Some remarks on Paul J. Nahin's “When Least is Best” (WLiB)

France Dacar, Jožef Stefan Institute

January 12, 2010

Maximization of $x(C-x)$, slightly generalized

The problem of dividing a constant into two parts so that their product is maximum, discussed on page 7 of WLiB, can be perceived as a special case of maximizing the product $f(x)f(a-x)$, where f is a positive, logarithmically concave function.

Let a be a positive real number. Let $f(x)$ be defined and positive for $0 < x < a$, and suppose that $\ln f$ is concave. We want to maximize the product $f(x)f(a-x)$. We shall show that the product attains the absolute maximum at $x = \frac{1}{2}a$; moreover, if $\ln f$ is *strictly* concave[1], then it is only at $x = \frac{1}{2}a$ that the product has the maximum value.

Define $g := \ln f$. By concavity of g we have

$$\tfrac{1}{2}g(x) + \tfrac{1}{2}g(a-x) \;\leqslant\; g\bigl(\tfrac{1}{2}x + \tfrac{1}{2}(a-x)\bigr) \;=\; g\bigl(\tfrac{1}{2}a\bigr)\,, \tag{1}$$

that is, $g(x) + g(a-x) \leqslant 2g\bigl(\frac{1}{2}a\bigr)$. Writing $F(x) := f(x)f(a-x)$, we see that $F(x) \leqslant F\bigl(\frac{1}{2}a\bigr)$ for $0 < x < a$. Now suppose g is strictly convex, and let $0 < x < a$, $x \neq \frac{1}{2}a$; then $x \neq a - x$, the inequality in (1) is strict, and consequently $F(x) < F\bigl(\frac{1}{2}a\bigr)$. Done.

We now examine the behaviour of $G(x) := g(x) + g(a-x)$, and with it the behaviour of $F(x) = e^{G(x)}$, on the interval $0 < x < a$. First note the symmetry $G(a-x) = G(x)$, and hence $F(a-x) = F(x)$, relative to the midpoint $\frac{1}{2}a$ of the interval. Next observe that $g(a-x)$ is concave because $g(x)$ is concave, thus G is also concave; if g is strictly concave, then G is strictly concave. We claim that $G(x)$ increases (strictly, if g is strictly concave) for $0 < x \leqslant \frac{1}{2}a$, and, because of the symmetry noted above, (strictly) decreases for $\frac{1}{2}a \leqslant x < a$; from this it will follow that $F(x)$ behaves in the same way. Let $0 < x < y \leqslant \frac{1}{2}a$; we shall show that $G(x) \leqslant G(y)$, where the inequality is strict when g is strictly concave. We can write $y = (1-\lambda)x + \lambda\frac{1}{2}a$, where $\lambda = (y-x)/\bigl(\frac{1}{2}a - x\bigr)$ lies in the range $0 < \lambda \leqslant 1$. We have

$$G(y) \;\geqslant\; (1-\lambda)G(x) + \lambda G\bigl(\tfrac{1}{2}a\bigr)$$

by concavity of G; taking into account that $G\bigl(\frac{1}{2}a\bigr) \geqslant G(x)$, with the inequality being strict if g is strictly concave, we obtain the promised inequality.

In the situation we have just examined the function g is concave on an open interval, and because of that it is continuous, and so is f. Now assume that we have a function $f(x)$ that is defined and non-negative on the closed interval $0 \leqslant x \leqslant a$, while it is positive on the open interval $0 < x < a$ with $\ln f(x)$ concave on this open interval; moreover, assume $f(x)$ to be continuous at $x = 0$ and $x = a$ (besides being continuous for $0 < x < a$). Define $F(x) := f(x)f(a-x)$ for $0 \leqslant x \leqslant a$. From what we know about the behaviour of $F(x)$ on the open interval $0 < x < a$, and because of the continuity of $F(x)$, we conclude that

$$F(0) \;=\; \lim_{x\searrow 0} F(x) \;=\; F(a) \;=\; \lim_{x\nearrow a} F(x) \;=\; \inf_{0<x<a} F(x)\,,$$

and that $F(x)$ increases for $0 \leqslant x \leqslant \frac{1}{2}a$ while it decreases for $\frac{1}{2}a \leqslant x \leqslant a$ (strictly increases/decreases provided $\ln f$ is strictly concave on the open interval).

