Lectures On
Approximation By Polynomials

By

J.G. Burkill

Tata Institute of Fundamental Research,
Bombay
1959

Contents

Chapter 1

Weierstrass's Theorem

1 Approximation by Polynomials

A basic property of a polynomial $P(x) = \sum_{0}^{n} a_r x^r$ is that its value for **1**
a given x can be calculated (e.g. by a machine) in a finite number of steps. A central problem of mathematical analysis is the approximation to more general functions by polynomials an the estimation of how small the discrepancy can be made. A discussion of this problem should be included in any University course of analysis. Not only are the results important, but their proofs admirably illustrate a number of powerful methods.

This account will be confined to the leading theorems, stated in their fundamental rather than their most general forms. There are many excellent systematic presentations in the literature, to which this may serve as an introduction.

Variables and functions will be real. We say that $f(x)$ is $C(a, b)$ meaning that $f(x)$ is continuous for $a \leq x \leq b$.$p(x)$ or $q(x)$ always denotes a polynomial; $p_n(x)$ is a polynomial of degree at most n. In this course, the goodness (or badness!) of the fit of a particular polynomial $p(x)$ to the function $f(x)$ will always be measured by

$$\sup |f(x) - p(x)|,$$

where the sup. is taken over $a \leq x \leq b$.

There are other useful ways of defining a 'distance' between $f(x)$ and $p(x)$, e.g.

$$\int_a^b \{f(x) - p(x)\}^2 dx,$$

but we shall not deal with them here.

2 The interval (a, b) will commonly be taken to be $(0, 1)$ or $(-1, 1)$ as may be convenient in particular context; there will be no loss of generality. Our enquiry is restricted to finite intervals. The numbers ϵ will always be supposed greater than 0. The Halmos symbol /// denotes the end of a proof.

Theorem 1 (Weierstrass 1885). *If $f(x)$ is $C(a, b)$, then, given ε, we can find $p(x)$ such that*

$$\sup |f(x) - p(x)| < \varepsilon.$$

This is the fundamental theorem of the subject. An alternative statement of it is that a continuous function is the sum of a uniformly convergent series of polynomials. For let $p_{n_1}(x), p_{n_2}(x), \cdots (n_1 \leq n_2 \leq \cdots)$ be polynomials corresponding to $\varepsilon, \frac{1}{2}\varepsilon, \ldots, \varepsilon/2^n \ldots$. Then the series

$$p_{n_1}(x) + \{p_{n_2}(x) - p_{n_1}(x)\} + \cdots$$

converges uniformly to $f(x)$.

We shall give three proofs of Weierstrass's theorem. The first and simplest is that of Lebesgue (1898). It is based on a polynomial approximation to the particular function $|x|$ in $(-1, 1)$. We shall study this function closely in Chapter *III*, and shall learn a lot from it.

Lemma. *There is a sequence of polynomials converging uniformly to $|x|$ for $-1 \leq x \leq 1$.*

Proof. If $u = 1 - x^2$. then $|x| = \sqrt{(1 - u)}$, and $0 \leq u \leq 1$ corresponds to $1 \geq |x| \geq 0$. □

3 $\sqrt{(1 - u)}$ has a binomial expansion in which the term in u^n is $-c_n u^n$ where

$$c_n = \frac{1.3.5 \ldots (2n - 3)}{2.4.6 \ldots\ 2n} \qquad (n \geq 2)$$

We can prove that this series, which certainly converges for $|u| < 1$, also converges for $u = 1$. This follows either from Gauss's test applied to

$$\frac{c_n}{c_{n+1}} = 1 + \frac{3}{2n} + 0\left(\frac{1}{n^2}\right)$$

or by proving (on the lines of the Lemma following Theorem 2) that $c_n \sim \frac{A}{n\sqrt{n}}$.

By Abel's limit theorem, the series for $\sqrt{(1-u)}$ converges uniformly for $0 \leq u \leq 1$, i.e., $|x|$ is uniform limit of a sequence of polynomials for $-1 \leq x \leq 1$.

Corollary. *Let*

$$g(x) = 0 \textit{ for } x < 0$$
$$g(x) = 0 \textit{ for } 0 \leq x \leq k.$$

Then $g(x)$ is the limit of a uniformly convergent sequence of polynomials in $-k \leq x \leq k$.

Proof. Changing the variable by a factor k, we may suppose that k is 1. Then

$$g(x) = \frac{1}{2}(x + |x|).$$

□

Proof of theorem 1. Given ε, we can find a function $l(x)$ whose graph is a polygon with vertices at $(a, y_0), (x_1, y_1), \ldots, (x_i, y_i), \ldots, (b, y_n)$ such that

$$|f(x) - l(x)| < \frac{1}{2}\varepsilon.$$

Now $l(x)$ is the sum of constant multiples of functions of the type **4**
$g(x - x_i)$ defined in the Corollary, namely,

$$l(x) = y_0 + \sum_{0}^{n-1} c_i g(x - x_i).$$

For the right hand side is linear in each (x_i, x_{i+1}), and the c_i give the right value of $l(x)$ at the vertices if

$$y_1 = y_0 + c_0(x_1 - x_0)$$

$$\cdots\cdots$$

$$y_i = y_0 + \sum_{k=0}^{i-1} c(x_i - x_k).$$

By the lemma and corollary, we can find a polynomial $p(x)$ such that

$$|l(x) - p(x)| < \frac{1}{2}\varepsilon, \ a \le x \le b$$

and this gives $$|f(x) - p(x)| < \varepsilon, \ a \le x \le b.$$

2 Singular Integrals and Landau's Proof

Weierstrass's own proof of Theorem 1 rested on the limit as $n \to \infty$ of the 'singular integral'

$$\frac{n}{\sqrt{\pi}} \int_{-\infty}^{\infty} \exp\left\{-n^2(t-x)^2\right\} f(t)dt.$$

The essence of the argument is that, if n is large, the exponential 'kernel' is small except in a small interval round $t = x$, and so the integral is nearly equal to $f(x)$. This integral is not, however, a polynomial in x and, to complete the proof, Weierstrass had to approximate to the exponential by the sum of a finite number of terms of its series. A natural step, taken, by Landau and by de la Vallee Poussin, was to start with a singular integral which is a polynomial in x. An appropriate kernel to replace Weierstrass's exponential factor is

$$\{1 - (t-x)^2\}^n$$

5 which (for large n) falls away rapidly from the value 1 as t moves away form x.We need a theorem about the convergence of singular integrals, and this is best stated for a general kernel $K_n(t - x)$.

Theorem 2. *Let*

$$J_n = \int_{-1}^{1} K_n(u)du$$

$$L_n(\delta) = \int_{-\delta}^{\delta} K_n(u)du \qquad (0 < \delta < 1)$$

Suppose that

(i) $K_n(u) \geq 0$

(ii) for each fixed $\delta, L_n(\delta)/J_n \to 1$, *as* $n \to \infty$.

Suppose that $f(x)$ *is* $C(0,1)$ *and* $0 < a < b < 1$. *Then, as* $n \to \infty$,

$$I_n(x) = \frac{1}{J_n} \int_0^1 K_n(t-x) f(t) dt \to f(x)$$

uniformly for $a \leq x \leq b$.

Proof. In $I_n(x)$, we shall split up the integral over $(0,1)$

$$I_n(x) = \frac{1}{J_n} \left\{ \int_0^{x-\delta} + \int_{x-\delta}^{x+\delta} + \int_{x+\delta}^{1} \right\}, \tag{1}$$

where $0 < x-\delta < x+\delta < 1$. Consider first the integral over $(x-\delta, x+\delta)$. Given ε, we can, by the continuity of $f(x)$, find $\delta = \delta(\varepsilon)$ such that

$$|f(t) - f(x)| < \varepsilon \text{ if } a \leq x \leq b, |t-x| \leq \delta.$$

□

Suppose further that $\delta < \min(a, 1-b)$. Then the middle term on the R. H. S. of (1)

$$\begin{aligned} &= \frac{1}{J_n} \int_{-\delta}^{\delta} K_n(u) f(x+u) du \\ &= \frac{L_n(\delta)}{J_n} f(x) + \frac{1}{J_n} \int_{-\delta}^{\delta} K_n(u) \{ f(x+u) - f(x) \} du. \end{aligned}$$

The first term in the last line tends to $f(x)$, from (ii) of the hypothe- **6**
sis. The second term is, by (i), numerically less than $\varepsilon L_n(\delta)/J_n$, that is, less than ε.

Now return to equation (1) and consider the first term on the R. H. S. Let $M = \sup |f(x)| in$ $(0,1)$.

$$\left| \frac{1}{J_n} \int_0^{x-\delta} K_n(t-x) f(t) dt \right| \leq \frac{M}{J_n} \int_{-x}^{-\delta} K_n(u) du$$

$$\leq M\left\{1 - \frac{L_n(\delta)}{J_n}\right\}$$

$\to 0$ as $n \to \infty$.

A similar estimate holds for the third term of (1).

All the inequalities in the above argument are independent of x, and, collecting the results, we have proved that $I_n(x) \to f(x)$ uniformly for $a \leq x \leq b$.

If, in Theorem 2, we take, following Landau

$$K_n(u) = (1 - u^2)^n,$$

then $I_n(x)$ is a polynomial in x of degree $2n$. We have, therefore, a second proof of Theorem 1 as soon as we have proved, as we do in the following Lemma, that this $K_n(u)$ satisfies the conditions of Theorem 2.

Lemma. *In Theorem* 2, $K_n(u)$ *may be taken to be* $(1 - u^2)^n$.

Proof.

$$J_n = \int_{-1}^{1}(1 - u^2)^n du = 2\int^{\frac{1}{2}0\pi} 0\sin^{2n+1}\theta d\theta$$

$$= 2S_{2n+1}, \text{ say.}$$

$$S_{2n+1} = \frac{2.4\ldots 2n}{3.5\ldots(2n+1)}$$

7 From the inequalities

$$S_{2n} > S_{2n+1} > S_{2n+2},$$

it is easily proved that

$$J_n \sim \sqrt{\frac{\pi}{n}} \text{ and } J_n > \sqrt{\frac{\pi}{n+1}}.$$

□

Then

$$1 - \frac{L_n(\delta)}{J_n} = \frac{2\int_\delta^1(1 - u^2)^n du}{J_n}$$

$$< 2(1 - \delta^2)^n\sqrt{\frac{n+1}{\pi}} \to 0 \text{ as } n \to \infty.$$

3 Bernstein Polynomials

We shall give a third proof of Theorem 1. It has the advantage of embodying a definite construction for the approximating polynomials.

Definition. Write $l_{n,m}(x) = \binom{n}{m} x^m (1-x)^{n-m}, 0 \leq m \leq n$. The n^{th} Bernstein polynomials of $f(x)$ in $(0, 1)$ is defined to be

$$B_n(x) = B_n(f; x) = \sum_{m=0}^{n} f(m/n) l_{n,m}(x).$$

$B_n(x)$ has degree n (at most).

Theorem 3. *Let $f(x)$ be $C(0, 1)$. Then, as $n \to \infty, B_n(x) \to f(x)$ uniformly.*

Note. We can see what underlies this. $l_{n,m}(x)$ has a maximum at $x = m/n$. So the terms of $B_n(x)$ for which m/n is near to x are those which contribute most. It is, in fact, the analogue for a finite sum of the 'singular integral' notion. Then two schemes, for sum and integral, could be combined into one by using a Steltjes integral.

Lemmas on $l_{n,m}(x)$.

The sums on the R.H.S. being taken for values of m such that $0 \leq$ **8**
$m \leq n$,

$$1 = \sum l_{n,m}(x)$$
$$nx = \sum m l_{n,m}(x)$$
$$nx(1-x) = \sum (nx - m)^2 l_{n,m}(x).$$

Proof. With a view to differentiating with regard to y, we write

$$\{e^y + (1-x)\}^n = \sum \binom{n}{m} e^{my} (1-x)^{n-m}$$

Put $e^y = x$ and we have the first result. Differentiate with regard to y and put $e^y = x$ and we have second. Differentiating again gives

$$nx + n(n-1)x^2 = \sum m^2 l_{n,m}(x)$$

Multiply the three equations in turn by $n^2x^2, -2x, 1$ and add. This gives the third result in the lemma. □

Proof of theorem 2. Given ε, there is δ such that $|f(x_1) - f(x_2)| < \varepsilon$ if $|x_1 - x_2| < \delta$. Now,

$$f(x) - B_n(x) = \sum_{m=0}^{n} \{f(x) - f(m/n)\} l_{n,m}(x).$$

Divide the sum on the R.H.S. into parts: $\sum_l$ taken over those values of m for which $|x - \frac{m}{n}| < \delta$, and $\sum_2$ the rest. Then $|\sum_1| \leq \varepsilon \sum_1 l_{n,m}(x) \leq \varepsilon \sum_0^n l_{n.m}(x) = \varepsilon$. If M is $\sup|f(x)|$ in $0 \leq x \leq 1$,

$$\begin{aligned} |\sum_2| &\leq 2M \sum_2 l_{n,m}(x) \\ &\leq 2M \sum_2 \frac{(nx-m)^2}{n^2\delta^2} l_{n,m}(x) \\ &\leq 2Mnx(1-x)/n^2\delta^2, \text{ from the Lemma} \\ &\leq M/2n\delta^2 \end{aligned}$$

9 So $|f(x) - B_n(x)| \leq |\sum_1| + |\sum_2| \leq \varepsilon + M/2n\delta^2$. Choose $n > M/2\varepsilon\delta^2$ and the R.H.S. $\leq 2\varepsilon$.

Remarks on Bernstein polynomials.

(1) They have applications to the theory of probability, moment problems and the summation of series. See Lorentz, Bernstein polynomials, (Toronto 1953).

(2) In questions of polynomial approximation, it is a disadvantage that the Bernstein polynomial of a polynomial $p_n(x)$ is not, in general, $p_n(x)$, e.g.

$$\text{for } x^2, B_2(x) \text{ is } \frac{1}{2}x(1+x)$$
$$\text{for } x(1-x), B_2(x) \text{ is } \frac{1}{2}x(1-x).$$

For most of the useful systems of polynomials, the approximation within the system to a given $p_n(x)$ is $p_n(x)$, e.g. with Legendre polynomials,

$$x^2 = \frac{1}{3}P_0(x) + \frac{2}{3}P_2(x).$$

Notes on Chapter I

Notes at the end of a chapter may include exercises (with hints for solutions), extensions of the theorem and suggestions for further reading. **10**

1. In the Lemma of §1, prove that the polynomial consisting of the terms up to x^{2n} in the expansion of $\sqrt{\{1 - (1 - x^2)\}}$ approximates to $|x|$ in $(-1, 1)$ with a greatest error which $\sim A/\sqrt{n}$.

