

When Cauchy and Hölder Met Minkowski: A Tour through Well-Known Inequalities

GERHARD J. WOEGINGER

Department of Mathematics

TU Eindhoven

The Netherlands

gwoegi@win.tue.nl

Many classical inequalities are just statements about the convexity or concavity of certain (hidden) underlying functions. This is nicely illustrated by Hardy, Littlewood, and Pólya [5] whose Chapter III deals with “*Mean values with an arbitrary function and the theory of convex functions*,” and by Steele [12] whose Chapter 6 is called “*Convexity—The third pillar*.” Yet another illustration is the following proof of the arithmetic-mean-geometric-mean inequality (which goes back to Pólya). The inequality states that the arithmetic mean of n positive real numbers a_1, \dots, a_n is always greater or equal to their geometric mean:

$$\frac{1}{n} \cdot \sum_{i=1}^n a_i \geq \left(\prod_{i=1}^n a_i \right)^{1/n}. \quad (1)$$

The substitution $x_i = \ln a_i$ shows that (1) is equivalent to the inequality

$$\frac{1}{n} \cdot \sum_{i=1}^n e^{x_i} \geq e^{\frac{1}{n} \sum_{i=1}^n x_i}. \quad (2)$$

The correctness of (2) is easily seen from the following two observations. First: $f(x) = e^x$ is a convex function. And second: Jensen’s inequality [7] states that any convex function $f : \mathbb{R} \rightarrow \mathbb{R}$ and any real numbers x_1, \dots, x_n satisfy

$$\frac{1}{n} \cdot \sum_{i=1}^n f(x_i) \geq f\left(\frac{1}{n} \sum_{i=1}^n x_i\right). \quad (3)$$

But we do not want to give the impression that this article is centered around *convexity* and that it perhaps deals with Jensen’s inequality. No, no, no, quite to the contrary: This article is centered around *concavity*, and it deals with the Cauchy inequality, the Hölder inequality, the Minkowski inequality, and with Milne’s inequality. We present simple, concise, and uniform proofs for these four classical inequalities. All our proofs proceed in exactly the same fashion, by exactly the same type of argument, and they all follow from the concavity of a certain underlying function in exactly the same way. Loosely speaking, we shall see that

Cauchy corresponds to the concave function \sqrt{x} ,

Hölder corresponds to the concave function $x^{1/p}$ with $p > 1$,

Minkowski to the concave function $(x^{1/p} + 1)^p$ with $p > 1$, and

Milne corresponds to the concave function $x/(1+x)$.

Interestingly, the cases of equality for all four inequalities fall out from our discussion in a very natural way and come almost for free. Now let us set the stage for concavity and explain the general approach.

Concavity and the master theorem

Here are some very basic definitions. Throughout we use \mathbb{R} and \mathbb{R}_+ to denote the set of real numbers and the set of positive real numbers, respectively. A function $g : \mathbb{R}_+ \rightarrow \mathbb{R}$ is *concave* if it satisfies

$$\lambda \cdot g(x) + (1 - \lambda) \cdot g(y) \leq g(\lambda x + (1 - \lambda)y) \tag{4}$$

for all $x, y \in \mathbb{R}_+$ and for all real λ with $0 < \lambda < 1$. In other words, for any x and y the line segment connecting point $(x, g(x))$ to the point $(y, g(y))$ must lie below the graph of function g ; FIGURE 1 illustrates this. A concave function g is *strictly concave*, if equality in (4) is equivalent to $x = y$. A function g is *convex* (*strictly convex*) if the function $-g$ is concave (strictly concave). For twice-differentiable functions g there are simple criteria for checking these properties: A twice-differentiable function g is concave (strictly concave, convex, strictly convex) if and only if its second derivative is nonpositive (negative, nonnegative, positive) everywhere.

