Maclaurin's Inequality and a Generalized Bernoulli Inequality

Iddo Ben-Ari
University of Connecticut
196 Auditorium Rd
Storrs, CT 06269-3009
iddo.ben-ari@uconn.edu

Keith Conrad
University of Connecticut
196 Auditorium Rd
Storrs, CT 06269-3009
kconrad@math.uconn.edu

Introduction

One of the most famous inequalities in mathematics is the *arithmetic-geometric mean inequality*: for every positive integer n and $x_1, \ldots, x_n > 0$,

$$\frac{x_1 + x_2 + \cdots + x_n}{n} \geq \sqrt[n]{x_1 x_2 \cdots x_n}, \tag{1}$$

and the inequality is strict unless the x_i's are all equal. Did you know there is an extension of (1) that interpolates terms between the average on the left and the nth root on the right? It was first stated by Maclaurin in 1729 [7, pp. 80–81], but remains relatively unknown outside of aficionados of inequalities. By comparison, even students who are not active users of inequalities will know (or should know!) the arithmetic-geometric mean inequality.

To interpolate terms in (1), we need to use the *elementary symmetric polynomials* in $x_1, \ldots, x_n$, which are

$$e_k(x_1, \ldots, x_n) = \sum_{1 \leq i_1 < i_2 < \cdots < i_k \leq n} x_{i_1} x_{i_2} \cdots x_{i_k} = \sum_{\substack{I \subset \{1,\ldots,n\} \\ \#I = k}} \prod_{i \in I} x_i$$

for $1 \leq k \leq n$. For instance, when $n = 3$

$$e_1(x, y, z) = x + y + z, \quad e_2(x, y, z) = xy + xz + yz, \quad e_3(x, y, z) = xyz.$$

In general $e_1(x_1, \ldots, x_n) = x_1 + \cdots + x_n$ and $e_n(x_1, \ldots, x_n) = x_1 \cdots x_n$, so the elementary symmetric polynomials interpolate between the sum of n numbers and the product of n numbers. These polynomials naturally arise as the coefficients of the polynomial whose roots are $x_1, x_2, \ldots, x_n$:

$$(T - x_1)(T - x_2) \cdots (T - x_n) = T^n - e_1 T^{n-1} + e_2 T^{n-2} - \cdots + (-1)^n e_n.$$

Each $e_k(x_1, \ldots, x_n)$ is a sum of $\binom{n}{k}$ terms, and its average

$$E_k(x_1, \ldots, x_n) := \frac{e_k(x_1, \ldots, x_n)}{e_k(1, \ldots, 1)} = \frac{e_k(x_1, \ldots, x_n)}{\binom{n}{k}}$$

is called the kth *elementary symmetric mean* of $x_1, \ldots, x_n$. When $n = 3$,

$$E_1(x, y, z) = \frac{x + y + z}{3}, \quad E_2(x, y, z) = \frac{xy + xz + yz}{3}, \quad E_3(x, y, z) = xyz.$$

Now we can state Maclaurin's inequality: for positive $x_1, \ldots, x_n$,

$$\boxed{\frac{x_1 + \cdots + x_n}{n} \geq \sqrt{\frac{\sum_{1 \leq i < j \leq n} x_i x_j}{\binom{n}{2}}} \geq \sqrt[3]{\frac{\sum_{1 \leq i < j < k \leq n} x_i x_j x_k}{\binom{n}{3}}} \geq \cdots \geq \sqrt[n]{x_1 x_2 \cdots x_n},}$$

or equivalently

$$E_1(x_1, \ldots, x_n) \geq \sqrt{E_2(x_1, \ldots, x_n)} \geq \sqrt[3]{E_3(x_1, \ldots, x_n)} \geq \cdots \geq \sqrt[n]{E_n(x_1, \ldots, x_n)}. \tag{2}$$

Moreover, the inequalities are all strict unless the x_i's are all equal. For example, when $n = 3$ Maclaurin's inequality says for positive x, y, and z that

$$\frac{x + y + z}{3} \geq \sqrt{\frac{xy + xz + yz}{3}} \geq \sqrt[3]{xyz}$$

and both inequalities are strict unless $x = y = z$.

The arithmetic-geometric mean inequality is a consequence of Maclaurin's inequality (look at the first and last terms), and these two inequalities are linked historically: the paper in which Maclaurin stated his inequality is also where the arithmetic-geometric mean inequality for n terms, not just 2 terms, first appeared [7, pp. 78–79].

In mathematics there are many "named" inequalities, such as the Cauchy–Schwarz inequality (in linear algebra), Chebyshev's inequality (in probability), Hölder's inequality (in real analysis), and Maclaurin's inequality. Recently Maligranda [9] (see also [8, Theorem 3]) showed the arithmetic-geometric mean inequality is equivalent to another named inequality, *Bernoulli's inequality*:

$$(1+t)^n \geq 1 + nt \tag{3}$$

for every positive integer n and real number $t > -1$, with the inequality strict for $n > 1$ unless $t = 0$. Since the arithmetic-geometric mean inequality is interpolated by Maclaurin's inequality, it's natural to wonder if there is an interpolated form of Bernoulli's inequality that would fill in the diagram below.

$$\begin{array}{rcl} \text{Arithmetic-Geometric Mean Inequality} & \iff & \text{Bernoulli's Inequality} \\ \text{Maclaurin's Inequality} & \iff & ??? \end{array}$$

One benefit of finding an interpolated Bernoulli's inequality is that it will lead to a new proof of Maclaurin's inequality. Before we go in that direction, though, we want to develop two reasons you should care about Maclaurin's inequality in case its statement alone is not immediately attractive: an open problem about recursive sequences and a probabilistic interpretation.

First Application: Convergence of a Recursive Sequence

The most interesting (to us) application of Maclaurin's inequality is to a recursion in n variables that generalizes Gauss's arithmetic-geometric mean recursion in 2 variables.

For a pair of positive numbers x and y, define the sequence of pairs (x_j, y_j) recursively by $x_0 = x$, $y_0 = y$, and

$$x_{j+1} = \frac{x_j + y_j}{2}, \quad y_{j+1} = \sqrt{x_j y_j}. \tag{4}$$

Example 1. If $x_0 = 1$ and $y_0 = 2$, Table 1 shows the first few values of x_j and y_j to 16 digits after the decimal point.

j	x_j	y_j
0	1	2
1	1.5	1.4142135623730950
2	1.4571067811865475	1.4564753151219702
3	1.4567910481542588	1.4567910139395549
4	1.4567910310469069	1.4567910310469068

Table 1: Iterating arithmetic and geometric means.

