Journal of Mathematical Analysis and Applications **249,** 118–134 (2000)
doi:10.1006/jmaa.2000.6934, available online at http://www.idealibrary.com on IDEAL®

An Essay on the History of Inequalities

A. M. Fink

Department of Mathematics, Iowa State University, Ames, Iowa 50011

Received February 14, 2000

We look at the rise of a discipline of "inequalities," trying to answer several questions. When did it become respectable to write a paper whose intent was the proof of an inequality? When were inequalities first given names? Who is responsible for the growth in the inequality literatures? What are the events that shaped the discipline? We give first answers to these questions. © 2000 Academic Press

Key Words: history; inequalities; Hardy.

1. INTRODUCTION

"Every mathematician loves an inequality" is part of the folklore of mathematics. It has not always been so. For example, one might peruse the many books on the history of mathematics without finding the word "inequality." But how is it that there are now two journals that are "journals of inequalities"? How is it that there are sections in the *Mathematical Reviews* and the *Zentralblatt für Mathematik* devoted to inequalities? Is there now a discipline called "inequalities"? Surely every analyst is well acquainted with a number of inequalities, but does one publish a paper whose sole purpose is to prove an inequality? The answer is a resounding yes and here is the story of how this developed. In short, in this paper I trace the development of a discipline known as "inequalities."

It is of course obvious that any paper that uses hard analysis will probably include some estimates of numbers, functions, or integrals. I want to loosely distinguish between an ad hoc inequality that is used in the proof of some theorem and an inequality that is in some sense general and can be applied in a different context. It is the latter that I have in mind. Of course, even this, is not precise. For example, is the inequality $1 + x \leq \exp(x)$ a general inequality or a special one, even though it holds for all real x?

0022-247X/00 $35.00

For this context, I would call it a particular inequality. A general inequality should be for a class of functions. The notion of function occurs rather late in the history of mathematics, so it is not surprising that few general inequalities predate, say, Newton or even Cauchy. (I exclude from our discussion inequalities in complex variables and results from approximation theory, where inequalities are everywhere dense.) Later I will make some comments on geometric inequalities and inequalities in functional analysis.

In a sense, the history of inequalities has already been written. The book "Inequalities" by Hardy *et al.* [1] organizes known inequalities and gives correct attribution to their source. The book is admittedly not comprehensive. Other books include those of Beckenbach and Bellman [2] and the early one of Mitrinović [3]. Each of these carefully delineates the origin of the inequalities discussed. However, more recently, as a crowning achievement of his mathematical career, D.S. Mitrinović undertook to publish and document *all* inequalities. This has lead to a five-volume set of books designed to be comprehensive and to give some idea of the evolution and application of the inequalities. He has had various collaborators in this effort; the present author was privileged to work on two of these volumes. The four in print [4–7], published in 1988–1993 will be followed by the fifth [8], to be published posthumously. It will consist of "particular inequalities." My story here will be mostly about some of those inequalities that appear in the first four volumes. This is not to say that particular inequalities are not important or that they will be excluded from these reflections.

2. A BRIEF OUTLINE

This is the history in broad strokes. Only a few inequalities came from the ancient traditions. Nothing much happened until Newton and a century later Cauchy. Four threads of inequalities emerged in the first eight decades of the 19th century. In the last two decades several inequalities were proved that had names attached to them and these inequalities began to be considered part of the basic knowledge of analysts. Moreover, certain papers whose sole purpose was to prove an inequality were published, including one by Chebyshev. Beginning in the third decade of this century, Hardy and some of his collaborators began to develop a systematic study of inequalities. This effort ended with the publishing of [1]. It became respectable to write papers on inequalities and there were very many of them. *Mathematical Reviews* began a subsection entitled "Inequalities." 26D, under *Real Variables* in 1980. The five-volume set by Mitrinović *et al.* now has thousands of entries in the bibliography. (Hardy, Littlewood, and Polya had 199.) Even this recent set is already out of date, with many new papers having appeared since its publication. In 1997 two new journals were begun for the dissemination of inequality articles.

3. THE ANCIENTS

Having given the broad outlines, we will now give more details in a rough chronological order. What did the ancients know? They knew the triangle inequality as a geometric fact. This is a general inequality since it applies to all triangles. A second general inequality in the geometrical context is the arithmetic-geometric mean inequality for two numbers recorded in Euclid. A geometric construction proves this inequality. A third inequality is what we now call the "isoperimetric inequality in the plane," which was known to Archimedes as well as earlier Greek mathematicians. The first rigorous proof awaited at least 2000 years of effort. Steiner (19th century C.E.) is given credit for a rigorous proof that no region except a circle could possibly give the maximum quotient. It was left to Weierstrass to complete the proof by showing that there exists an extremal region. It is fairly clear that the Hindu and Chinese mathematical traditions probably also knew these three inequalities.

How about the notion of inequality itself? Euclid, for example, could say that one area is larger than another using the words "falls short" of or "is in excess of," but no arithmetic of inequalities for numbers is indicated by any of the ancient traditions. The partial exception was efforts to approximate π, found in all of the ancient traditions. The best known is Archimedes' result that $223/71 < \pi < 22/7$. Van der Waerden [9, pp. 143ff] argues that Archimedes and others in the ancient traditions were able to improve such estimates as the result of an algorithm akin to continued fractions. The significant point is that in all these cases, along with the approximation being given, it was also known whether it was an upper or lower bound, thereby suggesting some knowledge of manipulation of inequalities. Early users of continued fractions realized that the successive approximations often were alternately upper and lower bounds. These properties could be viewed as a general inequality. Arabic mathematians seem to have used this principle most frequently (see Brezinski [10]).

