Approximating Lipschitz and continuous functions by polynomials; Jackson's theorem

G. J. O. JAMESON

Introduction

The celebrated theorem of Weierstrass, dating from 1885, states that continuous functions can be uniformly approximated by polynomials on any bounded, closed interval. But just how well can we approximate by polynomials of a certain degree? Let us introduce some notation to facilitate the discussion. For an interval I (which will usually be $[-1, 1]$ or $[0, 1]$), denote by $C(I)$ the space of continuous functions on I, and write $\|f\|$ for $\sup\{|f(x) : x \in I|\}$ (the notation $\|f\|_\infty$ is often used). Uniform convergence of f_n to f (on I) equates to the statement that $\|f_n - f\| \to 0$ as $n \to \infty$. Denote by $\prod_n$ the space of polynomials of degree not more than n. This is a linear subspace of $C(I)$ of dimension $n + 1$. We write

$$\operatorname{dist}(f, \textstyle\prod_n) = \inf\{\|f - p\| : p \in \textstyle\prod_n\}.$$

Since $\prod_n$ is finite-dimensional, there is a p for which the infimum is attained, the 'best uniform approximation' to f in $\prod_n$ (we will use this fact, but it is not really essential for our exposition). Our objective is to find estimations of $\operatorname{dist}(f, \prod_n)$, given information about f.

We shall see that particularly pleasant answers are available for *Lipschitz* functions. Recall that a function on $\mathbb{R}$, or an interval I, is 'Lipschitz' if there exists M such that $|f(x_2) - f(x_1)| \leqslant M|x_2 - x_1|$ for all x_1, x_2 in I: we denote the smallest such M by $\operatorname{Lip}(f)$. Of course, a sufficient (but not necessary) condition is that f is differentiable, with $\|f'\| \leqslant M$.

A neat application of the intermediate value theorem sets a clear limit to what can be achieved. (There is a slight gain in simplicity by stating it for $n - 1$ instead of n.)

Theorem 1: For each n, there is a function f_n on $[0, 1]$ such that $\operatorname{Lip}(f_n) = 1$ and $\operatorname{dist}(f_n, \prod_{n-1}) = \frac{1}{2n}$.

Proof: Define f_n as follows:

$$f\left(\frac{r}{n}\right) = \frac{(-1)^r}{2n} \text{ for } 0 \leqslant r \leqslant n,$$

so that $f_n(x)$ is alternately $\pm\frac{1}{2n}$ at the points $\frac{r}{n}$. Define $f_n(x)$ by linear interpolation between these points. Then f_n has gradients ± 1 on successive intervals of length $\frac{1}{n}$, so $\operatorname{Lip}(f_n) = 1$.

Now suppose that p is a polynomial with $\|f_n - p\| < \frac{1}{2n}$. Then $p\left(\frac{r}{n}\right)$ is strictly positive for r even and strictly negative for r odd. By the

https://doi.org/10.1017/mag.2023.58 Published online by Cambridge University Press

intermediate value theorem, p has a zero in each interval $\left(\frac{r-1}{n}, \frac{r}{n}\right)$, hence at least n zeros. So the degree of p is at least n. In other words, if $p \in \prod_{n-1}$, then $\|f_n - p\| \geqslant \frac{1}{2n}$. Equality occurs with $p = 0$.

If f_n is stretched to the interval $[-1, 1]$, the same applies, but $\text{Lip}(f_n)$ becomes $\frac{1}{2}$.

A previous article [1] described the rate of approximation achieved by the *Bernstein polynomials* $B_n(f)$ derived from a given function f: it was shown that $\|B_n(f) - f\| \leqslant \frac{\text{Lip}(f)}{2\sqrt{n}}$. However, the correct order of magnitude of $\text{dist}(f, \prod_n)$ for Lipschitz functions is actually $O\left(\frac{1}{n}\right)$, as seen in Theorem 1. This was proved in 1912 by the American mathematician Dunham Jackson; his article [2] built on a doctoral thesis written in Göttingen under the guidance of Edmund Landau. Here we will give an exposition of this result, which surely deserves to be better known. We present a clever proof devised later by P. Korovkin [3].

Jackson's theorem actually applies in the first instance to the parallel problem of approximation to periodic functions by trigonometric polynomials; of course this is the domain of Fourier analysis. The connection is easily explained. We denote by $C_{2\pi}$ the space of all continuous, 2-periodic functions on $\mathbb{R}$. Of course, to specify such a function it is enough to give the values on $[\pi, -\pi]$, necessarily with $f(-\pi) = f(\pi)$. Denote by T_n the set of all trigonometric polynomials of degree n, i.e. functions of the form $a_0 + \sum_{k=1}^{n} (a_k \cos kt + b_k \sin kt)$.