Taking one last look at our slightly generalized maximization problem, suppose that $f(x)$ is defined, positive, and differentiable on the open interval $0 < x < a$, and that $\ln f$ is strictly

[1] A real function $g(x)$ defined on an open interval $a < x < b$ (where possibly $a = -\infty$ and/or $b = +\infty$) is **convex** (**concave**) if $g\bigl((1-\lambda)x + \lambda y\bigr) \leqslant (1-\lambda)g(x) + \lambda g(y)$ (resp. $g\bigl((1-\lambda)x + \lambda y\bigr) \geqslant (1-\lambda)g(x) + \lambda g(y)$) for any x, y, and λ such that $a < x, y < b$ and $0 \leqslant \lambda \leqslant 1$, and is **strictly convex** (**strictly concave**) if $g\bigl((1-\lambda)x + \lambda y\bigr) < (1-\lambda)g(x) + \lambda g(y)$ (resp. $g\bigl((1-\lambda)x + \lambda y\bigr) > (1-\lambda)g(x) + \lambda g(y)$) for any x, y, and λ such that $a < x, y < b$, $x \neq y$, and $0 < \lambda < 1$. It is not quite clear whether strict convexity (concavity), as defined on page 336, is the same notion as the strict convexity (strict concavity) by the usual definition, the one given here. For example, it is somewhat peculiar that the Jensen's inequality, given on pages 336 and 337, is said to hold for a strictly convex (or concave) function, while it is known to hold for *any* convex (concave) function.

concave. Strict concavity of $\ln f$ is equivalent to f'/f being strictly decreasing, and it follows that the derivative

$$\big(f(x)f(a-x)\big)' \;=\; f(x)f(a-x)\left(\frac{f'(x)}{f(x)} - \frac{f'(a-x)}{f(a-x)}\right)$$

is positive for $0 < x < \frac{1}{2}a$ and is negative for $\frac{1}{2}a < x < a$.

As an example, consider the problem of determining the shape of the right triangle with the largest possible isoperimetric quotient; the answer is, of course, the isosceles right triangle, but let us prove this.

The shape of a right triangle with sides a, b, and c (where c is the hypotenuse) is determined by the interior angle α opposite the side a. We can assume that $c = 1$; then $a = \sin\alpha$, $b = \cos\alpha$, the perimeter is $P(\alpha) = 1 + \sin\alpha + \cos\alpha$, and the area is $A(\alpha) = \frac{1}{2}\sin\alpha\cos\alpha$. We want to maximize $Q(\alpha) = 4\pi A(\alpha)/P(\alpha)^2$. Since

$$\begin{aligned} P(\alpha)^2 &= (1 + \sin\alpha + \cos\alpha)^2 \\ &= 1 + \sin^2\alpha + \cos^2\alpha + 2\sin\alpha + 2\cos\alpha + 2\sin\alpha\cos\alpha \\ &= 2(1 + \sin\alpha + \cos\alpha + \sin\alpha\cos\alpha) \\ &= 2(1 + \sin\alpha)(1 + \cos\alpha)\,, \end{aligned}$$

we have

$$Q(\alpha) \;=\; \pi\,\frac{\sin\alpha}{1+\cos\alpha}\,\frac{\cos\alpha}{1+\sin\alpha} \;=\; \pi f(\alpha) f\big(\tfrac{1}{2}\pi - \alpha\big)\,,$$

where

$$f(\alpha) \;=\; \frac{\sin\alpha}{1+\cos\alpha} \;=\; \tan\big(\tfrac{1}{2}\alpha\big)\,.$$

Since

$$\frac{f'(\alpha)}{f(\alpha)} \;=\; \frac{1}{\sin\alpha}$$

is strictly decreasing for $0 < \alpha < \frac{1}{2}\pi$, it follows that $Q(\alpha)$ is unimodal with the unique maximum at $\alpha = \frac{1}{4}\pi$.

Maximization of $x(C-x)$, generalized slightly more

In this section we seek the maximum of the product $f(x_1)\cdots f(x_m)$, where f is a positive, logarithmically concave function.

We start on the level where the logarithm of f lives. Let $a > 0$, let the real-valued function $g(t)$ be defined and concave for $0 < t < a$, and let $m \geqslant 2$ be an integer. For any $x = (x_1, \ldots, x_m)$, where $x_1 > 0$, ..., $x_m > 0$ and $x_1 + \cdots + x_m = a$, define

$$G(x) \;:=\; g(x_1) + \cdots + g(x_m)\,; \tag{2}$$

here $x = (x_1, \ldots, x_m)$ is a point in the interior Δ° of the $(m-1)$-simplex Δ with the m vertices $(a, 0, \ldots, 0)$, $(0, a, \ldots, 0)$, ..., $(0, 0, \ldots, a)$. By concavity of g we have

$$\tfrac{1}{m}g(x_1) + \cdots + \tfrac{1}{m}g(x_m) \;\leqslant\; g\big(\tfrac{1}{m}x_1 + \cdots + \tfrac{1}{m}x_m\big) \;=\; g\big(\tfrac{1}{m}a\big)\,, \tag{3}$$

and multiplying this inequality by m we get

$$G(x) \;\leqslant\; G(b)\,, \qquad \text{where } b = \big(\tfrac{1}{m}a, \ldots, \tfrac{1}{m}a\big)\,. \tag{4}$$

If g is strictly concave, and $x \neq b$, then the inequality (3) is strict, and so is the inequality (4). Thus G attains its absolute maximum at the center b of the simplex Δ, and when g is strictly convex, the center b is the only point of the open simplex Δ° at which the value of G is the absolute maximum.