4. THE MIDDLE YEARS

The next general inequality did not appear for many years. Perhaps it is a result of Newton: let p_r be the average of the elementary symmetric function of the positive quantities $a_1 \ldots a_n$ of order r, i.e., the average of the sums of all products of the a_i taken r at a time. Then Newton showed that $p_{r-1}p_{r+1} < p_r^2$ for $1 \leq r < n$ unless all of the a_i are equal. Maclaurin later observed that $p_1 > p_2^{1/2} > \cdots > p_n^{1/n}$. As Hardy, Littlewood, and Polya remark [1, p. 52], this latter result is a corollary of Newton's result. The extremes of this sequence are the arithmetic mean and the geometric mean, establishing a now famous inequality. Newton missed this; Cauchy's famous induction and descent proof of the arithmetic-geometric mean inequality for n numbers is usually the first one cited. One might infer that maybe he was the first to write it down and prove it. It isn't so. Maclaurin proved it in 1729. References [1–3], for example, all make the correct attribution but it is easy to miss it.

Maclaurin plays a prominent role in the field of inequalities, but he did not originate named general inequalities. I am indebted for this observation to the fine article by Grabiner [11]. Maclaurin gave what amounts to epsilon–delta proofs for various limits and there are strong indications that this had an influence on the continental mathematicians who were beginning to use inequality-based proofs for analysis.

Interestingly enough, the century or so between Maclaurin and Cauchy did not give rise to inequalities. No inequality I know is due to Euler, for example. The single inequality from these times appears to be Bernoulli's Inequality, which routinely appears in advanced calculus texts and some undergraduate texts. I know no reference for this inequality and do not know which Bernoulli is to be credited or how it got its name.

As mentioned above, Cauchy is usually cited for his ingenious proof of the arithmetic-geometric mean inequality found in his course notes on analysis. The same note (Note II in [12]) begins by proving the usual elementary rules of arithmetic of inequalities, and includes the proof of Cauchy's Inequality, the finite sum version of what is now called the Cauchy–Schwarz–Buniakowski Inequality (Schwarz, 1884, gave the integral version, and Buniakowski, 1859, both versions). The note ends with the arithmetic-geometric mean inequality.

Cauchy's book on analysis is dated 1821. The next named inequality appears some time in the last two decades of the 19th century. What happened in between? Grabiner [11] argues that people were largely not greatly interested in mathematics outside the applicable areas during the 18th century. This is also true of the first few decades of the 19th century. Not much abstract mathematics was done; in particular, no one was interested in an inequality for its own sake. There were, however, four threads of inequality results in the period between Newton and approximately 1900.

The first thread was attached to Newton's inequality and the Maclaurin extension. A number of papers looked at relationships between the symmetric functions, viewed in relation to roots of polynomials. Papers by Fort, Campbell, Sylvester, C. Smith, Hamy, and Darboux appeared during this time. About 1890 these results began to appear in algebra books, such as Weber's in 1896. See [4] for references and commentary.

The second thread was means; see [4] again. There are about 15 references for the 19th century. Prominent among these are the 1840 paper by Bienayme who first considered power means with weights, Schlömlich's long and comprehensive paper of 1858, and Winckler's papers of 1860 and 1866, which dealt with moments and integral means. Schiaparelli wrote on means in 1868 and 1871, but a later paper published in 1907 was the first one that gave axioms for means. Muirhead published several papers on means and has a theorem named after him from 1903; see [1, p. 44]. About 20 more papers on means appeared in the first two decades of the present century. Now the number of papers is enormous; see [4] for evidence. It is also interesting that for Bullen *et al.* [4], means are discrete means of n numbers, integral means being mentioned only in passing in the last seven pages of the book.

The third thread is related to the arithmetic–geometric mean inequality. As one might guess, everybody had his favorite proof of the arithmetic-geometric mean inequality. Among the 25 or so papers in the 19th century, there are several proofs, including ones by Cauchy, Goursat, Liouville, and Darboux, to give some familiar names. The tradition of giving "new" proofs continues, the book [4] containing 52 of them. Also prominent is the algorithm of Gauss for computing certain elliptic integrals, the so-called geometric-arithmetic algorithm defined by $a_{n+1} = (a_n + b_n)/2, \quad b_{n+1} = (a_n b_n)^{1/2}$. Various authors wrote about it and gave variants.

The last thread is for the isoperimetric inequality: If IQ = area/(perimeter)^2, then what figure among some class has the highest IQ? Kazarinoff [13] claims that Euclid knew that the square solves the isoperimetric problem for rectangles, probably quoting some earlier mathematicians (Pythagoras?). A book, "Isoperimetric Figures" by Zendoros is lost. The results are recovered in Pappus' works. Zendoros lived in the first century before or after the common era and Pappus dates about 300 C.E. Lhuilier [14] and Steiner [15] gave arguments that the circle gave a better isoperimetric quotient than any polygon. Weierstrass is credited with pointing out that the existence of an extremal was needed to make Steiner's argument into a rigorous proof that the circle has the highest IQ. Historical remarks on the isoperimetric problem are given in Bonnesen and Fenchel, (see [16, p. 117] (original German version 1934)), where most of the rigorous results are given dates after 1900. Other references to history articles are given there as well.

