INEQUALITIES

BY

EDWIN F. BECKENBACH

AND

RICHARD BELLMAN

INEQUALITIES

BY

EDWIN F. BECKENBACH

AND

RICHARD BELLMAN

WITH 6 FIGURES

SPRINGER-VERLAG
BERLIN · GÖTTINGEN · HEIDELBERG
1961

Preface

Since the classic work on inequalities by HARDY, LITTLEWOOD, and PÓLYA in 1934, an enormous amount of effort has been devoted to the sharpening and extension of the classical inequalities, to the discovery of new types of inequalities, and to the application of inqualities in many parts of analysis. As examples, let us cite the fields of ordinary and partial differential equations, which are dominated by inequalities and variational principles involving functions and their derivatives; the many applications of linear inequalities to game theory and mathematical economics, which have triggered a renewed interest in convexity and moment-space theory; and the growing uses of digital computers, which have given impetus to a systematic study of error estimates involving much sophisticated matrix theory and operator theory.

The results presented in the following pages reflect to some extent these ramifications of inequalities into contiguous regions of analysis, but to a greater extent our concern is with inequalities in their native habitat. Since it is clearly impossible to give a connected account of the burst of analytic activity of the last twenty-five years centering about inequalities, we have decided to limit our attention to those topics that have particularly delighted and intrigued us, and to the study of which we have contributed.

We have tried to furnish a sufficient number of references to allow the reader to pursue a subject backward in time or forward in complexity, but we have made no attempt to be encyclopedic in covering a field either in the text or in the bibliography at the end of the separate chapters.

As with most authors, we have imposed upon our friends. To KY FAN we extend our sincere gratitude for reading the manuscript through several times and for furnishing us the most detailed suggestions. For the reading of individual chapters and for many valuable comments and references, we wish to thank R. P. BOAS, P. LAX, L. NIRENBERG, I. OLKIN, and O. TAUSSKY.

Our hope is that the reading of this book will furnish as much pleasure to others as the writing did to us.

Los Angeles and Santa Monica, 1961

EDWIN F. BECKENBACH
RICHARD BELLMAN

Printed in Germany

Druck der Brühlschen Universitätsdruckerei Gießen

Contents

Chapter 1

The Fundamental Inequalities and Related Matters

§ 1. Introduction

In this initial chapter, we shall present many of the fundamental results and techniques of the theory of inequalities. Some of the results are important in themselves, and some are required for use in subsequent chapters; others are included, as are multiple proofs, on the basis of their elegance and unusual flavor [1].

We shall begin with the Cauchy inequality and the Lagrange identity, both of which will be substantially extended in this and the following chapter. From this we turn to a topic to which a monograph could be devoted in itself — namely, the famous inequality connecting the arithmetic and geometric means of n nonnegative numbers. Twelve proofs will be given of this basic result, not to suggest any lack of confidence in any single proof but rather to illustrate the wide range of techniques that the algebraist and analyst have at their disposal in treating inequalities. Of particular interest are the proofs of CAUCHY, HURWITZ, and BOHR.

Leaving this topic, albeit reluctantly, we shall establish the work-horses of analysis, the inequalities of HÖLDER and MINKOWSKI, in both discrete and continuous versions.

Subsequently, we shall establish some related, but more complex, results of BECKENBACH and DRESHER. These will be obtained with the aid of the important technique of quasi linearization, a method initiated by MINKOWSKI, developed by MAHLER, and used by YOUNG, ZYGMUND, and BELLMAN.

From this, we jump to the transformations of SCHUR involving doubly stochastic matrices, and to some results of KARAMATA, OSTROWSKI, and HARDY, LITTLEWOOD, and PÓLYA, pertaining to majorizing sequences. Continuous versions due to FAN and LORENTZ are also mentioned.

Our next port of call is in the domain of the elementary symmetric functions. Here, the results of MARCUS and LOPES are considerably more difficult to establish than might be suspected. Perhaps the most elegant proof of their inequalities is one that rests on the Minkowski theory of mixed volumes, a theory we shall discuss at length in our second volume on inequalities. Results due to WHITELEY are also presented.