Let x be any point in the open simplex Δ°, and $y = (1-\lambda)x + \lambda b$, $0 \leqslant \lambda \leqslant 1$, a point of the line segment $[x\text{--}b]$; then

$$\begin{aligned} G(y) &= G\big((1-\lambda)x + \lambda b\big) \\ &= g\big((1-\lambda)x_1 + \lambda \tfrac{1}{m} a\big) + \cdots + g\big((1-\lambda)x_m + \lambda \tfrac{1}{m} a\big) \\ &\geqslant (1-\lambda)g(x_1) + \lambda g(\tfrac{1}{m}a) + \cdots + (1-\lambda)g(x_m) + \lambda g(\tfrac{1}{m}a) \\ &= (1-\lambda)G(x) + \lambda G(b) \\ &\geqslant (1-\lambda)G(x) + \lambda G(x) \\ &= G(x)\,. \end{aligned} \tag{5}$$

If g is strictly concave and $y \neq x$ (hence $x \neq b$), then $G(b) > G(x)$, $\lambda > 0$, $\lambda G(b) > \lambda G(x)$, the second inequality in (5) is strict, and we have $G(y) > G(x)$. Therefore, if g is strictly concave, then $G(x)$ is radially unimodal (it strictly decreases along any line segment in Δ° starting at the point b.)

We shall show that with $g(t)$ as above (defined and concave for $0 < t < a$), the value $g(t)$ tends to a finite value or decreases to $-\infty$ as t decreases to 0, and the same is true when t increases to a. Choose real numbers a' and a'' so that $0 < a' < a'' < a$, and let $\ell(t)$ be the linear function defined on the open interval $0 < t < a$ such that $\ell(a') = g(a')$ and $\ell(a'') = g(a'')$. The difference $h(t) = g(t) - \ell(t)$ is a concave function which has $h(a') = h(a'') = 0$. By concavity of h we have $h(t) \geqslant 0$ for $a' \leqslant t \leqslant a''$. Let $0 < t \leqslant a'$; then $a' = (1-\lambda)t + \lambda a''$ for a suitable λ with $0 \leqslant \lambda < 1$, and therefore $0 = h(a') \geqslant (1-\lambda)h(t) + \lambda h(a'') = (1-\lambda)h(t)$, whence $h(t) \leqslant 0$ because $1 - \lambda > 0$. Similarly we find that $h(t) \leqslant 0$ for $a'' \leqslant t < a$. There exists some c between a' and a'' such that $h(c) \geqslant h(t)$ for $a' \leqslant t \leqslant a''$, and since $h(c) \geqslant 0$, we have $h(c) \geqslant h(t)$ for every t in the interval $0 < t < a$. Moreover, $h(t)$ increases for $0 < t \leqslant c$ and decreases for $c \leqslant t < a$, again because of the concavity of h. It follows that if $t \searrow 0$ (or $t \nearrow a$), then $h(t)$ either decreases to $-\infty$ or converges to a finite limit, and that the same holds for the function $g(t)$.

Now we consider a real-valued function $f(t)$, defined and non-negative for $0 \leqslant t \leqslant a$, where $f(t) > 0$ and $g(t) := \ln f(t)$ is defined and concave for $0 < t < a$. We know that $g(t)$, and hence $f(t)$, is continuous for $0 < t < a$, and that the limits $f(0+) = \lim_{t\searrow 0} f(t)$ and $f(a-) = \lim_{t\nearrow a} f(t)$ exist; we are assuming that $f(t)$ is continuous at $t = 0$ and $t = a$, i.e. that $f(0) = f(0+)$ and $f(a) = f(a-)$. For any point $x = (x_1, \ldots, x_m)$ of the closed simplex Δ define

$$F(x) \;:=\; f(x_1)\cdots f(x_m)\,.$$

Then F is continuous on the closed simplex Δ, and $F(x) = e^{G(x)}$ for every point x of the open simplex Δ°, with $G(x)$ defined as in (2). From what we know of the behaviour of the function G, we conclude that $F(x) \leqslant F(b)$ for every point x of the simplex Δ (where b is the center of the simplex), and that $F(x) < F(b)$ if $x \neq b$ and g is strictly concave. Moreover, for any point $x \neq b$ of the simplex Δ, $F(y)$ decreases along the line segment $[b\text{--}x]$, and it strictly decreases along this line segment when g is strictly concave.