5. NAMED INEQUALITIES

Now we come to the decades of named inequalities. "Named inequalities" may be a tricky term because it requires two things. First, an inequality must be attributable to someone who stated it or proved it. Further, someone must judge that it is important enough to be referred to by name and others must find attribution important. The propensity to do this varies from mathematician to mathematician and this propensity may be a decreasing function of time. All the books that we have previously referenced, [1–7], are careful to try to meticulously attribute results to the correct person and to establish priority, although this is difficult when a special case of the general inequality appears earlier than the general named inequality.

The most obviously important named inequalities are those of Hölder [17] and Minkowski [18]; but the watershed paper, in my estimation, is the paper of Chebyshev [19]. This paper was submitted to the editorial committee of the Han'kovshov University. Journal for the volumes in 1883. But the editorial committee found that it was such an exciting paper that they placed it in the last volume of 1882. This paper contains statements of a sequence of inequalities, the first of which is now given the name "Chebyshev's Inequality." A paper in 1883 [20] contains the proofs. If f, g, and p are integrable functions (Riemann here) with f and g having derivatives that do not change sign and the sign in both cases is the same and $p \geq 0$, then

$$\int_a^b f(x)g(x)p(x)dx \int_a^b p(x)dx \geq \int_a^b f(x)p(x)dx \int_a^b g(x)p(x)dx.$$

As mentioned above, priority is a debatable issue. Chrystal [21, p. 50] credits Laplace (1749–1827) with a discrete version for two positive decreasing sequences. Winckler [22] gave a version with $p(x) = 1$. Hermite [23] gave a short proof due to Picard in his course. Korkine [24], in a letter to Hermite, gave a proof for the discrete case based on an identity, while Andreief [25] gave another identity that proved the statement given above. In fact, he showed that the proper hypothesis was that $[f(x) - f(y)][g(x) - g(y)] \geq 0$ for all x and y (that is, f and g are similarly ordered). These papers on inequalities attracted a lot of immediate attention, some by well–known (to us) mathematicians but written without any apparent application in mind. Chebyshev himself, however, did have stated applications to probability in mind. More on the history of this inequality and the false conclusion on priority are give in [7, pp. 240ff]. The chapter on Chebyshev's Inequality in [7] contains 217 references!

Inequalities became a respectable topic for a paper. Hadamard [26] wrote a paper on determinants and their inequalities, one of which today bears his name, to wit

$$\det \{x_{ij}\} \leq (\sum_{ij}(x_{ij})^2)^{1/2}.$$

He wrote this paper without any application in mind. In fact, he missed an important application. He did not think the result was overly important but he was wrong. It was the basic tool for Fredholm's theory of integral equations. Hadamard wrote, "I had been attracted by a question on determinants in 1893. When solving it, I had no suspicion of any definite use it might have, only feeling that it deserved interest; then in 1900 appeared Fredholm's theory, for which the result obtained in 1893 happens to be essential. This is the theory I failed to discover. It has been a consolation for my self-esteem to have brought a necessary link to Fredholm's argument" [27].

One other named inequality appeared in the 19th century. Gram [28] proved that $\det\{(x_i, x_j)\} \geq 0$, where the inner product is in R^n. This, of course, has been generalized, and other inequalities for the Grammian have been given; see [7].

Perhaps the most widely known and widely used inequality is Jensen's inequality,

$$f\Big(\sum_1^n r_i x_i\Big) \leq \sum_1^n r_i f(x_i),$$

whenever the weights r_i are positive and add to 1 and f is a convex function. Jensen [29] was the first to define convex functions. His definition was that $f((x+y)/2) \leq (f(x)+f(y))/2$, so he was restricted to using rational weights. (For continuous functions this is equivalent to the general two-weight case.) This inequality has been generalized to its familiar form in integrals, without the weights being necessarily rational or positive. This is another instance in which special cases were proved earlier. Hölder [17] proved this inequality for functions for which f'' exists and is non-negative (not giving such functions a special name). I believe that priority for Jensen's inequality should be given to Hölder and it should be named after him, but this will not happen. Hadamard also had his hand in this circle of ideas. In [30], he showed that if f has an increasing derivative then

$$f((x+y)/2) \leq \frac{1}{y-x}\int_x^y f(t)dt.$$

This inequality is also capable of being generalized and better understood; see [31], for example.

6. MORE NAMED INEQUALITIES

The era of named inequalities continues with several more inequalities established in the first two decades of the present century. The first was Hilbert's double series inequality,

$$\sum_{m,n=1}^{\infty} \frac{a_m b_n}{m+n} \leq \pi\Big(\sum_1^{\infty} a_m^2\Big)^{\frac{1}{2}}\Big(\sum_1^{\infty} b_n^2\Big)^{\frac{1}{2}}.$$

Hilbert first proved this inequality in his course on integral equations, without specifying the constant on the right. Weyl [32] published a proof in his dissertation of 1908. Schur [33] gave the correct best constant π. Hardy and Marcel Riesz later gave p-norm versions of both the discrete and continuous versions [34]. See [6] for more details.

