АКАДЕМИЯ НАУК СССР

С.Н. БЕРНШТЕЙН

СОБРАНИЕ СОЧИНЕНИЙ

ТОМ
I

КОНСТРУКТИВНАЯ ТЕОРИЯ ФУНКЦИЙ
[1905-1930]

ИЗДАТЕЛЬСТВО
АКАДЕМИИ НАУК СССР
1952

3

О НАИЛУЧШЕМ ПРИБЛИЖЕНИИ НЕПРЕРЫВНЫХ ФУНКЦИЙ ПОСРЕДСТВОМ МНОГОЧЛЕНОВ ДАННОЙ СТЕПЕНИ *

ВВЕДЕНИЕ

Вопрос о приближении непрерывных функций посредством многочленов или других простых выражений определенного вида, равнозначный вопросу о разложении функций в соответствующие ряды, является основным в теории функций вещественной переменной. Я не буду излагать здесь истории этого вопроса, поучительной во многих отношениях; напомню лишь важнейшие ее моменты.

Теория разложений функций в ряды обязана своим возникновением задачам математической физики, которые великие геометры XVIII столетия пытались решать при помощи бесконечных рядов. Разумеется, в исследованиях этого времени, когда даже разница между сходящимися и расходящимися рядами была не ясна, о точности в современном смысле этого слова не может быть и речи. Только в первой половине XIX столетия Дирихле и Коши доказали сходимость некоторых разложений для весьма обширного класса функций и положили таким образом основу современной строго математической теории функций вещественной переменной.

Но прошло еще полстолетия, прежде чем Вейерштрасс в 1885 г. доказал, пользуясь одним интегралом из теории теплоты, что всякая непрерывная функция может быть разложена в равномерно сходящийся ряд многочленов, и вместе с тем указал прием, хотя и довольно сложный, для построения многочленов, сколь угодно мало отличающихся от данной произвольной непрерывной функции. Открытие этой замечательной по своей общности теоремы определило дальнейший ход развития анализа; с этого момента теория функций комплексной переменной, достигшая в то же время своего величайшего расцвета, постепенно

---

* «Сообщ. Харьк. матем. об-ва», серия 2, т. 13 (1912), стр. 49—194 (43*).

отходит на задний план, выдвигая вперед изучение функций вещественной переменной.

После Вейерштрасса многими математиками были предложены более или менее простые доказательства его теоремы[1], дающие возможность, при всяком значении ε, найти для данной на некотором отрезке AB непрерывной функции $f(x)$ приближенные многочлены $P_n(x)$ достаточно высокой степени n, чтобы уклонение $|f(x) - P_n(x)|$ оставалось не более ε на данном отрезке.

Сопоставление различных методов естественно выдвинуло задачу: каково для данной функции $f(x)$ наилучшее приближение, которого можно достигнуть при помощи многочленов данной степени, или, точнее говоря, каково наименьшее возможное значение $E_n[f(x)]$ уклонения ε при данном n?

Эта задача была поставлена П. Л. Чебышевым более пятидесяти лет тому назад, т. е. задолго еще до открытия Вейерштрасса. Оригинальный алгебраический метод великого русского математика привел его к весьма замечательным свойствам многочленов, наименее уклоняющихся от данной функции $f(x)$, и в некоторых частных случаях позволил ему дать полное решение задачи. Однако в общем случае мы не находим у Чебышева никаких указаний относительно величины наименьшего уклонения $E_n[f(x)]$, и этим главным образом объясняется, почему в свое время исследования Чебышева не оказали влияния на развитие теории функций.

Настоящее сочинение представляет собой попытку приближенного вычисления наименьшего уклонения $E_n[f(x)]$ и исследования связи между законом убывания $E_n[f(x)]$ и дифференциальными свойствами рассматриваемой функции. Чтобы можно было судить о том, насколько простой и глубокой оказывается эта связь, достаточно будет указать, например, два предложения[2]: для того чтобы функция вещественной переменной $f(x)$ была аналитической на некотором отрезке AB, необходимо и достаточно, чтобы наименьшее уклонение $E_n[f(x)]$ на отрезке AB убывало с возрастанием n быстрее, чем члены некоторой убывающей

---

[1] См. Borel, Leçons sur les fonctions de variables réelles et les développements en séries de polynomes. 1905.

[2] Эти предложения и несколько других были мною указаны в заметке, представленной Французской Академии наук 28 февраля 1911 г. [2]. Из предшествующих этой заметке работ в том же направлении следует указать важные сочинения Лебега и Валле Пуссена, на которые в соответствующих местах будут сделаны ссылки. Более подробные библиографические указания читатель найдет в работе Д. Джексона (D. Jackson) «Über die Genauigkeit der Annäherung stetiger Funktionen durch ganze rationale Funktionen» (Preisschrift und Inaugural-Dissertation), Гёттинген. Автор этой интересной работы, появившейся в июле 1911 г., получил самостоятельно некоторые из результатов моей заметки, которую он цитирует на стр. 12 и 15. Вместе с тем считаю нужным заметить, что настоящая моя работа, за исключением двух «Добавлений» к IV и V главам [3.1], представляет, с незначительными редакционными изменениями, перевод мемуара [3.2] под тем же заглавием, удостоенного премии Бельгийской академии, куда он был направлен мною в июне 1911 г.

геометрической прогрессии; для того чтобы функция $f(x)$ имела производные всех порядков, необходимо и достаточно, чтобы при всяком p имело место равенство $\lim\limits_{n\to\infty} E_n[f(x)]\, n^p = 0$. Вообще, чем проще дифференциальная природа функции, тем быстрее убывает E_n, и обратно.

Таким образом, рассмотрение наименьшей возможной погрешности при приближении функции посредством многочленов возрастающих степеней дает совершенно общее основание для последовательной классификации и исследования всех непрерывных функций вещественной переменной.

Часть первая

О НЕКОТОРЫХ ОБЩИХ СВОЙСТВАХ РЯДОВ МНОГОЧЛЕНОВ

Глава I

ПРЕДВАРИТЕЛЬНЫЕ ТЕОРЕМЫ О МНОГОЧЛЕНАХ

1. Многочлены, наименее уклоняющиеся от нуля. В своих знаменитых исследованиях о приближенных многочленах Чебышев построил многочлены, наименее уклоняющиеся от нуля в данном промежутке, а именно, он доказал, что из всех многочленов вида

$$Ax^n + p_1x^{n-1} + \ldots + p_n,$$

где A — данная величина, а остальные коэффициенты произвольны, наименее уклоняется от нуля в промежутке $(-h, +h)$ многочлен

$$\frac{Ah^n}{2^{n-1}} T_n\left(\frac{x}{h}\right) = \frac{Ah^n \cos n \arccos \frac{x}{h}}{2^{n-1}} = \frac{A}{2^n}\left[\left(x+\sqrt{x^2-h^2}\right)^n + \left(x-\sqrt{x^2-h^2}\right)^n\right]. \quad (1)$$

Для краткости мы будем в дальнейшем называть выражения вида $cT_n(x)$, где c — постоянная величина, *тригонометрическими многочленами* и выведем некоторые их свойства, аналогичные свойству, открытому Чебышевым.

2. Теорема. *Если многочлен* $P_n(x) = p_0x^n + p_1x^{n-1} + \ldots + p_n$ *обладает свойством, что* $|P_n'(x)\sqrt{1-x^2}|$ *достигает в промежутке* $(-1, +1)$ *значения* M, *то* $|P_n(x)|$ *не может в этом промежутке оставаться менее* M/n; *эта последняя величина не будет превзойдена лишь в случае, когда* $P_n(x)$ — *тригонометрический многочлен.*

Чебышев допускал без доказательства существование многочленов данной степени, наименее уклоняющихся от данной функции. Но современный анализ требует этого доказательства, так как немало есть задач о минимуме, например в вариационном исчислении, которые не имеют решений. Ввиду этого нам необходимо сделать несколько предварительных замечаний, для того чтобы показать, что среди рассматриваемых многочленов существует действительно один или несколько таких многочленов, для которых максимум $|P_n(x)|$ *достигает* наименьшего возможного значения.

Рассмотрим вообще произведение $|P'_n(x)\varphi(x)|$, где $\varphi(x)$ — какая-нибудь непрерывная функция (голоморфная при всех значениях x данного промежутка, кроме тех, может быть, где $\varphi(x)=0$). Максимум этого произведения $m(p_0, p_1, \ldots, p_{n-1})$ есть непрерывная однородная функция первой степени коэффициентов $p_0, p_1, \ldots, p_{n-1}$, так как при умножении их на одно и то же число k, m будет умножено на то же число k. Значения коэффициентов, удовлетворяющие уравнению $m=M$, где M — данная величина, можно разделить на n групп: в первой $|p_0| \geqslant |p_i|$, во второй $|p_1| \geqslant |p_i|$ и т. д. $(i=0, 1, \ldots, n-1)$.

Рассмотрим, например, значения первой группы; в данном случае уравнение $m=M$ можем написать так:

$$p_0 \cdot m\left(1, \frac{p_1}{p_0}, \frac{p_2}{p_0}, \ldots, \frac{p_{n-1}}{p_0}\right)=M,$$

или, полагая $\frac{p_1}{p_0}=\lambda_1$, $\frac{p_2}{p_0}=\lambda_2$ и т. д.,

$$p_0=\frac{M}{m(1, \lambda_1, \lambda_2, \ldots, \lambda_{n-1})}.$$

Таким образом, p_0 есть конечная и непрерывная функция переменных $\lambda_1, \lambda_2, \ldots, \lambda_{n-1}$, которые по абсолютному значению не превышают единицы; поэтому максимум $|p_0(x^n+\lambda_1 x^{n-1}+\ldots+\lambda_{n-1}x)+p_n|=|P_n(x)$ есть непрерывная функция $\lambda_1, \lambda_2, \ldots, \lambda_{n-1}, p_n$; при этом, очевидно,| можно ограничиться рассмотрением значений $|p_n|$, не превышающих некоторого числа H. Но непрерывная функция n переменных $\lambda_1, \lambda_2, \ldots$ $\ldots, \lambda_{n-1}, p_n$, принимающих всевозможные значения из некоторой замкнутой области, достигает своего минимума для определенных значений переменных в этой области. Аналогичным образом можно доказать существование многочленов, наименее уклоняющихся от нуля, соответствующих каждой из n групп коэффициентов. Выбирая ту из групп, которая дает наименьшее значение для максимума $|P_n(x)|$, мы убеждаемся, наконец, что среди многочленов, для которых $m=M$, действительно, есть один или несколько таких, которых максимум $|P_n(x)|$ равен наименьшему возможному значению.

Итак, пусть $P(x)$ будет тот из подлежащих сравнению многочленов степени n, который наименее уклоняется от нуля. Обозначим через $x_1, x_2, \ldots, x_k$ точки, в которых модуль $P(x)$ получает наибольшее значение L, и через ζ — ту точку, где $P'(x)\varphi(x)$ достигает максимума M.

Я говорю, что нельзя найти такого многочлена $F_n(x)$ степени n, который удовлетворял бы уравнениям

$$F_n(x_1)=P(x_1), \quad F_n(x_2)=P(x_2), \ldots, \quad F_n(x_k)=P(x_k), \quad F'_n(\zeta)\varphi(\zeta)=0. \quad (2)$$

В самом деле, если бы равенства (2) были осуществлены, то можно было бы построить многочлен $P-\lambda F_n$ степени не выше n, выбрав положительное число λ следующим образом. Окружим точки $x_1, x_2, \ldots, x_k$

промежутками достаточно малыми, чтобы $P(x)$ и $F_n(x)$ сохраняли в каждом из них тот же самый знак, и отнимем эти промежутки из отрезка $[-1, +1]$; тогда в оставшейся части отрезка будем иметь $|P(x)| < L - \delta$, где δ — некоторое определенное положительное число (меньшее, если хотим, чем $L/2$); после этого мы выберем положительное количество λ настолько малым, чтобы было $\lambda|F_n(x)| < \delta$. В таком случае оказалось бы, что многочлен $P - \lambda F_n$ по абсолютному значению всегда менее (и никогда не равен) L, так как в отнятых промежутках $|P - \lambda F_n| < |P| \leqslant L$, а в оставшейся части отрезка $|P - \lambda F_n| < (L - \delta) + \delta = L$, причем $[P'(\zeta) - \lambda F'_n(\zeta)]\,\varphi(\zeta) = M$. Поэтому, обозначая через M_1 $(M_1 \geqslant M)$ максимум $|[P'(x) - \lambda F'_n(x)]\,\varphi(x)|$, убеждаемся, что многочлен $[P(x) - \lambda F_n(x)]\frac{M}{M_1}$ подлежал бы сравнению, уклоняясь от нуля менее, чем $P(x)$, что противоречило бы нашему допущению, что среди подлежащих сравнению многочленов нет такого, который уклоняется от нуля менее, чем $P(x)$.

