A Taste of Erdös on Interpolation

D.S. Lubinsky

15 October 1999

Abstract

We discuss a few of Erdös' results on Lagrange interpolation, and then focus on some of the ramifications of the Erdös-Turan Theorem on Mean Convergence of Lagrange interpolation.

1 Introduction

Interpolation by polynomials is a subject as old as mathematics itself: between any two points (x_1, y_1) and (x_2, y_2) in the plane with $x_1 \neq x_2$, we can fit a unique (and non-vertical) straight line. Of course this line may be described by a linear polynomial

$$y = ax + b.$$

What happens when we have more than two points? Newton attended to this question in the 1670's. Let $(x_j, y_j), 1 \leq j \leq n$ be n points, with distinct and *equally spaced* $\{x_j\}$. In the course of attempting to predict the location of comets at arbitrary times, from their location at equally spaced times, Newton found a formula for the polynomial P of degree$\leq n-1$ satisfying

$$P(x_j) = y_j, 1 \leq j \leq n. \tag{1}$$

His student Cotes, calculated the coefficients in this formula, and used it in the Newton-Cotes quadrature rule, for approximating integrals.

Remarkably enough, it was only in 1795 that a formula was given for the case of non-equally spaced $\{x_j\}$. Lagrange defined the *fundamental polynomials*

$$\ell_j(x) := \prod_{k=1, k\neq j}^{n} \left(\frac{x - x_k}{x_j - x_k} \right), 1 \leq j \leq n,$$

which satisfy

$$\ell_j(x_k) = \delta_{jk} \forall j, k.$$

He used these as the basis of the *Lagrange interpolation polynomial,*

$$P(x) := \sum_{j=1}^{n} y_j \ell_j(x),$$

which clearly satisfies (1). Amongst his many deep contributions to mathematics, this formula must have been the simplest. But he certainly valued it, and its impact has been enormous.

An immediate question is the validity of sampling a given function $f : \mathbb{R} \to \mathbb{R}$ at a growing number of points and forming the corresponding Lagrange interpolation polynomials. Will they converge to the underlying function? To elucidate this further, we need some notation.

Suppose that for $n \geq 1$, we are given n distinct real numbers

$$x_{nn} < x_{n-1,n} < x_{n-2,n} < ... < x_{1n}.$$

(Note the unusual practice that n is the second index - this seems to be part of interpolatory culture). We call the triangular array

$$X := \left\{ \begin{array}{llll} x_{11} & & & \\ x_{22} & x_{12} & & \\ x_{33} & x_{23} & x_{13} & \\ \vdots & \vdots & \vdots & \ddots \end{array} \right\} \tag{2}$$

an *array of interpolation points*. For a real valued function f defined at each point of the array X, we may define the nth Lagrange interpolation polynomial

$$L_n [f] (x) := \sum_{j=1}^{n} f (x_{jn}) \ell_{jn} (x), \tag{3}$$

where the fundamental polynomials $\{\ell_{jn}\}_{j=1}^{n}$ are given by

$$\ell_{jn} (x) := \prod_{k=1, k \neq j}^{n} \left(\frac{x - x_{kn}}{x_{jn} - x_{kn}} \right). \tag{4}$$

One unfortunate feature of this notation is possible confusion between L_p norms and Lagrange interpolation polynomials. However, the latter will alway be distinguished by the dependence on the function being interpolated. (Mathematics is a context sensitive language!) Thus, for example,

$$\| L_n [f] \|_{L_p[-1,1]} = \left(\int_{-1}^{1} |L_n [f]|^p \right)^{1/p}.$$

Throughout, we use $C, C_1, C_2, ...$ to denote positive constants independent of n, polynomials P of degree $\leq n$, and of x. If we wish to emphasise independence, we shall write $C \neq C(n, P)$.

While interpolation polynomials (and more generally, even interpolating rational functions) were widely used in the nineteenth century, there was not much rigorous analysis of their convergence. No less a father of rigour than Cauchy investigated interpolation by rational functions in the 1820's, but failed to notice problems with their existence, let alone their convergence.

Perhaps the first significant negative result is due to Meray in 1884, in the complex plane. He observed that the polynomial of degree $\leq n-1$ that interpolates to $f(z) := z^{-1}$ at the nth roots of unity is $P_n (z) := z^{n-1}$, and clearly this does not converge to f as $n \to \infty$, on the unit circle or off. An obvious criticism of this example is that f has a pole at 0, and so no sequence of polynomials - interpolatory or not - can converge to f uniformly on the unit circle. (If they did they would converge to an analytic function inside the unit circle also, thereby analytically continuing f to the unit ball).

A more genuine example was given by Carl Runge (of Runge's theorem fame) in 1901. He showed that if we let

$$f(x) := \frac{1}{c^2 + x^2}, x \in [-1, 1],$$

with c small enough (for example $c \leq \frac{1}{5}$), and if $L_n[f]$ denotes the interpolation polynomial to f at n equally spaced points in $[-1, 1]$, then

$$\| f - L_n[f] \|_{L_\infty[-1,1]} \to \infty, n \to \infty,$$

even with geometric rate [67], [68]. This was especially worrisome, as it had always been assumed that interpolation at equally spaced points is a good idea, and after all this f is analytic on the real axis, though it does have poles at $\pm ic$, which can be quite close to $[-1, 1]$. As it subsequently turned out, the Newton-Cotes rules for numerical integration based on equally spaced points also diverges on "nice" functions, including the Runge example [66]. Thus, sampling a function at equally spaced points and then interpolating by Lagrange interpolation polynomials is a bad idea.

If this shook faith in interpolation by polynomials as a means of approximation, a bigger shock was to come in 1914. Faber [42] showed that given any array X of interpolation points, they diverge on some continuous function. (The more famous S.N. Bernstein [5] also obtained part of this result). More precisely, let $\| L_n \|$ denote the norm of L_n as a linear operator on $C[-1, 1]$, the space of continuous functions $f[-1, 1] \to \mathbb{R}$, with uniform norm. Thus,

$$\| L_n \| = \sup \left\{ \frac{\| L_n[f] \|_{L_\infty[-1,1]}}{\| f \|_{L_\infty[-1,1]}} : f \in C[-1, 1] \right\}.$$

Then, both Bernstein and Faber showed that

$$\| L_n \| \geq \frac{1}{12} \log n \tag{5}$$

and in particular

$$\sup_n \| L_n \| = \infty.$$

The uniform boundedness principle then provides a continuous function f for which

$$\sup_n \| f - L_n[f] \|_{L_\infty[-1,1]} = \infty.$$

While that principle may have not yet been available to Faber, he was able to construct explicitly an f with this last property.

The fainthearted would have concluded from Runge and Faber-Bernstein's work that interpolation by polynomials is an inherently flawed process. But as an encouraging positive result, S.N. Bernstein showed in 1912 [4] that if for some $\rho > 1$, $f : [-1, 1] \to \mathbb{R}$ admits an analytic continuation to the interior of the ellipse

$$\mathcal{E}_\rho := \left\{ \frac{1}{2} \left(z + z^{-1} \right) : |z| = \rho \right\},$$

which has foci at ± 1, and if we interpolate to f at the zeros

$$x_{jn} := \cos\left(\left(j - \frac{1}{2}\right)\frac{\pi}{n}\right), 1 \le j \le n, \tag{6}$$

of the Chebyshev polynomial

$$T_n(x) := \cos(n \arccos x),$$

then $L_n[f]$ converges to f with a geometric rate:

$$\limsup_{n\to\infty} \parallel f - L_n[f] \parallel^{1/n}_{L_\infty[-1,1]} \le 1/\rho < 1.$$

The ramifications of this for interpolation of analytic functions were later explored by many mathematicians, notably Kalmar, Walsh, ... [87].

What was clear in Bernstein's proof, was the importance of the location of the interpolation points, and the analyticity of f, which enabled use of Cauchy's integral formula. There seemed to be no hope of ever proving convergence - in any sense - of Lagrange interpolation for functions that are not analytic.

How remarkable then was the 1937 result of Erdös and Turan, in the first of three seminal papers on Lagrange interpolation published in Annals of Mathematics [30]. Let $w : [-1,1] \to [0,\infty)$ be measurable, with

$$0 < \int_{-1}^{1} w < \infty.$$

Corresponding to w, there are orthonormal polynomials $\{p_n\}_{n=0}^{\infty}$, where

$$p_n(x) = \gamma_n x^n + ..., \gamma_n > 0,$$

has degree n, and

$$\int_{-1}^{1} p_n p_m w = \delta_{mn}.$$

It is a straightforward consequence of orthogonality that p_n has n distinct zeros that lie in $[-1,1]$.

