye, "le fascinant nombre π" page 166.

Replaçons ce théorème dans son contexte historique. Liouville a construit les premiers nombres transcendants comme application de son théorème de 1844 sur les approximations rationnelles d'un irrationnel algébrique. Exemple $\sum_{n=0}^{+\infty} \frac{1}{10^{n!}}$ est transcendant. Dès lors on eut envie de savoir si les constantes usuelles étaient transcendantes ou pas. Par exemple dans la formule d'Euler $e^{i\pi} + 1 = 0$ les constantes 0, 1, i sont algébriques, qu'en est-il de e et π ?
Charles Hermite prouva en 1873, à la suite d'un travail difficile que e était transcendant. Depuis des simplifications ont été apportées par Weierstrass, Hilbert, Hurwitz et Gordan et la preuve assez simple donnée ici suit la méthode de A.Baker dans *transcendental number theory, Cambridge Mathematical Library,1975-79-90.*
Enfin Lindemann en 1882 prouva que π est transcendant.
N'oublions pas Cantor qui en 1874 montra que l'ensemble des nombres réels n'était pas dénombrable alors que Dedekind avait auparavant prouvé que l'ensemble des réels algébriques était dénombrable et par conséquent que les transcendants étaient les "plus nombreux" parmi les nombres réels.

1 Préliminaires

f désignant un polynôme à coefficients réels de degré m on pose pour $t \geq 0 \quad I(t) = \int_0^t e^{t-u} f(u)du$. D'où en intégrant par parties une première fois on obtient

$$I(t) = \int_0^t e^t f(u)d(-e^{-u}) = [-e^{t-u} f(u)]_0^t + \int_0^t e^{t-u} f'(u)du =$$

$$e^t f(0) - f(t) + \int_0^t e^{t-u} f'(u)du$$

En intégrant une deuxième fois par parties il est clair que l'on obtiendra

$$I(t) = e^t f(0) - f(t) + e^t f'(0) - f'(t) + \int_0^t e^{t-u} f''(u) du$$

Et en recommençant jusqu'à l'ordre de dérivation $m+1$ on aura

$$I(t) = e^t \sum_{j=0}^{m} f^{(j)}(0) - \sum_{j=0}^{m} f^{(j)}(t) + \int_0^t e^{t-u} f^{(m+1)}(u) du$$

Or $f^{(m+1)} = 0$ vu que degré de $f = m$ d'où finalement

$$I(t) = e^t \sum_{j=0}^{m} f^{(j)}(0) - \sum_{j=0}^{m} f^{(j)}(t)$$

D'autre part si on note $f^*(x)$ le polynôme $f(x)$ où chaque coefficient est remplacé par sa valeur absolue on aura pour tout $x \geq 0 \quad |f(x)| \leq f^*(x)$ et f^* est une fonction croissante de x sur $[0, +\infty[$. D'où pour $t \geq 0$

$$|I(t)| \leq \int_0^t e^{t-u} |f(u)| du \leq e^t \int_0^t \frac{1}{e^u} f^*(u) du \leq t e^t f^*(t)$$

Enfin montrons que pour 2 polynômes f et g on a pour tout $x \geq 0$

$$(fg)^*(x) \leq f^*(x) g^*(x)$$

En effet

$$(fg)(x) = f(x)g(x) =$$

$$(\sum_i a_i x^i)(\sum_j b_j x^j) = \sum_k c_k x^k ; c_k = \sum_{r=0}^{k} a_r b_{k-r} \Longrightarrow |c_k| \leq \sum_{r=0}^{k} |a_r||b_{k-r}|$$

d'où

$$(fg)^*(x) = \sum_k |c_k| x^k \leq \sum_k \sum_{r=0}^{k} |a_r||b_{k-r}| x^k$$

or

$$f^*(x) g^*(x) = (\sum_i |a_i| x^i)(\sum_j |b_j| x^j) = \sum_k \sum_{r=0}^{k} |a_r||b_{k-r}| x^k \geq (fg)^*(x)$$

Ce qu'il fallait démontrer.