The AM-GM inequality says that for any integer $m \geqslant 1$ and any non-negative real numbers $x_1, \ldots, x_m$ we have

$$\frac{x_1 + \cdots + x_m}{m} \;\geqslant\; \sqrt[m]{x_1 \cdots x_m}\,,$$

where the equality holds if and only if $x_1 = \cdots = x_m$. We obtain the AM-GM inequality as a special case of the inequality $F(x) \leqslant F(b)$, as follows. We choose $f(t) = t$; $\ln f(t) = \ln t$ is strictly concave for $t > 0$. The AM-GM inequality is trivially true if $m = 1$ or all x_i are 0, so assume that $m \geqslant 2$ and some x_i is positive, whence $a := x_1 + \cdots + x_m > 0$. Then the inequality $F(x) \leqslant F(b)$ reads as

$$x_1 \cdots x_m \;\leqslant\; \left(\tfrac{1}{m}a\right)^m,$$

where the equality holds if and only if $x_1 = \cdots = x_m = \frac{1}{m}a$; but this statement is clearly equivalent to the AM-GM inequality.

As in the preceding section, we give an example: determine the shape of a tetrahedron with a right-angled corner which maximizes the isoperimetric quotient. As in the previsous section the answer is 'obvious', namely the isosceles right-angled tetrahedron (the three sides emanating from the right-angled corner have equal lengths).

Let the vertices of the tetrahedron be the origin O of the Euclidean coordinate system Oxyz, and the points A $= (x, 0, 0)$, B $= (0, y, 0)$, and C $= (0, 0, z)$ on the coordinate axes, where the coordinates x, y, z are positive. Denote by a, b, c, and d the areas of the triangles (in the same order) OBC, OAC, OAB, and ABC. We know that a, b, c, and d are positive, and that $a^2 + b^2 + c^2 = d^2$; since the isoperimetric quotient is invariant under scaling, we can assume that $d = 1$. For those bodies in a three-dimensional Euclidean space that have a surface area S and a volume V, the isoperimetric quotient is defined as

$$Q \;:=\; 36\pi \frac{V^2}{S^3}$$

In our case the surface area is $S = 1 + a + b + c$. To compute the volume, we first determine x, y, and z from $a = \frac{1}{2}yz$, $b = \frac{1}{2}xz$, $c = \frac{1}{2}xy$; we have $abc = \frac{1}{8}(xyz)^2$, thus $\frac{1}{2}xyz = \sqrt{2abc}$, and $x = \left(\frac{1}{2}xyz\right)/a = \sqrt{2bc/a}$, $y = \sqrt{2ac/b}$, $z = \sqrt{2ab/c}$. The volume is $V = \frac{1}{3}ax = \frac{1}{3}\sqrt{2abc}$, and the isoperimetric quotient is

$$Q(a, b, c) \;=\; 8\pi \frac{abc}{(1 + a + b + c)^3}\,.$$

By the AM-GM inequality,

$$\tfrac{1}{3}(a + b + c) \;\geqslant\; (abc)^{1/3} \;=:\; u\,,$$

where the equality holds if and only if $a = b = c = u$. Then we have the inequality

$$Q(a, b, c) \;\leqslant\; 8\pi \frac{abc}{\left(1 + 3(abc)^{1/3}\right)^3} \;=\; 8\pi \left(\frac{u}{1 + 3u}\right)^3,$$

which is an equality if and only if $a = b = c = u$. Note that

$$\frac{u}{1 + 3u} \;=\; \frac{1}{3} - \frac{1}{3(1 + 3u)}$$

is an increasing function for $u > -\frac{1}{3}$. Now $u^6 = a^2b^2c^2$, where $a^2 + b^2 + c^2 = 1$, is the largest when $a^2 = b^2 = c^2 = 1/3$, thus $u_{\max} = 1/\sqrt{3}$, and we have

$$Q(a, b, c) \;\leqslant\; 8\pi \left(\frac{u_{\max}}{1 + 3u_{\max}}\right)^3 \;=\; Q\left(1/\sqrt{3},\, 1/\sqrt{3},\, 1/\sqrt{3}\right),$$

where the equality holds if and only if $a = b = c = 1/\sqrt{3}$, that is, if and only if $x = y = z$.

Jensen's inequality

In the preceding section we used Jensen's inequality for a concave function, together with the necessary and sufficient conditions for the inequality to be strict when the function is strictly concave. The formulation of Jensen's inequality in Appendix B of WLiB lacks the criteria for strict inequality (equivalently, for equality), though it assumes that the function is strictly convex (or strictly concave). For this reason we shall spell Jensen's inequality out in full, this time only for a (strictly) convex function, and shall also provide a proof.

Jensen's inequality. *Let Y be a nonempty interval of real numbers (i.e. a nonempty convex subset of the real line), and let f be a real-valued function defined and convex on Y. Let $n \geqslant 1$ be an integer. If $c_1, \ldots, c_n$ are any non-negative real numbers which sum up to* 1, *and $x_1, \ldots, x_n$ are any real numbers lying in the interval Y, then*

$$f(c_1x_1 + \cdots + c_nx_n) \;\leqslant\; c_1f(x_1) + \cdots + c_nf(x_n)\,. \tag{6}$$

Moreover, if f is strictly convex, then the inequality (6) *is an equality if and only if those numbers x_i, whose corresponding coefficients c_i are nonzero, are all equal to each other.*

Proof by induction on n. In the case $n = 1$ the above assertion is trivially true.