Hardy is no doubt the major player in the story. His view of "pure mathematics" as the ultimate in mathematics was germane to his interest in inequalities. It is fitting that one of the named inequalities bears his name. In [35] in research aimed at giving a simple proof of Hilbert's inequality cited above, he found the following two theorems:

$$\text{If } p > 1, a_n \geq 0, \text{ and } A_n = a_1 + \cdots + a_n, \text{ then}$$

$$\sum_1^\infty (A_n/n)^p < C \sum_1^\infty a_n^p.$$

$$\text{If } p > 1, f \geq 0, \text{ and } F(x) = \int_0^x f(t)dt, \text{ then}$$

$$\int_0^\infty (F(x)/x)^p dx < C \int_0^\infty f^p dx.$$

The constant C, being $(p/(p-1))^p$, is the same in both cases. Hardy supplied it for the integral version and Landau [36] supplied it for the discrete version. Hardy's inequality has proven to be very important in singular integral theory and its extensions and applications have been deeply studied. A book by Opic and Kufner [37] is devoted to this study. This is one of several cases in which single inequalities have inspired books.

Two inequalities, which continue to spawn a great number of papers and for which there are still many open problems, are theorems of Landau [38]: *If f and f'' have absolute values less than one on an interval whose length is not less than 2 then $|f'(x)| \leq 2$ and 2 is the best constant. If $M_i = \sup|f^{(i)}|$ where the* sup *is taken over R, then $M_0 = M_2 = 1$ implies that $M_1 \leq 2^{1/2}$ and this is the best possible constant.*

Of course, the first result is of the same form if the sup is taken over the appropriate interval but the constant depends on the size of the interval. Hadamard [39] formulated this result as the homogenized inequality

$$M_1 \leq 2M_0^{1/2}M_2^{1/2}$$

and its generalization

$$M_k \leq C(k, n, p)M_0^{1-k/n}M_n^{k/n}$$

and showed that $C(1, n, \infty) \leq 2^{n-1}$. Here p is the norm that is being used. Hardy and Littlewood [40] showed that these constants exist and later found exact values for $p = 2$ and the interval either R or $[0, \infty)$. Landau [41] showed the usefulness of these results for differential equations. The most important open problem was the determination of the exact constants. The first major breakthrough came in 1938 when Kolmogorov in [42,43] completely solved the problem for $p = \infty$ and the interval R. What is amazing is that for each n and each k, the extremal $y^{(n)}$ is the same function, the square wave. The problem for the interval $[0, \infty)$ (named Landau's problem) turned out to be considerably more difficult. The answer was found by Schoenberg and Cavaretta [44] in 1970. The book [6] listed 217 references on this inequality and its generalizations (including function spaces). The monograph by Kwong and Zettl [45] is all about the problems of computing the constants for various p and n and gives a good indication of the remaining open problems.

Gronwall [46] proved that

If $0 \leq x(t) \leq \int_0^b (a + bx(s))ds$ *then on the same interval* $x(t) \leq at \exp(bt)$, *if a and b are positive constants.*

Bellman proved his lemma [47] in 1943:

If x and k are positive functions and a is a positive constant then $x(t) \leq a + \int_a^t k(s)x(s)ds$, $t \geq c$ *implies that* $x(t) \leq a \exp\left(\int_a^t k(s)ds\right)$ *on the same interval.*

For some time Bellman's lemma was the standard reference, although there is evidence that the lemma was also proved earlier by Reid. In any case, improvements were quickly made. The function $x(t)$ and the constant need not be non-negative for the result to hold. One only has to change the proofs. The version where a is replaced by a function is usually called Gronwall's inequality in differential equation papers. The higher dimensional versions (for example, a double integral for a function of two variables) are called Wendroff inequalities by Beckenbach and Bellman [2, p. 154], although Wendroff never published it (and again the version has non-negative hypotheses that are not required). Books on differential inequalities routinely use a lot of time proving and using Gronwall's type inequalities. See, for example, Walter [48] or Bainov and Simeonov [49] for some recent books.

Various other named inequalities merit comment. Steffensen's inequality, proved in [50], was somehow missed by Hardy *et al.* [1]. This inequality gives an estimate for the constant in some versions of the mean value theorem for integrals.

If f is a decreasing function and $0 \leq g(t) \leq 1$ *then*

$$\int_{b-\lambda}^{b} f(t)dt \leq \int_a^b f(t)g(t)dt \leq \int_a^{a+\lambda} f(t)dt,$$

where $\lambda = \int_a^b g(t)dt$. Gauss [51], in his theory of errors, had an earlier specific version of this. See [7] for more on this.

