Richard Askey

How Can Mathematicians and Mathematical Historians Help Each Other?

1. Introduction

This paper could have a slightly different title with the word *How* dropped, but I could argue both sides of that question. The present title presumes the optimistic answer, and while we all hope that this is the correct answer, the present time may not be the right one for this answer.

The history of mathematics is not an easy field, and it takes a rare person to be good at it. Keynes supposedly said it took a rare person to be a great economist: one must be a second-rate historian, mathematician, and philosopher. For him second-rate was very good, but not great. The same is probably true about the history of mathematics, except one may not have to be a philosopher. Some great mathematicians have made important contributions to the history of mathematics, but very few have spent enough time and thought on the history of mathematics to be able to write a first-rate historical account of part of mathematics. One recent exception is A. Weil, who has added an excellent historical account of number theory before Gauss (Weil 1984) to the historical notes he wrote for the Bourbaki volumes (Bourbaki 1974). His paper from the Helsinki Congress (Weil 1980) should also be read by anyone interested in the history of mathematics.

Since my training is in mathematics, and my reading of history has been almost random, I have found it useful to think about what history is and how the history of mathematics differs from cultural or political history. Collingwood starts his book *The Idea of History* (1956) with the following four questions: What is history, what is it about, how does it proceed, and what is it for? Of the many answers that could be given to these questions, he gave general ones that others could probably agree with, although most would think the answers were incomplete. For the first question, he wrote that "history is a kind of research or inquiry" (1956, 9).

This is so general that it means little without a more specific statement about what *type* of research. He gave this, and I have nothing to add. Here there is no real difference between the history of mathematics and other types of history.

Collingwood rephrased the second question to ask: What kind of things does history find out? His answer: "actions of human beings that have been done in the past" (1956, 9). Here the answer for the history of mathematics would be slightly different. Beethoven and Shakespeare did unique work, and we would not have *Fidelio* or *King Lear* if they have not lived. However, symmetric functions would have been discovered by others if Newton had not lived. Most mathematicians are Platonists who believe that we discover mathematics rather than invent or create it. The American Revolution and the government that came after would have been significantly different if another group of people had been involved, but most of us believe that calculus would be essentially the same no matter when or where it was developed. I am not trying to claim that our way of thinking of the real numbers would always be the same, for that is clearly not true, but that the essence of calculus would be the same. The essence of calculus has not really changed since Euler's time, and many of us would be happy to teach from his texts if our students were significantly better. As Rota wrote about Mark Kac (Kac 1985, xi): "He warned them [students] that axioms will change with the whims of time, but an application is forever." The applications of calculus, and the facts in Euler, are the real essence of calculus.

Skipping the third question, where I have little to add, we are left with last: What is history for? Collingwood answered that it is "'for' human knowledge" (1956, 10). For the history of mathematics this could be changed to "for mathematical knowledge," except that does not distinguish the history of mathematics from doing mathematics. As a substitute, consider the related but different question raised by Weil in (1980, 227); as he remarked, it is a question that has been discussed by others in the past. The question is: Who is history for? Weil made a number of useful comments on this question for the history of mathematics. My view, which is fairly narrow, is summed up nicely by one of the great historians of early mathematics, O. Neugebauer (1956). More will be quoted than is necessary to try to entice the reader to look at his complete note: "I always felt that its total lack of mathematical competence as well as its moralizing and anecdotal attitude seriously discredited the history of mathematics in the eyes of mathematicians, for whom, after all, the history of mathematics has to be written."

If the history of mathematics is to be written for mathematicians (exclusively, or even in part), then historians of mathematics need to know a lot of mathematics. One cannot form an adequate picture of what is really important on the basis of the current undergraduate curriculum and first-year graduate courses. In particular, I think there is far too much emphasis on the emergence of rigor and on the foundations of mathematics in much of what is published on the history of mathematics. In an interesting article, Jeremy Gray (1985, 18) wrote: "The foundations of analysis does not emerge as the central topic in mathematics that one might think it was from the historians of mathematics." He was writing about the period of one hundred years ago, and his conclusion was that the central topics then were the theory of differential equations and a variety of topics in geometry, including the theory of algebraic curves.

2. How Can Mathematicians Help?

Most mathematicians have little or no training in the ways of thought that historians have developed, so it is unrealistic to expect many of them to write papers or books that will satisfy mathematical historians. However, some mathematicians are tempted to write a paper on the history of a topic they have studied for years. I was tempted and did this over ten years ago. I had found a few series identities in papers that had been forgotten, and in one case an important result usually attributed to Saalschütz (1890) had been found by Pfaff (1797b) almost one hundred years earlier. Actually, I did not find this paper but read about it in (Jacobi 1848). The identity of Pfaff and Saalschütz is

$$(2.1)\qquad \sum_{k=0}^{n} \frac{(-n)_k(a)_k(b)_k}{(c)_k(a+b+1-n-c)_k k!} = \frac{(c-a)_n(c-b)_n}{(c)_n(c-a-b)_n},$$

where the shifted factorial $(a)n$ is defined by

$$(2.2)\qquad \begin{aligned} (a)_n &= a(a+1)\ldots(a+n-1), & n &= 1,2,\ldots, \\ &= 1 \qquad\qquad\qquad\qquad , & n &= 0. \end{aligned}$$