Consider the case $n > 1$. If $c_i = 0$ for some i, then we have a case (n−1), and the assertion for this case, which is true by induction hypothesis, is at the same time the assertion for our case n with $c_i = 0$.

It remains to prove the assertion assuming that all c_i are nonzero. Then

$$0 \;<\; c_1 + \cdots + c_{n-1} \;=\; 1 - c_n \;<\; 1\,,$$

and we have

$$\begin{aligned} f(c_1x_1 + \cdots + c_nx_n) \;&=\; f\left((1-c_n)\frac{c_1x_1 + \cdots + c_{n-1}x_{n-1}}{1-c_n} \;+\; c_nx_n\right) \\ &\leqslant\; (1-c_n)\,f\left(\frac{c_1}{1-c_n}x_1 + \cdots + \frac{c_{n-1}}{1-c_n}x_{n-1}\right) \;+\; c_nf(x_n) \\ &\leqslant\; (1-c_n)\left(\frac{c_1}{1-c_n}f(x_1) + \cdots + \frac{c_{n-1}}{1-c_n}f(x_{n-1})\right) \;+\; c_nf(x_n) \\ &=\; c_1f(x_1) + \cdots + c_{n-1}f(x_{n-1}) + c_nf(x_n)\,, \end{aligned} \tag{7}$$

where the first inequality holds because f is convex, and the second inequality holds by induction hypothesis

Suppose f is strictly convex. The first and the last expressions of the derivation (7) are equal if and only if both inequalities in the derivation are equalities. The first inequality is an equality if and only if

$$\frac{c_1x_1 + \cdots + c_{n-1}x_{n-1}}{1-c_n} \;=\; x_n \tag{8}$$

because f is strictly convex, and the second inequality is an equality if and only if

$$x_1 \;=\; \cdots \;=\; x_{n-1}\,, \tag{9}$$

this by induction hypothesis. The conjuction of conditions (8) and (9) is equivalent to

$$x_1 \;=\; \cdots \;=\; x_{n-1} \;=\; x_n\,.$$

Done.

The AM-QM inequality without tears

The direct proof of AM-QM inequality in Appendix B of WLiB is quite unnecessarily complicated. Why not give the well-known elegant[2] proof?

THE AM-QM INEQUALITY. *For any integer $n \geqslant 1$, if $c_1, \ldots, c_n$ are positive real numbers summing up to 1, and $x_1, \ldots, x_n$ are any real numbers, then*

$$\left|c_1 x_1 + \cdots + c_n x_n\right| \leqslant \sqrt{c_1 x_1^2 + \cdots + c_n x_n^2}\,, \tag{10}$$

where the inequality is an equality if and only if $x_1 = \cdots = x_n$.

Write

$$A := c_1 x_1 + \cdots + c_n x_n\,, \qquad B := c_1 x_1^2 + \cdots + c_n x_n^2\,.$$

The quadratic polynomial

$$t^2 - 2At + B = c_1(t - x_1)^2 + \cdots + c_n(t - x_n)^2 \tag{11}$$

is nonnegative for all real t, therefore its discriminant $A^2 - B$ is non-positive, that is, $A^2 \leqslant B$, or equivalently,

$$|A| \leqslant \sqrt{B}\,, \tag{12}$$

which is the AM-QM inequality (10).

The inequality in (12) is an equality if and only if the discriminant is 0, if and only if the quadratic polynomial (11) has a zero, if and only if $x_1 = \cdots = x_n$.

Done.

The AM-QM inequality is a special case of Schwarz inequality.

SCHWARZ INEQUALITY. *For any integer $n \geqslant 1$, if $c_1, \ldots, c_n$ are any positive real numbers, and $x_1, \ldots, x_n, y_1, \ldots, y_n$ are any real numbers, then*

$$\left|c_1 x_1 y_1 + \cdots + c_n x_n y_n\right| \leqslant \sqrt{c_1 x_1^2 + \cdots + c_n x_n^2} \cdot \sqrt{c_1 y_1^2 + \cdots + c_n y_n^2} \tag{13}$$

where the inequality is an equality if and only if there exist real numbers λ and μ, not both 0, such that $\lambda x_1 = \mu y_1, \ldots, \lambda x_n = \mu y_n$.

Schwarz inequality is derived from non-negativity of the (at most) quadratic polynomial

$$c_1(y_1 t - x_1)^2 + \cdots + c_n(y_n t - x_n)^2\,.$$

When $c_1 + \cdots + c_n = 1$ and $y_1 = \cdots = y_n = 1$, Schwarz inequality is the AM-QM inequality.

The triangle inequality is a consequence of Schwarz inequality.

