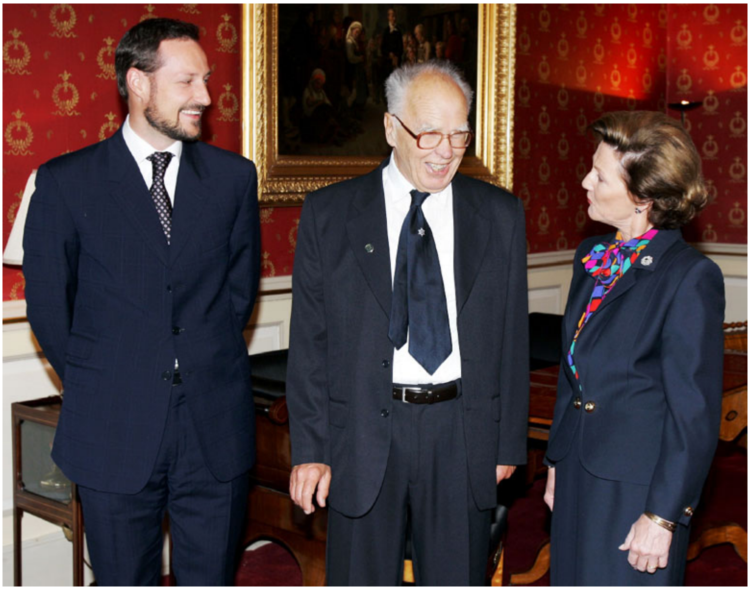
2006

Lennart Carleson



ABEL PRISEN

Reflections on My Life as a Mathematician

Lennart Carleson

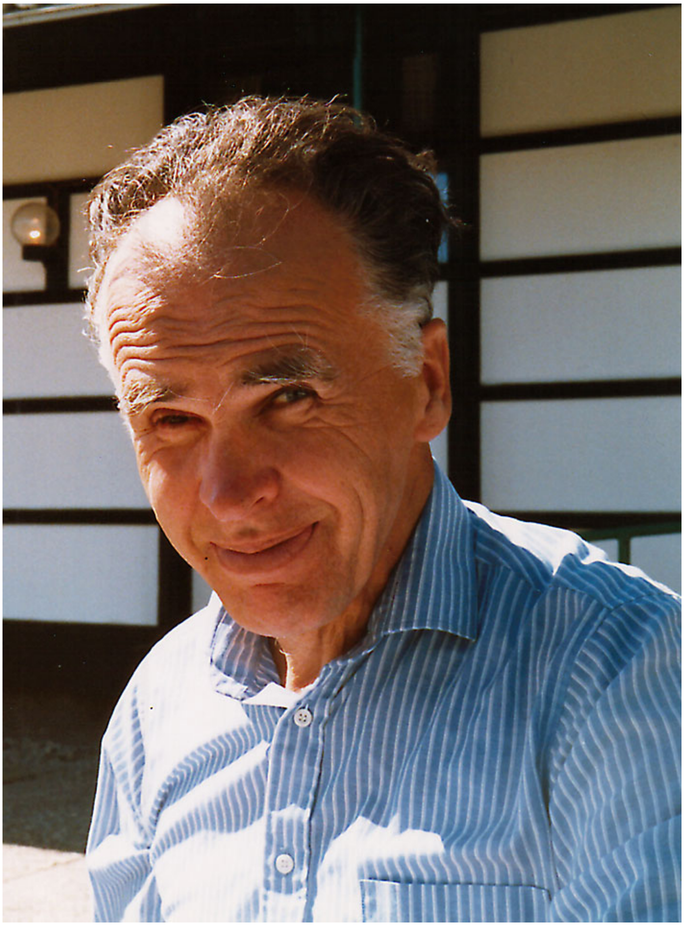
My ancestors are all from the Swedish bourgeois society. Among them are rather prominent civil servants and officers in the army. My father was originally a civil engineer and he became executive director of a large steel mill. My adolescence was very privileged and as a Swede I was fortunate to escape from the sufferings of the Second World War. There was no academic tradition in my family but I grew up in an intellectually stimulating environment where political matters were the main topics. I went to the local schools and skipped one grade, so that I finished high school barely 17 years old. I had then started to read mathematics textbooks, for the first undergraduate level, but I was in no way determined to become a mathematician. Most of my reading during the three last years of high school was classical literature, and when I now think back to that period I am amazed at how much I managed to read. This was actually the only period of my life that I could find time for a general education.

I graduated from high school in 1945 and started that year my mathematics studies at the university of Uppsala. That I should study mathematics was very clear to me, even if I had no thoughts on doing research. The subject seemed fascinating, I was intrigued by the strange equalities for infinite series, the mysterious imaginary numbers and by the problem solving. I would say to-day that my interests were very superficial and a little romantic. It was easy for me to learn how to handle the problems related to the texts and therefore to pass the exams. It seems to me (to-day) that there is nothing wrong with this rather superficial way of learning. It is similar to how you learn to ride a bicycle. It simply works and extensive explanations only complicate the matter. It will take you a long way; in my case I did not really understand how research in mathematics is created until after I was already a university professor.

But I am now way ahead of myself. During my second year at the university of Uppsala I had the opportunity to take a first course on analytic functions given

L. Carleson (✉)
Division of Mathematics, Royal Institute of Technology, 100 44 Stockholm, Sweden
e-mail: carleson@kth.se

H. Holden, R. Piene (eds.), *The Abel Prize*,
DOI 10.1007/978-3-642-01373-7_10,



(a)

by Beurling. This was a decisive moment in my life. The subject was mysterious and not easy to understand and reproduce. This is of course a completely natural reaction. Now that I have a life long relation to analytic functions the world seems natural and familiar, but I can still relive my original enchantment. Beurling was also a charismatic lecturer, very impressive in his appearance and in his dedication to his research. Even now that I have met most of this world's leading mathematicians, I can only think of Kolmogorov as comparable to him.

Beurling offered me a small job at the department and this was when I decided to continue and try to do research in mathematics. So I asked what I should read, and he suggested some books from "collection Borel" which contained books by Borel, Lebesgue, Vallée-Poussin, and Bernstein. There was no mention of examinations and I don't remember any exams up to Ph.D. I also read Zygmund's book on trigonometric series which, with its economic style, was very much to my lik-

(b) Carleson with S.-T. Yau

ing. This was when I learned about the problem of convergence of Fourier series. It seemed amazing that it was not known if the series for a continuous function need converge at any point at all.

Beurling suggested as a topic for my dissertation the study of a subclass of the class of meromorphic functions on the unit disc with bounded Nevanlinna characteristic functions. It was a good thesis problem in the sense that the general strategy was clear but some innovations were needed. My defense of the thesis was in 1950, precisely on my 22nd birthday. The following academic year was spent at Harvard where I met Lars Ahlfors. There was also the international congress. This was my first trip abroad; Sweden had been isolated during and after the war. It was a great experience to meet the wealth of mathematics and to see how little I knew. I listened to Zariski's lectures at Harvard and saw a world I did not know existed. I visited analysis seminars in several universities and realized how backward my type of analysis was considered to be. Who cares about the unit circle when one can solve problems for a locally compact topological space?

Nevertheless, I stuck to my type of problems and continued to go to Beurling's seminars. He remained in Uppsala until 1953 and gave a seminar on a new topic every (second?) Tuesday. These seminars can be admired in his Collected Works, and it was a great education for the small group attending them, even if for me it still was a rather passive experience. During this time Beurling told me about the

(c) Carleson in Oslo at the Abel Lectures with Björn Engquist, Erling Størmer, Peter D. Lax, Ingrid Daubechies, and Sun-Yung Alice Chang

Corona problem. I don't know from whom he got the problem; it was not his usual style. I was intrigued because it is a closure problem very different from the kind I was familiar with where linear functionals are the tool.

It is interesting to speculate on the flow of ideas in the mind. As a graduate student I was encouraged to try to understand the convergence and divergence in the absolute sense of Borel series. These are series when the nth term is given by $a_n|x - x_n|^{-\alpha}$. The divergence set is described by Hausdorff measures. Very similar series had been used by Otto Frostman in his famous thesis in potential theory and reoccurred in my thesis for the classes of meromorphic functions that I studied. Here there is an inner relation between the subjects that makes it all natural. What follows below is more mysterious.

During the year after my dissertation I worked on different problems in classic analysis without much originality. Nevertheless I was appointed professor first in Stockholm (1954) and then in Uppsala (1955). I decided to study interpolation in a general context, i.e., the interpolated values a_n should be taken from a linear space. A natural example is bounded analytic functions in $|z| < 1$ and $\{a_n\}$ all bounded sequences. The solution of this problem is in my opinion my first somewhat original result. I also considered a Fourier coefficient version which is usually associated to Sidon's name. One can also try to prescribe the nth derivative at $x = 0$ in an arbitrary way bounded by A_n in absolute value and try to continue the function to $(-1, 1)$ with nth derivative bounded by B_n. This problem has a complete, quite elegant solution of course related to the famous quasi-analytic functions. It was studied at the same time by Leon Ehrenpreis. E.g., if $A_n = n^{\alpha n}$, $\alpha > 1$, then B_n can be equal to A_n. But if $A_n = n^n(\log n)^\alpha$, then B_n must be larger, namely $n^n(\log n)^{\alpha+1}$.

My interest in interpolation inspired me to look at the Corona problem from this point of view and something surprising happened. The necessary interpolation which is implied by the Corona statement for two functions is also sufficient, and secondly, much of the needed argument was present in the interpolation problem for bounded analytic functions that I mentioned. In the center of the argument is a class of measures. My argument was very roundabout and got its present elegance by Eli Stein. How this all came about is an interesting riddle.

My next major project was to come back to the problem of my early years in Uppsala, the convergence of Fourier series, and I thought I had made a decisive progress: I knew why a Fourier series could diverge everywhere and still be square integrable This is what the experts in the area, in particular A. Zygmund, had been advocating for many years. I heard it from Zygmund during my 1950/51 stay at Harvard, and he encouraged me to use my knowledge of bounded analytic functions for this purpose. It did not work of course, and I put the problem aside. But now in 1963 I knew. And then I could prove that this optimal method to make a counterexample could not work! This created a strong conviction with me that the convergence is actually true. This is a very generally true statement in the psychology of mathematics. If you are to succeed to prove a statement, you must be convinced (for whatever reason, usually because your teacher says so) that it is true. It does not work to alternate: try to prove it one week, try to make a counterexample next week. If you have no clear idea, my advice is always to try to make a counterexample. This makes you familiar with the problem even if you find nothing.

My experience with analytic functions was helpful in the proof and some new technique concerning Hilbert transforms and maximal functions were used. But basically the proof could have been done in the 1930's. Most likely not so many people tried since all 'knew' that it was the converse statement that holds.

The solution of this 150 years old question was noticed in the mathematical community and I received several offers from universities in the USA. At the same time, the future of the Mittag-Leffler Institute was under debate. It had during its 50 years of existence suffered from lack of funds and now one option was to sell the property in Djursholm and move the remarkable library and the remaining funds to Stockholm. In my opinion, shared by many others, this would be a cultural crime and I decided to use my offers as leverage in negotiations for support from the government to the institute. The plan succeeded and the institute also got a large grant from the Wallenberg Foundation and yearly subsidies from the Scandinavian countries. Uppsala got a new position and I started to work at the institute. Activities there have now been going on for 40 years. Was this a good investment?

The institute serves as a symbol of mathematical research in Sweden on the international scene. This is enhanced by the fact that the institute publishes two journals: *Acta Mathematica* and *Arkiv för matematik*. The *Acta* was started 1882 by Mittag-Leffler, and I had taken over the editing in 1956 and did this first in Uppsala. The *Arkiv* belonged to the Swedish Academy and was about to be discontinued about 1970. It was moved to the institute as an outlet for Swedish mathematics primarily.

The institute should bring the best international mathematicians in some specific field every year and provide outstanding environment for collaboration. The institute should finally be a center for post-graduate education for Scandinavian mathematicians. New fields of mathematics, which are not represented in the small country of Sweden, could be presented. In this direction we had analytic number theory and statistical mechanics during the first few years and it certainly made a lasting impression on me.

So my answer is that an institute such as Mittag-Leffler's is an essential part of the organization of research and higher education, in particular in a small country.