From these matters, we turn to the fascinating questions of converses and refinements of the classical inequalities. Rather than follow the methods of BLASCHKE and PICK, and of BÜCKNER, or use moment-space arguments (the principal content of Chapter 3), we shall employ a method based on differential equations due to BELLMAN for establishing converse results. As far as the refinements are concerned, we shall merely mention some results and refer the reader to the original sources.

The last part of the chapter is devoted to some inequalities involving terms with alternating signs, discussed by WEINBERGER, SZEGÖ, OLKIN, BELLMAN, and others, all of which turn out to be particular cases of a novel inequality of STEFFENSEN.

§ 2. The Cauchy Inequality

The most basic inequality is the one stating that the square of any real number is nonnegative. To make effective use of this statement, we choose as our real number the quantity $y_1 - y_2$, where y_1 and y_2 are real. Then the inequality $(y_1 - y_2)^2 \geqq 0$ yields, upon multiplying out,

$$y_1^2 + y_2^2 \geqq 2 y_1 y_2. \tag{1}$$

The sign of equality holds if and only if $y_1 = y_2$. This is the simplest version of the inequality connecting the arithmetic and geometric means; following CAUCHY, we shall subsequently base one proof of the full result on this.

To make more effective use of the nonnegativity of squares, we form the sum

$$\sum_{i=1}^{n} (x_i u + y_i v)^2 = u^2 \sum_{i=1}^{n} x_i^2 + 2uv \sum_{i=1}^{n} x_i y_i + v^2 \sum_{i=1}^{n} y_i^2 , \tag{2}$$

where all quantities involved are real.

Since the foregoing quadratic form in u and v is nonnegative for all real values of u and v, its discriminant must be nonnegative, a fact expressed by the *Cauchy inequality* [1]:

$$\left(\sum_{i=1}^{n} x_i y_i\right)^2 \leqq \left(\sum_{i=1}^{n} x_i^2\right)\left(\sum_{i=1}^{n} y_i^2\right). \tag{3}$$

This inequality may be considered as expressing the result that, in euclidean space of any number of dimensions, the cosine of an angle is less than or equal to 1 in absolute value. Equality holds if and only if the sets (x_i) and (y_i) are proportional, that is, if and only if there are numbers λ and μ, not both 0, such that

$$\lambda x_i + \mu y_i = 0 , \qquad i = 1, 2, \ldots, n.$$

Still more general results can be obtained by applying the foregoing argument not merely to an n-dimensional euclidean space, but to a general linear space S possessing an inner product for any two elements x and y, written (x, y), with the following properties:

$$\begin{aligned} &\text{(a)} \quad (x, x) \geqq 0 \text{ for each } x \in S, \\ &\text{(b)} \quad (x, y) = (y, x), \\ &\text{(c)} \quad (x, uy + vw) = u(x, y) + v(x, w) \text{ for all real scalars } u \text{ and } v. \end{aligned} \tag{4}$$

These properties enable us to conclude that the quadratic form in u and v,

$$(ux + vy, ux + vy) = u^2(x, x) + 2uv\,(x, y) + v^2(y, y)\,, \tag{5}$$

is nonnegative for all real u and v.

Hence, as above, we obtain the inequality

$$(x, y)^2 \leqq (x, x)\,(y, y)\,, \tag{6}$$

a result that is, in turn, a particular case of more general results we shall derive in Chapter 2; see § 2.6.

A large number of results may now be obtained in a routine way by a choice of S and the inner product (x, y). Thus, we may take

$$(x, y) = \int_a^b x(t)\, y(t)\, dG(t)\,, \tag{7}$$

a Riemann-Stieltjes integral with $G(t)$ nondecreasing for $a \leqq t \leqq b$, or

$$(x, y) = \sum_{i,j=1}^{n} a_{ij} x_i y_j, \tag{8}$$

where $A = (a_{ij})$ is a positive definite matrix, and so on.

