CHAPTER I

INTRODUCTION

1.1. Finite, infinite, and integral inequalities. It will be convenient to take some particular and typical inequality as a text for the general remarks which occupy this chapter; and we select a remarkable theorem due to Cauchy and usually known as 'Cauchy's inequality'.

Cauchy's inequality (Theorem 7) is

$$(a_1 b_1 + a_2 b_2 + \dots + a_n b_n)^2 \leqq (a_1^2 + a_2^2 + \dots + a_n^2)(b_1^2 + b_2^2 + \dots + b_n^2) \tag{1.1.1}$$

or

$$(\sum_1^n a_\nu b_\nu)^2 \leqq \sum_1^n a_\nu^2 \sum_1^n b_\nu^2, \tag{1.1.2}$$

and is true for all real values of $a_1, a_2, \dots, a_n, b_1, b_2, \dots, b_n$. We call $a_1, \dots, b_1, \dots$ the *variables* of the inequality. Here the number of variables is finite, and the inequality states a relation between certain finite sums. We call such an inequality an *elementary* or *finite* inequality.

The most fundamental inequalities are finite, but we shall also be concerned with inequalities which are not finite and involve generalisations of the notion of a sum. The most important of such generalisations are the infinite sums

$$\sum_1^\infty a_\nu, \quad \sum_{-\infty}^\infty a_\nu \tag{1.1.3}$$

and the integral

$$\int_a^b f(x)\,dx \tag{1.1.4}$$

(where a and b may be finite or infinite). The analogues of (1.1.2) corresponding to these generalisations are

$$(\sum_1^\infty a_\nu b_\nu)^2 \leqq \sum_1^\infty a_\nu^2 \sum_1^\infty b_\nu^2 \tag{1.1.5}$$

(or the similar formula in which both limits of summation are infinite), and

$$(1.1.6) \qquad \left(\int_a^b f(x)\,g(x)\,dx\right)^2 \leqq \int_a^b f^2(x)\,dx \int_a^b g^2(x)\,dx.$$

We call (1.1.5) an *infinite*, and (1.1.6) an *integral*, inequality.

1.2. Notations. We have often to distinguish between different *sets* of the variables. Thus in (1.1.2) we distinguish the two sets $a_1, a_2, \ldots, a_n$ and $b_1, b_2, \ldots, b_n$. It is convenient to have a shorter notation for sets of variables, and often, instead of writing 'the set $a_1, a_2, \ldots, a_n$' we shall write 'the set (a)' or simply 'the a'.

We shall habitually drop suffixes and limits in summations, when there is no risk of ambiguity. Thus we shall write

$$\Sigma a$$

for any of $\quad \sum_1^n a_\nu, \quad \sum_1^\infty a_\nu, \quad \sum_{-\infty}^{\infty} a_\nu;$

so that, for example,

$$(1.2.1) \qquad (\Sigma ab)^2 \leqq \Sigma a^2 \, \Sigma b^2$$

may mean either of (1.1.2) or (1.1.5), according to the context.

In integral inequalities, the *set* is replaced by a *function*; thus in passing from (1.1.2) to (1.1.6), (a) and (b) are replaced by f and g. We shall also often omit variables and limits in integrals, writing

$$\int f\,dx$$

for (1.1.4): so that (1.1.6), for example, will be written as

$$(1.2.2) \qquad \left(\int fg\,dx\right)^2 \leqq \int f^2 dx \int g^2 dx.$$

The ranges of the variables, whether in sums or integrals, are prescribed at the beginnings of chapters or sections, or may be inferred unambiguously from the context.