There were two reasons I wanted to call attention to Pfaff's paper. One is historical, and should have been of interest to historians. When $n \to \infty$

in (2.1), the result is

$$(2.3) \qquad \sum_{k=0}^{\infty} \frac{(a)_k(b)_k}{(c)_k k!} = \frac{\Gamma(c)\Gamma(c-a-b)}{\Gamma(c-a)\Gamma(c-b)}$$

if Re(c-a-b) > 0 and $c \neq 0, -1, \ldots$. This result was proved by Gauss in his published paper on hypergeometric series (1813). Notice the publication date: Gauss lived in Pfaff's home for a few months in 1797, and this was the year in which Pfaff published (1797b) and also published a book (1797a), the middle third of which contains the most comprehensive treatment of hypergeometric functions that appeared before Gauss's work. In addition to the published paper mentioned above, Gauss wrote a sequel that was only published posthumously (1866). Felix Klein was aware of Pfaff's book, and he raised the question of connection between Pfaff's work and Gauss's later work (Klein 1933). The sum (2.1) is not in Pfaff's book, so one can also ask if Gauss had seen this paper by Pfaff before he did not work in (1813). If he had, it is clear he was not then aware of the importance of hypergeometric functions, for he would have easily seen that (2.1) implies (2.3), and so not have thought of (2.3) as a new result. If he saw (2.1) before he appreciated the importance of hypergeometric functions, it is very likely he would have forgotten it. We will probably never know whether Gauss saw this work by Pfaff, but it is worth pointing out the possible influence, for a mathematical historian may find an annotated book or offprint and not appreciate the importance unless told why this influence is interesting.

The second reason interests me but might not interest others. There is a second sum that seems similar to (2.1) and that is attributed to Dixon (1903);

$$(2.4) \qquad \sum_{k=0}^{\infty} \frac{(a)_k(b)_k(c)_k}{k!(a+1-b)_k(a+1-c)_k}$$

$$= \frac{\Gamma(1+\frac{a}{2})\Gamma(1+a-b)\Gamma(1+a-c)\Gamma(1+\frac{a}{2}-b-c)}{\Gamma(1+a)\Gamma(1+\frac{a}{2}-b)\Gamma(1+\frac{a}{2}-c)\Gamma(1+a-b-c)}.$$

A much more general result was given by Rogers (1895, sect. 8), and a few special cases of (2.4) were found earlier, but the earliest special case I know appeared in (Dixon 1891). This is almost one hundred years after Pfaff proved (2.1), but only one year after Saalschütz rediscovered it (1890). These two results are often given at the same time. For example, Knuth (1968, 70, #31; 73, #62) has both of them as problems. He assigned 20 points to (2.1) and 38 points to (2.4), on a scale of 0 to 50, so it is clear he knew that (2.4) is deeper than (2.1). This depth is also illustrated by the approximately one hundred year difference in time when they were discovered. If the reader thought the time difference was one year (for the special case) or thirteen years (for the general case), these results would seem to be of the same depth. It took me a number of years before I appreciated the difference in depth.

I wanted to make one other point in this paper. Many people who write about (2.1) and (2.4) do not really understand what these identities say. For example, Knuth did not write either of these identities in the above form. He used binomial coefficients rather than shifted factorials. To explain the reason behind this difference, I will have to get technical.

In elementary calculus, the favorite test for convergence of an infinite series is the ratio test. It is easy to use and works on most power series that are given to students at this level. For these power series, the ratio between successive terms is a rational function of the index of the term. For example, for a power series about $x = 0$, if

$$f(x) = \sum_{n=0}^{\infty} c_n$$

then for $f(x) = \exp x$ the ratio is

$$\frac{c_{n+1}}{c_n} = \frac{x}{n+1},$$

for $f(x) = (1-x)^{-a}$ the ratio is

$$\frac{c_{n+1}}{c_n} = \frac{(n+a)}{(n+1)} x,$$

and for $\log(1+x)$ the ratio is

$$\frac{c_{n+1}}{c_n} = \frac{-n}{(n+1)} x, \qquad n = 1,2, \ldots .$$

A generalized hypergeometric series is a series

$$\sum c_n$$

with $c_n + 1/c_n$ a rational function of n. This rational function is usually factored as

(2.5) $$\frac{c_{n+1}}{c_n} = \frac{(n+a_1)\ldots(n+a_p)x}{(n+b_1)\ldots(n+b_q)(n+1)}$$

and the series is written as

(2.6) $${}_pF_q\left[\begin{matrix}a_1, \ldots, a_p \\ b_1, \ldots, b_q\end{matrix}; x\right] = \sum_{n=0}^{\infty} \frac{(a_1)_n \ldots (a_p)_n}{(b_1)_n \ldots (b_q)_n} \frac{x^n}{n!}.$$

The special case $p = 2$, $q = 1$ is often called the hypergeometric series, and the analytic continuation of this function is called the hypergeometric function. It is single valued on the plane cut on $[1,\infty)$ and on an appropriate Riemann surface.