TRIANGLE INEQUALITY. *For any integer $n \geqslant 1$, if $c_1, \ldots, c_n$ are any positive real numbers, and $x_1, \ldots, x_n, y_1, \ldots, y_n$ are any real numbers, then*

$$\sqrt{c_1(x_1+y_1)^2 + \cdots + c_n(x_n+y_n)^2} \leqslant \sqrt{c_1 x_1^2 + \cdots + c_n x_n^2} + \sqrt{c_1 y_1^2 + \cdots + c_n y_n^2}$$

where the inequality is an equality if and only if there exist non-negative real numbers λ and μ, not both 0, such that $\lambda x_1 = \mu y_1, \ldots, \lambda x_n = \mu y_n$.

[2]There are those who would sneer at the proof as being 'slick'.

Which is larger, e^π or π^e?

Section 5.1 of WLiB. The third paragraph on page 141 begins:

> Start by defining the function $h(x) = \big(\ln(x)/x\big)$ (thinking of this definition is the "hard" part of the problem!).

The function $(\ln x)/x$ need not be snatched out of thin air: we can 'discover' it by manipulating the problem into a more solver-friendly form. First, we generalize the particular question "Which is larger, e^π or π^e?" to "Given real numbers x, $y > 0$, which is larger, x^y or y^x?". Since raising positive reals to the power $1/(xy)$ is a strictly increasing function, the inequality $x^y < y^x$ is equivalent to the inequality $x^{1/x} < y^{1/y}$; taking logarithms, and remembering that $\ln x$, defined on positive reals numbers x, is strictly increasing, we obtain another equivalent inequality,

$$\frac{\ln x}{x} < \frac{\ln y}{y}\,.$$

And here we have it, the function $h(x) = (\ln x)/x$. Observing the sign of the first derivative

$$h'(x) = (\ln x)'\frac{1}{x} + (\ln x)\Big(\frac{1}{x}\Big)' = \frac{1-\ln x}{x^2}\,,$$

we see that $h(x)$ strictly increases when $0 < x \leqslant e$, and strictly decreases when $x \geqslant e$, thus $h(x)$ attains the maximum $1/e$ at $x = e$, and nowhere else, and hence $x^{1/x} = e^{h(x)}$ atains the maximum $e^{1/e}$ at $x = e$, and nowhere else. It follows that if x, $y \leqslant e$, then $x^y < y^x$ iff $x < y$, and if x, $y \geqslant e$, then $x^y < y^x$ iff $x > y$. In particular, we get the answer to the original question:

$$\pi^e < e^\pi\,.$$

The function $x^{1/x}$ strictly increases to $1^{1/1} = 1$ for $0 < x \leqslant 1$, then strictly increases to $e^{1/e}$ for $1 \leqslant x \leqslant e$, and after that strictly decreases for $x \geqslant e$, approaching 1 as $x \to +\infty$. Thus, if $0 < x \leqslant 1$ and $y > 1$, then $x^y < y^x$; we can see this directly, since $x^y \leqslant 1 < y^x$. When $1 < x < e$ and $y > e$, we can gain no information from what we know about the behaviour of $h(x)$ that would help us compare x^y with y^x. There is a unique bijection $j\colon (1\,..\,e] \to [e\,..+\infty)$ such that $j(x)^{1/j(x)} = x^{1/x}$ for every x, $1 < x \leqslant e$. If $1 < x < e$ and $y > e$, then $x^y < y^x$ iff $j(x) > y$; alas, this is of some help only if we intend to compare x^y with y^x for many $y > e$, but for some fixed x, $1 < x < e$, since then we have to compute $j(x)$ (which is quite a task) only once, to be used in many comparisons.

We can use what we have learned about the behaviour of $x^{1/x}$ to easily solve the following problem: find all integer solutions x, y, with $0 < x < y$, of the equation $x^y = y^x$. Rewriting the equation as $x^{1/x} = y^{1/y}$, we see that we must have $x < e$ and $y > e$. The only two integers less that e are 1 and 2, and we cannot have $x = 1$ because $y^{1/y} > 1$ for $y > 1$, thus $x = 2$ is the only possibility. Testing integers $y = 3, 4, \ldots$ we find that $4^{1/4} = 2^{1/2}$ (that is, $j(2) = 4$), and we have the answer: $2^4 = 4^2$ is the only solution. If we try to solve this problem as an arithmetic problem about integers, we soon find ourselves sinking into the quicksand of comparing the counts of primes appearing as factors in x^y and y^x,[3] or doing something similarly futile.