2. Let $f(x) = \frac{1}{2} - |x - \frac{1}{2}|$ in $(0, 1)$. (This is an adaptation of $|x|$ to the interval $(0, 1)$). As in 1, investigate the order of magnitude of the error at $x = \frac{1}{2}$ given by (*a*) the Landau singular integral, (*b*) the Bernstein, approximations to $f(x)$.

3. Theorem 1 can be extended to a function of two (or more) variables, say $f(x, y)$ for $0 \leq x \leq 1, 0 \leq y \leq 1$. Suggest a method of proof.

4. If $f'(x)$ is continuous, then

$$\frac{d}{dx}B_n(f; x) \to f'(x) \text{ uniformly.}$$

   A similar result for the Landau integral.

5. Readers who like to place theorems on analysis in an abstract setting will be interested in Stone's extension of Theorem 1. See Math. Magazine 21 (1948) 167 and 237, or Lorentz, 9, or Rudin, Principles of Mathematical Analysis (New York 1953), 134.

Hints for 1 – 4

11 1. All the c_n are positive $n \geq 2$). Error is greatest when $u = 1$, i.e., $x = 0$, and is $\sum\limits_{n+1}^{\infty} c_r$.

   This $\sim \sum\limits_{n+1}^{\infty} \dfrac{A}{r\sqrt{r}} \sim A \int_n^{\infty} \dfrac{dx}{x\sqrt{x}} \sim A/\sqrt{n}$.

2. $A/\sqrt{n}$. For (b), approximate to factorials by Stirling's formula.

3. Could use

$$\frac{\int_0^1 \int_0^1 \left\{1-(t-x)^2\right\}^n \ \left\{1-(u-y)^2\right\}^n \ f(t,u)dtdu}{\left\{\int_{-1}^1 (1-t^2)^n dt\right\}^2}$$

   or (with some labour extend Bernstein's, as in P. L. Butzer, Canadian Journal of Mathematics, 5(1953), 107, or Lorentz, 51.

4. Lorentz, 26.

   For Landau, with $K_n(u) = (1-u^2)^n$ in Theorem 2,

$$\frac{d}{dx}\int_0^1 K_n(t-x)f(t)dt = \int_0^1 \frac{\partial K_n}{\partial x} f(t)dt = -\int_0^1 \frac{\partial K_n}{\partial_t} f(t)dt$$

   and integrate by parts.

Chapter 2

The Polynomial of Best Approximation Chebyshev Polynomials

4 The Lagrange Polynomial

We are given $n + 1$ values of x, **12**

$$x_0, x_1, \ldots, x_n$$

and $n + 1$ constants $c_0, c_1, \ldots, c_n$.

Write $\prod(x) = (x - x_0) \cdots (x - x_n)$.

The polynomial $p(x)$ of degree at most n which takes the values c_i at x_i is, by the partial- fraction rule for $p(x)/\prod(x)$,

$$\prod(x) \sum_0^n \frac{1}{x - x_i} \frac{c_i}{\prod(x_i)}.$$

If the c_i are the values at x_i of a function $f(x)$, we call $p(x)$ the Lagrange polynomial of $f(x)$ at the x_i. Here we follow the usual terminology, although Waring (1779) used the polynomial before Lagrange (1795) and indeed it is clear that the formula was known to Newton.

Suppose that the values $x_0, \ldots, x_n$ are fixed. The following lemmas follow from the definition of the Lagrange polynomial.

Lemma 1. *Given an aggregate of polynomials $p_\alpha(x)$ of degree at most n, where α runs through an index-set I, such that*

$$|p_\alpha(x_i)| \leq A, \ \alpha \text{ in } I; i = 0, \ldots, n.$$

Then, if $a_{\alpha,r}$ is the coefficient of x^r in $p_\alpha(x)$,

$$|a_{\alpha,r}| \leq AB,$$

where B is independent of α.

Proof. Write $p_\alpha(x_i)$ for c_i. We have a B depending only on the x_i. □

13 **Lemma 2.** *(For brevity of expression, we translate de la Vallee Poussin, Leçns, 74). If, at $n+1$ given points, two polynomials of degree at most n take 'infinitely close' values, their corresponding coefficients are infinitely close.*

Proof. Given ε, we have two polynomials say $p_\alpha(x)$, $q_\alpha(x)$ which differ by at most ε for each of the values $x_0, \ldots, x_n$. By Lemma 1, their corresponding coefficients differ by at most $B\varepsilon$. □

5 Best Approximation

Let P_n be the set of polynomials $p(x)$ of degree less than or equal to n. Then

$$P_0 \subset P_1 \subset P_2 \cdots$$

Define, for any particular p in P_n,

$$d(p, f) = \sup |f(x) - p(x)| \ for \ a \leq x \leq b.$$

Let $d = d_n = d_n(f) = \inf d(p, f)$ for all p in P_n. Then $d \geq 0$. Our first aim is to prove that there exists a p in P_n for which the inf. is attained, i.e., that, given $f(x)$ of $C(a, b)$, there is a *polynomial of degree n of best approximation.* Later we shall prove uniqueness.

If f is given,

$$d_0 \geq d_1 \geq d_2 \ldots,$$

and Theorem 1 asserts that $\lim d_n = 0$.

The existence of a polynomial of best approximation was known to Chebyshev (or Tschebyscheff) (1821 - 1894) who was one of the founders of the subject. The necessary proof was supplied by Borel (1905).

Theorem 4. *There is a polynomial* $p(x)$ *in* P_n *for which* **14**

$$\sup |f(x) - p(x)| = d(= d_n).$$

Proof. All our polynomials being P_n, we do not need the suffix n to denote degree, and the suffixes in $p_1, p_2, \ldots$ will be used to specify particular polynomials of P_n. As here, we shall commonly omit the variable x form a polynomial $p(x)$ or a function $f(x)$. □

By definition of d, there is a polynomial p_m with

$$d \leq d(p_m, f) < d + \frac{1}{m}.$$

For all m and $a \leq x \leq b$,

$$|p_m(x)| \leq d + 1 + \sup |f(x)| = A.$$

By §4, Lemma 1, the $n + 1$ coefficients of powers $x^0, x^1, \ldots, x^n$ in the $p_m(x)$ all lie in a bounded region of $n + 1$ space. This set of points in $n + 1$ space has at least one limit point, defining a polynomial $p(x)$ for which

$$\begin{aligned} d(f, p) &\leq d(f, p_m) + d(p_m, p) \\ &\leq d + \frac{1}{m} + \varepsilon \end{aligned}$$

where $\varepsilon \to 0$ as $m \to \infty$ through a sub-sequence for which there is convergence of the coefficients to their limits.

Therefore $d(f, p) = d$.

Theorem 5. *If* $f(x)$ *is* $C(a, b)$ *and* $p(x)$ *satisfies Theorem 4 there are* $n + 2$ *values (or more) at which*

$$f(x) - p(x) = \pm d,$$

with alternating sign.

Proof. $g(x) = f(x) - p(x)$ is continuous. Divide (a, b) into sub-intervals 15
such that $g(x)$ does not take the value 0 in any (closed) sub-interval in which it takes the value $\pm d$. Denote by $l_1, l_2, \ldots, l_m$ (numbered from left to right) those of the sub-intervals in which $g(x)$ takes the value $+d$ or $-d$. Define $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_m$ to be $+1$ or -1 according as the value is $+d$ or $-d$. We have to prove that there are at least are at least $n + 1$ changes of sign in the sequence of $\varepsilon' s$. Suppose there are fewer. We shall obtain a contradiction by constructing a polynomial of better approximation than $p(x)$. □

If all the $\varepsilon' s$ have the same sign, say $+$, add a small constant to $p(x)$. This gives a polynomial of better approximation.

Generally, suppose that there are k changes of sign in the sequence of $\varepsilon' s$, where $k \leq n$. Let $\varepsilon_i, \varepsilon_{i+1}$ be different. Then l_i and l_{i+1} cannot abut (since $g(x)$ does not vanish in either), so we can choose a value of x; lying between them. We have thus k values of x; call them

$$x_1, x_2, \ldots, x_k.$$

Define $h(x) = \varepsilon_1(x_1 - x)(x_2 - x) \cdots (x_k - x).h(x)$ has the same sign as $g(x)$ in each of sub-intervals l. We shall prove that, if η is small enough, the polynomial of P_n

$$p(x) + \eta h(x)$$

has better approximation to $f(x)$ than $p(x)$ has.

In those intervals of the original subdivision which are not $l' s$,

$$\sup |g(x)| = d' \text{ (say) } < d.$$

16 Choose η to make $|\eta h(x)| < d - d' (a \leq x \leq b)$, now,

$$|f - p - \eta h| = |g - \eta h|.$$

In the $l' s$, this is less than d, since g, h have the same sign.

And, in the sub-intervals other than $l' s$,

$$|g - \eta h| \leq |g| + |\eta h| < d' + (d - d') = d.$$

So $p + \eta h$ approximates better to f than p does.

Theorem 6. *The polynomial $p(x)$ of Theorem 4 is unique.*

Proof. Suppose that two polynomials p, q satisfy Theorem 4. Let $r = \frac{1}{2}(p+q)$. Then $f - r = \frac{1}{2}(f-p) + \frac{1}{2}(f-q)$. Therefore r satisfies Theorem 4, and so, by Theorem 5,

$$f - r = \pm d$$

for $n+2$ values of x. □

But $f - r = d$ only if $f - p = f - q = d$. Therefore there are $n+2$ values of x for which $p(x)$ and $q(x)$, polynomials of degree at most n, are equal. Therefore $p(x) \equiv q(x)$.

In future we can (by Theorem 6) describe *as the best* P_n that polynomial $p(x)$ in P_n for which

$$\sup |f(x) - p(x)| = d,$$

where $d = \inf d(f, q)$ for all $q(x)$ in P_n. The number d (or d_n if it is necessary to make then n explicit) may be called *the best approximation*.

Theorem 7. *Suppose f is $C(a, b)$ and q is in P_n. Let there be $n+2$ values of x at which $f - q$ takes values alternating in sign*

$$d_1, -d_2, d_3, \ldots, (-1)^{n+1} d_{n+2}.$$

Then the best approximation d satisfies **17**

$$d \geq \min d_i$$

Proof. Suppose that $d < d_i (i = 1, \ldots, n+2)$ and let p be the best P_n. Then $p - q = (f - q) - (f - p)$ takes alternate signs at the $n+2$ values in the hypothesis. Therefore $p - q$ (which is in P_n) has at least $n+1$ zeros. This is a contradiction. □

Corollary. *Let q be in P_n and let*

$$\sup |f - q| = d'.$$

Suppose that $f - q$ takes the values $\pm d'$ alternately for $n+2$ values of x. Then $d' = d$ and q is the best P_n.

Proof. By theorem 6, $d \geq d'$. But $d \leq d'$, since d is the best approximation. □

6 Chebyshev polynomials

Theorem 4 guarantees the existence of the best P_n for a given f. It is only in special cases that the explicit calculation of this polynomial is practicable. Theorem 7 and its corollary can often be turned to use.

Easy exercises:

For x^2 in (0, 1), the best P_0 is $\frac{1}{2}$, the best P_1 is $x + \frac{1}{8}$.

For x^4 in $(-1, 1)$, the best P_3 is $x^2 + \frac{1}{8}$.

Consider now the general problem.

A. Among all $p_n(x)$ with coefficient of x^n equal to 1, find that which deviates least from 0 in $(-1, 1)$ in other words, that for which $\sup|p_n(x)| = d$ is least.

This problem can be stated in the equivalent form.

B. Find the best approximation in P_{n-1} to x^n in $(-1, 1)$.

18 From Theorem 7 (Corollary) we wish to find a $p_n(x)$ which takes the values $\pm d$ alternately at $n+1$ points (why not $n+2$ points?). Enlightened guessing soon leads to the answer

$$p_n(x) = d \cos n\theta \text{ where } x = \cos\theta.$$

It is worth while to give Chebyshev's own proof of this, which does not depend on guesswork.

Theorem 8. *Among all $p_n(x)$ with coefficient of x^n equal to* 1*, the polynomial*

$$2^{-n+1} \cos n\theta \textit{ where } x = \cos\theta$$

deviates least from 0 *in* $(-1, 1)$.

Proof. let $p(x) = x^n + \cdots$ be the required polynomial, and $d = \sup|p(x)|$. □

By Theorem 5, there are $n+1$ values of x (at least) where $p(x) = \pm d$. These may be end-points or interior points of $(-1, 1)$. At such a point which is an interior point, $p(x)$ has a maximum or minimum and so $p'(x) = 0$. Since $p'(x)$ has degree $n - 1$, the $n + 1$ values must be $1, -1$ and $n - 1$ others, say $x_1, x_2, \dots, x_{n-1}$.

The two polynomials of degree $2n$

$$d^2 - p^2 \text{ and } (1 - x^2)p'^2$$

have the same zeros, namely, $1, -1$ and each of $x_1, x_2, \dots, x_{n-1}$ doubly. Comparing the coefficients of x^{2n} we have

$$n^2(d^2 - p^2) = (1 - x^2)p'^2.$$

Solving this differential equation for p we find, putting $x = \cos\theta$,

$$p = d\cos(n\theta + C)$$

Since $p(x)$ is a polynomial, $C = 0$. But

$$\cos n\theta = 2^{n-1}\cos^n\theta + \text{lower powers of } \cos\theta,$$

and so $$= 2^{-n+1}.$$

The polynomials revealed by Theorem 8 are named after Chebyshev **19**
and (following the alternative spelling of his name) we define

$$T_n(x) = \cos(n \text{ arc } \cos x).$$

The early members of the sequence are

$$T_0(x) = 1, \qquad T_1(x) = x \qquad T_2(x) = 2x^2 - 1$$
$$T_3(x) = 4x^3 - 3x, \qquad T_4(x) = 8x^4 - 8x^2 + 1.$$

Their mode of definition is restricted to $(-1, 1)$ and it is in that interval that their utility mainly lies. But many of their properties hold for all values of x. Some useful results are collected in Theorem 9; the proofs can easily be supplied.