2 Démonstration

Par l'absurde. Si le nombre e était algébrique de degré $n \quad (n \geq 2$ car e est irrationnel) et $m(x) \in \mathbb{Q}[x]$ son polynôme minimal, par définition $m(x)$ serait annulateur de e, unitaire et de degré minimun n alors il existerait un polynôme $M = q_0 + q_1 x + q_2 x^2 + ... + q_n x^n \in \mathbb{Z}[x]$ annulateur de e avec $q_0 \neq 0$. sinon si $q_0 = 0$ il est clair que l'on aurait un polynôme annulateur de e de degré inférieur à n d'où une contradiction. De plus comme n est le degré de M on a $q_n \neq 0$.
On choisit un entier premier p tel que $p > n$ et $p > |q_0|$ et on pose

$$f(x) = x^{p-1}(x-1)^p(x-2)^p...(x-n)^p$$

Dans la suite on utilisera aussi $f(x)$ sous la forme $f(x) = (x-k)^p A(x)$ pour $k = 1$ à n ou encore $f(x) = x^{p-1}B(x)$, les notations étant évidentes.
On pose aussi $m = deg(f) = (n+1)p - 1$ puis

$$J = q_0 I(o) + q_1 I(1) + ... q_n I(n)$$

Montrons que

$$-J = \sum_{j=0}^{m} \sum_{k=0}^{n} q_k f^{(j)}(k)$$

C'est facile à partir de l'égalité $I(t) = e^t \sum_{j=0}^{m} f^{(j)}(0) - \sum_{j=0}^{m} f^{(j)}(t)$ en additionnant membre à membre les égalités pour $t = 0, 1, 2...n$ après multiplication par respectivement $q_0, q_1, q_2...q_n$ et en remarquant que $q_0 + q_1 e + q_2 e^2 + ...q_n e^n = M(e) = 0$. On peut réécrire $-J$ sous la forme

$$-J = \sum_{j=0}^{m} (\sum_{k=1}^{n} q_k f^{(j)}(k) + q_0 f^{(j)}(0)) = \sum_{j=0}^{m} \sum_{k=1}^{n} q_k f^{(j)}(k) + \sum_{j=0}^{m} q_0 f^{(j)}(0))$$

Quelques résultats sur les $f^{(j)}(k)$ en vue de simplifier $-J$
• Il est clair que $f^{(j)}(k) = 0$ si $j < p$ et $k = 1, 2, 3...$ ou n.
En effet $f(x) = (x-k)^p A(x)$ d'où d'après le théorème de Leibniz avec D désignant l'opérateur de dérivation par rapport à x

$$f^{(j)}(x) = \sum_{r=0}^{j} C_j^r D^r((x-k)^p) D^{j-r} A(x)$$

Or pour $j < p$ on a $0 \leq r \leq j < p$ d'où $D^r((x-k)^p)|_{x=k} = 0$ et finalement on a bien $f^{(j)}(k) = 0$ pour $j < p$ et $k = 1, 2, 3...$ ou n.
• Il est aussi clair que $f^{(j)}(0) = 0$ si $j < p - 1$.
En effet $f(x) = x^{p-1}B(x)$ d'où $f^{(j)}(x) = \sum_{r=0}^{j} C_r^j D^r(x^{p-1}) D^{j-r} B(x)$.
Or pour $j < p - 1$ on a $0 \leq r \leq j < p - 1$ d'où $D^r(x^{p-1})|_{x=0} = 0$ et finalement on a bien $f^{(j)}(0) = 0$ si $j < p - 1$.

D'où une réécriture de $-J$ grâce à ces deux derniers résultats.

$$-J = \sum_{j=p}^{m}\sum_{k=1}^{n} q_k f^{(j)}(k) + \sum_{j=p-1}^{m} q_o f^{(j)}(0) \quad (I)$$

Quelques résultats sur les $f^{(j)}(k)$ **restants qui figurent dans** (I)

• Si $j \geq p$ et $k = 1, 2, ...$ ou n on a $p!|f^{(j)}(k)$ ce qui impliquera bien sûr que

$$p!|\sum_{j=p}^{m}\sum_{k=1}^{n} q_k f^{(j)}(k)$$

En effet

$$f^{(j)}(x) = \sum_{r=0}^{j} C_j^r D^r((x-k)^p)D^{j-r}A(x)$$

et comme $j \geq p$ pour $x = k$ le seul terme non nul de la somme correspond à $r = p$ vu que si $0 \leq r < p$ alors $D^r((x-k)^p)|_{x=k} = 0$ et si $p < r \leq j$ on a $D^r((x-k)^p) = 0$. Donc $f^{(j)}(k) = C_j^p p! D^{j-p}A(k) \in \mathbb{N}$ et ainsi $p!|f^{(j)}(k)$.