Now given $f \in C[-1,1]$, define $F \in C_{2\pi}$ by $F(t) = f(\cos t)$. Note that $\text{Lip}(F) \leqslant \text{Lip}(f)$, since $|\cos t_2 - \cos t_1| \leqslant |t_2 - t_1|$. Choose $P \in T_n$ such that $\|F - P\| = \text{dist}(F, T_n)$. Since F is an even function, we can take $P(t)$ to be of the form $\sum_{r=0}^{n} a_r \cos rt$, without the sine terms (replace $P(t)$ by $\frac{1}{2}P(t) + \frac{1}{2}P(-t)$). For each r, there is a polynomial q_r of degree r such that $\cos rt = q_r(\cos t)$ (these are the 'Chebyshev polynomials', but we do not need to know anything about them beyond their existence). Let $p(x) = \sum_{r=0}^{n} a_r q_r(x)$: then $p \in \prod_n$ and $p(\cos t) = P(t)$, so $\|f - p\| = \|F - P\|$, hence $\text{dist}(f, \prod_n) \leqslant \text{dist}(F, T_n)$.

It will help to make sense of the proof of Jackson's theorem if we first outline the simpler and better known case of modified Fourier series, as given by the Fejér kernel. The version for Lipschitz functions is both simple and attractive; surprisingly, it is not mentioned in many books. The method then adapts easily to continuous functions, using the concept of 'modulus of continuity'. Similar steps, but with some extra work, serve to establish Jackson's theorem.

A final twist is that we can apply Jackson's theorem to deduce estimations of the rate of convergence of actual Fourier series, showing that for Lipschitz functions it is similar to the rate for the Fejér kernel. This also would seem to be not very well known.

The modulus of continuity

This notion will enable us to adapt the results for Lipschitz functions to the class of all continuous functions.

The *modulus of continuity* of a function f on an interval I is the function $\omega_f(\delta)$ defined for $\delta > 0$ by

$$\omega_f(\delta) = \sup\left\{\left|f(x_2) - f(x_1)\right| : x_1, x_2 \in I \text{ and } \left|x_2 - x_1\right| \leqslant \delta\right\}.$$

Uniform continuity of f on I equates to the statement that $\omega_f(\delta) \to 0$ as $\delta \to 0^+$. Clearly, if $\text{Lip}(f) = M$, then $\omega_f(\delta) \leqslant M\delta$. If $F(t) = f(\cos t)$, then $\omega_F(\delta) \leqslant \omega_f(\delta)$. As an example, if $f(x) = |x|^\alpha$, where $0 < \alpha \leqslant 1$, then one can show that $\omega_f(\delta) = \delta^\alpha$.

For points further apart, we can estimate $\left|f(x_2) - f(x_1)\right|$ as follows.

Lemma 1: For all x_1, x_2 in I, we have

$$\left|f(x_2) - f(x_1)\right| \leqslant \left(\frac{1}{\delta}\left|x_2 - x_1\right| + 1\right)\omega_f(\delta).$$

Proof: We may assume that $x_2 > x_1$. Let $x_2 - x_1 = (k + \theta)\delta$, where k is an integer and $0 \leqslant \theta < 1$, and let $y_r = x_1 + r\delta$ for each r. Then $\left|f(y_r) - f(y_{r-1})\right| \leqslant \omega_f(\delta)$ for $1 \leqslant r \leqslant k$, and $\left|f(x_2) - f(y_k)\right| \leqslant \omega_f(\delta)$. So $\left|f(x_2) - f(x_1)\right| \leqslant (k + 1)\,\omega_f(\delta)$. The statement follows.

The Dirichlet kernel and generalisations

We need some basic notions from Fourier analysis; details can be seen in many books. For any integrable 2π-periodic function, the *Fourier coefficients* are

$$a_n(f) = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\cos nt\,dt, \qquad b_n(f) = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\sin nt\,dt.$$

We define the function $S_n(f)$ (or S_nf) by

$$\left(S_nf\right)(x) = \tfrac{1}{2}a_0(f) + \sum_{k=1}^{n}\left(a_k(f)\cos kx + b_k(f)\sin kx\right). \tag{1}$$

This is the nth partial sum of the Fourier series for f. Clearly S_n is a linear operator mapping $C_{2\pi}$ to T_n, with $S_n(f) = f$ for $f \in T_n$: in other words, it is a *projection* onto T_n.