Example 1 illustrates a general phenomenon: for all choices of x and y, the sequences $\{x_j\}$ and $\{y_j\}$ converge and their limits are the same. Gauss called the common limit of the sequences $\{x_j\}$ and $\{y_j\}$ produced from (4) the *arithmetic-geometric mean* of x and y, denoted $M(x, y)$. By Table 1, it looks like $M(1, 2) \approx 1.456791031046906$.

To establish the existence of $M(x, y)$, we will use the arithmetic-geometric mean inequality for two terms, which tells us $x_j \geq y_j$ for all $j \geq 1$. Feeding this into (4), we have $x_{j+1} \leq x_j$ and $y_{j+1} \geq y_j$ for $j \geq 1$. Hence

$$x_1 \geq x_2 \geq \cdots \geq x_j \geq \cdots \geq y_j \geq \cdots \geq y_2 \geq y_1.$$

Since $\{x_j\}_{j\geq 1}$ is decreasing and bounded below (by y_1) and $\{y_j\}_{j\geq 1}$ is increasing and bounded above (by x_1), both $\{x_j\}$ and $\{y_j\}$ converge. Call the limits X and Y, so $x_1 \geq X \geq Y \geq y_1$. Letting $j \to \infty$ in (4), we get $X = (X + Y)/2$ and $Y = \sqrt{XY}$. Either of these equations implies $X = Y$.

From numerical calculations, Gauss discovered that up to 11 decimal digits

$$M(1, \sqrt{2}) = \frac{\pi}{2\int_0^1 du/\sqrt{1-u^4}}.$$

He then proved the general formula

$$\frac{1}{M(x,y)} = \frac{2}{\pi}\int_0^{\pi/2} \frac{dt}{\sqrt{x^2\cos^2 t + y^2 \sin^2 t}}, \tag{5}$$

where the integrand is not symmetric in x and y even though $M(x,y)$, by its definition, is symmetric. Under the change of variables $u = y\tan t$,

$$\frac{1}{M(x,y)} = \frac{2}{\pi}\int_0^\infty \frac{du}{\sqrt{(u^2+x^2)(u^2+y^2)}}, \tag{6}$$

where the integrand is now symmetric in x and y. The significance of $M(x,y)$ for 19th century analysis is described in [1], [3], and [4], where the proofs of the convergence of $\{x_j\}$ and $\{y_j\}$ to $M(x,y)$ show that it is very rapid, as we saw in Table 1.

Using elementary symmetric means we can generalize the recursion (4) from two numbers to n numbers: for $x_1,\ldots,x_n > 0$, define n-tuples $\{(x_{1,j}, x_{2,j},\ldots,x_{n,j})\}$ for $j \geq 0$ by

$$x_{k,0} = x_k \quad\text{and}\quad x_{k,j} = \sqrt[k]{E_k(x_{1,j-1},\ldots,x_{n,j-1})} \quad\text{for}\quad j \geq 1. \tag{7}$$

Example 2. Let $x_1 = 1$, $x_2 = 2$, and $x_3 = 3$. Table 2 lists the first few iterations to 16 digits after the decimal point. Although $x_{1,0} < x_{2,0} < x_{3,0}$, we have $x_{1,j} > x_{2,j} > x_{3,j}$ for $j > 0$.

j	$x_{1,j}$	$x_{2,j}$	$x_{3,j}$
0	1	2	3
1	2	1.9148542155126762	1.8171205928321396
2	1.9106582694482719	1.9099276289927102	1.9091929427097283
3	1.9099262803835701	1.9099262335408387	1.9099261866980376
4	1.9099262335408155	1.9099262335408153	1.9099262335408151

Table 2: An iteration of E_1, $\sqrt{E_2}$, and $\sqrt[3]{E_3}$ on three numbers.

Example 3. When $x_1 = 1$, $x_2 = 2$, and $x_3 = 5$, the initial iterations are in Table 3.

From these two examples, it will be no surprise that, for all positive numbers $x_1,\ldots,x_n$, the n sequences sequences $x_{k,0}, x_{k,1}, x_{k,2},\ldots$ in (7), for $1 \leq k \leq n$, all converge and have the same limit. To demonstrate this, we first observe from Maclaurin's inequality that $x_{1,j} \geq x_{2,j} \geq \cdots \geq x_{n,j}$ for all $j \geq 1$ (perhaps not at $j = 0$). Therefore

j	$x_{1,j}$	$x_{2,j}$	$x_{3,j}$
0	1	2	5
1	2.6666666666666666	2.3804761428476166	2.1544346900318837
2	2.4005258331820556	2.3959462942846843	2.3914169307949695
3	2.3959630194205698	2.3959615764905558	2.3959601335780796
4	2.3959615764964017	2.3959615764962569	2.3959615764961121

Table 3: Another iteration of E_1, $\sqrt{E_2}$, and $\sqrt[3]{E_3}$ on three numbers.

it suffices to prove the outer sequences $\{x_{1,j}\}$ and $\{x_{n,j}\}$ converge and have a common limit. Our argument will be based on W. Sawin's proof on the web page [10].