§ 3. The Lagrange Identity

A problem of much interest and difficulty with surprising ramifications is that of replacing any given valid inequality by an identity that makes the inequality obvious. The inequality (2.3) can be derived immediately from the identity

$$\left(\sum_{i=1}^{n} x_i^2\right)\left(\sum_{i=1}^{n} y_i^2\right) - \left(\sum_{i=1}^{n} x_i y_i\right)^2 = \sum_{\substack{i,j=1 \\ i<j}}^{n} (x_i y_j - x_j y_i)^2. \tag{1}$$

This also is a special case of a more general identity discussed in § 6 of Chapter 2.

§ 4. The Arithmetic-mean — Geometric-mean Inequality

We shall begin our consideration of results less on the surface by discussing what is probably the most important inequality, and certainly a keystone of the theory of inequalities — namely, the arithmetic-mean — geometric-mean inequality. The result, of singular elegance, follows:

Theorem 1. *Let $x_1\, x_2, \ldots, x_n$ be a set of n nonnegative quantities, $n \geqq 1$. Then*

$$\frac{x_1 + x_2 + \cdots + x_n}{n} \geqq (x_1 x_2 \ldots x_n)^{1/n}. \tag{1}$$

There is strict inequality unless the x_i are all equal.

Twelve proofs of this basic result will be presented in §§ 5—16, each based on a different principle or at least using a different device. There are a number of extensions of (1), involving weights. Amusingly enough, they are actually particularizations of the inequality, together with limiting cases. See § 14, below; a full discussion will be found also in [1.1].

§ 5. Induction — Forward and Backward

The following classical proof of Theorem 1 is due to CAUCHY [2.1]. As noted in (2.1), for any two quantities y_1 and y_2 we have

$$y_1^2 + y_2^2 \geqq 2 y_1 y_2 \,. \tag{1}$$

Setting $y_1^2 = x_1$, $y_2^2 = x_2$ in this last inequality, we obtain

$$\frac{x_1 + x_2}{2} \geqq \sqrt{x_1 x_2}\,, \tag{2}$$

valid for any two nonnegative quantities x_1 and x_2. Referring to (2.1), we see that equality holds if and only if $x_1 = x_2$.

Now replace x_1 by the new variable $(x_1 + x_2)/2$, and x_2 by $(x_3 + x_4)/2$. Then (2), together with its repetition, yields

$$\begin{aligned} \frac{x_1 + x_2 + x_3 + x_4}{4} &\geqq \left[\frac{(x_1 + x_2)}{2}\,\frac{(x_3 + x_4)}{2}\right]^{1/2} \\ &\geqq \left[(x_1 x_2)^{1/2} (x_3 x_4)^{1/2}\right]^{1/2} = (x_1 x_2 x_3 x_4)^{1/4}. \end{aligned} \tag{3}$$

Proceeding in this way, we readily see that we can establish the inequality (4.1) for $n = 1, 2, 4, \ldots$, and, generally, for n a power of 2. This is a *forward* induction.

Let us now use *backward* induction. We shall show that if the inequality holds for n, then it holds for $n - 1$. In (4.1), replace x_n by the value

$$x_n = \frac{x_1 + x_2 + \cdots + x_{n-1}}{n - 1}\,, \tag{4}$$

$n \geqq 2$, and leave the other x_i unchanged. Then, from (4.1), we obtain the inequality

$$\begin{aligned} &\frac{x_1 + x_2 + \cdots + x_{n-1} + \dfrac{x_1 + x_2 + \cdots + x_{n-1}}{n-1}}{n} \\ &\geqq (x_1 x_2 \ldots x_{n-1})^{1/n} \left(\frac{x_1 + x_2 + \cdots + x_{n-1}}{n - 1}\right)^{1/n}, \end{aligned} \tag{5}$$

or

$$\frac{x_1+x_2+\cdots+x_{n-1}}{n-1} \geqq (x_1 x_2 \ldots x_{n-1})^{1/n} \left(\frac{x_1+x_2+\cdots+x_{n-1}}{n-1}\right)^{1/n}. \tag{6}$$

Simplifying, we obtain

$$\left(\frac{x_1+x_2+\cdots+x_{n-1}}{n-1}\right) \geqq (x_1 x_2 \ldots x_{n-1})^{1/(n-1)}, \tag{7}$$

the desired inequality.

Combining the result for powers of 2 with this last result, we have an inductive proof of the theorem.