The years at the Mittag-Leffler Institute, 1968–1984, were very rewarding. I became involved in the international politics of mathematics and was president of the IMU during the period 1979–1982. As time went by it became harder for me to organize years in new topics, and I needed help. This is how Peter Jones came to spend two years at the institute. We organized a year of dynamical systems and in particular complex dynamics 1983/84. At this time the field was just emerging and much seemed mysterious, at least to me. Michael Herman was among the visitors, and his thesis on the conjugation of circle diffeomorphisms is a paper which made a lasting impression on me. Here was new analysis in an elementary context. Mystifying papers were also Feigenbaum's paper on period doubling, Jacobson's on the logistic map (how could something so simple be so complicated?) and a short note by Ruelle on the Hénon map (pure magic!). So Michael Bendicks and I decided to demystify, first the Jacobsen logistic map and then the Hénon map. This actually was a ten years project and was only partially successful. One could say that the good Lord keeps some of his secrets to himself. After all, iteration of maps, is precisely his method of building the organic world.

One of the most rewarding aspects of the life of a university professor is the contact with students. In my case, I have not had to do so much class room teaching but have had some 30 thesis students. Many of them are now university professor, some even retired. Finding a good thesis subject is a rather delicate matter. It has to fit the interest of the student, it should have different levels so that the work can be finished before the subject is exhausted. You have to feel certain that the project is doable, without having actually done it all. Many of my students have turned out to be outstanding mathematicians. Let me mention here two examples. In the 1960's I suggested to Hans Brolin that he should do a modernized vision of the Julia–Fatou theory of rational iterations from about 1920 and add some new results. When the subject became fashionable around 1980 his thesis was the standard reference for a numbers of years. About 10 years ago, I suggested to Warwick Tucker to prove that the Lorenz attractor exists, using a combination of computer computations and a mathematical analysis of the fixed point. He was able to do this and in this way solved a long open problem.

When I review my main contributions to mathematics as I did above, I can see that they are all depending on one and the same idea. This is the idea of renormalization as physicists say, or stopping time, as probabilists say. Actually, the probabilistic idea is much older, due to Paul Lévy in the 1920's, while renormalization appeared in physics only in the 1960's. Namely, in the Corona proof, you make a geometric construction that is stopped and restarted by conditions on the analytic functions which are given. In the Fourier series proof you localize to an interval around the point that you consider. The size of this interval is determined by a stopping rule on the Fourier coefficients, and you repeat the argument on this small interval. In the Hénon map, finally, the times for stopping are determined by the expansion. You restart when you can disregard the past and this is precisely renormalization.

It is interesting to speculate on the reason for this similarity. The most natural explanation would be that the problems are closely related. But they certainly appear very different and they come from completely different branches of analysis. I was certainly not aware of any connection when I worked on them. It seems more likely that a mathematician develops a certain personal way of thinking and that this approach can be applied in very diverse situations. The method of solving, or rather analyzing, a problem would then be far from uniquely determined by the problem. There are certainly many examples of this, when, e.g., one proof comes from algebra and another from analysis. Think of Abel's and Galois' approaches to the quintic equation! Also you can recognize the person behind the proof (sometimes) much as you recognize a particular composer when you hear one of his symphonies.

Carleson for Beginners

Tom Körner

When I was just beginning as a research student, an older mathematical friend invited me to dine at Trinity. At dessert I was seated next to Littlewood, Hardy's collaborator and a legendary figure in modern British analysis. With old fashioned politeness, Littlewood set himself out to entertain me. He talked about a recent comet and recalled how fifty years earlier he had viewed Halley's comet in company with a Trinity fellow who had himself seen its previous visitation. He then spoke about Carleson's recent proof of the convergence theorem, what a marvellous result it was, how surprising it was that it turned out that Lusin's conjecture was true, how many people known to him had thought about the problem for a long time without success and how much he regretted being too old to take up the task of understanding the details of the proof.

I shall try to explain the importance of the result at the level of a well informed and able first year mathematics student. Non-mathematicians should not worry about the mathematics but just read the story. Mathematicians should note that I will make no attempt to explain the proof itself. I shall then talk more briefly about two other famous results of Carleson. Finally, in the last four or five pages, I shall address non-mathematicians directly.

Students often think of mathematics as consisting of definitions and theorems. It would be more correct to think of mathematics as consisting of examples and methods. A definition tells us nothing unless it is illuminated by concrete examples and counterexamples and can lead nowhere unless we have accompanying methods of proof.

The early 19th century saw the ultimately successful attempt to base calculus on a new type of 'epsilon, delta' definition. In 1829, Dirichlet gave one of the first and most impressive examples of the 'new analysis' by applying it to 'Fourier series'. In his work on heat conduction Fourier had given plausible arguments to show that

T. Körner (✉)
Department of Pure Mathematics and Mathematical Statistics, Centre for Mathematical Sciences, University of Cambridge, Clarkson Road, Cambridge, UK
e-mail: twk@dpmms.cam.ac.uk

H. Holden, R. Piene (eds.), *The Abel Prize*,
DOI 10.1007/978-3-642-01373-7_11,

any reasonable periodic function (with period L) could be represented by its Fourier series

$$f(x) = \frac{A_0}{2} + \sum_{r=-\infty}^{\infty} \left[A_r \cos\left(\frac{2\pi r x}{L}\right) + B_r \sin\left(\frac{2\pi r x}{L}\right)\right].$$

We shall take $L = 1$ and use the equivalent formulation

$$f(x) = \sum_{r=-\infty}^{\infty} a_r \exp(2\pi i r x).$$

As Fourier and others before him had noted, the only reasonable choice for the a_r is

$$a_r = \hat{f}(r) = \int_{-1/2}^{1/2} f(x) \exp(-2\pi i r x)\, dx.$$

'Fourier sums', and the closely related Fourier integrals, occur naturally in optics, communication theory and more generally in any physical problem involving oscillation or waves. They also occur in unexpected places like number theory and methods for fast machine computation.

If we write

$$S_n(f,t) = \sum_{r=-n}^{n} \hat{f}(r) \exp(2\pi i r t),$$

then the key problem facing Dirichlet and his successors was to find wide conditions under which

$$S_n(f,t) \to f(t)$$

as $n \to \infty$ and to prove that these conditions were indeed sufficient. The first observation to make is that

$$\begin{aligned}\sum_{r=-n}^{n} \hat{f}(r) \exp 2\pi i r t &= \sum_{r=-n}^{n} \int_{-1/2}^{1/2} f(x) \exp(-2\pi i r x)\, dx \exp(i r t)\\ &= \int_{-1/2}^{1/2} f(x) \sum_{r=-n}^{n} \exp\big(2\pi i r (t-x)\big)\, dx\\ &= \int_{-1/2}^{1/2} f(x) K_n(t-x)\, dx,\end{aligned}$$

where (summing a geometric series)

$$K_n(s) = \sum_{r=-n}^{n} \exp(2\pi i r s) = \frac{\sin(2\pi(n+\frac{1}{2})s)}{\sin \pi s}.$$

(To avoid division by zero, we set $K_n(0) = 2n+1$, but this creates no problems.)

Unfortunately K_n is not very well behaved. The reader who sketches the graph of K_{10} will see that K_n must be highly oscillatory and will not be surprised to learn that

$$\int_{-1/2}^{1/2} |K_n(s)|\,ds \to \infty$$

as $n \to \infty$ (indeed the integral grows at the same rate as $\log n$). However we do have

$$\int_{-1/2}^{1/2} K_n(s)\,ds = 1,$$

so, if f is constant, then the 'oscillations cancel' and

$$S_n(f, t) = f(t).$$

We can therefore hope that, if f is sufficiently well behaved, then, as n gets larger and the oscillations 'crowd together', they will start to cancel and

$$S_n(f, t) = \int_{-1/2}^{1/2} f(x)K_n(t - x)\,dx \to f(t)$$

as $n \to \infty$. By careful estimation, Dirichlet was able to show that if f is continuous and has only a finite number of maxima and minima, this is indeed the case.

It is possible that Dirichlet may, for a time, have thought that he could extend his result to all continuous functions. It is certain that most mathematicians and physicists thought that such a result was true. However, in 1873, du Bois-Reymond gave an example of a continuous function f such that $S_n(f, 0)$ fails to converge. In retrospect, the finding of such an example illustrates how, during the 19th century, mathematicians acquired a menagerie of continuous functions, came to understand the freedom of behaviour that epsilon-delta definitions allowed and developed new techniques for controlling that freedom.

Observe that it is easy to find a three times differentiable periodic function f_n such that $|f_n(t)| \le 1$ for all t and

$$S_n(f_n, 0) = \int_{-1/2}^{1/2} f_n(x)K_n(-x)\,dx \ge \frac{1}{2}\int_{-1/2}^{1/2} |K_n(-x)|\,dx.$$

It is plausible that, provided $N(j)$ increases fast enough,

$$f(t) = \sum_{j=1}^{\infty} 2^{-j} f_{n(j)}(t)$$

will define a continuous function such that

$$S_{n(j)}(f, 0) \to \infty,$$

and this is indeed the case although the details of the proof require some care.

There are two remarks which are worth making. The first is that a natural approach to the proof is via the notion of uniform convergence. From the modern point of view, this involves treating continuous functions as points in (or if the reader prefers, elements of) a space of continuous functions and considering the 'distance' between two points (or elements) F and G say to be given by

$$d(F,G) = \|F - G\|_\infty = \sup_t |F(t) - G(t)|.$$

Our second remark is the following. Let the sequence x_j contain each rational number infinitely often. If we modify the argument of the previous paragraph by considering

$$f(t) = \sum_{j=1}^{\infty} 2^{-j} f_{n(j)}(t - x_j),$$

then, provided we choose the n_j increasing fast enough, it is both plausible and true that we will define a continuous function f

$$S_{n(j)}(f, x_j) \to \infty,$$

and so the Fourier sum $S_n(f,t)$ will fail to converge at every rational point t. A look at Dirichlet's argument shows that this means that f has an infinite number of maxima and minima in each interval.

Of course, we can restrict ourselves to functions which satisfy Dirichlet's conditions or something similar. (If we only look at functions which are once continuously differentiable, then their Fourier sums are absolutely convergent and most of the analytic difficulties vanish.) However, in many cases, when we try to apply the results to natural problems, we find that much of our effort goes into dealing with the peripheral difficulties caused by these restrictions. It is not surprising that the subject began to stagnate. Although the point of view of this essay is very different from that of Klein, we can use his words to describe the state of Fourier Analysis at the end of the 19th century as 'like a large weapon shop in peace time. The store window is filled with showpieces whose ingenious artful and pleasing design enchants the connoisseur. The real purpose of these things, to attack and defeat the enemy, has retreated so far into the background of consciousness as to be forgotten' [2].

Fourier Analysis and analysis in general was revivified by the invention by Lebesgue and others of measure theory. One way of understanding this new idea is to consider the relation of the rational numbers to the reals. The rational numbers form a readily comprehensible algebraic system with an obvious notion of distance. Unfortunately this system behaves badly when we apply limiting processes. As a simple example consider the

$$1,\ 1 + \frac{1}{1!},\ 1 + \frac{1}{1!} + \frac{1}{2!},\ 1 + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!},\ \ldots$$

which 'ought to converge' but does not. To avoid this problem we embed the rationals in the larger system of the reals where everything that 'ought to converge' does. (More formally, every Cauchy sequence converges.) However, the new system of reals remains strongly linked to the old since every real number is the limit of a sequence of rationals. We say that the real numbers form a 'completion' of the rationals and that the rationals are 'dense in the reals'. Since the rationals form a subset of the reals we can prove results about the rationals by using results about the reals, and since the rationals are dense in the reals, we can prove results about the reals by using results about the rationals.

Lebesgue showed that the notion of integration can be extended to cover a very large class (called L^1) of functions on the interval $[-1/2, 1/2]$. Any bounded function that the reader can write down explicitly will represent an object in L^1, but the chief merit of L^1 is that it bears much the same relation to the continuous functions as the real numbers do to the rationals. Let us measure the 'distance' between two functions f and g in L^1 by

$$d_1(f, g) = \|f - g\|_1 = \int_{-1/2}^{1/2} |f(t) - g(t)|\, dt.$$

Then every sequence in L^1 that 'ought to converge' does, and so it is easy to do analysis in L^1, but every $f \in L^1$ is the limit of continuous functions, that is to say, given $f \in L^1$ we can find continuous functions f_n with

$$\|f - f_n\|_1 \to 0.$$

Since 'the continuous functions are dense in L^1', we can prove results about L^1 by using results about continuous functions and since every continuous function lies in L^1, every result on L^1 implies a result on continuous functions.