Using Maclaurin's inequality again, for $j \geq 1$

$$x_{1,j+1} = E_1(x_{1,j}, x_{2,j}, \ldots, x_{n,j}) \leq E_1(x_{1,j}, x_{1,j}, \ldots, x_{1,j}) = x_{1,j}$$

and

$$x_{n,j+1} = \sqrt[n]{E_n(x_{1,j}, x_{2,j}, \ldots, x_{n,j})} \geq \sqrt[n]{E_n(x_{n,j}, x_{n,j}, \ldots, x_{n,j})} = \sqrt[n]{x_{n,j}^n} = x_{n,j},$$

so the sequences $\{x_{1,j}\}$ and $\{x_{n,j}\}$ for $j \geq 1$ satisfy

$$x_{1,1} \geq x_{1,2} \geq \cdots \geq x_{1,j} \geq \cdots \geq x_{n,j} \geq \cdots \geq x_{n,2} \geq x_{n,1}.$$

Therefore the sequences $\{x_{1,j}\}$ and $\{x_{n,j}\}$ each converge. Call the respective limits X_1 and X_n, so $X_1 \geq X_n$. To prove the reverse inequality, for $j \geq 1$ we have

$$x_{1,j+1} = \frac{1}{n}(x_{1,j} + \cdots + x_{n,j}) \leq \frac{1}{n}((n-1)x_{1,j} + x_{n,j}), \tag{8}$$

and letting $j \to \infty$ in (8) gives us $X_1 \leq \frac{1}{n}((n-1)X_1 + X_n)$, so $X_1 \leq X_n$. Thus $X_1 = X_n$. This proves all n sequences $x_{k,0}, x_{k,1}, x_{k,2}, \ldots$ converge to the same number.

For $x_1, \ldots, x_n > 0$ the common limit of the sequences $\{x_{k,0}, x_{k,1}, x_{k,2}, \ldots\}$, where $x_{k,0} = x_k$, is called the *symmetric mean* $M(x_1, \ldots, x_n)$. The name is reasonable since it is a symmetric function of the x_i's. Another property is $M(tx_1, tx_2, \ldots tx_n) = tM(x_1, \ldots, x_n)$ for all $t > 0$; this is called being homogeneous of degree 1.

Unlike (5), for $n \geq 3$ no general explicit formula for $M(x_1, \ldots, x_n)$ is known! The case $n = 3$ was first investigated by Meissel [13, Sect. 5] in the 19th century, although not in any way conclusively. A plausible guess at a formula for $M(x_1, \ldots, x_n)$, in an attempt to generalize (6), is

$$\frac{1}{M(x_1, \ldots, x_n)} \stackrel{?}{=} c \int_0^\infty \frac{u^{n-2} du}{\sqrt[r]{(u^r + x_1^r)(u^r + x_2^r) \cdots (u^r + x_n^r)}}$$

for some $c > 0$ and some integer $r \geq 2$ that would need to be determined. We place u^{n-2} in the numerator of the integral to make the right side homogeneous in $x_1, \ldots, x_n$ of degree -1, like the left side (*i.e.*, replacing x_i with tx_i on both sides has $1/t$ pulled out, using the change of variables $v = tu$ in the integral). The right side is symmetric in the x_i's, like the left side is. Since $M(1, 1, \ldots, 1) = 1$, c is determined from r by setting each x_i equal to 1:

$$1 = c \int_0^\infty \frac{u^{n-2} du}{(u^r + 1)^{n/r}}.$$

Alas, using the approximations for $M(1, 2, 3)$ and $M(1, 2, 5)$ from Examples 2 and 3, such a formula for $1/M(x, y, z)$ as an integral is wrong for $r = 2, 3$, and 4, and gets worse as r grows. Can you find a formula for $M(x_1, \ldots, x_n)$ when $n \geq 3$?

Second Application: Products of Random Variables

This section requires some familiarity with several basic notions in probability theory, in particular random variables and their expectation. Readers unfamiliar with these topics can find a treatment in any undergraduate textbook in probability, e.g. [14].

The first elementary symmetric mean $E_1(x_1, \ldots, x_n)$ is the average of $x_1, \ldots, x_n$. The other symmetric means $E_k(x_1, \ldots, x_n)$ are averages of k-fold products of the x_i's. This suggests there should be a role for Maclaurin's inequality in probability theory: we seek random variables $X_1, X_2, \ldots, X_n$ for which the expectation $\mathbb{E}(X_1 \cdots X_k)$ is the kth symmetric mean $E_k(x_1, \ldots, x_n)$.

Fix positive numbers $x_1, \ldots, x_n$ and consider an urn containing n balls, labeled by the x_i's. Suppose we select balls randomly from the urn, one after another, without replacement until all n balls are picked. Let X_j be the label of the jth ball that is sampled, so each X_j is a random variable with values in $\{x_1, \ldots, x_n\}$. The outcome of such sampling is a sequence of numbers $(X_1, \ldots, X_n)$. Because we sample without replacement, the value of X_j is affected by the values of $X_1, \ldots, X_{j-1}$, so the random variables $X_1, \ldots, X_n$ are not independent if the labels x_i are not all the same.

Example 4. If we have three balls, numbered as 1, 2, 3, they can be selected in 6 possible ways: 123, 132, 213, 231, 312, 321. If balls 1 and 2 have label x and ball 3 has label y, where $y \neq x$, then the labels we see when selecting the balls in all possible ways are xxy, xyx, xxy, xyx, yxx, yxx. In the first sampling, $X_1 = X_2 = x$ and $X_3 = y$. In the second sampling, $X_1 = X_3 = x$ and $X_2 = y$. Looking at how often x and y occur as a label for the first sampled ball, the second sampled ball, and the third sampled ball, we get x four times and y two times in each position, so X_1, X_2, and X_3 all have the same distribution: $\mathrm{Prob}(X_j = x) = 2/3$ and $\mathrm{Prob}(X_j = y) = 1/3$.

Since the actual selection of the balls one after another doesn't see the labels, when sampling the n balls without replacement and considering them as n distinct objects any of the $n!$ possible sequences are equally likely. Consequently, given a label x_i, and some $j \in \{1, \ldots, n\}$, the number of sequences of n balls in which the first ball selected has label x_i is the same as the number of sequences of n balls in which the jth ball selected has label x_i. Therefore, as Example 4 illustrates, the X_j's are identically distributed (with the hypergeometric distribution, or multivariate hypergeometric distribution if some x_i's are equal). For $1 \leq k \leq n$, $\mathbb{E}(X_1 \cdots X_k) = E_k(x_1, \ldots, x_n)$.