Следовательно, система уравнений (2) не имеет решения, а потому либо число уравнений $(k+1)$ больше числа неизвестных коэффициентов $(n+1)$, т. е. $k > n$, либо $k \leqslant n$ и все определители $(k+1)$-го порядка матрицы

$$\left\|\begin{matrix} x_1^n & x_1^{n-1} \ldots x_1 & 1 \\ \cdot\;\cdot\;\cdot & \cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot & \cdot \\ \cdot\;\cdot\;\cdot & \cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot & \cdot \\ \cdot\;\cdot\;\cdot & \cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot & \cdot \\ x_k^n & x_k^{n-1} \ldots x_k & 1 \\ n\zeta^{n-1} & \ldots \quad 1 & 0 \end{matrix}\right\|$$

равны нулю (так как, очевидно, $\varphi(\zeta) \gtreqless 0$).

В первом случае $P(x)$ есть тригонометрический многочлен. В самом деле, так как степень многочлена $P(x)$ равна n, то во всяком случае $k \leqslant n+1$; поэтому при допущении, что $k > n$, находим $k = n+1$. Таким образом, из значений $x_1, x_2, \ldots, x_k$ два равны $+1$ и -1, а остальные суть $n-1$ корней уравнения $P'(x) = 0$. Так как, с другой стороны, все эти значения обращают в нуль $P^2(x) - L^2$, то все корни $P'(x) = 0$ суть двойные корни уравнения $P^2(x) - L^2 = 0$, имеющего еще всего два простых корня $+1$ и -1. Отсюда выводим дифференциальное уравнение Чебышева

$$P^2(x) - L^2 = \frac{(x^2-1)\,[P'(x)]^2}{n^2}, \tag{3}$$

единственным рациональным решением которого служит $L \cos n \arccos x$. Следовательно, $P(x) = L \cos n \arccos x$.

Во втором случае $k = n$. В самом деле, если бы было $k < n$, то $P(x) + (ax+b)R(x)$, где $R(x) = (x-x_1)(x-x_2)\ldots(x-x_k)$, был бы многочленом степени не выше n. Но, полагая $F_n(x) = P(x) + (ax+b)R(x)$, мы можем, очевидно, выбрать коэффициенты a и b так, чтобы все уравнения (2) были удовлетворены; для этого достаточно удовлетворить уравнению

$$P'(\zeta) + aR(\zeta) + (a\zeta + b)R'(\zeta) = 0, \tag{2'}$$

к которому приводится последнее из уравнений (2), между тем как первые k уравнений удовлетворены тождественно. Уравнение же (2′) всегда разрешимо, ибо не может быть одновременно $R'(\zeta)=0$ и $R(\zeta)=0$. Но так как, по доказанному, уравнения (2) несовместимы, следовательно $k=n$.

Однако, как мы увидим, для функций $\varphi(x)$, рассматриваемых нами, второй случай вообще не может представиться. Для этого перейдем к следствиям, вытекающим из предположения, что $k=n$.

Прежде всего мы замечаем, полагая $F_n(x)=P(x)+bR(x)$, что уравнения (2) приводятся к единственному уравнению $P'(\zeta)+bR'(\zeta)=0$, которое будет неразрешимо лишь в случае, когда $R'(\zeta)=0$. Таким образом, ζ есть корень уравнения

$$R'(x)=0.$$

Но

$$R(x)=C\frac{(x^2-1)\,P'(x)}{x-\beta},$$

где C — постоянный множитель, β — тот из корней уравнения

$$(x^2-1)\,P'(x)=0,$$

которого нехватает уравнению $R(x)=(x-x_1)\dots(x-x_k)=0$. Поэтому ζ удовлетворяет одновременно уравнениям

$$\frac{d}{dx}\left[\frac{(x^2-1)\,P'(x)}{x-\beta}\right]=0 \text{ и } (x^2-1)\frac{d}{dx}[P'(x)\,\varphi(x)]=0. \tag{4}$$

Легко обнаруживается несовместность этих уравнений, если $\varphi(x)=1-x^2$. Тогда, очевидно, $\zeta^2-1<0$, так что второе уравнение обращается в

$$\frac{d}{dx}[P'(x)(1-x^2)]=0,$$

вследствие чего первое уравнение приводится к $P'(x)=0$, что невозможно, так как $|P'(x)(1-x^2)|$ при $x=\zeta$, по предположению, достигает своего наибольшего значения M.

Докажем, что случай $k=n$ также не представляется, если $\varphi(x)=\sqrt{1-x^2}$, как это имеет место в условии теоремы. Если мы положим $P'(x)\sqrt{1-x^2}=P_1(x)$, то уравнения (4) примут форму

$$\frac{d}{dx}\left[P_1\frac{\sqrt{1-x^2}}{x-\beta}\right]=0, \qquad \frac{dP_1}{dx}=0,$$

или

$$\frac{dP_1}{dx}(1-x^2)(x-\beta)+P_1\cdot(\beta x-1)=0, \qquad \frac{dP_1}{dx}=0,$$

откуда

$$\beta x-1=0,$$

поэтому $\zeta=\frac{1}{\beta}$. И так как $|\zeta|<1$, то, следовательно,

$$|\beta|>1.$$

С другой стороны, легко убедиться, что $P(x)$ удовлетворяет дифференциальному уравнению вида

$$P^2 - L^2 = \frac{(P')^2 (x^2 - 1)(x^2 + bx + c)}{n^2 (x - \beta)^2}. \tag{5}$$

Действительно, многочлен $P^2 - L^2$ степени $2n$ имеет двойными корнями те из значений $x_1, x_2, \ldots, x_n$, которые отличны от ± 1 (так как они обращают в нуль P'), и простыми корнями те из значений, которые равны ± 1. Поэтому $P^2 - L^2$ делится на многочлен $(2n - 2)$-й степени $\frac{(P')^2 (x^2 - 1)}{(x - \beta)^2}$, и так как коэффициент первого члена делимого в n^2 раз меньше коэффициента первого члена делителя, то частное имеет форму $\frac{x^2 + bx + c}{n^2}$, откуда вытекает уравнение (5).

Я говорю, что корни уравнения

$$x^2 + bx + c = 0$$

вещественны, имеют тот же знак, что и β, и больше этого числа по абсолютному значению. В самом деле, допустим для определенности, что $\beta > 0$; в таком случае $\beta > 1$. Если x, возрастая от единицы, достигает значения β, где P' обращается в нуль, P^2 возрастает от L^2 до некоторого числа L_1^2, затем P^2 убывает; но так как P' более не меняет знака, то P^2, пройдя, при $x = \gamma > \beta$, через значение L^2, обращается в нуль, и после этого возрастает до бесконечности, проходя снова через значение L^2 при $x = \delta > \gamma > \beta$. Очевидно, что γ и δ суть корни уравнения $x^2 + bx + c = 0$. Итак уравнение (5) можем написать в виде

$$P^2 - L^2 = \frac{(x^2 - 1)(x - \gamma)(x - \delta)}{n^2 (x - \beta)^2} (P')^2, \tag{6}$$

причем $\gamma > \beta > 0$ и $\delta > \beta > 0$. (То же самое рассуждение привело бы, при $\beta < 0$, к неравенствам $\gamma < \beta < 0$ и $\delta < \beta < 0$.)

Следовательно, для $|x| < 1$ имеем

$$\theta^2 (L^2 - P^2) = \frac{(1 - x^2)(P')^2}{n^2}, \tag{6'}$$

где $\theta < 1$; поэтому

$$|P' \sqrt{1 - x^2}| \leqslant n\theta L.$$

Таким образом, если бы k было равно n, то, несомненно, наибольшее значение L модуля $P(x)$ удовлетворяло бы неравенству $L \geqslant \frac{M}{n\theta} > \frac{M}{n}$. Напротив, при $k = n + 1$ мы нашли, что $P(x) = L \cos n \arccos x$, откуда $|P' \sqrt{1 - x^2}| = Ln |\sin n \arccos x|$, так что в этом случае $L = \frac{M}{n}$. Следовательно, только случай, когда $P(x)$ есть тригонометрический многочлен, приводит к наименьшему значению для L, причем $L = \frac{M}{n}$, что и требовалось доказать.

3. Следствия. а) *Если на отрезке* $[-h, +h]$ *произведение* $|P_n'(x) \sqrt{h^2 - x^2}|$ *достигает значения* M, *то, при предположении, что*

$P_n(x)$ есть многочлен степени n, $|P_n(x)|$ не может на рассматриваемом отрезке оставаться менее M/n.

В самом деле, положим $x = hx_1$. В таком случае

$$P_n(x) = P_n(hx_1) = Q_n(x_1) \text{ и } P_n'(x)\sqrt{h^2 - x^2} = Q_n'(x_1)\sqrt{1 - x_1^2}.$$

Применяя к $Q_n(x_1)$ только что доказанную теорему, заключаем, что, так как на отрезке $[-1, +1]$ выражение $|Q_n'(x_1)\sqrt{1 - x_1^2}|$ достигает значения M, то, следовательно, $|Q_n(x_1)|$ на том же отрезке $[-1, +1]$, а $|P_n(x)|$ на отрезке $[-h, +h]$ не может оставаться меньше M/n.

б) *Если на отрезке $[a, b]$ произведение $|P_n'\sqrt{(a-x)(x-b)}|$ достигает значения M, то $|P_n(x)|$ на этом отрезке не остается менее M/n.*

Это вытекает из доказанного следствия, если положим $x_1 = x - \frac{a+b}{2}$.

в) *Если $|P_n(x)| \leqslant L$ на отрезке $[a, b]$, то на том же отрезке $|P_n'(x)\sqrt{(a-x)(x-b)}| \leqslant nL$.*

В самом деле, если бы $|P_n'(x)\sqrt{(a-x)(x-b)}|$ достигало значения $M = nL + \varepsilon$, то $|P_n(x)|$, в силу предыдущего следствия, получало бы значение $\frac{M}{n} = L + \frac{\varepsilon}{n}$, что противоречит условию.

4. Теорема А. А. Маркова[1]. *Многочлен n-ой степени $P_n(x)$ на отрезке $[-1, +1]$ не остается менее M/n^2 по абсолютному значению, если на том же отрезке $|P_n'(x)|$ достигает M.*

Очевидно, что вся первая часть доказательства теоремы § 2 (до специализации функции $\varphi(x)$) остается в силе. В данном случае мы должны положить в уравнениях (4) $\varphi(x) = 1$; таким образом, если многочлен $P(x)$, дающий наименьшее отклонение L, — не тригонометрический многочлен и достигает максимального отклонения только в n точках, то значение ζ, при котором $|P'(x)|$ достигает максимума M, удовлетворяет уравнениям

$$\frac{dP'}{dx}(x^2-1)(x-\beta) + P'[2x(x-\beta) - (x^2-1)] = 0; \qquad (x^2-1)\frac{dP'}{dx} = 0. \quad (4')$$

Следовательно, либо $\zeta = \pm 1$, тогда $\beta = \zeta$; либо $|\zeta| < 1$, тогда

$$\zeta^2 - 2\beta\zeta + 1 = 0,$$

так что $|\beta| > 1$.