There is a small price to pay for extending our interest from the continuous functions to L^1, but it is a price well worth paying. We say that a subset E of $[-1/2, 1/2]$ has measure zero if, given any $\epsilon > 0$, we can cover E with a countable collection of intervals of total length less than ϵ. (Another way to think of this is that, if you drop a dart at random on the interval $[-1/2, 1/2]$, the probability of hitting E is zero.) It turns out that if two functions f and g in L^1 only differ on a set of measure zero then $\|f - g\|_1 = 0$ and f and g are indistinguishable as objects in L^1.

The set L^1 has a very important subset L^2 (also called the space of square integrable functions) consisting of those $f \in L^1$ such that

$$\int_{-1/2}^{1/2} |f(t)|^2\, dt < \infty.$$

A simple modification of the argument used to prove the Cauchy–Schwarz inequality shows that we can introduce a new 'distance' between functions f and g in L^2 given by

$$d_2(f, g) = \|f - g\|_2 = \left(\int_{-1/2}^{1/2} |f(t) - g(t)|^2\, dt\right)^{1/2}.$$

The distance d_2 is a natural generalisation of Euclidean distance and L^2 turns out to give the most natural generalisation of Euclidean ideas to infinite dimensional spaces.

Although the distances d_1 and d_2 are very different, it turns out that every sequence in L^2 which ought to converge under the new metric d_2 does indeed converge to a function in L^2 and that the continuous functions are dense in L^2 under the new metric.

It requires a fair amount of work to set up Lebesgue's theory, but, once this is done, many results which were 'almost true' in the old analysis become 'exactly true' and many arguments that 'almost worked' now work and work easily. For example Fourier analysis within L^2 can be condensed into four easily proved statements. (We write $e_n(t) = \exp(2\pi i n t)$.)

(1) If $f \in L^2$, then $\sum_{j=-\infty}^{\infty} |\hat{f}(j)|^2 < \infty$.

(2) If $\sum_{j=-\infty}^{\infty} |a_j|^2 < \infty$, then we can find an $f \in L^2$ with $\hat{f}(j) = a_j$ for all j.

(3) If $f, g \in L^2$ and $\hat{f}(j) = \hat{g}(j)$ for all j, then the set of t with $f(t) \neq g(t)$ has measure zero (that is to say f and g are identical from the point of view of Lebesgue's theory).

(4) If $f \in L^2$, then

$$\left\| \sum_{j=-n}^{n} \hat{f}(j) e_j - f \right\|_2 \to 0$$

as $n \to \infty$.

The behaviour of Fourier series in L^1 is not so simple, but, here again, Lebesgue's theory made many of the old rough paths smooth. In mathematics, as in other branches of learning, the adoption of new methods must often wait until a new generation replaces the old. However, even established mathematicians like Hardy and Birkhoff adopted the new ideas with enthusiasm. For a short time it appeared that difficulties could be confined to a set of measure zero and that any set of measure zero could be ignored.

If we look at the kind of functions f which appear in constructions modeled on those of du Bois-Reymond we see that, although the set E on which $S_n(f,t)$ fails to converge may be very complicated, it always seems to have measure zero. It is natural to conjecture that, if f is continuous, $S_n(f,t) \to f(t)$ as $n \to \infty$ except when t belongs to some set of measure zero. (More briefly $S_n(f,t) \to f(t)$ except on a set of measure zero.) Lusin strengthened this conjecture to cover all $f \in L^2$. (The reader should not assign Lusin the role of a lucky bit player in the drama. He was a deep mathematician who among other things was the first to construct a sequence $a_j \to 0$ as $|j| \to \infty$ but $\sum_{j=-n}^{n} a_j \exp(2\pi i j t)$ diverges at every point t.)

In 1922, at the age of 19, Kolmogorov shot to international fame by showing that Lusin's conjecture is false if we replace L^2 by L^1. In its final form, his result showed that there is an $f \in L^1$ such that $S_n(f,t)$ diverges for all t. His proof illustrates an important property of the Dirichlet kernel K_n. If we fix $1/2 > \delta > 0$ then, although $\int_{|s| \geq \delta} |K_n(s)|\, ds$ remains bounded as $n \to \infty$, it does not tend to zero. The increasing oscillation of K_n still means that

$$\int_{|s|\geq\delta} K_n(s) f(s)\, ds \to 0,$$

and so the limiting behaviour of $S_n(f, 0)$ does not depend on the values of $f(s)$ with $1 \geq |s| \geq \delta$. (This is the Riemann localisation principle and states that *ultimate behaviour* of the Fourier sum for f at a point only depends on the value of f near that point.) Kolmogorov's construction depends, in part, on the fact that for $f \in L^1$ it may take an arbitrarily long time for localisation to assert itself.

The reader should note that (though the actual behaviour is much more complicated) the divergence in Kolmogorov's example is more akin to the behaviour of g_n where

$$g_{2^m+r}(t) = \begin{cases} n & \text{for } 2^{-1} + r2^{-m} \leq t \leq 2^{-1} + (r+1)2^{-m}, \text{ and } 0 \leq r \leq 2^m - 1, \\ 0 & \text{otherwise,} \end{cases}$$

(so that, at each point t, long periods of good behaviour are interrupted by occasional periods of bad behaviour) than h_n where $h_n(t) = (-1)^n n$ and bad behaviour occurs everywhere all the time.

It seems likely that for the next forty years most mathematicians expected that some sort of tweaking of Kolmogorov's example would produce a counterexample to Lusin's conjecture. It is certainly true that, although several important theorems were obtained which proved the convergence of Fourier sums under various conditions, none of these results suggested that Lusin's conjecture was true.

However, there were advances in other parts of analysis which, we now see, shed light on the problem. Earlier, I said that it was possible to extend results from the rational numbers to the reals by a density argument. However, we must exercise caution. Consider the function $f : \mathbb{Q} \to \mathbb{Q}$ given by

$$f(x) = \begin{cases} 1 & \text{if } x^2 < 2, \\ 0 & \text{otherwise.} \end{cases}$$

Although f is continuous (the reader who doubts this is asked to find a point $t \in \mathbb{Q}$ where f is not continuous), we cannot find a continuous extension $\tilde{f}$ to the reals. (In other words, we cannot find a continuous $\tilde{f} : \mathbb{R} \to \mathbb{R}$ with $\tilde{f}(x) = f(x)$ for all $x \in \mathbb{Q}$.) On the other hand, if $g : \mathbb{Q} \to \mathbb{Q}$ is *uniformly* continuous, then it is easy to see that it has a continuous (indeed a uniformly continuous) extension $\tilde{g}$ to the reals.

Simple minded attempts to use this idea to tackle the kind of problem we are considering are bound to fail. The trigonometric polynomials (that is to say functions of the form $\sum_{j=-n}^{n} a_j \exp(2\pi i j t)$) are dense in L^2, but it is not true that every function in L^2 is a trigonometric polynomial.

A way forward was provided by the introduction by Hardy and Littlewood of the idea of a maximal function and the more general idea of a 'maximal inequality'. One outstanding result proved by these means was Birkhoff's ergodic theorem. Suppose that $T : [-1/2, 1/2] \to [-1/2, 1/2]$ is a bijection which 'preserves measure' in the sense that

$$\int_{-1/2}^{1/2} g(Tx)\, dx = \int_{-1/2}^{1/2} g(x)\, dx$$

for all $g \in L^1$. Then Birkhoff's ergodic theorem tells us that, if $f \in L^1$, then

$$\frac{f(x) + f(Tx) + \cdots + f(T^n x)}{n+1}$$

tends to a limit for all x not in some set of measure zero. The standard proof of the ergodic theorem depends on identifying a set of 'good' functions for which the result is easy to prove and which are dense in L^1 and then using a 'maximal inequality' to extend it to all functions in L^1.

We illustrate these ideas in a very simple case, but, for once in this essay, instead of taking about mathematics, we shall actually do some. Our object is to prove the following theorem.

Theorem 1 *Suppose $a_j \in \mathbb{R}$ and $\sum_{j=1}^{\infty} a_j^2$ converges. If $X_1, X_2, \ldots$ are independent identically distributed random variables with*

$$\Pr(X_j = 1) = \Pr(X_j = -1) = 1/2,$$

then $\sum_{j=1}^{\infty} a_j X_j$ converges with probability 1.

Informally, if we have a real sequence a_j with $\sum_{j=1}^{\infty} a_j^2$ convergent, then, if we assign signs at random, $\sum_{j=1}^{\infty} \pm a_j$ will converge with probability 1.

Let us call a sequence a_j 'good' if only finitely many of the a_j are non-zero. It is obvious that every good sequence satisfies the conclusions of Theorem 1. Our object is to use this fact to prove Theorem 1 in general. We use the following 'reflection' lemma.

Lemma 2 *Let a_j and X_j satisfy the conditions of Theorem 1. We write $\alpha = \sup_{j\geq 1} |a_j|$ and take $\lambda > \alpha$.*

(i) *If $N \geq 1$,*

$$\Pr\left(\max_{1\leq n\leq N} \sum_{j=1}^{n} a_j X_j \geq \lambda\right) \leq 2 \Pr\left(\sum_{j=1}^{N} a_j X_j \geq \lambda - \alpha\right).$$

(ii) *If $N \geq 1$,*

$$\Pr\left(\max_{1\leq n\leq N} \left|\sum_{j=1}^{n} a_j X_j\right| \geq \lambda\right) \leq 4 \Pr\left(\sum_{j=1}^{N} a_j X_j \geq \lambda - \alpha\right).$$

(iii) *If* $N \geq 1$,

$$\Pr\left(\max_{1\leq n\leq N}\left|\sum_{j=1}^{n} a_j X_j\right| \geq \lambda\right) \leq 2\frac{\sum_{j=1}^{N} a_j^2}{(\lambda-\alpha)^2}.$$

Proof Let $x_j = \pm 1$. If $\max_{1\leq n\leq N}\sum_{j=1}^{n} a_j x_j \geq \lambda$, then there exists an $1 \leq M \leq N-1$ such that

$$\sum_{j=1}^{M+1} a_j x_j \geq \lambda \quad \text{but} \quad \sum_{j=1}^{m} a_j x_j < \lambda \quad \text{for } 1 \leq m \leq M.$$

If we write $\mu = \sum_{j=1}^{M} a_j x_j$, then, automatically, $\mu \geq \lambda - \alpha$.

We now observe (and this is where the 'reflection' occurs) that

$$\sum_{j=M+1}^{N} a_j x_j \geq 0 \quad \Leftrightarrow \quad \sum_{j=M+1}^{N} a_j(-x_j) \leq 0,$$

so that

$$\sum_{j=1}^{M} a_j x_j + \sum_{j=M+1}^{N} a_j x_j \geq \mu \quad \Leftrightarrow \quad \sum_{j=1}^{M} a_j x_j + \sum_{j=M+1}^{N} a_j(-x_j) \leq \mu.$$

Thus

$$\max\left\{\sum_{j=1}^{M} a_j x_j + \sum_{j=M+1}^{N} a_j x_j, \sum_{j=1}^{M} a_j x_j + \sum_{j=M+1}^{N} a_j(-x_j)\right\} \geq \mu \geq \lambda - \alpha.$$

We have shown that at least half of the possible choices of $x_j = \pm 1$ which yield $\max_{1\leq n\leq N}\sum_{j=1}^{n} a_j x_j \geq \lambda$ also yield $\sum_{j=1}^{N} a_j x_j \geq \lambda - \alpha$. Since the X_j are independent random variables with $\Pr(X_j = 1) = \Pr(X_j = -1) = 1/2$, the required result follows.

(ii) By symmetry

$$\Pr\left(\max_{1\leq n\leq N}\sum_{j=1}^{n} a_j X_j \geq \lambda\right) = \Pr\left(\min_{1\leq n\leq N}\sum_{j=1}^{n} a_j X_j \leq -\lambda\right).$$

(iii) By symmetry and Tchebychev's inequality,

$$2\Pr\left(\sum_{j=1}^{N} a_j X_j \geq \lambda - \alpha\right) = \left(\left|\sum_{j=1}^{N} a_j X_j\right| \geq \lambda - \alpha\right)$$
$$\leq \frac{\operatorname{var} \sum_{j=1}^{N} a_j X_j}{(\lambda-\alpha)^2}$$
$$= \frac{\sum_{j=1}^{N} a_j^2}{(\lambda-\alpha)^2},$$

and the result follows. □

We now introduce the 'maximal function'

$$S^*(\mathbf{X}) = \sup_{1\leq n}\left|\sum_{j=1}^{n} a_j X_j\right|.$$

(Note that S^* can take the value ∞, but, as the next lemma shows, the probability that this occurs is zero.) Although the formal proof involves the generalisation of Lebesgue's theory called 'measure theory' the reader should have no difficulty in accepting that the result follows from Lemma 2.

