A Chapter on Inequalities *

P.S.Bullen
Department of Mathematics, University of British Columbia,
Vancouver, Canada V6T 1Z2

A survey is given of some of the methods used to prove various classical inequalities.

§1. Introduction

The object of this paper is to interest the reader in the topic of inequalities. It is a subject that requires only the most basic mathematical knowledge and skills, although more sophisticated ideas can play a part as well. While it is probably true to say that any mathematician can read a paper on inequalities, it is hoped that this survey will persuade others, who are too modest to call themselves mathematicians, that they too can read such papers, and further that they can add to our knowledge of the subject.

The results to be discussed are all classical, and can be found in most of the standard works listed in the references;[1]—[5], [7],[8],[10],[11]. They will be used to illustrate methods of proof, to show how simple inequalities can be generalized, and the ability of inequalities to assume almost inpenetrable disguises; that is, two inequalities that are in fact identical may look very different; further to this see [3], p.139, and p.155. Most of the time we will use elementary algebra, and sometimes elementary calculus, yet the results discussed are of extreme importance, and are of great use in applications; see [8],[14] and section 4 below.

---

* This paper is based on talks given at the University of Malaya, Kuala Lumpur, and Chulalongkorn University, Bangkok, in the spring of 1993. The author wishes to thank the mathematics departments of those universities for their invitations.

§2. Bernouilli's Inequality

Consider the following statement:

$$\alpha > 0,\ \beta > 0,\ n \geq 1 \Longrightarrow \frac{\alpha^n}{\beta^{n-1}} \geq n\alpha - (n-1)\beta; \qquad (1)$$

and the two sides are equal only if either $n = 1$ or $\alpha = \beta$.

Here, as throughout, we are making statements about real numbers, and in addition we require n to be a positive integer. So (1) is short for: for all positive α, β, and for all positive integers n the inequality in (1) holds, further it is strict except in the two situations mentioned.

How can this be proved? One standard technique is to re-write and, or re-interpret an inequality such as (1) and see if this gives us any idea of how to proceed. We will do this in two ways; the first following Nanjundiah, [12], and secondly in a way that is more standard.

The advantages of several proofs of the same statement are various; it may give more insight into just what is being said, and perhaps when trying to generalize a statement only one of the proofs will suggest how to do this.

Nanjundiah's approach is to re-interpret (1) as a statement about two sequences. If

$\underline{L}$ is the sequence : $L_1 = \dfrac{\alpha}{1}, L_2 = \dfrac{\alpha^2}{\beta}, L_3 = \dfrac{\alpha^3}{\beta^2}, \ldots, L_n = \dfrac{\alpha^n}{\beta^{n-1}}, \ldots$

and,

$\underline{R}$ is the sequence : $R_1 = \alpha, R_2 = 2\alpha - \beta, R_3 = 3\alpha - 2\beta, \ldots,$
$R_n = n\alpha - (n-1)\beta, \ldots$

then inequality (1) just says that the sequence $\underline{L}$ dominates the sequence $\underline{R}$; that is

$$L_n \geq R_n, \qquad n = 1, 2, \ldots \qquad (2)$$

or even more shortly.

$$\underline{L} \geq \underline{R}. \qquad (3)$$

Now since $L_1 = R_1$ both sequences start at the same place, so (2), or (3), will certainly hold if each move along the sequence $\underline{L}$ is at least as big as the corresponding move along the sequence $\underline{R}$; that is if

$$L_n - L_{n-1} \geq R_n - R_{n-1} \qquad n = 2, 3, \ldots$$

or equivalently if

$$(L_n - L_{n-1}) - (R_n - R_{n-1}) \geq 0 \qquad n = 2, 3, \ldots \qquad (4)$$

It is easily seen see that

$$L_n - L_{n-1} = \frac{\alpha^n}{\beta^n}(\alpha - \beta),$$

$$R_n - R_{n-1} = (\alpha - \beta).$$

So, substituting in the left hand side of (4),

$$(L_n - L_{n-1}) - (R_n - R_{n-1}) = \left(\frac{\alpha^n}{\beta^n} - 1\right)(\alpha - \beta).$$

Since either $\alpha > \beta$, when the last two factors are both positive, or $\alpha < \beta$, when the last two factors are both negative, or $\alpha = \beta$ when the last both factors are zero, this leads to (4).

So we have proved (1), together with the cases of equality.

Another method is to re-write inequality (1) to see that it is nothing else but a disguised form of *Bernouilli's Inequality*, a result that goes back to the seventeenth century. To see this let us rearrange and generally play around with inequality (1) to see what we get. At each stage we get a different looking, but completely equivalent inequality.

(i) Division by $\beta > 0$ is possible and will not change the direction of the inequality. Hence:

$$\frac{\alpha^n}{\beta^{n-1}} \geq n\alpha - (n-1)\beta \iff \left(\frac{\alpha}{\beta}\right)^n \geq n\left(\frac{\alpha}{\beta}\right) - (n-1).$$

This shows that our inequality does not depend on two variables α and β but only on their ratio α/β.

(ii) Let us recognise this by putting $x = \alpha/\beta$ to get:

$$\frac{\alpha^n}{\beta^{n-1}} \geq n\alpha - (n-1)\beta \iff x^n \geq nx - (n-1)$$

$$\iff x^n \geq n(x-1) + 1;$$

The conditions on α, β translate into $x > 0$, and the conditions for equality into $n = 1$, or $x = 1$.

(iii) Now why not see what happens if we simplify the right hand side of the last expression by putting $y = x - 1$. This results in the following statement that is equivalent to that in (1):

$$y > -1 \implies (1+y)^n \geq 1 + ny, \tag{B}$$

with equality only if either $n = 1$ or $y = 0$.

This is the well-known Bernouilli Inequality; and there are many ways to prove it as can be in the references [1], [3] or [10].

One way of looking for a proof of an inequality such as (B) is to experiment with it by checking special cases. Let us do that, seeing what happens when $n = 1, 2, 3, \ldots$.