Inequalities of Young [52] and Lypanov [53] were other named inequalities that came from this era. Another named inequality, although not from this era, has generated an entire book [54]. This is Opial's inequality [55],

$$\textit{if } f(0) = f(a) = 0 \textit{ and } f \geq 0, \textit{then}$$

$$\int_0^a |f(x)f'(x)|dx \leq (a/4) \int_0^a f'(x)^2 dx.$$

Wirtinger's inequality must also be mentioned as a significant named inequality. Its history has an unusual twist. Wirtinger may not have proved it and we do not know why the name has stuck. The first mention of it is in the book by Blaschke [56]. The inequality is one of the form

$$\int f^2 dx \leq C \int (f')^2 dx$$

under various boundary conditions which include zeros of the function or the average of f being zero and the interval varies from $[0, \pi/2]$, $[0, \pi]$, or $[0, 2\pi]$. The book [1] gives three versions and calls all of them "Wirtinger's Inequality." But the history is more complicated; see [6, pp. 65ff]. Several earlier versions were missed by Hardy, Littlewood, and Polya. In fact, the earliest one seems to be Ahmansi [57] in 1905. This inequality can be generalized by replacing the first derivative by some higher order derivative with various n-dimensional boundary conditions, and finally by replacing the 2-norms by general norms. Brink [58] and Fink [59, 60] have the beginnings of a general theory.

7. THE WATERSHED EVENT

Having written about the first two decades of the present century, a bit about the third, and having talked about specific named inequalities, we now turn to what is the second watershed event in the history of inequalities. This is the retiring presidential (London Mathematical Society) address of G. H. Hardy on November 8, 1928, "Prolegomena to a Chapter on Inequalities" [61] I recommend it to everyone, and quote extensively here.

After pointing out various advances of the Society under his presidency, he says (Insert reprinted with the permission of the London Mathematical Society), *"The most important event in the history of the Society during this period has been the foundation of the* Journal, *and for this at any rate I could reasonably claim a good share of the credit. I had been anxious for many years, as my colleagues in the Cambridge Philosophical Sociey may remember, to see an English mathematical periodical run on some such lines, and now that the* Journal *has come into existence, I am indeed delighted to think that it is an obvious success."* He added, *"If the* Journal *is open to criticism, it will be on the ground that it is one-sided, that subjects are not properly represented, and in particular that it is overweighted with 'analysis'. Its most striking feature has been a series of papers, twenty or more by now, on elementary inequalities and series of positive terms.*

For my own part I do not regret that the Journal *should show a tendency to specialise in one or two particular directions. It is not at all a bad thing for a new periodical to gain the reputation of being particularly interesting on some special subject. In this case it is quite obvious, from the foreign contributions which we receive, that the* Journal *is already regarded as a particularly appropriate medium for the publication of notes on inequalities. The subject is 'bright' and amusing, and intelligible without large reserves of knowledge; and it affords unlimited opportunities for that expertness in elementary technique which is supposed, rightly or wrongly, to be one of the characteristic results of English mathematical education. If then the* Journal *is one-sided, it is one-sided in a way which I like."* [He goes on to say that the fundamental inequalities are elementary and adds], *"Anyone who has read the* Journal *at all regularly will realise the great amount of thought and ingenuity that has been expended in recent years in this apparently restricted field.*

I would say in passsing that this ingenuity has certainly not been wasted. A thorough mastery of elementary inequalities is today one of the first necessary qualifications for research in the theory of functions, at any rate, in function theory of the 'hard, sharp, narrow' kind as opposed to the 'soft, large, vague' kind (I do not use any of these adjectives as words either of praise or blame), the function-theory of Bohr, Landau, or Littlewood, as opposed to the function-theory of Birkhoff or Koebe. It is essential to anyone working in this field to be master both of the main results and of the tricks of the trade." [He then talks of the Cauchy–Schwarz inequality and its various extensions.] *"In this particular case, of course, everything is easy;.... There are, however, plenty of inequalities which are really hard to prove; Littlewood and I have had any amount of practice during the last few years, and we have found quite a number of which there seems to be no really easy proof. It has been our unvarying experience that the real crux, the real difficulty of idea, is encountered at the very beginning. It is very curious indeed how in this field the old-fashioned 'tripos trick' comes into its own again. There was a time, perhaps, when undertanding was what an analyst needed most.*

The elementary inequalities thus form the subject-matter of one of the first fundamental chapters in the theory of functions. But this chapter has never been properly written; the subject is one of which it is impossible to find a really scientific or coherent account. I think that it was Harald Bohr who remarked to me that 'all analysts spend half their time hunting through the literature for inequalities which they want to use and cannot prove'. I will give a few examples. [He then goes on to cite Hölder's and Minkowski's inequalities along with the arithmetic-geometric mean inequality.] *To find all three theorems in one volume is, I think, strictly impossible. Even Pólya–Szegö, which is better on inequalties than any other book I know, does not include the general form of Hölder's inequality.*

For these reasons Littlewood, Pölya, and myself have undertaken to contribute a tract on inequalities to the Cambridge series, and I am sure that we shall deserve the thanks of the mathematical world, even if we do not do it particularly well.

> [Hardy discusses a few other inequalities and concludes with,] *Any serious student of inequalities ought to read the* Journal*: I have confined myself in this summary to quite familiar theorems, but I hope that I have said enough to make this clear. The obligation will seem the more pressing when we remember that the* Journal *contains quite a number of new and very interesting inequalities to which I have never referred. I admitted at the beginning that we have published a great number of papers in a very special field; I am so far from regretting it that I hope that we shall publish very many more.*

Was Hardy's hope realized? The rate at which the *Journal* published inequality papers after Hardy's talk actually decreased after 1929, but it increased slightly again after the publication of [1] in 1934. But what happened is told by other journals and books. When Beckenbach and Bellman published their book [2] in 1961, the bibliography for Chapter 1 included 21 titles whose publication date preceded publication of the book [1] and 72 published later. This first chapter is about the elementary and classical inequalities so a fair number of old papers are expected to be referenced. However, the remaining chapters, most of which contain more modern inequalities, have very few references to early papers. Most of the references are to papers written after 1940. The book [3], published in 1970 has about the same division between old and new references. Clearly something happened after 1934 and what happened is that lots of people proved inequalities and recorded them in the mathematical literature.