This urn model provides us with a probabilistic interpretation of Maclaurin's inequality. For the dependent random variables $X_1, \ldots, X_n$, Maclaurin's inequality is equivalent to

$$\mathbb{E}(X_1) \geq \sqrt{\mathbb{E}(X_1 X_2)} \geq \cdots \geq \sqrt[n]{\mathbb{E}(X_1 \cdots X_n)}. \tag{9}$$

To get a probabilistic feel for (9), it should be contrasted with the case of of n independent and identically distributed random variables $X_1, \ldots, X_n$ with positive values, for which $\mathbb{E}(X_1 \cdots X_k) = (\mathbb{E}(X_1))^k$, so

$$\mathbb{E}(X_1) = \sqrt{\mathbb{E}(X_1 X_2)} = \cdots = \sqrt[n]{\mathbb{E}(X_1 \cdots X_n)}.$$

And if X is a single random variable with positive values, its powers $X, X^2, \ldots, X^n$

are usually not independent or identically distributed and

$$\mathbb{E}(X) \leq \sqrt{\mathbb{E}(X^2)} \leq \cdots \leq \sqrt[n]{\mathbb{E}(X^n)}$$

by Jensen's inequality (another named inequality). This is the reverse of (9)!

Maclaurin's inequality also gives us information about the covariance of products of the X_j's. The *covariance* of two random variables X and Y, denoted $\mathrm{cov}(X,Y)$, is $\mathbb{E}\big((X-\mathbb{E}(X))(Y-\mathbb{E}(Y)\big) = \mathbb{E}(XY) - (\mathbb{E}X)(\mathbb{E}Y)$, where the equality follows from the linearity of the expectation. As its definition immediately suggests, $\mathrm{cov}(X,Y)$ is a mathematically-tractable measure of how the two random variables X and Y jointly deviate from their respective expectations. Positive covariance, also known as positive correlation, is intuitively the statement that X and Y tend to deviate from their respective expectations in similar patterns: "typically", when one random variable is above its expectation so is the other. Negative covariance, also known as negative correlation, corresponds to the intuitive statement that the random variables tend to deviate from their respective expectations in opposite directions: when one is above its expectation, the other is below its expectation. Zero covariance, in which case the random variables are called uncorrelated, is intuitively the statement that knowing one of X or Y is above or below its expectation does not say much about the other. Independent random variables are uncorrelated, but the converse is not true in general: uncorrelated random variables could be dependent. (Can you find an example?)

For positive integers ℓ_1 and ℓ_2 such that $\ell_1 + \ell_2 \leq n$, set $Y_1 = X_1 \cdots X_{\ell_1}$ and $Y_2 = X_{\ell_1+1} \cdots X_{\ell_1+\ell_2}$, where the X_j's are from our urn model. Since Y_2 has the same distribution as $X_1 \cdots X_{\ell_2}$, $\mathbb{E}(Y_2) = \mathbb{E}(X_1 \cdots X_{\ell_2})$. If $\ell_1 \leq \ell_2$ then Maclaurin's inequality implies $\mathbb{E}(Y_2)^{\ell_1/\ell_2} \leq \mathbb{E}(Y_1)$, so

$$\mathbb{E}(Y_1Y_2) = \mathbb{E}(X_1 \cdots X_{\ell_1+\ell_2}) \leq \mathbb{E}(Y_2)^{(\ell_1+\ell_2)/\ell_2} = \mathbb{E}(Y_2)^{\ell_1/\ell_2}\mathbb{E}(Y_2) \leq \mathbb{E}(Y_1)\mathbb{E}(Y_2). \quad (10)$$

If $\ell_2 \leq \ell_1$ then Maclaurin's inequality implies $\mathbb{E}(Y_1)^{\ell_2/\ell_1} \leq \mathbb{E}(Y_2)$, so

$$\mathbb{E}(Y_1Y_2) \leq \mathbb{E}(Y_1)^{(\ell_1+\ell_2)/\ell_1} = \mathbb{E}(Y_1)\mathbb{E}(Y_1)^{\ell_2/\ell_1} \leq \mathbb{E}(Y_1)\mathbb{E}(Y_2). \quad (11)$$

Thus $\mathbb{E}(Y_1Y_2) \leq \mathbb{E}(Y_1)\mathbb{E}(Y_2)$ either way, so

$$\mathrm{cov}(Y_1, Y_2) = \mathbb{E}(Y_1Y_2) - \mathbb{E}(Y_1)\mathbb{E}(Y_2) \leq 0.$$

Furthermore, Maclaurin's inequality tells us that $\operatorname{cov}(Y_1, Y_2) = 0$ if and only if all the labels x_i on the balls are equal. If the x_i's are all equal then Y_1 and Y_2 each take just one value, so $\operatorname{cov}(Y_1, Y_2) = 0$. Conversely, if $\operatorname{cov}(Y_1, Y_2) = 0$ then both inequalities in (10) or (11) – depending on whether $\ell_1 \leq \ell_2$ or $\ell_2 \leq \ell_1$ – are equalities. The first inequality in (10) or (11) as an equality, in terms of elementary symmetric means, says $E_{\ell_1+\ell_2}(x_1, \ldots, x_n)^{1/(\ell_1+\ell_2)}$ equals $E_{\ell_2}(x_1, \ldots, x_n)^{1/\ell_2}$ or $E_{\ell_1}(x_1, \ldots, x_n)^{1/\ell_1}$. Either one implies $x_1 = \cdots = x_n$ by the rule for strict inequality in Maclaurin's inequality.

Connection to a Generalized Bernoulli Inequality

We hope you now believe Maclaurin's inequality is interesting. How is the inequality proved? The standard proof (see [2, pp. 10–11], [5, p. 52], [11, Thm. 4, p. 97], or [15, Chap. 12]) is based on Newton's inequality, which says

$$E_{k-1}(x_1, \ldots, x_n)E_{k+1}(x_1, \ldots, x_n) \leq E_k(x_1, \ldots, x_n)^2$$

for $x_1, \ldots, x_n > 0$ and $1 \leq k \leq n-1$, where $E_0(x_1, \ldots, x_n) = 1$. We will present a different approach, based on an extension of Bernoulli's inequality (3). When the right side of (3) is less than or equal to 0, which is when $t \leq -1/n$, Bernoulli's inequality is trivial. When $t > -1/n$ and we set $x = nt$, Bernoulli's inequality can be reformulated as

$$1 + \frac{1}{n}x \geq \sqrt[n]{1+x} \tag{12}$$

for $x > -1$. Doesn't that remind you of the arithmetic-geometric mean inequality? The following extension of (12), which we call the *generalized Bernoulli inequality*, should remind you of Maclaurin's inequality: for each positive integer n and $x > -1$,

$$\boxed{1 + \frac{1}{n}x \geq \sqrt{1 + \frac{2}{n}x} \geq \sqrt[3]{1 + \frac{3}{n}x} \geq \cdots \geq \sqrt[n]{1 + \frac{n}{n}x},} \tag{13}$$

with the inequalities all strict unless $x = 0$.