https://doi.org/10.1017/mag.2023.58 Published online by Cambridge University Press

Substituting for $a_n(f)$ and $b_n(f)$, we see that

$$\begin{aligned}(S_nf)(x) &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)\,dt + \frac{1}{\pi}\sum_{k=1}^{n}\int_{-\pi}^{\pi} f(t)\big(\cos kx \cos kt + \sin kx \sin kt\big)\,dt \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t) D_n(x-t)\,dt, \qquad (2)\end{aligned}$$

where D_n is the Dirichlet kernel

$$D_n(t) = \tfrac{1}{2} + \sum_{k=1}^{n} \cos kt. \qquad (3)$$

Note that $D_n(-t) = D_n(t)$. From (3) it is clear that $\frac{1}{\pi}\int_{-\pi}^{\pi} D_n(t) = 1$. Using the identity $\sin\left(k + \frac{1}{2}\right)t - \sin\left(k - \frac{1}{2}\right)t = 2\cos kt \sin\frac{1}{2}t$, we find that for $t \neq 2k\pi$,

$$D_n = \frac{\sin\left(n + \frac{1}{2}\right)t}{2\sin\frac{1}{2}t}. \qquad (4)$$

Identity (2) says that $S_n(f) = f * D_n$, where the convolution $f * g$ is defined for 2π-periodic functions by

$$(f * g)(x) = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\,g(x - t)\,dt.$$

By substituting $x - t = u$ and equating the integrals on $[-\pi, x - \pi]$ and $[\pi, x + \pi]$, we see that $f * g = g * f$ for 2π-periodic functions, so we have equally

$$(S_nf)(x) = (D * f)(x) = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x - t) D_n(t)\,dt. \qquad (5)$$

From now on we will use the notation $D_n * f$ rather than $S_n(f)$.

We now describe a weighted version of this construction. Given non-negative numbers $\rho_{k,n}$, let

$$R_n(t) = \tfrac{1}{2} + \sum_{k=1}^{n} \rho_{k,n} \cos kt. \qquad (6)$$

In the same way as for D_n, we have

$$(R * f) = \tfrac{1}{2}a_0(f) + \sum_{k=1}^{n} \rho_{k,n}\big(a_k(f)\cos kx + b_k(f)\sin kx\big). \qquad (7)$$

In general, these expressions are no longer partial sums of a series, and $R_n * f$ does not equal f for $f \in T_n$: for example, if $f(x) = \cos kx$, then $(R_n * f)(x) = \rho_{k,n}\cos kx$. Because of the constant term $\frac{1}{2}$, we still have

$$\frac{1}{\pi}\int_{-\pi}^{\pi} R_n(t)\,dt = 1, \qquad (8)$$

and hence for all x,

https://doi.org/10.1017/mag.2023.58 Published online by Cambridge University Press

$$\big(R_n * f\big)(x) - f(x) = \frac{1}{\pi}\int_{-\pi}^{\pi}\big[f(x-t) - f(x)\big]\,R_n(t)\,dt. \tag{9}$$

This identity will be the starting point for all our approximation theorems. Note also that if $R_n(t) \geqslant 0$ for all t, then (8) implies that $\|R_n * f\| \leqslant \|f\|$ for all $f \in C_{2\pi}$.

The Fejér operator

An important special case of the above is the *Fejér kernel* $K_n(t)$ defined for $n \geqslant 1$ by

$$K_n = \frac{1}{n}\big(D_0 + D_1 + \dots + D_{n-1}\big)$$

(also $K_0 = D_0 = \frac{1}{2}$). This equates to taking $\rho_{k,n} = 1 - \frac{k}{n}$ in (6). Of course, this operator actually maps into T_{n-1}. Using (4) and the identity

$$\sin^2\tfrac{1}{2}(k+1)t - \sin^2\tfrac{1}{2}kt = \sin\big(k+\tfrac{1}{2}\big)t\,\sin\tfrac{1}{2}t,$$

one can show that

$$K_n(t) = \frac{\sin^2\frac{1}{2}nt}{2n\,\sin^2\frac{1}{2}t}. \tag{10}$$

for $t \neq 2k\pi$. Also $K_n(0) = \frac{1}{2}n$, so $K_n(t) \geqslant 0$ for all t. This positivity allows some estimations that do not work for D_n. It has been gained at the cost of not preserving elements of T_{n-1}.