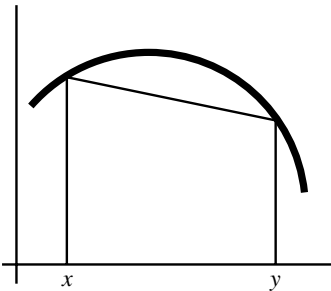


Figure 1 A concave function

Most of our arguments are based on the following theorem which we dub *the master theorem* (although admittedly, it rather is a simple observation on concavity, whose proof is only slightly longer than its statement). We would guess that the statement was known already before the Second World War, but its exact origin is unknown to us. Walther Janous pointed out to us that Godunova [4] used the idea in 1967.

MASTER THEOREM. *Let $g : \mathbb{R}_+ \rightarrow \mathbb{R}$ be a strictly concave function, and let $f : \mathbb{R}_+^2 \rightarrow \mathbb{R}$ be the function defined by*

$$f(x, y) = y \cdot g\left(\frac{x}{y}\right). \tag{5}$$

Then all positive real numbers x_1, \dots, x_n and y_1, \dots, y_n satisfy the inequality

$$\sum_{i=1}^n f(x_i, y_i) \leq f\left(\sum_{i=1}^n x_i, \sum_{i=1}^n y_i\right). \tag{6}$$

Equality holds in (6) if and only if the two sequences x_i and y_i are proportional (that is, if and only if there is a real number t such that $x_i/y_i = t$ for all i).

Proof. The proof is by induction on n . For $n = 1$, the inequality (6) becomes an equation. Since the two sequences have length one, they are trivially proportional. For $n = 2$, we use the concavity of g : From (4) with $\lambda = y_1/(y_1 + y_2)$, we derive that

$$\begin{aligned}
f(x_1, y_1) + f(x_2, y_2) &= y_1 \cdot g\left(\frac{x_1}{y_1}\right) + y_2 \cdot g\left(\frac{x_2}{y_2}\right) \\
&= (y_1 + y_2) \left\{ \frac{y_1}{y_1 + y_2} \cdot g\left(\frac{x_1}{y_1}\right) + \frac{y_2}{y_1 + y_2} \cdot g\left(\frac{x_2}{y_2}\right) \right\} \\
&\leq (y_1 + y_2) \cdot g\left(\frac{x_1 + x_2}{y_1 + y_2}\right) \\
&= f(x_1 + x_2, y_1 + y_2).
\end{aligned} \tag{7}$$

Since g is strictly concave, equality holds in this chain if and only if $x_1/y_1 = x_2/y_2$. The inductive step for $n \geq 3$ follows easily from (7), and the proof is complete. ■

Here are two brief remarks before we proceed. First, if the function g in the theorem is just concave (but not strictly concave), then inequality (6) is still valid, but we lose control over the situation where equality holds. The cases of equality are no longer limited to proportional sequences, and can be quite arbitrary. Second, if g is strictly convex (instead of strictly concave), then the inequality (6) follows with a greater-or-equal sign instead of a less-or-equal sign.

Our next goal is to derive four well-known inequalities by four applications of the master theorem with four appropriately chosen strictly concave functions. As a propaedeutic exercise the reader should recall that the functions \sqrt{x} and $x/(1+x)$ are strictly concave. Furthermore, for any fixed real $p > 1$ the functions $x^{1/p}$ and $(x^{1/p} + 1)^p$ are strictly concave. Throughout a_1, \dots, a_n and b_1, \dots, b_n will denote sequences of positive real numbers.

Cauchy Augustin-Louis Cauchy [3] published his famous inequality in 1821. Then in 1859, Viktor Yakovlevich Bunyakovsky [2] derived a corresponding inequality for integrals, and in 1885 Hermann Schwarz [11] proved a corresponding version for inner-product spaces. Therefore the Cauchy inequality sometimes also shows up under the name Schwarz inequality, or Cauchy-Schwarz inequality, or Cauchy-Bunyakovsky-Schwarz inequality. In any case, its discrete version states that

$$\sum_{i=1}^n a_i b_i \leq \sqrt{\sum_{i=1}^n a_i^2} \cdot \sqrt{\sum_{i=1}^n b_i^2}. \tag{8}$$