Theorem 9. *(1)* $y = T_n(x)$ *satisfies the differential equation*

$$(1 - x^2)y'' - xy' + n^2 y = 0.$$

(2) *$T_n(x)$ is the coefficient of t^n in the expansion of the generating function*

$$\frac{1 - tx}{1 - 2tx + t^2}$$

(3) *the recurrence relation*

$$T_n(x) = 2xT_{n-1}(x) - T_{n-2}(x) \qquad (n \geq 2).$$

(4) *an explicit formula for the coefficients*

$$T_n(x) = \sum (-1)^k \frac{n}{n-k} \binom{n-k}{n} 2^{n-2k-1} x^{n-2k}$$

summed for $0 \leq k \leq \left[\frac{n}{2}\right]$

(5) *orthogonality with the weight-function* $1/\sqrt{(1 - x^2)}$

$$\int_{-1}^{1} \frac{T_n(x)T_m(x)}{\sqrt{(1 - x^2)}} dx = \begin{cases} 0(m \neq n) \\ \frac{1}{2}\pi(m = n) \end{cases}$$

(6) *for* $|x| > 1$,

$$2T_n(x) = \{x + \sqrt{(x^2 - 1)}\}^n + \{x + \sqrt{(x^2 - 1)}\}^n.$$

Note

20 The calculation of polynomials of best approximation is in practice troublesome. See de la Vallee Poussin, Chapter VI. For a method of calculation by a convergent sequence, see Polya, Comptes Rendus (Paris) 157 (1913), 840.

Chapter 3

Approximations to $|x|$

7

We now take up the central problem of polynomial approximation, namely 21

Given a function $f(x)$, how high is the degree of the polynomial necessary to approach it with an assigned accuracy?

The answer may well depend on structural properties of $f(x)$. For instance, we may guess (rightly) that we can predict a lower degree if $f(x)$ is assumed to be differentiable instead of only continuous. The best theorems on these matter lie fairly deep. We shall go through some heuristic motion of finding from particular cases what truth appears to be and then deciding how to try to establish it.

A useful function to study with care is $|x|$ in $(-1, 1)$. This function was the basis of Lebesgue's proof of Theorem 1.

From Exercises 1, 2 at the end of Chapter *I*, the deviation from $|x|$ of a polynomial of degree n of any of the three types used in proving Theorem 1 is of order $1/\sqrt{n}$.

Let us clarify our mode of speech. If, for some $p(x)$ in P_n,

$$|f(x) - p(x)| = 0\{\varphi(n)\}$$

we will say that the approximation is $0\{\varphi(n)\}$. If, moreover, there is no $p(x)$ in P_n for which

$$|f(x) - p(x)| = 0\{\varphi(n)\}$$

we will say that the approximation is *actually* $0\{\varphi(n)\}$.

Study of the proofs of Theorem 1 might lead us to conjecture that the best approximation in p_n to $|x|$ is actually $0(1/\sqrt{n})$.

22 We proceed to show that, in fact, it is actually $0(1/n)$. This will be proved, following Bernstein, Lecons Ch.I, by elementary (though rather lengthy) algebra.

To approximation to $|x|$ in $(-1, 1)$ is the same thing as to approximate to x in $(0, 1)$ by polynomials whose exponents are all even, and this is what we shall do.

If d_{2n} is the best approximation to x in $(0, 1)$ by

$$a_0 + a_1 x^2 + \cdots + a_n x^{2n}$$

we shall prove that

$$\frac{1}{2n+1} > d_{2n} > \frac{1}{4(1+\sqrt{2})} \cdot \frac{1}{2n-1}.$$

Bernstein went further and proved that $d_{2n} \sim C/n$, where C is a constant which he evaluated as 0.282 ± 0.004.

The theorems of this Chapter will not be used later in the course, and any one who wishes may note above inequalities for d_{2n} and pass on to Chapter IV.

8 Oscillating polynomials

Definition. If $0 \leq \alpha_0 < \alpha_1 \cdots < \alpha_n$ and $A_i \neq 0$ (all i), we say that

$$p(x) = A_0 x^{\alpha_0} + A_1 x^{\alpha_1} + \cdots + A_n x^{\alpha_n}$$

is an *oscillating polynomial* in $(0, 1)$ if $\sup |p(x)|$ is attained for $n + 1$ values of x in $0 \leq x \leq 1$. We shall suppose the $\alpha' s$ integers.

Illustrations.

(1) $\alpha_i = 2i + 1$ $T_{2n+1}(x)$ satisfies

(2) $\alpha_i = i$ $T_{2n}(\sqrt{x})$.

23 **Lemma 1.** *The polynomial $p(x)$ in the definition has at most n positive zeros. If it has n the coefficients alternate in sign.*

From Descrates' rule fo signs.

Lemma 2. *The coefficients of an oscillating polynomial alternate in sign.*

Proof. (1) Let $\alpha_0 = 0$ (and $A_0 \neq 0$). There are at most $n-1$ changes of sign in the coefficients of $p'(x)$. Therefore $p'(x)$ has at most $n-1$ positive zeros. The $n+1$ values of x at which $\sup|p(x)|$ is attained must be $n-1$ zeros of $p'(x)$ and $x = 0, x = 1$. So $p'(x)$ has $n-1$ zeros, say $x_1, x_2, \ldots, x_{n-1}$ lying inside $(0, 1)$ and its coefficients $A_1, \ldots, A_{n-1}$ alternate in sign.

$p(x)$ has no maxima or minima other than these $n-1$. Therefore

$$p(0), p(x_1), \ldots, p(x_{n-1}), p(1)$$

alternate in sign. Therefore $p(x)$ has n zeros. Therefore $A_0, A_1, \ldots, A_n$ alternate in sign

(2) Let $\alpha_0 > 0$. Then $p(0) = 0$. So $\sup|p(x)|$ is attained at n points inside $(0, 1)$ which are roots of $p'(x) = 0$. Therefore the coefficients alternate in sign.

□

Corollary. *$p(x)$ takes the values $\pm \sup|p(x)|$ with $+$ and $-$ sign alternately.*

Theorem 10. *$p(x) = \sum_{i=0}^{n} A_i x^{\alpha_i}$ is an oscillating polynomial in $(0, 1)$.$q(x)$ is another polynomial $\sum B_i x^{\alpha_i}$ with the same exponents. One coefficient of p is the same as the corresponding one of q (say $A_j = B_j$), where $\alpha_j > 0$. Then*

$$\sup|q| > \sup|p|.$$

Proof. If not, $p - q$ takes alternate signs (may be 0) for the $n+1$ values 24
of x for which p takes its numerically greatest value. Therefore $p - q$ has at least n zeros in $0 \leq x \leq 1$. But, since $A_j = B_j$ it has only n terms,

and so at most $n-1$ changes of sign in its coefficients and so (by Lemma 1) at most $n-1$ positive zeros. This is a contradiction. □

Converse of Theorem 10. $p(x)$ and $q(x)$ are two polynomials with the same exponents and one coefficient the same ($A_j = B_j$, where $\alpha_j > 0$). If

$$\sup |p| < \sup |q|$$

for every such q, then p is an oscillating polynomial.

Proof. We gives the gist of the proof, without setting out all the detail in full. It uses a 'deformation' argument like that of theorem 5. □

Suppose that $p(x)$ is not an oscillating polynomial. Then $p(x)$ takes the values $\pm M$, where $M = \sup |p(x)|$, at h points, say $x_k (k = 1, \ldots, h)$, where $h < n + 1$. We can construct a polynomial $r(x) = \sum C_i x^{\alpha_i}$ with $C_j = 0$ and $r(x_k) = p(x_k)$. (The $n + 1$ coefficients C_i have to satisfy at most $n + 1$ equations; the determinant can be proved $\neq 0$).

We can take ε and intervals round the x_k, outside which $|p(x)| < M - \varepsilon$ and inside each of which $p(x)$ and $r(x)$ have the same sign.

$$\text{Choose } \lambda \text{ to make } \lambda |r(x)| < \varepsilon \text{ for } 0 \leq x \leq 1.$$
$$\text{Then} \qquad \sup |p - \lambda r| < \sup |p|.$$

But $p - \lambda r$ satisfies the conditions for a q, giving a contradiction.

25 Apply theorem 10, taking $p(x)$ to be a constant multiple of one of the oscillating polynomials $T_{2n}(\sqrt{x})$ and $T_{2n+1}(x)$. We obtain

Corollary 1. *If* $q(x) = a_0 + a_1 x + \cdots a_n x^n$ *and* $M = \sup |q(x)|$ *in* $(0, 1)$, *then*

$$|a_i| \leq M |t_i| (i = 0, 1, \ldots, n),$$

26 *where* t_i *is the coefficient of* x^i *in* $T_{2n}(\sqrt{x})$.

Corollary 2. *If* $q(x) = a_0 x + a_1 x^3 + \cdots a_n x^{2n+1}$ *and* $M = \sup |q(x)|$ *in* $(0, 1)$, *then*

$$|a_i| \leq M |t_i| (i = 0, \ldots, n)$$

where t_i *is the coefficient of* x^{2i+1} *in* $T_{2n+1}(x)$.

Theorem 11. *To a given set of exponents there corresponds an oscillating polynomial in* $(0, 1)$*, which is unique except for a constant factor.*

Proof. Let $\alpha_0, \alpha_1, \ldots, \alpha_n$ be the given exponents in ascending order. Suppose that the coefficient of x^{α_k} is given to be K. ◻

We need to prove that among all the polynomials with the given exponents

$$q(x) = B_0 x^{\alpha_0} + \cdots + B_{k-1} x^{\alpha_{k-1}} + k x^{\alpha_k} + \cdots + B_n x^{\alpha_n},$$

there is a unique $q(x)$ for which $\sup\limits_{0 \le x \le 1} |q(x)|$ attains its lower bound.

Clearly $\sup |q(x)|$ is a continuous function of the n variables $(B_0 \cdots, B_{k-1}, B_{k+1}, \ldots, B_n)$. Its lower bound is greater than 0 by Corollary 1 of Theorem 10. It is less than or equal to K, as is seen by taking the $B's$ to be small. Again by Corollary 1, we need only consider values of B_i for which

$$|B_i| \le K|t_i| \quad (i = 0, 1, \ldots, k-1, k+1, \ldots, n),$$

where t_i is the coefficient of x^{α_i} in $T_{2\alpha_n}(\sqrt{x})$.

The B_i lie in a bounded closed region of n space, and so they have at least one set of values for which $\sup |q(x)|$ attains its lower bound. This proves the existence of an oscillating polynomial. Uniqueness follows from Theorem 10.

Theorem 12. *If*

$$p(x) = x^{\alpha_0} + A_1 x_{\beta_1}^{\alpha_1} + \cdots + A_n x_{\beta_n}^{\alpha_n}$$

and

$$q(x) = x^{\alpha_0} + B_1 x^{\beta_1} + \cdots + B_n x^{\beta_n}$$

are both oscillating polynomials in $(0, 1)$ *where*

$$0 < \alpha_0 < \beta_1 < \alpha_1 < \beta_2 < \cdots < \beta_n < \alpha_n,$$

then $\quad \sup |p(x)| > \sup |q(x)|.$

Proof. By Lemma 2, the coefficients fo of $p(x)$ alternate in sign and so do those of $q(x)$.

$$q(x) - p(x) = B_1 x^{\beta_1} - A_1 x^{\alpha_1} + B_2 x^{\beta_2} - \cdots - A x_n^{\alpha_n}$$

has n variation of sign, and so the equation

$$q(x) - p(x) = 0$$

has at most n positive roots. □

Suppose the theorem false and

$$\sup |p(x)| \leq \sup |q(x)|.$$

Then $q(x) - p(x)$ has the sign of $q(x)$ (it may be 0) for the values $x_n (k = 1, \ldots, n+1)$ at which $|q(x)|$ takes its maximum value. Therefore $q(x) - p(x)$ vanishes for n values $\xi_1, \ldots, \xi_n$ such that

$$x_i \leq \xi_1 \leq x_2 \leq \xi_2 \leq \cdots \leq \xi_n \leq x_{n+1}.$$

27 Moreover, there are $n + 1 x's$ and only $n\xi's$ so at least one x, say x_i must satisfy $\xi_{i-1} < x_i < \xi_i$ (giving meaning to ξ_0, ξ_{n+1}).

We shall now compute the sign of $q(x_i)$ by two different methods and obtain contradictory results.

Firstly, in $(0, \xi_1), q(x) - p(x)$ has the sign of its dominant term $B_1 x^{\beta_1}$, which is negative. By following the changes of sign along the sequence, $q(x) - p(x)$ has sign $(-1)^i$ in (ξ_{i-1}, ξ_i). At $x_i, q(x) - p(x)$ and also $q(x)$ have the sign $(-1)^i$.

Secondly, for small values of $x, q(x)$ had the sign of its first term, which is positive. Therefore $q(x_1) > 0$. So $q(x_2) < 0$, and generally, $q(x_i)$ has the sign $(-1)^{i+1}$.

This is a contradiction.

The same arguments can be used to prove

Theorem 12 (Extension). *If*

$$p(x) = A_0 x^{\alpha_0} + \cdots + A_{i-1} x^{\alpha_{i-1}} + x^m + A_{i+1} x^{\alpha_{i+1}} + \cdots + A_n x^{\alpha_n}$$

$$q(x) = B_0 x^{\beta_0} + \cdots + B_{i-1} x^{\beta_{i-1}} + x^m + B_{i+1} x^{\beta_{i+1}} + \cdots + B_n x^{\beta_n}$$

are both oscillating polynomials in $(0, 1)$*, where*

$$0 \leq \alpha_0 < \beta_0 < \cdots < \alpha_{i-1} < \beta_{i-1} < m < \beta_{i+1} < \alpha_{i+1} \cdots < \beta_n < \alpha_n,$$

then

$$\sup |p(x)| > \sup |q(x)|.$$

9 Approximation to $|x|$

Theorem 13. *If*

$$p(x) = x + a_1 x^2 + a_2 x^4 + \cdots + a_n x^{2n}$$

is an oscillating polynomial in $(0, 1)$*, then*

$$\frac{1}{2n+1} > \sup |p(x)| > \frac{1}{2(1+\sqrt{2})(2n-1)} (n > 1).$$

Note. If $n = 1$, the second inequality is to be replaced by equality. The oscillating polynomial is $x - \left(\frac{1}{2} + \frac{1}{\sqrt{2}}\right) x^2$. 28

Proof. Take $n > 1$. By Theorem 12, $\sup |p(x)|$ is less than the supremum of the oscillating polynomial

$$x + b_1 x^3 + \cdots$$

with exponents $1, 3, 5, \ldots 2n + 1$. But that polynomial is $(-1)^n T_{2n+1}(x)/(2n + 1)$. This gives the first inequality of the theorem. □

By Theorems 12 and 10, the oscillating polynomial $x + b_1 x^3 + \cdots + b_{n-1} x^{2n-1}$ has smaller maximum modulus than the polynomial $x + c_2 x^4 + \cdots + c_r x^{2n}$ with exponents $1, 4, 6, \ldots, 2n$. But the former polynomial is $(-1)^{n-1} T_{2n-1}(x)/(2n - 1)$, with maximum modules $1/(2n - 1)$. We shall now construct a polynomial of the latter form (with no term in x^2).