First note that if $n = 1$ then both sides of (B) are the same, while if

$$\begin{aligned} n = 2: \qquad \text{left hand side of (B)} &= (1+y)^2 \\ &= 1 + 2y + y^2 \\ &\geq 1 + 2y = \text{the right hand side of (B)}. \end{aligned}$$

Note that so far we have not needed $y > -1$; also, in the case $n = 2$ equality only occurs if $y = 0$, as stated. If now,

$$\begin{aligned} n = 3: \qquad \text{left hand side of (B)} &= (1+y)^3 \\ &= 1 + 3y + y^2(y+3) \\ &\geq 1 + 3y = \text{the right hand side of (B)}. \end{aligned}$$

provided $y > -3$, and again there is equality only if $y = 0$.

This suggests we apply the Binomial Theorem in the general case; if

$$\begin{aligned} n \geq 2: \qquad \text{left hand side of (B)} &= (1+y)^n \\ &= 1 + ny + y^2\left(\frac{n(n-1)}{2} + \cdots + y^{n-2}\right) \\ &\geq 1 + ny = \text{the right hand side of (B)}, \end{aligned}$$

provided the last bracket is non-negative. However it is not easy to see when this is so. We can only state with certainty that this bracket is non-negative if $y \geq 0$, which is not quite the statement we would like.

Since the Binomial Theorem is proved by *mathematical induction*, this "almost proof" is also an example of proof by induction. If we are a little more careful we can use induction to get the right answer. So let us continue our experiments. Consider the case $n = 4$, and assume as we may that $y \neq 0$; if

$$\begin{aligned} n = 4: \qquad \text{left hand side of (B)} &= (1+y)^4 \\ &= (1+y)^3(1+y) \\ &> (1+3y)(1+y), \quad \text{by the } n = 3 \text{ case} \\ &\qquad \text{just proved, and provided } 1 + y > 0, \\ &= 1 + 4y + 3y^2 > 1 + 4y \\ &= \text{right hand side of (B);} \end{aligned}$$

which gives proof under the right conditions.

A similar argument can be given if $n = 5$ and $y > -1, y \neq 0$, but appealing to the case $n = 4$ that has just been proved.

This can now be formalised into a proof by induction of (B).

A question remains: is the condition $y > -1$ as result of our method of proof or is it essential? A simple example shows that it is necessary:

$$-32 = (-2)^5 = \big(1 + (-3)\big)^5 < 1 + 5(-3) = -14.$$

It would appear that none of the above approaches to (B) suggests how to prove the obvious generalisation in which the positive integer n is replaced by an arbitrary real number r. However note that in such an extension the condition $y > -1$ would be automatically required since the general rth power is not defined for zero or negative numbers. We clearly get equality if $r = 0$ or $r = 1$, the last case being covered by (B) itself; but what is true otherwise, $r = 1/2$ or $r = -1/3$ say? Since rth powers are more sophisticated than ordinary algebra it is only natural to expect to use calculus here.

If we remember that the Binomial Theorem is a special case of *Taylor's Theorem* it is natural to try to use this result to generalize (B). We appeal to Taylors's Theorem with remainder, omitting the trivial cases of $r = 0, 1$, and $y = 0$; then for some z between 0 and y,

$$(1+y)^r = 1 + ry + \frac{r(r-1)}{2}(1+z)^{r-2}. \qquad (T)$$

So, on simple inspection of this we get

$$(1+y)^r \begin{cases} > 1 + ry, & \text{if } r(r-1) > 0; \\ < 1 + ry, & \text{if } r(r-1) < 0; \end{cases}$$

Since the condition $r(r-1)$ is equivalent to $r > 1$ or $r < 0$, and the condition $r(r-1) < 0$ is equivalent to $0 < r < 1$ we get, on collecting the trivial cases the following complete generalisation of (B):

$$y > -1, y \neq 0 \quad \Longrightarrow \quad (1+y)^r \begin{cases} > 1 + ry & \text{if } r > 1 \text{ or } r < 0, \\ < 1 + ry & \text{if } 0 < r < 1, \\ = 1 + ry & \text{if } r = 1 \text{ or } r = 0. \end{cases} \qquad (B_g)$$

However we do not need to use Taylor's Theorem; more elementary calculus will suffice. Consider the following function

$$f(y) = (1+y)^r - (1+ry), \qquad y > -1,$$

and assume that $r \neq 0, 1$.

Easy calculations will show that $f(0) = f'(0) = 0$, and that further $y = 0$ is the only zero of f'; finally $f''(0) = r(r-1)$. It follows that $y = 0$ is the maximum of f if $r(r-1) > 0$, and the minimum of f if $r(r-1) < 0$; so in the first case f is a negative function except at $y = 0$, while in the second case f is a positive function except at $y = 0$. This gives another proof of (B_g).

Of course, by reversing the steps that led to (B) from (1) we can get, from (B_g) the following generalization of (1).

$$\alpha > 0, \beta > 0, \quad \Longrightarrow \quad \frac{\alpha^r}{\beta^{r-1}} \geq r\alpha - (r-1)\beta; \qquad (1_g)$$

if $r \geq 1$ or $r \leq 0$;if $0 \leq r \leq 1$ the oppposite inequality holds; there is equality only if $r = 1, r = 0$, or if $\alpha = \beta$.

§3. The Geometric and Arithmetic Means

If $a_1, \ldots, a_n$ are n positive numbers put $\underline{a} = (a_1, \ldots, a_n)$ and write

$$A_n(\underline{a}) = A_n(a_1, \ldots, a_n) = \frac{a_1 + a_2 + \cdots + a_n}{n};$$

$$G_n(\underline{a}) = G_n(a_1, \ldots, a_n) = \sqrt[n]{a_1 a_2 \cdots a_n};$$

These quantites are the *arithmetic and geometric means* of $\underline{a}$, respectively.

Is it true that these two means are *comparable*? That is, is it true that for all positve $\underline{a}$, and all positive integers n that we either always have that

$$G_n(\underline{a}) = \sqrt[n]{a_1 a_2 \cdots a_n} \leq \frac{a_1 + a_2 + \cdots + a_n}{n} = A_n(\underline{a}), \tag{G-A}$$

or always have the opposite inequality?