We now estimate $\|(K_n * f) - f\|$, first for Lipschitz functions. We apply formula (9), with $R_n = K_n$. If $\mathrm{Lip}(f) = M$, then $|f(x-t) - f(x)| \leqslant M\,|\,t\,|$. Since $K_n(t) \geqslant 0$, we have

$$\|(K_n * f) - f\| \leqslant \frac{M}{\pi}\int_{-\pi}^{\pi}|t|\,K_n(t)\,dt.$$

So we need an estimation of this integral. We shall derive one from an estimate of $\int_0^\pi tD_n(t)\,dt$.

Lemma 2: For each $n \geqslant 1$,

$$\int_0^\pi tD_n(t)\,dt \leqslant \frac{1}{n}. \tag{11}$$

Proof: Integrating by parts, we find

$$\int_0^\pi t\,\cos kt\,dt = \begin{cases} -\frac{2}{k^2} & \text{for } k \text{ odd,} \\ 0 & \text{for } k \text{ even.} \end{cases} \tag{12}$$

So, using (3) and picking out the odd terms, we have

$$\int_0^\pi tD_{2n-1}(t)\,dt = \int_0^\pi tD_{2n}(t)\,dt = \frac{\pi^2}{4} - \sum_{k=1}^{n}\frac{2}{(2k-1)^2}.$$

https://doi.org/10.1017/mag.2023.58 Published online by Cambridge University Press

Denote these integrals by J_n. Since $\displaystyle\sum_{k=1}^{\infty} \frac{2}{(2k-1)^2} = \frac{\pi^2}{4}$,

$$J_n = \sum_{k=n+1}^{\infty} \frac{2}{(2k-1)^2}.$$

It is easily checked that

$$\frac{2}{(2k-1)^2} \leqslant \frac{1}{2k-2} - \frac{1}{2k},$$

so by cancellation $J_n \leqslant \frac{1}{2n}$. The statement (both for n even and n odd) follows.

Conversely, it is clear that $J_n \geqslant \dfrac{1}{2n+1}$, so that $\displaystyle\int_0^{\pi} tD_n(t)\,dt \geqslant \frac{1}{n+2}$. Recall that $D_n(t)$ has positive and negative values: this is *not* an estimation of $\displaystyle\int_0^{\pi} t\left|D_n(t)\right| dt$.

Write H_n for the harmonic sum $\displaystyle\sum_{r=1}^{n} \frac{1}{r}$. Of course, $\log n < H_n < \log n + 1$.

Lemma 3: For all $n \geqslant 1$,

$$\int_{-\pi}^{\pi} |t| K_n(t)\,dt \leqslant \frac{2L_n}{n}, \tag{13}$$

where

$$L_n = H_{n-1} + \frac{\pi^2}{4}. \tag{14}$$

Proof: This follows at once from (11) and $K_n = \frac{1}{n}\left(\frac{1}{2} + D_1 + \ldots D_{n-1}\right)$.

So we have completed the proof of the following estimation for Lipschitz functions:

Theorem 2: If $f \in C_{2\pi}$ and $\mathrm{Lip}(f) = M$, then

$$\left\|(K_n * f) - f\right\| < 2M\frac{L_n}{n\pi}. \tag{15}$$

Very few books that I have seen include an explicit statement of Theorem 2. One that does is [4, p. 49], where it is attributed to Bernstein.

The same reasoning serves to show that the estimate (15) is essentially optimal:

Example 1. Let $f(x) = |x|$ on $[\pi, -\pi]$. Then $\mathrm{Lip}(f) = 1$, and by (9) again,

$$(K_n * f)(0) - f(0) = \frac{1}{\pi}\int_{-\pi}^{\pi} |t| K_n(t)\,dt.$$

https://doi.org/10.1017/mag.2023.58 Published online by Cambridge University Press

By the remark following Lemma 2, with the term $D_0 = \frac{1}{2}$ separated out,

$$\int_0^\pi tK_n(t) > \frac{1}{n}\left(\frac{\pi^2}{4} + \frac{1}{3} + \ldots + \frac{1}{n+1}\right) > \frac{H_n}{n},$$

hence $(K_n * f)(0) - f(0) > \dfrac{2H_n}{n\pi}$.

With minimal further work, the notion of modulus of continuity enables us to adapt the reasoning to continuous functions.

Theorem 3: For all $f \in C_{2\pi}$, with L_n as in (14), we have

$$\|(K_n * f) - f\| \leqslant 3\omega_f\left(\frac{L_n}{n\pi}\right). \tag{16}$$

Proof. Again we apply (9) and (13), using Lemma 1 to estimate $|f(x-t) - f(x)|$. For any $\delta > 0$,

$$|(K_n * f)(x) - f(x)| \leqslant \frac{1}{\pi}\omega_f(\delta)\int_{-\pi}^{\pi}\left(\frac{|t|}{\delta} + 1\right)K_n(t)\,dt \leqslant \omega_f(\delta)\left(\frac{2L_n}{n\pi\delta} + 1\right).$$

Taking $\delta = L_n/(n\pi)$, we obtain (16).