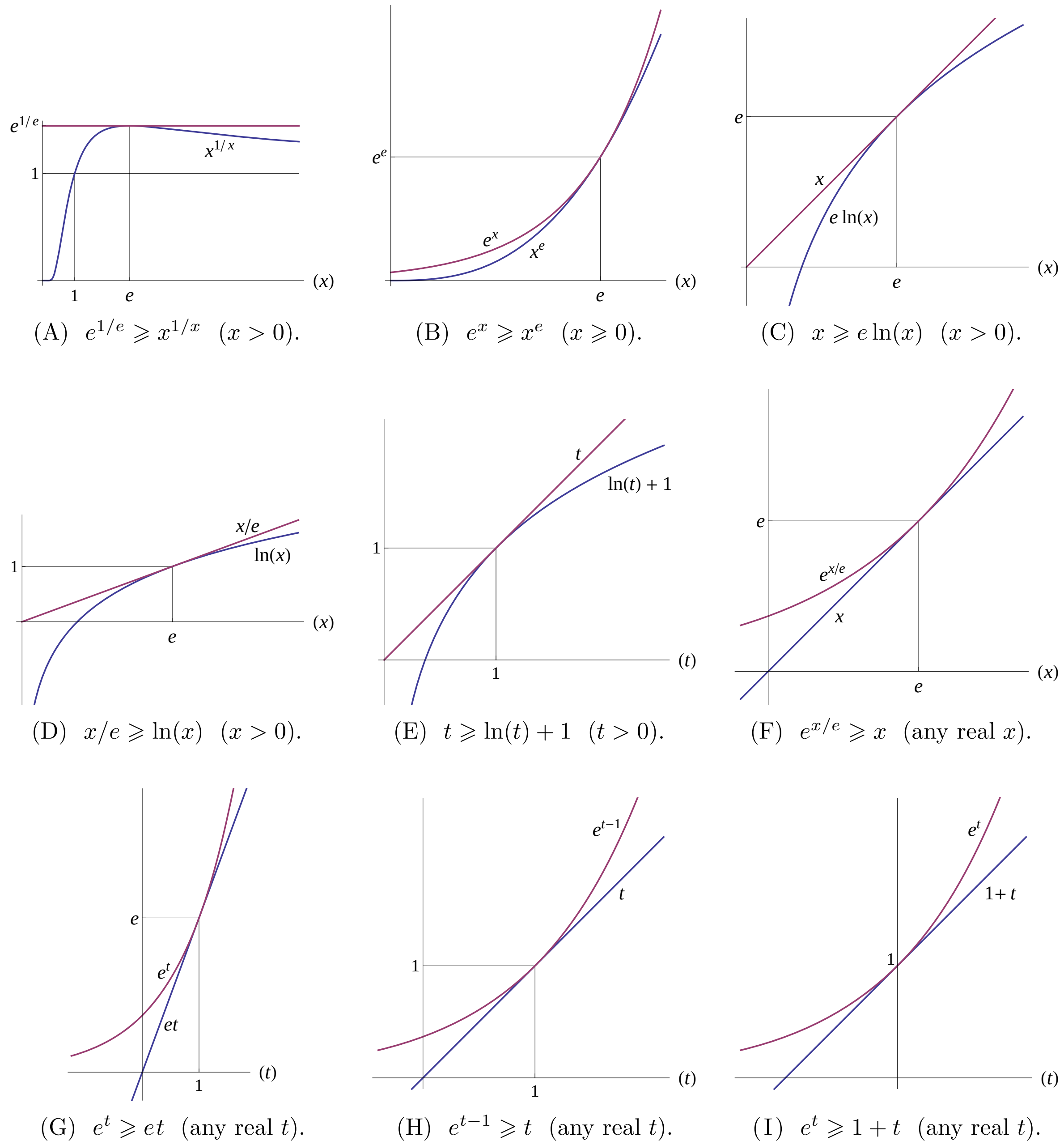
[3] Correction: there *is* a nice solution just along these lines. Here goes. We must have $x > 1$, since $x = 1$ in $x^y = y^x$ imply $y = 1$, contradicting $x < y$. If p is a prime, if p^e is the highest power of p dividing x, and if p^f is the highest power of p dividing y, then $ey = fx$, whence $f = (y/x)e \geqslant e$. It follows that y is divisible by x, say $y = xz$, where $z > 1$ is an integer. From $x^{xz} = (xz)^x$ we find, by extracting x-th roots and dividing by x, that $x^{z-1} = z$. Estimating $x^{z-1} = \big(1 + (x-1)\big)^{z-1}$ from below by the first two terms of the binomial expansion, we obtain $z = x^{z-1} \geqslant 1+(z-1)(x-1) \geqslant 1+(z-1) = z$, where both inequalities must be equalities. Since the first inequality is an equality if and only if $z = 2$ (otherwise $z > 2$ and there are further nonzero terms in the binomial expansion), and the second inequality is an equality if and only if $x = 2$, we conclude that $x = 2$, $y = 4$ is the only solution of $x^y = y^x$ in integers $0 < x < y$.

In Z. A. Melzak's "Companion to Concrete Mathematics", section 2.1 in Volume II opens with a nice short proof of the fact that $e^{1/e} \geqslant x^{1/x}$ for every $x > 0$, with equality iff $x = e$. Here is how it goes. Since $(e^t)'' = e^t > 0$ for all real t, the diagram of the exponential function e^t is strictly above any of its tangets except at the point of contact; in particular, considering the tangent at the point $(0,1)$, we get the inequality

$$e^t \geqslant 1 + t\,,$$

which holds for every real t, with equality iff $t = 0$.[4] Now, given any $x > 0$, take $t = x/e - 1$, multiply both sides of the inequality by e, then raise them to the power $1/x$.