To prove (13), all terms are equal when $x = 0$. For $x \neq 0$, we want to show

$$\sqrt[k]{1 + \frac{k}{n}x} > \sqrt[k+1]{1 + \frac{k+1}{n}x} \tag{14}$$

when $1 \leq k \leq n-1$. Equivalently, we want to show

$$\frac{1}{k} \log \left(1 + \frac{k}{n} x\right) > \frac{1}{k+1} \log \left(1 + \frac{k+1}{n} x\right),$$

where log is the natural logarithm. We will derive this inequality on the values of logarithms from the fact that $\log t$ is *strictly concave*:

$$\log(\lambda u + (1-\lambda)v) > \lambda \log u + (1-\lambda) \log v \qquad (15)$$

if u and v are distinct positive numbers and $0 < \lambda < 1$. See Figure 1.

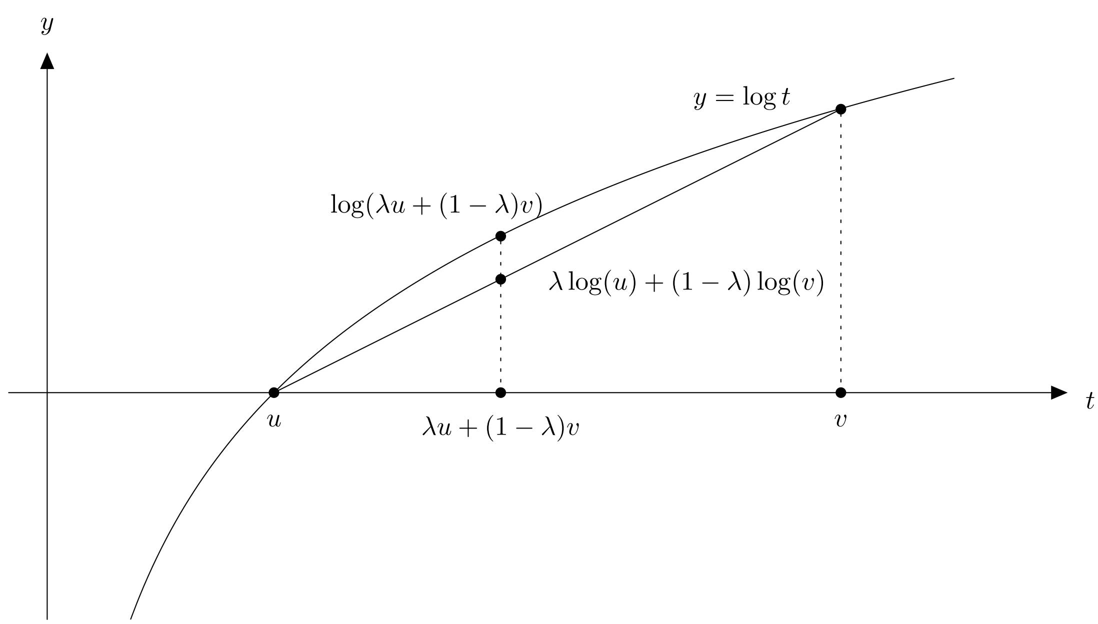


Figure 1: Strict concavity of $\log t$.

Since $1 + \frac{k}{n}x$ lies strictly between $u := 1$ and $v := 1 + \frac{k+1}{n}x$, let's write $1 + \frac{k}{n}x$ as a convex combination of the other terms:

$$1 + \frac{k}{n}x = \lambda u + (1-\lambda)v, \qquad \text{for } \lambda = \frac{1}{k+1}.$$

Then

$$\begin{aligned}
\frac{1}{k}\log\left(1+\frac{k}{n}x\right) &= \frac{1}{k}\log\left(\lambda u+(1-\lambda)v\right)\\
&> \frac{1}{k}\left(\lambda\log u+(1-\lambda)\log v\right)\\
&= \frac{1-\lambda}{k}\log\left(1+\frac{k+1}{n}x\right)\\
&= \frac{1}{k+1}\log\left(1+\frac{k+1}{n}x\right).
\end{aligned}$$

That completes the proof of the generalized Bernoulli inequality in (13).

To derive Maclaurin's inequality from the generalized Bernoulli inequality, we will need a recursive formula for the elementary symmetric means:

$$E_k(x_1,\ldots,x_n)=\left(1-\frac{k}{n}\right)E_k(x_1,\ldots,x_{n-1})+\frac{k}{n}E_{k-1}(x_1,\ldots,x_{n-1})x_n, \tag{16}$$

for $1\le k\le n$, where we set $E_0(x_1,\ldots,x_{n-1})=1$ if $k=1$ and $E_n(x_1,\ldots,x_{n-1})=0$ if $k=n$. This recursion follows from a recursive formula for elementary symmetric polynomials:

$$e_k(x_1,\ldots,x_n)=e_k(x_1,\ldots,x_{n-1})+e_{k-1}(x_1,\ldots,x_{n-1})x_n, \tag{17}$$

where we set $e_0(x_1,\ldots,x_{n-1})=1$ if $k=1$ and $e_n(x_1,\ldots,x_{n-1})=0$ if $k=n$. Dividing both sides of (17) by $\binom{n}{k}$,

$$\begin{aligned}
E_k(x_1,\ldots,x_n) &= \frac{e_k(x_1,\ldots,x_n)}{\binom{n}{k}}\\
&= \frac{e_k(x_1,\ldots,x_{n-1})+e_{k-1}(x_1,\ldots,x_{n-1})x_n}{\binom{n}{k}} \quad \text{by (17)}\\
&= \frac{\binom{n-1}{k}E_k(x_1,\ldots,x_{n-1})+\binom{n-1}{k-1}E_{k-1}(x_1,\ldots,x_{n-1})x_n}{\binom{n}{k}}\\
&= \left(1-\frac{k}{n}\right)E_k(x_1,\ldots,x_{n-1})+\frac{k}{n}E_{k-1}(x_1,\ldots,x_{n-1})x_n.
\end{aligned}$$