This implies uniform convergence of $K_n * f$ to f, since $\omega_f(L_n/n\pi) \to 0$ as $n \to \infty$.

Other choices of δ deliver variants of (16). For example, the choice $\delta = \frac{1}{n}$ gives

$$\|(K_n * f) - f\| \leqslant \omega_f\left(\tfrac{1}{n}\right)\left(\frac{2L_n}{\pi} + 1\right).$$

In some cases this gives a stronger bound than (16). However, it does not imply uniform convergence, because there are continuous functions for which $L_n\omega_f\left(\frac{1}{n}\right)$ does not tend to 0.

Note: The more usual proof of uniform convergence for the Fejér operator, seen in many books, separates the interval $[-\delta, \delta]$ and uses (10). With the notion of $\omega_f(\delta)$ incorporated, the method leads to an estimate of the following type, clearly weaker than (16):

$$\|(K_n * f) - f\| \leqslant \omega_f\left(\frac{1}{n^{1/3}}\right) + \frac{8}{n^{1/3}}\|f\|.$$

Jackson's theorem

We now present Korovkin's method for Jackson's theorem, roughly following [5, p. 17-24]. Let $R_n(t)$ be as in (6), with $\rho_{k,n}$ to be chosen. We start with some Lemmas.

Lemma 4: We have $t^2 \leqslant \dfrac{\pi^2}{2}(1 - \cos t)$ for $|t| \leqslant \pi$.

https://doi.org/10.1017/mag.2023.58 Published online by Cambridge University Press

Proof: It is clearly sufficient to prove this for $0 \leqslant t \leqslant \pi$. By concavity of the sine function, $\sin t \geqslant \frac{2}{\pi}t$ for $0 \leqslant t \leqslant \frac{\pi}{2}$, hence $t \leqslant \pi \sin \frac{1}{2}t$ for $0 \leqslant t \leqslant \pi$. So

$$t^2 \leqslant \pi^2 \sin^2 \tfrac{1}{2}t = \tfrac{1}{2}\pi^2 (1 - \cos t).$$

Lemma 5: If $R_n(t) \geqslant 0$ for all t, then

$$\frac{1}{\pi} \int_{-\pi}^{\pi} |t| R_n(t)\, dt \leqslant \frac{\pi}{\sqrt{2}} \sqrt{1 - \rho_{1,n}}.$$

Proof: By the Cauchy-Schwarz inequality for integrals, with (8), we have

$$\frac{1}{\pi} \int_{-\pi}^{\pi} |t| R_n(t)\, dt \leqslant \left(\frac{1}{\pi} \int_{-\pi}^{\pi} t^2 R_n(t)\, dt\right)^{1/2} \left(\frac{1}{\pi} \int_{-\pi}^{\pi} R_n(t)\, dt\right)^{1/2}$$
$$= \left(\frac{1}{\pi} \int_{-\pi}^{\pi} t^2 R_n(t)\, dt\right)^{1/2}.$$

Recall that $\int_{-\pi}^{\pi} \cos^2 t\, dt = \pi$ and $\int_{-\pi}^{\pi} \cos t \cos kt\, dt = 0$ for $k > 1$ and $k = 0$. By this, (8) and Lemma 4,

$$\frac{1}{\pi} \int_{-\pi}^{\pi} t^2 R_n(t)\, dt \leqslant \frac{\pi}{2} \int_{-\pi}^{\pi} (1 - \cos t) R_n(t)\, dt$$
$$= \frac{\pi^2}{2} - \frac{\pi}{4} \int_{-\pi}^{\pi} \cos t\, dt - \frac{\pi}{2} \sum_{k=1}^{n} \int_{-\pi}^{\pi} \rho_{k,n} \cos t \cos kt dt$$
$$= \frac{\pi^2}{2} (1 - \rho_{1,n}).$$

The challenge is to choose the $\rho_{k,n}$ in a way that makes $R_n(t)$ non-negative. One way to ensure this is as follows. For any real c_r $(0 \leqslant r \leqslant n)$,

$$\left(\sum_{r=0}^{n} c_r e^{irt}\right)\left(\sum_{r=0}^{n} c_r e^{-irt}\right) \geqslant 0.$$

This equates to $\displaystyle\sum_{k=0}^{n} \rho_{k,n} \cos kt$, where $\rho_{0,n} = \displaystyle\sum_{r=0}^{n} c_r^2$ and

$$\rho_{k,n} = 2 \sum_{r=0}^{n-k} c_r c_{r+k} \qquad (k \geqslant 1).$$

We require $\rho_{0,n} = \frac{1}{2}$, hence $\displaystyle\sum_{r=0}^{n} c_r^2 = \frac{1}{2}$. We will describe Korovkin's ingenious choice of c_r.