Lemma 3 *Let a_j and X_j satisfy the conditions of Theorem 1. If $\lambda > \alpha$, where $\alpha = \sup_{j\geq 1}|a_j|$, then*

$$\Pr(S^*(\mathbf{X}) \geq \lambda) \leq 2\frac{\sum_{j=1}^{\infty} a_j^2}{(\lambda-\alpha)^2}.$$

This maximal lemma now gives us a proof of Theorem 1 as follows.

Lemma 4 *Let a_j and X_j satisfy the conditions of Theorem 1.*
(i) *There exists an $N(k)$ such that*

$$\Pr\left(\left|\sum_{j=n}^{m} a_j X_j\right| \leq 2^{-k} \quad \textit{for all } m \geq n \geq N(k)\right) \geq 1 - 2^k.$$

(ii) *If $r \geq 1$, then $\sum_{j=1}^{\infty} a_j X_j$ converges with probability at least $1 - 2^{-r}$.*
(iii) *$\sum_{j=1}^{\infty} a_j X_j$ converges with probability 1.*

Proof (i) Chose $N(k)$ so that $\sum_{j=N(k)}^{\infty} a_j^2 \leq 2^{-3k-6}$ and so, in particular $\sup_{j\geq N(k)}|a_j| \leq 2^{-k-2}$. Consider the sequence b_j defined by $b_j = 0$ for $1 \leq j \leq N(k)$ and $b_j = a_j$ for $j \geq N(k)$. (Note that, if we write $c_j = a_j$ for $1 \leq j \leq N(k)$ and $c_j = 0$ for $j \geq N(k)$, then (c_j) is a 'good' sequence, (b_j) is a sequence which is 'close to the zero sequence' and $a_j = b_j + c_j$.)

We write

$$T^*(\mathbf{X}) = \sup_{1 \le n} \left| \sum_{j=1}^{n} b_j X_j \right|$$

and observe that, if $m \ge n \ge N(k)$, we have

$$\begin{aligned} \left| \sum_{j=n}^{m} a_j X_j \right| &\le \left| \sum_{j=N(k)}^{m} a_j X_j \right| + \left| \sum_{j=N(k)}^{n} a_j X_j \right| \\ &= \left| \sum_{j=N(k)}^{m} b_j X_j \right| + \left| \sum_{j=N(k)}^{n} b_j X_j \right| \\ &\le 2T^*(\mathbf{X}). \end{aligned}$$

By the maximal inequality of Lemma 3,

$$\Pr(T^*(\mathbf{X}) \ge 2^{-k-1}) \le 2 \frac{\sum_{j=1}^{\infty} b_j^2}{(2^{-(k+1)} - 2^{-k-2})^2} \le 2^{-k-1},$$

so, combining the two results obtained in this paragraph,

$$\Pr\left(\left| \sum_{j=n}^{m} a_j X_j \right| \le 2^{-k} \quad \text{for all } m \ge n \ge N(k) \right) \ge 1 - 2^{-k}.$$

(ii) By (i), we can find $N(k)$ such that

$$\Pr\left(\left| \sum_{j=n}^{m} a_j X_j \right| \le 2^{-k} \quad \text{for all } m \ge n \ge N(k) \right) \ge 1 - 2^{-k}$$

for each $k \ge r + 1$. If the probability of an event A_k happening is at most 2^{-k}, then the reader will readily accept (as can be proved formally using measure theory) that the chance of at least one of $A_{r+1}, A_{r+2}, \dots$ occurring is at most $\sum_{k=r+1}^{\infty} 2^{-k} = 2^{-r}$. Thus, with probability at least $1 - 2^{-r}$, we have

$$\left| \sum_{j=n}^{m} a_j X_j \right| \le 2^{-k} \quad \text{for all } m \ge n \ge N(k) \text{ and all } k \ge r + 1.$$

Using the general principle of convergence, it follows, with probability at least $1 - 2^{-r}$, that $\sum_{j=1}^{\infty} a_j X_j$ converges.

(iii) The result follows from the fact that (ii) holds for every $r \ge 1$. $\square$

To see that the convergence is very different from that which we see in elementary analysis, observe that there is a strictly positive probability that all the random variables in a sequence $X_n, X_{n+1}, \dots, X_{n+q}$ take preassigned values.

Exercise 5 Suppose that a_j and X_j satisfy the conditions of Theorem 1 and, in addition, that $\sum_{j=1}^{\infty} |a_j|$ diverges. Show that, given any n, we can find an $m \geq n$ such that

$$\Pr\left(\left|\sum_{j=n}^{m} a_j X_j\right| \geq 1\right) > 0.$$

The repeated successes of the maximal inequality technique prompted Banach to prove a theorem which showed, in effect, that any convergence theorem of the type we are discussing implies a maximal inequality. We need to introduce some notation. If E is a subset of $[-1/2, 1/2]$, we say that E has measure at most K if, given any $\epsilon > 0$, we can cover E with a countable collection of intervals of total length less than $K + \epsilon$. (Strictly speaking, we should talk about 'outer measure' but, for all the sets we talk about, the distinction does not matter. Parallelling our discussion of sets of measure zero, one way of thinking about a set of measure at most K is that, if you drop a dart at random on the interval $[-1/2, 1/2]$, the probability of hitting E is at most K.)

If we write

$$S^*(f, t) = \sup_{n \geq 0} \left| \sum_{j=-n}^{n} \hat{f}(j) \exp(2\pi i j) \right|,$$

then Banach's theorem tells us that Lusin's conjecture is true if and only if there exists a positive function $B \colon \mathbb{R} \to \mathbb{R}$ with $B(\lambda) \to 0$ as $\lambda \to \infty$ such that the measure of the set

$$\{t \in [-1/2, 1/2] \colon S^*(f, t) \geq \lambda \|f\|_2\}$$

is less than $B(\lambda)$. The proof that this inequality implies Lusin's conjecture runs along the lines set out in the proof of Lemma 3 and the proof of the converse is a 'grandchild' of the method by which we proved du Bois-Reymond's result. Although Banach's result is not very deep in itself, it confirms that, if Lusin's conjecture were true, it would be a reasonable strategy to try to prove it via a maximal inequality of the type we have just proved. It turns out that the correct form of B for our particular problem is $B(\lambda) = K\lambda^2$ for some constant K. Rewriting the previous formula, we now know that Lusin's conjecture is true if and only if there exists a K such that the set

$$\{t \in [-1/2, 1/2] \colon S^*(f, t) \geq \lambda\}$$

has measure less than $K\lambda^{-2}\|f\|_2^2$ for all $\lambda > 0$ and all $f \in L^2$.

Simple limiting arguments show that Lusin's conjecture can be restated as follows.

Conjecture 6 *There exists a constant K with the following property. Given N a positive integer, $a_j \in \mathbb{C}$ $[|j| \leq N]$ and $\lambda > 0$, we can find intervals $I_1, I_2, \ldots, I_m$ of* *total length less than* $K\lambda^{-2}\sum_{j=-N}^{N}|a_j|^2$ *such that*

$$\left|\sum_{j=-r}^{r} a_j \exp 2\pi i j t\right| \leq \lambda$$

for all $0 \leq r \leq N$ *and all* $-1/2 \leq t \leq 1/2$ *with* $t \notin \bigcup_{k=1}^{m} I_k$.

Theorem 1 can be easily restated in a non probabilistic way. Recall that, if x is real, we write

$$\operatorname{sgn} x = \begin{cases} 1 & \text{if } x > 0, \\ 0 & \text{if } x = 0, \\ -1 & \text{if } x = 0. \end{cases}$$

We set $H_r(t) = \operatorname{sgn}(\sin 2\pi r t)$. The reader should sketch H_r for $r = 1$, 2, 4 and for $r = 5$.

Theorem 7 *Suppose* $a_j \in \mathbb{R}$ *and* $\sum_{j=1}^{\infty} a_j^2$ *converges. If we define* H_r *as above then* $\sum_{j=1}^{\infty} a_j H_{2^j}(t)$ *converges for all* t *outside a set of measure zero.*

To see why Theorem 7 is just Theorem 1 in disguise, think of the points t where $H_{2^j}(t) = 1$ as corresponding to $X_j = 1$ (the jth throw of a fair coin comes down heads) and the points t where $H_{2^j}(t) = 0$ as corresponding to $X_j = 1$ (the jth throw of a fair coin comes down tails). The set of points where $H_{2^j}(t) = 0$ has measure zero (equivalently have probability zero and correspond to the coin coming down on its edge) and may be ignored.

Theorem 7 suggests very strongly that, if $a_j \in \mathbb{R}$ and $\sum_{j=1}^{\infty} a_j^2$ converges, then $\sum_{j=1}^{\infty} a_j \sin(2\pi 2^j t)$ will converge for all t outside a set of measure zero. Readers who try to prove this may already find the task harder than they expect, but it will be helpful to add an extra layer of complexity by considering the following almost equally plausible result: If $a_j \in \mathbb{R}$ and $\sum_{j=1}^{\infty} a_j^2$ converges, then $\sum_{j=1}^{\infty} a_j \sin(2\pi(2^j + 1)t)$ will converge for all t outside a set of measure zero.

Why is the new problem harder than the old? When we considered $\sum_{j=1}^{\infty} a_j X_j$, knowing the values of $X_1, X_2, \dots, X_n$ told us nothing about the value of X_{n+1} (or indeed the sequence $X_{n+1}, X_{n+2}, \dots$). In the same way when we considered $\sum_{j=1}^{\infty} a_j H_{2^j}(t)$, then, with some trivial exceptions, knowledge of $H_{2^n}(t)$ told us nothing about the value of $H_{2^{n+1}}(t)$. On the other hand, if we know $\sin(2\pi(2^n + 1)t)$ exactly, then we know that t must take one of a finite number of values and so the sequence $\sin(2\pi(2^n + 1)t)$ must be one of a finite number of possible sequences.

Matters are not quite so bad as they seem. If we only know the value of $\sin(2\pi(2^n + 1)t)$ to within an accuracy δ, then, speaking very roughly, we only know t to within an accuracy of $2^{-n}\delta$ and so when $2^{m-n}\delta$ is substantially larger than 1, we know practically nothing about $\sin(2\pi(2^m + 1)t)$. This suggests that if $\beta > 1$, $n(r+1)/n(r) > \beta$ the sequence of functions $s_{n(r)}$ given by $s_{n(r)}(t) = \sin(2\pi n(r)t)$

behaves like a sequence of 'almost independent random variables' with terms far apart in the sequence being 'essentially independent'. Although this heuristic argument does not point to the best proof, the result it suggests is true. If $\beta > 1$ and $n(r+1)/n(r) > \beta$ for all $j \geq 1$, then if $a_j \in \mathbb{C}$ and $\sum_{j=-\infty}^{\infty} |a_j|^2$ converges then, as $r \to \infty$, $\sum_{j=-r}^{r} a_j \sin(2\pi n(j)t)$ converges outside a set of measure zero. Similarly, the same hypotheses imply that $\sum_{j=-r}^{r} a_j \exp(in(j)t)$.

If we were very adventurous and prepared to wave our hand very wildly indeed, we could go further and, instead of talking about the $\exp(in(j)t)$ as being in some vague sense 'almost independent', we might hope that 'block sums'

$$B_r(t) = \sum_{n(r+1) \geq |j| \geq n(r)+1} a_j \exp(2\pi ijt)$$

might be almost independent in the sense that, if we work to the appropriate level of accuracy, knowledge of B_r tells us very little about B_s when r is not close to s. Whether the arm waving is justified or not, Kolmogorov proved that, if $\beta > 1$ and $n(r+1)/n(r) > \beta$ for all $j \geq 1$, then if $a_j \in \mathbb{C}$ and $\sum_{j=-\infty}^{\infty} |a_j|^2$ converges then, as $r \to \infty$, $\sum_{j=-n(r)}^{n(r)} a_j \exp(2\pi ijt)$ converges outside a set of measure zero. Because of the close link between square summable sequences and L^2, this tells us at once that

$$\sum_{j=-n(r)}^{n(r)} \hat{f}(j) \exp(2\pi ijt) \to f(t)$$

outside a set of measure zero for all $f \in L^2$.