Now we are ready to prove Maclaurin's inequality, by induction on n, from the generalized Bernoulli inequality. Maclaurin's inequality for $n=1$ is trivial, and for

$n = 2$ it is the arithmetic-geometric mean inequality for two terms, which can be proved in many ways. Let's derive it from the generalized Bernoulli inequality when $n = 2$, which says $1 + \frac{1}{2}x \geq \sqrt{1+x}$ for $x > -1$. For positive x_1 and x_2,

$$\begin{aligned} E_1(x_1, x_2) &= \frac{x_1 + x_2}{2} \\ &= x_2\left(\frac{x_1/x_2}{2} + \frac{1}{2}\right) \\ &= x_2\left(1 + \frac{1}{2}\left(\frac{x_1}{x_2} - 1\right)\right) \\ &\overset{\text{gen. Bern.}}{\geq} x_2\sqrt{1 + \left(\frac{x_1}{x_2} - 1\right)} \\ &= \sqrt{x_1 x_2} \\ &= \sqrt{E_2(x_1, x_2)}, \end{aligned}$$

and by the generalized Bernoulli inequality this inequality is strict unless $x_1/x_2 - 1 = 0$, that is, $x_1 = x_2$.

Assume that (2) holds for $n-1$ variables, where $n \geq 3$. We want to show it holds for n variables. Since each $E_k(x_1, \ldots, x_n)$ in symmetric in the x_i's we may assume without loss of generality that x_n is the maximal x_i. To simplify notation, write

$$E_k := E_k(x_1, \ldots, x_n) \text{ for } 1 \leq k \leq n, \quad \varepsilon_k := E_k(x_1, \ldots, x_{n-1}) \text{ for } 1 \leq k \leq n-1,$$

and set $\varepsilon_0 := 1$ and $\varepsilon_n := 0$. The recursion (16) can be rewritten as

$$E_k = \left(1 - \frac{k}{n}\right)\varepsilon_k + \frac{k}{n}\varepsilon_{k-1}x_n \tag{18}$$

for $1 \leq k \leq n$. By the induction hypothesis,

$$\varepsilon_{k-1}^{1/(k-1)} \geq \varepsilon_k^{1/k} \text{ for } 2 \leq k \leq n-1.$$

We can rewrite this in two ways:

$$\varepsilon_{k-1} \geq \varepsilon_k^{(k-1)/k} \text{ and } \varepsilon_{k+1} \leq \varepsilon_k^{(k+1)/k} \text{ for } 1 \leq k \leq n-1. \tag{19}$$

(The first inequality holds at $k = 1$ by the definition of ε_0 and the second inequality holds at $k = n-1$ by the definition of ε_n.) Combining (18) and (19), when $1 \leq k \leq n-1$ we get

$$\begin{aligned} E_k &\geq \left(1 - \frac{k}{n}\right)\varepsilon_k + \frac{k}{n}\varepsilon_k^{(k-1)/k} x_n \\ &= \varepsilon_k \left(1 + \frac{k}{n}\left(\varepsilon_k^{-1/k} x_n - 1\right)\right) \end{aligned} \tag{20}$$

and

$$\begin{aligned} E_{k+1} &= \left(1 - \frac{k+1}{n}\right)\varepsilon_{k+1} + \frac{k+1}{n}\varepsilon_k x_n \\ &\leq \left(1 - \frac{k+1}{n}\right)\varepsilon_k^{(k+1)/k} + \frac{k+1}{n}\varepsilon_k x_n \\ &= \varepsilon_k^{(k+1)/k}\left(1 + \frac{k+1}{n}\left(\varepsilon_k^{-1/k} x_n - 1\right)\right). \end{aligned} \tag{21}$$

Letting c_k denote the (positive) term $\varepsilon_k^{-1/k} x_n$ in (20) and (21),

$$E_k^{1/k} \overset{(20)}{\geq} \varepsilon_k^{1/k} \sqrt[k]{1 + \frac{k}{n}(c_k - 1)} \overset{\text{gen. Bern.}}{\geq} \varepsilon_k^{1/k} \sqrt[k+1]{1 + \frac{k+1}{n}(c_k - 1)} \overset{(21)}{\geq} E_{k+1}^{1/(k+1)},$$

which proves (2) for n variables and that completes the induction.

When does equality occur in Maclaurin's inequality? From the way we used the generalized Bernoulli inequality just above, $E_k^{1/k} > E_{k+1}^{1/(k+1)}$ if $c_k - 1 \neq 0$. How can $c_k - 1 = 0$, or equivalently, how can $\varepsilon_k^{1/k} = x_n$? Since $\varepsilon_k^{1/k} \leq \varepsilon_1 = \frac{1}{n-1}(x_1 + \cdots + x_{n-1})$ and x_n is the maximal x_i, if some x_i is less than x_n then $\varepsilon_1 < x_n$, so $\varepsilon_k^{1/k} < x_n$. Therefore the inequalities in (2) are all strict unless every x_i is x_n, in which case each E_k is x_n^k and then the inequalities in (2) are all equalities.

The generalized Bernoulli inequality not only implies Maclaurin's inequality, but follows from it. Fix $x > -1$ and let $x_1 = \cdots = x_{n-1} = 1$ and $x_n = 1 + x$. By (16), for $1 \leq k \leq n-1$

$$E_k(1, \ldots, 1, 1+x) = \left(1 - \frac{k}{n}\right) E_k(1, \ldots, 1) + \frac{k}{n} E_{k-1}(1, \ldots, 1)(1 + x),$$

where E_k and E_{k-1} on the right have $n-1$ 1's in them. Since $k < n$, $E_k(1, \ldots, 1) = 1$ and $E_{k-1}(1, \ldots, 1) = 1$. Therefore $E_k(1, \ldots, 1, 1+x) = (1-k/n) + (k/n)(1+x) = 1+(k/n)x$. Also $E_n(1, \ldots, 1, 1+x) = 1+x = 1+(n/n)x$. Thus Maclaurin's inequality when $x_1 = \cdots = x_{n-1} = 1$ and $x_n = 1+x$ is the generalized Bernoulli inequality. Furthermore, if we know that the inequalities in Maclaurin's inequality are all strict unless $x_1 = x_2 = \cdots = x_n$, then the inequalities in the generalized Bernoulli inequality are all strict unless $x = 0$.