В первом случае, полагая для определенности $\beta = \zeta = 1$, увидим, что наибольшее значение $|P(x)|$ достигается в $n - 1$ внутренних точках, где $P'(x) = 0$, и в точке $x = -1$. Поэтому $P(x)$ будет удовлетворять дифференциальному уравнению

$$P^2 - L^2 = \frac{(1+x)(x-\alpha)}{n^2}(P')^2, \tag{7}$$

---

[1] А. Марков. Об одном вопросе Д. И. Менделеева. 1889. Из доказательства А. А. Маркова вытекает в сущности также и теорема § 2, хотя она и не формулирована в упомянутой статье.

причем $\alpha > 1$, так как, при $x = 1$, $P^2 < L^2$. Следовательно,

$$P = L \cos n \arccos \frac{2x + \alpha - 1}{\alpha + 1}.$$

Во втором случае многочлен P также должен удовлетворять уравнению (6) с соблюдением тех же неравенств относительно β, γ, δ. Поэтому попрежнему

$$\theta^2 (L^2 - P^2) = \frac{(1 - x^2)(P')^2}{n^2}, \tag{6'}$$

где $\theta < 1$.

Наконец, в случае, когда $k = n + 1$, $P(x)$ есть тригонометрический многочлен, удовлетворяющий, как мы видели, уравнению (3), которое можно получить из уравнения (6'), полагая в последнем $\theta = 1$.

Введем новые переменные, определяемые уравнениями

$$P = \pm L \cos z; \qquad x = \cos t$$

(знак + возьмем, если, при $x = 1$, будет $P = L$, в противном случае возьмем знак —); тогда уравнение (6') преобразуется в

$$\theta^2 L^2 \sin^2 z = \frac{L^2 \sin^2 z}{n^2} \left(\frac{dz}{dt}\right)^2,$$

откуда $\left|\frac{dz}{dt}\right| = n\theta$. Так как, при $x = 1$, $P = \pm L$, то можно положить $z = 0$ при $t = 0$. Следовательно, $z = n\theta_1 t$, где $|\theta_1| < 1$ для уравнения (6') и $\theta_1 = 1$ для уравнения (3). Откуда

$$\theta^2 L^2 \sin^2 n\theta_1 t = \frac{\sin^2 t}{n^2} (P')^2;$$

поэтому

$$L = \left|\frac{\sin t}{\theta n \sin n\theta_1 t} P'\right| > \frac{M}{n^2},$$

если $\theta < 1$, $|\theta_1| < 1$, и

$$L = \frac{M}{n^2},$$

если $\theta = \theta_1 = 1$, так как $|P'|$, очевидно, принимает наибольшее значение M, когда $\left|\frac{\sin t}{\theta n \sin n\, \theta_1 t}\right|$ получает наименьшее значение (которое больше, чем $1/n^2$, в первом случае, и равно $1/n^2$ во втором случае). Итак, отклонение L тригонометрического многочлена при том же M было бы менее отклонения многочлена $P(x)$, удовлетворяющего уравнению (6'); а потому $P(x)$ не может удовлетворять и уравнению (7), ибо в этом случае было бы

$$P(x) = L \cos n \arccos \frac{2x + \alpha - 1}{\alpha + 1},$$

$$P'(x) = \frac{2nL \sin n \arccos \frac{2x + \alpha - 1}{\alpha + 1}}{(\alpha + 1) \sin n \arccos \frac{2x + \alpha - 1}{\alpha + 1}},$$

откуда получилось бы $M < n^2 L$, так как $\alpha > 1$.

Таким образом, $|P_n(x)|$ остается возможно малым, если $P_n(x)$ — тригонометрический многочлен; но даже в этом случае многочлен $P(x)$ достигает абсолютного значения $L = \frac{M}{n^2}$, что и требовалось доказать.

5. Следствия. а) *Из всех многочленов степени n, производная которых достигает данного абсолютного значения на отрезке $[-1, +1]$, наименее уклоняется от нуля на этом отрезке тригонометрический многочлен.*

б) *Если на отрезке $[a, b]$ производная многочлена n-й степени $P_n(x)$ достигает абсолютного значения M, то $|P_n(x)|$ на этом отрезке не остается менее $\frac{|b-a|M}{2n^2}$.*

Для того чтобы в этом убедиться, достаточно сделать линейное преобразование $x = \frac{b-a}{2}x_1 + \frac{b+a}{2}$.

в) *Если*[1] *на отрезке $[a, b]$ многочлен n-й степени $P_n(x)$ не превышает по абсолютному значению L, то $|P_n'(x)|$ на том же отрезке не превышает $\frac{2n^2L}{b-a}$.*

г) *Если на отрезке $[a, b]$ $|P_n(x)|$ не превышает L, то $\left|\frac{d^k P_n(x)}{dx^k}\right|$ не превышает $\left(\frac{2}{b-a}\right)^k n^2(n-1)^2\ldots(n-k+1)^2 L$ на том же отрезке*[2].

Это вытекает из k-кратного повторения предыдущего следствия.

6. Теорема. *Из всех многочленов степени n, принимающих в данной вещественной точке, не лежащей на отрезке $[-1, +1]$, абсолютное значение M, наименее уклоняется от нуля на этом отрезке тригонометрический многочлен* *.

В самом деле, посредством соображений, совершенно аналогичных приведенным при доказательстве теоремы § 2, убеждаемся, что среди многочленов, подлежащих рассмотрению, существует такой $P(x)$, который достигает наименьшего отклонения L. Обозначая через $x_1, x_2, \ldots, x_n$

---

[1] Это есть формулировка теоремы А. А. Маркова, данная им в выше упомянутой статье; к сожалению, с этой работой, так же, как и с сочинением В. А. Маркова «О функциях, наименее уклоняющихся от нуля» (1892), я познакомился лишь после того, как предварительные алгебраические теоремы, составляющие содержание настоящей главы, были мной самостоятельно найдены и доказаны. Несомненно, более ранее знакомство с идеями этих ученых упростило бы мою задачу, а также, быть может, и изложение этой главы. Но изменять уже вполне законченные доказательства я не счел нужным ввиду вспомогательной роли упомянутых теорем и так как мне казалось, кроме того, что применение общего метода В. А. Маркова, могущего дать даже больше того, что нам здесь нужно, не упростило бы изложения. Рассуждения же А. А. Маркова, которыми в некоторых случаях, быть может, было бы целесообразно воспользоваться, в других случаях, повидимому, нуждались бы в значительных дополнениях.

[2] В упомянутой выше работе В. А. Марков, подобно тому как это уже было сделано для первой производной, дает максимум, которого k-я производная действительно может достигнуть. Мы же указываем здесь лишь верхнюю границу этого максимума, вполне, однако, достаточную для тех выводов, которые будут сделаны в следующей главе.

* См. статью [10]. *(Ред.)*

значения, где $|P(x)| = L$, а через ξ данное значение, где $P(\xi) = M$, находим, подобно предыдущему, что никакой многочлен $F_n(x)$ степени n не может удовлетворить уравнениям

$$F_n(x_1) = P(x_1),\ F_n(x_2) = P(x_2), \ldots,\ F_n(x_k) = P(x_k),\ F_n(\xi) = 0,$$

что будет иметь место лишь тогда, когда $k > n$. Следовательно, $k = n + 1$, и $P(x)$ есть тригонометрический многочлен; что и требовалось доказать.

7. Следствия. а) *Если на отрезке* $[-1, +1]$ *многочлен степени* n *достигает максимума* L, *то наибольшее абсолютное значение, какое он может получить в вещественной точке* ξ (*не лежащей на этом отрезке*), *есть*

$$M = L\left|\frac{(\xi+\sqrt{\xi^2-1})^n + (\xi-\sqrt{\xi^2-1})^n}{2}\right|. \tag{8}$$

В самом деле, указанное значение M есть абсолютное значение, получаемое в точке ξ соответствующим тригонометрическим многочленом.

б) *Если обозначить через* R *полусумму осей эллипса, проходящего через точку* ξ *и имеющего фокусами* -1, $+1$, *то имеет место неравенство*

$$M < LR^n. \tag{9}$$

В самом деле, положим

$$\xi = \frac{1}{2}\left[(e^b + e^{-b})\cos a + i(e^b - e^{-b})\sin a\right] = \cos(a - bi).$$

В таком случае, если b получает определенное положительное значение, ξ находится на эллипсе, имеющем фокусами -1, $+1$, а осями $e^b + e^{-b}$ и $e^b - e^{-b}$. С другой стороны,

$$M = L\,|\cos n(a - bi)| = \frac{L}{2}\,|\cos na\,(e^{nb} + e^{-nb}) + i\sin na\,(e^{nb} - e^{-nb})| =$$
$$= \frac{L}{2}\sqrt{e^{2nb} + e^{-2nb} + 2\cos 2na}.$$

Следовательно,

$$\frac{L}{2}(e^{nb} - e^{-nb}) \leqslant M \leqslant \frac{L}{2}(e^{nb} + e^{-nb}) < Le^{nb}.$$

Но

$$e^b = \frac{(e^b + e^{-b}) + (e^b - e^{-b})}{2} = R$$

есть полусумма осей рассматриваемого эллипса. Откуда

$$M < LR^n.$$

Примечание. Легко проверить, что неравенство (9) останется в силе, если отрезок $[-1, +1]$ заменить любым отрезком $[\alpha, \beta]$; только R будет тогда обозначать отношение суммы осей эллипса, проходящего через ξ и имеющего фокусами α, β, к фокусному расстоянию.

8. Теорема. *Если* $P_n(x)$ *есть многочлен* n*-й степени, и на отрезке* $[-1, +1]$ *существуют значения* x, y, *для которых*

$$E(x, y) = \left|\frac{P_n(x) - P_n(y)}{(x-y)^\alpha}\right|(1-x^2)^{\alpha/2}(1-y^2)^{\alpha/2} = M,$$

при $0 < \alpha < 1$, *то* $|P_n(x)|$ *не остается менее* $\frac{M}{n^\alpha 2^{1-\alpha}}$ *на этом отрезке.*

В самом деле, подобно предыдущему, убеждаемся в существовании многочлена $P(x)$, для которого максимум $|P(x)|$ достигает наименьшего возможного значения. Кроме того, если (x, y) суть значения, для которых $E(x, y)$ — максимум, и $x_1, x_2, \ldots, x_k$ — значения, где $|P(x)|$ — максимум, то уравнения

$$F_n(x_1) = P(x_1), \ldots, \quad F_n(x_k) = P(x_k), \quad F_n(x) - F_n(y) = 0 \tag{10}$$

несовместимы. Поэтому, если P не тригонометрический многочлен, то k равно n, и, полагая

$$F_n(x) = P(x) + b(x-x_1)(x-x_2)\ldots(x-x_n) = P(x) + bR(x),$$

находим, что уравнения (10) приводятся к одному

$$P(x) - P(y) + b[R(x) - R(y)] = 0,$$

которое будет неразрешимо только если

$$R(x) = R(y),$$

т. е. если

$$\frac{P'(x)(x^2-1)}{x-\beta} = \frac{P'(y)(y^2-1)}{y-\beta}.$$

Но, с другой стороны, числа x, y удовлетворяют уравнениям, выражающим, что $|E(x, y)|$ — максимум:

$$(1-x^2)\{P'(x)(x-y) - \alpha[P(x) - P(y)]\} - \alpha x(x-y)[P(x) - P(y)] = 0,$$

$$(1-y^2)\{P'(y)(y-x) - \alpha[P(y) - P(x)]\} - \alpha y(y-x)[P(y) - P(x)] = 0,$$

или

$$P'(x)(1-x^2) = \alpha\frac{P(x)-P(y)}{x-y}(1-xy) = A \gtreqless 0$$

и

$$P'(y)(1-y^2) = \alpha\frac{P(x)-P(y)}{x-y}(1-xy) = A \gtreqless 0.$$

Таким образом,

$$\frac{A}{x-\beta} = \frac{A}{y-\beta},$$

что невозможно, так как $x \gtreqless y$. Следовательно, P есть тригонометрический многочлен, $P = L\cos n \arccos x$.