Cauchy's original proof of (8) rewrites it into the equivalent and obviously true

$$0 \leq \sum_{1 \leq i < j \leq n} (a_i b_j - a_j b_i)^2.$$

We give another very short proof of (8) by deducing it from the master theorem: We use the strictly concave function $g(x) = \sqrt{x}$, which yields $f(x, y) = \sqrt{x} \sqrt{y}$. Then (6) turns into

$$\sum_{i=1}^n \sqrt{x_i} \sqrt{y_i} \leq \sqrt{\sum_{i=1}^n x_i} \cdot \sqrt{\sum_{i=1}^n y_i}.$$

Finally, setting $x_i = a_i^2$ and $y_i = b_i^2$ for $1 \leq i \leq n$ yields the Cauchy inequality (8). Furthermore equality holds in (8) if and only if the a_i and the b_i are proportional.

Hölder We turn to the Hölder inequality, which was first derived in 1888 by Leonard James Rogers [10], and then in 1889 in a different way by Otto Ludwig Hölder [6].

This inequality is built around two real numbers $p, q > 1$ with $1/p + 1/q = 1$. It states that

$$\sum_{i=1}^n a_i b_i \leq \left(\sum_{i=1}^n a_i^p \right)^{1/p} \left(\sum_{i=1}^n b_i^q \right)^{1/q} . \tag{9}$$

Note that the Cauchy inequality is the special case of the Hölder inequality with $p = q = 2$. One standard proof of (9) is based on Young’s inequality, which gives $xy \leq x^p/p + y^q/q$ for all real $x, y > 0$ and for all real $p, q > 1$ with $1/p + 1/q = 1$.

But let us deduce the Hölder inequality from the master theorem. We set $g(x) = x^{1/p}$, which is strictly concave if $p > 1$. Then the corresponding function f is given by $f(x, y) = x^{1/p} y^{1/q}$, and inequality (6) becomes

$$\sum_{i=1}^n x_i^{1/p} y_i^{1/q} \leq \left(\sum_{i=1}^n x_i \right)^{1/p} \left(\sum_{i=1}^n y_i \right)^{1/q} .$$

We substitute $x_i = a_i^p$ and $y_i = b_i^q$ for $1 \leq i \leq n$, and get the Hölder inequality (9). Clearly, equality holds in (9) if and only if the a_i^p and the b_i^q are proportional.

Minkowski The Minkowski inequality was established in 1896 by Hermann Minkowski [9] in his book *Geometrie der Zahlen* (Geometry of Numbers). It uses a real parameter $p > 1$, and states that

$$\left(\sum_{i=1}^n (a_i + b_i)^p \right)^{1/p} \leq \left(\sum_{i=1}^n a_i^p \right)^{1/p} + \left(\sum_{i=1}^n b_i^p \right)^{1/p} . \tag{10}$$

The special case of (10) with $p = 2$ is the triangle inequality $\|a + b\|_2 \leq \|a\|_2 + \|b\|_2$ in Euclidean spaces. Once again we exhibit a very short proof via the master theorem. We choose $g(x) = (x^{1/p} + 1)^p$. Since $p > 1$, this function g is strictly concave. The corresponding function f is given by $f(x, y) = (x^{1/p} + y^{1/p})^p$. Then the inequality in (6) becomes

$$\sum_{i=1}^n (x_i^{1/p} + y_i^{1/p})^p \leq \left(\left(\sum_{i=1}^n x_i \right)^{1/p} + \left(\sum_{i=1}^n y_i \right)^{1/p} \right)^p .$$

By setting $x_i = a_i^p$ and $y_i = b_i^p$ for $1 \leq i \leq n$ and by taking the p th root on both sides, we produce the Minkowski inequality (10). Furthermore equality holds in (10), if and only if the a_i^p and the b_i^p are proportional, which happens if and only if the a_i and the b_i are proportional.