With the notation of the hypothesis for $p(x)$, write $\sup |p(x)| = m$.

Then, if $\mu > 0$,

$$\left|\frac{x}{1+\mu} + a_1\left(\frac{x}{1+\mu}\right)^2 \cdots + a_n\left(\frac{x}{1+\mu}\right)^{2n}\right| \leq m.$$

Therefore

$$|x(1+\mu) + a_1x^2 + a_2'x^4 + \cdots + a_n'x^{2n}| \leq m(1+\mu)^2$$

i.e.,

$$|p(x) + \mu(x + c_2x^4 + \cdots + c_nx^{2n})| \leq m(1+\mu)^2.$$

Therefore

$$|\mu(x + c_2x^4 + \cdots + c_nx^{2n})| \leq m\left\{(1+\mu)^2 + 1\right\}$$

and so

$$|x + c_2x^4 + \cdots + c_nx^{2n}| \leq m\left\{(1+\mu)^2 + 1\right\}/\mu$$

29 This is true for all positive values of μ, and so we can replace the right-hand side by its minimum, which is $2m(1+\sqrt{2})$.

As we said, the maximum modulus of a polynomial with exponents $1, 4, 6, \ldots, 2n$ is greater than $1/(2n-1)$ and therefore

$$m > \frac{1}{2(1+\sqrt{2})} . \frac{1}{2n-1}$$

We have now all the material for the final result announced at the end of §7.

Theorem 14. *If d_{2n} is the best approximation to x in $(0, 1)$ by*

$$a_0 + a_1x^2 + \cdots + a_nx^{2n},$$

then

$$\frac{1}{2n+1} > d_{2n} > \frac{1}{4(1+\sqrt{2})} \cdot \frac{1}{2n-1}.$$

Proof. d_{2n} is the maximum modulus of the oscillating polynomial

$$A_0 + x + A_1x^2 + A_2x^4 + \cdots + A_nx^{2n}.$$

Let $p(x)$ and m have the same meanings as in Theorem 13. □

By Theorem 10, $d_{2n} < m$.

Write $q(x) - A_0 = x + A_1 x^2 + \cdots + A_n x^{2n}$.

So, by Theorem 10,

$$\sup |q(x) - A_0|$$

is greater than the maximum modulus of $p(x)$, the oscillating polynomial with exponents $1, 2, 4, \ldots, 2n$ and coefficient of x equal to 1; that is to say, is greater than m.

But $\sup |q(x) - A_0| \leq d_{2n} + |A_0| \leq 2d_{2n}$.

Therefore $2d_{2n} > m$.

The inequalities of Theorem 13 for m give the started inequalities for d_{2n}.

Notes

1. Example. Find the polynomial in P_n for which the coefficient of x^k is 1 and which deviates least from 0. **30**

2. The definition on page 22 of an oscillating polynomial can be extended to a system

$$A_0\varphi_0 + \cdots + A_n\varphi_n(x),$$

if the $\varphi' s$ satisfy certain conditions. See Bernstein, Lecons, 1 or Aschieser, 67

Hint

1. If k, n are both even or both odd, consider $T_n(x)$, otherwise $T_{2n}(\sqrt{x})$.

Chapter 4

Trigonometric Polynomials

10 Trigonometric polynomials. Modulus of Continuity

The central problem of approximation, namely the degree of the polynomial required an assigned closeness to a given function, yields more easily to trigonometric than to algebraic treatment. Trigonometric series and in particular Fourier series have been in the fore-front of Analysis for something like a century, and knowledge about them has been available for any problem of approximation. 31

A trigonometric polynomials is

$t(x) = \frac{1}{2}a_0 + (a_1 \cos x + b_1 \sin x) + \cdots + (a_n \cos nx + b_n \sin nx)$. This can be written $t_n(x)$ if $a_n \neq 0$ or $b_n \neq 0$ and we wish to display the *order* of the polynomial. We can denote by T_n the set of all polynomials which are sums of multiples of $\cos kx$ and $\sin kx$ for $<\leq k \leq n$. (There will be no confusion with the Chebyshev polynomials $T_n(x)$ of §6).

The function $t(x)$ is periodic with period 2π (and, in general, with no smaller period). We say that $f(x)$ is $C(2\pi)$ if it is continuous with period 2π.

The problem of approximating to a $C(2\pi)$ function by a trigonometric polynomial is essentially the same as that of approximating to a $C(a, b)$ function by an algebraic polynomials. In the first place, the analogue of Theorem 1 holds.

Theorem 15 (Weierstrass). *If $f(x)$ is $C(2\pi)$ then, given ε, there is $t(x)$ such that*

$$|f(x) - t(x)| < \varepsilon \quad (\textit{all } x)$$

This will emerge as a by-product of theorem 18, and we shall give an independent proof here. You should, however, read Notes 1 − 3 *at the end of this chapter.*

32 In statements about periodic functions, values of x differing by multiple of 2π will be regarded as the same.

Lemma 1. *The equation $t_n(x) = 0$ has at most $2n$ roots.*

(Prove by expressing in term of $\tan\frac{1}{2}x$ or of $\exp ix$).

Corollary 1. *Two t_n's which take the same values at $2n + 1$ points are identical.*

Corollary 2. *If two t_n's have $2n$ common zeros one is a consult multiple of the order*

The reader should verify that there is an analogue of the Lagrange polynomial of §4, namely

The polynomial in T_n which takes the values c_i at $x_i (i = 0, 1, \ldots, 2n)$ is

$$P(x) \sum \frac{1}{2\sin\frac{x-xi}{2}} \frac{c_i}{P'(x_i)}$$

$$\text{where} \quad P(x) = \prod \sin\frac{x - xi}{2}.$$

We shall take for granted the trigonometric analogues of Theorem 4 − 7 (pages 14 − 17) about best approximation. Briefly, for a given $f(x)$ in $C(2\pi)$, there is unique $t(x)$ of best approximation in T_n which is characterized by $f(x) - t(x)$ taking its greatest numerical value, with alternating sign, for alt least $2n + 2$ values of x. Proofs can be found in the book of de la Vallee Poussin or Natanson.

Illustrations:

1) If $f(x) = t_{n-1}(x) + (a_n \cos nx + b_n \sin nx)$. then $t_{n-1}(x)$ gives the best approximation in T_{n-1} to $f(x)$.

Proof. $f - t_{n-1}$ takes the values $\pm\sqrt{(a_n^2 + b_n^2)}$ alternately at $2n$ points. □

2) An interesting example is Weiertrass's non-differentiable function **33**

$$f(x) = \sum_{r=0}^{\infty} a^r \cos b^r x$$

where $0 < a < 1, b$ is an odd integer and $ab > 1$. We shall prove that the best approximation in T_n to $f(x)$ is

$$t(x) = \sum_{r=0}^{k} a^r \cos b^r x, \text{ where } b^k \leq n < b^{k+1}.$$

Proof. $f(x) - t(x) = \sum_{k+1}^{\infty} a^r \cos b^r x$. □

This takes its greatest value $\sum\limits_{k+1}^{\infty} a^r$ at $x = 0$. Cos $b^{k+1}x$ takes the values ± 1 alternately at integral multiples of π/b^{k+1}, of which there are $2b^{k+1}$ in a period.

Since b is an odd integer, $\cos b^r x$ for $r > k + 1$ takes the same values at those points as $\cos b^{k+1} x$.

Now $2b^{k+1} \geq 2n + 2$ and so $f(x) - t(x)$ takes its numerically greatest value for at least $2n + 2$ values of$x$.

Corollary. *The approximation given by this $t(x)$ is A/n^α, where α = $\log(1/a)/\log b$.*

Proof. The approximation is

$$\frac{a^{k+1}}{1-a} = \frac{b^{-\alpha(k+1)}}{1-a} \sim \frac{1}{1-a} \cdot \frac{1}{n^\alpha}$$

□

Modulus of continuity. Let $f(x)$ be $C(a, b)$ and define

$$\omega(\delta) = \sup |f(x_2) - f(x_1)| \text{ for } |x_2 - x_1| \leq \delta.$$

Then $\omega(\delta)$ is continuous, increases as δ increases, and tends to 0 as δ tends to 0. We shall find that the rapidity with which $\omega(1/n)$ tends to 0 as $n \to \infty$ gives the clue to the approximation to $f(x)$ attainable in P_n or T_n.

34 If $f(x)$ is $C(2\pi)$, the same definition of $\omega(\delta)$ holds. Observe that now the greatest value of$\omega(\delta)$ is $\omega(\pi)$

Properties of $\omega(\delta)$ are collected in the following theorem.

Theorem 16. *(1) If n is an integer,*

$$\omega(n\delta) \leq n\omega(\delta).$$

(2) If $k > 0, \omega(k\delta) \leq (k+1)\omega(\delta)$.

(3) If $\omega(\delta) =$ *for some* $\delta > 0$, *then* $f(x)$ *is a constant.*

Proof. (1) $f(x + nh) - f(x) = \sum_{k=o}^{n-1} \left\{ f(x + kh + h) - f(x + kh) \right\}$.

For $h \leq \delta$, the R.H.S. is numerically at most $n\omega(\delta)$.

(2) If k is not an integer, let n be the integer next greater. Then

$$\omega(k\delta) \leq \omega(n\delta) \leq n\omega(\delta) \leq (k+1)\omega(\delta).$$

(3) $f(x)$ is constant in any interval less that δ, and so everywhere.

□

Lipschitz condition. Def. $f(x)$ satisfies the Lipschitz condition of order α(briefly, is Lip. α) in a given interval, if for every x_1, x_2 in it,

$$|f(x_2) - f(x_1)| \leq A|x_2 - x_1|^\alpha.$$

It follows that $\omega(\delta) \leq A\delta^\alpha$.

In this, $0 < \alpha \leq 1$. If $\alpha > 1$, $f(x)$ can only be a constant,be a constant, because then

$$\omega(\delta) \leq n\omega(\delta/n) \leq A\delta^\alpha/n^{\alpha-1}.$$

Making $n \to \infty$, we have $\omega(\delta) = 0$.

11 Fourier and Fejer Sums

We collect for reference in Theorem 17 some well-known facts. Proofs 35
can be found in any text-book of analysis which includes a chapter on Fourier Series.

Theorem 17. (1) *The sum*

$$S_n = \frac{1}{2}a_0 + \sum_{r=1}^{n}(a_r \cos rx + b_r \sin rx)$$

for the Fourier Series of $f(x)$ *is equal to*

$$\frac{1}{\pi}\int_0^{\pi}\{f(x+t)+f(x-t)\}\frac{\sin\left(n+\frac{1}{2}\right)t}{2\sin\frac{1}{2}t}dt.$$

(2) $|f(x) - S_n(x)| < M(A\log n + B)$, *where* $M = \sup|f(x)|$ *and* A, B *are constants.*

(3) *If* $\sigma_n = (S_0 + S_1 \cdots + S_{n-1})/n$

is the Fejer $(C1)$ *sum of the Fourier series of* $f(x)$, *then*

$$\sigma_n = \frac{1}{n\pi}\int_0^{\pi} f(x+2t)\left(\frac{\sin nt}{\sin t}\right)^2 dt.$$

(4) $\dfrac{1}{\sin^2 t} = \sum_{-\infty}^{\infty}\dfrac{1}{(t+k\pi)^2} \qquad (t \neq k\pi).$

The result (2), which cannot be improved, shows that, in general, the Fourier series of a function gives a poor approximation in the sense measured by $\sup|f(x) - S_n(x)|$. As the R.H.S. of (2) tends to infinity with n, (2) does not include Weierstrass's Theorem 15. The sense in which the Fourier series does give the best approximation is the mean-square sense (omitted here). It is known that the Fejer sums σ_n of (3) behave more regularly than the Fourier sums S_n; this is due to the kernel $(\sin nt/\sin t)^2$ in the integral for ${}^{\sigma}_{n}$ being positive, whereas the kernel in S_n takes both signs. The next theorem gives the approximation to $f(x)$ afforded by $({}^{\sigma}_{n}(x)$.

Theorem 18. *If $f(x)$ has modulus of continuity $\omega(\delta)$, then* **36**

$$|f(x) - \sigma_n(x)| \leq A\omega(1/n)|log\omega(1/n)|.$$

Proof. We first put $\sigma_n(x)$ into a form more convenient than that of Theorem 17(3). Since f is periodic

$$\int_0^\pi f(x+2t)\frac{\sin^2 nt}{(t+k\pi)^2}dt = \int_{k\pi}^{(k+1)\pi} f(x+2t)\frac{\sin^2 nt}{(t^2)}dt.$$

□

Then, from (3) and (4) of Theorem 17,

$$\sigma_n = \frac{1}{n\pi}\int_0^\pi f(x+2t)\frac{\sin^2 nt}{\sin^2 t}dt = \frac{1}{n\pi}\int_{-\infty}^\infty f(x+2t)\frac{\sin^2 nt}{t^2}dt,$$

and so, by changing the variable from t to t/n,

$$\sigma_n = \frac{1}{\pi}\int_{-\infty}^\infty f(x+\frac{2t}{n})\frac{\sin^2 t}{t^2}dt$$
$$\sigma_n - f = \frac{1}{\pi}\int_{-\infty}^\infty \left\{f(x+\frac{2t}{n}) - f(x)\right\}\frac{\sin^2 t}{t^2}dt.$$

Therefore

$$|f - \sigma_n| \leq \frac{2}{\pi}\int_0^\infty \omega(2t/n)\frac{\sin^2 t}{t^2}dt.$$

The integral on the R.H.S. is the sum of the integrals over $(0,1)$, $(1,X)$ and (X,∞). This gives

$$\begin{aligned}|f - \sigma_n| &\leq \frac{2}{\pi}\left\{\omega(2/n) + \int_1^X \omega(2t/n)\frac{dt}{t^2} + \omega(\pi)\int_X^\infty \frac{dt}{t^2}\right\}\\ &\leq \frac{2}{\pi}\left\{\omega(2/n) + \omega(2/n)\int_1^X \frac{t+1}{t^2}dt + \frac{\omega(\pi)}{X}\right\}\\ &\leq \frac{2}{\pi}\left\{\omega(2/n)(2+\log X) + \frac{\omega(\pi)}{X}\right\}.\end{aligned}$$

Choose $X = 1/\omega(2/n)$ and we have a result equivalent to that stated.