Остается вычислить максимум $|E(x, y)|$ для этого многочлена. С этой целью полагаем

$$x = \cos\theta, \qquad y = \cos\varphi \qquad (0 < \theta < \pi;\ \ 0 < \varphi < \pi).$$

В таком случае

$$E(x, y) = L\left|\frac{\cos n\theta - \cos n\varphi}{(\cos\theta - \cos\varphi)^{\alpha}}(\sin\theta\sin\varphi)^{\alpha}\right| =$$

$$= L\left|\left(\frac{\cos n\theta - \cos n\varphi}{\cos\theta - \cos\varphi}\right)^{\alpha}(\cos n\theta - \cos n\varphi)^{1-\alpha}(\sin\theta\sin\varphi)^{\alpha}\right| \leqslant$$

$$\leqslant L\left|\left(\frac{\sin\frac{n}{2}(\theta-\varphi)}{\sin\frac{1}{2}(\theta-\varphi)}\right)^{\alpha}\left(\frac{\sin\theta\sin\varphi}{\sin\frac{1}{2}(\theta+\varphi)}\right)^{\alpha}(\cos n\theta - \cos n\varphi)^{1-\alpha}\right| < 2^{1-\alpha}n^{\alpha}L.$$

Поэтому

$$M < 2^{1-\alpha}n^{\alpha}L,$$

откуда

$$L > \frac{M}{2^{1-\alpha}n^{\alpha}},$$

что и требовалось доказать.

Примечание. Аналогичным образом получим, что наибольшее значение

$$\left|\frac{P_n(x) - P_n(y)}{(x-y)^{\alpha}}\right|\sqrt{(h^2 - x^2)^{\alpha}(h^2 - y^2)^{\alpha}}$$

на отрезке $[-h,\ +h]$ меньше, чем $L \cdot 2^{1-\alpha}(nh)^{\alpha}$.

9. Теорема. *Произведение* $|P_n(x)\sqrt{1-x^2}|$, *где* $P_n(x)$ — *многочлен* n-*й степени, не может оставаться менее* $\frac{M}{n+1}$ *на отрезке* $[-1,\ +1]$, *если* $\left|\frac{d}{dx}(P_n(x)\sqrt{1-x^2})\right|\sqrt{1-x^2}$ *достигает значения* M *на этом отрезке.*

В самом деле, подобно предыдущему, убеждаемся в существовании многочлена $P(x)$, осуществляющего минимальное отклонение, а также и в том, что число k точек, где оно имеет место, более или равно n. Случай $k = n$, вследствие несовместности уравнений

$$F_n(x_1) = P(x_1), \ldots,\ \ F_n(x_k) = P(x_k),\ \ \frac{d}{dx}[F_n(\zeta)\sqrt{1-\zeta^2}] = 0,$$

приводит к невозможности уравнения

$$\frac{d}{dx}[P(\zeta)\sqrt{1-\zeta^2}] + b\frac{d}{dx}[R(\zeta)\sqrt{1-\zeta^2}] = 0,$$

где

$$R(x) = (x - x_1)(x - x_2)\ldots(x - x_k),$$

откуда следует, что ζ удовлетворяет уравнению

$$\frac{d}{dx}[R(x)\sqrt{1-x^2}] = 0. \tag{11}$$

При этом нужно заметить, что $|P(x)\sqrt{1-x^2}|$ достигает максимума лишь во внутренних точках, обращающих в нуль выражение

$$\frac{d}{dx}[P(x)\sqrt{1-x^2}]=\frac{P'(x)(1-x^2)-xP(x)}{\sqrt{1-x^2}}=\frac{Q(x)}{\sqrt{1-x^2}},$$

т. е. не более чем в $n+1$ точках, удовлетворяющих уравнению $P'(x)(1-x^2)-xP(x)=0$; поэтому

$$R(x)=\frac{CQ(x)}{x-\beta},$$

так что уравнение (11) превращается в

$$\frac{d}{dx}\left(\frac{Q(x)\sqrt{1-x^2}}{x-\beta}\right)=0. \tag{11'}$$

Но M, по предположению, — наибольшее значение

$$\left[\frac{d}{dx}(P\sqrt{1-x^2})\right]\sqrt{1-x^2}=Q(x);$$

поэтому ζ удовлетворяет также уравнению

$$Q'(x)=0.$$

Следовательно, уравнение (11') приводится (как в теореме § 2) к виду

$$Q(x)(\beta x-1)=0,$$

и так как $Q(\zeta)\gtrless 0$, то $\beta=\frac{1}{\zeta}$, откуда $|\beta|>1$.

Замечая далее, что $S=P\sqrt{1-x^2}$ достигает n раз наибольшего абсолютного значения L, заключаем, что

$$S^2-L^2=\frac{(S')^2}{(n+1)^2}\cdot\frac{(x^2-1)(x-\gamma)(x-\delta)}{(x-\beta)^2}.$$

Я утверждаю [3.3], что *корни γ и δ, которые комплексны и сопряжены, так как $S^2-L^2\leqslant 0$ при всех вещественных значениях x, имеют вещественную часть того же знака, что и β, и бо́льшую, чем β, по абсолютной величине.*

Предположим для определенности, что $\beta>0$; тогда, вследствие сказанного ранее, $\beta>1$. В таком случае $iS(x)=P(x)\sqrt{x^2-1}$, обратившись в 0 в точке $x=1$, достигнет при возрастании $x>1$ единственного максимума (или минимума) в точке $x=\beta$ и, бесконечно возрастая при $x\to\infty$, обратится сначала еще раз в 0 в точке $x=\alpha>\beta$ и примет значение $+L$ или $-L$ точке $x=\lambda>\alpha$. Следовательно, многочлен S^2, все корни которого вещественны (корни ± 1, $n-1$ двойных корней внутри отрезка $(-1,+1)$ и двойной корень $\alpha>\beta$), принимает значение $-L^2$ в точке λ. Через точку λ пройдет, таким образом, в плоскости комплексной переменной x кривая, на которой $|S^2(x)|=L^2$, окружающая по крайней мере один нуль многочлена $S^2(x)$. Двигаясь вдоль этой кривой из λ, например в верхней полуплоскости комплексной переменной x, мы встретим перпендикуляр, восстановленный к вещественной оси из точки α, в некоторой точке M, в которой аргумент $S^2(x)$ изменится на угол больший, чем π, так как все корни $S^2(x)$ лежат слева от α. Следовательно, раньше чем мы попадем в точку M, мы встретим точку γ с вещественной частью, большей $\alpha > \beta$, в которой $S^2(\gamma) = L^2$. Этим наше утверждение доказано.

Поэтому

$$\theta^2 (S^2 - L^2) = \frac{(x^2 - 1)\, S'^2}{(n+1)^2},$$

где $\theta < 1$. Откуда

$$|S' \sqrt{1 - x^2}| = |Q(x)| < (n + 1) L, \quad \text{т. е.} \quad L > \frac{M}{n+1}.$$

Напротив, если $k = n + 1$, то

$$S^2 - L^2 = \frac{(x^2 - 1) S'^2}{n + 1},$$

так что

$$S = L \sin (n + 1) \arccos x \quad \text{и} \quad P = \frac{L \sin (n + 1) \arccos x}{\sqrt{1 - x^2}},$$

следовательно, $L = \dfrac{M}{n+1}$, что и требовалось доказать.

10. Применение предыдущего к тригонометрическим суммам. Условимся называть тригонометрической суммой n-го порядка выражение вида

$$A_0 + A_1 \cos t + B_1 \sin t + \ldots + A_n \cos nt + B_n \sin nt;$$

если все B_i равны нулю, то выражение будет называться (тригонометрическою) суммою конусов n-го порядка; если же все A_i равны нулю, то это будет сумма синусов того же порядка. Все выше доказанные теоремы приводят к аналогичным предложениям относительно тригонометрических сумм, если положить $x = \cos t$ и заметить, что всегда возможно, с одной стороны, отожествить выражения

$$a_0 + a_1 \cos t + \ldots + a_n \cos^n t \quad \text{и} \quad A_0 + A_1 \cos t + \ldots + A_n \cos nt,$$

и, с другой стороны, отожествить

$$\sin t\, [b_0 + b_1 \cos t + \ldots + b_n \cos^n t] \quad \text{и} \quad B_0 \sin t + \ldots + B_n \sin (n + 1) t.$$

Ограничимся лишь формулировкой предложений, соответствующих теоремам §§ 2 и 9.

Если абсолютное значение суммы косинусов n-го порядка

$$w = A_0 + A_1 \cos t + \ldots + A_n \cos nt$$

не превышает L, то абсолютное значение ее производной

$$-(A_1 \sin t + \ldots + nA_n \sin nt)$$

не превышает никогда nL; последнее значение достигается только при $A_0 = A_1 = \ldots = A_{n-1} = 0$.

Действительно, полагая $x = \cos t$, мы превращаем w в многочлен n-й степени $P_n(x)$; при этом

$$\frac{dw}{dt} = -P'(x)\sqrt{1-x^2}.$$

Таким же точно образом легко вывести из теоремы § 9 предложение: *Если абсолютное значение суммы синусов n-го порядка*

$$B_1 \sin t + B_2 \sin 2t + \ldots + B_n \sin nt$$

не превышает L, то абсолютное значение ее производной

$$B_1 \cos t + \ldots + nB_n \cos nt$$

не превышает nL; последнее значение достигается только при $B_1 = B_2 = \ldots$ $\ldots = B_{n-1} = 0$.

Эти два предложения можно обобщить следующим образом [3.4]:

Если абсолютное значение тригонометрической суммы n-го порядка удовлетворяет неравенству

$$|f(t)| = |A_0 + A_1 \cos t + B_1 \sin t + \ldots + A_n \cos nt + B_n \sin nt| \leqslant L,$$

то производная ее $f'(t)$ удовлетворяет неравенству

$$|f'(t)| = |-A_1 \sin t + B_1 \cos t + \ldots - nA_n \sin nt + nB_n \cos nt| \leqslant nL.$$

Действительно, пусть

$$|f(t)| = |A_0 + A_1 \cos t + B_1 \sin t + \ldots + A_n \cos nt + B_n \sin nt| \leqslant L;$$

тогда будем иметь также

$$\frac{1}{2}|f(t+t_0) - f(t-t_0)| \leqslant L.$$

Но

$$\frac{1}{2}[f(t+t_0) - f(t-t_0)] = \sum_{k=1}^{n} (-A_k \sin kt_0 + B_k \cos kt_0) \sin kt$$

является суммой синусов и, следовательно, вследствие предыдущего предложения, относящегося к сумме синусов, имеем

$$\left|\sum_{k=1}^{n} k(-A_k \sin kt_0 + B_k \cos kt_0) \cos kt\right| \leqslant nL.$$

Так как это неравенство справедливо при любом t, то, полагая $t = 0$, заключаем, что

$$\left|\sum_{k=1}^{n} k(-A_k \sin kt_0 + B_k \cos kt_0)\right| \leqslant nL,$$

т. е.

$$|f'(t)| \leqslant nL,$$

что и требовалась доказать.

11. Производные высших порядков. Из первых двух предложений § 10 вытекает, что *если L есть наибольшее абсолютное значение суммы* $A_0 + A_1 \cos t + \ldots + A_n \cos nt$ $\left(\textit{или} \sum\limits_{k=1}^{n} A_k \sin k\theta\right)$, *то наибольшее абсолютное значение ее p-ой производной не превышает* n^pL (*случай равенства имеет место только при* $A_0 = A_1 = \ldots = A_{n-1} = 0$).

Этот результат можно преобразовать, возвращаясь снова к многочленам. Именно, полагая, что $|P_n(x)| = |a_0 + \ldots + a_nx^n|$ на отрезке $[-1, +1]$ менее L, мы должны заключить, что

$$\left|\frac{d^kP_n(x)}{(d \arccos x)^k}\right| \leqslant n^kL,$$

или

$$|P'_n(x)\sqrt{1-x^2}| \leqslant nL,$$

$$|P''_n(x)(1-x^2) - xP'_n(x)| \leqslant n^2L,$$

и т. д.