Let us visualize possible sequences of simple steps that transform the inequality $e^t \geqslant 1+t$,



(A) $e^{1/e} \geqslant x^{1/x}$ $(x > 0)$. (B) $e^x \geqslant x^e$ $(x \geqslant 0)$. (C) $x \geqslant e\ln(x)$ $(x > 0)$.

(D) $x/e \geqslant \ln(x)$ $(x > 0)$. (E) $t \geqslant \ln(t) + 1$ $(t > 0)$. (F) $e^{x/e} \geqslant x$ (any real x).

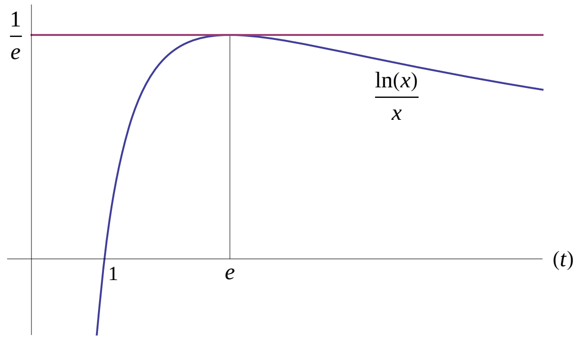
(G) $e^t \geqslant et$ (any real t). (H) $e^{t-1} \geqslant t$ (any real t). (I) $e^t \geqslant 1 + t$ (any real t).

Figure 1: Some morphs of the inequality $e^{1/e} \geqslant x^{1/x}$ (for $x > 0$, with equality iff $x = e$).

and some other inequalities, to the inequality $e^{1/e} \geqslant x^{1/x}$. The nine plates in Figure 1 display

[4] We can also prove this inequality as follows: put $f(t) = e^t - 1 - t$; then $f(0) = 0$, $f'(t) = e^t - 1 < 0$ for $t < 0$, and $f'(t) > 0$ for $t > 0$.

diagrams of nine 'morphs' of the inequality $e^{1/e} \geqslant x^{1/x}$, that is, of nine inequalities equivalent to it; Figure 2 has the diagram of one more morph, namely of the inequality $1/e \geqslant \ln(x)/x$,



(J) $1/e \geqslant \ln(x)/x \quad (x > 0)$

Figure 2: One more morph of the inequality $e^{1/e} \geqslant x^{1/x}$.

that we obtained by analyzing the behaviour of the function $(\ln x)/x$. In Figure 3 we have another sort of diagram, which represents simple transitions between the ten inequalities.

$$
\begin{tikzcd}
(\mathrm{I}) \arrow[r, "t \mapsto t-1"] & (\mathrm{H}) \arrow[r, "e \cdot (-)"] \arrow[d, leftrightarrow, "inv"'] & (\mathrm{G}) \arrow[r, "t \mapsto x/e"] \arrow[d, leftrightarrow, "inv"'] & (\mathrm{F}) \arrow[r, "(-)^{1/x}"] \arrow[d, leftrightarrow, "inv"'] \arrow[dr, "(-)^e"] & (\mathrm{A}) \arrow[r, "\ln(-)"] & (\mathrm{J}) \\
& (\mathrm{E}) \arrow[r, "t \mapsto x/e"] & (\mathrm{D}) \arrow[r, "e \cdot (-)"] & (\mathrm{C}) \arrow[r, "e^{(-)}"] & (\mathrm{B}) \arrow[u, "(-)^{1/(ex)}"'] &
\end{tikzcd}
$$

Figure 3: Metamorphoses of the inequality (A) $e^{1/e} \geqslant x^{1/x}$.

Some transitions are omitted from this diagram, say (J) $\xrightarrow{xe\cdot(-)}$ (C) and (J) $\xrightarrow{x\cdot(-)}$ (D), also all the transitions that are inverses of the shown one-way transitions. A two-way arrows labeled by *inv* represents a transition from an inequality between two strictly increasing (continuous) functions to the opposite inequality between their inverses; having the diagram of one of the two inequalities that are related in this manner, we obtain the diagram of the other inequality by exchanging the coordinate axes.

The (not so) amazing identity

In section 6.2 of WLiB (page 206), the derivation of the fact that the factor in parentheses on the right hand side of the equality

$$T_\theta \;=\; T_D \left(\sqrt{2}\, \frac{\sin\left(45^\circ - \tfrac{1}{2}\theta\right)}{\sqrt{1 - \sin\theta}} \right)$$

is 1, for every θ in the range $0 \leqslant \theta < 90^\circ$, takes up half a page. It can be considerably shortened by observing that

$$1 - \sin\theta \;=\; 1 - \cos(90^\circ - \theta) \;=\; 2\sin^2\left(45^\circ - \tfrac{1}{2}\theta\right).$$