It is easy to see that the statement concerning strict inequality can also be established inductively.

Another interesting inequality that can be established by forward and backward induction is the following unpublished result due to Ky Fan:

"If $0 < x_i \leqq 1/2$ for $i = 1, 2, \ldots, n$, then

$$\frac{\prod_{i=1}^{n} x_i}{\left(\sum_{i=1}^{n} x_i\right)^n} \leqq \frac{\prod_{i=1}^{n} (1-x_i)}{\left[\sum_{i=1}^{n} (1-x_i)\right]^n}, \tag{8}$$

with equality only if all the x_i are equal."

§ 6. Calculus and Lagrange Multipliers

Let us now approach the arithmetic-mean — geometric-mean inequality as a problem in calculus. We wish to minimize the function $x_1+x_2+\cdots+x_n$ over all nonnegative x_i satisfying the normalizing condition

$$x_1 x_2 \ldots x_n = 1. \tag{1}$$

Since the minimum clearly is not assumed at a boundary point, we can utilize the Lagrange-multiplier approach to determine the local minima. For the function

$$f(x_1, x_2, \ldots, x_n) = x_1 x_2 \ldots x_n - \lambda (x_1+x_2+\cdots+x_n), \tag{2}$$

the variational equations

$$\frac{\partial f}{\partial x_i} = \frac{x_1 x_2 x_3 \ldots x_n}{x_i} - \lambda = 0, \quad i = 1, 2, \ldots, n, \tag{3}$$

yield the result that $x_1 = x_2 = \cdots = x_n$. From this we readily see that $x_i = 1/n$, $i = 1, 2, \ldots, n$, is the unique minimizing point, and thus we obtain the inequality (4.1).

§ 7. Functional Equations

Theorem 1 can also be established through the functional-equation approach of dynamic programming [1]. We begin with the problem of maximizing $x_1 x_2 \dots x_n$ subject to the constraints

$$x_1 + x_2 + \cdots + x_n = a, \quad x_i \geq 0 .$$

Denote this maximum value by $f_n(a)$, for $n = 1, 2, \dots$, and $a \geq 0$. In order to obtain a recurrence relationship connecting the functions $f_n(a)$ and $f_{n-1}(a)$, we observe that once x_n has been chosen, the problem that remains is that of choosing $x_1, x_2, \dots, x_{n-1}$ subject to the constraints

$$x_1 + x_2 + \cdots + x_{n-1} = a - x_n , \quad x_i \geq 0 , \tag{1}$$

so as to maximize the product $x_1 x_2 \dots x_{n-1}$.

It follows that

$$f_n(a) = \max_{0 \leq x_n \leq a} [x_n f_{n-1}(a - x_n)] , \quad n = 2, 3, \dots , \tag{2}$$

with $f_1(a) = a$.

The change of variable $x_i = a y_i$, $i = 1, 2, \dots, n$, enables us to conclude that

$$f_n(a) = a^n f_n(1) . \tag{3}$$

Using this functional form in (2), we see that

$$f_n(1) = f_{n-1}(1) \left[\max_{0 \leq y \leq 1} y(1-y)^{n-1} \right] = \frac{f_{n-1}(1)(n-1)^{n-1}}{n^n} . \tag{4}$$

Since $f_1(1) = 1$, it follows that $f_n(1) = 1/n^n$, which is equivalent to (4.1).

§ 8. Concavity

Let us now present a proof of Theorem 1 by means of a geometric argument [1, 2, 3, 4]. Consider the curve $y = \log x$, shown in Fig. 1. Differentiation shows that the curve is concave, so that the chord joining any two of its points lies beneath the curve. Hence, for $x_1, x_2 > 0$,

$$\log\left(\frac{x_1 + x_2}{2}\right) \geq \frac{\log x_1 + \log x_2}{2} , \tag{1}$$

with strict inequality unless $x_1 = x_2$.