Euler, Gauss, Kummer, Riemann, and many other mathematicians studied the general ${}_2F_1$, and facts about this function are important enough to be collected in many handbooks. Limiting or special cases such as Bessel functions, which are essentially ${}_0F_1$'s; confluent hypergeometric functions, which are ${}_1F_1$'s or linear combinations of two ${}_1F_1$'s; and Legendre functions, which are ${}_2F_1$'s with one or two parameters specialized in appropriate ways, have also been studied extensively. Facts about them are given in very large books, such as Watson's book on Bessel functions (1944) Robin's three volumes on Legendre functions (1957-59), and Hobson's book on spherical harmonics (1931), as well as the standard handbooks (Abramowitz and Stegun 1965; Erdélyi 1953-55). Thus one cannot claim this material is not well known, at least in some circles. However, there are many other circles where this work is almost completely unknown. The best example is combinatorics. One of the most important sets of mathematics books written in the last twenty years in Knuth's *The Art of Computer Programming* (1968, 1969, 1973). In this set of books, and in many others of a combinatorial nature, binomial coefficients occur regularly. The binomial coefficient $\left[\begin{matrix}n\\k\end{matrix}\right]$ is given by

$$\left[\begin{matrix}n\\k\end{matrix}\right] = \frac{n!}{k!(n-k)!}$$

and it counts the number of ways k identical objects can be put in n spots. Knuth wrote:

> There are literally thousands of identities involving binomial coefficients, and for centuries many people have been pleased to discover them. However, there are so many relations present that when someone finds a new identity, there aren't many people who get excited about

it any more, except the discoverer! In order to manipulate the formulas which arise in the analysis of algorithms, a facility for handling binomial coefficients is a must.... (1968, sect. 1.2.6, 52-53)

When a mathematician who is as good as Knuth write nonsense like the above (except for the last sentence, where he is probably right), then one must look seriously at what he wrote and try to understand why he missed the essence of what is really true. There are actually very few identities of the sort Knuth gave in this section—there just seem to be many because he does not know how to write them. For example, in (1968, sect. 1.2.6I) he gave six sums (21)-(26) and then wrote that (21) is by far the most important. What he did not point out is that five of these identities, (21)-(25), are all just disguised versions of

$$ {}_2F_1\left[\begin{matrix} -n,a \\ c \end{matrix} ; 1\right] = \frac{(c-a)_n}{(c)_n}. $$

In other words, they are all the same identity. Binomial coefficients are important, since they count things; but when one has a series of products of binomial coefficients, the right thing to do is to translate the sum to the hypergeometric series for (2.6). Translation is almost always easy (there can be some problems that require limits when division by zero arises), and it has been known for a long time that this is the right way to handle sums of products of binomial coefficients. Andrews spelled this out in detail in (1974, sect. 5), but the realization that hypergeometric series are just series with term ratio a rational function of n is very old. Horn (1889) used this as the definition of a hypergeometric series in two variables. R. Narasimhan told me that he found a definition of "comfortable" series in one of the late volumes of Euler's collected works. For Euler, a comfortable series is a power series whose term ratio is a rational function of n. When I asked Narasimhan to give me a specific reference, he was unable to find it again. I will be very pleased to pay \$50 U.S. for this reference, for it would be worth that to know that Euler's insight was also good here. An even earlier place one might look for this insight would be in Newton's work. In any case, by the time of Kummer's early work (1836), some mathematicians started to look at higher hypergeometric series and write them as

$$ 1 + \frac{\alpha\beta\lambda}{1\cdot\gamma\nu}x + \frac{\alpha(\alpha+1)\beta(\beta+1)\lambda(\lambda+1)}{1\cdot 2\cdot\gamma(\gamma+1)\nu(\nu+1)}x^2 + \ldots $$

which is (2.6) when $p = 3$, $q = 2$. Clausen (1828) used the same notation even earlier.

I wrote a short note (1975) mentioning many of the above facts and sent it to *Historia Mathematica*. It was sent to two referees, who disliked it because I had not written a history paper. The editor then sent it to two other referees, who also did not like it. One thought there were the makings of a reasonable paper if I only did some serious historical research, whereas the other thought the paper was silly. I had written that this material was treated badly in books of mathematical history and was not part of the standard curriculum. His view was essentially the following: if this work is not part of the curriculum, then it probably is not very important.

I would like to quote one paragraph from this technical report, changing it slightly to make it readable without including the earlier text, and then give the surprising sequel:

> The real reason for the obscurity of this material seems to be that it plays a minor role in most problems. Often this role is essential, but there is usually some other idea involved in the solution of a problem which seems to be more central (it usually is) and the explicit sum which is necessary remains a lemma. These sums are easy enough to derive so that a mathematician who has been able to come up with other ideas on how to solve a problem can also rediscover the required sum. But this has not always been true. For example, Good obtained the sum
>
> $$d_{0,2s} = \sum_{\nu=0}^{s} (-1)^{\nu} \binom{\beta}{\nu} \binom{\beta+s-\nu}{\beta} \frac{\alpha}{\alpha+s-\nu}.$$
>
> Then he says "(The sum of the series on the right must be non-negative if $\alpha > \beta$, an inequality that is not obvious directly.) If $\beta = 0$,
>
> $$d_{0,2s} = \frac{\alpha}{\alpha+s}."$$
>
> See Good (1958). What he did not notice is that this series can be translated into hypergeometric form and summed by a special case of (2.1). The result is
>
> $$d_{0,2s} = \frac{(\alpha-\beta)_s}{(\alpha+1)_s}.$$
>
> Now the positivity for $\alpha > \beta$ is obvious. Recently some very complicated formulas for hypergeometric series have been used by Gasper (1975) to obtain some inequalities for integrals which have not been obtained by any other method. Askey and Gasper (1977) used some other deep facts about hypergeometric series to extend a result of Szegö (1933). If more complicated problems of this sort are to be solved in other areas then mathematicians are going to have to realize that even the subject