The example of

$$g_{2^m+r}(t) = \begin{cases} n & \text{for } 2^{-1} + r2^{-m} \leq t \leq 2^{-1} + (r+1)2^{-m} \text{ and } 0 \leq r \leq 2^m - 1, \\ 0 & \text{otherwise,} \end{cases}$$

which we discussed earlier, shows that the result we have just described can not be used to prove Lusin's conjecture.

Since we have followed what now seems the natural path from sums of random variables to Fourier sums, it is interesting to note that we have reversed the historic path. As Kolmogorov recalled 'Such topics as conditions for the validity of the law of large numbers and conditions for convergence of series of independent random variables were actually tackled by methods developed by Lusin and his pupils in the general theory of trigonometric series'.

At the end of the 1950's, two giants of the subject published treatises summarising the state of Fourier Analysis. Bari's *A Treatise on Trigonometric Series* [1] ran to 900 pages and Zygmund's second edition of *Trigonometric Series* [3] to 700. Both authors devoted a substantial part of their works to aspects of the problem of convergence of Fourier sums. Zygmund's introduction specifies 'the problem of the existence of a continuous function with everywhere divergent series' as one of the two main open problems in the Fourier Analysis but does not mention Lusin's conjecture. Bari gives a discussion of the reasons (needless to say substantially deeper than those I have discussed) which, she believed, led Lusin to his conjecture but states that '[these] arguments ... cannot have any force today'.

It came as an enormous surprise when, in 1966, Carleson announced his proof of Lusin's conjecture. Although Carleson was already known for a spectacular proof of another classical conjecture (the Corona theorem which we discuss later) many mathematicians were dubious. In my fourth year of university studies, I inherited a set of notes on Hilbert space containing the marginal notation 'A Swede called Carleson claims to have proved pointwise convergence but nobody believes him'. (The note must have been written in 1967.)

Even when it became clear that the proof was correct, the surprise remained. Carleson had essentially proved Lusin's conjecture by a frontal attack on the version given in Conjecture 6. But, although no one would have thought of stating Conjecture 6 until the 1930's, its statement would have been understood by Fourier and the proof of Lusin's conjecture (at least in so far as it concerns continuous functions) from Conjecture 6 would have been understood by Dirichlet.

In our discussion we saw that it was reasonable to divide the sum into blocks of the type

$$\sum_{j=2^n}^{2^{n+1}-1} a_j \exp(2\pi i j t) = \exp(2\pi i 2^n t) \sum_{j=0}^{2^n} a_{j-2^n} \exp(2\pi i j t),$$

which we may think of as 'frequency blocks' at scale 2^n. Carleson's proof divides the interval $[-1/2, 1/2]$ into intervals $[r2^{-n}, (r+1)2^{-n}]$ which we may think of as space blocks at scale 2^{-n}. It will, I hope, appear natural that we try to study the behaviour of frequency blocks at scale 2^n on space blocks at scale 2^{-n}. If the behaviour on a space block at scale 2^{-n} is good, then we retain it. If it is bad, then it will be one of the intervals I_k rejected in Conjecture 6. If we cannot decide, then we split the interval in two and examine the results at the appropriate new scale.

In his address [4] to the Nice International Congress, Hunt shows how the proof works for a model system which bears much the same relation to the trigonometric problem as the sum $\sum_{j=1}^{\infty} a_j H_{2^j}(t)$ in Theorem 7 does to the sum $\sum_{j=1}^{\infty} a_j \sin(2\pi(2^j+1)t)$. The reader who wishes to go further is recommended to consult this expository tour de force. However, even knowing how the proof works in the model case, it is entirely unclear how or if it can be made to work in a system where the 'exact' relations of the model must be replace by the 'approximate' relations of the trigonometric case. Carleson's proof overcomes these difficulties in a magnificent example of what mathematicians call 'power' and which they value far above 'cleverness'.

The difficulty of his paper is not simply due to the depth of the ideas but also to an expository problem which dogs all papers of this type. As I tried to indicate earlier the 'strategic' vision of the proof is useless without a large number of 'tactical' decisions to overcome various problems as they arise. However, there are often several ways of tackling the difficulty. As a simple example, it does not matter whether the

C in Conjecture 6 is 10 or 10^{10}, but readers may well be baffled if we use this freedom to make estimates which are weaker than they know are available. (At a trivial level, suppose that a real variable proof requires only that $(a+b)^2 \le K(a^2+b^2)$ for some constant K. If we write

$$(a+b)^2 \le \big(2\max\{|a|,|b|\}\big)^2 = 4\max\{a^2,b^2\} \le 4(a^2+b^2),$$

the reader who spots a better estimate may well be worried.) A paper which requires a large number of decisions, where more than one 'correct choice' is possible but no choice is 'natural', will always be difficult to read. Hunt's lecture to the Nice Conference is made much easier because the model system chosen has a fair number of 'natural choices'.

Since every continuous function is an L^2 function, Carleson's theorem settles the convergence problem as it would have appeared to Fourier and his successors but, in view of the difficulty of Carleson's L^2 proof, it may be asked whether it might not be possible to extract an easier proof which merely applies to continuous functions. Two years after Carleson's result, Kahane and Katznelson produced a construction (which may not have been an example of 'power' but was certainly one of extreme 'cleverness') which showed that given any set E of measure zero there exists a continuous function f whose Fourier sum diverged at every point of E (and possibly others). Thus, at least from this point of view, the Fourier sums of continuous functions exhibit the same wildness as general L^2 functions. Although the theorems of mathematicians can, usually, be relied on, as we have seen, their opinions are just opinions. However, it is hard to see how a simpler 'continuous function proof' would work.

In theory, mathematical proofs, even those as complicated as Carleson's can be checked line by line. However, it is much more satisfactory if the underlying ideas can be tested by applying them in new situations. Hunt showed that Carleson's results applied not merely to those L^1 functions for which $\int_{-1/2}^{1/2} |f(t)|^2\,dt < \infty$ but also to the much larger class of L^1 functions for which $\int_{-1/2}^{1/2} |f(t)|^p\,dt < \infty$ for some $p > 1$ and Sjölin carried this idea even further.

Many of the ideas and methods introduced by Carleson (for example in the solution of the Corona problem) were quickly absorbed into main stream of mathematics. (Some received that rather backhanded tribute that the goddess of mathematics pays to favoured worshipers by disappearing into the unattributed common stock.) In the case of the convergence theorem it looked for some time as though no successor could be found to bend the bow of Ulysses. However a new generation of brilliant harmonic analysts have taken up Carleson's ideas and developed them further.

It is much harder to give a clear idea of the other major problems tackled by Carleson using only the ideas available to a first or second year mathematics undergraduate and from now on my discussion will be both briefer and much vaguer.

Earlier I said that, for a short time, it was possible to hope that sets of measure zero could always be ignored. For example, it was known that for many sets E of measure zero the statement

$$\sum_{j=-\infty}^{\infty} a_j \exp(2\pi ijt) = 0 \text{ for } t \notin E$$

implied that $a_j = 0$ for all j and it was expected that this result would turn out to be true for all sets E of measure zero. In 1917, Mensov showed that this was not the case and mathematicians began to study the sets of measure zero in earnest.

Physicists have always been happy to consider 'point masses' and 'charges living on the surface of a sphere'. Lebesgue's successors created a theory of 'general measures' which included such objects as special cases. It turned out these 'general measures' were ideal tools for studying continuous functions (so that, in some sense, if we knew everything about general measures, we would know everything about continuous functions and vice-versa). Just as a point mass lives on one point, so many important general measures live on sets of measure zero (in the Lebesgue sense). Much of Carleson's early work is concerned with the study of sets of Lebesgue measure zero and the general measures that live on them. His results left a permanent mark on the subject, but he soon moved into other fields.

Much of mathematics follows the kind of pattern we have sketched out. A concrete problem turns out to require new tools for its solution. Reflection shows that these new tools are themselves concrete 'graspable' objects and their study raises new concrete problems. However, it is often profitable to seek theories which join many different concrete problems and their solutions. These abstract theories often provide more powerful tools or suggest new problems.

Consider the following problems.

Problem A. Given $a_j \in \mathbb{C}$ such that $\sum_{n=0}^{\infty} a_n \exp(2\pi int)$ converges for all real t, when can we find $b_j \in \mathbb{C}$ such that $\sum_{n=0}^{\infty} b_j \exp(2\pi int)$ converges for all real t and

$$\sum_{n=0}^{\infty} a_j \exp(2\pi int) \sum_{n=0}^{\infty} b_n \exp(2\pi int) = 1$$

for all real t.

Problem B. Given $a_n \in \mathbb{C}$ such that $\sum_{n=-\infty}^{\infty} |a_n|$ converges, when can we find $b_n \in \mathbb{C}$ such that $\sum_{n=-\infty}^{\infty} |b_n|$ converges,

$$\sum_{j=-\infty}^{\infty} a_{r-j} b_j = 0 \quad \text{for } r \neq 0 \text{ and } \sum_{j=-\infty}^{\infty} a_{-j} b_j = 1.$$

In 1820 mathematicians would have understood both problems (though they might have considered them a bit odd) but would not have been able to resolve them. By 1850, the discovery of complex analysis had produced a simple answer to Problem A. Since $\sum_{n=0}^{\infty} a_n z^n$ converges for all $|z| = 1$, the power series $\sum_{n=0}^{\infty} a_n z^n$ converges for all $|z| \leq 1$ to a function f which is analytic in the open disc $\{z\colon\ |z| < 1\}$ and continuous on $\{z\colon\ |z| < 1\}$. Conversely, if a function g is analytic in the open disc $\{z\colon\ |z| < 1\}$ and continuous on the closed disc $\{z\colon\ |z| < 1\}$, then we can write

$g(z)=\sum_{n=0}^{\infty} b_n z^n$ for $|z|\le 1$ and $\sum_{n=0}^{\infty} b_n \exp(2\pi n t)$ will converge for all t. With a little extra work it is now possible to restate Problem A as follows.
Problem A′. Given a function f analytic in the open disc $\{z\colon\ |z|<1\}$ and continuous on the closed disc $\{z\colon\ |z|\le 1\}$, when can we find a function g analytic in the open disc $\{z\colon\ |z|<1\}$ and continuous on $\{z\colon\ |z|\le 1\}$ such that $f(z)g(z)=1$ for all $|z|\le 1$.

Elementary theorems of complex analysis tell us that we can find the required g if and only $f(z)\neq 0$ for all $|z|\le 1$.

Problem B was resolved by Wiener in the 1930's. Primed by our discussion of Problem A, the reader will have no difficulty as recasting it in the following form.
Problem B′. Given $a_n\in\mathbb{C}$ such that $\sum_{n=-\infty}^{\infty}|a_n|$ converges, when can we find $b_n\in\mathbb{C}$ such that $\sum_{n=-\infty}^{\infty}|b_n|$ converges and

$$\sum_{n=0}^{\infty} a_n \exp(2\pi i n t)\sum_{m=0}^{\infty} b_m \exp(2\pi i m t)=1$$

for all real t.

Wiener showed that we can find the required b_j if and only if

$$\sum_{n=-\infty}^{\infty} a_n \exp(2\pi i n t)\neq 0$$

for all real t. We may recast his result as follows.

Theorem 8 *Given $a_n\in\mathbb{C}$ such that $\sum_{n=-\infty}^{\infty}|a_n|$ converges, we can find $b_n\in\mathbb{C}$ such that $\sum_{n=-\infty}^{\infty}|b_n|$ and*

$$\sum_{n=-\infty}^{\infty} a_n z^n \sum_{m=-\infty}^{\infty} b_m z^m = 1$$

if and only if $\sum_{n=-\infty}^{\infty} a_n z^n \neq 0$ for all $|z|=1$.

Here is a closely related result which reflects our answer to Problem A.

Theorem 9 *Given $a_n\in\mathbb{C}$ such that $\sum_{n=0}^{\infty}|a_n|$ converges, we can find $b_n\in\mathbb{C}$ such that $\sum_{n=0}^{\infty}|b_n|$ converges and*

$$\sum_{n=-\infty}^{\infty} a_n z^n \sum_{m=-\infty}^{\infty} b_m z^m = 1$$

if and only if $\sum_{n=-\infty}^{\infty} a_n z^n \neq 0$ for all $|z|\le 1$.