• si $j = p-1$ on a $(p-1)!|f^{p-1}(0)$ et $p! \nmid f^{p-1}(0)$
En effet $f^{(p-1)}(x) = \sum_{r=0}^{p-1} C_{p-1}^r D^r(x^{p-1})D^{p-1-r}B(x)$
or $D^r(x^{p-1})|_{x=0} = 0$ si $r = 0, 1, ..., p-2$ et $D^{p-1}(x^{p-1}) = (p-1)!$ d'où $f^{p-1}(0) = C_{p-1}^{p-1}(p-1)!D^{(0)}B(0) = (p-1)!B(0) = (p-1)!(-1)^{np}(n!)^p$ Il est donc clair que $(p-1)!|f^{p-1}(0)$. De plus $p! \nmid f^{p-1}(0)$ car sinon on aurait $p!|(-1)^{np}(p-1)!(n!)^p$ d'où on aurait $p|(n!)^p$ et comme p est premier on aurait $p|n! = 1.2.3...n$ et finalement p devrait diviser un des nombres 1, 2,...n ce qui contredirait l'hypothèse $p > n$
• Si $j > p-1$ alors $p!|f^{(j)}(0)$
En effet $f^{(j)}(x) = \sum_{r=0}^{j} C_j^r D^r(x^{p-1})D^{j-r}B(x)$ et pour $x = 0$ tous les termes sont nuls sauf pour $r = p-1$ d'où $f^{(j)}(0) = C_j^{p-1}(p-1)!D^{j-p+1}B(0)$ or $j > p-1$ donc $j-p \geq 0$ et comme $B = (x-1)^p(x-2)^p...(x-n)^p$ on a $DB(x) = pL(x)$, donc $D^{j-p+1}B(x) = D^{j-p}(DB)(x) = D^{j-p}(pL)(x) = pD^{j-p}L(x)$ et finalement $D^{j-p+1}B(0) = pD^{j-p}L(0)$ ce qui entraîne bien $p!|f^{(j)}(o)$ si $s > p-1$.

Dès lors on est en mesure d'affirmer que $J \neq 0$
Sinon on aurait d'après (I)

$$\sum_{j=p}^{m}\sum_{k=1}^{n} q_k f^{(j)}(k) + \sum_{j=p-1}^{m} q_0 f^{(j)}(0) = 0$$

ou encore

$$-q_0 f^{(p-1)}(0) = \sum_{j=p}^{m}\sum_{k=1}^{n} q_k f^{(j)}(k) + \sum_{j>p-1}^{m} q_0 f^{(j)}(0)$$

Or, d'après ce qui précède, $p!$ diviserait le membre de droite de la dernière égalité et donc $p!|q_0 f^{(p-1)}(0) = q_0(-1)^{np}(p-1)!(n!)^p$ ce qui reviendrait à $p|q_0(n!)^p$. Mais on a supposé $p > |q_0|$ et $q_0 \neq 0$ donc $p \nmid q_0$ (sinon on aurait $p||q_0|$ et $|q_0| \neq 0 \Longrightarrow p \leq |q_0|$, d'où une contradiction) et comme p est premier on en déduirait que $p|(n!)^p$ puis $p|n! = 1.2.3..n$ et enfin p diviserait un des nombres 1, 2, ...n ce qui contredirait l'hypothèse $p > n$.
Par conséquent on a bien $J \neq 0$ et comme d'après (I) tous les termes du membre de droite de (I) sont divisibles par $p!$ sauf un : $q_0 f^{(p-1)}(0)$ qui lui est divisible par $(p-1)!$ on a $(p-1)!|J$ puis enfin la minoration de $|J|$

$$|J| \geq (p-1)!$$

Une majoration de $|J|$
Montrons qu'il existe une constante $c > 0$ telle que pour un entier premier p suffisament grand on ait $|J| \leq c^p$.
D'abord de $(fg)^*(x) \leq f^*(x)g^*(x)$ pour $x \geq 0$ et en généralisant à un produit de plus de 2 facteurs on tire

$$f^*(x) \leq (x^{p-1})^*\{[(x-1)(x-2)...(x-n)]^p\}^* \leq x^{p-1}[(x-1)^*(x-2)^*...(x-n)^*]^p$$