Однако этими неравенствами мы в дальнейшем пользоваться не будем и заменим их менее точными, но более удобными. С этой целью замечаем, что

$$|P'_n(x)| \leqslant \frac{nL}{\sqrt{1-x^2}};$$

но в таком случае P'_n — многочлен $(n-1)$-й степени, который в промежутке $(-x_1, +x_1)$ менее, чем $\frac{nL}{\sqrt{1-x_1^2}}$, а потому

$$|P''_n(x)| < \frac{n(n-1)L}{\sqrt{(x_1^2-x^2)(1-x_1^2)}},$$

и, повторяя то же рассуждение, найдем

$$|P_n^{(k)}(x)| < \frac{n(n-1)\ldots(n-k+1)L}{\sqrt{(x_{k-1}^2-x^2)\ldots(1-x_1^2)}}.$$

Полагая же $1 - x_1^2 = x_1^2 - x_2^2 = \ldots = x_{k-1}^2 - x^2 = \frac{1-x^2}{k}$, получим, наконец,

$$|P_n^{(k)}(x)| < \left(\frac{k}{1-x^2}\right)^{k/2} n(n-1)\ldots(n-k+1)L. \tag{12}$$

Аналогичным образом можно проверить правильность неравенства

$$\left|\frac{P_n^{(k)}(z) - P_n^{(k)}(z_1)}{(z-z_1)^\alpha}\right| < \left(\frac{k+1}{1-x^2}\right)^{\frac{k}{2}+\alpha} 2n(n-1)\ldots(n-k+1)\left(\frac{n-k}{2}\right)^\alpha L \tag{12'}$$

при условиях

$$|z| \leqslant x, \qquad |z_1| \leqslant x.$$

Глава II [3.5]

ОПРЕДЕЛЕНИЕ НИЗШЕГО ПРЕДЕЛА УКЛОНЕНИЯ НЕПРЕРЫВНОЙ ФУНКЦИИ ОТ МНОГОЧЛЕНА ДАННОЙ СТЕПЕНИ

12. Теорема. *Пусть дан ряд*

$$f(x) = u_1 + \ldots + u_n + \ldots,$$

где $u_n(x)$ *— многочлен степени не выше* n. *Если этот ряд сходится на отрезке* $[-1, +1]$, *и притом*

$$|u_n + u_{n+1} + \ldots| < \frac{A}{n^p},$$

где A *— постоянная величина, то* $f(x)$ *имеет во всякой точке внутри отрезка* $[-1, +1]$ *непрерывную и конечную производную* k*-го порядка, где* k *— наибольшее целое число, меньшее, чем* p; *кроме того, эта производная удовлетворяет условиям Липшица степеней* α *сколь угодно близких к* $p - k$.

В самом деле, полагая

$$R_n = u_n + u_{n+1} + \ldots$$

имеем, по условию,

$$|R_n| < \frac{A}{n^p};$$

поэтому, в частности,

$$|R_{2^m}| < \frac{A}{2^{mp}}, \qquad |R_{2^{m+1}}| < \frac{A}{2^{(m+1)p}}.$$

Следовательно, если обозначим через v_m многочлен степени $2^{m+1} - 1$

$$v_m = R_{2^m} - R_{2^{m+1}} = u_{2^m} + u_{2^m+1} + \ldots + u_{2^{m+1}-1}, \tag{13}$$

то

$$|v_m| < \frac{A}{2^{mp}} + \frac{A}{2^{(m+1)p}} < \frac{2^{p+1}A}{2^{(m+1)p}}; \tag{14}$$

таким образом, указанной группировкой членов мы превращаем ряд (13) в абсолютно сходящийся ряд

$$f(x) = v_0 + v_1 + \ldots + v_m + \ldots,$$

каждый член которого есть многочлен степени $2^{m+1} - 1$.

Дифференцируем почленно k раз полученный ряд, замечая, что, вследствие неравенств (12) и (14),

$$|v_m^{(k)}(x)| < \left(\frac{k}{1-x^2}\right)^{k/2} 2^{(m+1)k} \frac{2^{p+1}A}{2^{(m+1)p}} = 2^{p+1}\left(\frac{k}{1-x^2}\right)^{k/2} A \left(\frac{1}{2^{p-k}}\right)^{m+1}.$$

Следовательно,

$$|\rho_m^{(k)}| = |v_m^{(k)}(x) + v_{m+1}^{(k)}(x) + \ldots| <$$

$$< 2^{p+1}\left(\frac{k}{1-x^2}\right)^{k/2} A\left[\left(\frac{1}{2^{p-k}}\right)^{m+1} + \left(\frac{1}{2^{p-k}}\right)^{m+2} + \ldots\right] = \frac{2^{p+1}}{2^{p-k}-1}\left(\frac{k}{1-x^2}\right)^{k/2}\frac{A}{2^{(p-k)m}},$$

а потому ряд

$$f^{(k)} = v_0^{(k)}(x) + \ldots + v_m^{(k)}(x) + \ldots$$

равномерно (и абсолютно) сходится во всяком промежутке внутри отрезка $[-1, +1]$. Отсюда вытекают существование конечной k-й производной и ее непрерывность.

Вторая часть теоремы получится, если вместо неравенства (12) мы воспользуемся неравенством (12′). Полагая $p > k + \alpha$, находим

$$\left| \frac{f^{(k)}(z) - f^{(k)}(z_1)}{(z - z_1)^\alpha} \right| < \sum_{m=0}^{\infty} \left| \frac{v_m^{(k)}(z) - v_m^{(k)}(z_1)}{(z - z_1)^\alpha} \right| <$$

$$< \left(\frac{k+1}{1-x^2}\right)^{\frac{k}{2}+\alpha} 2^{p+2} A \sum_{m=0}^{\infty} \left(\frac{1}{2^{p-k-\alpha}}\right)^{m+1} = \frac{2^{p+2} A}{2^{p-k-\alpha} - 1} \left(\frac{k+1}{1-x^2}\right)^{\frac{k}{2}+\alpha},$$

если $|z| \leqslant x$ и $|z_1| \leqslant x$, что и требовалось доказать.

Примечание. Применяя следствие (г) § 5, мы таким же образом убедились бы в конечности k-й производной и в концах отрезка, если только $k < \frac{p}{2}$.

13. Следствие. *Ряд* (13) *может быть дифференцируем почленно* k *раз, если* $k < p$.

В самом деле, из предшествующего доказательства видно, что это дифференцирование возможно при условии соединения в одну группу членов $u_{2m} + u_{2^m+1} + \ldots + u_{2^{m+1}-1} = v_m$. Но группировка (необходимая вообще для *абсолютной* сходимости) не является необходимой для равномерной сходимости, ибо легко видеть, что при всяком $N < 2^m$

$$\left| u_{2^m}^{(k)} + \ldots + u_{2^m+N}^{(k)} \right| < 2^{p+1} \left(\frac{k}{1-x^2}\right)^{k/2} A \left(\frac{1}{2^{p-k}}\right)^{m+1}.$$

14. Теорема. *Если (при прежних обозначениях)*

$$|u_n + u_{n+1} + \ldots| < \frac{A_n}{n^p},$$

и ряд

$$S = A_1 + A_2 + A_4 + \ldots + A_{2^m} + \ldots \tag{15}$$

сходящийся, то, при p *целом, функция* $f(x)$ *имеет конечную и непрерывную производную* p*-го порядка во всякой точке внутри отрезка* $[-1, +1]$; *в случае же, когда* $p = k + \alpha$, *где* k — *наибольшее целое число, меньшее, чем* p, k*-я производная удовлетворяет во всяком промежутке внутри того же отрезка условию Липшица степени* α.

Ограничимся случаем, когда p — целое число, так как вторая часть теоремы доказывается таким же образом.

Полагая, как в предыдущем параграфе,

$$v_m = u_{2^m} + \ldots + u_{2^{m+1}-1},$$

находим, что

$$|v_m| < \frac{A_{2m+1}}{2^{(m+1)p}} + \frac{A_{2m}}{2^{mp}}.$$

А потому, пользуясь неравенством (12), заключаем, что

$$|f^{(p)}(x)| \leqslant |v_0^{(p)}(x)| + |v_1^{(p)}(x)| + \ldots + |v_m^{(p)}(x)| + \ldots <$$

$$< \left(\frac{p}{1-x^2}\right)^{p/2}(2^p+1)(A_1 + A_2 + A_4 + \ldots + A_{2m}) = \left(\frac{p}{1-x^2}\right)^{p/2}(2^p+1)\,S.$$

Что и требовалось доказать.

15. Следствия. В условии только что доказанной теоремы не сделано никаких предположений относительно чисел A_n, кроме сходимости ряда (15).

Однако мы можем заметить, что, группируя, если это понадобится, члены ряда (13), всегда возможно превратить его в ряд того же вида, но обладающий свойством, что числа A_n/n^p идут *не возрастая* с возрастанием n; другими словами, рассматривая конечную сумму $u_1 + u_2 + \ldots$ $\ldots + u_n$ как приближенный многочлен степени n функции $f(x)$, мы можем не вводить $(n+1)$-го члена, если он не увеличивает приближени и тогда $u_{n+1} = 0$ и $\frac{A_{n+1}}{(n+1)^p} = \frac{A_n}{n^p}$, и ввести затем сразу группу членов, действительно улучшающих приближение.

В таком случае легко убедиться в следующем.

Если есть такое число p, *что*

$$\frac{A_{n+1}}{(n+1)^p} \leqslant \frac{A_n}{n^p},$$

то ряды

$$S = A_1 + A_2 + A_4 + \ldots + A_{2^n} + \ldots \quad \text{и} \quad \Sigma = A_1 + \frac{A_2}{2} + \frac{A_3}{3} + \ldots + \frac{A_n}{n} + \ldots$$

или оба сходящиеся, или оба расходящиеся.

Действительно, если $p \geqslant 1$, то

$$I_n = \frac{A_n}{n} + \frac{A_{n+1}}{n+1} + \ldots + \frac{A_{2n-1}}{2n-1} =$$

$$= \frac{A_n}{n^p} n^{p-1} + \frac{A_{n+1}}{(n+1)^p}(n+1)^{p-1} + \ldots + \frac{A_{2n-1}}{(2n-1)^p}(2n-1)^{p-1} =$$

$$= n^{p-1}\left[\frac{A_n}{n^p} + \frac{A_{n+1}}{(n+1)^p}\left(\frac{n+1}{n}\right)^{p-1} + \ldots + \frac{A_{2n-1}}{(2n-1)^p}\left(\frac{2n-1}{n}\right)^{p-1}\right] < A_n \cdot 2^{p-1}$$

и, с другой стороны,

$$I_n = n^{p-1}\left[\frac{A_n}{n^p} + \ldots + \frac{A_{2n-1}}{(2n-1)^p}\left(\frac{2n-1}{n}\right)^{p-1}\right] > \left(\frac{n}{2n-1}\right)^p A_{2n-1} \geqslant$$

$$\geqslant n^p \frac{A_{2n}}{(2n)^p} = \frac{A_{2n}}{2^p}.$$

Таким образом,

$$\frac{A_{2n}}{2^p} < \frac{A_n}{n} + \frac{A_{n+1}}{n+1} + \ldots + \frac{A_{2n-1}}{2n-1} < A_n \cdot 2^{p-1},$$

и, следовательно,

$$\frac{S}{2^p} - A_1 < \Sigma < 2^{p-1}S \qquad (p \geqslant 1).$$

Если $p \leqslant 1$, то подобным же образом получим

$$\frac{S}{2} - A_1 < \Sigma < S \qquad (p \leqslant 1).$$

Итак, при предположении, что $\frac{A_{n+1}}{(n+1)^p} \leqslant \frac{A_n}{n^p}$, *условие сходимости ряда* S *в теореме* § 14 *может быть заменено равнозначным ему условием сходимости ряда*

$$\Sigma = A_1 + \frac{A_2}{2} + \ldots + \frac{A_n}{n} + \ldots \tag{15'}$$

Для практического применения теоремы § 14 можем воспользоваться различными достаточными условиями сходимости. Таким образом, условие сходимости ряда (15) или (15′) может быть заменено более специальными условиями (не равнозначными предыдущим), а именно, например, условием, чтобы было

$$A_n < \frac{1}{(\log n)^{1+\varepsilon}} \quad \text{или} \quad A_n < \frac{1}{\log n\,(\log\log n)^{1+\varepsilon}} \quad \text{и т. д.,}$$

где ε — некоторое положительное число.