The book Hardy promised appeared in 1934, and as they say, the rest is history. I would be hard pressed to find a paper or a book which has been more frequently cited (and continues to be cited). One of the interesting aspects of the book is the philosophy of inequalities, presented in the introduction: generally an inequality that is elementary should be given an elementary proof, the proof should be "inside" the theory it belongs to, and finally the proof should try to settle the cases of equality. This introductory chapter is recommended reading with ideas that are still applicable today.

The paper [62] attempts a beginning of an addition to the philosophy of inequalities, following up on [63, 64]. That is, one should try to state and prove an inequality so that it cannot be generalized. Whether this is fruitful remains to be seen.

8. OTHER INEQUALITIES

Two other inequalities are the bases of books. One is found in Hardy [65], in which the inequality of Muirhead and a paper by Schur are combined to prove the first theorem using majorization. This is the theorem that *a necessary and sufficient condition that*

$$\sum_1^n f(x_i) \le \sum_1^n f(y_i) \qquad \textit{for all convex } f$$

is that the vector x is majorized by the vector y, i.e., that the partial sums of the decreasing rearrangement of x are less than or equal the corresponding partial sums of the decreasing rearrangement of y, with the sums of all components being equal. This inequality cannot be generalized in the Fink and Jodeit sense because "for a fixed (piecewise continuous) f the inequality holds for all x majorized by y if and only if f is convex." The concept of majorization has been fruitful. There are very many applications of this idea; see the wonderful book by Marshall and Olkin [66].

The other notable inequality that I would like to mention is Shannon's inequality [67], which is the basis of information theory:

$$\sum_1^n p_i \log p_i \geq \sum_1^n p_i \log q_i \qquad \text{unless } p = q.$$

Here p and q are probability vectors.

I now want to discuss some other lines of enquiry in inequalities. First there are geometric inequalities, the first of which, the isoperimetric inequality, we have mentioned earlier. These inequalities are relations between the various elements of a triangle, for example. The book [5] "Recent Advances in Geometric Inequalities" has an astonishing number of inequalities. However, the bibliography only contains about ten entries published before 1930. The authors mention that most of these results were rediscovered more recently. One gets the impression from this book that the majority of the papers cited were published after the appearance of the book "Geometric Inequalities" [70], by Bottema *et al.* in 1969. I think that here is a clear case where a single publication was very instrumental in the blossoming of an area of research.

A second line of inequalities is those associated with matrices. The book "Matrix Inequalities" [71] by Marcus and Ming in 1964 has only a handful of inequalities that were proved earlier than 1940. It is unclear what prompted the later work.

The third set of inequalities one might mention are those that center around the concrete functional analysis of singular integrals. Hardy and Littlewood's maximal function inequalities were proved in 1930 and Marcinkiewicz's weak type inequalities also were proved in the 1930s. But the Calderone–Zygmund theory was first published in 1950. See Stein's book [72] for an introduction to a theory that is based almost entirely on inequalities.

Finally there are the inequalities associated with partial differential equations. The maximum principle of the Laplacian was known to Gauss, for example, and various other versions of maximum principles were known early in the 20th century. See Protter and Weinberger [73] for a exposition of this line of inequalities. Poincaré proved an inequality in 1894 that bears his name. See [6, Chap. II]. It is a two-dimensional version of Wirtinger's

Inequality relating the norm of a function to the norm of its gradient. Then there are the functional analysis inequalities of Sobolev first published by him in 1938. Much of the modern theory and study of partial differential equations centers around the correct spaces in which to do the analysis and these are related to embedding theory and inequalities of the Sobolev type. See Adams [74] for an introduction to these inequalities.

9. CONCLUSION

The latest chapter of this story is now the formation of two journals devoted to inequalities and their applications. They are *Journal of Inequalities and Applications* [68] with the first volume in 1997 and *Mathematical Inequalities and Applications* [69] with the first volume in 1998. The editorial boards of these two journals are largely disjoint.

In conclusion, I have sketched the history of inequalities from meager beginnings in ancient times to the awakening of inequality analysis in the early 18th century. The two watershed events were the papers of Chebysev in 1882 and the presidential address of Hardy in 1928 followed by the publication of [1]. I am not a professional historian or a professional historian of mathematics, and I hope that the reader will have found this essay entertaining enough to comment on it and perhaps improve on it on another occasion.