We have derived Maclaurin's inequality from Bernoulli's inequality and then seen that they are in fact equivalent (including conditions on when they become equalities). There is an additional equivalence worth bringing out. The strict concavity (15) for $\log t$, illustrated in Figure 1, was used with $\lambda = 1/(k+1)$ to prove the generalized Bernoulli inequality, which in turn implied Maclaurin's inequality, which has the arithmetic-geometric mean inequality as a special case. Let's complete the cycle by using the arithmetic-geometric mean inequality to prove (15) with rational $\lambda \in (0,1)$, so Maclaurin's inequality and the generalized Bernoulli inequality are equivalent to (15) with rational λ.

Let $0 < u < v$ and let $\lambda \in (0,1)$ be rational. Then $\lambda = \frac{k}{n}$ for some integer $n \geq 2$ and $k \in \{1, \ldots, n-1\}$. Let $x_1 = x_2 = \cdots = x_k = u$ and $x_{k+1} = \cdots = x_n = v$. By the arithmetic-geometric mean inequality,

$$\begin{aligned} \lambda u + (1-\lambda)v &= \frac{(x_1 + \cdots + x_k) + (x_{k+1} + \cdots + x_n)}{n} \\ &> \sqrt[n]{(x_1 \cdots x_k)(x_{k+1} \cdots x_n)} \\ &= u^\lambda v^{1-\lambda}, \end{aligned}$$

where the inequality is strict since $x_1 \neq x_n$. Therefore

$$\log(\lambda u + (1-\lambda)v) > \lambda \log u + (1-\lambda) \log v$$

for all rational $\lambda \in (0,1)$.

Earlier we stated Maclaurin's inequality in probabilistic terms, in (9). It would be fantastic if a reader could develop a proof of Maclaurin's inequality based on probability!

Graph-theoretic Inequalities

A *graph* is an object consisting of vertices that are connected by edges. A typical example of a graph is in Figure 2, where we see that some vertices may not be the endpoint of any edge. The *complete graph* on n vertices, denoted K_n, is the graph with n vertices that has an edge connecting every pair of vertices. The graph K_4 is in Figure 3 (we don't consider the intersection of the two diagonal edges to be a vertex in the graph; to avoid the edge intersection think of K_4 in space as the edges of a tetrahedron). Graphs have applications in the study of networks as well as in pure math, such as algebraic topology. And they are studied in their own right as a branch of combinatorics.

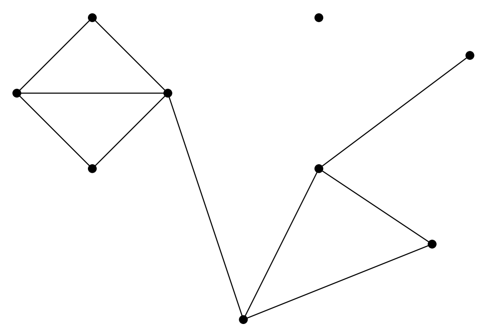


Figure 2: A graph.

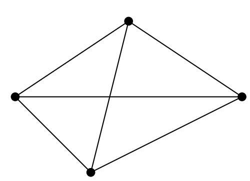


Figure 3: The complete graph on 4 vertices.

Let G be a graph with $n \geq 2$ vertices. We assume it has no edge that starts and ends at the same point (that is, no loops) and there is at most one edge between any two vertices (no multiple edges). A subgraph G' of G is called a *clique* if it is a complete subgraph: every two vertices of G' are connected by an edge of G'. A clique with k vertices is called a k-clique. For instance, a 1-clique is a vertex in G, a 2-clique is a pair of vertices in G and an edge connecting them, and a 3-clique is a set of 3 vertices in G and an edge connecting each pair of these vertices. Figure 2 has 1-cliques, 2-cliques, and 3-cliques, but no k-cliques for $k > 3$. That is, the largest complete subgraph in Figure 2 has 3 vertices.

Let $m = m_G$ be the largest integer k such that G has a k-clique, so $m \leq n$, and $m = n$ if and only if G is the complete graph on n vertices. Assign to each vertex v of G a variable X_v. Let $\mathbf{X}$ be the vector of these variables and for $1 \leq k \leq m$ set

$$e_{k,G}(\mathbf{X}) = \sum_{k-\text{cliques } G_k} \prod_{v \in G_k} X_v.$$

This is a polynomial in the X_v's. Set

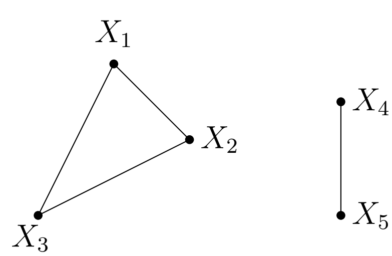
$$E_{k,G}(\mathbf{X}) = \frac{e_{k,G}(\mathbf{X})}{\binom{m}{k}}.$$

If $G = K_n$ then this is the kth elementary symmetric mean $E_k(X_1, \ldots, X_n)$.

Example 5. In the graph below, with two components, $m = 3$. Using the vertex labels from the picture,

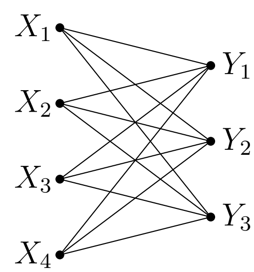
$$E_{1,G} = \frac{X_1 + X_2 + X_3 + X_4 + X_5}{3}, \quad E_{2,G} = \frac{X_1X_2 + X_1X_3 + X_2X_3 + X_4X_5}{3},$$

and $E_{3,G} = X_1X_2X_3$.