37 **Corollary 1.** *Theorem 15*

Corollary 2. *If $\omega(\delta) < A\delta^{\alpha}(0 < \alpha < 1)$, then $|f - \sigma_n| < \frac{AB}{n^{\alpha}}$, where $B = B(\alpha)$ is independent of f.*

Proof.

$$\begin{aligned} |f - \sigma_n| &\leq \frac{2}{\pi}\int_0^{\infty} \omega(2t/n)\frac{\sin^2 t}{t^2}dt \\ &\leq \frac{2^{\alpha+1}A}{\pi n^{\alpha}}\int_0^{\infty} t^{\alpha}\frac{\sin^2 t}{t^2}dt \end{aligned}$$

The estimates in Chapter III would lead us to suspect that, if we can find a $t(x)$ which approximates to $f(x)$ more closely than $\sigma_n(x)$ does, we may get rid of the $\log\omega(1/n)$ on the R.H.S. of Theorem 18. It is easy to see how to try to do this. The logarithm arises from integrating a term in $1/t$. The Fejer sum is

$$F_r(x,n) = \frac{1}{J_r}\int_{-\infty}^{\infty} f\left(x + \frac{2t}{n}\right)\left(\frac{\sin t}{t}\right)2r\,dt$$

for $r = 1$ and $J_r = \pi$. If $r \geq 2$, there will be no term in $1/t$. We shall achieve our purpose by taking $r = 2$. ◻

Lemma 1. *(1)* $J_2 = \overset{\infty}{\int} -\infty \left(\frac{\sin t}{t}\right)^4 dt = \frac{2\pi}{3}$.

(2) $F_2(x,n)$ is in T_{2n-1}.

Proof. (1)

$$\begin{aligned} J_2 &= \int_0^{\pi} \sin^4 t \sum_{-\infty}^{\infty} \frac{1}{(t+k\pi)^4}dt \\ &= \frac{1}{6}\int_0^{\pi} \sin^4 t \frac{d^2}{dt^2}\left(\frac{1}{\sin^2 t}\right)dt \\ &= \frac{1}{6}\int_0^{\pi} \sin^4 t\left(\frac{6}{\sin^4 t} - \frac{4}{\sin^2 t}\right)dt = \frac{2\pi}{3}. \end{aligned}$$

(2) Reversing the steps by which $F_1(x,n)$ was obtained in the first part of the proof of Theorem 18. we have

$$F_2(x,n) = \frac{3}{2\pi}\frac{3}{6n}\int_0^{\pi} f(x+2t)\sin^4 nt \frac{d^2}{dt^2}\left(\frac{1}{\sin^2 t}\right)dt.$$

38 Then $\sin^4 nt \dfrac{d^2}{dt^2}\left(\dfrac{1}{\sin^2 t}\right) = \sin^4 nt\left(\dfrac{6}{\sin^4 t} - \dfrac{4}{\sin^2 t}\right).$ □

Now $\dfrac{\sin nt}{\sin t}$ is the sum of multiples of $\cos kt$ where $k \leq n-1$. Hence $\sin^4 nt \dfrac{d^2}{dt^2}\left(\dfrac{1}{\sin^2 t}\right)$ is the sum of multiples of $\cos kt$ where $k \leq 4n-2$. Moreover, the expression is even and has period π, so k takes only even values, 21 say, where $1 \leq 2n-1$. Finally,

$$\int_0^{\pi} f(x+2n)\cos 21tdt = \frac{1}{2}\int_0^{2\pi} f(u)\cos l(u-x)du$$

and $F_2(x,n)$ is in T_{2n-1}.

Theorem 19. $|f(x) - F_2(x,n)| \leq 3\omega(1/n)$.

Proof. $F_2(x,n) - f(x) = \dfrac{3}{2\pi}\displaystyle\int_{-\infty}^{\infty}\left\{f(x+\frac{2t}{n}) - f(x)\right\}\left(\frac{\sin t}{t}\right)^4 dt.$ □

Now $|f(x+\dfrac{2t}{n}) - f(x)| \leq \omega\left(\dfrac{2|t|}{n}\right) \leq (2|t|+1)\omega(1/n)$ from Theorem 16 (2). Therefore

$$|f(x) - F_2(x,n)| \leq \omega(1/n)\frac{3}{\pi}\int_0^{\infty}(2t+1)\left(\frac{\sin t}{t}\right)^4 dt = A\omega(1/n),$$

where $A = 1 + \dfrac{6}{\pi}\displaystyle\int_0^{\infty}\frac{\sin^4 t}{t^3}dt$. But

$$\int_0^{\infty}\frac{\sin^4 t}{t^3}dt \int_0^{\infty}|\frac{\sin^3 t}{t^2}|dt = \frac{1}{2}\int_0^{\pi}\frac{\sin^3 t}{\sin^2 t}dt = 1$$

(again by use of Theorem 17(4)).

So $\quad A < 1 + \dfrac{6}{\pi} < 3.$

Theorem 20. *If $f(x)$ is $C(2\pi)$ and $f'(x)$ is continuous with modulus of continuity $\omega_1(\delta)$, then*

$$|f(x) - F_2(x,n)| \leq \frac{A}{n}\omega_1(1/n) \text{ where } A < 5/2.$$

Proof. $F_2(x,n) - f(x) = \dfrac{3}{2\pi}\int\limits_0^{\infty}\left\{f\left(x+\dfrac{2t}{n}\right) + f\left(x-\dfrac{2t}{n}\right) - 2f(x)\right\}\left(\dfrac{\sin t}{t}\right)^4 dt.$

The modulus of the term within { } is 39

$$\begin{aligned}
&|\frac{2}{n}\int_0^t\left\{f'\left(x+\frac{2u}{n}\right) - f'\left(x-\frac{2u}{n}\right)\right\}du| \\
&\leq \frac{2}{n}\int\limits_0^t \omega_1\left(\frac{4u}{n}\right)du \\
&\leq \frac{2}{n}\omega_1\left(\frac{1}{n}\right)\int\limits_0^t (4u+1)du \\
&= \frac{2}{n}\omega_1\left(\frac{1}{n}\right)(2t^2+t)
\end{aligned}$$

□

Therefore

$$|f(x) - F_2(x,n)| \leq \frac{A}{n}\omega_1\left(\frac{1}{n}\right),$$

where

$$A = \frac{3}{\pi}\int\limits_0^{\infty}(2t^2+t)\left(\frac{\sin t}{t}\right)^4 dt = \frac{3}{\pi}\sin^2 t dt + \frac{3}{\pi}\int\limits_0^{\infty}\frac{\sin^4 t}{t^3}dt < \frac{3}{\pi}\left(\frac{\pi}{2}+1\right) < \frac{5}{2}$$

Theorem 20 can be extended to higher derivatives. If $f(x)$ has an r-th derivative with modulus of continuity $\omega_r(\delta)$, the approximation attainable in T_n is a constant multiple of $n^{-r}\omega_r(1/n)$.

Notes on Chapter IV 40

1. Use the singular integral (de la Vallee Poussin)

$$\frac{1}{J_n}\int_{-\pi}^{\pi}\cos^{2n}\frac{1}{2}(t-x)f(t)dt,$$

where J_n is the value of the integral when $f(t) = 1$, to give a direct proof of Theorem 15.

2. Assuming Theorem 15 proved, deduce Theorem 1 from it.

3. Deduce Theorem 15 from Theorem 1 as follows:

   (a) Prove that, if $f(x)$ is $C(0,\pi)$, it can be approximated uniformly by a $t(x)$ containing cosines only.

   (b) By applying (a) to the even functions

$$2g(x) = f(x) + f(-x)$$
$$2h(x) = \left\{f(x) - f(-x)\right\}\sin x,$$

   deduce that $f(x)$ is uniformly approximately by a $t(x)$.

4. With the notation of Illustration (2), Corollary, page 31, prove that Weierstrass's function $\sum a^r \cos b^r x$ satisfies a Lipschitz condition of order α.

41 **Hints**

1 Follow Theorem 2. Detail is in Natanson, 10.

2 Approximate to $\cos kx$ and $\sin kx$ by a finite number of terms of their expansions in powers of x.

3 (a) Put $y = \cos x$.

(b) $g(x), h(x)$ are uniformly approximately in $(-\pi,\pi)$. So is $g(x)\sin^2 x + h(x)\sin^2 x$. So is $f(x)\cos^2 x$, and hence $f(x)(\sin^2 x + \cos^2 x)$.

4 Given h, choose n so that $b^n h \leq 1 < b^{n+1} h$.

$$f(x+h) - f(x-h) = -2 \sum_1^\infty a^r \sin b^r h \sin b^r x$$

$$= \sum_1^n + \sum_{n+1}^\infty$$

$$\left| \sum_{n+1}^\infty \right| \leq 2 \sum_{n+1}^\infty a^r = \frac{2a^{n+1}}{1-a}$$

$$\left| \sum_1^n \right| \leq 2h \sum_1^n a^r b^r = 2abh \frac{a^n b^n - 1}{ab-1} < \frac{2ba^{n+1}}{ab-1}$$

But $\quad a^{n+1} = b^{-\alpha(n+1)} < h^\alpha$.

Hence $\quad |f(x+h) - f(x-h)| < Ah^\alpha$.

With more trouble (Aschieser and Krein, 167) this can be proved best possible.

Chapter 5

Inequalities, etc.

12 Bernstein's and Markoff's Inequalities

Theorem 21 (Bernstein)**.** *If* $t(x) = \frac{1}{2}a_o + \sum_1^n (a_k \cos kx + b_k \sin kx)$ *then* **42**
$|t'(x)| \leq n \sup |t(x)|$.

Proof. Suppose, on the contrary, that

$$\sup |t'(x)| = nl,$$

where $$l > \sup |t(x)|.$$ □

$t'(x)$, being continuous, attains its bounds and so, for some $c, t'(a) = \pm nl$ and we will suppose that

$$t'(c) = nl.$$

Since nl is a maximum value of $t'(x)$,

$$t''(c) = 0.$$

Define $$S(x) = l \sin n(x - c) - t(x).$$

Then $r(x) = S'(x) = nl \cos n(x - c) - t'(x)$.

$S(x)$ and $r(x)$ both have order n.

Consider the points

$$u_o = c + \pi/2n, u_k = u_o + k\pi/n (1 \leq k \leq 2n).$$

Then

$$S(u_o) = 1 - t(u_o) > 0$$
$$S(u_1) = 1 - t(u_1) < 0$$
$$\cdots\cdots$$
$$S(u_{2n}) = 1 - t(u_{2n}) > 0$$

Each of the $2n$ intervals $(u_o, u_1), (u_1, u_2), \ldots, (u_{2n-1}, u_{2n})$ then contains a zero of $S(x)$.say

$$S(y_i) = 0,$$

43 where $u_i < y_i < u_{i+1}, (0 \le i \le 2n - 1)$. Clearly

$$y_{2n-1} < y_o + 2\pi.$$

Write $$y_{2n} = y_o + 2\pi.$$

Then $$S(y_{2n}) = S(y_o) = 0.$$

By Rolle's Theorem, there is a zero x_i of $r(x)$ inside each interval (y_i, y_{i+1}) where $0 \le i \le 2n - 1$. Clearly

$$x_{2n-1} < x_o + 2\pi.$$

Now $r(c) = nl - t'(c) = 0.$

Since the polynomial $r(x)$ of order n has at most $2n$ zeros, it follows that, for some k,

$$c \equiv x_k \pmod{2\pi}.$$

But $r'(c) = -t''(c) = 0.$

Therefore c(and so x_k) is a double zero (at least) of $r(x)$.

Therefore the $x_i(0 \le i \le 2n - 1)$ provide at least $2n + 1$ zeros of $r(x)$. This is only possible if $r(x) \equiv 0$, and so $S(x)$ is a constant. But $S(u_o) > 0$ and $S(u_1) < 0$ and we have a contradiction

Corollary 1. $t(x) = \sin nx$ *shows that the result is the best possible.*

Corollary 2. *The algebraic equivalent is- If* $p(x)$ *has degree* n *and* $|p(x)| \le M$ *in* $(-1, 1)$*, then*

$$|p'(x)| \le nM\sqrt{(1 - x^2)}.$$

Proof. Put

$$\begin{aligned} t(\theta) &= p(\cos\theta) \\ t'(\theta) &= -p'(\cos\theta)\sin(\theta). \end{aligned}$$

The bound for $|p'(x)|$ given in Corollary 2 fails at the end-points ± 1. A **44**
better result, due to Markoff, is

$$|p'(x)| \le Mn^2,$$

as will be proved in Theorem 22. □

Lemma 1. *Let*

$$x_k = \cos\frac{(2k-1)\pi}{2n} \qquad (k = 1, 2, \ldots, n)$$

be the zeros of the Chebyschev polynomial $T_n(x)$. If $q(x)$ is in P_{n-1}, then

$$q(x) = \frac{1}{n}\sum_{k=1}^{n}(-1)^{k-1}\sqrt{(1-x_k^2)}\,q(x_k)\cdot\frac{T_n(x)}{x-x_k}.$$

Proof. Both sides are in P_{n-1} and so it is sufficient to show that they agree for the n values x_k. As $x \to x_k$,

$$\begin{aligned} \frac{T_n(x)}{x-x_k} \to T_n'(x_k) &= \frac{n}{\sqrt{(1-x_k^2)}}\sin(n\ \text{arc}\ \cos x_k) \\ &= \frac{n(-1)^{k-1}}{\sqrt{(1-x_k^2)}} \end{aligned}$$

Also, for $x = x_k$, every term on the R.H.S. except the k-th vanishes. □

Lemma 2. *Suppose that $q(x)$ is in P_{n-1} and $|q(x)| \le \dfrac{1}{\sqrt{(1-x^2)}}(-1 < x < 1)$.*

Then $|q(x)| \le n\,(-1 \le x \le 1)$.