A few numerical experiments will suggest that if anything is true then it is (G-A), not the opposite, that will hold. Let us see if we can prove this by first looking at some special values of n. In the case of $n = 2$, and writing $a_1 = x, a_2 = y$, (G-A) is

$$\sqrt{xy} \leq \frac{x+y}{2}. \tag{5}$$

As before, let us try rewriting and see if anything is suggested:

$$\begin{aligned} \sqrt{xy} \leq \frac{x+y}{2} \quad &\Longleftrightarrow \quad 2\sqrt{xy} \leq x + y \\ &\Longleftrightarrow \quad 0 \leq x - 2\sqrt{xy} + y \\ &\Longleftrightarrow \quad 0 \leq (\sqrt{x} - \sqrt{y})^2; \end{aligned}$$

but this last statement is obvious. Further this argument shows that (5) is strict unless $x = y$.

However if we now try to get a proof of the case $n = 3$ in a similar way we run into difficulties so let us follow Cauchy and proceed to $n = 4$; there is some merit in this as fourth powers are related to squares in a nice way. Put $a_1 = a, a_2 = b, a_3 = c, a_4 = d$ when (G-A) is

$$\sqrt[4]{abcd} \leq \frac{a+b+c+d}{4}. \tag{6}$$

Trying to use what we know rewrite the right hand side of (6) as follows

$$\begin{aligned} \frac{a+b+c+d}{4} &= \frac{\dfrac{a+b}{2} + \dfrac{c+d}{2}}{2} \\ &\geq \frac{\sqrt{ab} + \sqrt{cd}}{2}, \text{by (5), by the case } n = 2; \\ &\geq \sqrt{(\sqrt{ab})(\sqrt{cd})}, \text{again by (5), the case } n = 2; \\ &= \sqrt[4]{abcd}; \end{aligned}$$

which is (6). To get equality we need, from the first use of the $n=2$ case that $a=b$ and $c=d$, and from the second use we need that $ab=cd$; that is we need to have $a=b=c=d$.

Now we can go back to the case $n=3$:

$$\frac{a+b+c}{3}=\frac{a+b+c+\dfrac{a+b+c}{3}}{4}$$

$$\geq\left(abc\left(\frac{a+b+c}{3}\right)\right)^{1/4}, \text{ by (6), the case } n=4.$$

This, by simple algebra, is just

$$\sqrt[3]{abc}\leq\frac{a+b+c}{3},$$

with equality only if $a=b=c$.

This procedure suggests that next we consider the case $n=8$, and then come back to the cases $n=5,6,7$, then move to the case $n=16$, and come back to the cases $n=9,\ldots,15$. It can be checked that this will in fact work, the method is sometimes known as *backward induction.* Of course we do not have to prove each case missed separately; for instance we can handle an arbitrary $k, 9\leq k\leq 15$ by writing

$$\frac{a_1+\cdots+a_k}{k} \tag{7}$$
$$=\frac{a_1+\cdots+a_k+\left(\dfrac{a_1+\cdots+a_k}{k}\right)+\cdots+\left(\dfrac{a_1+\cdots+a_k}{k}\right)}{16},$$

there being $16-k$ terms equal to $(a_1+\cdots+a_k)/k$ in the numerator on the right, giving a sum of 16 terms to which the $n=16$ case can be applied.

So (G-A) holds with equality only when $a_1=\cdots=a_n$. There are many proof of this important result; 52 are given in [3], most of them extremely elementary.

Now the question of generalizations arises. If we rewrite (G-A) as

$$A_n(\underline{a})-G_n(\underline{a})\geq 0 \quad \text{or} \quad \frac{A_n(\underline{a})}{G_n(\underline{a})}\geq 1. \tag{8}$$

then since equality in the inequalities (8) can occur only in the trivial case of $a_1 = \cdots = a_n$ we could ask if bigger right hand sides are possible if not all of the $a_i, 1 \le i \le n$ are equal? To answer this question we follow Nanjundiah, [12] and use (1).

Let us agree to the notation for $1 \le k \le n$,

$$A_k(\underline{a}) = \frac{a_1 + \cdots + a_k}{k} \qquad G_k(\underline{a}) = \left(a_1 \cdots a_k\right)^{1/k}.$$

If

$$\alpha = A_n(\underline{a}),\ \beta = A_{n-1}(\underline{a}) \quad \text{easy calculations give that} \quad n\alpha - (n-1)\beta = a_n;$$

while if

$$\alpha' = G_n(\underline{a}),\ \beta' = G_{n-1}(\underline{a}), \quad \text{easy calculations give that} \quad \frac{\alpha'^n}{\beta'^{n-1}} = a_n.$$

This enables us to use (1) twice as follows:

$$\frac{\alpha^n}{\beta^{n-1}} \ge n\alpha - (n-1)\beta = a_n = \frac{\alpha'^n}{\beta'^{n-1}} \ge n\alpha' - (n-1)\beta'.$$

From this we can deduce the two inequalities,

$$\frac{\alpha^n}{\beta^{n-1}} \ge \frac{\alpha'^n}{\beta'^{n-1}} \qquad \text{and} \qquad n\alpha - (n-1)\beta \ge n\alpha' - (n-1)\beta'.;$$

which on rewriting give

$$\left(\frac{\alpha}{\alpha'}\right)^n \ge \left(\frac{\beta}{\beta'}\right)^{n-1} \qquad \text{and} \qquad n(\alpha - \alpha') \ge (n-1)(\beta - \beta').$$

If the values given to $\alpha, \alpha', \beta, \beta'$ are substituted into these last in terms the following inequalities, called the *inequalities of Popviciu and Rado*, respectively, are obtained.

$$\begin{gathered} \left(\frac{A_n(\underline{a})}{G_n(\underline{a})}\right)^n \ge \left(\frac{A_{n-1}(\underline{a})}{G_{n-1}(\underline{a})}\right)^{n-1} \\ \textit{and} \\ n\big(A_n(\underline{a}) - G_n(\underline{a})\big) \ge (n-1)\big(A_{n-1}(\underline{a}) - G_{n-1}(\underline{a})\big). \end{gathered} \tag{9}$$

Inequalities (9) improve inequalities (8) since unless $a_1 = \cdots = a_{n-1}$ the right hand sides are respectively bigger than zero, bigger than 1. In the deduction of (9) we used (1) and so there will be equality in (9) only if there is equality in the uses of (1); that is only if $\alpha = \beta$ for Popoviciu's inequality and $\alpha' = \beta'$ for Rado's inequality; in other words, unless $a_n = A_{n-1}(\underline{a})$, and $a_n = G_{n-1}(\underline{a})$, repectively.