Erdös-Turan Theorem on Mean Convergence of Lagrange Interpolation (1937)

Let $f : [-1,1] \to \mathbb{R}$ be continuous. For $n \ge 1$, let $L_n[f]$ denote the Lagrange interpolation polynomial to f at the zeros of p_n. Then

$$\lim_{n\to\infty} \int_{-1}^{1} (f - L_n[f])^2 w = 0. \tag{7}$$

Thus, at least in a Hilbert space setting, interpolation at zeros of orthogonal polynomials (for any weight on $[-1,1]$) is a perfectly respectable thing to do! The ramifications of this result continue to be explored to this day. In this author's opinion, it is the most important positive result of Erdös on interpolation, and the most important positive result ever on convergence of polynomial interpolation of continuous functions.

We shall shortly give Shohat's 1939 proof, but this requires a little background, which also helps to explain why interpolation at zeros of orthogonal polynomials may be more easily analysed than the general case. A key ingredient of Shohat's proof is the Gauss quadrature formula for the weight w. Given $n \geq 1$, let $\{x_{jn}\}_{j=1}^{n}$ denote the zeros of p_n. Then there are positive numbers $\{\lambda_{jn}\}_{j=1}^{n}$, called Christoffel or Gauss-Christoffel numbers such that

$$G_n[P] := \sum_{j=1}^{n} \lambda_{jn} P(x_{jn}) = \int P w, \tag{8}$$

whenever P is a polynomial of degree at most $2n-1$.

That quadrature formulae should play a role in analysing Lagrange interpolation is scarcely surprising, as the two topics are first cousins. Indeed very many quadrature formulae are derived by integrating Lagrange interpolation polynomials (so called interpolatory quadrature formulae), and convergence of the quadrature rules is equivalent to

$$\lim_{n\to\infty} \int_{-1}^{1} (f - L_n[f])\, w = 0.$$

It was T.J. Stieltjes who first investigated this for general weights w. It must be emphasised, however, that inserting a square (or absolute value) converts the problem from one about the theory of integration, into a far more difficult analysis problem.

There are several expressions for the Christoffel numbers. Some involve the reproducing kernel

$$K_n(x,t) := \sum_{j=0}^{n-1} p_j(x)\, p_j(t).$$

There is a compact formula for $K_n(x,t)$, the Christoffel-Darboux fomula:

$$K_n(x,t) = \frac{\gamma_{n-1}}{\gamma_n} \frac{p_n(x)\, p_{n-1}(t) - p_{n-1}(x)\, p_n(t)}{x-t},$$

which when $x = t$, takes the confluent form (use l'Hôpital)

$$K_n(x,x) = \frac{\gamma_{n-1}}{\gamma_n} \left(p_n'(x)\, p_{n-1}(x) - p_{n-1}'(x)\, p_n(x)\right).$$

It is known that

$$\lambda_{jn} = 1/K_n(x_{jn}, x_{jn}) = 1/\left(\frac{\gamma_{n-1}}{\gamma_n} p_n'(x_{jn})\, p_{n-1}(x_{jn})\right).$$

Not only do the Christoffel numbers admit a representation in terms of the reproducing kernel, but the fundamental polynomials $\{\ell_{jn}\}_{j=1}^{n}$ also do:

$$\ell_{jn}(x) = \lambda_{jn} K_n(x, x_{jn}) = K_n(x, x_{jn}) / K_n(x_{jn}, x_{jn}).$$

Indeed the polynomial on the right-hand side takes the value 1 at $x = x_{jn}$, and by the Christoffel-Darboux formula, it vanishes at $x_{kn}, k \neq j$. See [46], [74] ... for more of this.

These identities are very suggestive of the relationship between Lagrange interpolation and orthonormal expansions. We see that we may write

$$\begin{aligned} L_n[f](x) &= \sum_{j=1}^{n} \lambda_{jn} f(x_{jn}) K_n(x, x_{jn}) \\ &= G_n[f(\cdot) K_n(x, \cdot)] \\ &\approx \int f(t) K_n(x,t) w(t)\, dt \\ &= S_n[f](x), \end{aligned}$$

where $S_n[f]$ is the nth partial sum of the orthonormal expansion of f in $\{p_n\}_{n=0}^{\infty}$. (We use $\approx$ to mean approximately equal to). Recall that that expansion has the form

$$f \sim \sum_{j=0}^{\infty} \left(\int_{-1}^{1} f p_j w \right) p_j$$

and so the nth partial sum is

$$S_n[f](x) = \sum_{j=0}^{n-1} \left(\int_{-1}^{1} f p_j w \right) p_j(x) = \int_{-1}^{1} f(t) K_n(x,t) w(t)\, dt. \tag{9}$$

At this stage, the reader's patience with the meandering aside from the Erdös-Turan theorem should be rewarded. Almost every course in functional analysis contains a treatment of orthonormal expansions in Hilbert space and the least squares property of their partial sums. That in turn leads to the limit relation

$$\lim_{n\to\infty} \int_{-1}^{1} (f - S_n[f])^2\, w = 0, \tag{10}$$

under mild conditions on f, w. For example, if w is supported on $[-1,1]$, the polynomials are dense in the relevant Hilbert space, and (10) is valid for every Lebesgue measurable f for which

$$\int_{-1}^{1} f^2 w < \infty.$$

So in retrospect, the Erdös-Turan theorem may be viewed as a discrete analogue of the mean square convergence of orthonormal expansions, inasmuch as $L_n[f]$ is a discretisation of $S_n[f]$. It is not clear from their papers whether Erdös and Turan ever took any motivation from (10).

We can now present

Shohat's Proof of the Erdös-Turan theorem (1939)

Let f be continuous on $[-1,1]$ and $\varepsilon > 0$. By Weierstrass' Theorem, we can choose a polynomial P such that

$$\| f - P \|_{L_\infty[-1,1]} < \varepsilon.$$

Since L_n is a linear projection onto the polynomials of degree $< n$, we have for

$n > \deg(P)$,

$$\begin{aligned}
&\int_{-1}^{1} (f - L_n[f])^2 w \\
= &\int_{-1}^{1} (f - P + L_n[P - f])^2 w \\
\leq \; &2\int_{-1}^{1} (f-P)^2 w + 2\int_{-1}^{1} L_n^2[P-f]\, w. \qquad (11)
\end{aligned}$$

Here we have used the inequality

$$(x+y)^2 \leq 2\left(x^2+y^2\right),$$

instead of opening up the square, which might have been more expected. Since $L_n^2[P-f]$ is a polynomial of degree $\leq 2n-2$, we may use the Gauss quadrature formula and then the interpolation property to deduce that

$$\begin{aligned}
\int_{-1}^{1} L_n^2[P-f]\, w &= \sum_{j=1}^{n} \lambda_{jn} (P-f)^2 (x_{jn}) \\
&\leq \varepsilon^2 \sum_{j=1}^{n} \lambda_{jn} = \varepsilon^2 \int_{-1}^{1} w.
\end{aligned}$$

Since the first term in the right-hand side of (11) is bounded by $2\varepsilon^2 \int_{-1}^{1} w$, the proof is complete. □

If we define the errors of polynomial approximation

$$E_n[f] := \min_{\deg(P)\leq n} \| f - P \|_{L_\infty[-1,1]},$$

then it is easy to see that the above proof actually yields the rate

$$\int_{-1}^{1} (f - L_n[f])^2 w \leq 4\left(\int_{-1}^{1} w\right)(E_{n-1}[f])^2. \qquad (12)$$

Of course once a great idea like this comes along, it is pushed and pulled in every conceivable direction. Continuity of f has been weakened to Riemann integrability of f and that has been weakened to improper Riemann integrability, thereby allowing f to have a number of infinities in the interval. Why not Lebesgue integrability? Oh yes, we are only sampling f at a countable set of values $\{x_{jn}\}_{j,n}$.

And, of course, you would have guessed that no effort has been spared to replace the L_2 norm by an L_p norm. We shall discuss some of that in greater detail in the next section. The extension from weights with compact support to those with non-compact support, came early, and is due to Shohat. The latter was very interested in the moment problem, which involves orthogonal polynomials for weights supported on $(0,\infty)$ or $\mathbb{R}$. We shall provide a few references to work on unbounded intervals in the next section.

One thing that was clear at the conference associated with these proceedings, is that a lot of the very greatest work of Erdös (and especially in collaboration with Turan) was done in the late 1930's and early 1940's, a time of great difficulty for both of them. In addition to the above theorem, their three Annals of Mathematics papers on interpolation contained the greatest advance in the theory of orthogonal polynomials since Szegö's work of the 1920's. (This again reminds one that the importance of a subject is often measured by its external impact. Certainly interpolation á lá Erdös has had an impact on many other areas of mathematics).