Proof. With the notation of Lemma 1, if $-x_1 = x_n \le x \le x_1$,

$$\sqrt{(1-x^2)} \ge \sqrt{(1-x_1^2)} = \sin\frac{\pi}{2n} \ge \frac{1}{n}.$$

Therefore Lemma 2 is true for $x_n \le x \le x_1$. If $x_1 < x \le 1$ (or $-1 \le x < x_n$) Lemma 1 gives

$$|q(x)| \le \frac{1}{n}\Big|\sum \frac{T_n(x)}{x-x_k}\Big|,$$

45 since all the $x - x_k$ are positive (or all negative). Now

$$T_n(x) = 2^{n-1}\prod(x-x_k),$$

and so
$$\frac{T_n'(x)}{T_n(x)} = \sum\frac{1}{x-x_k}.$$

□

Therefore $$|q(x)| \le \frac{1}{n}|T_n'(x)|.$$

But, if $x = \cos\theta$, $T_n'(x) = \dfrac{n\sin n\theta}{\sin\theta}$, which gives

$$|T_n'(x)| \le n^2.$$

Theorem 22 (Markoff)**.** *If $p(x)$ is in P_n, then*

$$|p'(x)| \le n^2 \sup|p(x)| \qquad -1 \le x \le 1.$$

Proof. If $\sup|p(x)| = M$, take in Lemma 2,

$$q(x) = \frac{p'(x)}{M_n}.$$

□

Corollary. *$p(x) = T_n(x)$ shows that the result is the best possible.*

13 Structural Properties Depend on the closeness of the approximation

Theorem 21 can be used to prove theorems of a type converse to Theorems 18-20. Theorem 23, which is complementary to Theorem 18, Corollary 2, will suffice to show the method.

Theorem 23. *Let $f(x)$ be $C(2\pi)$. Suppose that, for all n, the best approximation in T_n to $f(x)$ is less than A/n^α, where $0 < \alpha < 1$. Then $f(x)$ is Lip.α.*

Proof. Let $t_n(x)$ satisfy

$$|f(x) - t_n(x)| \leq \frac{A}{n^\alpha}.$$

Define
$$u_0(x) = t_1(x)$$
$$a_n(x)t_{2^n}(x) - t_{2^{n-1}}(x) \qquad (n \geq 1).$$

□

Then $f(x)$ is the sum of the uniformly convergent series $\sum\limits_0^\infty u_n(x)$. 46
Choose δ with $0 < \delta \leq \frac{1}{2}$, and m such that

$$2^{m-1} \leq \frac{1}{\delta} < 2^m.$$

Suppose $|x - y| \leq \delta$. We have

$$|f(x) - f(y)| \leq \sum_0^{m-1} |u_n(x) - u_n(y)| + \sum_m^\infty |u_n(x)| + \sum_m^\infty |u_n(y)|.$$

We shall find upper bounds for the terms on the R.H.S.

$$|u_n(x)| \leq |t_{2^n}(x) - f(x)| + |f(x) - t_{2^{n-1}}(x)|$$
$$\leq \frac{A}{2^{n\alpha}} + \frac{A}{2^{(n-1)}\alpha} = \frac{A(1+2^\alpha)}{2^{n\alpha}}.$$

Therefore

$$\sum_{m}^{\infty} |u_n(x)| \leq A(1+2^{\alpha}) \sum_{m}^{\infty} \frac{1}{2^{n\alpha}} = \frac{A(1+2^{\alpha})}{1-2^{-\alpha}} \frac{1}{2^{m\alpha}}.$$

This gives

$$|f(x) - f(y)| \leq \sum_{o}^{m-1} |u_n(x) - u_n(y)| + \frac{B}{2^{m\alpha}}.$$

Theorem 21 applied to $u_n(x)$ gives

$$|u_n'(x)| \leq 2^n \sup |u_n(x)| \leq A(1+2^{\alpha}) 2^{n(1-\alpha)}.$$

By the mean-value theorem,

$$|u_n(x) - u_n(y)| \leq |u_n'(\xi)||x-y| \leq A(1+2^{\alpha}) 2^{n(1-\alpha)} \delta$$

Therefore

$$|f(x) - f(y)| \leq A(1+2^{\alpha}) \delta \sum_{o}^{m-1} 2^{n(1-\alpha)} + \frac{B}{2^{m\alpha}}.$$

Putting $C = A(1+2^{\alpha})$ and using $\frac{1}{2^m} < \delta$, we have

$$\omega(\delta) \leq C\delta \sum_{o}^{m-1} 2^{n(1-\alpha)} + B\delta^{\alpha}.$$

47 If now $\alpha < 1$,

$$\sum_{o}^{m-1} 2^{n(1-\alpha)} = \frac{2^{m(1-\alpha)} - 1}{2^{1-\alpha} - 1} < \frac{2^{m(1-\alpha)}}{2^{1-\alpha} - 1}.$$

Use now $2^m \leq \frac{2}{\delta}$ and we find

$$\omega(\delta) < \left(\frac{2^{1-\alpha}}{2^{1-\alpha} - 1} C + B \right) \delta^{\alpha}$$

See Notes 1-4 at the end of the Chapter.

14 Divergence of the Lagrange Sequence

There is a sense in which the Lagrange polynomial of degree n (§4) fitted to a function $f(x)$ at $n+1$ points equally spaced through an interval follows the function closely. It is natural to expect that, by increasing n, the approximation would improve and we might, for instance, find another proof of Theorem 1 on these lines. Such expectations are falsified. Unless heavy restrictions are laid on $f(x)$, the sequence of Lagrange polynomials diverges except for certain special values of x.

We shall construct an example of this phenomenon.

Lemma. *Let $p(x)$ be the Lagrange polynomial which takes the value 0 at the $2m$ values of x*

$$k/m \qquad (-m \le k \le m; k \ne 2)$$

and takes the value $1/m$ when $x = 2/m$. Then, if m is odd, $|p\left(\frac{1}{2}\right)| \to \infty$ as $m \to \infty$.

Proof. The polynomial $p(x)$, of degree $2m$, is

$$\frac{1}{m}\frac{(x+1)(x+\frac{m-1}{m})\ldots x(x-1/m)(x-3/m)\ldots(x-1)}{(2/m+1)(2/m+\frac{m-1}{m})\ldots 2/m(2/m-1/m)(2/m-3/m)\ldots(2/m-1)}$$

This gives for $|p\left(\frac{1}{2}\right)|$ **48**

$$\frac{1}{m}\frac{3m(3m-2)\ldots m(m-2)(m-6)\ldots 1.1.3.\ldots.m}{2^{2m}(m+1)(m+1)\ldots 2.1.1.2\ldots \qquad (m-2)}$$

□

This can be estimated by forming it into factorials and using Stirbing's theorem. More simply, we can prove that it tends to ∞ by grouping the factors as follows:

$$|p\left(\frac{1}{2}\right)| = \frac{m-1}{2(m+1)(m+2)(m-4)}A^2BC,$$

where
$$A = \frac{3.54\ldots m}{2.4\ldots(m-1)}$$
$$B = \frac{(m+2)(m+4)\ldots(2m+1)}{(m+1)(m+3)\ldots 2m}$$
$$C = \left(-\frac{2m+3}{m+1}\right)\left(\frac{3m+5}{m+3}\right)\cdots\left(\frac{3m}{2m-2}\right).$$

Here $A > 1, B > 1$, and the factors of C decrease from left to right, the last being greater than $3/2$. So $C > (3/2)^{m-1}$.

Note. $x = \frac{1}{2}$ has been taken for ease of calculation. The conclusion holds for other values of x.

Theorem 24 (Borel)**.** *There is $f(x)$ in $C(-1,1)$ whose nth Lagrange polynomial does not converge to $f(x)$ as $n \to \infty$.*

Proof. Define a continuous curve C_k which coincides with Ox outside the interval $(3^{-k-1}, 3^{-k})$ and has maximum 3^{-k-1} at the midpoint of that interval. For example we can define C_k by

$$y = 3^{-k-1}\sin\left\{(3^{k+1}x-1)\pi/2\right\}.$$

□

We shall use the C_k to construct a curve S.$P_{k,S}(x)$ will denote the Lagrange polynomial which takes the same values as S for the values $x = 1/3^k$, where $-3^k \leq 1$ (integer) $\leq 3^k$.

49 We shall construct S so that $P_{k,S}\left(\frac{1}{2}\right)$ does not converge to the point on S where $x = \frac{1}{2}$. Observe first that $P_{k,C_{k-1}}$ is the Lagrange polynomial in the Lemma with $m = 3^k$. From the Lemma, given A, there is h_1 such that

$$|P_{k,C_{k-1}}\left(\frac{1}{2}\right)| > 2A \text{ if } k-1 > h_1.$$

There are two possibilities:

(a) With h_1 fixed, $P_{k,C_{h1}}\left(\frac{1}{2}\right)$ does not tend to 0 as $k \to \infty$. Then S can be taken to be C_{h1}; or

(b) there exists r such that

$$|P_{k,C_{h1}}\left(\frac{1}{2}\right)| < \frac{1}{2}A \text{ for all } k > r_1.$$

Choose $h_2 > \max(h_1, r_1)$.

Define $D_{2,k}$ to be the sine-curves in C_{h1} and $C_{k-1}(k-1 \geq h_2)$, and, for the rest, the x-axis in $(-1, 1)$.

$D_{2,k}$ is a continuous curve; its ordinate for $x = \dfrac{1}{2}$ is 0, and

$$P_{k,D_{2,k}} = P_{k,C_{h_1}} + P_{k,C_{k-1}}.$$

From above, since $k - 1 \geq h_2$,

$$|P_{k,D_{2,k}}\left(\frac{1}{2}\right)| > 2A - \frac{1}{2}A.$$

Again, there are two possibilities:

(a) With h_2 fixed, $P_{k,D_{2,k}}\left(\dfrac{1}{2}\right)$ does not tend to 0 as $k \to \infty$. Then S can be taken to be D_{2,h_2}; or

(b) there exists r_2 such that

$$|P_{k,D_{2,k_{h_2}}}\left(\frac{1}{2}\right)| < \frac{1}{4} \cdot A \text{ for all } k > r_2.$$

Choose $h_3 > \max(h_2, r_2)$. **50**

Define $D_{3,k}$ to be the sine-curves in C_{h_1}, C_{h_2} and $C_{k-1}(k-1 \geq h_3)$ and, for the rest, the -axis in $(-1, 1)$. After n repetitions, there are two possibilities:

(a) There is a D_{n,h_n} for which the k th Lagrange polynomial does not tend to 0 at $x = \dfrac{1}{2}$; and this serves for S; or

(b) there is an infinite sequence D_{n,h_n} for which

$$|P_{h_n+1,D_{n,h_n}}\left(\frac{1}{2}\right)| > 2A - \frac{1}{2}A - \frac{1}{4}A - \ldots - \frac{A}{2^{n-1}} > A.$$

As $n \to \infty, D_{n,h_n}$ defines a continuous curve S whose ordinate for $x = \dfrac{1}{2}$ is 0. Its Lagrange polynomial takes values greater than A for $x = \dfrac{1}{2}$ when its degree is $h_1, h_2, \ldots, h_n, \ldots$.

51

Notes

1. Weierstrass's function $\sum a^r \cos b^r x$ illustrates Theorem 18 (Corollary 2) and Theorem 23. See Chapter IV, note 4.

2. If $\alpha = 1$, the best that can be proved in Theorem 23 is that $\omega(\delta) < A\delta \log(1/\delta)$. The latter part of the argument can be adapted for this purpose (Natanson, 91).

   The function $\sum_1^\infty \frac{\sin nx}{n^2}$ satisfies $d_n < \dfrac{1}{n}$, but is not in Lip.1 (Natanson, 93).

3. A condition which is necessary and sufficient of $d_n < A/n$ is that

$$|f(x+h) - 2f(x) + f(x-h)| < Bh.$$

   (Zymund, Duke Mathematical Journal, 12(1945)47 or Natanson, 96).

4. (Extension of Theorem 23). If, for $f(x)$, $d_n < A/n^{p+\alpha}$ (p = integer, $0 < \alpha < 1$), then $f(x)$ has a pth derivative $f^{(p)}(x)$ in Lip α.

5. For further 'negative results' like Theorem 24, see Natanson, 369 – 388. For example, the Lagrange polynomial taking the values of $|x|$ at n equally spaced points in $(-1, 1)$ converges to $|x|$ as $n \to \infty$ for no value of x expect $0, \pm 1$.

Chapter 6

Approximation in Terms of Differences

15

This is the only chapter in the course, of which the results are not classical. The point of view here might lead to a re-orientation towards algebraic rather than trigonometric polynomials. **52**

In Theorem 20 (and its known extensions) the approximation attainable in P_n or T_n to a differentiable function $f(x)$ is expressed in terms of its first higher derivative. We shall now give simple examples which lead us to suppose that bounds of *differences* of $f(x)$ rather than derivatives may be more directly related to the closeness of the approximation.

Example 1. If $f(x)$ attains its greatest value at x_2 and its least at x_1, then the best approximation in P_0 is

$$\frac{1}{2}\{f(x_1)+f(x_2)\}$$

and
$$d_o = \frac{1}{2}\{f(x_2)-f(x_1)\} = \frac{1}{2}\sup|\Delta f|$$

This depends solely on the first difference of $f(x)$; the derivative of $f(x)$- if it exists-has no bearing on it.

Now raise the degree by one.

Example 2. If $f(x)$ is $C(0,1)$, there is a linear function $p(x)$ for which

$$|f(x) - p(x)| \leq \sup |f(x+2h) - 2f(x+h) + f(x)|,$$

the sup being taken over all x, h such that

$$0 \leq x \leq x + 2h \leq 1.$$

53 *Proof.* Define $p(x)$ to be equal to $f(x)$ at $x = 0$ and $x = 1$ Write

$$g(x) = f(x) - p(x).$$

□

Then $|g(x)|$ attains its maximum, M, for $0 \leq x \leq 1$ at x_1, say.
If $0 \leq x_1 \leq \dfrac{1}{2}$, take $x = 0, h = x_1$. Then

$$|g(x+2h) - 2g(x+h) + g(x)| = |g(2x_1) - 2g(x_1) + g(0)| \geq M.$$

If $\dfrac{1}{2} < x_1 \leq 1$ take $x = 2x_1 - 1$, $h = 1 - x_1$. Then

$$|g(x+2h) - 2g(x+h) + g(x)| = |g(1) - 2g(x_1) + g(1 - x_1)| \geq M.$$

But the second difference of $g(x)$ and $f(x)$ are equal.

By a longer argument it is possible to prove the corresponding result for P_2 and the third difference.

The general result was conjectured in 1949 by H. Burkill.