Milne In 1925 Milne [8] used the following inequality (11) to analyze the biases inherent in certain measurements of stellar radiation:

$$\left(\sum_{i=1}^n (a_i + b_i) \right) \left(\sum_{i=1}^n \frac{a_i b_i}{a_i + b_i} \right) \leq \left(\sum_{i=1}^n a_i \right) \left(\sum_{i=1}^n b_i \right) . \tag{11}$$

Milne’s inequality is fairly well known, but of course the inequalities of Cauchy, Hölder, and Minkowski are in a completely different league—both in terms of relevance and in terms of publicity. Milne’s inequality is also discussed on page 61 of Hardy, Littlewood, and Pólya [5]. The problem corner in *Crux Mathematicorum* [1] lists three simple proofs that are due to Ardila, to Lau, and to Murty, respectively.

Murty’s proof is particularly simple and rewrites (11) into the equivalent

$$0 \leq \sum_{1 \leq i < j \leq n} \frac{(a_i b_j - a_j b_i)^2}{(a_i + b_i)(a_j + b_j)}.$$

And here is our proof: This time we use the strictly concave function $g(x) = x/(1 + x)$, which yields $f(x, y) = xy/(x + y)$. The resulting version of (6) yields

$$\sum_{i=1}^n \frac{a_i b_i}{a_i + b_i} \leq \left(\sum_{i=1}^n a_i \right) \left(\sum_{i=1}^n b_i \right) / \left(\sum_{i=1}^n a_i + \sum_{i=1}^n b_i \right),$$

which is equivalent to (11). Once again equality holds if and only if the a_i and the b_i are proportional.

A generalization of the master theorem

We now generalize the master theorem to higher dimensions. This is a fairly easy enterprise, since all concepts and arguments for the higher-dimensional case run perfectly in parallel to the one-dimensional case. For instance, a function $g : \mathbb{R}_+^d \rightarrow \mathbb{R}$ is *concave* if it satisfies

$$\lambda \cdot g(\vec{x}) + (1 - \lambda) \cdot g(\vec{y}) \leq g(\lambda \vec{x} + (1 - \lambda) \vec{y}) \tag{12}$$

for all $\vec{x}, \vec{y} \in \mathbb{R}_+^d$ and for all real numbers λ with $0 < \lambda < 1$. A concave function g is *strictly concave*, if equality in (12) is equivalent to $\vec{x} = \vec{y}$. It is known that a twice-differentiable function g is concave (strictly concave) if and only if its Hessian matrix is negative semidefinite (negative definite) for all $\vec{x} \in \mathbb{R}_+^d$.

Here is the higher-dimensional version of the master theorem. Note that by setting $d = 2$ in the new theorem we recover the master theorem.

HIGHER-DIMENSIONAL MASTER THEOREM. *Let $d \geq 2$ be an integer, and let $g : \mathbb{R}_+^{d-1} \rightarrow \mathbb{R}$ be a strictly concave function. Let $f : \mathbb{R}_+^d \rightarrow \mathbb{R}$ be the function defined by*

$$f(x_1, x_2, \dots, x_d) = x_d \cdot g\left(\frac{x_1}{x_d}, \frac{x_2}{x_d}, \dots, \frac{x_{d-1}}{x_d}\right). \tag{13}$$

Then any $n \times d$ matrix $Z = (z_{i,j})$ with positive real entries satisfies the inequality

$$\sum_{i=1}^n f(z_{i,1}, z_{i,2}, \dots, z_{i,d}) \leq f\left(\sum_{i=1}^n z_{i,1}, \sum_{i=1}^n z_{i,2}, \dots, \sum_{i=1}^n z_{i,d}\right). \tag{14}$$

Equality holds in (14) if and only if matrix Z has rank 1 (that is, if and only if there exist real numbers s_1, \dots, s_n and t_1, \dots, t_d such that $z_{i,j} = s_i t_j$ for all i, j).

Proof. The proof closely follows the proof of the master theorem. As in (7), we observe that all positive real numbers a_1, \dots, a_d and b_1, \dots, b_d satisfy

$$f(a_1, \dots, a_d) + f(b_1, \dots, b_d) \leq f(a_1 + b_1, a_2 + b_2, \dots, a_d + b_d).$$

Equality holds if and only if the a_i and the b_i are proportional. Then an inductive argument based on this observation yields the statement in the theorem, and completes the proof. ■

We conclude this article by posing two exercises to the reader that both can be settled through the higher-dimensional master theorem. Each exercise deals with inequalities for three sequences a_1, \dots, a_n , b_1, \dots, b_n , and c_1, \dots, c_n of positive real numbers.