of explicit evaluation of sums will have to be looked at in a more systematic way. It will be interesting to see if the knowledge of these important results spreads more widely in the next few decades. If one judges by past history the prospects are poor.

This is probably the most prophetic paragraph I have ever written. In the winter of 1984, L. de Branges reduced the Bieberbach conjecture to showing that a certain integral of a ${}_2F_1$ is positive. He had a new idea about how to use the Loewner machine, and he reduced a generalization of the Bieberbach conjecture—the Milin conjecture—to the following inequality:

$$(2.7) \qquad \int_0^1 t^{n-k-1/2}\, {}_2F_1\left[\begin{matrix} -k, 2n-k+2 \\ 2n-2k+1 \end{matrix}; tx\right] dt \geq 0,$$

$$0 < x < 1,\ k = 0,1, \ldots n-1.$$

For the sake of those who do not know these conjectures, the Bieberbach conjecture is the following. A function $f(z)$ is univalent if it is one to one on its domain. Let $f(z)$ by analytic for $|z| < 1$ and normalized by

$$f(z) = z + \sum_{n=2}^{\infty} a_n z^n, \qquad |z| < 1.$$

If f is univalent, Bieberbach (1916) showed that $|a_2| \leq 2$ and conjectured that $|an| \leq n$. There is equality when $f(z) = z(1\text{-}z)^{-2}$. This had been proved for $n = 3, 4, 5, 6$ and for all n for some subclasses of univalent functions, such as those with real coefficients an, but the general case was open and thought to be very hard. The Milin conjecture is too technical to state here, but it was known to imply the Bieberbach conjecture, and many experts on univalent functions thought it was false.

The final step in the proof of these conjectures has been described by the participants (Askey 1986; de Branges 1986; Gautschi 1986). Briefly, it is as follows. De Branges asked Walter Gautschi for aid in seeing when (2.7) holds; Gautschi is a colleague of de Branges's and probably the best person in the world to approach when asking for numerical results on the integral in (2.7). De Branges had heard that many experts on univalent functions thought the Bieberbach conjecture was false for odd values of n starting at 17 or 19, and Gautschi was skeptical that this approach would work; so the two of them were very excited when the numbers Gautschi found seemed to say that (2.7) held for n up to 30. This strongly suggested that the Milin conjecture and the Bieberbach conjecture were true for these

n and that it should be relatively easy to obtain error estimates to show that (2.7) held for those n, and probably a good deal higher. In fact, J. Hummel (personal communication) had independently done calculation on the integrated form of (2.7):

$$(2.8)\quad {}_3F_2\left[\begin{matrix}-k, 2n-k+2, n-k+\frac{1}{2} \\ n-k+\frac{3}{2}, 2n-2k+1\end{matrix}; x\right] > 0, \quad 0 < x < 1,$$

$$k = 0, 1, \ldots n-1,$$

and had shown that this is true for n up to about 20, and thus that the Milin and Bieberbach conjectures were true for these values of n.

The Bieberbach conjecture was a big problem, there was a nice new idea, and the final step dealt with hypergeometric functions, as in my outline quoted above. However, this time the hypergeometric function work was harder than just one identity. There is a relatively simple proof of (2.8), but it requires three identities, only one of which was well known. A second one is contained in the best handbooks. The third one is over a hundred years old and is contained in a few books, but not in the standard handbooks. Gautschi eventually called me and asked if I knew how to prove (2.7). I looked at it that evening, changed it to (2.8), and found this in the first place I looked (Askey and Gasper 1976). George Gasper and I had needed this inequality to prove a conjecture I had made, and Gasper had proved it.

De Branges's paper has now appeared (1985). After a version of de Branges's proof was available in preprint form, many mathematicians went through the details of his argument, gave talks on it, and some wrote their own accounts (Aharonov 1984; Anonymous [Maynooth] 1985; Fitz Gerald and Pommerenke 1985; Korevaar 1985; Milin 1984). However, only two of these accounts gave a complete proof of (2.7), and the general consensus was that the proof of (2.7) was magic and that it would be nice to have a more conceptual, less computational proof. One always wants simple proofs, but if one is willing to admit that Euler, Gauss, Kummer, Riemann et al. knew what they were doing when they studied hypergeometric functions, then this proof seems very natural. To prove that something is nonnegative, one tries to write it as a square or the sum of squares with nonnegative coefficients. The proof Gasper found is just that. What I am afraid of is that the last part of the quotation above will also be prophetic and that, even with this striking use of hypergeometric functions, knowledge of them will not spread to the mathematical community at large as it should.