In the 1940's Gelfand and others showed that a wide range of similar problems had a similar solution. Given an appropriate collection A of complex valued functions on a space X we can find a space $Y\supseteq X$ and an extension of each $f\in A$

to a function $\tilde{f}$ on Y such that the following result is true. Given an $f \in A$ we can find a $g \in A$ with $f(x)g(x) = 1$ for all $x \in X$ if and only if $\tilde{f}(y) \neq 0$ for all $y \in Y$. Gelfand's theory is a marvellous display of the power of abstract methods but, though the theory tells us that the appropriate Y always exists, it does not always tell us how to find it.

If we look at Theorems 8 and 9, we see that they both concern similar collections of functions which live on the boundary $X = \{z\colon\ |z| = 1\}$ of the unit disc but in one case $Y = X$ whilst in the other case we have $Y = \{z\colon\ |z| \le 1\}$, the entire closed unit disc. In more general problems the extending space Y may not be readily describable in terms of X.

Readers are strongly advised to work through the next exercise.

Exercise 10 Consider the collection A of uniformly continuous functions $f\colon [0, 1] \cap \mathbb{Q} \to \mathbb{C}$ (that is to say, the set of uniformly continuous functions on the space X of rational numbers in $[0, 1]$).

(i) Given an $f \in A$, we can find a $g \in A$ with $f(x)g(x) = 1$ for all $x \in X$ if and only if the continuous extension $\tilde{f}$ of f to $Y = [0, 1]$ is nowhere zero.

(ii) Given an $f \in A$, we can find a $g \in A$ with $f(x)g(x) = 1$ for all $x \in X$ if and only if there exists a $\delta > 0$ such that $|f(x)| > \delta$ for all $x \in X$.

Here we have an example in which, although $Y \neq X$, we have X dense in Y and (as we have discussed before) knowledge of X gives us a very strong hold on the behaviour of Y.

The Corona theorem concerns the space H^∞ of functions $f(z) = \sum_{n=0}^{\infty} a_n z^n$ where the series converges for all $|z| < 1$ to a bounded function. In our notation, $X = \{z\colon\ |z| < 1\}$ and $A = H^\infty$. Anyone who has a slight acquaintance of the unpleasant behaviour of power series on their circle of convergence will not be surprised to learn that A contains functions which cannot be extended continuously to $D = \{z\colon\ |z| \le 1\}$ and that there is no way in which we can identify the Gelfand space Y with X, D or any other subset of $\mathbb{C}$. However, if X were dense in Y (in the appropriate sense), we would have very strong hold on the behaviour of Y.

The reader must be warned that when I speak of 'X being dense in Y in the appropriate sense', I am probably (I hope, for the first time in this essay) overstretching analogy, but this rather abstract formulation can be replaced by the following concrete result (compare parts (i) and (ii) of Exercise 10).

Theorem 11 *If $f_1, f_2, \dots, f_n \in H^\infty$ and there exists a $\delta > 0$ such that $\sum_{j=1}^{n} |f_j(z)| \ge \delta$ for all $|z| < 1$ then we can find $f_1, f_2, \dots, f_n \in H^\infty$ such that*

$$\sum_{j=1}^{n} f_j(z) g_j(z) = 1$$

for all $|z| < 1$.

This is Carleson's Corona Theorem. Important though the theorem is, the 'Carleson measures' that he introduced in order to prove it have turned out to be even more important and now appear in many branches of analysis.

It is characteristic of major mathematicians that they move to new fields and attack new problems throughout their career. At an age when most mathematicians have retired or, at least settled into routine teaching and research, Carleson helped establish the existence of 'strange attractors in the Hénon family of planar maps'. Incomprehensible as the phrase sounds it may be placed on the map of humanity's intellectual interests.

Mathematicians, like humanity in general, have always been interested in the long term behaviour of systems. For a mathematician this often means the study of differential equations like $\dot{x}(t) = f\big(x(t), t\big)$. Unfortunately such equations rarely have explicit solutions, so we must settle for numerical computation (which, at best, tell us about one particular case) or seek to establish general properties of the solutions of a given type of differential equation.

The natural way to seek such general principles is to look at those differential equations that we can solve and at actual physical systems described by differential equations of the appropriate type. The main class of differential equations that we can solve exactly is the class of linear differential equations with constant coefficients, that is to say, differential equations like

$$\ddot{x} + a\dot{x} + bx = 0.$$

The solutions of such equations either explode, settle down or oscillate in a periodic fashion.

If we think of a physical model, we see that, in explosions, points which start off close together separate very rapidly. Small changes in initial conditions very rapidly produce very different solutions and long term behaviour is essentially unpredictable. If you try to balance a billiard cue on its tip then, although there must be some position of balance, the slightest deviation rapidly leads to disaster. We speak of 'unpredictability', but it is better to think in terms of a soothsayer who will tell us our fortune one second into the future for 10 euros, two seconds into the future for 100 euros, three seconds into the future for 1000 euros and so on. In theory, the soothsayer will look one minute into the future, if we so desire, but in practice we cannot afford the fee.

If the system frictional, then everything settles down and small changes in initial conditions make very little difference. If we drop two steel balls fairly close together at much the same time into a deep vat of treacle (so we model an equation like $\ddot{x} + a\dot{x} + bx = c$), the two balls will remain close together as they fall through the treacle.

The development of electrical engineering greatly extended the variety of physical models which could be drawn on by mathematicians. In the 1920's and 30's this was exploited by Van der Pol who investigated the equation

$$\ddot{x} - k(1 - x^2)\dot{x} + x = E\sin(\omega t)$$

experimentally. (The term $k(1-x^2)$ can be thought of as the resistance of the circuit, but it is allowed to be *negative*.) Van der Pol observed many strange phenomena. The system could settle down to periodic behaviour but the period could be a proper integer multiple of the period of the forcing term $E\sin(\omega t)$. If the parameters of the circuit changed this, the period of the system would jump discontinuously to another period (and the new period might depend not only on the present values of the parameters but on the way in which the present values had been arrived at). The 'mathematical reality' of these observations was proved by Cartwright and Littlewood in a paper famous for its depth, length and difficulty.

Any proper discussion of these topics would need to distinguish carefully between what was known (or guessed) by particular individuals (such as Poincaré and Littlewood) or even particular mathematical schools (such as the Russian groups studying non-linearity) and what was generally known by the mathematical community. However, from the point of view of the scientific community as a whole, a turning point came with the work of Lorenz published in 1963.

Lorenz set up a system of differential equations to provide a simple model of atmospheric convection and then solved them numerically using an early desk top computer. He observed that the solutions seemed to settle down to some sort of periodic repetition only to suddenly jump to new almost periodic pattern. The new pattern would persist for some time and then the system would jump again and so on. The jumps appeared random or, more properly, showed extreme sensitivity to initial conditions.

Mathematicians were of course aware of the phenomenon of 'extreme sensitivity to initial conditions' in cases like explosions or balancing billiard cues. However Lorenz's system exhibited apparent stability (the fairly periodic patterns) combined with real instability (the essentially unpredictable jumps) in a totally unexpected way.

Differential equations have discrete analogues like $x_{n+2} = F(x_{n+1}, x_n)$. (Thus we start with x_0 and x_1, compute $x_2 = F(x_0, x_1)$ as $x_3 = F(x_1, x_2)$ and so on.) Unfortunately it proved as difficult to find explicit solutions of such equations as it is to find explicit solutions of the corresponding differential equations and much more difficult to find physical analogues. For these reasons, difference equations were very much regarded as the poor relations of differential equations.

All this changed when electronic computers were introduced. When we calculate (an approximation to) the solution of a differential equation on a computer, we do so by replacing the differential equation by a difference equation and solving the difference equation step by step. Even more importantly, it became possible for anyone with access to a hand calculator to follow the behaviour of a particular solution through many hundred steps and anyone with a computer to follow it through millions of steps. This *experimental mathematics* gave the same sort of insight that a physical experiment might give to a classical mathematician. It is a pity that the word 'chaos' is used in this connection because if the picture revealed by these numerical experiments had been truly chaotic, it would have been of little interest to mathematicians. Instead the experimenters saw a world full of unexplained patterns, and it became the job of mathematicians to find out if these patterns were truly there (they might be artifacts caused by subtle problems with the experiments) and, if so, to explain those patterns.

Mathematical progress seldom follows a unique path. Difference equations may be considered a special case of iteration in which we take a map T from a set X to itself and consider the *orbit* $x, Tx, T^2x, \dots$ traced out by images of x under repeated application of T. (In other words, we take $x_0 = x$ and consider the sequence x_0, x_1, $x_2, \dots$, with $x_{n+1} = Tx_n$.) Birkhoff's pointwise ergodic theorem which I mentioned earlier, is an example of a very deep theorem on iteration.

At the beginning of the 20th century, the French mathematicians Julia and Fatou investigated iteration when $X = \mathbb{C}$ and T took various simple forms. The following result was already known and mathematical readers may enjoy treating it as an exercise.

Exercise 12 Let $\mathbb{C}^* = \mathbb{C} \cup \{\infty\}$. Consider the collection $\mathcal{M}$ of Möbius maps $T\colon \mathbb{C}^* \to \mathbb{C}^*$ given by

$$Tz = \frac{az+b}{cz+d}$$

with $ad - bc \neq 0$.

(i) Show that if a Möbius map is not the identity, then it has one or two fixed points.

(ii) Suppose $T \in \mathcal{M}$ has exactly two fixed points z_1 and z_2. If $S \in \mathcal{M}$ and $Sz_1 = 0$, $Sz_2 = \infty$, show that STS^{-1} is a Möbius map which fixes 0 and ∞. Conclude that $STS^{-1}(z) = Az$ for some $A \in \mathbb{C}$ with $A \neq 0$.

(iii) If $|A| > 1$ show that, unless $w = z_1$, we have $T^n w \to z_2$. What happens if $|A| < 1$? What happens if $A = \exp(2\pi\alpha)$ and α is rational? What happens if $A = \exp(2\pi\alpha)$ and α is irrational?

(iv) Suppose $T \in \mathcal{M}$ has exactly one fixed point z_1. Show that there is a Möbius map S such that STS^{-1} fixes ∞. Show that $STS^{-1}(z) = z + B$ with $B \neq 0$. Conclude that, if $w \in \mathbb{C}^*$, we have $T^n w \to z_1$.

We might expect that similar, but slightly more complicated, results will hold for similar, but slightly more complicated, families of functions. In fact, as Julia, Fatou and their successors showed, the moment we consider slightly more complicated functions the behaviour of iterates becomes very much more complicated. A deservedly popular example (though coming from yet another intellectual source) is the logistic map.

Exercise 13 Consider the system $x_{n+1} = rx_n(1 - x_n)$. Use a pocket calculator or a computer to trace the behaviour of the system for various values of r and various initial values of n. (There are Internet programs that will do this for you, but it is more interesting to retrace the steps of the pioneers and do it yourself.)

(i) For $0 \leq r \leq 1$ you should find (and it is easy to prove) that $x_n \to 0$.

(ii) For $1 < r \leq 3$ you should find (and it is easy to prove) that x_n tends to a unique value.

(iii) For r a little bigger than 3, x_n will (usually) oscillate between two values.

(iv) For $r = 3.5$, x_n will (usually) oscillate between four values.

(v) As r increases beyond this, matters rapidly become very complicated. (Of course, looking up the 'logistic map' on the Internet will produce clear accounts of the matter but the readers may well prefer to see how confusing things are before doing this.) Try $r = 3.9$.

The Hénon map is a map $T : \mathbb{R}^2 \to \mathbb{R}^2$ given by

$$T(x, y) = (y + 1 - ax^2, bx).$$

It was inspired by the Lorenz differential equation and also appeared to exhibit very odd properties for certain values of a and b. For these values the iterates $\mathbf{x}_{n+1} = T\mathbf{x}_n$ either move off to infinity or move towards a well defined infinite set A. However, the sequence does not settle down towards one point of A but moves around the entire set. The following artificial example may give some idea of what is involved.

Exercise 14 Consider the map $S : \mathbb{C} \to \mathbb{C}$ given by

$$Sz = \frac{z}{|z|^{1/2}} \exp(i)$$

for $z \neq 0$, $S0 = 0$. If $z_0 \neq 0$ and $z_{n+1} = Sz_n$, describe the behaviour of the sequences $z_0, z_1, z_2, \ldots$.

However, the behaviour of the Hénon map is much more interesting than that of our example, since it appears to exhibit sensitive dependence on initial conditions. In other words, although we can be sure that after a large number of iterations we will be close to the set A, we cannot tell which part of A we will be close to.