Generalized Hölder The first exercise concerns the generalized Hölder inequality, which is built around three real numbers $p, q, r > 1$ with $1/p + 1/q + 1/r = 1$. It states that

$$\sum_{i=1}^n a_i b_i c_i \leq \left(\sum_{i=1}^n a_i^p \right)^{1/p} \left(\sum_{i=1}^n b_i^q \right)^{1/q} \left(\sum_{i=1}^n c_i^r \right)^{1/r}. \quad (15)$$

The reader is asked to deduce inequality (15) from the higher-dimensional master theorem (perhaps by using the function $g(x, y) = x^{1/p} y^{1/q}$), and to identify the cases of equality.

Generalized Milne Problem #68 on page 62 of Hardy, Littlewood, and Pólya [5] concerns the following generalization of Milne's inequality (11) to three sequences.

$$\begin{aligned} & \left(\sum_{i=1}^n a_i \right) \left(\sum_{i=1}^n b_i \right) \left(\sum_{i=1}^n c_i \right) \\ & \geq \left(\sum_{i=1}^n (a_i + b_i + c_i) \right) \left(\sum_{i=1}^n \frac{a_i b_i + b_i c_i + a_i c_i}{a_i + b_i + c_i} \right) \left(\sum_{i=1}^n \frac{a_i b_i c_i}{a_i b_i + b_i c_i + a_i c_i} \right) \end{aligned}$$

We ask the reader to deduce it from the higher-dimensional master theorem, and to describe the cases of equality. One possible proof goes through two steps, where the first step uses $g(x, y) = xy/(xy + x + y)$, and the second step uses the function $g(x, y) = (xy + x + y)/(x + y + 1)$.

REFERENCES

1. F. Ardila, K. W. Lau, and V. N. Murty, Solution to problem 2113, *Crux Mathematicorum* **23** (1997) 112–114.
2. V. Y. Bunyakovsky, Sur quelques inégalités concernant les intégrales aux différences finies, *Mémoires de l'Académie impériale des sciences de St.-Petersbourg* **1**(9) (1859), 4.
3. A. L. Cauchy, *Cours d'Analyse de l'École Royale Polytechnique, Première Partie, Analyse Algébrique*, De-bure frères, Paris, 1821.
4. E. K. Godunova, Convexity of complex functions and its use in proving inequalities (in Russian), *Matematicheskie Zametki* **1** (1967) 495–500. English translation in *Mathematical Notes* **1** (1967) 326–329.
5. G. H. Hardy, J. E. Littlewood, and G. Pólya, *Inequalities*, Cambridge University Press, Cambridge, 1934.
6. O. L. Hölder, Über einen Mittelwerthsatz, *Nachrichten von der Königl. Gesellschaft der Wissenschaften und der Georg-Augusts-Universität zu Göttingen* (1889) 38–47.
7. J. L. W. V. Jensen, Sur les fonctions convexes et les inégalités entre les valeurs moyennes, *Acta Mathematica* **30** (1906) 175–193.
8. E. A. Milne, Note on Rosseland's integral for the stellar absorption coefficient, *Monthly Notices of the Royal Astronomical Society* **85** (1925) 979–984.
9. H. Minkowski, *Geometrie der Zahlen*, Teubner, Leipzig, 1896.
10. L. J. Rogers, An extension of a certain theorem in inequalities, *Messenger of Mathematics* **17** (1888) 145–150.
11. H. A. Schwarz, Über ein Flächen kleinsten Flächeninhalts betreffendes Problem der Variationsrechnung, *Acta Societatis Scientiarum Fennicae* **15** (1885) 315–362.
12. J. M. Steele, *The Cauchy-Schwarz Master Class: An Introduction to the Art of Mathematical Inequalities*, Cambridge University Press, Cambridge, 2004.