They proved in their third Annals paper [32] that if w is a weight positive a.e. in $[-1,1]$, then the associated orthogonal polynomials $\{p_n\}_{n=0}^{\infty}$ satisfy an nth root asymptotic,

$$\lim_{n\to\infty} |p_n(z)|^{1/n} = \left|z+\sqrt{z^2-1}\right|, z \in \mathbb{C}\backslash[-1,1], \tag{13}$$

the convergence being uniform in compact subsets. Here the branch of the square root is chosen so that $\sqrt{z^2-1}$ is single valued and analytic in $\mathbb{C}\backslash[-1,1]$, and is positive for $z \in (1,\infty)$. The function $z+\sqrt{z^2-1}$ maps $\mathbb{C}\backslash[-1,1]$ conformally onto the exterior of the unit ball. Many years later, E.A. Rakhmanov [76], [77] showed that the positivity a.e. of w actually implies the stronger ratio asymptotic

$$\lim_{n\to\infty} \frac{p_{n+1}(z)}{p_n(z)} = z+\sqrt{z^2-1}, z \in \mathbb{C}\backslash[-1,1].$$

The quest for a characterization of weights w that allow an nth root asymptotic has lasted at least fifty years. Among those involved have been Erdös, J.L. Ullman, G. Freud, H. Widom, H. Stahl, V. Totik, A. Ambroladze, ... The recent monograph of Stahl and Totik [81] may be viewed as a culmination of this research, though of course not all the problems are solved: the boundaries of research problems have a fractal character.

Nth root asymptotics have intrinsic interest, but also have applications in studying geometric convergence of Lagrange interpolation processes, in potential theory, distribution of zeros, ... In the third of their three great papers, they deduced from (13) that the zeros of $\{p_n\}_{n=0}^{\infty}$ have arcsin distribution: given $-1 \le a < b \le 1$, let $N_n[a,b]$ denote the total number of zeros of p_n in $[a,b]$. Then (13) is equivalent to

$$\lim_{n\to\infty} \frac{N_n[a,b]}{n} = \int_a^b \frac{dx}{\pi\sqrt{1-x^2}} \forall -1 \le a < b \le 1.$$

This may be more elegantly reformulated for

$$\theta_{jn} := \arccos x_{jn} \in (0,\pi).$$

If $\mathcal{N}_n[\alpha,\beta]$ denotes the number of j with $1 \le j \le n$ and $\theta_{jn} \in [\alpha,\beta]$ then

$$\lim_{n\to\infty} \frac{\mathcal{N}_n[\alpha,\beta]}{n} = \frac{\beta-\alpha}{\pi} \forall 0 \le \alpha < \beta \le \pi.$$

In [32], the great duo acknowledged their indebtedness to the work of Fejér, and also to the work of Kalmar, another great Hungarian interpolator. His work dealt primarily with functions with analyticity in some region.

Having shown that Hilbert space is a good setting for convergence of Lagrange interpolation, it seemed natural to determine what can be salvaged in the sup norm. The obvious answer (in retrospect) is to weaken the interpolatory requirement. In 1916, L. Fejér [43] had given the first result of this type, thereby providing the first interpolatory proof of Weierstrass' Theorem. Let $\{x_{jn}\}_{j=1}^{n}$ denote the zeros of $T_n(x)$, and define the Hermite-Fejér polynomial $H_n[f](x)$ by the $2n$ conditions

$$\begin{aligned} H_n[f](x_{jn}) &= f(x_{jn}), 1 \leq j \leq n; \\ H_n[f]'(x_{jn}) &= 0, 1 \leq j \leq n. \end{aligned}$$

This is a polynomial of degree $\leq 2n-1$. Fejér showed that for every $f \in C[-1,1]$,

$$\lim_{n\to\infty} \| f - H_n[f] \|_{L_\infty[-1,1]} = 0,$$

thereby initiating the topic of Hermite-Fejér interpolation. A little later (B. Shekhtman informed the author of this) the great S.N. Bernstein showed that we can replace $2n-1$ by $(1+\varepsilon)n$ for fixed $\varepsilon > 0$. More precisely, there exists an interpolation array X with the following property: for $f \in C[-1,1]$, and $n \geq 1$, there exist polynomials P_n of degree $\leq n(1+\varepsilon)$ such that

$$P_n(x_{jn}) = f(x_{jn}), 1 \leq j \leq n$$

and

$$\lim_{n\to\infty} \| f - P_n \|_{L_\infty[-1,1]} = 0. \tag{14}$$

Many years later, Erdös, Kroó and Szabados [26] determined what conditions on the array X allow a result of this type. They also strengthened the convergence (14) in terms of errors of best approximation, defined above. Let $[x]$ denote the greatest integer $\leq x$. Moreover, let $\mathcal{N}_n(I)$ denote the number of j with $\theta_{jn} \in I$, as above, and let $|I|$ denote the length of I.

Theorem of Erdös, Kroó and Szabados (1989)
Let X be an array of interpolation points. Let $\varepsilon > 0$. The following are equivalent:
(I) There exists $C > 0$ such that $\forall f \in C[-1,1]$, there exist polynomials p_n of degree at most $n(1+\varepsilon)$ such that

$$P_n(x_{jn}) = f(x_{jn}), 1 \leq j \leq n$$

and

$$\| f - P_n \|_{L_\infty[-1,1]} \leq C E_{[n(1+\varepsilon)]}[f], n \geq 1.$$

(II) For every sequence of intervals $\{I_n\}_{n=1}^{\infty}$ with

$$\lim_{n\to\infty} n|I_n| = \infty,$$

we have

$$\limsup_{n\to\infty} \frac{\mathcal{N}_n(I_n)}{n|I_n|} \leq \frac{1}{\pi}.$$

Moreover,

$$\liminf_{n\to\infty} \left[n \min_{1\leq k\leq n-1} (\theta_{k+1,n} - \theta_{k,n}) \right] > 0.$$

In essence the requirements of (II) is that the $\{\theta_{jn}\}$ are very uniformly distributed in $[0,\pi]$. This was one of the last papers of Erdös on interpolation, and despite its importance, somewhat off his main focus. A far more enduring interest of his was the size of the Lebesgue function. This is the pointwise norm of $L_n[f](x)$ as a linear functional from $C[-1,1]$ to $\mathbb{R}$. Thus

$$\lambda_n(x) := \sup\left\{\frac{|L_n[f](x)|}{\| f \|_{L_\infty[-1,1]}} : f \in C[-1,1]\right\}.$$

It is easy to see that $\lambda_n(x)$ admits the following representation in terms of the fundamental polynomials:

$$\lambda_n(x) = \sum_{j=1}^{n} |\ell_{jn}(x)|.$$

The Lebesgue constant is

$$\Lambda_n := \Lambda_n(X) := \| \lambda_n \|_{L_\infty[-1,1]} = \| L_n \|,$$

the norm of L_n as an operator on $C[-1,1]$. We already noted that Bernstein and Faber independently showed that for any array Λ_n must grow at least fast as $\frac{1}{12}\log n$.

Amongst the more obvious questions is which array(s) X give the minimal Lebesgue constant

$$\Lambda_n^* := \min_X \Lambda_n(X).$$

To this day there is no explicit representation for the nth row of the optimal array, and in all likelihood there never will be. The first insight into the size of Λ_n^* was given by L. Fejér [43], showing that if we take X to be the array whose nth row consists of the zeros of the nth Chebyshev polynomial, then

$$\Lambda_n(X) = \frac{2}{\pi}\log n + O(1). \tag{15}$$

Many of the estimations of Λ_n^* have in some way involved comparison with the Chebyshev polynomials. Indeed their small Lebesgue constant, and simple representation suggests that "when in doubt, interpolate at the zeros of T_n".

Since the results of Bernstein and Faber, many contributed to the estimation of Λ_n^* and related quantities, including Bernstein, Erdös, Vértesi, Szabados, Brutman, Kilgore, Pinkus, de Boor, Güntter, ... Amongst the most significant advances was a 1961 result of Erdös, which showed that for any array X,

$$\Lambda_n(X) \geq \frac{2}{\pi}\log n - O(1).$$

Together with (15), this shows that

$$\left|\Lambda_n^* - \frac{2}{\pi}\log n\right| \leq C.$$

For surveys of subsequent improvements, see [9], [18], [83]. Important conjectures of Bernstein (1931 vintage) and Erdös (1950 vintage) regarding the characteristion of the optimal arrays have been resolved by Kilgore [47], de Boor and Pinkus [6].