Incidentally not only do (9) improve (G-A), they also give a proof of this inequality. Consider the case of the Rado inequality, and re-apply it to get

$$\begin{aligned} n\big(A_n(\underline{a}) - G_n(\underline{a})\big) &\geq (n-1)\big(A_{n-1}(\underline{a}) - G_{n-1}(\underline{a})\big) \\ &\geq (n-2)\big(A_{n-2}(\underline{a}) - G_{n-2}(\underline{a})\big) \\ &\geq \cdots \geq 2\big(A_2(\underline{a}) - G_2(\underline{a})\big) \geq \big(A_1(\underline{a}) - G_1(\underline{a})\big) = 0, \end{aligned} \tag{10}$$

giving (G-A) in the first form in (8). If we stop one stage earlier in (10) we get

$$A_n(\underline{a}) - G_n(\underline{a}) \geq \frac{2}{n}\big(A_2(\underline{a}) - G_2(\underline{a})\big) = \frac{\left(\sqrt{a_1} - \sqrt{a_n}\right)^2}{n};$$

However, the order of the terms in $\underline{a}$ do not matter, so by re-ordering if necessary, we get the following improvement of (G-A)

$$A_n(\underline{a}) - G_n(\underline{a}) \geq \left(\frac{1}{n}\right) \max_{1 \leq i,j \leq n} \left(\sqrt{a_i} - \sqrt{a_j}\right)^2.$$

Now let us turn to a generalization of (G-A) suggested by (7); the right hand side of (7) can be rewritten as

$$\frac{a_1 + \cdots + a_k + (16-k)\dfrac{a_1 + \cdots + a_k}{k}}{16},$$

where now the numerator has $k+1$ terms but one of the terms is given a weight of $16-k$, compared to a weight of 1 for all the other k terms. So now we introduce n positive numbers $\underline{w} = (w_1, \ldots, w_n)$, write $W_n = w_1 + \cdots + w_k$, $1 \leq k \leq n$, and define

$$A_n(\underline{a};\underline{w}) = A_n(a_1, \ldots, a_n; w_1, \ldots, w_n) = \frac{a_1 w_1 + a_2 w_2 + \cdots + a_n w_n}{W_n};$$

$$G_n(\underline{a};\underline{w}_n) = G_n(a_1, \ldots, a_n; w_1, \ldots, w_n) = \left(a_1^{w_1} a_2^{w_2} \cdots a_n^{w_n}\right)^{1/W_n};$$

the *arithmetic and geometric means of $\underline{a}$ with weights $\underline{w}$*, respectively. The natural generalisation of (G-A) is

$$G_n(\underline{a};\underline{w}) = \left(a_1^{w_1}\cdots a_n^{w_n}\right)^{1/W_n} \le \frac{a_1w_1+\cdots+a_nw_n}{W_n} = A_n(\underline{a};\underline{w}). \tag{G-A$_g$}$$

This can be proved, even improved, by repeating the above argument that used (1), but now use (1_g) with

$$\alpha = A_n(\underline{a};\underline{w}), \beta = A_{n-1}(\underline{a};\underline{w}), \alpha' = G_n(\underline{a};\underline{w}), \beta' = G_n(\underline{a};\underline{w}),$$
and $r = W_n/w_n$, when $r-1 = W_{n-1}/w_n$.

This leads to the following generalization of (9),

$$\left(\frac{A_n(\underline{a};\underline{w})}{G_n(\underline{a};\underline{w})}\right)^{W_n} \ge \left(\frac{A_{n-1}(\underline{a};\underline{w})}{G_{n-1}(\underline{a};\underline{w})}\right)^{W_{n-1}}$$
and
$$W_n\big(A_n(\underline{a};\underline{w}) - G_n(\underline{a};\underline{w})\big) \ge W_{n-1}\big(A_{n-1}(\underline{a};\underline{w}) - G_{n-1}(\underline{a};\underline{w})\big). \tag{9_g}$$

§4. Applications of the Inequality between the Arithmetic and Geometric Means

(a) The name mean is justified since it is very easy to prove that

$$\min \underline{a} \le A_n(\underline{a};\underline{w}) \le \max \underline{a} \qquad \min \underline{a} \le G_n(\underline{a};\underline{w}) \le \max \underline{a}; \tag{11}$$

further these inequalities are strict unless $a_1 = \cdots a_n$.

(b) Let us agree to a simple extension of our notation, writing $\underline{a} = (a_1, a_2, \ldots, a_n \ldots)$, and $\underline{w} = (w_1, w_2, \ldots, w_n, \ldots)$ for two sequences of positive numbers; then we will get two sequences of means,

$$\underline{A} = \big(A_1(\underline{a};\underline{w}), A_2(\underline{a};\underline{w}), \ldots, A_n(\underline{a};\underline{w}), \ldots\big)$$
$$\underline{G} = \big(G_1(\underline{a};\underline{w}), G_2(\underline{a};\underline{w}), \ldots, G_n(\underline{a};\underline{w}), \ldots\big);$$

when of course (G-A) just says that the sequence $\underline{A}$ dominates the sequence $\underline{G}$, or $\underline{G} \le \underline{A}$

A question that can be asked is whether properties of a sequence are inherited by the sequence of its means? One classical theorem of Cauchy says that if a sequence converges then so does the sequence of its arithmetic means, and further the limit is the same; see [6] p.90. A modification of the proof of this result will easily show that the same is true for the geometric means; alternatively we could use the simple identity:

$$G_n(\underline{a};\underline{w}) = \exp\Big(A_n(\log \underline{a};\underline{w})\Big), \tag{12}$$

where we have introduced the useful notation $f(\underline{a}) = (f(a_1), f(a_2), \ldots, f(a_n), \ldots)$.