16. Теорема. *Пусть попрежнему*

$$f(x) = u_1 + u_2 + \ldots + u_n + \ldots,$$

где u_n — *многочлен степени не выше* n, *и на отрезке* $[-1, +1]$

$$|u_n + u_{n+1} + \ldots| < \frac{A_n}{n^p},$$

где числа A_n *идут не возрастая; в таком случае, для всякого целого значения* $p_1 < p$,

$$f^{(p_1)}(x) = w_{p_1} + w_{p_1+1} + \ldots + w_n + \ldots,$$

где w_n — *многочлен степени не выше* $n - p_1$, *причем*

$$(1 - x^2)^{p_1/2}\,|w_{2n} + w_{2n+1} + \ldots| < \frac{p_1^{p_1/2} \cdot 2^{p+1}}{2^{p-p_1} - 1} \cdot \frac{A_n}{n^{p-p_1}}, \tag{16}$$

и, при $2p_1 < p$,

$$|w_{2n} + w_{2n+1} + \ldots| < \frac{2^{p+1}}{2^{p-2p_1} - 1} \cdot \frac{A_n}{n^{p-2p_1}}. \tag{16'}$$

В самом деле,

$$f^{(p_1)}(x) = v_0^{(p_1)} + \ldots + v_m^{(p_1)} + \ldots,$$

причем, вследствие неравенства (12),

$$\left| v_m^{(p_1)} \right| + \left| v_{m+1}^{(p_1)} \right| + \ldots <$$

$$< \left(\frac{p_1}{1 - x^2}\right)^{p_1/2} \cdot 2^{p+1} \left[A_{2m} \cdot 2^{(m+1)(p_1 - p)} + A_{2m+1} \cdot 2^{(m+2)(p_1 - p)} + \ldots\right] \leqslant$$

$$\leqslant \left(\frac{p_1}{1 - x^2}\right)^{p_1/2} \frac{2^{p+1} 2^{(m+1)(p_1 - p)}}{1 - 2^{p_1 - p}} A_{2m}.$$

Поэтому, полагая

$$w_n = v_m^{(p_1)}, \qquad \text{если} \qquad n = 2^{m+1} - 1,$$

и

$$w_n = 0, \qquad \text{если} \qquad n \gtrless 2^{m+1} - 1,$$

находим:

$$f^{(p_1)}(x) = w_{p_1} + \ldots + w_n + \ldots,$$

где w_n — многочлен степени не выше $n - p_1$, причем

$$(1 - x^2)^{p_1/2} \left| w_{2n} + w_{2n+1} + \ldots \right| < \frac{p_1^{p_1/2} \cdot 2^{p+1}}{2^{p - p_1} - 1} \cdot \frac{A_n}{n^{p - p_1}}.$$

Точно так же из следствия (г) § 5 заключаем, что

$$\left| v_m^{(p_1)} \right| + \left| v_{m+1}^{(p_1)} \right| + \ldots < 2^{p+1} \left[A_{2m} 2^{(m+1)(2p_1 - p)} + A_{2m+1} 2^{(m+2)(2p_1 - p)} + \ldots\right] \leqslant$$

$$\leqslant \frac{2^{p+1} 2^{(m+1)(2p_1 - p)}}{1 - 2^{2p_1 - p}} \cdot A_{2m},$$

откуда

$$\left| w_{2n} + w_{2n+1} + \ldots \right| < \frac{2^{p+1}}{2^{p - 2p_1} - 1} \cdot \frac{A_n}{n^{p - 2p_1}}.$$

Примечания. а) Теорема, в частности, применима, если $A_n = A$ постоянная величина.

б) Заметим также, что $\left| u_{2n+l}^{(p_1)} + \ldots \right|$ удовлетворяют тем же неравенствам, что и $\left| w_{2n} + w_{2n+1} + \ldots \right|$, при всяком $l > 0$.

в) Аналогичные неравенства имеют место, если, вместо производных, брать отношения $\frac{f(x) - f(a)}{(x - a)^p}$, при $p < 1$.

17. Тригонометрические ряды. Принимая во внимание результаты § 10, легко видеть, что предыдущие теоремы остаются в силе, если в ряду

$$f(x) = u_1 + u_2 + \ldots + u_n + \ldots \tag{13'}$$

функции u_n будут тригонометрическими суммами n-го порядка. Таким образом:

Если $\left| u_n + u_{n+1} + \ldots \right| < \frac{A_n}{n^p}$, *где* p — *целое число, и ряд* $S = A_1 + A_2 +$

$+A_4+\ldots+A_{2m}+\ldots$ *сходящийся, то p-я производная* $|f^{(p)}(x)|$ *будет непрерывна и* $|f^{(p)}(x)|<2^p(2^p+1)S$. *В случае, когда все* u_n *содержат только косинусы или только синусы,* $|f^{(p)}(x)|\leqslant(2^p+1)S$.

Эта теорема доказывается совершенно так же, как и теорема § 14; и подобно ей, mutatis mutandis, получаются и другие эквивалентные теоремы, если многочлены заменяются тригонометрическими суммами.

18. Теорема. *Если внутри отрезка* $[-1, +1]$ *есть по крайней мере одна точка, где p-я производная* $f^{(p)}(x)$ *некоторой функции* $f(x)$ *не непрерывна и наилучшее приближение* E_{n-1} *функции* $f(x)$ *на этом отрезке при помощи многочлена степени* $n-1$ *равно* $\frac{A_n}{n^p}$, *то ряд* $\Sigma=A_1+\frac{A_2}{2}+\ldots$ $\ldots+\frac{A_n}{n}+\ldots$ *расходящийся. Обратно, каковы бы ни были данные положительные числа* $A'_1, A'_2, \ldots, A'_n, \ldots$, *удовлетворяющие условию* $\frac{A'_n}{n^p}\geqslant\frac{A'_{n+1}}{(n+1)^p}$, *если ряд* $\Sigma'=A'_1+\ldots+\frac{A'_n}{n}+\ldots$ *расходящийся, то можно построить функцию* $f(x)$, *p-я производная которой* $f^{(p)}(x)$ *не непрерывна внутри отрезка, причем для всякого n наилучшее приближение удовлетворяет неравенству* $E_{n-1}<\frac{A'_n}{n^p}$. (Аналогичная теорема для тригонометрических сумм.)

Первая часть теоремы непосредственно вытекает из формулировки, данной в § 15 теореме § 14, так как, если бы ряд Σ сходился, то $f^{(p)}(x)$ была бы непрерывна и конечна внутри отрезка $(-1, +1)$.

Допустим далее, что ряд $\Sigma'=A'_1+\ldots+\frac{A'_n}{n}+\ldots$ расходящийся, и рассмотрим два случая. Пусть, во-первых, начиная от некоторого n_1, все $A'_n\geqslant 1$. В таком случае можно выбрать (см. § 69) численный коэффициент α так, чтобы функция $\varphi(x)=\alpha|x|^p$ удовлетворяла требованию теоремы, а именно:

$$\text{при } n\leqslant n_1,\ E_{n-1}<\alpha<\frac{A'_n}{n^p} \text{ и, при } n>n_1,\ E_{n-1}<\frac{1}{n^p}\leqslant\frac{A'_n}{n^p}.$$

Во втором случае, среди чисел A'_{4m+1} есть бесчисленное множество удовлетворяющих условию $A'_{4m+1}<1+\varepsilon$, как бы мал ни был ε. Пусть, для определенности, p будет нечетно; построим функцию

$$f(x)=\frac{1}{5^p}\sum_{m=1}^{\infty}\left[\frac{A'_{4m+1}}{(4m-3)^p}-\frac{A'_{4m+5}}{(4m+1)^p}\right]\cos(4m+1)x=\sum_{n=1}^{\infty}u_n(x). \qquad (17)$$

Таким образом, тригонометрическая сумма (n_1-1)-го порядка $\sum_{n=1}^{n_1-1}u_n(x)$, при $4m-2\leqslant n_1<4m+2$, удовлетворяет неравенству

$$\left|f(x)-\sum_{n=1}^{n_1-1}u_n(x)\right|\leqslant\frac{A'_{4m+1}}{(4m+1)^p}\leqslant\frac{A'_{n_1}}{n_1^p}.$$

Следовательно, тригонометрическое приближение функции $f(x)$ (от которого мы затем легко перейдем к многочленам) удовлетворяет условию теоремы. Поэтому достаточно будет показать, что p-я производная $f^{(p)}(x)$ в некоторой точке, а именно в $x = \frac{\pi}{2}$, безгранично возрастает. В самом деле, заметив, что все коэффициенты в ряде (17) положительны, дифференцируем его почленно; получим

$$\pm\frac{1}{5^p}\sum_{m=1}^{\infty}\left[A'_{4m+1}\left(1+\frac{4}{4m-3}\right)^p - A'_{4m+5}\right]\sin(4m+1)x$$

и, полагая $x = \frac{\pi}{2}$, находим бесконечно возрастающую сумму положительных членов

$$\frac{1}{5^p}\sum_{m=1}^{\infty}\left(A'_{4m+1}-A'_{4m+5}\right)+\left(\frac{4p}{4m-3}+\ldots\right)A'_{4m+1};$$

но этого не могло бы быть, если бы в рассматриваемой точке $f^{(p)}(x)$ была непрерывна, ибо в таком случае был бы применим способ суммирования тригонометрических рядов Фейера[1], который дал бы $f^{(p)}\left(\frac{\pi}{2}\right) = \infty$; следовательно, при $x = \frac{\pi}{2}$, $f^{(p)}(x)$ не непрерывна. Для того чтобы распространить полученный вывод на многочлены, полагаем $z = \cos x$; тогда $f(x) = \varphi(z)$, и приближение E_{n-1} функции $\varphi(z)$ в промежутке $(-1, +1)$ удовлетворяет условию теоремы. Но ясно, что точке $x = \frac{\pi}{2}$ соответствует $z = 0$, где $\varphi^{(p)}(z)$ не может также быть непрерывна.

19. Добавление к предшествующей теореме. Метод, которым мы пользуемся в этой главе, не может дать никаких указаний относительно верхней границы E_n. Поэтому для полноты картины нам необходимо упомянуть о некоторых результатах, которые будут доказаны лишь в третьей части. А именно, *если $f(x)$ имеет на отрезке $[-1, +1]$ конечную производную p-го порядка, то можно указать такое положительное число k, чтобы, при всяком $n > 0$, имело место неравенство*

$$E_n < \frac{k\log(n+1)}{(n+1)^p};$$

если же эта p-я производная удовлетворяет условию Липшица степени α, то, при всяком $n > 0$,

$$E_n < \frac{k\log(n+1)}{(n+1)^{p+\alpha}} < \frac{k_1}{(n+1)^{p+\alpha_1}},$$

где α_1 $(\alpha_1 < \alpha)$ положительное число, сколь угодно близкое к α.

Отсюда следует, что, *если p-я производная непрерывна и, кроме того, удовлетворяет какому-нибудь условию Липшица, то ряд*

$$\Sigma = A_1 + \frac{A_2}{2} + \ldots + \frac{A_n}{n} + \ldots < \frac{k\log 2}{2^{1+\alpha}} + \ldots + \frac{k\log n}{n^{1+\alpha}} + \ldots$$

сходящийся.

---

[1] Lebesgue. Leçons sur les séries trigonométriques.

Напротив, если p-я производная только непрерывна, то первый из упомянутых результатов дает только[1]

$$\Sigma < \frac{k\log 2}{2} + \ldots + \frac{k\log n}{n} + \ldots$$

и не дает таким образом права заключать о сходимости ряда Σ.