Once we know the size of Λ_n in the optimal case, it seems natural to ask on how large a set of x, can $\lambda_n(x)$ grow like $\log n$? Hopefully such a set would be small, so that for most x, the Lebesgue function is not too large, and in that case, for most x, Lagrange interpolation is not that much worse than best polynomial interpolation. Unfortunately, this is not the case. Erdös and Vertesi [40] showed in 1981 that most of the time, $\lambda_n(x)$ grows like $\log n$:

Theorem of Erdös and Vertesi on the Lebesgue function (1981)
Let X be an array of interpolation points in $[-1,1]$ and $\varepsilon > 0$. There exists $\eta > 0$ and for $n \geq 1$, a set $\mathcal{S}_n \subset [-1,1]$ such that

$$\lambda_n(x) > \eta \log n, x \in [-1,1] \backslash \mathcal{S}_n$$

and

$$|\mathcal{S}_n| < \varepsilon.$$

This suggests that for any array X, one should be able to find $f \in C[-1,1]$ for which $L_n[f]$ does not converge for most x. This would dash hopes of any analogue of the famous Carleson result that the Fourier series of an L_2 function converges a.e. Indeed, in very sharp contrast, Erdös and Vertesi [38], [39] showed in 1981 that there is always $f \in C[-1,1]$ for which $\{L_n[f]\}$ diverges a.e.:

Theorem of Erdös and Vertesi on a.e. divergence (1981)
Let X be an array of interpolation points in $[-1,1]$. There exists $f \in C[-1,1]$ such that

$$\limsup_{n\to\infty} |f(x) - L_n[f](x)| = \infty, \text{ a.e. } x \in [-1,1].$$

It is with these two negative, but very deep and impressive, results that we end our sample of Erdös' work on interpolation. For a more detailed survey of his work on interpolation, see [19]. For surveys and further results on Lebesgue functions, see [7-9], [12], [13], [50], [68], [83], [85-86]. For other surveys on Lagrange interpolation, see the monograph of Szabados and Vertesi [83], but also [2], [59], [84]. Nevai's surveys of 1976 and 1986 [69], [74] are still relevant and up to date for some aspects of pointwise and mean convergence of Lagrange interpolation.

2 Convergence in L_p Norms

When one looks at the complexity of some of the proofs that Lagrange interpolation converges in norms other than the L_2 norm, one is tempted to paraphrase Kronecker and say "God created L_2 and man created all else". But, just as the investigation of the boundedness of operators in L_p spaces (such as the Hilbert transform) has greatly enriched mathematics, so have the techniques developed for Lagrange interpolation.

In 1936-7 Erdös and Feldheim and independently Marcinkiewicz [58] (there does not seem to be a joint paper of Erdös and Feldheim on this, so I am not sure where it appeared) proved that interpolation at zeros of Chebyshev polynomials is a good idea in weighted L_p for any $1 < p < \infty$:

Theorem of Erdös and Feldheim; and Marcinkiewicz (1936-7)
Let $0 < p < \infty, f \in C[-1,1]$ and for $n \geq 1$, let $L_n[f]$ denote the interpolation polynomial to f at the zeros (6) of T_n. Then

$$\lim_{n\to\infty} \int_{-1}^{1} |f(x) - L_n[f](x)|^p \left(1-x^2\right)^{-1/2} dx = 0. \tag{16}$$

This may be viewed as a discrete analogue of the theorem of Riesz that the Fourier series of a function in $L_p[0,2\pi]$ converges in the norm of that space. This analogy is hardly incidental: mean convergence of orthogonal expansions has been the main tool in analysing mean convergence of Lagrange interpolation for many years.

The obvious next step was to investigate what happens when the Chebyshev weight $\left(1-x^2\right)^{-1/2}$ is replaced by more general weights. An immediate question is to what extent the convergence in L_p persists for all, or some, p. In this direction, P. Turan posed the question of finding a weight w on $(-1,1)$, for which the analogue of (16) fails for every $p > 2$. More precisely, let us form the interpolation polynomials $\{L_n[f]\}_{n=1}^{\infty}$ at the zeros of the orthogonal polynomials for w (in this section, L_n is always associated with w in this way, unless otherwise specified).

Turan's Problem I
Find a weight w such that for every $p > 2$, there exists $f \in C[-1,1]$ (depending on p), such that

$$\limsup_{n\to\infty} \int_{-1}^{1} |f - L_n[f]|^p \, w > 0.$$

Turan's Problem II
Find a function f such that this last divergence takes place for every $p > 2$.

What about $p < 2$? Well via an application of Hölder's inequality, the convergence for $p < 2$ follows from the Erdös-Turan Theorem.

Of course, (II) is a stronger form of (I). These problems of Turan have provided a framework for investigating mean convergence for the latter part of this century. On the one hand, for specific weights, the range of p admitting convergence has been established, and for general weights, necessary conditions have been established.

The obvious first step in generalising the Erdös-Feldheim results is to replace the Chebyshev weight $\left(1-x^2\right)^{-1/2}$ by the more general Jacobi weight $(1-x)^{\alpha}(1+x)^{\beta}, \alpha, \beta > -1$. This task was taken up by R. Askey [1] in the 1970's; in a slightly different setting, earlier work was due to Turan, Hollo, ...

The ideas of Askey's proof involve duality and boundedness of orthonormal expansions. Let $1 < p < \infty$ and let $q := \frac{p}{p-1}$ and for the next while, let us use the notation

$$\| g \|_{p,w} := \left(\int_{-1}^{1} |g|^p \, w \right)^{1/p}.$$

Moreover, let $\mathcal{L}_p(w)$ be the space of all g for which this norm is defined and finite. By duality,

$$\| L_n[f] \|_{p,w} = \sup_{\|g\|_{q,w}=1} \int_{-1}^{1} L_n[f] g w. \tag{17}$$

Here if $S_n[g]$ denotes the nth partial sum of the orthonormal expansion of g as at (9), we note the orthogonality property

$$\int_{-1}^{1} (g - S_n[g]) P w = 0,$$

for all polynomials P of degree$\leq n-1$. Thus

$$\begin{aligned} \int_{-1}^{1} L_n[f] g w &= \int_{-1}^{1} L_n[f] S_n[g] w \\ &= \sum_{j=1}^{n} \lambda_{jn} f(x_{jn}) S_n[g](x_{jn}) \\ &\leq \left(\sum_{j=1}^{n} \lambda_{jn} |f(x_{jn})|^p \right)^{1/p} \left(\sum_{j=1}^{n} \lambda_{jn} |S_n[g](x_{jn})|^q \right)^{1/q} . \end{aligned} \tag{18}$$

Here we have used the Gauss quadrature formula, and Hölder's inequality. Suppose now that for the given q, we have an inequality of the form

$$\int_{-1}^{1} \lambda_{jn} |P(x_{jn})|^q \leq C^q \int_{-1}^{1} |P|^q w, \tag{19}$$

valid for $n \geq 1$ and polynomials P of degree $\leq n-1$, where $C \neq C(n, P)$. Such an inequality is often called a *Marcinkiewicz-Zygmund inequality.* Then we may continue (17), (18) as

$$\| L_n[f] \|_{p,w} \leq C \left(\sum_{j=1}^{n} \lambda_{jn} |f(x_{jn})|^p \right)^{1/p} \sup_{\|g\|_{q,w}=1} \| S_n[g] \|_{q,w} .$$

Assuming now that $\{S_n\}_{n=1}^{\infty}$ is a uniformly bounded sequence of operators on the space $\mathcal{L}_q(w)$, that is, for some $C_1 \neq C_1(n, g)$,

$$\| S_n[g] \|_{q,w} \leq C_1 \| g \|_{q,w}, \tag{20}$$

we obtain a *converse Marcinkiewicz-Zygmund inequality*

$$\| L_n[f] \|_{p,w} \leq C C_1 \left(\sum_{j=1}^{n} \lambda_{jn} |f(x_{jn})|^p \right)^{1/p} . \tag{21}$$

Note that if we choose f to be a polynomial P of degree $\leq n-1$, this gives

$$\| P \|_{p,w} \leq C_2 \left(\sum_{j=1}^{n} \lambda_{jn} |P(x_{jn})|^p \right)^{1/p} , \tag{22}$$

with $C_2 \neq C_2(n, P)$. Once we have (21), we are basically done, since the convergence of the Gauss quadrature rule on Riemann integrable f shows that

$$\lim_{n \to \infty} \left(\sum_{j=1}^{n} \lambda_{jn} \left| f(x_{jn}) \right|^p \right)^{1/p} = \| f \|_{p,w} .$$

At this stage, we obtain for every $f \in C[-1,1]$,

$$\| L_n[f] \|_{p,w} \leq C_3 \| f \|_{p,w} \leq C_4 \| f \|_{L_\infty[-1,1]} .$$

Here C_4 is independent of n. Initially it may depend on f, but uniform boundedness shows that we may take C_4 independent of f also. The projection property

$$L_n[P] = P, \deg(P) \leq n-1,$$

then gives

$$\lim_{n \to \infty} \| f - L_n[f] \|_{p,w} = 0, \tag{23}$$

for every $f \in C[-1,1]$.

Let us summarize what we have:

Theorem on Convergence of Lagrange Interpolation via Orthonormal Expansions
Let $1 < p < \infty, q = \frac{p}{p-1}$, and assume that we have the Marcinkiewicz-Zygmund inequality (19), as well as the mean boundedness (20) in $\mathcal{L}_q(w)$ of the orthonormal expansion. Then for every $f \in C[-1,1]$, we have the convergence (23).