Then if $\lim_{n\to\infty} a_n = a$, we have from (12), using Cauchy's result just quoted that

$$\lim_{n\to\infty} G_n(\underline{a};\underline{w}) = \exp\Big(\lim_{n\to\infty}\big(A_n(\log \underline{a};\underline{w})\big)\Big) = \exp\log a = a.$$

Another property that is inherited is monotonicity; if $\underline{a}$ is increasing then so is $\underline{A}$, and $\underline{G}$; further if the sequence is strictly increasing so are the mean sequences. First note the useful identities:

$$\begin{aligned} A_n(\underline{a};\underline{w}) &= \frac{W_{n-1}}{W_n}A_{n-1}(\underline{a};\underline{w}) + \frac{w_n}{W_n}a_n \\ &= A_2\big(A_{n-1}(\underline{a};\underline{w}), a_n; W_{n-1}, w_n\big); \\ G_n(\underline{a};\underline{w}) &= G_{n-1}(\underline{a};\underline{w})^{W_{n-1}/W_n} a_n^{w_n/W_n} \\ &= G_2\big(G_{n-1}(\underline{a};\underline{w}), a_n; W_{n-1}, w_n\big). \end{aligned} \tag{13}$$

Suppose now that the sequence $\underline{a}$ is strictly increasing; by (11), $A_{n-1}(\underline{a};\underline{w}) < a_{n-1}$ and $G_{n-1}(\underline{a};\underline{w}) < a_{n-1}$; so by (11) and (13), $A_{n-1}(\underline{a};\underline{w}) < A_n(\underline{a};\underline{w})$ and $G_{n-1}(\underline{a};\underline{w}) < G_n(\underline{a};\underline{w})$, that is, the sequences of the means are also strictly increasing.

(c) If now $e_n = (1+1/n)^n, n = 1, \ldots$ an essential step in showing that $\lim_{n\to\infty} e_n = e$ is to show that the sequence $e_n, n = 1, \ldots$ is strictly increasing. We can use (G-A) to do this, using an argument of Melzak; [9]. Let

$$a_1 = 1, a_k = 1 + \frac{1}{n-1}, k = 2, \ldots, n;$$

then

$$G_n(\underline{a}) = \left(1 + \frac{1}{n-1}\right)^{\frac{n-1}{n}} < A_n(\underline{a}) = \frac{1 + (n-1)(1 + \frac{1}{n-1})}{n} = 1 + \frac{1}{n};$$

which is just $e_{n-1} < e_n$.

(d) Now we prove that

$$\lim_{n\to\infty} \frac{n}{\sqrt[n]{n!}} = e. \tag{14}$$

To do this it suffices to note that if

$$a_1 = 1, a_k = \frac{k^k}{k(k-1)^{k-1}} = e_{k-1}, k = 2, \ldots, n,$$

then

$$\epsilon_n = \frac{n}{\sqrt[n]{n!}} = G_n(\underline{a}).$$

Since, as we have just remarked the sequence $e_n, n = 1, \ldots$ increases strictly to the limit e, so also, by the above discussion, must the sequence $\epsilon_n, n = 1, \ldots$ increase strictly to the same limit.

§5. The Power Means

Another mean of great antiquity is the *harmonic mean*;

$$H_n(\underline{a}; \underline{w}) = \frac{W_n}{(w_1/a_1) + \cdots + (w_n/a_n)};$$

or, with the above notation,

$$H_n(\underline{a}; \underline{w}) = \left(A_n(\underline{a}^{-1}; \underline{w}) \right)^{-1}. \tag{15}$$

Then a simple application of (G-A$_g$) gives

$$\left(A_n(\underline{a}^{-1}; \underline{w}) \right)^{-1} \leq \left(G_n(\underline{a}^{-1}; \underline{w}) \right)^{-1} = G_n(\underline{a}; \underline{w}),$$

or

$$H_n(\underline{a}; \underline{w}) \leq G_n(\underline{a}; \underline{w}). \tag{16}$$

The relation (15) is analogous to (12) in that in (12) we use the inverse functions exponential and logarithm, while in (15) we use the reciprocal function twice, it being its own inverse. This suggests a generalization in which we use the rth power function and its inverse the $(1/r)$th power function. This would give the rth *power mean*, $-\infty < r < \infty, r \neq 0$;

$$M_n^{[r]} = \left(A_n(\underline{a}^r;\underline{w})\right)^{1/r} = \left(\frac{a_1^r w_1 + \cdots + a_n^r w_n}{W_n}\right)^{1/r}.$$

Clearly then $r = 1, r = -1$ are just the arithmetic and harmonic means respectively. While by (12) the geometric mean has a similar form it does not seem to fit into this scheme. However, we will show that

$$\lim_{r\to 0} M_n^{[r]}(\underline{a};\underline{w}) = G_n(\underline{a};\underline{w}), \tag{17}$$

so that we could reasonably define

$$M_n^{[0]}(\underline{a};\underline{w}) = G_n(\underline{a};\underline{w}).$$

To prove (17) first note that we can without loss of generality assume that $a_i > 1, 1 \le i \le n$ for if not consider $\lambda\underline{a}$, where λ is large enough to ensure that $\lambda a_i > 1, 1 \le i \le n$, and note that $M_n^{[r]}(\lambda\underline{a};\underline{w}) = \lambda M_n^{[r]}(\underline{a};\underline{w})$. With this in mind consider

$$\log M_n^{[r]}(\underline{a};\underline{w}) = \frac{-\log W_n + \log(a_1^r w_1 + \cdots + a_n^r w_n)}{r};$$

As $r \to 0$ both the numerator and the denominator of the right hand side tend to zero so we can apply L'Hôpital's Rule to get:

$$\begin{aligned}\lim_{r\to 0}\log M_n^{[r]}(\underline{a};\underline{w}) &= \lim_{r\to 0}\frac{w_1 a_1^r \log a_a + \cdots + w_n a_n^r \log a_n}{w_1 a_1^r + \cdots + w_n a_n^r}\\ &= \frac{w_1\log a_1 + \cdots + w_n \log a_n}{Wn}\\ &= A_n(\log \underline{a};\underline{w}) = \log G_n(\underline{a};\underline{w}), \text{ by (12)}\end{aligned}$$

A similar line of reasoning will show that

$$\lim_{r\to\infty} M_n^{[r]}(\underline{a};\underline{w}) = \max \underline{a} \qquad \lim_{r\to-\infty} M_n^{[r]}(\underline{a};\underline{w}) = \min \underline{a};$$

so it is reasonable to define $M_n^{[\infty]}(\underline{a};\underline{w})$ and $M_n^{[-\infty]}(\underline{a};\underline{w})$ by these limit values.