$$f^*(x) \leq x^{p-1}[(x+1)(x+2)...(x+n)]^p$$

D'où pour $k = 1, 2, ...$ ou n

$$f^*(k) \leq k^{p-1}[(k+1)(k+2)...(k+n)]^p \leq n^{p-1}[(n+1)(n+2)...(n+n)]^p$$

$$f^*(k) \leq (2n)^{p-1}[(2n)(2n)...(2n)]^p = (2n)^{p-1}(2n)^{np} = (2n)^{(n+1)p-1}$$

$$f^*(k) \leq (2n)^m$$

Dès lors vu que $I(0) = 0$ on a

$$|J| \leq |q_1||I(1)| + |q_2||I(2)| + ...|q_n||I(n)|$$

et comme pour $t \geq 0 \quad |I(t)| \leq te^t f^*(t)$ on aura pour $k = 1, 2, ..$ ou n

$$|I(k)| \leq ke^k f^*(k) \leq ke^k(2n)^m$$

d'où

$$|J| \leq (|q_1|e + |q_1|2e^2 + |q_3|3e^3 + ...|q_n|ne^n)(2n)^m = C_n(2n)^m \leq c^p$$

où on a posé $C_n = |q_1|e + |q_1|2e^2 + |q_3|3e^3 + ...|q_n|ne^n$
On cherche une constante $c > 0$ vérifiant la dernière inégalité pour p assez grand.

$$C_n(2n)^m \leq c^p \Longleftrightarrow C_n(2n)^{(n+1)p-1} \leq c^p \Longleftrightarrow \frac{1}{2n}C_n(2n)^{(n+1)p} \leq c^p$$

La condition $C_n(2n)^{(n+1)p} \leq c^p$ est donc suffisante pour avoir l'inégalité cherchée vu que $\frac{1}{2n} < 1$.

$$C_n(2n)^{(n+1)p} \leq c^p \Longleftrightarrow ln(C_n) + (n+1)pln(2n) \leq pln(c)$$

$$C_n(2n)^{(n+1)p} \leq c^p \Longleftrightarrow \frac{ln(C_n)}{p} + (n+1)ln(2n) \leq ln(c)$$

Or $ln(C_n)$ est une constante d'où à partir d'un certain rang pour p on a $\frac{ln(C_n)}{p} \leq 1$. Par conséquent si on choisit pour c la valeur $c = e(2n)^{(n+1)}$ l'inégalité cherchée est vraie pour p assez grand.
En effet $e(2n)^{(n+1)} = c \Longrightarrow ln(e)+(n+1)ln(2n) = 1+(n+1)ln(2n) = ln(c)$.
or $\frac{ln(C_n)}{p} \leq 1$ et donc $\frac{ln(C_n)}{p} + (n+1)ln(2n) \leq 1 + (n+1)ln(2n) = ln(c)$.
Ce qu'il fallait démontrer.

Réduction à l'absurde
On a donc l'encadrement $(p-1)! \leq |J| \leq c^p$ pour p assez grand et en faisant tendre p vers $+\infty$ une contradiction va apparaître.

$$(p-1)! \leq c^p \Longleftrightarrow p! \leq pc^p \Longleftrightarrow \frac{p!}{pc^p} \leq 1$$

Or avec la formule de Stirling on a

$$p! \sim (\frac{p}{e})^p\sqrt{2\pi p} \Longrightarrow \frac{p!}{pc^p} \sim (\frac{p}{e})^p\sqrt{2\pi p}\frac{1}{pc^p} = (\frac{p}{ec})^p\sqrt{2\pi}\frac{1}{\sqrt{p}} =$$

$$(\frac{p}{ec})^{p-1}\frac{p}{ec}\sqrt{2\pi}\frac{1}{\sqrt{p}} = (\frac{p}{ec})^{p-1}\frac{\sqrt{2\pi}}{ec}\sqrt{p}$$

D'où $\frac{p!}{pc^p} \sim (\frac{p}{ec})^{p-1}\frac{\sqrt{2\pi}}{ec}\sqrt{p} \mapsto +\infty$ quand $p \mapsto +\infty$ ce qui contredirait $\frac{p!}{pc^p} \leq 1$.

Finalement le nombre $e = \sum_{n=0}^{+\infty} \frac{1}{n!}$ est bien transcendant.

Pour tout contact$\rightsquigarrow$ guyphilippe2@aol.com