I said that the Hénon map 'appeared to exhibit very odd properties' since, as the reader will appreciate, it is very hard to distinguish between 'nearly chaotic' behaviour and 'truly chaotic' behaviour. There is a further point of difficulty. Even if we could prove that what seems to happen actually happens for some particular values of a and b, this might be a property of the specially chosen a and b. Since in 'real life' we can only specify a and b to a certain accuracy, this would not really settle the question of whether what we seem to see is actually taking place. What Carleson and Benedicks did is to show that it really does take place over a substantial range of choices of a and b.

As the prize citation makes clear, Carleson has served mathematics in many other ways, through teaching, through encouragement of younger mathematicians, through influential books, service on committees, and much else. However, mathematicians are remembered for their theorems and not for their other services to mathematics. This is as it should be, but as one of those who benefited from Carleson's revival of the Mittag-Leffler Institute, I should add my thanks for nine gloriously happy months I spent there as a young mathematician.

So far, I have been talking to the tyro mathematician. Now let me address more general readers. In my mind's eye I see those readers in turn as engineers who wonder whether the matters considered here have any real connection with nature, politicians who wonder if they have any real use and lay people who wonder if they have any real interest.

Engineers may use the while worn jibe that they 'would never fly an aeroplane which depended on the study of sets of measure zero' (though, oddly enough, they are quite happy to fly warplanes which depend on imaginary numbers). They may claim that they use Fourier methods all the time but the kind of things we have discussed are 'pathological' and cannot bear any relation to the matters discussed here.

It is, of course, true that, at university, students 'solve' the problem of the plucked violin string by writing the form of the string as a Fourier sum with an infinite number of terms and then performing various 'formal' (that is to say 'magical') operations but the object of the exercise is to pass examinations and not to find out what actually happens to the string. In reality we must accept that we do not know the exact shape of the string at any time. Thus our task is to show that if we start with a shape which is close to our 'idealised initial shape' then our prediction will be close to the observed shape and in order to do this we must decide what it means for one function to be close to another, that is to say we must define a distance between functions.

In our previous discussion we used the fact that we could find nice functions arbitrarily close to nasty functions. Unfortunately the reverse is also true for all reasonable notions of distance. So long as we do not perform any numerical calculations, we can shut our eyes to this disagreeable fact but the moment we engage in any serious attempt to make numerical predictions (as in weather forecasting) what was merely 'some highbrow pure mathematical concern' becomes a major problem for the numerical analyst. The du Bois-Reymond example suggests that simply taking the maximum pointwise difference, that is to say, using the distance

$$d_\infty(f, g) = \sup_{-1/2 \le t \le 1/2} |f(t) - g(t)|$$

may not always be the best way forward and more direct evidence from numerical analysis confirms this.

In many circumstances the best notion of distance to use is

$$d_2(f, g) = \|f - g\|_2 = \left(\int_{-1/2}^{1/2} |f(t) - g(t)|^2 \, dt \right)^{1/2}$$

since, in some sense, it represents the 'power difference between two signals f and g'. Engineers who use this metric may claim that they are only interested in the case when f and g are continuous. But this is exactly the same as claiming that we are only interested in rational numbers and so the real numbers are of no concern to us. Just as ordinary calculus involves all real numbers whether we like it or not, any serious calculus using d_2 automatically involves all L^2 functions whether we acknowledge this or not.

There is another way in which we can try to avoid mathematical problems. It is to assert that Nature only deals in smooth functions. Unfortunately this is not always true. It is a common experience that methods of numerical analysis designed to take advantage of high differentiability rarely work on real data, and I would certainly refuse to cross a bridge designed by an engineer who believed that wind speed varied in a nice smooth manner. The behaviour of noise in electrical circuits and the prices of stocks and shares may indeed be modeled by continuous functions, but it is notorious that those functions are nowhere differentiable. (It is worth noting that they also have an infinite number of maxima and minima in each interval so they fail to satisfy the conditions of Dirichlet's theorem.)

None of this means that engineers have to study Carleson's theorem, but it does suggest that they should be aware of the issues raised. After all, it shows a certain disrespect for Nature to think that all the problems she sets us can be resolved by the methods of the 19th century.

Next I address politicians who ask why society should pay for mathematicians to study the matters discussed here. Let me say straight away that I know no way in which Carleson's L^2 theorem contributes anything to the material satisfactions of mankind. For several hundred years the mathematicians of Europe pursued the goal of solving high order polynomial equations by radicals (essentially this just means finding a simple formula for the solution of standard equations). To the non-mathematical reader this must seem an obviously useful goal and so it seemed to mathematicians for the first couple of centuries. However, by the time Abel, Galois and others resolved the problem by showing that no simple formula can exist, it was clear that the answer was irrelevant for all practical purposes. In the same way, although throughout the nineteenth century most users of Fourier series thought the resolution of the pointwise convergence problem would have wide ranging practical use, the twentieth century has shown that, in fact, most practical applications involve notions of the distance between functions and appropriate approximation rather than pointwise convergence.

Jacobi wrote that Fourier reproached 'Abel and myself for not having given priority to our research in the theory of heat conduction. ... It is true that Fourier was of the opinion that the chief end of mathematics was the public good and the explanation of natural phenomena; but a philosopher such as he was should have known that the only goal is the honour of the human spirit and in this respect a question in the theory of the numbers is as valuable as a problem in physics'. With the complete, unexpected and beautiful resolution of a central problem in mathematics as well as his other successful solution of other major problems Carleson has honoured the human spirit and it is entirely appropriate that he should receive a prize named after another great upholder of the human spirit.

When we speak of great mathematics and great mathematicians, we may surely say that a society without mathematicians like a society without poets would not be aware of their absence but would, none the less, be a poorer society.

The politician may grant the truth of all I have said but remark that very few of the mathematicians supported by society reach the heights of a Carleson. Although most mathematicians do mathematics because they enjoy it and very few do it for the public good, it is possible to make a plausible defence of pure mathematics (that is to say mathematics for its own sake) on the grounds that, historically, some pure mathematics has later turned out to be useful. For example, although the problem of solution by radicals turned out to have no practical use, the methods developed to solve it gave rise to group theory (useful in Quantum Mechanics, crystallography and World War II code breaking) and the theory of fields (useful in communication theory).

One problem with such a historical defence is that it may come to resemble one of those patriotic histories in which everything is invented by an Englishman (or Russian or Frenchman according to taste). In fact ideas and methods flow in both directions across the ill defined frontiers of mathematics, physics and engineering. One can invent a history in which all the ideas of mathematics come from mathematicians but in reality mathematics has benefited from major contributions from physics and engineering and vice versa.

Having said all this, the reader may still be interested to know that Carleson's Corona theorem has fairly direct links with problems in control theory (which deals with the control of machines, electrical circuits and so on). The type of 'hard classical' Fourier analysis pursued by Carleson played a vital role in the emergence of Wavelet Theory which provides a new way of analysing complex data like photograph. (Wavelets are used in many areas of physics, engineering and biology but readers may be more interested to learn of their use in restoring the 'true sound' from old gramophone records.)

When we discussed strange attractors, we noted that Lorenz's work arose in considering weather forecasting. Although he worked with a 'toy model', the behaviour it showed, provided meteorologists with insight into why weather forecasting must have a time horizon beyond which 'sensitivity to small changes' (the so called 'butterfly effect') make accurate forecasting impractical. Developing this insight, they have realised that this time horizon is not fixed but depends on the state of the weather. Sometimes the weather permits accurate forecasting over long periods of time and sometimes it does not. Weather forecasters run their computer models many times making slight changes to the initial conditions and observe how fast the various forecasts spread out (ensemble forecasting). If they diverge rapidly, forecasts will probably only be accurate over a short time. If they diverge slowly, the forecasts can probably be relied on over a much longer period.

The outpouring of popular presentations of 'chaos theory' has obscured a fact known to aeronautical engineers since the earliest days of human flight. Insensitivity to small changes gives stability and predictability but reduces controlability. Early gliders were built for stability with the result that things went wrong only rarely but when they went wrong they stayed wrong. Controlability requires that only small changes are needed to give large effects. It is nearly impossible to balance a billiard cue on a table but very small motions allow me to keep it vertical on the tip of one finger. Provided that we do not overdo matters, 'chaotic' situations may represent an opportunity rather than a problem.

An interesting example of this comes from space travel. Since the time of Newton, mathematicians have been interested in the long term behaviour of the solar system. (The reader may say that it must be stable, otherwise we would not be here, but Newton and his followers had the option of invoking a clockmaker to reset the clock when it went badly wrong.) At the end of the 19th century, it is probable that most people who thought about the matter would have said that the solar system and similar objects were stable though mathematics could not yet prove it. As a result of the work of Poincaré, Kolmogorov and many others we might give a more nuanced answer. Most of the time, most systems of the type we are interested in show high stability (small changes in position and velocity make very little difference to the behaviour of the system even over long periods of time) but situations can arise (for example in near collisions) in which small changes can produce very different outcomes. If we deliberately place our spacecraft in such a situation, then a very small expenditure of energy can achieve changes in direction and velocity which would otherwise be impossible. This technique is routinely used for projects such as *Voyager 1*.

Modern technology has added to the number of complex systems whose long term behaviour is of obvious interest. Internet type systems which link many computers may work in satisfactory mode for several days and then freeze. Traffic on motorways can flow freely and then suddenly jam for no obvious reason. Using a mixture of advanced mathematics and 'engineering intuition' we can make sense of what happens to Internet systems and at least delay their collapse. Motorway traffic involves human beings and raises non-mathematical problems. ('If the traffic jams on a motor way, the motorist blames bad luck, if we install traffic lights at the motorway entrances, the motorist blames the government'.)

Although there exist many books and articles proclaiming the imminent understanding of general complex systems, my suspicion is that we are still at the stage of examining particular properties and particular systems. Birkhoff's ergodic theorem is a striking example of a particular type of property and Carleson's work a deep example of the study a particular type of system. Even if the optimistic hope that we (or rather our great grand children) will be able to extract general properties which hold for general systems is not realised, the study of each particular system or property will increase our ability to deal with others. Carleson's convergence theorem lies at the end of one research program, his work on Hénon maps is one of the markers at the beginning of another.

Finally, I should answer the general member of the public. Unlike engineers or the politicians, though on even less evidence, the general public has an even higher opinion of the cleverness and usefulness of mathematicians than they have themselves. The general public is perfectly happy to support mathematics and has no doubt that, in some undefined way, the work of mathematicians leads to a general increase in prosperity. However, the man or woman in the street is puzzled as to what satisfaction someone like Carleson can derive from doing work which can be understood by so few people.

To answer this question he or she should consider the satisfaction that a cook has in presenting a good dinner or a string quartet playing privately for their own enjoyment. It is not necessary to hear applause to feel the satisfaction of a job well done. The amateur painter or, at a lower level, the completer of a large jigsaw feels satisfaction whether or not anyone else is there to praise them. For mathematicians the satisfaction of resolving a problem is a reward in itself. (Fortunately for us, the satisfaction is only slightly dependent on the importance of the puzzle. I have seen distinguished mathematicians head over heels with delight at results which they know will interest no one else.)

The virtuoso violinist is rewarded by his or her own satisfaction, the praise of his or her peers and the applause of the multitude. The mathematician will never have the applause of the multitude (though an Abel prize comes close to it), but Carleson can enjoy the unstinted admiration of the mathematical world in which his life has been lived.

The mathematician does have one advantage over many other artists. Authors know that their books may be acclaimed by a generation of readers and critics alike and yet be condemned by the taste of the next generation. There is no sure test for poetic greatness. But someone who solves a problem that has baffled the finest mathematicians for a century and a half, knows that he has done a great thing. Names like Dirichlet, Riemann and Kolmogorov may mean nothing to the general reader but to mathematicians they are heroes. It has been given to Carleson to be the hero of a story in which the other actors are his own heroes.

References

1. Bari, M.K.: A Treatise on Trigonometric Series. Pergamon, Elmsford (1964). (Two volumes, English translation)
2. Klein, F.: Development of Mathematics in the Nineteenth Century. Math. Sci. Press, Brookline (1979) (English translation by M. Ackerman)
3. Zygmund, A.: Trigonometric Series, vol. I, 2nd. ed. Cambridge Univ. Press, New York (1959)
4. Actes du Congrès International des Mathématiciens, 1970, Nice. Gauthier–Villars, Paris (1971)

List of Publications for Lennart Carleson

1951

[1] On Bernstein approximation problem. *Proc. Amer. Math. Soc.*, 2:953–961.
[2] On null-sets for continuous analytic functions. *Ark. Mat.*, 1:311–318.