This result is equivalent to

$$\frac{x_1 + x_2}{2} \geq \sqrt{x_1 x_2} . \tag{2}$$

The same reasoning shows (see page 17) that

$$\log\left(\frac{x_1+x_2+\cdots+x_n}{n}\right) \geqq \frac{\log x_1+\log x_2+\cdots+\log x_n}{n}, \tag{3}$$

for $x_1, x_2, \ldots, x_n > 0$, and, generally, that

$$\log\frac{\lambda_1 x_1+\lambda_2 x_2+\cdots+\lambda_n x_n}{\lambda_1+\lambda_2+\cdots+\lambda_n} \geqq \frac{\lambda_1\log x_1+\lambda_2\log x_2+\cdots+\lambda_n\log x_n}{\lambda_1+\lambda_2+\cdots+\lambda_n}, \tag{4}$$

for any combination of values $x_i \geqq 0$, $\lambda_i > 0$.

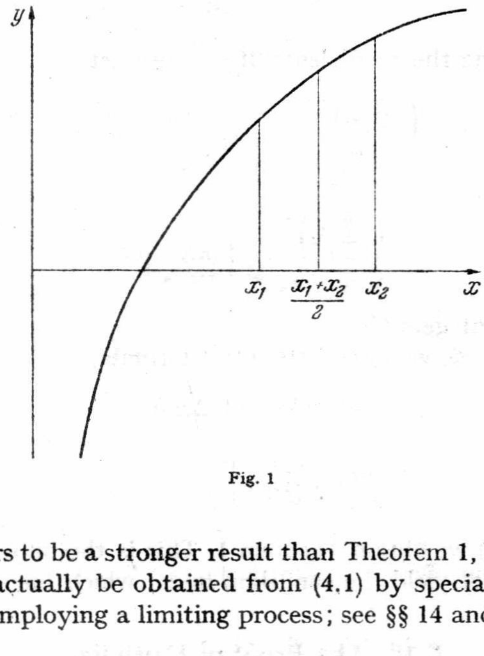

Fig. 1

This appears to be a stronger result than Theorem 1, but, as remarked in § 4, it can actually be obtained from (4.1) by specializing the values of the x_i and employing a limiting process; see §§ 14 and 16, below.

§ 9. Majorization — The Proof of Bohr

An amusing proof of Theorem 1 is due to H. Bohr [1].

To begin with, let us introduce the concept of majorization. Let $f(y)$ and $g(y)$ be two formal power series,

$$f(y) = \sum_{n=0}^{\infty} a_n y^n, \quad g(y) = \sum_{n=0}^{\infty} b_n y^n, \tag{1}$$

where $a_n, b_n \geqq 0$ for $n \geqq 0$.

If $a_n \geqq b_n$ for $n \geqq 0$, we write

$$f(y) \gg g(y) . \tag{2}$$

If $f_1(y) \gg g_1(y)$ and $f_2(y) \gg g_2(y)$, then clearly $f_1(y) f_2(y) \gg g_1(y) g_2(y)$.

Beginning with the obvious relationship

$$e^{xy} \gg \frac{x^N y^N}{N!} , \tag{3}$$

for $N = 1, 2, \ldots$, and $x, y \geqq 0$, we obtain

$$e^{y \sum_{i=1}^{n} x_i} \gg \frac{(x_1 x_2 \ldots x_n)^N y^{nN}}{(N!)^n} . \tag{4}$$

Hence, comparing the coefficients of y^{nN}, we get

$$\frac{\left(\sum_{i=1}^{n} x_i\right)^{nN}}{(nN)!} \geqq \frac{(x_1 x_2 \ldots x_n)^N}{(N!)^n} , \tag{5}$$

or

$$\frac{\left(\sum_{i=1}^{n} x_i\right)^{n}}{x_1 x_2 \ldots x_n} \geqq \left[\frac{(nN)!}{(N!)^n}\right]^{1/N} \tag{6}$$

for all positive integers N.

Since, as $k \to \infty$, we have STIRLING's formula,

$$k! \sim k^k e^{-k} \sqrt{2\pi k} , \tag{7}$$

we see that

$$\lim_{N \to \infty} \left[\frac{(nN)!}{(N!)^n}\right]^{1/N} = n^n . \tag{8}$$

From (6) and (8) we obtain Theorem 1. This is the only proof we shall give that does not yield the condition under which the sign of equality holds.