It is a simple exercise to extend (11) as:

if $-\infty < r < \infty$ *then*

$$M_n^{[-\infty]}(\underline{a};\underline{w}) \le M_n^{[r]}(\underline{a};\underline{w}) \le M_n^{[\infty]}(\underline{a};\underline{w}), \tag{18}$$

with equality only if $a_1 = \cdots = a_n$; further (G-A$_g$) and (16) give:

$$M_n^{[-1]}(\underline{a};\underline{w}) \le M_n^{[0]}(\underline{a};\underline{w}) \le M_n^{[1]}(\underline{a};\underline{w}),$$

with equality only if $a_1 = \cdots = a_n$.

In fact these last two sets of inequalities are just particular cases of a another generalization of (G-A);

if $-\infty \le r < s \le \infty$ *then*

$$M_n^{[r]}(\underline{a};\underline{w}) \le M_n^{[s]}(\underline{a};\underline{w}), \tag{$r;s$}$$

with equality only if $a_1 = \cdots = a_n$.

There are many proofs of this fundamental result, see [1], [3], [4], [10], [11]. We show now that rewriting will reduce (r;s) to a few special cases.

(I) The cases when one at least of r, s is not finite are already contained in (18).

(II) If one of r or s is zero, and the other is finite, the result is equivalent to (G-A$_g$).

(i) $(0;s)$); put $\underline{b} = \underline{a}^s$ when, since $s > 0$,

$$
\begin{aligned}
(0;s) \quad &\iff \quad G_n(\underline{b}^{1/s};\underline{w}) \le \Big(A_n(\underline{b};\underline{w})\Big)^{1/s} \\
&\iff \quad G_n(\underline{b};\underline{w}) \le A_n(\underline{b};\underline{w}).
\end{aligned}
$$

(ii) $(r;0)$; substitute $\underline{b} = \underline{a}^r$ when, since $r < 0$,

$$
\begin{aligned}
(r;0) \quad &\iff \quad \Big(A_n(\underline{b};\underline{w})\Big)^{1/r} \le G_n(\underline{b}^{1/r};\underline{w}) \\
&\iff \quad G_n(\underline{b});\underline{w}) \le A_n(\underline{b};\underline{w}).
\end{aligned}
$$

(III) The cases $-\infty < r \le 0 \le s < \infty$ follow from (II) and (G-A$_g$).

(IV) The cases $-\infty < r < s < 0$ follow from the cases $0 < r < s < \infty$. Suppose that r, s are negative and put $\rho = -r, \sigma = -s$ and $\underline{b} = 1/\underline{a}$, then $0 < \sigma < \rho < \infty$ and

$$
\begin{aligned}
(r;s) \quad &\iff \quad \Big(M_n^{[\rho]}(\underline{b};\underline{w})\Big)^{-1} \le \Big(M_n^{[\sigma]}(\underline{b};\underline{w})\Big)^{-1} \\
&\iff \quad M_n^{[\sigma]}(\underline{b};\underline{w}) \le M_n^{[\rho]}(\underline{b};\underline{w}). \\
&\iff \quad (\sigma;\rho).
\end{aligned}
$$

(V) The cases $0 < r < s < \infty$ follow from the cases $(1;t)$. If $0 < r < s < \infty$ put $\underline{b} = \underline{a}^r, t = s/r$, when $t > 1$ and,

$$(r;s) \iff A_n(\underline{b};\underline{w}) \le M_n^{[t]}(\underline{b};\underline{w}) \iff (1;t). \qquad (19)$$

This long but elementary argument reduces the consideration of (r;s) to (G-A_g),(II), already discussed, and the single case (1;t),(V), (19). To complete the discussion we turn to yet another generalization.

§6. Convex Functions

Let us reconsider the proof of (B_g); the use of Taylor's Theorem there allows us to prove much more. Suppose that ϕ is any function that is twice differentiable in some open interval containing 1; if then $1 + y$ is in that interval Taylor's Theorem says that for some z between 0 and y,

$$\phi(1+y) = \phi(1) + y\phi'(1) + \frac{y^2}{2}\phi''(1+z), \qquad (T_g)$$

an equation that is a direct generalization of (T). So now if ϕ has a second derivative that is non-negative we can deduce the inequality

$$\phi(1+y) \ge \phi(1) + y\phi'(1); \qquad (20)$$

and (B_g) is just (20) in the case $\phi(x) = x^r, r \le 0, r \ge 1$; further if the second derivative of ϕ is positive the only case of equality in (20) is $y = 0$. If the second derivative is non-positive then the opposite inequality to (20) holds, with equality again only when $y = 0$ if this derivative is negative,

By repeating the discussion of the first paragraph we get the following generalization of (1_g), for ϕ with $\phi'' \ge 0$;

$$\beta\phi(\alpha/\beta) \ge \alpha\phi'(1) + \beta\big(\phi(1) - \phi'(1)\big).$$

The generality of (20) can be seen in that it contains

$$e^{1+y} > 1 + y, \quad y \ne 0$$

by taking $\phi(x) = e^x$; while taking $\phi(x) = \log x$, and using the opposite inequality to (20) we get

$$\log(1+y) < 1 + y, \quad y \ne 0.$$

It would seem that functions ϕ for which $\phi'' \geq 0$ are useful; they are called *convex functions*, while if $\phi'' > 0$ we call them *strictly convex*. Strictly speaking the class of convex functions can be defined without assuming second order differentiability, but that is not important here, [16]. Functions ϕ for which $\phi'' \leq 0$, (< 0) are called *concave, (strictly concave), functions.* The importance of these classes is further enhanced by the fact that all such functions they satisfy an important inequality, *Jensen's Inequality.* Let us see that this is the case. We will only discuss the case of convex functions, the cases of strictly convex, concave and strictly concave functions are essentially similar.