1952

[3] On bounded analytic functions and closure problems. *Ark. Mat.*, 2:283–291.
[4] On the zeros of functions with bounded Dirichlet integrals. *Math. Z.*, 56:289–295.
[5] Sets of uniqueness for functions regular in the unit circle. *Acta Math.*, 87:325–345.

1953

[6] On infinite differential equations with constant coefficients. I. *Math. Scand.*, 1:31–38.

1954

[7] A proof of an inequality of Carleman. *Proc. Amer. Math. Soc.*, 5:932–933.

1956

[8] Random sequences and additive number theory. *Math. Scand.*, 4:303–308.

1957

[9] A mathematical model for highway traffic. *Nordisk Mat. Tidskr.*, 5:176–180, 213.
[10] Representations of continuous functions. *Math. Z.*, 66:447–451.

1958

[11] An interpolation problem for bounded analytic functions. *Amer. J. Math.*, 80:921–930.
[12] Two remarks on the basic theorems of information theory. *Math. Scand.*, 6:175–180.
[13] On the connection between Hausdorff measures and capacity. *Ark. Mat.*, 3:403–406.

1959

[14] (with E.J. Akutowicz). The analytic continuation of interpolatory functions. *J. Analyse Math.*, 7:223–247.

H. Holden, R. Piene (eds.), *The Abel Prize*,
DOI 10.1007/978-3-642-01373-7_17,

1960

[15] A representation formula for the Dirichlet integral. *Math. Z.*, 73:190–196.

1961

[16] On universal moment problems. *Math. Scand.*, 9:197–206.
[17] A remark on Picard's theorem. *Bull. Amer. Math. Soc.*, 67:142–144.

1962

[18] On the existence of boundary values for harmonic functions in several variables. *Ark. Mat.*, 4:393–399 (1962).
[19] Interpolations by bounded analytic functions and the corona problem. *Ann. of Math. (2)*, 76:547–559.

1963

[20] Interpolations by bounded analytic functions and the Corona problem. In *Proc. Internat. Congr. Mathematicians (Stockholm, 1962)*, pages 314–316. Inst. Mittag-Leffler, Djursholm.
[21] Removable singularities of continuous harmonic functions in R^m. *Math. Scand.*, 12:15–18.

1964

[22] Mergelyan's theorem on uniform polynomial approximation. *Math. Scand.*, 15:167–175.

1965

[23] Maximal functions and capacities. *Ann. Inst. Fourier (Grenoble)*, 15(fasc. 1):59–64.

1966

[24] On convergence and growth of partial sums of Fourier series. *Acta Math.*, 116:135–157.

1967

[25] *Selected problems on exceptional sets*. Van Nostrand Mathematical Studies, No. 13. D. Van Nostrand Co., Inc., Princeton, N.J.-Toronto, Ont.-London. Also available in Russian. Izdat. "Mir", Moscow, 1971.
[26] On mappings, conformal at the boundary. *J. Analyse Math.*, 19:1–13.

1968

[27] Convergence and summability of Fourier series. In *Proc. Internat. Congr. Math. (Moscow, 1966)*, pages 83–88. Izdat. "Mir", Moscow.

1970

[28] The Corona theorem. In *Proceedings of the Fifteenth Scandinavian Congress (Oslo, 1968), Lecture Notes in Mathematics, Vol. 118*, pages 121–132, Berlin. Springer.

1972

[29] (with P. Sjölin). Oscillatory integrals and a multiplier problem for the disc. *Studia Math.*, 44:287–299. (errata insert). Collection of articles honoring the completion by Antoni Zygmund of 50 years of scientific activity, III.
[30] (with S. Jacobs). Best uniform approximation by analytic functions. *Ark. Mat.*, 10:219–229.

1973

[31] On the distortion of sets on a Jordan curve under conformal mapping. *Duke Math. J.*, 40:547–559.

[32] Projections métriques de L^p sur H^p. *C. R. Acad. Sci. Paris Sér. A–B*, 276:A1159–A1160.

1974

[33] The extension problem for quasiconformal mappings. In *Contributions to analysis (a collection of papers dedicated to Lipman Bers)*, pages 39–47. Academic Press, New York.

1976

[34] Two remarks on H^1 and BMO. *Advances in Math.*, 22(3):269–277.

[35] Asymptotic paths for subharmonic functions in $\mathbb{R}^n$. *Ann. Acad. Sci. Fenn. Ser. A I Math.*, 2:35–39.

1978

[36] Some problems in harmonic analysis related to statistical mechanics. In *Fourier analysis and approximation theory (Proc. Colloq., Budapest, 1976), Vol. I*, volume 19 of *Colloq. Math. Soc. János Bolyai*, pages 213–217. North-Holland, Amsterdam.

1979

[37] An example concerning analytic functions with finite Dirichlet integrals. *Zap. Nauchn. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI)*, 92:283–287, 326.

[38] A remark on Denjoy's inequality and Herman's theorem. *Inst. Hautes Études Sci. Publ. Math.*, (49):235–241.

1980

[39] Some analytic problems related to statistical mechanics. In *Euclidean harmonic analysis (Proc. Sem., Univ. Maryland, College Park, Md., 1979)*, volume 779 of *Lecture Notes in Math.*, pages 5–45. Springer, Berlin.

[40] The work of Charles Fefferman. In *Proceedings of the International Congress of Mathematicians (Helsinki, 1978)*, pages 53–56. Also available in *Fiz.-Mat. Spis. B"lgar. Akad. Nauk.*, 23(56)(3):247–253 (in Bulgarian).

[41] An explicit unconditional basis in H^1. *Bull. Sci. Math. (2)*, 104(4):405–416.

1981

[42] BMO—10 years' development. In *18th Scandinavian Congress of Mathematicians (Aarhus, 1980)*, volume 11 of *Progr. Math.*, pages 3–21. Birkhäuser Boston, Mass.

1982

[43] Estimates of harmonic measures. *Ann. Acad. Sci. Fenn. Ser. A I Math.*, 7(1):25–32.

[44] (with L. Gårding). Obituary on Niels Erik Nørlund. *Acta Math.*, 148:1.

1983

[45] J.-P. Kahane and Y. Katznelson. Séries de Fourier des fonctions bornées. In *Studies in pure mathematics*, pages 395–413. Birkhäuser, Basel. With an appendix by L. Carleson.

[46] Selected problems on exceptional sets. In *Selected reprints*, Wadsworth Math. Ser., pages iv+100. Wadsworth, Belmont, CA.

[47] On H^∞ in multiply connected domains. In *Conference on harmonic analysis in honor of Antoni Zygmund, Vol. I, II (Chicago, Ill., 1981)*, Wadsworth Math. Ser., pages 349–372. Wadsworth, Belmont, CA.

1985

[48] W.K. Hayman. Local value distribution of functions bounded in a half-plane. *Ann. Acad. Sci. Fenn. Ser. A I Math.*, 10:227–237. With an appendix by L. Carleson.

[49] On the support of harmonic measure for sets of Cantor type. *Ann. Acad. Sci. Fenn. Ser. A I Math.*, 10:113–123.

[50] (with M. Benedicks). On iterations of $1 - ax^2$ on $(-1, 1)$. *Ann. of Math. (2)*, 122(1):1–25.

1986

[51] (with S.-Y.A. Chang). On the existence of an extremal function for an inequality of J. Moser. *Bull. Sci. Math. (2)*, 110(2):113–127.

1988

[52] (with L. Ahlfors). Arne Beurling in memoriam. *Acta Math.*, 161(1–2):1–9.

1989

[53] A. Beurling. *The collected works of Arne Beurling. Vol. 1*. Contemporary Mathematicians. Birkhäuser Boston Inc., Boston, MA. Edited by L. Carleson, P. Malliavin, J. Neuberger and J. Wermer.

[54] A. Beurling. *The collected works of Arne Beurling. Vol. 2*. Contemporary Mathematicians. Birkhäuser Boston Inc., Boston, MA. Edited by L. Carleson, P. Malliavin, J. Neuberger and J. Wermer.

[55] (with M. Benedicks). On the Hénon attractor. In *IXth International Congress on Mathematical Physics (Swansea, 1988)*, pages 498–500. Hilger, Bristol.

[56] (with C.J. Bishop, J.B. Garnett, and P.W. Jones). Harmonic measures supported on curves. *Pacific J. Math.*, 138(2):233–236.

1991

[57] The dynamics of non-uniformly hyperbolic systems in two variables. In *Proceedings of the International Congress of Mathematicians, Vol. I, II (Kyoto, 1990)*, pages 1241–1247, Math. Soc. Japan, Tokyo.

[58] Stochastic models of some dynamical systems. In *Geometric aspects of functional analysis (1989–90)*, volume 1469 of *Lecture Notes in Math.*, pages 1–12. Springer, Berlin.

[59] (with M. Benedicks). The dynamics of the Hénon map. *Ann. of Math. (2)*, 133(1):73–169.

1992

[60] (with P.W. Jones). On coefficient problems for univalent functions and conformal dimension. *Duke Math. J.*, 66(2):169–206.

1993

[61] (with T.W. Gamelin). *Complex dynamics*. Universitext: Tracts in Mathematics. Springer-Verlag, New York.

1994

[62] (with N.G. Makarov). Some results connected with Brennan's conjecture. *Ark. Mat.*, 32(1):33–62.

[63] (with P.W. Jones and J.-C. Yoccoz). Julia and John. *Bol. Soc. Brasil. Mat. (N.S.)*, 25(1):1–30.

1997

[64] Wiener's Tauberian theorem. In *The Legacy of Norbert Wiener: A Centennial Symposium (Cambridge, MA, 1994)*, volume 60 of *Proc. Sympos. Pure Math.*, pages 65–70. Amer. Math. Soc., Providence, RI.

2000

[65] Lars Ahlfors and the Painlevé problem. In *In the tradition of Ahlfors and Bers (Stony Brook, NY, 1998)*, volume 256 of *Contemp. Math.*, pages 5–10. Amer. Math. Soc., Providence, RI.

2001

[66] (with S.-Y.A. Chang, P.W. Jones, M. Keel, P.D. Lax, N. Makarov, D. Sarason, W. Schlag, and B. Simon). Thomas H. Wolff (1954–2000). *Notices Amer. Math. Soc.*, 48(5):482–490.

[67] (with N. Makarov). Aggregation in the plane and Loewner's equation. *Comm. Math. Phys.*, 216(3):583–607.

2002

[68] (with N. Makarov). Laplacian path models. *J. Anal. Math.*, 87:103–150.

2004

[69] (with V. Totik). Hölder continuity of Green's functions. *Acta Sci. Math. (Szeged)*, 70(3-4):557–608.

2006

[70] Computations in pure mathematics (outline of lecture at the B. Engquist conference). *BIT*, 46(suppl.):S19–S20.

Curriculum Vitae for Lennart Axel Edvard Carleson

Born:	March 18, 1928, Stockholm, Sweden
Degrees/education:	Fil. dr. and docent Uppsala University, 1950
Positions:	Docent, Uppsala University, 1951–54 Professor, Stockholm Högskola, 1954–55 Professor, Uppsala University, 1955–93 Professor, University of California at Los Angeles, 1991–
Visiting positions:	Guest professor, MIT, 1957 Visiting member, Institute for Advanced Study, Princeton, 1961–62 Guest professor, Stanford, 1965–66 Guest professor, MIT, 1974–75 Distinguished visiting professor, Institute for Advanced Study, Princeton, 1988–89
Memberships:	Royal Swedish Academy of Sciences, 1957 American Academy of Arts and Sciences, USA, 1967 Finnish Academy of Science and Letters, 1970 Royal Danish Society of Sciences and Letters, 1970 Royal Norwegian Society of Sciences and Letters, 1979 Russian Academy of Sciences, 1982 Norwegian Academy of Science and Letters, 1983 Hungarian Academy of Sciences, Budapest, 1986 French Academy, Paris, 1992 Royal Society, 1993 National Academy of Sciences, Washington, 2006 European Academy of Sciences

Awards and prizes:	Steele Prize, American Mathematical Society, 1984 Wolf Prize, 1992 Lomonosov Gold Medal, Russian Academy of Science, 2002 Sylvester Medal, Royal Society, 2003 Abel Prize, 2006
Honorary degrees:	University of Helsinki, 1982 University of Paris, 1988 Royal Institute of Technology, 1989
Presidencies:	Director of the Mittag-Leffler Institute, 1968–84 Editor, Acta Mathematica, 1956–79 President International Mathematical Union, 1978–82