1.3. Positive inequalities. We are interested primarily in 'positive' inequalities[a]. A finite or infinite inequality is *positive* if all variables $a, b, \ldots$ involved in it are real and non-negative. An inequality of this type usually carries with it, as a trivial

[a] There are exceptions, as for example in §§ 8.8–8.17. There the 'positive' cases of the theorems discussed are relatively trivial.

corollary, an apparently more general inequality valid for all real, or even complex, a, b, ... Thus from (1.1.2) and the inequality

$$|\Sigma u| \leqq \Sigma |u|, \tag{1.3.1}$$

valid for all real or complex u, we deduce

$$|\Sigma ab|^2 \leqq (\Sigma |a||b|)^2 \leqq \Sigma |a|^2 \Sigma |b|^2, \tag{1.3.2}$$

where the a and b are arbitrary complex numbers. We shall usually be content to state our theorems in the fundamental 'positive' form and to leave the derived results to the reader. Occasionally, however, when the inequality in question is very important, we state it explicitly in its most general form.

Similar remarks apply to integral inequalities. The independent variable x will be real, but will (like the variable of summation ν) take either positive or negative values; while the functions $f(x)$, $g(x)$, ... will generally assume non-negative values only. To such an inequality as (1.1.6), true for non-negative f, g, corresponds the more general inequality

$$|\int fg\,dx|^2 \leqq \int |f|^2 dx \int |g|^2 dx, \tag{1.3.3}$$

valid for arbitrary complex functions f, g of the real variable x.

Numbers k, l, r, s, ... occurring as indices in our theorems are real but in general capable of either sign.

1.4. Homogeneous inequalities. The two sides of (1.1.2) are homogeneous functions of degree 2 of the a and also of the b; and generally both sides of our inequalities will be homogeneous functions, of the same degree, of certain sets of variables. Since homogeneous functions of positive degree vanish when all their arguments vanish, both sides, if of positive degree, will vanish, and so be equal, when the sets concerned consist entirely of 0's. Thus (1.1.2) reduces to an equality if all the a, or all the b, are 0.

A set consisting entirely of 0's is called a *null set*, or *the null set*, if the context is unambiguous. In general the '$\leqq$' or '$\geqq$' of our theorems will reduce to '$=$' when one or all of the sets involved is null. Sometimes this will be the *only* case of equality. More usually there will be other cases; thus plainly '$=$' occurs in (1.1.2) if every a is equal to the corresponding b. We shall be careful, wherever it is possible, to pick out explicitly such cases of equality.

The homogeneity of an inequality in certain sets of variables often enables us to simplify our proofs by imposing an additional restriction (a *normalisation*) on them. Thus the means $\mathfrak{M}_r(a)$ of § 2.2 are homogeneous, of degree 0, in the weights p, and we may always suppose, if we please, that $\Sigma p = 1$. Again, if we wish to prove that

$$(1.4.1) \qquad (a_1^s + a_2^s + \dots + a_n^s)^{1/s} \leqq (a_1^r + a_2^r + \dots + a_n^r)^{1/r}$$

when $0 < r < s$ (Theorem 19), we may suppose (since both sides are homogeneous in the a of degree 1) that $\Sigma a^r = 1$. We have then

$$a_\nu^r \leqq 1, \quad a_\nu^s = (a_\nu^r)^{s/r} \leqq a_\nu^r,$$

and so $\Sigma a^s \leqq \Sigma a^r = 1$. Without this preliminary normalisation, our proof would run

$$\frac{(\Sigma a^s)^{1/s}}{(\Sigma a^r)^{1/r}} = \left\{\Sigma \frac{a^s}{(\Sigma a^r)^{s/r}}\right\}^{1/s} = \left\{\Sigma \left(\frac{a^r}{\Sigma a^r}\right)^{s/r}\right\}^{1/s} \leqq \left(\Sigma \frac{a^r}{\Sigma a^r}\right)^{1/s} = 1.$$

There is another sense of 'homogeneity' which is sometimes important. Let us compare (1.4.1) above, which may be written as

$$(1.4.2) \qquad (\Sigma a^s)^{1/s} \leqq (\Sigma a^r)^{1/r},$$

with (1.1.2). Both inequalities are homogeneous in the variables, but (1.1.2) has a further homogeneity which (1.4.2) has not. It is, as we may say, 'homogeneous in Σ'; Σ, if treated as a number, would occur to the same power on the two sides of the inequality.