REFERENCES

1. G. H. Hardy, J. E. Littlewood, and G. Polya, “Inequalities,” Cambridge Univ. Press, Cambridge, UK, 1934.
2. E. F. Beckenbach and R. Bellman, “Inequalities,” Springer-Verlag, New York/Berlin, 1961.
3. D. S. Mitrinović, “Analytic Inequalities,” Springer-Verlag, New York/Berlin. 1970.
4. P. S. Bullen, D. S. Mitrinović, and P. M. Vasić, “Means and Their Inequalities,” Kluwer Academic, Dordrecht, 1988.
5. D. S. Mitrinović, J. E. Pečarić, and V. Volence, “Recent Advances in Geometric Inequalities,” Kluwer Academic, Dordrecht, 1989.
6. D. S. Mitrinović, J. E. Pečarić, and A. M. Fink, “Inequalities Involving Functions and Their Integrals and Derivatives,” Kluwer Academic, Dordrecht, 1991.
7. D. S. Mitrinović, J. E. Pečarić, and A. M. Fink, “Classical and New Inequalities in Analysis,” Kluwer Academic, Dordrecht, 1993.
8. D. S. Mitrinović *et al.*, to appear.
9. B. L. van der Waerden, “Geometry and Algebra in Ancient Civilizations,” Springer-Verlag, New York/Berlin, 1983.
10. C. Brezinski, “History of Continued Fractions and Pade Approximates,” Springer-Verlag, New York/Berlin, 1991.
11. J. V. Grabiner, Was Newton’s calculus a dead end? The continental influence of Maclaurin’s treatise of fluxions *Amer. Math. Monthly* **104** (1997), 393–410.
12. A. L. Cauchy, “Cours d’analyse de l’École Royale Polytechnique,” I^re partie, Analyse algébrique, Paris, 1821.

13. N. D. Kazarinoff, “Geometric inequaliites,” Yale Univ. Press, New Haven, CT, 1961.
14. S. Lhuilier, “Polygonométrie, ou de la mesure des figures rectilignes, et abrégé d’isopérimétrie élémentaire,” Genéva and Paris, 1789.
15. J. Steiner, Einfache Beweis der isoperimetrische Hauptsätze, *J. Reine Angew. Math.* **18** (1838), 289–296.
16. T. Bonnesen and W. Fenchel, “Theory of Convex Bodies,” B.C.S. Associates 1987. [Translation from the original German 1934].
17. O. Hölder, Über einen Mittelwertsatz *Göttinger Nachr.* (1889), 38–47.
18. H. Minkowski, “Geometrie der Zahlen, I,” Leipzig 1896.
19. P. L. Chebyshev, O priblizennyh vyraženijah odnih integralov čerez drugie *Soobshch. Protok. Zasekamiĭ Mat. Obšestva Imperator. Har’kovskom Univ.* **2** (1882), 93–98.
20. P. L. Chebyshev, Ob odnom rjade, dostavljajuščem predel’nye veličny integralov pri razloženii podintegral’noi funkcii na množeteli, *Prilož. Zapisok Imp. Akad. Nauk.* **4** (1883).
21. G. Chrystal, “Algebra, II,” 2nd ed. London 1900.
22. A. Winckler, Allgemeine Sätze zur Theorie der unregelmässigen Beobachtungsfehler, *Sitzungsber. Wiener Akad.* **53** (1866), 6–41.
23. M. Hermite, Cours de M. Hermite, professé pendant le l^e semestre, 1881–1882; redigé par M. Androyer, pp. 48–49, second tirage, Hermite, Paris, 1883.
24. A. N. Korkine, Sur un theórème de M. Tchebychef *C. R. Acad. Sci. Paris* **96** (1883), 326–327.
25. K. A. Andréief, Notes sur une relation entre les intégrales définies des produits des fonctions *Mém. Soc. Sci. Bordeaux (3)* **2** (1883), 1–14.
26. J. Hadamard, Résolution d’une question relative aux déterminants *Bull. Sci. Math (2)* **17** (1893), 240–248.
27. J. Hadamard, “The Psychology of Invention in the Mathematical Field,” Princeton Univ. Press, Princeton, NJ, 1949.
28. J. P. Gram, Über die Entwicklung reellen Funktionen in Reihen mittelst der Methode der kleinsten Quadrate, *J. Reine Angew. Math.* **94** (1881), 41–73.
29. J. L. W. V. Jensen, Om konvexe funktioner og uligheder mellem middelvaerdier *Nyt. Tidsskr. Math. B* **16** (1905), 49–69.
30. J. Hadamard, Étude sur les propriétés des fonctions entières et en particulier d’une fonction considérée par Riemann *J. Math. Pures Appl.* **58** (1893), 171–215.
31. A. M. Fink, A best possible Hadamard inequality, submitted for publication.
32. H. Weyl, “Singuläre Integralgleichungen mit besonderer Berücksichtigung des Fourierschen Integraltheorems,” Inaugural-Dissertation, Göttingen, 1908.
33. I. Schur, Bemerkungen zur Theorie der beschränkten Bilinearformen mit unendlich vielen Veränderlichen *J. Math.* **140** (1911), 1–28.
34. G. H. Hardy, Note on a theorem of Hilbert concerrning series of positive terms *Proc. London Math. Soc. (2)* **23** (1925).
35. I. Schur, Note on a theorem of Hilbert *Math. Z.* (1920), 314–317.
36. E. Landau, A note on a theorem concerning series of positive terms *J. London Math. Soc.* **1** (1926), 38–39.
37. B. Opic and A. Kufner, “Hardy-type Inequalities,” Longman, Harlow/New York, 1990.
38. E. Landau, Einige Ungleichungen für zweimal differentiierbaren Funktionen *Proc. London Math. Soc. (2)* **13** (1913), 43–49.
39. J. Hadamard, Sur le module maximum d’une fonction et de ses dérivees *C.R. Soci. Math. France* (1914), 68–72.
40. G. H. Hardy and J. E. Littlewood, Contributions to the arithmetic theory of series *Proc. London Math. Soc.* (1913), 411–478.
41. E. Landau, Über einem Satz von Herrn Esclangon *Math. Ann.* **102** (1929), 177–188.