If $\phi'' \geq 0$ then ϕ' is increasing so that if $x < z < y$ we must have the following inequality between chord slopes

$$\frac{\phi(z) - \phi(x)}{z - x} \leq \frac{\phi(y) - \phi(z)}{y - z}. \tag{21}$$

This is because, by the Mean Value Theorem of Differentiation the left hand side of (21) is $\phi'(u)$ for some u, $x < u < z$, while the right hand side of (21) is $\phi'(v)$ for some v, $z < v < y$, and $u < v$ means that $\phi'(u) \leq \phi'(v)$. Now consider the simple identity between various chord slopes:

$$\frac{\phi(y) - \phi(x)}{y - x} = \left(\frac{y - z}{y - x}\right)\left(\frac{\phi(z) - \ qf(x)}{z - x}\right) + \left(\frac{z - x}{y - x}\right)\left(\frac{\phi(y) - \phi(z)}{y - z}\right).$$

Since the coefficients on the right hand side add to 1 and are positive, this exhibits the chord slope on left hand side as a weighted arithmetic mean of the two chord slopes on the right hand side. So either immediately, or as an application of (11), we see that from (21) we get that

$$\frac{\phi(z) - \phi(x)}{z - x} \leq \frac{\phi(y) - \phi(x)}{y - x} \leq \frac{\phi(y) - \phi(z)}{y - z}. \tag{22}$$

Since we chose x, y, z with $x < y < z$ but otherwise arbitrarily it follows that (21) implies that the chord slope $(\phi(y) - \phi(x))/(y - x)$ increases with either x or with y. In particular in (22) we let z tend to x, and independently let z tend to y we deduce that $\phi'(x) \leq \phi'(y)$, or that ϕ' is increasing, that is $\phi'' \geq 0$, or ϕ is convex. So that (21) is equivalent to convexity.

Rewriting (21) we get that

$$\phi(z) \leq \frac{y - z}{y - x}\phi(x) + \frac{z - x}{y - x}\phi(y); \tag{23}$$

now note that

$$z = \frac{y-z}{y-x}x + \frac{z-x}{y-x}y$$

and that if

$$\lambda = \frac{z-x}{y-x} \quad \text{then}, 0 < \lambda < 1, \quad \text{and} \quad 1-\lambda = \frac{y-z}{y-x}$$

so that (23) can be written;
if $0 < \lambda < 1$

$$\phi((1-\lambda)x + \lambda y) \le (1-\lambda)\phi(x) + \lambda\phi(y);$$

or (J)

$$\phi\Bigg(A_2(x, y; 1-\lambda, \lambda) \Bigg) \le A_2\Bigg(\phi(x), \phi(y); 1-\lambda, \lambda \Bigg).$$

This is Jensen's inequality, and as we have seen is equivalent to convexity; it has a simple geometric interpretation: the points on the chord are above the points of the graph see [3], p.27. The discussion of the strictly convex case would show then we can only get equality in (J) if $x = y$.

To see the power of (J) note the following two cases:

(a): $\phi(x) = x^t, t > 1$, is a strictly convex function; so by (J) in this case

$$\Bigg(A_2(x, y; 1-\lambda, \lambda) \Bigg)^t \le A_2(x^t, y^t; 1-\lambda, \lambda).$$

or, with a slight change of notation

$$A_2(\underline{a}; \underline{w}) \le M_2^{[t]}(\underline{a}; \underline{w}),$$

the special case $n = 2$ of (19).

(b): $\phi(x) = \log x$, a strictly concave function; so by the opposite inequality to (J) in this case

$$\log\Bigg(A_2(x, y; 1-\lambda, \lambda) \Bigg) \ge A_2\Bigg(\log x, \log y, 1-\lambda, \lambda \Bigg);$$

or, by (11), again with a slight change of notation

$$A_2(\underline{a}; \underline{w}) \ge G_2(\underline{a}; \underline{w}),$$

the special case $n = 2$ of (G-A$_g$).

To be completely useful we need to generalize (J) :
if ϕ is convex then

$$\phi(A_n(\underline{a};\underline{w}) \geq A_n(\phi(\underline{a});\underline{w}) \qquad (J_g),$$

further if ϕ is strictly convex there is equality only in the case $a_1 = \cdots = a_n$.

The proof is a simple application of induction and the special case (J); consider (J_g),

$$\begin{aligned}\text{right hand side} =& A_n(\phi(\underline{a});\underline{w}) = \frac{W_{n-1}}{W_n}A_{n-1}(\phi(\underline{a});\underline{w}) + \frac{w_n}{W_n}\phi(a_n), \quad \text{by (13)}\\ \geq& \frac{W_{n-1}}{W_n}\phi(A_{n-1}(\underline{a};\underline{w})) + \frac{w_n}{W_n}\phi(a_n),\\ & \text{by the induction hypothesis}\\ \geq& \phi\left(\frac{W_{n-1}}{W_n}A_{n-1}(\underline{a};\underline{w}) + \frac{w_n}{W_n}a_n\right), \text{ by (J)}\\ =& \phi(A_n(\underline{a};\underline{w})) = \text{left hand side.}\end{aligned}$$

Now using (J_g) we can use the arguments in (a) to complete the proof of (19), and of (b) to give another proof of (G-A$_g$).

This is but an introduction to the many applications of convex functions and of Jensen's inequality; see [3], [4], [8], [10],[11], [14], [16].

§7. Nanjundiah's Inequalities

It might be thought that no further generalizations of (B) are possible. However generalizations come in two types; those like (J$_g$) are extremely useful because of the many important special cases, those like the inequalities of Rado and Popoviciu are important because they give more insight into an important inequality. On of the most remarkable generalizations of the second kind is due to Nanjundiah,[13], remarkable and neglected. This is in part due to no proof having been published, the proof in [3], p.121, being wrong, as was pointed out in a private communication by Alzer. Professor Nanjundiah has told the author that he hopes to publish his proofs in the near future.

What are these remarkable results. For simplicity we will only consider the case of equal weights. Given a sequence of positive numbers, $\underline{a}$ the rth power means of the first n terms of this sequence define a new sequence,

$$\underline{M}^{[r]} = (M_1^{[r]}(\underline{a}), M_2^{[r]}(\underline{a}), \ldots, M_n^{[r]}(\underline{a}), \ldots);$$

the cases $r = 0, 1$ have already been used in section 4.