20. Пример функции, имеющей непрерывную производную при расходящемся ряде Σ. Действительно, можно указать пример функции, для которой ряд Σ расходится, хотя производная везде непрерывна. Этим свойством обладает, например, функция[2]

$$f(x) = \sum_{n=1}^{\infty} \frac{\alpha_n \cos nx}{n^2},$$

если выбрать соответствующим образом числа $\alpha_1 > \alpha_2 > \ldots$, причем $\lim\limits_{n\to\infty} \alpha_n = 0$. В самом деле, дифференцируя, получим равномерно сходящийся ряд

$$f'(x) = -\sum_{n=1}^{\infty} \frac{\alpha_n \sin nx}{n},$$

ибо можно указать определенную постоянную A так, чтобы, при всяком n', было

$$\left| \sum_{n=n'}^{\infty} \frac{\sin nx}{n} \right| < A.$$

Но, с другой стороны,

$$\sum_{n=n'}^{\infty} \frac{\alpha_n}{n^2} = \frac{\alpha_{n'}}{n'^2} + \frac{\alpha_{n'+1}}{(n'+1)^2} + \ldots > \frac{\alpha_{2n'}}{4n'} + \frac{\alpha_{4n'}}{8n'} + \ldots > \frac{\alpha_{n'}}{6n'},$$

если только $\alpha_{2n} > \frac{\alpha_n}{2}$. Отсюда можно заключить, как будет доказано в третьей части, что

$$E_n[f(x)] > \frac{k\alpha_{n+1}}{(n+1)\log(n+1)}.$$

Поэтому, если числа α_n убывают достаточно медленно, например $\alpha_n = \frac{1}{\log\log(n+1)}$, то ряд Σ будет расходящимся.

---

[1] Из работы Джексона, упомянутой в сноске[2] на стр. 12 вытекает, что

$$E_n < \frac{k}{(n+1)^p}, \quad \text{т. е.} \quad \Sigma < \frac{k}{2} + \ldots + \frac{k}{n} + \ldots;$$

я полагаю, что в случае *непрерывности* p-й производной можно даже показать, что

$$E_n < \frac{k_{n+1}}{(n+1)^p},$$

где k_n стремится к нулю; но и этого недостаточно для сходимости ряда Σ.

[2] Подобно предыдущему, от тригонометрического ряда к строке многочленов можно перейти при помощи подстановки $t = \cos x$.

Из предыдущего видно, что вообще функции, имеющие непрерывную производную, допускают лучшее приближение при помощи многочленов данной степени, чем функции, не имеющие производной; но тем не менее есть среди функций, имеющих непрерывные производные, особый класс функций $f(x)$, для которых, при всяком n, $E_n[f(x)] > E_n[\varphi(x)]$, где $\varphi(x)$ — некоторая функция, не имеющая непрерывной производной.

21. Применение к функции $|x|$. Производная функции $|x|$ имеет точку разрыва $x = 0$. Отсюда следует, что ряд

$$\Sigma = A_1 + \frac{A_2}{2} + \ldots + \frac{A_n}{n} + \ldots = E_0 + E_1 + \ldots + E_n + \ldots$$

расходящийся, обозначая через $E_n = \frac{A_{n+1}}{n+1}$ наилучшее приближение $|x|$ на отрезке $[-1, +1]$ при помощи многочлена степени n. Никаких заключений о каждом определенном E_n отсюда нельзя вывести. Единственное, что можно сказать, — что при всяком ε будет бесчисленное множество значений n, для которых

$$E_{n-1} > \frac{1}{n(\log n)^{1+\varepsilon}}, \quad E_{n-1} > \frac{1}{n \log n (\log\log n)^{1+\varepsilon}}, \text{ и т. д.}$$

Напротив, одного факта, что производная $|x|$ не непрерывна, *недостаточно* для того, чтобы утверждать, что будет бесчисленное множество значений, для которых $E_{n-1} > \frac{1}{n \log n}$, так как мы видели, что есть функции, не обладающие непрерывной производной, для которых все E_{n-1} менее членов любого расходящегося ряда.

22. Теорема. *Условие необходимое и достаточное для того, чтобы функция $f(x)$ на всем отрезке $[-1, +1]$ имела конечные и непрерывные производные всех порядков, заключается в том, чтобы при всяком p существовало число α_p, не зависящее от n, обладающее свойством, что для всех n*

$$E_n n^p < \alpha_p.$$

В самом деле, условие достаточно, так как из примечания к теореме § 12 вытекает существование конечной производной k-го порядка на всем отрезке, если $k < \frac{p}{2}$. С другой стороны, условие необходимо вследствие § 19.

23. Пример функции, для которой E_n убывает неправильно. Из предыдущего следует, что, если условие $E_n < \frac{\alpha_p}{n^p}$ соблюдается для всякого n, то функция имеет производные всех порядков. Нельзя того же сказать, если неравенство это соблюдено хотя и для бесчисленного множества, но не для всех значений n.

В самом деле, рассмотрим функцию

$$f(x) = \sum_{m=1}^{\infty} \frac{\cos 2^{m!} x}{2^{m!}}. \tag{18}$$

Полагая $n = 2^{m!} > 2^{p!}$, находим:

$$E_n \leqslant \left[\frac{1}{2^{(m+1)!}} + \frac{1}{2^{(m+2)!}} + \dots\right] < \frac{2}{2^{(m+1)!}} = \frac{2}{(2^{m!})^{m+1}} = \frac{2}{n^{m+1}} \leqslant \frac{1}{n^m} < \frac{1}{n^p}.$$

Однако легко убедиться, что функция $f(x)$ не имеет производной.

24. Обобщение условий Липшица. Предыдущий пример естественно наводит на мысль о выяснении дифференциальной природы функций, которые не для всех, но для бесчисленного множества значений n допускают приближение того же порядка, что и функции, обладающие производными. Как мы увидим, эти функции обладают свойствами, аналогичными условиям Липшица.

Пусть $f(x)$ будет некоторая непрерывная на отрезке AB функция. Обозначим через $\delta_1(\varepsilon)$ максимум колебания функции $f(x)$ в любом промежутке длины ε на отрезке, или, другими словами, максимум разности $|f(x+h) - f(x)|$ при $|h| \leqslant \varepsilon$. Функция $\delta_1(\varepsilon)$ будет, очевидно, непрерывной, неотрицательной и монотонной (неубывающей); при этом $\delta_1(0) = 0$. Обыкновенное условие Липшица степени s выражает, что существует такое определенное число k, что при всяком ε

$$\delta_1(\varepsilon) < k\varepsilon^s. \tag{19}$$

Мы скажем, что функция $f(x)$ удовлетворяет обобщенному условию Липшица степени s, если существует бесчисленное множество значений ε, для которых неравенство (19) соблюдено.

Точно так же вместо максимума первой разности $|f(x+h) - f(x)|$, при $|h| \leqslant \varepsilon$, можно рассматривать максимумы последовательных разностей: $\delta_2(\varepsilon) = \max|f(x+2h) - 2f(x+h) + f(x)|$, $\delta_3(\varepsilon) = \max|f(x+3h) - 3f(x+2h) + 3f(x+h) - f(x)|$ и т. д. при $|h| \leqslant \varepsilon$.

Если для бесчисленного множества значений ε имеет место неравенство

$$\delta_i(\varepsilon) < k\varepsilon^s, \tag{19'}$$

то мы будем говорить, что функция $f(x)$ удовлетворяет на отрезке AB обобщенному условию Липшица i-го вида степени s. Легко убедиться, что если $\delta_i(\varepsilon) > 0$ (для всех $\varepsilon > 0$), то $s \leqslant i$. Заметим, что в случае существования конечной производной i-го порядка на отрезке AB условие (19′) соблюдается для всех ε при $s = i$.

25. Теорема. *Если существует бесчисленное множество значений n, для которых наилучшее приближение*[1] *E_n на отрезке $[-1, +1]$ удовлетворяет неравенству $E_n < \frac{A}{n^p}$, то функция $f(x)$ на всяком отрезке ab внутри отрезка $[-1, +1]$ удовлетворяет обобщенным условиям Липшица i-го вида степени $s_i = \frac{ip}{i+p}$.*

---

[1] При помощи многочленов степени n. Та же теорема (см. § 17) остается в силе и для тригонометрических сумм.

Рассмотрим сначала функцию $\delta_1(\varepsilon)$. Обозначая через P_n приближенный многочлен степени n, удовлетворяющий неравенству

$$|f(x) - P_n(x)| < \frac{A}{n^p}, \tag{20}$$

будем, очевидно, иметь для бесчисленного множества значений n

$$|P_n(x)| < M + \frac{A}{n^p} < 2M,$$

где M — максимум $|f(x)|$.

В таком случае на всяком определенном отрезке ab внутри отрезка $[-1, +1]$

$$|P'_n(x)| < RMn,$$

где R — некоторый численный множитель (§ 3).

Поэтому

$$|P_n(x_1) - P_n(x_2)| < RMn\varepsilon,$$

если $|x_1 - x_2| \leqslant \varepsilon$. Но значения x_1 и x_2 можно выбрать так, что

$$|f(x_1) - f(x_2)| = \delta_1(\varepsilon).$$

Следовательно,

$$|f(x_1) - P_n(x_1) + P_n(x_2) - f(x_2)| > \delta_1(\varepsilon) - RMn\varepsilon.$$

Сопоставляя это неравенство с неравенством (20), находим

$$\frac{2A}{n^p} > \delta_1(\varepsilon) - RMn\varepsilon,$$

или

$$\delta_1(\varepsilon) < \frac{2A}{n^p} + RMn\varepsilon. \tag{21}$$

Положим в этом неравенстве $\varepsilon = \frac{1}{n^{1+p}}$. Получим

$$\delta_1(\varepsilon) < \frac{2A}{n^p} + \frac{RM}{n^p} = (2A + RM)\,\varepsilon^{\frac{p}{1+p}}.$$

Таким образом, для $i = 1$, теорема доказана.

Достаточно будет рассмотреть еще случай $i = 2$, чтобы убедиться, что тот же прием доказательства применим для всякого i.

На основании § 11 имеем $|P''_n(x)| < R_1Mn^2$, где R_1 — численный коэффициент, зависящий только от отрезка AB. Поэтому, при $|h| \leqslant \varepsilon$,

$$|P_n(x + 2h) - 2P_n(x + h) + P_n(x)| < 2R_1Mn^2\varepsilon^2;$$

но, выбирая x соответствующим образом, имеем

$$|f(x + 2h) - 2f(x + h) + f(x)| = \delta_2(\varepsilon),$$

откуда

$$|f(x + 2h) - P_n(x + 2h) - 2[f(x + h) - P(x + h)] + f(x) - P_n(x)| >$$
$$> \delta_2(\varepsilon) - 2R_1Mn^2\varepsilon^2.$$

Следовательно,

$$\frac{4A}{n^p} > \delta_2(\varepsilon) - 2R_1Mn^2\varepsilon^2,$$

или

$$\delta_2(\varepsilon) < \frac{4A}{n^p} + 2R_1Mn^2\varepsilon^2. \tag{22}$$

Полагая в неравенстве (22)

$$\varepsilon = \frac{1}{n^{1+\frac{p}{2}}},$$

получим

$$\delta_2(\varepsilon) < \frac{4A}{n^p} + \frac{2R_1M}{n^p} = (4A + 2R_1M)\,\varepsilon^{\frac{2p}{2+p}}.$$

Таким образом, теорема доказана также для $i = 2$, и ясно, что то же рассуждение применимо для всякого i.

26. Приложения предшествующей теоремы. Функция, рассмотренная нами в § 23, обладала свойством, что при всяком p есть бесчисленное множество значений n, для которых $E_n < \frac{1}{n^p}$. Таким образом, вследствие только что доказанной теоремы заключаем, что указанная функция удовлетворяет обобщенному условию Липшица вида i любой степени $s < i$.

Не останавливаясь на более детальном изучении этих своеобразных функций, применим предыдущую теорему к определению низшего предела $E_n|x|$. Для этого заметим, что ни при каком i функция $|x|$ не удовлетворяет обобщенному условию Липшица степени выше первой. В самом деле, при $x = -h$,

$$|x + nh| - n|x + (n-1)h| + \ldots + (-1)^n|x| = (-1)^n 2h,$$

так что $\delta_i(\varepsilon) \geqslant 2\varepsilon$.