In his work for Jacobi weights, Askey established the Marcinkiewicz inequality (19) using the positivity of suitable order Cesaro means of the $\{S_n\}$, as well as Jensen's inequality. For (20), he used results of Pollard on mean convergence. In particular, Askey [1] showed that the convergence (23) takes place for the Jacobi weight

$$w(x) = (1-x)^\alpha (1+x)^\beta ,$$

if $\alpha = \beta \geq -\frac{1}{2}$ or $-\beta = \alpha \in \left(0, \frac{1}{2}\right)$ provided

$$0 < p < 4\frac{\alpha+1}{2\alpha+1}.$$

The next great set of works on mean convergence of Lagrange interpolation were undertaken by P. Nevai in the 1970's and 1980's. In his landmark AMS memoir [70], he solved Turan's problem I, by showing that the Pollaczek type weight

$$w(t) = \exp\left(-\left(1-t^2\right)^{-1/2}\right), t \in (-1,1)$$

does not allow the convergence (23). Subsequently [73], he showed that this and more general weights solve Turan's problem II, through the following theorem: (take $u = w$ below to solve Turan's problem)

Theorem of Nevai on Turan's Problem II (1985)
Assume that w satisfies Szegö's condition

$$\int_{-1}^{1} \frac{\log w(x)}{\sqrt{1-x^2}} dx > -\infty. \tag{24}$$

Let $1 \leq p_0 < \infty$ and let $u : [-1,1] \to [0,\infty)$ be integrable. Suppose that for every $p > p_0$,

$$\int_{-1}^{1} \left[w(t) \sqrt{1-t^2} \right]^{-p/2} u(t)\, dt = \infty.$$

Then there exists $f \in C[-1,1]$ such that for every $p > p_0$,

$$\limsup_{n\to\infty} \int_{-1}^{1} |L_n[f]|^p u = \infty. \tag{25}$$

Nevai's memoir and papers contain methods to establish (19), and have had a great impact on the theory and application of orthogonal polynomials, not just on Lagrange interpolation.

What is notable in the above theorem is that all we require of the weight w is that it satisfies Szegö's condition (24). This is possible, because the theorem deals with necessary conditions for convergence. Sufficient conditions generally require pointwise bounds on the orthogonal polynomials, which are far more special. But in the past twenty years, such bounds have become available for a variety of weights.

In his 1972 paper, Askey made some conjectures about the mean convergence of orthogonal expansions associated with what are now called generalized Jacobi weights. These have the form

$$w(x) = (1-x)^{\alpha} (1+x)^{\beta} \prod_{j=1}^{m} |x - t_j|^{\rho_j}, x \in (-1,1), \tag{26}$$

where $\alpha, \beta, \rho_j > -1$, and $t_j \in (-1,1), 1 \leq j \leq m$. Thus the weight is a Jacobi weight with added interior singularities. Quite often, w is multiplied by a positive continuous function that has some mild smoothness property, for example a Dini type condition. Badkov [3] essentially resolved Askey's conjectures, in the process obtaining pointwise bounds on the orthonormal polynomials $\{p_n\}_{n=0}^{\infty}$ for w. The latter were one of the important ingredients in Paul Nevai's landmark 1984 paper [72] on Lagrange interpolation. Here is a small sample of what was proved there [72,Thm6,p.695]:

Theorem of Nevai on Generalized Jacobi Weights (1984)
Assume that w is a generalized Jacobi weight (26). Let $0 < p < \infty$. Let u be another generalized Jacobi weight. Then

$$\lim_{n\to\infty} \int_{-1}^{1} |f - L_n[f]|^p u = 0 \ \forall f \in C[-1,1]$$

iff

$$\int_{-1}^{1} \left[w(x) \sqrt{1-x^2} \right]^{-p} u(x)\, dx < \infty.$$

In [72], new ideas were introduced that avoided the use of orthonormal expansions, though the Marcinkiewicz-Zygmund inequalities (19) still played a role. These were subsequently used (together with other new ideas) by G. Mastroianni and his coworkers to study far reaching generalisations of the generalised Jacobi weights. They considered weights of the form

$$w(x) := \prod_{j=0}^{m+1} \left[|x - t_j|^{\rho_j} \omega_j \left(|x - t_j|^{\sigma_j} \right) \right], \tag{27}$$

where

$$-1 = t_0 < t_1 < t_2 < ... < t_m < t_{m+1} = 1,$$

each $\rho_j > -1$ and $\sigma_j = \frac{1}{2}$ or 1 according as $t_j = \pm 1$ or lies inside $(-1, 1)$. Moreover, each ω_j is either identically 1 or is a concave modulus of continuity satisfying, amongst other things,

$$\lim_{x \to 0+} \frac{\omega_j(t)}{t^\varepsilon} = \infty, \text{ for each } \varepsilon > 0.$$

For example, one can take

$$\omega_j(t) = \left[\log \frac{4}{t} \right]^{-\alpha},$$

if $\alpha > 0$. The weights (27) were called *generalised Ditzian Totik weights* because of the use of the Ditzian-Totik modulus of continuity [].

In a series of papers [10], [11], [16], [60-62], [64], Mastroianni and his coworkers established necessary and sufficient conditions for the existence of converse Marcinkiewicz-Zygmund inequalities for generalised Ditzian Totik weights. They then deduced boundedness of Lagrange interpolation in various weighted spaces and also investigated error estimates that are analogous to (12). Obviously the technical nature of the weights complicates the formulation to some degree. Let us state a special case of their earlier error estimates [62]. Let

$$\varphi(x) := \sqrt{1 - x^2}, x \in (-1, 1).$$

Theorem of Mastroianni-Vertesi on the Degree of Mean Convergence (1995)
Let w be a generalised Jacobi weight of the form (26), and u be another generalised Jacobi weight. Let $1 < p < \infty$, and $q := \frac{p}{p-1}$. Assume that

$$\begin{aligned} u &\in L_p[-1,1]; \frac{u}{\sqrt{w\varphi}} \in L_p[-1,1]; \\ \frac{\sqrt{w}}{u\sqrt{\varphi}} &\in L_q[-1,1]; \frac{w}{u} \in L_q[-1,1]. \end{aligned}$$

Then for $n \geq 1$ and every bounded and measurable function $f : [-1, 1] \to \mathbb{R}$,

$$\| (f - L_n[f]) u \|_{L_p[-1,1]} \leq C \widetilde{E}_{n-1}[f]$$

where $C \neq C(n, f)$ and

$$\widetilde{E}_m[f] := \inf \left\{ \| \left(P^+ - P^- \right) u \|_{L_p[-1,1]} : P^- \leq f \leq P^+ \text{ in } (-1,1), \deg \left(P^\pm \right) \leq n \right\}.$$

In recent years, converse Marcinkiewicz-Zygmund inequalities have clearly emerged as the main ingredient of proofs on mean convergence of Lagrange interpolation. It is clear that (22) immediately gives (21) and then the problem of convergence of L_n is reduced to the far easier problem of convergence of the quadrature rule.

These types of inequalities have been surveyed in the author's article [51], in the setting of weights on the whole real line, so we discuss them only briefly here. Amongst the methods for proving the forward estimate (19), are
(i) Nevai's method, which involves estimates for Christoffel functions, the fundamental theorem of calculus, and a Markov-Bernstein inequality for derivatives of polynomials;
(ii) The Large Sieve Method, involving ideas from the large sieve of number theory, and the reproducing kernel for the Chebyshev weight;
(iii) The duality method, in which one starts with a converse inequality such as (22), and then uses duality to pass to a forward inequality.
(iv) Complex Methods, including the use of subharmonicity of $|P|^p$, for P a polynomial, and Carleson measures, which enable one to pass from integral estimates in the upper half plane to estimates on the real line.

Amongst the methods used to prove converse inequalities (22) are
(i) The duality method, which starts from a forward inequality, and proceeds as from (17) to (22) above.
(ii) König's method, which effectively regards $L_n[f]$ as a discrete Hilbert transform. Very very roughly, we write

$$\begin{aligned} L_n[f](x) &= p_n(x) \sum_{j=1}^{n} \frac{f(x_{jn})}{p_n'(x_{jn})(x - x_{jn})} \\ &\approx p_n(x) \sum_{j=1}^{n} \int_{x_{jn}}^{x_{j-1,n}} \frac{f(t)\, w(t)\, \psi_n(t)}{x-t} dt \\ &= p_n(x) \int_{-1}^{1} \frac{f(t)\, w(t)\, \psi_n(t)}{x-t} dt, \end{aligned}$$

where the integral must be interpreted in a principal value sense (it is a Hilbert transform) and ψ_n is a suitable function. The main advantage of this is that one can now use boundedness of the Hilbert transform in a suitable setting.
(iii) Complex methods, again using Carleson measures, Cauchy's integral formula, subharmonicity, ...

Amongst those who used the duality method for converse estimates, especially with application to Lagrange interpolation, are Yuan Xu [88-90] who also applied them to Lagrange interpolation and extended Lagrange interpolation. In the process, Xu obtained impressive extensions of Badkov's results on mean convergence of orthogonal expansions. The work of Mastroianni et al. discussed above, extends much of Xu's work. At the time that the author wrote the survey [51], he did not appreciate that Nevai's paper also contained new ideas for proving converse Marcinkiewicz-Zygmund Inequalities. The latter were employed to great effect by Mastroianni and Russo [61].