In terms of these sequences the inequality (r;s) says that if $-\infty \le r < s \le \infty$ then the sequence $\underline{M}^{[s]}$ dominates the sequence $\underline{M}^{[r]}$, or $\underline{M}^{[r]} \le \underline{M}^{[s]}$.

Suppose now we take the smaller rth power mean of the larger sth sequence, how does it compare with the larger sth power mean of the smaller rth sequence; or are the two comparable? The remarkable result of Nanjundiah is

$$M_n^{[r]}(\underline{M}^{[s]}) \ge M_n^{[s]}(\underline{M}^{[r]}). \tag{N}$$

To get some idea of just how remarkable a result (N) is we consider first the special case $n = 2, 0 < r < s$ and put $a_1 = x, a_2 = y$, when (N) becomes

$$\left(\frac{x^r + (\frac{x^s + y^s}{2})^{r/s}}{2}\right)^{1/r} \ge \left(\frac{x^s + (\frac{x^r + y^r}{2})^{s/r}}{2}\right)^{1/s}. \tag{24}$$

We simplify this along the lines of our discussion of (r;s); put $a = x^r, b = y^r, t = s/r$ when of course $t > 1$ and

$$(24) \iff \frac{1}{2}\left(a + (\frac{a^t + b^t}{2})^{1/t}\right) \ge \left(a^t + (\frac{a+b}{2})^t\right)^{1/t};$$

now putting $z = b/a$ we reduce (21) to the equivalent

$$\frac{1}{2}\left(1 + (\frac{1+z^t}{2})^{1/t}\right) \ge \left(\frac{1 + (\frac{1+z}{2})^t}{2}\right)^{1/t}.$$

It turns out that this can be reduced to (J) ; put $T(z) = (1 + z^t)/2$, and $S(z) = \dfrac{1 + z^{1/t}}{2}$, then the last inequality can be written as

$$\frac{S \circ T(1) + S \circ T(z)}{2} \ge S \circ T(\frac{1+z}{2}), \tag{25}$$

where $S \circ T(z) = S(T(z)) = (1/2)(1 + (1 + z^t)/2)^{1/t}$. Simple calculations show that $S \circ T'' \geq 0$, so that $S \circ T$ is convex and (25) is jus a case of (J).

In a similar way if $n = 2, r = 0 < s$ then, with $a_1 = x, a_2 = y$, (N) is just

$$\sqrt{x(\frac{x+y}{2})} \geq \frac{x + \sqrt{xy}}{2},$$

or, putting $z = y/x$

$$\sqrt{\frac{1+z}{2}} \geq \frac{1 + \sqrt{z}}{2},$$

which is immediate, or can be considered a case of the inequality (1;2).

Similar proofs can be given in the $n = 3$ case, but for the general case a new idea along the lines suggested in [13], is needed.

Finally let us see how (N) can imply other well known inequalities. The case of $r = 0, s = 1$ is

$$G_n(\underline{A}) \geq A_n(\underline{G}),$$

or

$$\sqrt[n]{a_1(\frac{a_1+a_2}{2})\cdots(\frac{a_1+\cdots+a_n}{n})} \geq \frac{a_1 + \sqrt{a_1a_2} + \cdots + \sqrt[n]{a_1a_2\cdots a_n}}{n}.$$

This can be rewritten as

$$\frac{n}{\sqrt[n]{n!}}\sqrt[n]{a_1(a_1+a_2)\cdots(a_1+\cdots+a_n)} \geq a_1 + \sqrt{a_1a_2} + \cdots + \sqrt[n]{a_1a_2\cdots a_n}. \tag{26}$$

By (G-A)

$$\sqrt[n]{a_1(a_1+a_2)\cdots(a_1+\cdots+a_n)} < a_1 + \cdots + a_n, \quad \text{by (G-A)},$$

so using (14) inequality (26) implies the weaker

$$a_1 + \sqrt{a_1a_2} + \cdots + \sqrt[n]{a_1a_2\cdots a_n} < e(a_1 + \cdots + a_n),$$

or, in series form

$$\sum_{n=1}^{\infty} G_n(\underline{a}) < e\sum_{n=1}^{\infty} a_n,$$

a result known as *Carleman's Inequality*.

§8. References

[1] E.F. Beckenbach & R.Bellman, *Inequalities* , Springer Verlag, Berlin-Heidelberg-New York, 1961.

[2] E.F. Beckenbach & R.Bellman, *An Introduction to Inequalities,* New Math. Lib., Random House, 1961.

[3] P.S.Bullen, D.S.Mitrinović & P.M.Vasić, *Means and Their Inequalities,* Reidel Publishing Co. Dordrecht–Boston, 1988.

[4] G.H.Hardy, J.E.Littlewood & G.Pólya, *Inequalities* , Camb. Univ. Press, Cambridge, 1934.

[5] N.Kazarinoff, *Analytic Inequalities,* New York, 1961.

[6] K.Knopp, *Theory and Application of Infinite Series,* Blackie & Son, 1928.

[7] P.P.Korovkin, *Inequalities,* New York, 1961.

[8] A.W.Marshall & I.Olkin, *Inequalities ; Theory of Majorization and Applications,* New York—London, 1979.

[9] Z.A.Melzak On the exponential function, *Amer. Math. Monthly,* 82 (19756), 842–844.

[10] D.S.Mitrinović; (with P.M.Vasić), *Analytic Inequalities,* Springer Verlag, Berlin-Heidelberg-New York,1970

[11] D.S.Mitrinović, J.E.Pečarić & A.M.Fink, *Classical and New Inequalities in Analysis* , Kluwer Academic Pub,1993.

[12] T.S.Nanjundiah, *Inequalities relating to arithmetic and geometric means I, II,, J. Mysore Univ. Sect. B6,* (1946), 63–77, 107–113.

[13] T.S.Nanjundiah, Sharpening some classical inequalities, *Math. Student,* **20** (1952), 24–25.

[14] T. Needham, A visual explanation of Jensen's inequality, *Amer. Math. Monthly,* 100 (1993), 768-771.

[15] I.Niven, Maxima and Minima without Calculus, *Math. Expos.* No.6, Math. Assoc. Amer., 1981.

[16] A.W.Roberts & D.E.Varberg, *Convex Functions,* Academic Press, New York-London, 1973.