42. A. Kolmogorov, Une généralisation de l'intalgrite de J. Hadamard entre les bornes supérieurs des dérivés successives d'une fonction *C. R. Acad. Sci. Paris* **20** (1938), 763–765.
43. A. Kolmogorov, On inequalities between upper bounds of consecutive derivatives of an arbitrary function defined on an infinite interval, *Učen. Zap. Modov. Gos. Univ. Mt.* **30** (1939), 3–16. [In Russian]
44. I. J. Schoenberg and A. Cavaretta, Solutions of Landau's problem concerning higher derivatives on the half line *in* "Proc. Int. Conf. of Constr. Func. Th. Varna, 1970," pp. 19–25.
45. M. K. Kwong and A. Zettl, "Norm Inequalities for Derivatives and Difference," Springer-Verlag, New York/Berlin, 1980.
46. T. Gronwall, Note on the derivatives with respect to a parameter of the solutions of a system of differential equations, *Ann. Math. (2)* **20** (1919), 292–296.
47. R. Bellman, The stability of solutions of linear differential equations *Duke Math. J.* **10** (1943), 643–647.
48. W. Walter, "Differential and Integral Inequalities," Springer-Verlag, New York/Berlin, 1970.
49. D. Bainov and P. Simeonov, "Integral Inequalities and Applications," Kluwer Academic, Dordrecht, 1992.
50. J. F. Steffensen, On certain inequalities between mean values, and their application to actuarial problems, *Skand. Aktuarietids* (1918), 82–97.
51. C. F. Gauss, Werke, Vol. 4, Theoria combinationis observationum erroribus minimus obnoxiae pars prior, art. 7–10.
52. W. H. Young, On classes of summable functions and their Fourier series *Proc. Roy. Soc. London Ser. A* **87** (1912), 225–229.
53. A. M. Lyapanov, Problème général de la stabilité du mouvement *Ann. Fac. Sci. Toulouse Math. (2)* **9** (1907), 406.
54. R. P. Agarwal and P. Y. H. Pang, "Opial Inequalities with Applications in Differential and Difference Equations," Kluwer Academic, Dordrecht, 1995.
55. Z. Opial, Sur une inegalité *Ann. Polon. Math.* **8** (1960), 29–32.
56. W. Blaschke, "Kreis und Kugle," pp. 105–106, Teubner, Leipzig, 1916.
57. E. Ahmansi, Sopra una delle esperienze di Plateau *Ann. Mat. Pura Appl.* **12** (1905), 1–17.
58. J. Brink, Inequalities involving $||f||_p$ and $||f^{(n)}||$ for f with n zeros *Pacific J. Math.* **42** (1972), 289–311.
59. A. M. Fink, Conjugate inequalities for functions and their derivatives *SIAM J. Math. Anal.* **5** (1974), 399–411.
60. A. M. Fink, Differential inequalities and disconjugacy *J. Math. Anal. Appl.* **49** (1975), 758–772.
61. G. H. Hardy, Prolegmena to a chapter on inequalities *J. London Math. Soc.* **4** (1929), 61–78; **5** (1930), 80.
62. A. M. Fink, Toward a theory of best possible inequalities *Nieuw Arch. Wisk.* **12** (1994), 19–30.
63. A. M. Fink and M. Jodeit, Jr., On Chebyshev's other inequality, *in* "Inequalities in Statistics and Probability," Lecture Notes IMS, Vol. 5, pp. 115–120, Inst. Math. Statist. Hayward, CA, 1984.
64. A. M. Fink and M. Jodeit, Jr., Jensen's inequality for functions with higher monotonicities *Aequationes Math.* **40** (1990), 26–43.
65. G. H. Hardy, J. E. Littlewood, and G. Polya, Some simple inequalities satisfied by convex functions, *Messenger Math.* **58** (1929), 145–152.
66. A. W. Marshall and I. Olkin, "Inequalities: Theory of Majorization and Its Applications," Academic Press, San Diego, 1979.

67. C. E. Shannon, The mathematical theory of communication *Bell Syst. Tech. J.* **27** (1948), 379–423, 623–656.
68. J. Inequal. Appl., Gordon & Breach, New York, Founded 1997.
69. “Mathematical Inequalities and Applications,” Element 1998.
70. O. Bottema, R.Ž. Djordjević, R.R. Janić, D.S. Mitrinović, and P.M. Vasić, “Geometric Inequalities” Groningen, 1969.
71. M. Marcus and H. Ming, “A Survey of Matrix Theory and Matrix Inequalities,” Allyn & Bacon, Boston 1964.
72. E. Stein, “Singular Integrals and Differentiability Properties of Functions,” Princeton Univ. Press, Princeton, NJ, 1970.
73. M. Protter and H. Weinberger, “Maximum Principles in Differential Equations,” Springer-Verlag, New York 1984.
74. R. Adams, “Sobolev Spaces,” Academic Press, New York, 1975.