We note in passing that the choice $c_r = c$ for all r, with $(n+1)c^2 = \frac{1}{2}$, gives $\rho_{k,n} = 1 - \dfrac{k}{n+1}$, the Fejér kernel K_{n+1}.

https://doi.org/10.1017/mag.2023.58 Published online by Cambridge University Press

Lemma 6: Write $\frac{\pi}{n+2} = \theta_n$. Let $c_r = c \sin (r+1)\theta_n$ for $0 \leqslant r \leqslant n$, with c chosen so that $\sum_{r=0}^{n} c_r^2 = \frac{1}{2}$. Then $\rho_{1,n} = \cos \theta_n$.

Proof: Note that $\rho_{1,n} = 2 \sum_{r=0}^{n-1} c_r c_{r+1}$. Write $d_r = \sin (r+1)\theta_n$ and $S = \sum_{r=0}^{n-1} d_r d_{r+1}$, so that $c_r = cd_r$ and $\rho_{1,n} = 2c^2 S$. Since $d_{n+1} = 0$, we have also $S = \sum_{r=0}^{n} d_r d_{r+1}$. Also, substituting s for $r+1$ and adding a zero term $d_{-1}d_0$, we have $S = \sum_{s=0}^{n} d_{s-1}d_s$. So

$$2S = \sum_{r=0}^{n} d_r \left(d_{r-1} + d_{r+1}\right).$$

Since $\sin (a+b) + \sin (a-b) = 2 \sin a \cos b$, we have

$$d_{r-1} + d_{r+1} = \sin r\theta_n + \sin (r+2)\theta_n = 2d_r \cos \theta_n,$$

hence

$$S = \cos \theta_n \sum_{r=0}^{n} d_r^2.$$

Since $\sum_{r=0}^{n} c_r^2 = c^2 \sum_{r=0}^{n} d_r^2 = \frac{1}{2}$, we have $\rho_{1,n} = 2c^2 S = \cos \theta_n$.

We are now ready to state Jackson's theorem for Lipschitz functions.

Theorem 4: With R_n defined in this way, if $f \in C_{2\pi}$ and $\mathrm{Lip}(f) = M$, then

$$\left\|(R_n * f) - f\right\| \leqslant \frac{M\pi^2}{2n+4}. \tag{17}$$

Proof: By (9),

$$\left\|(R_n * f) - f\right\| \leqslant \frac{M}{\pi} \int_{-\pi}^{\pi} |t| R_n(t)\, dt.$$

With c_r defined as in Lemma 6, we have

$$1 - \rho_{1,n} = 1 - \cos \frac{\pi}{n+2} = 2 \sin^2 \frac{\pi}{2n+4} \leqslant \frac{2\pi^2}{(2n+4)^2}. \tag{18}$$

The statement follows, by Lemma 5.

We have actually established the bound $M\pi \sin\left(\frac{\pi}{2n+4}\right)$.

https://doi.org/10.1017/mag.2023.58 Published online by Cambridge University Press

The version for continuous functions is as follows.

Theorem 5: For all $f \in C_{2\pi}$, we have

$$\|(R_n * f) - f\| \leqslant C\omega_f\left(\frac{1}{n+2}\right), \tag{19}$$

where $C = \frac{1}{2}\pi^2 + 1$ (so $C < 6$).

Proof: By (9) and Lemma 1, for any $\delta > 0$,

$$\|(R_n * f) - f\| \leqslant \frac{\omega_f(\delta)}{\pi}\int_{-\pi}^{\pi}\left(\frac{|t|}{\delta} + 1\right)R_n(t)\,dt.$$

By Lemma 5 and (18), we have

$$\|(R_n * f) - f\| \leqslant \omega_f(\delta)\left(\frac{\pi^2}{(2n+4)\delta} + 1\right).$$

Taking $\delta = 1/(n+2)$, we obtain (19).