My note was not the only paper written by a mathematician about some early mathematics of interest today that was turned down by the editors of *Historia Mathematica.* Another was written by George Andrews, and he published it elsewhere (1982). It is also not really a history paper by the standards of *Historia Mathematica,* but it deals with historical documents—in his case, two unpublished letters of L. J. Rogers; it is also of more interest to quite a few mathematicians than many of the papers in *Historia Mathematica.* I suggest it would be useful to publish an occasional article like this. If there were more articles of interest to mathematicians, then more mathematicians would read this journal. A few more papers by mathematicians would also tell historians what mathematicians consider important and would suggest topics that should be looked at in detail by historians. After my experience, the next time I had a topic of historical interest I wrote my comments for publication elsewhere. This topic was the orthogonal polynomials that generalize the classical polynomials of Hermite, Laguerre, and Jacobi. Most of these are older than is generally known. For example, a set of polynomials that is orthogonal on $x = 0, 1, \ldots, N$ with respect to the function $\left[{x+\alpha \atop x}\right]\left[{N-x+\beta \atop N-x}\right]$ and that are known as Hahn polynomials were really discovered by Tchebychef (1875); see my comments to (Szegö 1968) in the reprinted version. There is a need for a historical treatment of orthogonal polynomials. Szegö (1968) wrote an outline, I added further comments, and a historical resource without equal exists in (Shohat et al. 1940). This bibliography is not complete, but when it was written it was probably the best bibliography of a part of mathematics, and I do not know of another that equals it in the coverage of the eighteenth and nineteenth centuries. My comments appeared in a place where some mathematicians will see them, but mathematical historians are unlikely to hear about them.

There should be some place where mathematicians can record historical observations that will be read by mathematical historians. For example, consider general history books, which mostly contain material copied by the author from other books or papers. Errors tend to be propagated from one book to another, some of the errors being historical and others mathematical. More frequently, they are errors of ignorance, where the real point of the work is not understood. There needs to be a place where these errors can be corrected so that they will not appear in future books. I will illustrate these errors by mentioning some in M. Kline's book (1972). It is used as an illustration because I agree with the following remark of Rota (1974): "It is easy to find something to criticize in a treatise 1,200 pages long and packed with information. But whatever we say for or against it, we had better treasure this book on our shelf, for as far as mathematical history goes, it is the best we have." If the best can have errors like the following, then it is clear that mathematical historians need all the help they can get.

First, a mathematical error. Kline attributed the following formula to Euler (Kline 1972, 489):

$$(2.9) \quad {}_2F_1\left[\begin{matrix}-n,b\\ c\end{matrix};z\right] = \frac{n!}{c(c+1)\ldots(c+n-1)} \cdot \int_0^1 t^{-n-1}(1-t)^{c+n-1}(1-tz)^{-b}dt.$$

He does not say what n is, but by implication he has n a nonnegative integer, for he does not define what $c(c+1)\cdots(c+n\text{-}1)$ means when n is anything else. However, when $n = 0, 1, \ldots$, the integral diverges. This integral was copied from (Slater 1966, 3). Euler had a formula like (2.9) that is correct. It is

$${}_2F_1\left[\begin{matrix}a,b\\ c\end{matrix};z\right] = \frac{\Gamma(c)}{\Gamma(a)\Gamma(c-a)}\int_0^1 t^{a-1}(1-t)^{c-a-1}(1-tz)^{-b}dt$$

when $\operatorname{Re} c > \operatorname{Re} a > 0$.

An example of a historical fact that is wrong is the following:

> The problem of solving ordinary differential equations over infinite interval or semi-infinite intervals and of obtaining expansions of arbitrary functions over such intervals was also tackled by many men during the second half of the century and such special functions as Hermite functions first introduced by Hermite in 1864 and Nikolai J. Sonine in 1880 serve to solve this problem. (Kline 1972, 714)

These functions were known more than a half century before Hermite. Laplace studied them in his work in probability theory (Laplace 1812) and gave as many facts about them as Hermite was to rediscover. The date of discovery is important, but there is a more important point that Kline could have illustrated here. The message he put across is that these functions were introduced to solve differential equations. That is not true. They arose for other reasons, such as their connection with Fourier transforms, and the differential equation was a minor fact to Laplace. Part of the power of mathematics is the surprising way in which mathematical ideas or results that arose in one area turn out to be useful in other areas. That is just as true of special functions as it is of other parts of mathematics.

Finally, Kline does not seem to be aware that the ${}_2F_1$ hypergeometric function is really the most important of the special functions that arise as solutions of differential equations that come from separation of variables. He even said that the Bessel equation is the most important ordinary differential equation that results from separation of variables (1972, 710). Historically it probably arose most frequently, but the hypergeometric equation is much more important, and it contains the Bessel equation and many others as limiting or special cases. Riemann was the first to give a structural reason why the hypergeometric equation is so important. He showed that it is the only linear, homogeneous, second-order differential equation with regular singularities at $z = 0, 1,$ and ∞, with every other point an ordinary point. The Bessel equation has a regular singular point at $z = 0$ and an irregular singular point at $z = \infty$. Riemann's result is well over one hundred year old and should now be in a comprehensive history book such as Kline's. Kline's treatment of Riemann's problem and related work is considerably better than any other treatment in a general book on mathematical history—he just did not draw the right conclusions about the ${}_2F_1$ and its limits.