Theorem 25. *There is a number K_n depending only on n such that given $f(x)$ in $C(a,b)$, there is a polynomial $p(x)$ in P_{n-1} for which*

$$|f(x) - p(x)| \leq K_n \sup |\Delta_n(f)|$$

(where the supremum is taken for all sets of $n+1$ points $x, \ldots, x+nh$ in (a,b))

The theorem looks innocent, but attempts at it failed until Whitney it in 1955. He took for his $p(x)$ the Lagrange polynomial for the points of division of (a, b) into $n - 1$ equal parts. His work does not yield an estimate of K_n for general n; in view of Theorem 24, we should hardly expect good value of K_n.

54 Whitney's elegant arguments are too long for reproduction here, and the reader is referred to his paper in journal de Mathematiques 36(1957), 67-95.

It is worth observing, however, that instead of the usual n^{th} difference with equal increments, we can take a more general n^{th} difference depending of the values of $f(x)$ at $n + 1$ arbitrary points. The difficulty then disappears and the polynomial of best approximation can be used instead of the Lagrange polynomial.

16 Definition and Properties of the n^{th} Difference

If

$$\varphi(u) = (u - h_o)(u - h_1) \cdots (u - h_n),$$

the n^{th} divided difference of $f(x)$ for the values specified is commonly defined by

$$D_n = D_n(f; h_o, \ldots, h_n) = \sum_{i=0}^{n} \frac{f(h_i)}{\varphi'(h_i)}.$$

In what follows it will be convenient to suppose that

$$h_o > h_1 > \cdots > h_n$$

To define an n^{th} difference Δ_n, as distinct from a divided difference, we naturally take

$$\Delta_n = H_n D_n,$$

where H_n is homogeneous of degree n in the $h's$.

The most suitable definition of H_n appears to be

$$H_n = 2^n / T_n,$$

where
$$T_n = T_n(h_o, h_1, \ldots, h_n) = \sum_{i=0}^{n} |\varphi'(h_i)|^{-1}.$$

55 In the special case of equal increments with $h_o - h_n = nh$, this gives $H_n = n!h^n$, which is right.

As a further check on the appropriateness of our H_n, we observe that if a function is numerically less than A, its n^{th} difference Δ_n is numerically less than $2^n A$.

In working with D_n, Δ_n, etc., we shall specify the function and the values of the variables only so far as is necessary for clarity.

We are now in a position to restate and prove Theorem 25, taking $\Delta_n(f)$ to be the difference $H_n D_n(f)$ just defined.

Theorem 25′. *Theorem 25 is true with $K_n = 2^{-n}$ and $\Delta_n(f)$ as just defined and the supremum taken over all values of $h_o, \ldots, h_n$ in (a, b).*

Proof. Given $f(x)$, take $p(x)$ to be its polynomial of best approximation of degree at most $n - 1$. Then $f(x) - p(x)$ takes its greatest numerical value at $n + 1$ points, with signs alternately + and −. These $n + 1$ points we takes as $h_o, \ldots, h_n$. □

Since the n^{th} difference of a polynomial of degree $n - 1$ is 0, we have

$$\Delta_n(f; h_o, \ldots, h_n) = H_n \sum_{i=o}^{n} \frac{f(h_i) - p(h_i)}{\varphi'(h_i)}$$

So
$$|\Delta_n| = H_{n^d} \sum |\varphi'(h_i)|^{-1} = 2^n d,$$

by definition of H_n.

Therefore, for all x in $(-1, 1)$,

$$|f(x) - p(x)| \le d \le 2^{-n} \sup |\Delta_n(f)|.$$

This proves Theorem 25′.

56 Alternatively we can prove Theorem 25′, starting from the upper bound of $|\Delta_n|$ instead of from the polynomial of best approximation.

Suppose, then, that $\sup|\Delta_n| = L$ and that the bound L is assumed for the values $h_o, h_1, \dots, h_n$ of the independent variable. Define points (h_i, y_i) for $i = 0, 1, \dots,$ by taking

$$y_i = f(h_i) - (-1)^k \frac{L}{2^n},$$

where k is i or $i+1$ according as $\Delta_n(f, h_o, \dots, h_n)$ is positive or negative.

Construct *a* $p(x)$ of P_{n-1} through the n points (h_i, y_i) for $i = 0, 1, \dots, n-1$. Write

$$g(x) = f(x) - p(x).$$

Since $\Delta_n(p) \equiv 0$, $|\Delta_n(g)| = |\Delta_n(f)|$ attains its upper bound for $h_0, h_1, \dots, h_n$. From the definition of Δ_n, the value of $g(h_n)$ which makes $|\Delta_n(g, h_0, h_1, \dots, h_n)| = L$ is $(-1)^k L/2^n$, where k is n or $n+1$ according as $\Delta_n(f, h_o, \dots, h_n)$ is positive or negative.

We prove that $|g(x)| \leq 2^{-n}L$ for all x. Suppose that $|g|$ takes values greater than $L/2^n$, say $g(h_i') > L/2^n$ for a value h_i' between h_{i-1} and h_{i+1} where $g(h_i) = 2^{-n}L$. Then, from the definition of T_n,

$$D_n(g, h_o, \dots, h_{i-1}, h_i', h_{i+1}, \dots, h_n) \\ > 2^{-n} L T_n(h_o, \dots, h_{i-1}, h_i', h_{i+1}, \dots, h_n)$$

and so, by definition of Δ_n,

$$\Delta_n(g, h_o, \dots, h_{i-1}, h_i', h_{i+1}, \dots, h_n) > L,$$

which is a contradiction. We have therefore

$$|f(x) - p(x)| = |g(x)| \leq 2^{-n} \sup|\Delta_n(f)|,$$

which is Theorem 25′. **57**

For the next result let us call the $n+1$ values

$$-1, -\cos\frac{\pi}{n}, \dots, \cos\frac{\pi}{n}, 1$$

at which the Chebyshev polynomial $\cos(n \operatorname{arc\,cos} x)$ assumes the values ± 1 the Chebyshev points of the interval $(-1, 1)$.

Theorem 26. *Suppose that* $1 \geq> h_o > h_1 \geq \cdots \geq h_n \geq -1$. *Then*

$$T_n(h_o, \ldots, h_n) \geq 2^{n-1}$$

and the sign = holds if and only if the h_i *are the Chebyshev points.*

Proof. The polynomial $q_n(x)$ of degree n which takes the value $(-1)^i$ at $h_i(i = 0, \ldots, n)$ is

$$q_n(x) = \varphi(x) \sum_{i=0}^{n} \frac{(-1)^i}{\varphi'(h_i)\,(x - h_i)}.$$

□

Then $q_n(x) = a_n x^n + \cdots + a_o$, where

$$a_n = \sum_{i=o}^{n} \frac{1}{|\varphi'(h_i)|} = T_n(h_o, \ldots, h_n).$$

Write $t_n(x) = 2^{-n+1} a_n \cos(n \text{ arc } \cos x)$.

Then $q_n(x) - t_n(x)$ has degree $n - 1$ at most.

If $a_n < 2^{n-1}$, then $|t_n(x)| < 1$ and $q_n(x) - t_n(x)$ has the sign of $q_n(x)$ for the $n + 1$ values $h_o, \ldots, h_1$. If $a_n = 2^{n-1}$ the same is true on the understanding that $q_n(x) - t_n(x)$ may vanish for any of these values. So the polynomial $q_n(x) - t_n(x)$, of degree at most $n - 1$, has n zeros. This is a contradiction if $a_n < 2^{n-1}$, and is only possible for $a_n = 2^{n-1}$ when $q_n(x) \equiv t_n(x)$. This proves the theorem.

Corollary. *For* $1 \geq h_o \geq \cdots \geq h_n \geq -1, H_n \leq 2$.

Bibliography

[1] Achieser,N.I.- (English by Hyman), Theory of Approximation, New Yor, 1956. 58

[2] Bertstein,S.- Lecons sur les properties extremales et la meilleure approximation des fucntions analytiques *d′une* variables reelle, paris, 1926

[3] Jackson,D.- The theory of approximation, (Ameriacan Mathematical Society Colloquium, *XI* New York, 1930.

[4] Natanson,I,P. - (German by Bogel), Konstruktive Funktionentheorie, Berlin, 1955.

[5] de la Vallee Poussin, G.J.- Lecons sur 1' approximation ses functions dune variable reelle, Paris, 1919.

APPENDIX
Approximation by Polynomials in the Complex Domain

1 Runge's Theorem

The problem considered till now was the approximation of a given continuous function on a finite closed interval by polynomials in a real variable. Even for functions of two variables, we considered only the problem of approximation by polynomials in two independent real variables x, y. in what follows, we shall consider the approximation of a function in a domain in the plane (open connected set) by polynomials in the complex variable $z = x + iy$ (which are analytic functions of the variable z). **59**

Let $p_n(z)$ be a sequence of polynomials and suppose that G (which we assume is not empty) is the largest open set in which $p_n(z)$ converges, *uniformly on every compact subset*. (This is the type of approximation we shall consider; the problem of approximation on closed sets more difficult). By Weierstrass's theorem the limit of the sequence $p_n(z)$ is an analytic function in G. There is moreover, a purely topological restriction on G, viz., every connected component D of G is simply connected: for, if C is a sample closed curve contained in D and B is its interior, then

(maximum modulus principle)

$$\sup_{z\in B\cup C} |p_n(z) - p_m(z)| = \sup_{z\in C} |p_n(z) - p_m(z)| \to \text{ as } 0n, m \to \infty$$

so that the sequence $p_n(z)$ converges uniformly on $B\cup C$. Hence $B\cup C \subset G$ and since D is a connected component and $C \subset D, B \subset D$.

60 The main theorem, which is the analogue of Weierstrass's approximation theorem ($Th.1, p.2$) and which includes a converse of the remarks made above, runs as follows.

Theorem A (Runge). *Let D be a domain in the plane and f an analytic function in D. Then f can be approximated, uniformly on every compact subset of D, by rational functions whose poles lie outside D. If D is simply connected, f may be approximated by polynomials.*

We begin by funding a sequence of open regions $G_n, n = 1, 2, 3, \ldots$ bounded by polygons such that G_n is relatively compact in G_{n+1}, whose limit is D. We may take G_n as a subsequence of the sequence B_m, where B_m is defined to be the interior of the union of those squares $\frac{k}{2^m} \leq x \leq \frac{k+1}{2^m}, \frac{1}{2^m} \leq y \leq \frac{l+1}{2^m}, k, 1$ integers, $|k|, |l| \leq 2^{2^m}$, which lie in D. The boundary of G_n can be split into a finite number of simple closed polygons $C_{n,k}$ can be joined by a simple are which does not meet G_n to a point on the boundary of D. If $C_n = \bigcup_k C_{n,k}$ is the boundary of G_n we have

$$f(z) = \frac{1}{2\pi i} \int_{C_n} \frac{f(t)}{t-z} dt, z \in G_n.$$

By the definition of the integral, we may approximate $f(z)$ uniformly in G_{n-1} by finite sums of the form

$$f_n(z) = \frac{1}{2\pi i} \sum \frac{f(t_r)}{t_r - z}(t_{r+1} - t_r)$$

where t_r are certain points on C_n. Hence if $\varepsilon_n \downarrow 0$, we can find a sequence of rational functions R_n such that R_n has poles at most on C_n and

(1) $$|f(z) - R_n(z)| < \varepsilon_n \text{ in } G_{n-1}.$$

The main idea in the proof of the theorem is contained in the next step, which we state as a separate lemma. 61

Lemma. *Let C be a simple are joining the points z_o and z_1 and K a compact set not meeting C. Then, any rational function whose only possible pole is at z_o can be approximated, uniformly on K, by rational functions which no poles expect possibly at z_1.*

Proof of the lemma. Let $\varepsilon > 0$ be given and $2d$ be the distance between C and K. We find points $a_o, \ldots, a_\mu$ on $C, a_o = z_o, a_\mu = z_1$ such that $|a_{k+1} - a_k| \leq d$. Let $R(z)$ be the given rational function. There are two polynomials p and q so that

$$R(z) = p(z) + q\left(\frac{1}{z - z_o}\right)$$

and we have only to approximate $q\left(\frac{1}{z - z_o}\right) = f(z)$. Since $f(z)$ is analytic in $|z - a_1| > d$ and is finite at ∞, the Laurent expansion of $f(z)$ about a_1 contains no positive powers of $z - a_1$ and converges uniformly on every compact subset of $|z - a_1| > d$. A suitable partial sum then gives us a polynomial p_1 with $\left|f(z) - p_1\left(\frac{1}{z - a_1}\right)\right| < \frac{\varepsilon}{\mu + 1}$ for $z \in K$. Repeating this process, we find successively, polynomials p_j, $j = 1, \ldots, \mu$ with

$$\left|p_j\left(\frac{1}{z - a_j}\right) - p_{j+1}\left(\frac{1}{z - a_{j+1}}\right)\right| < \frac{\varepsilon}{\mu + 1} \text{ for } z \in K.$$

Then clearly

$$\left|f(z) - p_\mu\left(\frac{1}{z - a_\mu}\right)\right| < \varepsilon, z \in K$$

and the lemma follows.

Proof of Theorem. Let D be any domain and f analytic in D. Let G_n be the sequence of regions exhausting D described above. There is a 62
rational function r_n with poles at most on C_{n+1} such that

$$\left|f(z) - r_n(z)\right| < \frac{1}{2n} \text{ on } G_n.$$

Every point of the boundary of G_{n+1} can be joined be an arc not meeting $\overline{G}_n$ to the boundary of D, so that, by the lemma, there is a rational function R_n with poles outside D such that $\left|R_n(z) - r_n(z)\right| < \dfrac{1}{2n}$ on G_n and $|f(z) - R_n(z)| < \dfrac{1}{n}$ on G_n. The first part of Runge's theorem is proved. If D is simply connected, then every connected component of the complement is unbounded (unless D is the whole plane in which case the theorem is trivial). Hence every point of the boundary of G_{n+1} can be joined to a point $z_1(|z_1| \geq 2r)$ by an arc which does not meet $\overline{G}_n$, r being such that G_n is contained in the circle $|z| < r$.