§ 10. The Proof of Hurwitz

Let us now present an interesting proof due to HURWITZ [1]. This result was published in 1891, six years before his famous paper on the generation of invariants by integration over groups [2], and one may see the germ of the later technique in his earlier analysis, which follows.

For the function $f(x_1, x_2, \ldots, x_n)$ of the n real variables $x_1, x_2, \ldots, x_n$, let us denote by $Pf(x_1, x_2, \ldots, x_n)$ the sum of f over the $n!$ quantities that result from all possible $n!$ permutations of the x_i. Thus

$$\begin{aligned} P x_1^n &= (n-1)! \, (x_1^n + x_2^n + \cdots + x_n^n) , \\ P x_1 x_2 \ldots x_n &= n! \, x_1 x_2 \ldots x_n . \end{aligned} \tag{1}$$

Consider the functions ϕ_k, $k = 1, 2, \ldots, n-1$, obtained in the following manner:

$$\begin{aligned} \phi_1 &= P\,[(x_1^{n-1} - x_2^{n-1})\,(x_1 - x_2)]\,, \\ \phi_2 &= P\,[(x_1^{n-2} - x_2^{n-2})\,(x_1 - x_2)\,x_3]\,, \\ \phi_3 &= P\,[(x_1^{n-3} - x_2^{n-3})\,(x_1 - x_2)\,x_3 x_4]\,, \\ &\vdots \\ \phi_{n-1} &= P\,[(x_1 - x_2)\,(x_1 - x_2)\,x_3 x_4 \ldots x_n]\,. \end{aligned} \tag{2}$$

We see that

$$\begin{aligned} \phi_1 &= P x_1^n + P x_2^n - P x_1^{n-1} x_2 - P x_2^{n-1} x_1 \\ &= 2 P x_1^n - 2 P x_1^{n-1} x_2\,. \end{aligned} \tag{3}$$

Similarly,

$$\begin{aligned} \phi_2 &= 2 P x_1^{n-1} x_2 - 2 P x_1^{n-2} x_2 x_3\,, \\ \phi_3 &= 2 P x_1^{n-2} x_2 x_3 - 2 P x_1^{n-3} x_2 x_3 x_4\,, \\ &\vdots \\ \phi_{n-1} &= 2 P x_1^2 x_2 x_3 \ldots x_{n-1} - 2 P x_1 x_2 \ldots x_n\,. \end{aligned} \tag{4}$$

Adding these results, we have

$$\phi_1 + \phi_2 + \cdots + \phi_{n-1} = 2 P x_1^n - 2 P x_1 x_2 \ldots x_n\,, \tag{5}$$

or, referring to (1),

$$\frac{x_1^n + x_2^n + \cdots + x_n^n}{n} - x_1 x_2 \ldots x_n = \frac{1}{2n!}\,(\phi_1 + \phi_2 + \cdots + \phi_n)\,. \tag{6}$$

It is easy to see that each of the functions $\phi_k(x)$ is nonnegative for $x_i \geqq 0$, since

$$\begin{aligned} \phi_k &= P\,[(x_1^{n-k} - x_2^{n-k})\,(x_1 - x_2)\,x_3 x_4 \ldots x_{k+1}] \\ &= P\,[(x_1 - x_2)^2\,(x_1^{n-k-1} + \cdots + x_2^{n-k-1})\,x_3 x_4 \ldots x_{k+1}]\,. \end{aligned} \tag{7}$$

Thus the difference appearing on the left-hand side of the identity (6) is nonnegative, whence Theorem 1 follows. This is the only proof we shall give that establishes the inequality (4.1) by means of an appropriate identity.

§ 11. A Proof of Ehlers

We shall prove Theorem 1 by showing that

$$x_1 x_2 \ldots x_n = 1,\ x_i \geqq 0,$$

implies that

$$x_1 + x_2 + \cdots + x_n \geqq n.$$

Assume that the result is valid for n, and let

$$x_1 x_1 \ldots x_n x_{n+1} = 1.$$

Let x_1 and x_2 be two of the x_i with the property that $x_1 \geqq 1$ and $x_2 \leqq 1$.