Kline mentions Gauss's published paper on the hypergeometric function twice but does not mention what I consider to be the major insight in this paper. In his first comment (p. 712), he mentions Gauss's sum (2.3) and the fact that Gauss recognized that special choices of the parameters will give many known elementary and higher transcendental functions. In the second comment (p. 962), he wrote about Gauss's work on convergence. Then he wrote the following, which I do not believe (p. 962): "The unusual rigor discouraged interest in the paper by mathematicians of the time."

The really new point of view that Gauss introduced in this paper was to treat

$$ {}_2F_1\left[\begin{matrix} \alpha,\beta \\ \gamma \end{matrix} ; x\right] $$

as a function of four variables, α,β and γ as well as x. This is very important and eventually led to some orthogonal polynomials that have been used in the quantum theory of angular momentum, in the analysis of pea growth, in the first proof of the irrationality of $\zeta(3)$, and in discrete analogues of spherical harmonics that are used in coding theory. Vorsselman de Heer drew upon Gauss's paper and some neglected early work of Euler for his thesis in (1833), and Kummer read Gauss's paper very carefully before he wrote (1836). Twenty years is not an exceedingly long time for a piece of work to be appreciated and extended. The work that Gasper and I did on positive sums of Jacobi polynomials was not used for ten years, and no one has extended or used Gasper's deeper work (1977), or even considered the dual problems. Gauss pushed the field ahead quite a few ways, and it took time to understand his work well enough so that the next steps could be made. The most significant extension of (Gauss 1813) was done by Heine (1845) but that is another story that leads to the Rogers-Ramanujan identities in partition theory and eventually to some beautiful and deep work of Baxter in statistical mechanics (1982, chap. 14).

3. How Can Mathematical Historians Help Mathematicians?

The answer to this question is easy: they can write papers on the history of mathematics that mathematicians care about. The American Mathematical Society has periodic special sessions on the history of mathematics, and these are very well attended. Mathematicians care about mathematics outside their own field, but they find a lack of good papers that contain the essence of the subject—namely, the problems that led to the field and the ideas that were used to attack these problems. This is really mathematical history. Please help us to learn more about mathematics by writing such papers. Eventually it may be possible to write a book like Kline's (1972) that will allow us to get a good overview of mathematical development. Before it is possible to have a good global view, however, we must have much better local views than we now have. As an illustration of a good local treatment of some parts of mathematics, look at some of the papers in (Dieudonné 1978) and the excellent treatment of the arithmetic-geometric mean in (Cox 1984). We need more historical work like these.

I mentioned work on orthogonal polynomials by R.. A. Fisher and a co-worker of his, F. Allan, in my comments to (Szegö 1968), but I did not say how I found this. There was a cryptic comment in Box's biography of Fisher (Box 1978) that mentioned the existence of some discrete orthogonal polynomials. She did not appreciate the importance of the representation found by Allan, and there is no reason she should have, but this representation was very important. If I had not rediscovered it a few years earlier, it would have been a good illustration of how a work on mathematical history could have aided a mathematician. I never would have come across Fisher's paper in the *Journal of Agricultural Science*, and was unlikely to have seen Allan's paper in the *Proceedings of the Royal Society of Edinburgh* (in 1930). However, a comment in a biography alerted me to the existence of these very interesting papers. Mathematicians read mathematical history outside the areas where they work, so this can be a good way of helping to break down the barriers that we have created around our work. Historians can help in ways they do not know, but only if they treat significant mathematics.

4. Conclusion

One of the referees of my note objected that it only contained facts and that history is much more than just facts. That is clearly true, but history built on incorrect facts is justly suspect. There needs to be a place where facts can be recorded, or corrected, and this needs to be a place that will be read by mathematical historians.

Weil (1980) set very high standards for what one would like to have in a mathematical historian, and then he demonstrated that he was not writing about an empty set in this century by living up to his standards. To be realistic, it is unlikely that there will be many others who match his description of a great mathematical historian, so we are going to have to help each other. Together we may able to do some useful history. Separately, some will be done, but not as much as is needed, and the quality will often be lower than it should be.

References

Abramowitz, M., and Stegun, I. 1965. *Handbook of Mathematical Functions*. New York: Dover.

Aharonov, D. 1984. *The de Branges Theorem on Univalent Functions*. Technion-Israel Institute of Technology, Haifa.

Andrews, G. 1974. Applications of Basic Hypergeometric Series. *SIAM Review* 16: 441-84.

——. 1982. L. J. Rogers and the Rogers-Ramanujan Identities. *Mathematical Chronicle* 11(2): 1-15.

Anonymous. 1985. *The Proof of the Bieberbach Conjecture*. Maynooth Mathematics Seminar, Maynooth University.

Askey, R. 1975. *A Note on the History of Series*. Mathematical Research Center Technical Report 1532. University of Wisconsin, Madison.