Amongst those who have used König's method, which first appeared in [48], [49] in a Banach space setting, are the author and S.B. Damelin [14], [15], [52],

[54]. The author believes that König's method offers the best hope of analysing Lagrange interpolation on the real line, since it is not wedded to orthogonal polynomials, and so may be applied to any interpolation array. Evidence of this is presented in [54], where general converse Marcinkiewicz-Zygmund Inequalities were presented via a modification of König's method. Here is a special case. To emphasise the independence of parameters, we drop the index n from the interpolation points:

Converse Marcinkiewicz-Zygmund Inequality via König's Method (1999)
Let $n \geq 1$ and

$$-1 =: x_{n+1} \leq x_n < x_{n-1} < ... < x_1 \leq x_0 := 1.$$

Let $\nu : [-1, 1] \to [0, \infty)$ be measurable, and let π_n be a polynomial of degree n, whose zeros are $\{x_j\}_{j=1}^n$, normalized so that

$$|\pi_n \nu| \leq 1 \text{ in } [-1, 1]. \tag{28}$$

Let

$$\delta_j := x_{j-1} - x_{j+1}, 1 \leq j \leq n,$$

and assume that there exists $\alpha > 0$ such that for $|j - k| \geq 2$, we have

$$|x_j - x_k| \geq \alpha |j - k|^{1/3} [1 + \log |j - k|]^{2/3} \delta_j. \tag{29}$$

Let $1 < p < \infty$. Then for polynomials P of degree $\leq n - 1$, we have

$$\int_{-1}^{1} |P\nu|^p \leq C \sum_{j=1}^{n} |P(x_j)|^p \left\{ \int_{x_j-\delta_j}^{x_j+\delta_j} |\ell_j \nu|^p + \frac{\delta_j}{[\delta_j |\pi_n'(x_j)|]^p} \right\}. \tag{30}$$

Here $\{\ell_j\}_{j=1}^n$ denote the fundamental polynomials of Lagrange interpolation. The constant C depends on α, p but is independent of $\nu, \{x_j\}_{j=1}^n, n, P$.

While (28) is essentially a normalisation condition, the main restriction is the spacing condition (29). It is satisfied for the zeros of all sets of orthonormal and extremal polynomials of which the author is aware (usually in a far stronger form). Another feature is that C is also independent of ν. In applications, especially to weights on the whole real line, one needs to vary ν. The author believes that this result and its extensions in [], can be used to unify a lot of mean convergence results on Lagrange interpolation. In most examples, one has appropriate estimates for the fundamental polynomials $\{\ell_j\}_{j=1}^n$ and for $\{\pi_n'(x_j)\}_{j=1}^n$, so that the technical right-hand side of (30) transforms into a useful estimate.

The specific example to which this estimate was applied in [54] was extended Lagrange interpolation for weights that decay rapidly at ± 1. Let us first discuss its precursor in [52]. Recall that one of Nevai's examples in resolving Turan's problem was the Pollaczek type weight $w(x) = \exp\left(-\left(1 - x^2\right)^{-1/2}\right)$. As it turns out from [52], the real reason that Turan's problem has a negative solution is that it is inappropriate to use $u = w$ as a weighting factor in (25): one should really use $u = w^{p/2}$. To avoid excessive notation, we shall deal with a subclass of the weights considered in [52]. Let

$$\exp_0 (x) := x$$

and for $k \geq 1$, define $\exp_k$ inductively by

$$\exp_k (x) := \exp \left(\exp_{k-1} (x) \right).$$

For $k \geq 0, a > 0$, define the weight

$$w_{k,\alpha} := \exp \left(-Q_{k,\alpha} \right), \tag{31}$$

where

$$Q_{k,\alpha} (x) := \exp_k \left(\left(1 - x^2 \right)^{-\alpha} \right), x \in (-1, 1). \tag{32}$$

The case $k = 0, \alpha = \frac{1}{2}$ is the Pollaczek type weight considered by Nevai.

In [52], König's method was used to derive a converse Marcinkiewicz inequality, and then the following was deduced:

Theorem on Lagrange Interpolation for Fast Decreasing Weights (1998) *Let $k \geq 1, \alpha > 0$ and $w_{k,\alpha}$ be defined by (31). Let $f : (-1, 1) \to \mathbb{R}$ be Riemann integrable in each compact subinterval of $(-1, 1)$, and assume that*

$$\lim_{|x| \to 1-} (f w_{k,\alpha}) (x) = 0.$$

Let $L_n [f]$ denote the Lagrange interpolation polynomial formed from interpolating f at the zeros of the nth orthonormal polynomial for w.
(a) If $1 < p < 4$,

$$\lim_{n \to \infty} \int_{-1}^{1} \left[|f - L_n [f]| \, w_{k,\alpha}^{1/2} \right]^p = 0.$$

(b) If $p \geq 4$, and

$$\Delta > \frac{1}{4} - \frac{1}{p},$$

$$\lim_{n \to \infty} \int_{-1}^{1} \left[|f - L_n [f]| \, w_{k,\alpha}^{1/2} \left(1 + Q_{k,\alpha}^{-1/3} Q'_{k,\alpha} \right)^{\Delta} \right]^p = 0.$$

Moreover, if $\Delta < \frac{1}{4} - \frac{1}{p}$, then there exists $f \in C [-1, 1]$ that vanishes outside $\left[-\frac{1}{2}, \frac{1}{2} \right]$ for which this last limit fails.

The breakpoint at $p = 4$ is not a weakness of the technique (as should be clear from the negative assertion in (b)), nor peculiar to these weights. It comes from the fact that, very roughly speaking, the orthonormal polynomials p_n for the weight $w_{k,\alpha}$ have the property that on a set of sufficiently large measure

$$\left| p_n (x) \, w_{k,\alpha}^{1/2} (x) \right| \approx \left(a_n^2 - x^2 \right)^{-1/4}.$$

Here $a_n \in (-1, 1)$ is an appropriately chosen nimber. The crux is that the integral of the pth power of $\left(a_n^2 - x^2 \right)^{-1/4}$ diverges if $p \geq 4$. A similar factor of $\left(1-x^2\right)^{-1/4}$ also appears in bounds for orthogonal polynomials for classical weights such as the Jacobi weights, and causes similar problems there. For $w_{k,\alpha}$ the new feature is that $w_{k,\alpha}$ vanishes so strongly at ± 1 that for convergence in L_p, one should replace $w_{k,\alpha}$ by $w_{k,\alpha}^{p/2}$.

We already alluded to results on extended Lagrange interpolation from [54]. It is an old idea that adding two extra points to the set of interpolation points, for example adding ± 1, can substantially improve the convergence properties. It was J. Szabados who used this idea of adding two carefully chosen points to reduce Lebesgue functions for exponential weights [82], and in the context of mean convergence, Szabados' idea was adopted in [12], [13], [56]. For $w_{k,\alpha}$, it was shown that if we add to the zeros $\{x_{jn}\}_{j=1}^{n}$ of p_n, the two points in $(-1,1)$ at which $\left|p_n w_{k,\alpha}^{1/2}\right|$ achieves its maximum, both of which are close to ± 1, then we can achieve a better result than that above: the conclusion of (a) above holds for $1<p<\infty$, so there is no need to damp with $1+Q_{k,\alpha}^{-1/3}Q_{k,\alpha}'$.

The reader will have doubtless observed that all the concrete positive conclusions above deal with interpolation at zeros of orthogonal polynomials, or something very similar. Is there nothing that can be said about general arrays X? Even (30), which ostensibly applies to general arrays, requires detailed bounds on the fundamental polynomials. Unfortunately, this seems to be inherent in the problem.

There are a range of necessary conditions for convergence for general arrays, notably the recent ones of Ying Guang Shi [78-79]. But sufficient conditions for mean convergence are hard to come by. What has recently become apparent is that convergence in L_p norms for $p<1$ is a somewhat easier problem. In essence this seems to be because

$$\int_{-1}^{1}\frac{dx}{|x-x_{jn}|^p}<\infty, p<1.$$

The author showed in [53] how distribution functions and Loomis' Lemma can be used to investigate the case $p<1$, and in a recent paper [55], how they can be used to solve the problem of whether there is convergence for at least one $p>0$:

Theorem on Necessary and Sufficient Conditions for Mean Convergence for General Arrays

Let X be an array of interpolation points in $[-1,1]$. For $n\geq 1$, let π_n be a polynomial of degree n whose zeros are $\{x_{jn}\}_{j=1}^{n}$. Let $v\in L_q[-1,1]$ for some $q>0$. The following are equivalent:
(I) There exists $p>0$ such that for every $f\in C[-1,1]$,

$$\lim_{n\to\infty}\int_{-1}^{1}|(f-L_n[f])\,v|^p=0. \tag{33}$$

(II) There exists $r>0$ such that

$$\sup_{n\geq 1}\left(\int_{-1}^{1}|\pi_n v|^r\right)^{1/r}\sum_{j=1}^{n}\frac{1}{|\pi_n'(x_{jn})|}<\infty. \tag{34}$$

What is refreshing is the relative simplicity: the single condition (34) guarantees convergence for all $f\in C[-1,1]$ in some L_p. Of course this is achieved at the price of apparently unrelated parameters p, r. In fact, it was Ying Guang Shi who showed [79] that the convergence (33) easily implies (34) for $r = p$. The author established the more difficult implication (34) $\Rightarrow$ (33) for $p < \min\left\{\frac{1}{2}, \frac{r}{2}, q\right\}$. Of course, the result admits generalisations, and one may replace $[-1, 1]$ by any compact set, and introduce a weight u to multiply $\pi'_n(x_{jn})$.