Now it follows as above that there is a rational function $R_n(z)$ with all its poles lying in $|z| \geq 2r$, with

$$\left|f(z) - R_n(z)\right| < \frac{1}{2n} \text{ on } G_n.$$

If we expand $R_n(z)$ in a Taylor series about $z = 0$ (which converges uniformly for $|z| \leq r$), then a suitable partial sum $p_n(z)$ satisfies

$$\left|R_n(z) - p_n(z)\right| < \frac{1}{2n} \text{ on } G_n$$

so that

$$\left|f(z) - p_n(z)\right| < \frac{1}{n} \text{ on } G_n.$$

This complete the proof of Runge's theorem

The same argument proves the following theorem

THEOREM A[1]**.** *Let D be any plane domain. From each connected component of the complement of D, choose a point z_α. Then any analytic*
63 *function in D can be approximated uniformly on every compact set in D by rational functions which have poles at most at the points z_α.*

Runge's theorem is of importance in the theory of functions. As an instance of its applicability we prove the following extension to an arbitrary of Mittag-Leffler's theorem

THEOREM. *Let D be a plane domain and a_ν, $\nu = 1, 2, \ldots$ sequence of points in D having no limit point in D. Let $p_\nu$be polynomials (without* *constant term). Then there is a meromorphic function f in D with poles at most at the a_ν such that $f(z) - p_\nu\left(\frac{1}{z-a_\nu}\right)$ is analytic at a_ν.*

Proof. We can construct a sequence $G_n, n = 1, 2, \ldots$ of regions so that G_n is relatively compact in G_{n+1}, $\bigcup_n G_n = D$ and so that any point of the boundary of G_n can be joined to a point not in D by an arc not meeting G_n. Let

$$f_n(z) = \sum_{a_\nu \in G_n} p_\nu\left(\frac{1}{z-a_\nu}\right)$$

the sum being over those (finitely many) a_ν which lie in G_n. Since $f_{n+2} - f_{n+1}$ is analytic in G_{n+1}, we can find a rational function R_{n+1} with no poles in D such that

$$\left| f_{n+2}(z) - f_{n+1}(z) - R_{n+1}(z) \right| < \frac{1}{2^n} \text{ for } z \text{ in } G_n.$$

□

Since the poles of the R_{n+1} lie outside D and the series

$$\sum_{n=n_o+1}^{\infty} (f_{n+1} - f_n - R_n)$$

converges uniformly in G_{n_o}, it follows easily that we may take **64**

$$f(z) = f_2(z) + \sum_{n=2}^{\infty} (f_{n+1}(z) - f_n(z) - R_n(z)).$$

2 Interpolation

For functions f of a real variable, if p_n is the (Lagrange) polynomial p of degree n which agrees with f at $n+1$ equally spaces points on an interval, the sequence p_n in general diverges as $n \to \infty$. The behaviour for functions of a complex variable is more satisfactory. We proceed to prove two of the main theorems.

Let C be a simple closed rectifiable curve and $f(z)$ a function analytic inside and on C. Let $t_1, \ldots, t_{n+1}$ be $n+1$ points inside C (not necessarily distinct). Then the polynomial $p_n(z)$ of degree n such that $p_n(t_i) = f(t_i), i = 1, \ldots, n+1$ (multiplicity being taken into account if some of the t_i coincide) is easily seen to be given by

$$f(z) - p_n(z) = \frac{1}{2\pi i} \int\limits_C \frac{(z-t_1)\cdots(z-t_{n+1})}{(t-t_1)\cdots(t-t_{n+1})} \frac{f(t)}{t-z} dt.$$

Our first theorem is as follows. It is also due to Runge.

THEOREM B. *Let $f(z)$ be analytic for $|z| < R, R > 1$ and $p_n(z)$ the polynomial of degree n with $p_n(z_i) = f(z_i), i = 0, 1, \ldots, n$, where the z_i are the $(n+1)^{th}$ roots of unity.*

Then

$$p_n(z) \to f(z) \text{ as } n \to \infty$$

uniformly for $|z| \leq \varrho < R$.

65 *Proof.* Let C be the circle $|z| = \rho', \rho' > \rho, \rho > 1$. Then, for $|z| \leq \rho$,

$$f(z) - p_n(z) = \frac{1}{2\pi i} \int_c \frac{z^{n+1} - 1}{t^{n+1} - 1} \frac{f(t)}{t-z} dt,$$

so that

$$\begin{aligned} |f(z) - p_n(z)| &= \frac{1}{2\pi} \left| \int_c \frac{z^{n+1} - 1}{t_{n+1} - 1} \frac{f(t)}{t-z} dt \right| \\ &\leq \frac{1 + \rho^{n+1}}{(\rho'^{n+1} - 1)(\rho' - \rho)} M \ (M = \sup_{z \in C} |f(z)|) \\ &\to 0 \text{ as } n \to \infty \text{ since } \rho' > \rho, \rho' > 1. \end{aligned}$$

□

The next theorem is due to Fejer and is considerably deeper; it contains Theorem *B*.

Let C be a simple closed curve and suppose that $w(z)$ maps the exterior of C one-one conformally onto $|w| > 1$ in such a way that the points at infinity correspond. Then, as is well known, $w(z)$ is one-one continuous on C. Let $\alpha_i^{(n)}, i = 0, 1, \ldots, n$ be the $n+1$ points of C corresponding to the $(n+1)^{th}$ roots of unity in the w-plane. Then we have

THEOREM C. *Let $f(z)$ be a function analytic inside and on C and $p_n(z)$ the polynomial of degree n which equals $f(z)$ at the points $\alpha_i^{(c)}$. Then $p_n(z) \to f(z)$, uniformly inside and on C.*

We begin with a lemma.

Lemma.

$$\lim_{n\to\infty} \prod_{i=0}^{n} \left| z - \alpha_i^{(n)} \right|^{\frac{1}{(n+1)}} = A|w(z)|.$$

uniformly on any compact set exterior to $C, A > 0$ being a constant (depending on C). **66**

Proof of the lemma. Let $z = z(w)$ be the inverse of $w = w(z)$ and let $w_o, \ldots, w_n$ be the $(n+1)^{th}$ roots of unity. We prove first that

1. $\lim\limits_{n\to\infty} \prod\limits_{i=0}^{n} \left| \frac{z(w)-z(w_i)}{w-w_i} \right|^{\frac{1}{(n+1)}} = A.$

The logarithm of the term on the left is

2. $\lim\limits_{n\to\infty} \frac{1}{n+1} \sum_{i=o}^{n} \log \left| \frac{z(w)-z(w_i)}{w-w_i} \right| = \frac{1}{2\pi} \int\limits_{o}^{2\pi} \log \left| \frac{z(w)-z(e^{i\theta})}{w-e^{i\theta}} \right| d\theta$

and the limit is uniform for w in a compact set in $|w| > 1$.

Now $\dfrac{z(w)-z(\zeta)}{w-\zeta}$ is an analytic function of ζ for $|\zeta| > 1$ and fixed w, including $\zeta = \infty, \zeta = w$. Hence the integral in (2) is equal to

$$\lim_{\zeta\to\infty} \log \left| \frac{z(w)-z(\zeta)}{w-\zeta} \right| = \log A, \text{ say}$$

(we have only to make the substitution $\zeta \to 1/\zeta$ and use the Poison integral).

From (1), it follows that

$$\lim_{n\to\infty} \prod_{i=o}^{n} \frac{|z(w)-z(w_i)|^{\frac{1}{(n+1)}}}{|w-w_i|^{1/(n+1)}} = \lim_{n\to\infty} \frac{\prod\limits_{i=0}^{n} |z(w)-z(w_i)|^{\frac{1}{(n+1)}}}{|w^{n+1}-1|^{1/(n+1)}}$$

$$= \frac{\lim\limits_{n\to\infty} \prod\limits_{i=0}^{n} |z(w) - z(w_i)^{\frac{1}{(n+1)}}}{|w|} = A$$

and the lemma follows on substituting $w = w(z)$.

67 **Proof of Theorem.** Let $C_R(R > 1)$ denote the image under $z(w)$ of the circle $|w| = R$; we can choose $R > 1$ such that f is analytic inside and on C_R. Let $1 < r_1 < r_2 < R$; and put

$$\pi_n(z) = \prod_{i=o}^{n} (z - \alpha_i^{(n)}).$$

We have

$$f(z) - p_n(z) = \frac{1}{2\pi i} \int\limits_{C_{r_2}} \frac{\pi_n(z)}{\pi_n(t)} \frac{f(t)}{t - z} dt.$$

If z is on C_{r_1} and t on C_{r_2},

$$\lim_{n\to\infty} \left|\frac{\pi_n(z)}{\pi_n(t)}\right|^{\frac{1}{n+1}} = \frac{r_1}{r_2} \text{ (by the lemma)}$$

so that
$$\overline{\lim_{n\to\infty}} \left[\sup_{z\in C_{r_1}} |f(z) - p_n(z)|\right]^{\frac{1}{n+1}} \leq \frac{r_1}{r_2} < 1.$$

Consequently $f(z) - p_n(z) \to o$ uniformly for z on C_{r_1}.

The theorem follows at once from the maximal modulus principle.

3 Best Approximation

In this section we shall consider the problem of best approximation.

68 Let K be a compact set containing infinitely many points and $f(z)$ a *continuous* function on K. Our aim is to prove the existence and uniqueness of a polynomial $p_n(z)$ of degree n such that

$$d(f, p_n) = \sup_{z\in K} \left| f(z) - p_n(z) \right|$$

is least. in general, of course, this minimum $d(f, p_n)$ does not tend to zero as $n \to \infty$.

Existence of a polynomial of best approximation.

Let P_n be the family of all polynomials of degree $\leq n$, and let $f(z)$ be a continuous function on the compact set K. Let

$$d(f) = d = \inf_{p \in P_n} d(f, p) = \inf_{p \in P_n} (\sup_{z \in K} |f(z) - p(z)|).$$

Then we have the

THEOREM D. *There exists a $p \in P_n$ with $d(f, p) = d$.*

Proof. Any polynomial $p \in P_n$ takes values 0 or 1 at n points at most. Hence P_n is a quasi-normal family of order n, (theorem of Montel, see [1] p. 67) i.e., given a sequence p_ν of polynomials in P_n, there is a subsequence p_{ν_k} and n points z_i such that p_{ν_k} converges, uniformly on every compact set not containing the z_i, either to a finite limit function or to ∞. In the first case it is clear that p_{ν_k} converges uniformly on any compact set (which may contain some of the z_i). □

There is a sequence $p^{(\nu)}$ of polynomials of P_n so that $d(f, p^{(\nu)}) \to d$. Then, clearly, if $z \in K, |p^{(\nu)}(z)| \leq d + 1 + \sup_{\zeta \in K} |f(\zeta)|$ for large ν (we may **69**
suppose that holds for all ν). Let $p^{(\nu_k)}$ be a subsequence converging outside n points z_i, uniformly on compact sets. Since K contains infinitely many points there are points of K not equal to any z_i, and it these points $z, |p^{(\nu_k)}(z)|$ is bounded. Hence the limit outside z_i is finite and consequently, $p^{(\nu_k)}$ converges uniformly on any compact set. From Cauchy's inequality, it follows then that the corresponding co-efficients of $p^{(\nu_k)}$ converge, so that $\lim_{k \to \infty} p^{(\nu_k)}(z) = p(z) \in P_n$. Then we have

$$d \leq d(f, p) \leq d(f, p^{(\nu_k)}) + d(p^{(\nu_k)}, p) \to d$$

so that $d(f, p) = d$.

[If K contains a circle $|z - a| < r, r > o$, the existence of a sequence $p^{(\nu_k)}$ converging uniformly on any compact set follows at once, as in the case $(Ch.II, Th.4, p.14)$ if we use the Cauchy inequalities.]

Uniqueness of the polynomial of best approximation.

We shall deduce the uniqueness fro the following theorem, as in the case of a real variable.

Let $p \in P_n$ satisfy $d(f, p) = d(f)$. Then $|f(z) - p(z)|$ attains its maximum at atleast $n + 2$ distinct points of K.

(The proof is similar in principle to the proof of $Th.5, p.14$).

Proof. Suppose that $f(z)-p(z) = g(z)$ attains its maximum modulus at m points $(m \leq n + 1)z_1, \ldots, z_m$ of K. Then, we can construct a polynomial
70 $q(z)$ of degree n such that $q(z_i) = g(z_i)$. Given $\varepsilon > o$, we can find $\delta > 0$ so that if

$$|\zeta_1 - \zeta_2| < \delta, |g(\zeta_1) - g(\zeta_2)| < \varepsilon, |q(\zeta_1) - q(\zeta_2)| < \varepsilon$$

□

Let K^1 be the set obtained from K by removing the points of the (open) discs $|z - z_i| < \delta$. Then

$$\sup_{z \in K^1} |g(z)| = d^1 < d = \sup_{z \in K} |g(z)|.$$

Let $1 > \eta > o$ be sufficiently small. Consider $g(z) - \eta q(z)$; then for $|z - z_i| < \delta, |g(z) - g(z_i)| < \varepsilon, |\eta q(z) - \eta q(z_i)| < \eta\varepsilon$, so that $|g(z) - \eta q(z)| = |\eta(g(z_i) - g(z)) + \eta g(z) - g(z) + \eta(q(z) - q(z_i))|$ (since $q(z_i) = g(z_i)$) $< \eta\varepsilon + \eta\varepsilon + (1 - \eta)d < d$ if $2\varepsilon < d$

If we choose η so small that

$$\sup_{z \in K^1} |g(z) - \eta q(z)| < d$$

we have
$$\sup_{z \in K} |g(z) - \eta q(z)| < d$$

and $d(f, p + \eta q) < d$, contradicting the definition of d.

THEOREM E. *The polynomial of degree $\leq n$ of best approximation is unique.*

Proof. Let $d(f, p) = d = d(f, q)$; let $r(z) = \dfrac{1}{2}(p(z) + q(z))$. □

Then

$$|f(z) - r(z)| = |\frac{1}{2}(f(z) - p(z)) + \frac{1}{2}(f(z) - q(z))| \le d$$

Let $z_1, \ldots, z_{n+2}$ be points at which $|f(z_i) - r(z_i)| = d$. Then, unless **71**
$f(z_i) - p(z_i) = f(z_i) - q(z_i) = w_i$ with $|w_i| = d$, $\frac{1}{2}|f(z_i) - p(z_i) + f(z_i) - q(z_i)| < d$. Hence, $p(z)$ and $q(z)$ take the same value at the $n + 2$ points z_i and since they are polynomials of degree n, $p(z) = q(z)$.

Bibliography

[1] Montel, P. : Lecons sur les familles normales de fonctions analytiques et leurs applications, Gauthiers-Villars, Paris, 1927

[2] Walsh,J.L. : Approximation by polynomials in the complex domain, Mem. des Sciences Mathem., Paris, 1935

[3] Walsh, J.L. : Interpolation and approximation by rational functions in the complex domain, Amer. Math. Soc. Colloquium Publications, $Vol.XX$, 1935.