——. 1986. My Reaction to de Branges's Proof of the Bieberbach Conjecture. In *The Bieberbach Conjecture: Proceedings of the Symposium on the Occasion of the Proof*, ed. A. Baernstein II, D. Drasin, P. Duren, and A. Marden. Providence, R. I.: American Mathematical Society, pp. 213-15.

Askey, R., and Gasper, G. 1976. Positive Jacobi Polynomial Sums. II. *American Journal of Mathematics* 98: 709-37.

——. 1977. Convolution Structures for Laguerre Polynomials. *Journal d'Analyse Mathématique* 31: 48-68.
Baxter, R. J. 1982. *Exactly Solved Models in Statistical Mechanics*. London: Academic Press.
Bieberbach, L. 1916. Über die Koeffizienten derjenigen Potenzreihen, welche eine schlichte Abbildung des Einheitskreises vermitteln. *Sitzungsberichte der Preussischen Akademie der Wissenschaften, physikalish-mathematische Klasse* 940-55.
Bourbaki, N. 1974. *Eléments d'histoire des mathématiques*. Paris: Hermann.
Box, J. F. 1978. *R. A. Fisher, the Life of a Scientist*. New York: Wiley.
Clausen, T. 1828. Ueber die Fälle, wenn die Reihe von der Form $y = 1 + (\alpha\cdot\beta)/(1\cdot\gamma)x + (\alpha\cdot\alpha+1)/(1\cdot2)\cdot(\beta\cdot\beta+1)/(\gamma\cdot\gamma+1)x^2$ + etc. ein Quadrat von der Form $z = 1 + (\alpha'\beta'\gamma')/(1\delta'\varepsilon')x$ + etc. hat. *Journal für die reine und angewandte Mathematik* 3: 89-91.
Collingwood, R. G. 1956. *The Idea of History*. New York: Oxford University Press.
Cox, D. A. 1984. The Arithmetic-Geometric Mean of Gauss. *L'Enseignement Mathématique* 30: 275-330.
de Branges, L. 1985. A Proof of the Bieberbach Conjecture. *Acta Mathematica* 154: 137-52.
——. 1986. The Story of the Verification of the Bieberbach Conjecture. In *The Bieberbach Conjecture: Proceedings of the Symposium on the Occasion of the Proof*, ed. A. Baernstein II, D. Drasin, P. Duren, and A. Marden. Providence, R.I.: American Mathematical Society, pp. 199-203.
Dieudonné, J. 1978. *Abrégé d'histoire des mathématiques, 1700-1900*. 2 vols. Paris: Hermann.
Dixon, A. C. 1891. On the Sum of the Cubes of the Coefficients in a Certain Expansion by the Binomial Theorem. *Messenger of Mathematics* 20: 79-80.
——. 1903. Summation of a Certain Series. *Proceedings of the London Mathematical Society* 35: 284-89.
Erdélyi, A. 1953-55. *Higher Transcendental Functions*. 3 vols. New York: McGraw-Hill.
FitzGerald, C., and Pommerenke, C. 1985. The de Branges Theorem on Univalent Functions. *Transactions of the American Mathematical Society* 290: 683-90.
Gasper, G. 1975. Positive Integrals of Bessel Functions. *SIAM Journal on Mathematical Analysis* 6: 868-81.
——. 1977. Positive Sums of the Classical Orthogonal Polynomials. *SIAM Journal on Mathematical Analysis* 8: 423-47.
Gauss, C. F. 1813. Disquisitiones generales circa seriem infinitam $1 + (\alpha\beta)/(1\cdot\gamma)x + [\alpha(\alpha+1)\beta(\beta+1)]/[1\cdot2\cdot\gamma(\gamma+1)]xx + [\alpha(\alpha+1)(\alpha+2)\beta(\beta+1)(\beta+2)]/[1\cdot2\cdot3\cdot\gamma(\gamma+1)(\gamma+2)]x^3$ + etc. Pars prior. *Commentationes societatis regiae scientiarum Gottingensis recentiores, B*. Vol. 2. Reprinted in *Carl Friedrich Gauss Werke*. Vol. 3: *Analysis*. Göttigen, 1866, pp. 123-62.
——. 1866. Determinatio seriei nostrae per aequationem differentialem secundi ordinis. In *Carl Friedrich Gauss Werke*. Vol. 3: *Analysis*. Göttingen, 1866, pp. 207-29.
Gautschi, W. 1986. Reminiscences of My Involvement in de Branges' Proof of the Bieberbach Conjecture. In *The Bieberbach Conjecture: Proceedings of the Symposium on the Occasion of the Proof*, ed. A. Baernstein II, D. Drasin, P. Duren, and A. Marden. Providence, R.I.: American Mathematical Society, pp. 205-11.
Good, I. J. 1958. Random Motion and Analytic Continued Fractions. *Proceedings of the Cambridge Philosophical Society* 54: 43-47.
Gray, J. 1985. Who Would Have Won the Fields Medals a Hundred Years Ago? *Mathematical Intelligencer* 7 (3): 10-19.
Heine, E. 1845. Untersuchungen über die Reihe $1 + [(1-q^\alpha)(1-q^\beta)]/[(1-q)(1-q^\gamma)]x$ *Journal für die reine und angewandte Mathematik* 34: 285-328.
Hobson, E. W. 1931. *The Theory of Spherical and Ellipsoidal Harmonics*. Cambridge: Cambridge University Press. Reprinted New York: Chelsea, 1955.
Horn, J. 1889. Ueber die Convergenz der hypergeometrischen Reihen zweier und dreier Veränderlichen. *Mathematische Annalen* 34: 544-600.
Jacobi, C. G. J. 1848. De seriebus ac differentiis observatiunculae. *Journal für die reine und angewandte Mathematik* 36: 135-42. Reprinted in *Gesammelte Werke*, vol. 6, New York: Chelsea, 1969, pp. 174-82.