There are many strands of this topic of mean convergence that we have omitted. For example, one may replace convergence on continuous, or bounded and Riemann integrable f, by convergence on functions with integrable singularities. We have avoided mean convergence of Hermite and Hermite-Fejér interpolation - for some of that see [74], [75], [84], [89], [90].

Still more importantly, we have neglected weights on infinite intervals, something which Shohat already considered in 1939. To give a proper treatment of these would require a lengthy survey on its own, and we rather refer the reader to some of the more important and more recent papers [14], [15], [46], [57], [65], [71], and the monographs of Freud [46] and Mhaskar [65].

3 Conclusion

The reader will observe the impressive amount of effort devoted to mean convergence of Lagrange interpolation in recent years. That activity takes place, for example, in China in the East; the USA in the West; South Africa in the south; and Italy in the north, and of course Hungary. This is testimony to the continuing impact of the Erdös-Turan theorem.

In the broader topic of Lagrange interpolation one cannot but marvel at the strength of the Hungarian school that Erdös helped to build. And again that broader enterprise is not restricted to Hungary, but stretches across the world, from Australia to China to India, from Russia to Israel to Western Europe, to the USA. No doubt, the travels of Erdös helped to spread interest.

Acknowledgement and Disclaimer

The author would like to thank the organisers of the conference for their very kind invitation, and their gracious hospitality. He also ackowledges the efforts of T. Mills, B. Shekhtman, J. Szabados and P. Vertesi to correct some of the more embarassing historical errors (fortunately prior to his talk). No doubt others are present in this survey.

The author is sure that important papers and ideas have been omitted, and credits for ideas may well be incomplete or distorted. Several of the references are included for completeness, and have not been discussed in the text. For all these omissions, I apologize in advance, but would welcome corrections to 036dsl@cosmos.wits.ac.za - in surveying a topic one often learns as much from the feedback after publication as from the effort of researching and writing the paper itself.

References

[1] R. Askey, Mean Convergence of Orthogonal Series and Lagrange Interpolation, 23(9173), 71-85.

[2] L. Babai, C. Pomerance, P. Vertesi, The Mathematics of Paul Erdös, Notices of the AMS, 45(1998), 19-31.

[3] V. Badkov, Convergence in Mean and Almost Everywhere of Fourier Series in Orthogonal Polynomials, Mat. Sbornik, 95(137)(1974), 229-262.

[4] S.N. Bernstein, Sur l' ordre de la meillure approximation des fonctions continues par des polynomes de degre donne, Memoires publies par la classe des sci. Acad. de Belgique, 4(1912), 1-103.

[5] S.N. Bernstein, Quelques remarques sur l'interpolation, Communications Math. Soc. Kharkov, 14(1914).

[6] C. de Boor and A. Pinkus, Proof of the Conectures of Bernstein and Erdös concerning the Optimal Nodes for Polynomial Interpolation, J. Approx. Theory, 24(1978), 289-303.

[7] L. Brutman, A. Pinkus, On the Erdös Conjecture Concerning Minimal Norm Interpolation on the Unit Circle, SIAM J. Numer. Anal., 17(1980), 373-375.

[8] L. Brutman, A Note on Polynomial Interpolation at the Chebyshev Extrema Nodes, J. Approx. Th., 42(1984), 283-292.

[9] L. Brutman, Lebesgue Functions for Polynomial Interpolation - A Survey, Annals of Numerical Mathematics, 4(1997), 111-127.

[10] M.R. Capobianco, G. Mastroianni, Uniform Boundedness of Lagrange Operator in Some Weighted Sobolev-type Spaces, Math. Nach., 187(1997), 1-17.

[11] G. Criscuolo, G. Mastroianni, Fourier and Lagrange Operators in Some Weighted Sobolev Type Spaces, Acta Sci. Math (Szeged), 60(1995), 131-146.

[12] S.B. Damelin, The Weighted Lebesgue Constant of Lagrange Interpolation for Non-Szegö Weights on $[-1,1]$, Acta Math. Hungarica, 81(1998), 211-228.

[13] S.B. Damelin, Lagrange Interpolation for non-Szegö Weights on $[-1,1]$, to appear in the Proceedings of the International Workshop on Approximation Theory and Numerical Analysis.

[14] S.B. Damelin and D.S. Lubinsky, Necessary and Sufficient Conditions for Mean Convergence of Lagrange Interpolation for Erdös Weights, Canadian Math. J., 40(1996), 710-736.

[15] S.B. Damelin and D.S. Lubinsky, Necessary and Sufficient Conditions for Mean Convergence of Lagrange Interpolation for Erdös Weights II, Canadian Math. J., 40(1996), 737-757.

[16] B. Della Vecchia, G. Mastroianni, P. Vertesi, Boundedness of Lagrange and Hermite Operators, (in) Approximation Theory IDOMAT95, Akademie Verlag, 1995, pp. 53-70.

[17] Z. Ditzian and V. Totik, Moduli of Smoothness, Springer, Berlin, 1987

[18] V.K. Dzadyk and V.V. Ivanov, On Asymptotics and Estimates for the Uniform Norms of the Lagrange Interpolation Polynomials Corresponding to the Chebyshev Nodal Points, Analysis Mathematica, 9(1983), 85-97.

[19] T. Erdelyi and P. Vertesi, Paul Erdös: In Memoriam, J. Approx. Theory, 94(1998), 1-41.

[20] P. Erdös, On divergence properties of the Lagrange interpolation parabolas, Annals. of Math. 42(1941), 309–315.

[21] P. Erdös, Problems and results on the theory of interpolation, I., Acta Math. Hungar. 9(1958), 381–388.

[22] P. Erdös, Problems and results on the convergence and divergence properties of the Lagrange interpolation polynomials and some extremal problems, Mathematica, Cluj 10(1968), 65–73.

[23] P. Erdös, Problems and results on polynomials and interpolation, (in) Aspects of contemporary complex analysis (eds. D.A. Brannan and J.G. Clunie), Academic Press, 1980, 383-391.

[24] P. Erdös and G. Grünwald, Note on an elementary problem of interpolation, Bull. Amer. Math. Soc. 44(1938), 515–518.

[25] P. Erdös and G. Halasz, On the arithmetic means of Lagrange interpolation, in: Approximation theory, (eds.: J. Szabados and K. Tandori), Colloq. Math. Soc. J. Bolyai 58(1991), 263–274.

[26] P. Erdös, A. Kroo and J. Szabados, On convergent interpolatory polynomials, J. Approx. Theory 58(1989), 232–241.

[27] P. Erdös and B. A. Lengyel, On fundamental functions of Lagrangean interpolation, Bull. Amer. Math. Soc. 44(1938), 828–834.

[28] P. Erdös, J. Szabados, A. Varma and P. Vertesi, On an interpolation theoretical extremal problem, Studia Sci. Math. Hungar. 29(1994), 55–60, (J. Szabados; A.K. Varma; P. Vértesi).

[29] P. Erdös, J. Szabados, and P. Vertesi, On the Integral of the Lebesgue Function of Interpolation II, Acta Math. Hungar., 68(1995), 1-6.

[30] P. Erdös and P. Turan, On interpolation, I. Quadrature and mean convergence in the Lagrange interpolation, Ann. of Math. 38(1937), 142–155.

[31] P. Erdös and P. Turan, On interpolation, II. On the distribution of the fundamental points of Lagrange and Hermite interpolation, Ann. of Math. 39(1938), 703–724.

[32] P. Erdös and P. Turan, On interpolation, III. Interpolatory theory of polynomials, Ann. of Math. 41(1940), 510–553.

[33] P. Erdös and P. Turan, On the uniformly dense distribution of certain sequences of points, Ann. of Math. 41(1940), 162–173.

[34] P. Erdös and P. Turan, On the uniform distribution of the roots of certain polynomials, Ann. of Math. 43(1942), 59–64.

[35] P. Erdös and P. Turan, On some convergence properties in the interpolation polynomials, Ann. of Math. 44(1943), 330–337.