Example 6. In the graph below, with 7 vertices, $m = 2$. Using the indicated vertex labels,

$$\begin{aligned} E_1 &= \frac{X_1 + X_2 + X_3 + X_4 + Y_1 + Y_2 + Y_3}{2}, \\ E_2 &= X_1Y_1 + X_1Y_2 + X_1Y_3 + X_2Y_1 + \cdots + X_4Y_3 \\ &= (X_1 + X_2 + X_3 + X_4)(Y_1 + Y_2 + Y_3). \end{aligned}$$



Khadzhiivanov [6] extended Maclaurin's inequality to graphs: if we pick numbers $x_v \geq 0$ for each vertex v in G and let $\mathbf{x}$ be the vector of these numbers, then he proved

$$E_{1,G}(\mathbf{x}) \geq \sqrt{E_{2,G}(\mathbf{x})} \geq \sqrt[3]{E_{3,G}(\mathbf{x})} \geq \cdots \geq \sqrt[m]{E_{m,G}(\mathbf{x})}. \tag{22}$$

When $G = K_n$, (22) is Maclaurin's inequality. Nikiforov [12] has given a recent account of (22), including the case of equality, which is more subtle for general G than when $G = K_n$.

In our treatment of Maclaurin's inequality and the generalized Bernoulli inequality, the latter was derived from the former by setting every variable equal to 1 except the last variable, which was set equal to $1+x$ with $x > -1$. This extends to variables indexed by the vertices of a graph G: if we set each x_v equal to 1 except for a single x_v, which we set equal to $1+x$ with $x > -1$, then the resulting inequality in (22) will be called a *Bernoulli inequality* for G. When $G = K_n$ this is the generalized Bernoulli inequality. Due to the asymmetry in most graphs, there are usually several Bernoulli inequalities for a graph, depending on which variable is set to $1 + x$.

Example 7. The graph in Example 5 has two Bernoulli inequalities. If we set X_1 (or X_2 or X_3) equal to $1 + x$ with $x > -1$ and the remaining 4 variables equal to 1, then (22) becomes

$$1+\frac{2+x}{3} \geq \sqrt{1+\frac{1+2x}{3}} \geq \sqrt[3]{1+x},$$

while if we set X_4 (or X_5) equal to $1+x$ with $x>-1$ and the other 4 variables equal to 1, then (22) is

$$1+\frac{2+x}{3} \geq \sqrt{1+\frac{1+x}{3}} \geq \sqrt[3]{1+x}.$$

The equivalence of Maclaurin's inequality and the generalized Bernoulli inequality, for all n, extends to the setting of graphs (without loops or multiple edges): (22) for all graphs and the Bernoulli inequalities for all graphs are equivalent. The reason is that (22) for all G and all $\mathbf{x}$ is equivalent to (22) for all G with all $x_v = 1$ (Bernoulli inequalities using $x = 0$). This is explained in [12]. As a special case, which gives the general flavor, let's derive the arithmetic-geometric inequality for two terms from (22) for all G with all $x_v = 1$. For positive integers a and b, build a graph with $a+b$ vertices and an edge connecting each of the first a vertices to each of the last b vertices (Example 6 is the case $a = 4$ and $b = 3$.) The inequality (22) for this graph says $(\sum_{i=1}^{a+b} x_i)/2 \geq \sqrt{(\sum_{i=1}^{a} x_i)(\sum_{j=a+1}^{a+b} x_j)}$, and when every x_i is 1 it is $(a+b)/2 \geq \sqrt{ab}$. From $(a+b)/2 \geq \sqrt{ab}$ for positive integers a and b we obtain $(a+b)/2 \geq \sqrt{ab}$ for positive rational a and b by introducing denominators: writing $a = A/C$ and $b = B/C$ for positive integers A, B, and C, $(a+b)/2 \geq \sqrt{ab}$ follows from $(A+B)/2 \geq \sqrt{AB}$ by dividing both sides by C. We get $(a+b)/2 \geq \sqrt{ab}$ for all positive a and b from the case of positive rational a and b by continuity of both sides.

If G is not a complete graph, so $m < n$, then (22) has less than n terms, so there doesn't seem to be an iterative process related to (22) that would be analogous to (7).

Abstract

Maclaurin's inequality is a natural, but nontrivial, generalization of the arithmetic-geometric mean inequality. We present a new proof that is based on an analogous generalization of Bernoulli's inequality. Applications of Maclaurin's inequality to iterative sequences and probability are discussed, along with a graph-theoretic version of the inequality.

References

[1] G. Almkvist and B. Berndt, Gauss, Landen, Ramanujan, the Arithmetic-Geometric Mean, Ellipses, π, and the Ladies Diary, *Amer. Math. Monthly* **95** (1988), 585–608.

[2] E. F. Beckenbach and R. Bellman, "Inequalities", Springer-Verlag, Berlin, 1983.

[3] J. M. Borwein and P. B. Borwein, "Pi and the AGM: A Study in Analytic Number Theory and Computational Complexity", Wiley, New York, 1987.

[4] D. Cox, The Arithmetic-Geometric Mean of Gauss, *L'Enseignement Mathématique* **30** (1984), 275–330.

[5] G. H. Hardy, J. E. Littlewood, G. Polya, "Inequalities", Cambridge Univ. Press, Cambridge, 1934.

[6] N. Khadzhiivanov, Inequalities for Graphs (Russian), C. R. Acad. Bulgare Sci. **30** (1977), 793–796.

[7] C. Maclaurin, A Second Letter from Mr. Colin McLaurin to Martin Folkes, Esq.; Concerning the Roots of Equations, with the Demonstration of Other Rules in Algebra, *Phil. Trans.* **36** (1729), 59–96.

[8] L. Maligranda, Why Hölder's Inequality Should be Called Rogers' Inequality, *Math. Inequal. Appl.* **1** (1998), 69–83.

[9] L. Maligranda, The AM-GM Inequality is Equivalent to the Bernoulli Inequality, *The Mathematical Intelligencer* **34** (2012), 1–2.

[10] Math Overflow, http://mathoverflow.net/questions/37576/nth-order-generalizations-of-the-arithmetic-geometric-mean.

[11] D. S. Mitrinovic, "Analytic Inequalities", Springer-Verlag, Berlin, 1970.

[12] V. Nikiforov, An Extension of Maclaurin's Inequality, arXiv:math/0608199v2.

[13] J. Peetre, Generalizing the Arithmetic-Geometric Mean – a Hapless Computer Experiment, *Internat. J. Math. & Math. Sci.* **12** (1989), 235–246.

[14] S. Ross, “A First Course in Probability” (8th ed.), Pearson, 2010.

[15] J. M. Steele, “The Cauchy–Schwarz Master Class: An Introduction to the Art of Inequalities”, Cambridge Univ. Press, Cambridge, 2004.