Of course, $\operatorname{dist}(f, T_n) \leqslant \|(R_n * f) - f\|$. As explained in the Introduction, the estimations in Theorems 4 and 5 translate without change to the corresponding problem for ordinary polynomials, thereby achieving our original objective:

Theorem 6: If f is Lipschitz on $[-1,1]$ with $\operatorname{Lip}(f) = M$, then

$$\operatorname{dist}(f, \Pi_n) \leqslant \frac{M\pi^2}{2n+4}. \tag{20}$$

If f is continuous on $[-1, 1]$, then, with C as in Theorem 5,

$$\operatorname{dist}(f, \Pi_n) \leqslant C\omega_f\left(\frac{1}{n+2}\right).$$

While Theorem 6 assures us of the existence of a polynomial approximating to f as stated, of course it does not deliver anything like an actual expression for it (in contrast to interpolating and Bernstein polynomials).

Do stronger conditions on f imply closer approximation by polynomials? We show that this is indeed the case, in two stages, both using Theorem 6.

Theorem 7: If f is differentiable and $n \geqslant 2$, then

$$\operatorname{dist}(f, \Pi_n) \leqslant \frac{\pi^2}{2n+4}\operatorname{dist}(f', \Pi_{n-1}).$$

Proof: Let $\operatorname{dist}(f', \Pi_{n-1}) = \rho$. Take p_{n-1} in Π_{n-1} with $\|f' - p_{n-1}\| = \rho$. Let p_n be an indefinite integral of p_{n-1}: then $p_n \in \Pi_n$ and $p_n' = p_{n-1}$. Then

$\operatorname{Lip}(f - p_n) \leqslant \|f' - p_n'\| = \rho$. So by (20), there exists $q_n \in \Pi_n$ satisfying

$$\|f - p_n - q_n\| \leqslant \frac{\rho\pi^2}{2n + 4}.$$

This proves the result, since $p_n + q_n \in \Pi_n$.

Theorem 8: If f is differentiable and $\operatorname{Lip}(f') \leqslant M_2$ (in particular, if $\|f''\| \leqslant M_2$, then

$$\operatorname{dist}(f, \Pi_n) \leqslant \frac{M_2\pi^4}{4(n + 1)(n + 2)}.$$

Proof: By (20) again,

$$\operatorname{dist}(f', \Pi_{n-1}) \leqslant \frac{M_2\pi^2}{2n + 2}.$$

The statement follows, by Theorem 7.

Of course, the reasoning can be repeated for higher derivatives.

The Dirichlet kernel: convergence of ordinary Fourier series

There is no future in trying to copy Theorem 2 for the Dirichlet kernel because $\int_{-\pi}^{\pi} |tD_n(t)|\, dt$ does not tend to zero as $n \to \infty$. But instead, we can deduce convergence results for D_n from Theorems 4 and 5, as follows. Write $\frac{1}{\pi}\int_{-\pi}^{\pi} |D_n(t)|\, dt = \Delta_n$. Then by (5), $\|D_n * f\| \leqslant \Delta \|f\|$ for all f in $C_{2\pi}$. Since $R_n * f \in T_n$, we have $D_n * (R_n * f) = R_n * f$, so

$$\|(D_n * f) - (R_n * f)\| \leqslant \|D_n * (f - R_n * f)\| \leqslant \Delta \|f - (R_n * f)\|,$$

hence

$$\begin{aligned}\|(D_n * f) - f\| &\leqslant \|(D_n * f) - (R_n * f)\| + \|(R_n * f) - f\| \\ &\leqslant (\Delta_n + 1)\|(R_n * f) - f\|.\end{aligned}$$

The estimation of Δ_n is quite well known and can be seen in various books, e.g. [6, p. 212-213]. So we state a version of it here without proof:

Lemma 7: We have

$$\Delta_n \leqslant \frac{4}{\pi^2}H_n + 3.$$

So we can deduce at once:

Theorem 9: If $f \in C_{2\pi}$ and $\operatorname{Lip}(f) = M$, then

$$\|(D_n * f) - f\| \leqslant \frac{2M}{n + 2}(H_n + \pi^2). \tag{22}$$

https://doi.org/10.1017/mag.2023.58 Published online by Cambridge University Press

In particular, $D_n * f \to f$ uniformly on $\mathbb{R}$.

For all $f \in C_{2\pi}$, with C as in Theorem 5, we have

$$\left\|\left(D_n * f\right) - f\right\| \leqslant 4C\left(\frac{H_n}{\pi^2} + 1\right)\omega_f\left(\frac{1}{n + 2}\right). \tag{23}$$

Explicit statements of the estimations in Theorem 9 are not easy to find in the literature. The traditional proof of uniform convergence for Lipschitz functions uses some form of the Riemann-Lebesgue lemma applied to (4). A version can be seen in [7, Theorem 2.5]. Careful scrutiny shows that this method delivers an estimate of the form $\left\|\left(D_n * f\right) - f\right\| \leqslant \dfrac{C}{n^{1/3}} \operatorname{Lip}(f)$.