The result of this homogeneity in Σ is that (1.1.2) remains true if every *sum* which occurs is replaced by the corresponding *mean*, i.e. if written in the form

$$\left(\frac{1}{n}\Sigma ab\right)^2 \leqq \left(\frac{1}{n}\Sigma a^2\right)\left(\frac{1}{n}\Sigma b^2\right).$$

The importance of this kind of homogeneity will appear very clearly in § 2.10 and § 6.4. Roughly, an inequality which possesses it has an integral analogue, while one which does not, like (1.4.2), has none.

1.5. The axiomatic basis of algebraic inequalities[a]. Our subject is difficult to define precisely, but belongs partly to 'algebra' and partly to 'analysis'. Algebra or analysis, like geometry, may be treated axiomatically. Instead of saying, as

[a] See Artin and Schreier (1).

for example in Dedekind's theory of real numbers, that we are concerned with such or such definite objects, we may say, as in projective geometry, that we are concerned with any system of objects which possesses certain properties specified in a set of axioms. We do not propose to consider the 'axiomatics' of different parts of the subject in detail, but it may be worth while to insert a few remarks concerning the axiomatic basis of those theorems which, like (1.1.2) and most of the theorems of Ch. II, belong properly to algebra.

We may take as the axioms of an algebra only the ordinary laws of addition and multiplication. All our theorems will then be true in many different fields, in real algebra, complex algebra, or the arithmetic of residues to any modulus. Or we may add axioms concerning the solubility of linear equations, axioms which secure the existence and uniqueness of difference and quotient. Our theorems will then be true in real or complex algebra or in arithmetic to a *prime* modulus.

In our present subject we are concerned with relations of *inequality*, a notion peculiar to *real* algebra. We can secure an axiomatic basis for theorems of inequality by taking, in addition to the 'indefinables' and axioms already referred to, one new indefinable and two new axioms. We take as indefinable the idea of a *positive* number, and as axioms the two propositions:

I. *Either a is 0 or a is positive or $-a$ is positive, and these possibilities are exclusive.*

II. *The sum and product of two positive numbers are positive.*

We say that a is *negative* if $-a$ is positive, and that a is *greater* (*less*) than b if $a-b$ is positive (negative). Any inequality of a purely algebraic type, such as (1.1.2), may be made to rest on this foundation.

1.6. Comparable functions. We may say that the functions

$$f(a)=f(a_1, a_2, \ldots, a_n), \quad g(a)=g(a_1, a_2, \ldots, a_n)$$

are *comparable* if there is an inequality between them valid for all non-negative real a, that is to say if either $f \leq g$ for all such a or

$f \geqq g$ for all such a. Two given functions are not usually comparable. Thus two positive homogeneous polynomials of different degrees are certainly not comparable[a]; if $0 \leqq f \leqq g$ for all non-negative a, and both sides are homogeneous, then f and g are certainly of the same degree.

The definition may naturally be extended to functions $f(a, b, \ldots)$ of several sets of variables.

We shall be occupied throughout this volume with problems concerning the comparability of functions. Thus the arithmetic and geometric means of the a are comparable: $\mathfrak{G}(a) \leqq \mathfrak{A}(a)$ (Theorem 9). The functions $\mathfrak{G}(a+b)$ and $\mathfrak{G}(a)+\mathfrak{G}(b)$ are comparable (Theorem 10). The functions $\mathfrak{A}(ab)$ and $\mathfrak{A}(a)\,\mathfrak{A}(b)$ are not comparable; their relative magnitude depends upon the relations of magnitude of the a and b (Theorem 43). The functions

$$\psi^{-1}(\Sigma p \psi(a)), \quad \chi^{-1}(\Sigma p \chi(a))$$

are comparable if and only if $\chi\psi^{-1}$ is convex or concave (Theorem 85).