[36] P. Erdös and P. Turan, On the role of the Lebesgue functions in the theory of the Lagrange interpolation, Acta Math. Hungar. 6(1955), 47–66.

[37] P. Erdös and P. Turan, An extremal problem in the theory of interpolation, Acta Math. Hungar. 12(1961), 221–234.

[38] P. Erdös and P. Vertesi, On the almost everywhere divergence of Lagrange interpolation polynomials for arbitrary systems of nodes, Acta Math. Acad. Sci. Hungar. 36 (1980), 71–99.

[39] P. Erdös and P. Vertesi, On the almost everywhere divergence of Lagrange interpolation (in) Approximation and function spaces (Gdańsk, 1979), pp. 270–278, North-Holland, Amsterdam-New York,1981.

[40] P. Erdös and P. Vertesi, On the Lebesgue functions of interpolation, Functional Analysis and Approximation (1981), 299–309.

[41] P. Erdös and P. Vertesi, Problems and results on polynomials and interpolation, in: Approximation theory, (eds.: J. Szabados; K. Tandori), Colloq. Math. Soc. J. Bolyai 58(1991), 253–261.

[42] G. Faber, Über die interpolatorische Darstellung stetiger Funktionen, Jahresbericht der Deutsch. Math. Ver., 23(1914), 190-210.

[43] L. Fejér, Über interpolation, Göttinger Nachrichten, (1916), 66-91.

[44] L. Fejér, Lagrangesche interpolation und die zugehörigen konjugierten Punkte, Mathematische Annalen, 106(1932), 1-55.

[45] E. Feldheim, Quelque recherches sur l'interpolation de Lagrange et d'Hermite, Mathematische Zeitschrift, 44(1938), 55-84.

[46] G. Freud, Orthogonal Polynomials, Akademiai Kiado/ Pergamon Press, Budapest, 1971.

[47] T. Kilgore, A Characterization of the Lagrange Interpolation Interpolating Projection with Minimal Chebyshev Norm, J. Approx. Theory, 24(1978), 273-288.

[48] H. König, Vector Valued Lagrange Interpolation and Mean Convergence of Hermite Series, Proc. Essen Conference on Functional Analysis, North Holland.

[49] H. König and N.J. Nielsen, Vector Valued L_p Convergence of Orthogonal Series and Lagrange Interpolation, Forum Math., 6(1994), 183-207.

[50] D. Kubayi, Lebesgue Functions, ...

[51] D.S. Lubinsky, Marcinkiewicz-Zygmund Inequalities: Methods and Results, (in) Recent Progress in Inequalities, A Volume dedicated to Prof. Mitrinovic, Kluwer, 1997, 1-28.

[52] D.S. Lubinsky, Mean Convergence of Lagrange Interpolation for Exponential Weights on $[-1, 1]$, Canadian Math. Journal, 50(1998), 1273-1297.

[53] D.S. Lubinsky, On Boundedness of Lagrange Interpolation in $L_p, p < 1$, J. Approx. Theory, 96(1999), 399–404.

[54] D.S. Lubinsky, On Converse Marcinkiewicz–Zygmund Inequalities in $L_p, p > 1$, Constr. Approx., 15(1999), 577-610.

[55] D.S. Lubinsky, On Mean Convergence of Lagrange Interpolation for General Arrays, to appear in J. Approx. Th.,

[56] D.S. Lubinsky and G. Mastroianni, Mean Convergence of Extended Lagrange Interpolation with Freud Weights, Acta Math Hungar., 84(1999), 47-63.

[57] D.S. Lubinsky and D. Matjila, Necessary and Sufficient Conditions for Mean Convergence of Lagrange Interpolation for Freud Weights, SIAM J. Math. Anal., 26(1995), 238-262.

[58] J. Marcinkiewicz, Quelques remarques sur l'interpolation, Acta Sci. Math. Szeged, 8(1937), 127-130.

[59] G. Mastroianni, Boundedness of Lagrange Operator in Some Functional Spaces, a Survey, (in) Approximation Theory and Function Spaces, Bolyai Math. Soc. Studies, 5(1996), 117-139.

[60] G. Mastroianni and M.G. Russo, Weighted Marcinkiewicz Inequalities and Boundedness of the Lagrange Operator, to appear.

[61] G. Mastroianni and M.G. Russo, Lagrange Interpolation in Weighted Besov Spaces, Constr. Approx., 15(1999), 257-289.

[62] G. Mastroianni and P. Vertesi, Mean Convergence of Lagrange Interpolation on an Arbitrary System of Nodes, Acta Sc. Math. Szeged, 57(1993), 429-441.

[63] G. Mastroianni and P. Vertesi, Weighted L_p Error of Lagrange Interpolation, J. Approx. Theory, 82(1995), 321-339.

[64] G. Mastroianni and P. Vertesi, Some Applications of Generalized Jacobi Weights, Acta Math. Hungar., 77(1997), 323-357.

[65] H.N. Mhaskar, Weighted Polynomial Approximation, World Scientific, Singapore, 19??.

[66] T.M. Mills and S.J. Smith, A Note on the Newton-Cotes Integration Formula, J. Approx. Th., 66(1991), 98-105.

[67] T.M. Mills and S.J. Smith, Lagrange Interpolation Polynomials Based on Equidistant Nodes, Math. Scientist, 16(1991), 107-117.

[68] T.M. Mills and S.J. Smith, On the Lebesgue Function for Lagrange Interpolation with Equidistant Nodes, J. Austral. Math. Soc., 52 (1992), 111-118.

[69] P. Nevai, Lagrange Interpolation at Zeros of Orthogonal Polynomials, (in) Approximation Theory II, (eds. G.G. Lorentz et al.), Academic Press, New York, 1976, pp. 163-201.

[70] P. Nevai, Orthogonal Polynomials, Memoirs of the Amer. Math. Soc., 213(1979).

[71] P. Nevai, Mean Convergence of Lagrange Interpolation II, J. Approx. Theory, 30(1980), 263-276.

[72] P. Nevai, Mean Convergence of Lagrange Interpolation III, Transactions of the Amer. Math. Soc., 282(1984), 669–698.

[73] P. Nevai, Solution of Turán's Problem on Divergence of Lagrange Interpolation in L^p with $p > 2$, J. Approx. Th., 43(1985), 190-193.

[74] P. Nevai, Geza Freud, Orthogonal Polynomials and Christoffel Functions: A Case Study, J. Approx. Theory, 48(1986), 3-167.

[75] P. Nevai and P. Vertesi, Mean Convergence of Hermite-Fejér Interpolation, J. Math. Anal. Applns., 105(1985), 26-58.

[76] E.A. Rakhmanov, On the Asymptotics of the Ratio of Orthogonal Polynomials, Math. USSR Sbornik, 32(1977), 199-213.

[77] E.A. Rakhmanov, On the Asymptotics of the Ratio of Orthogonal Polynomials II, Math. USSR Sbornik, 46(1983), 105-117.

[78] Y.G. Shi, Bounds and Inequalities for Arbitrary Orthogonal Polynomials on Finite Intervals, J. Approx. Theory, 81(1995), 217-230.

[79] Y.G. Shi, Mean Convergence of Interpolatory Processes on an Arbitrary System of Nodes, Acta Math. Hungar., 70(1996), 27–38.

[80] Y.G. Shi, On L_p Extremal Polynomials, Constr. Approx., 14(1998), 187-194.

[81] H. Stahl and V. Totik, General Orthogonal Polynomials, Cambridge University Press, Cambridge, 1992.

[82] J. Szabados, Weighted Lagrange and Hermite-Fejer Interpolation on the Real Line, J. of Inequal. and Applns., 1(1997), 99-123.

[83] J. Szabados and P. Vertesi, Interpolation of Functions, World Scientific, Singapore, 1990.

[84] J. Szabados, P. Vertesi, A Survey on Mean Convergence of Interpolatory Processes, J. Comp. Appl. Math., 43(1992), 3-18.

[85] P. Vertesi, On The Lebesgue Function of Weighted Lagrange Interpolation I (Freud Type Weights), to appear in Constr. Approx.

[86] P. Vertesi, On The Lebesgue Function of Weighted Lagrange Interpolation II J. Austral. Math. Soc., 65(1998), 145-162.

[87] J.L. Walsh, Interpolation and Approximation by Rational Functions in the Complex Domain, Amer. Math. Soc. Colloq. Publns., Vol. XX, (Fifth Edition), American Math. Soc., Providence, 1969

[88] Y. Xu, On the Marcinkiewicz-Zygmund Inequality, (in) "Progress in Approximation Theory", (P. Nevai and A. Pinkus, eds.), Academic Press, San Diego, 1991, pp.879-891.

[89] Y. Xu, Mean Convergence of Generalized Jacobi Series and Interpolating Polynomials I, J. Approx. Th., 72(1993), 237-251.

[90] Y. Xu, Mean Convergence of Generalized Jacobi Series and Interpolating Polynomials II, J. Approx. Th., 76(1994), 77-92.