Comparing (22) with (15), we see that for Lipschitz functions, the Dirichlet kernel delivers a rate of convergence similar to the one for the Fejér kernel. However, in contrast to the Fejér kernel, uniform convergence does not follow from (23), because $H_n\omega_f(1/(n+2))$ may fail to converge to 0 (note the subtle difference between (23) and (16)). Indeed, it has been known for a long time that there are functions f in $C_{2\pi}$ for which $D_n * f$ does not even converge pointwise to f: see [7, section 2.5] or [8, p. 161-162].

Example 2: Recall from Example 1 that for $f(x) = |x|$, we have $\left\|\left(K_n * f\right) - f\right\| > \dfrac{H_n}{n\pi}$. We show that, despite what we have just been saying, this is a case where $D_n * f$ converges more rapidly than $K_n * f$. By (12) and the original definition (1), we have

$$\left(D_{2n-1} * f\right)(x) = \left(D_{2n} * f\right)(x) = \frac{\pi}{2} - \frac{4}{\pi}\sum_{k=1}^{n}\frac{\cos(2k - 1)x}{(2k - 1)^2}.$$

Since $\left(D_n * f\right)(x)$ converges to $f(x)$, we deduce that

$$\left(D_{2n} * f\right)(x) - f(x) = \frac{4}{\pi}\sum_{k=n+1}^{\infty}\frac{\cos(2k - 1)x}{(2k - 1)^2},$$

so, estimating as in the proof of Lemma 2, we have for all x

$$\left|\left(D_{2n} * f\right)(x) - f(x)\right| \leqslant \frac{4}{\pi}\sum_{k=n+1}^{\infty}\frac{1}{(2k - 1)^2} \leqslant \frac{2}{\pi}J_n \leqslant \frac{1}{n\pi}.$$

Finally, in defence of D_n, it should be mentioned that there are other respects in which the operator $S_n(f) = D_n * f$ is both natural and optimal. First, it can be shown that among all projections of $C_{2\pi}$ onto T_n, it has the distinction of being the one with least norm. Secondly, consider the Hilbert space $L_2\left[-\pi, \pi\right]$, in which the norm is now defined by

$$\|f\|_2 = \left(\frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)^2\,dx\right)^{1/2}.$$

In this space, the functions $\cos kt$ and $\sin kt$ form an orthogonal basis of T_n. By standard Hilbert space theory, S_n is the orthogonal projection onto T_n, and $S_n(f)$ is therefore the best approximation to f in T_n measured by this norm.

References

1. G. J. O. Jameson, How close is the approximation by Bernstein polynomials? *Math. Gaz.* **104** (November 2020) pp. 484-494.
2. D. Jackson, On approximation by trigonometric sums and polynomials, *Trans. Amer. Math. Soc.* **13** (1912) pp. 491-515.
3. P. P. Korovkin, *Linear operators and approximation theory*, Hindustan (1960).
4. G. Meinardus, *Approximation of functions: theory and numerical methods*, Springer (1967).
5. Theodore J. Rivlin, *An introduction to the approximation of functions*, Dover (1969).
6. E. W. Cheney, *Introduction to approximation theory*, McGraw Hill (1966).
7. P. L. Walker, *The theory of fourier series and integrals*, Wiley (1986).
8. R. E. Edwards, *Fourier series: a modern introduction*, vol. 1, Springer (1979).

10.1017/mag.2023.58 © The Authors, 2023
Published by Cambridge University Press on
behalf of The Mathematical Association

G. J. O. JAMESON
13 Sandown Road,
Lancaster LA1 4LN
e-mail: *pgjameson@talktalk.net*

Farewell to 259

What a strange heading! Is 259 vanishing? Is 259 in any way special?

Some brief research tells us that $259 = 7 \times 37$ so it is one of the infinitely many numbers to carry the label of 'semiprime'. A little bit more digging produces $259 = 6^0 + 6^1 + 6^2 + 6^3$ so it is also a bearer of the label 'repdigit' (in base 6).

But 259 has extra significance for the MA, as its address for the last $49\,(= 7^2)$ years has been 259 London Road, Leicester. However the time has come to move. From now on, the MA address will be:

The Mathematical Association.
Charnwood Building, Holywell Park,
Loughborough University Science and Enterprise Park,
Leicestershire LE11 3AQ