An important general theorem concerning the comparability of two functions of the form

$$\Sigma a_1^{\alpha_1} a_2^{\alpha_2} \ldots a_n^{\alpha_n},$$

due to Muirhead, will be found in § 2.18.

1.7. Selection of proofs. The methods of proof which we use in different parts of the book will depend on very different sets of ideas, and we shall often, particularly in Ch. II, give a number of alternative proofs of the same theorem. It may be useful to call attention here to certain broad distinctions between the methods which we employ.

In the first place, many of the proofs of Ch. II are 'strictly elementary', since they depend solely on the ideas and processes of finite algebra. We have made it a principle to give at any rate one such proof of any really important theorem whose character permits it.

Next we have, even in Ch. II, many proofs which are not elementary in this sense because they involve considerations of

[a] Compare § 2.19.

limits and continuity. We have also, particularly in Ch. IV, proofs which depend upon the standard properties of differential coefficients, as for example upon Rolle's Theorem. All these proofs belong to the elements of the theory of functions of one real variable.

Later, when dealing with integrals in Ch. VI, we naturally make use of the theory of measure and of the integral of Lebesgue. This we take for granted, but we give a summary in §§ 6.1–6.3 of the parts of the theory which we require.

Occasionally we appeal to the more remote parts of the theory of functions of real variables; but we do this only in alternative proofs or in the proofs of theorems of considerable intrinsic difficulty. Thus in Ch. IV (§ 4.6) we use the theory of the maxima and minima of functions of several variables; in Ch. VII we use the methods of the Calculus of Variations; and in Ch. IX we use the theory of double and repeated integration. We make no use of complex function theory, although, in the last chapters, we refer to it occasionally for purposes of illustration. The sections in which we do this do not belong properly to the main body of the book.

We add a few further remarks of a more detailed character.

(i) Cauchy's inequality (1.1.2) is a proposition of finite algebra, as defined in § 1.5. It is a recognised principle that the proof of such a theorem should involve only the methods of the theory to which it belongs.

(ii) We shall be continually meeting theorems, such as Hölder's inequality

(1.7.1) $$\Sigma ab \leqq (\Sigma a^k)^{1/k} (\Sigma b^{k'})^{1/k'}$$

(Theorem **13**), whose status depends upon the value of a parameter k. If k is *rational*, the theorem is algebraical, and our remarks under (i) apply. If k is irrational, a^k is not an algebraical function, and it is obvious that there can be no strictly algebraical proof.

It is however reasonable to demand, when we are concerned with an inequality so fundamental as Hölder's, that our step outside algebra shall be the absolute minimum which the nature of the problem necessitates. It is plain that this step will depend upon our definition of a^k. We may define a^k as $\exp(k \log a)$, and in this case it is obviously legitimate and necessary to use the theory of the exponential and logarithmic functions. If, as is more usual, we define a^k as the limit of a^{k_n}, where k_n is an appropriate rational approximation to k, then *this* limiting process should be the *only* one to which we appeal.

(iii) Suppose that, adopting the last point of view, we have proved Hölder's inequality, for rational k, in the form (1.7.1.). We can infer its truth for irrational k by a passage to the limit.

Such a proof, however, is not usually sufficient for our purpose. We always wish to prove a theorem of a more precise type than (1.7.1), in which (as in Theorem **13**) we establish strict inequality except in certain specified special cases. When we pass to the limit, '$<$' becomes '$\leqq$', we lose touch with the cases of equality (though these are in fact the same as in the rational case), and our proof is incomplete. It is therefore necessary to arrange our proofs in such a manner as to avoid such passages to the limit wherever it is possible. The same point arises whenever we wish to pass from a finite inequality to the corresponding infinite or integral inequality. It recurs at intervals throughout the volume and has often determined our choice of a particular line of proof.