Kac, M. 1985. *Enigmas of Chance*. New York: Harper and Row.
Klein, F. 1933. *Vorlesungen über die hypergeometrische Funktion*. Berlin: Springer.
Kline, M. 1972. *Mathematical Thought from Ancient to Modern Times*. New York: Oxford University Press.
Knuth, D. 1968. *The Art of Computer Programming*. Vol. 1: *Fundamental Algorithms*. Reading, Mass.: Addison Wesley.
——. 1969. *The Art of Computer Programming*. Vol. 2: *Seminumerical Algorithms*. Reading, Mass.: Addison Wesley. 2d ed., 1981.
——. 1973. *The Art of Computer Programming*. Vol. 3: *Sorting and Searching*. Reading, Mass.: Addison Wesley.
Korevaar, J. 1985. *Ludwig Bieberbach's Conjecture and Its Proof by Louis de Branges*. Report 85-08, Mathematics Department, University of Amsterdam.
Kummer, E. E. 1836. Über die hypergeometrische Reine $1 + (\alpha \cdot \beta)/(1 \cdot \gamma)x + [\alpha(\alpha+1)\beta(\beta+1)]/[1 \cdot 2 \cdot \gamma(\gamma+1)]x^2 + \ldots$. *Journal für die reine und angewandte Mathematik* 15: 39-83, 127-72.
Laplace, P. S. 1812. *Théorie analytique des probabilitiés*. 2d ed. Paris.
Milin, I. M. 1984. *L. de Branges' Proof of the Bieberbach Hypothesis* (in Russian). Leningrad seminar notes.
Neugebauer, O. 1956. A Notice of Ingratitude. *Isis* 47: 58.
Pfaff, J. F. 1797a. *Disquisitiones Analyticae*. Helmstadt.
——. 1797b. Observationes analyticae ad L. Euler Institutiones Calculi Integralis, vol. 4, suppl. 2 et 4. *Histoire de 1793. Nova acta academiae scientiarum Petropolitanae* 11: 38-57. (Note that the history section is paged separately from the scientific section of this journal.)
Robin, L. 1957-59. *Fonctions sphériques de Legendre et fonctions sphéroidales*. 3 vols. Paris: Gauthier-Villars.
Rogers, L. J. 1895. Third Memoir on the Expansion of Certain Infinite Products. *Proceedings of the London Mathematical Society* 26: 15-32.
Rota, G.-C. 1974. Review of "Mathematical Thought from Ancient to Modern Times" by M. Kline. *Bulletin of the American Mathematical Society* 80: 805-7.
Saalschütz, L. 1890. Eine Summationsformel. *Zeitschrift für Mathematik und Physik* 35: 186-88.
Shohat, J., Hille, E., and Walsh, J. 1940. *A Bibliography on Orthogonal Polynomials*. National Research Council, Bull. 103. National Academy of Sciences, Washington, D.C.
Slater, L. J. 1966. *Generalized Hypergeometric Functions*. Cambridge: Cambridge University Press.
Szegö, G. 1933. Über gewisse Potenzreihen mit lauter positiven Koeffizienten. *Mathematische Zeitschrift* 37: 674-88. Reprinted in *Collected Papers*, vol. 2. Boston: Birkhäuser Boston, 1982, pp. 480-94.
——. 1968. An Outline on the History of Orthogonal Polynomials. In *Orthogonal Expansions and Their Continuous Analogues*, ed. D. Haimo. Carbondale: Southern Illinois University Press, pp. 3-11. Reprinted in *Collected Papers*, vol. 3. Boston: Birkhäuser Boston, 1982, pp. 857-65. Comment on this paper is on pp. 866-69.
Tchebychef, P. L. 1875. Sur l'interpolation des valeurs équidistants. *Zapiski Imperatorskoi Akademii Nauk* (Russia), vol. 25, suppl. 5 (in Russian). French translation appears in *Oeuvres*, vol. 2. New York: Chelsea, 1961, pp. 219-42.
Vorsselman de Heer, P. O. C. 1833. *Specimen in augurale de fractionibus continuis*. Thesis, Utrecht.
Watson, G. N. 1944. *Theory of Bessel Functions*. 2d ed. Cambridge: Cambridge University Press.
Weil, A. 1980. History of Mathematics: Why and How. *Proceedings of the International Congress of Mathematicians, Helsinki, 1978*. Vol. 1, Helskinki, pp. 227-36.
——. 1984. *Number Theory*. Boston: Birkhäuser Boston.