(iv) The general principles which have governed our choice of methods are as follows. When a theorem is simple and fundamental, like Theorems **7**, **9**, or **11**, we prove it by several different methods, and are careful that one of our methods at any rate shall conform to the canons laid down under (i) and (ii). When the theorem is subsidiary or difficult, or when a proof satisfying these conditions would be troublesome or long, we use whatever method seems to us simplest or most instructive.

1.8. Selection of subjects. The principles which have guided us in our selection of *subjects* may be summarised as follows.

(i) The first part of the book (Chs. II–VI[a]) contains a systematic treatment of a definite subject. Our object has been to

[a] Except perhaps some parts of Ch. IV.

discuss thoroughly (with their analogues and extensions) the simple inequalities which are 'in daily use' in analysis. Of these three are fundamental, viz.

(1) the theorem of the arithmetic and geometric means (Theorem **9**),

(2) Hölder's inequality (Theorem **11**),

(3) Minkowski's inequality (Theorem **24**);

and these three theorems dominate the first six chapters. We prove them in a variety of ways, in the finite case in Ch. II, in the infinite case in Ch. V, and in the integral case in Ch. VI; while Ch. III (which contains a general account of the theory of convex functions) is mainly occupied with their generalisations. In these chapters, of which the most important are II, III, and VI, we have aimed at a comprehensive and in some ways exhaustive treatment.

(ii) The rest of the book (Chs. VII–X) is written in a different spirit and must be judged by different standards. These chapters contain a series of essays on subjects suggested by the more systematic investigations which precede. In them there is very little attempt at system or completeness. They are intended as an introduction to certain fields of modern research, and we have allowed our personal interests to dominate our choice of topics.

In spite of this (or because of it) the chapters have a certain unity. There is much modern work, in real or complex function-theory, in the theory of Fourier series, or in the general theory of orthogonal developments, in which the 'Lebesgue classes L^k' occupy the central position. This work demands a considerable mastery of the technique of inequalities; Hölder's and Minkowski's inequalities, and other more modern and more sophisticated inequalities of the same general character, are required at every turn. Our object has been to write such an introduction to this field of analysis as may be made to hang naturally on the subject matter of the early chapters.

(iii) We are interested primarily in certain parts of *real analysis*, and not in arithmetic or in algebra for its own sake. The line between algebra and analysis is often difficult to draw, especially in the theory of quadratic or bilinear forms, and we have often doubted what to include or reject. We have however excluded all developments whose main interest seemed to us to be algebraical.

We have also excluded function-theory proper, real or complex. In the later chapters, however, we have sometimes tried to show the significance of our theorems by sketching the lines of some of their function-theoretic applications.

Thus (to give definite examples) our programme excludes

(1) inequalities of a definitely arithmetical character, such as those of the theory of primes, or those which give bounds for forms with integral variables;

(2) inequalities which belong properly to the algebraical theory of quadratic forms;

(3) inequalities, such as 'Bessel's inequality', which belong to the theory of orthogonal series;

(4) inequalities, such as 'Hadamard's three circle theorem', which belong to function-theory proper:

and there is no systematic discussion of geometrical inequalities, though we use them frequently for purposes of illustration.

It may be useful to end this introduction by a few words of advice to readers who are anxious to avoid unnecessary immersion in detail. The subject, attractive as it is, demands, for the writer at any rate, a great deal of attention to details of a rather tiresome kind. These details arise particularly in the exclusion of exceptional cases, the complete specification of cases of equality, and the conventional treatment of zero and infinite values. Such a reader as we have in mind may be content, in general, to simplify his task as follows. (1) He may ignore the distinction between *non-negative* and *positive*, so that the numbers and functions with which he is concerned are all positive in the narrow sense. (2) He may ignore our conventions concerning 'infinite values'. (3) He may assume that the parameter k or r of inequalities such as Hölder's and Minkowski's is greater than 1.