Следовательно, если есть *бесчисленное множество* значений n, для которых $E_n[|x|] < \frac{1}{n^p}$, то показатель p должен обладать свойством, что, при всяком i,

$$s = \frac{ip}{i+p} \leqslant 1,$$

откуда

$$p \leqslant \frac{i}{i-1}.$$

Таким образом, *p не может быть более единицы.*

27. Условие Дини и Липшица. Условием Дини и Липшица называют свойство (которым обладают некоторые непрерывные функции), заключающееся в том, что произведение

$$\delta_1(\varepsilon)\log\varepsilon$$

стремится к нулю вместе с ε. Мы будем говорить, что функция удовлетворяет обобщенному условию Дини-Липшица, если возможно выбрать бесчисленное множество значений ε таким образом, чтобы указанное произведение $\delta_1(\varepsilon)\log\varepsilon$ стремилось к нулю вместе с ε. Приняв эти определения, докажем, что *функция, для которой $E_n \log n$ стремится к нулю для бесчисленного множества значений n, удовлетворяет обобщенному условию Дини-Липшица; если $E_n \log n$ стремится к нулю при всех значениях n, то функция удовлетворяет обыкновенному условию Дини-Липшица.*

В самом деле, повторяя рассуждение § 25, приходим немедленно к обобщению неравенства (21)

$$\delta_1(\varepsilon) < 2E_n + kn\varepsilon, \tag{21'}$$

где k — постоянная (не зависящая от n). Применяя это неравенство к настоящему случаю, когда $E_n \log n = \beta_n$ стремится к нулю (причем, не нарушая общности, можно принять, что $\beta_n > \frac{\log n}{n^2}$), положим

$$\varepsilon = \frac{\beta_n}{n \log n};$$

тогда из (21') получим

$$\log n \cdot \delta_1(\varepsilon) < \beta_n(2+k),$$

и так как

$$|\log\varepsilon| < 2\log n,$$

то, следовательно,

$$|\log\varepsilon|\,\delta_1(\varepsilon) < \beta_n(4+2k) \tag{23}$$

для бесчисленного множества значений n. Таким образом, для соответствующего бесчисленного множества значений ε, произведение $\log\varepsilon \cdot \delta_1(\varepsilon)$ стремится к нулю. Если же неравенство (23) соблюдается для всякого целого n, то ясно, что $\log\varepsilon \cdot \delta_1(\varepsilon)$ будет всегда стремиться к нулю вместе с ε. Что и требовалось доказать.

28. Теорема Лебега. В своей большой работе[1] «Sur les intégrales singulières» Лебег доказывает следующую теорему: *Если рассматривается совокупность всех непрерывных функций $f(x)$, для которых $|f(x)| \leqslant M$, то, при всяком n, верхним пределом $E_n[f(x)]$ является M* (т. е. среди функций $f(x)$ есть такие, для которых $E_n[f(x)] > M - \alpha$, как бы мало ни было α, и, кроме того, для всех функций $E_n(f) \leqslant M$). При помощи неравенства (21') эту теорему чрезвычайно легко доказать. В самом деле, как бы мало ни было $\varepsilon = \frac{\alpha}{kn}$, среди рассматриваемых функций можно выбрать такую, что $\delta_1(\varepsilon) = 2M$. Поэтому, вследствие неравенства (21'), для этой функции

$$E_n > M - \alpha,$$

что и требовалось доказать (само собой понятно, что для всех функций рассматриваемой совокупности $E_n[f(x)] \leqslant M$).

Однако теорема Лебега оставляет открытым интересный вопрос: существует ли такой ряд чисел $\alpha_1, \alpha_2, \ldots, \alpha_n, \ldots$, имеющих пределом 0,

---

[1] «Ann. de Toulouse», 1909.

чтобы для всякой данной непрерывной функции $f(x)$ можно было указать независимое от n число R_f достаточно большое, чтобы было $E_n < R_f \alpha_n$?

На основании теоремы Лебега можно лишь утверждать, что если бы ряд чисел α_n существовал, то, для всей совокупности непрерывных функций $f(x)$, не превышающих M по абсолютному значению, множитель R_f не имел бы (конечного) верхнего предела. Действительно, легко убедиться, что теорема Лебега остается справедливой, если совокупность непрерывных функций заменить одними лишь многочленами; а между тем, каковы бы ни были числа α_n, например $\alpha_n = \frac{1}{2^n}$, для всякого определенного многочлена $P(x)$ возможно, конечно, указать число R_P так, чтобы $E_n < R_P \alpha_n$.

Неравенство (21') дает немедленно *отрицательный* ответ на поставленный вопрос. В самом деле, если для некоторой функции $E_n < R\alpha_n$, то $\delta_1(\varepsilon) < 2R\alpha_n + kn\varepsilon$. Полагая $\alpha_n > \frac{1}{n}$ (что мы вправе сделать, не нарушая общности), берем $\varepsilon = \frac{1}{n^2}$; в таком случае $\delta_1\left(\frac{1}{n^2}\right) < 2R\alpha_n + \frac{k}{n} < (2R + k)\alpha_n$. Но такому неравенству при всяком n не может удовлетворить, например, ни одна непрерывная функция $f(x)$, которая при $x = \frac{1}{n^2}$ обращается в $\sqrt{\alpha_n}$, так как для этой функции $\delta_1\left(\frac{1}{n^2}\right) \geqslant \sqrt{\alpha_n}$.

29. Теорема. *Если для всякого n наилучшее приближение функции $f(x)$ на отрезке $[-1, +1]$ удовлетворяет неравенству $E_n < M\rho^n$, то функция $f(x)$ голоморфна внутри эллипса, фокусами которого служат точки -1, $+1$, а полусумма осей равна $\frac{1}{\rho}$*:

В самом деле, обозначая через $P_n(x)$ многочлен степени n, для которого

$$|f(x) - P_n(x)| < M\rho^n,$$

можем написать

$$f(x) = P_1(x) + [P_2(x) - P_1(x)] + \ldots + [P_n(x) - P_{n-1}(x)] + \ldots; \qquad (24)$$

при этом

$$|P_n(x) - P_{n-1}(x)| < 2M\rho^{n-1}$$

на отрезке $[-1, +1]$. Поэтому во всякой точке H эллипса, сумма полуосей которого равна $\frac{1}{\rho_1} < \frac{1}{\rho}$, а фокусы находятся в точках -1, $+1$, имеем (§ 7)

$$|P_n(x) - P_{n-1}(x)| < \frac{2M}{\rho_1}\left(\frac{\rho}{\rho_1}\right)^{n-1}.$$

Следовательно, ряд (24) равномерно сходится во всякой области внутри эллипса, сумма полуосей которого равна $1/\rho$, а потому функция $f(x)$ голоморфна.

(Обратная теорема будет доказана в третьей части.)

Часть вторая

ПРИБЛИЖЕННОЕ ВЫЧИСЛЕНИЕ МНОГОЧЛЕНОВ, НАИМЕНЕЕ УКЛОНЯЮЩИХСЯ В ДАННОМ ПРОМЕЖУТКЕ ОТ ДАННОЙ ФУНКЦИИ

Глава III

ОБЩИЙ МЕТОД

30. Введение. Идея метода приближенного вычисления многочленов, наименее уклоняющихся от данной функции, которому посвящена эта глава, состоит в том, чтобы соответствующим образом использовать уже известные многочлены, наименее уклоняющиеся от некоторых данных функций. Иногда, вместо других функций, целесообразно будет вводить аналогичные многочленам выражения, наименее уклоняющиеся от той же самой функции. И в том и в другом случае непрерывный переход от известного к неизвестному совершается посредством аналитического продолжения; при этом как для практических применений, так и для теоретических выводов, весьма важно выбрать исходный пункт таким образом, чтобы первые же приближения обладали уже значительной точностью.

Напомним сначала классические результаты, вытекающие из исследований Чебышева *.

а) *Существует один и только один многочлен P_n степени не выше n, наименее уклоняющийся в промежутке AB от данной непрерывной функции $f(x)$.*

б) *Из всех многочленов степени не выше n только многочлен $P_n(x)$ обладает тем свойством, что разность $f(x)-P_n(x)$ достигает не менее чем $n+2$ раза своего максимума, с последовательно чередующимися знаками, в рассматриваемом промежутке.*

Из последнего предложения вытекает, что если бы $|f(x)-P_n(x)|$ достигало своего максимума более чем $n+2$ раза, а именно $n+2+k$ раз, то многочлен $P_n(x)$ был бы в то же время единственным наименее уклоняющимся от функции $f(x)$ среди всех многочленов степени не выше $n+k$. Таким образом, задача определения многочленов $P_n(x)$ по существу нисколько не суживается, если ограничимся только теми значениями n, для которых разность $|f(x)-P_n(x)|$ достигает своего максимума в $n+2$ точках.

31. Обобщения. Рассмотрим ряд степеней $x^{\alpha_0}, x^{\alpha_1}, \ldots, x^{\alpha_n}$, где $0 \leqslant \alpha_0 < \alpha_1 < \ldots < \alpha_n$, и составим суммы $A_0x^{\alpha_0}+\ldots+A_nx^{\alpha_n}$ с произвольными коэффициентами $A_0, \ldots, A_n$. Если сумма

$$R_n(x)=B_0x^{\alpha_0}+\ldots+B_nx^{\alpha_n}$$

из всех сумм указанного вида наименее уклоняется от функции $f(x)$ в промежутке AB, то $R_n(x)$ называется суммой вида $\sum_{i=0}^{n} A_ix^{\alpha_i}$, наименее

---

* Строгое доказательство этих результатов читатель может найти в монографии «Э. П.», глава I, §§ 4—5. (*Ред.*)

уклоняющейся от функции $f(x)$ в промежутке AB. Относительно отрезка AB необходимо ввести ограничение, а именно: на всем отрезке $x \geqslant 0$. Благодаря этому ограничению числа x^{α_i} будут всегда иметь вполне определенное арифметическое значение. Рассуждениями, совершенно подобными тем, которые читатель найдет в книге Бореля «Leçons sur les fonctions de variables réelles» для случая, когда $\alpha_i = i$, можно доказать существование суммы $R_n(x)$, наименее уклоняющейся от данной непрерывной функции $f(x)$, и в общем случае. Для доказательства же того, что эта сумма — единственная, нам необходимо доказать предварительно следующую лемму, являющуюся обобщением теоремы Декарта.

32. Лемма. *Число положительных корней уравнения*

$$Q(x) = a_0 x^{\alpha_0} + a_1 x^{\alpha_1} + \ldots + a_n x^{\alpha_n} = 0, \tag{25}$$

где $0 \leqslant \alpha_0 < \alpha_1 < \ldots < \alpha_n$, *не может превышать числа перемен знаков ряда* $a_0, a_1, \ldots, a_n$.

В случае, когда числа α_i целые, высказанное предложение является прямым следствием из известной теоремы Декарта. Точно так же случай, когда числа α_i рациональные, посредством подстановки $x^{1/p} = y$ приводится к предшествующему.

Предполагая далее, что числа α_i какие угодно, будем, однако, считать, что все положительные корни уравнения (25) различны между собой. Если бесконечно мало изменить показатели уравнения, то бесконечно мало изменятся и корни; поэтому каждому положительному корню данного уравнения будет соответствовать один положительный корень измененного уравнения и обратно, ибо комплексные корни вещественного уравнения — всегда попарно сопряженные. Таким образом, число положительных корней данного уравнения то же, что измененного, но в этом последнем всегда можно предположить показатели рациональными. Следовательно, число простых положительных корней уравнения (25) не может превышать числа перемен знаков ряда $a_0, a_1, \ldots, a_n$.

Тем же способом убеждаемся, что число различных положительных корней нечетной кратности не может превышать числа перемен знаков ряда $a_0, a_1, \ldots, a_n$. Но нам остается еще показать, что число корней, взятых с их степенью кратности, также не превышает упомянутого числа. Для этого составляем уравнение

$$\frac{d}{dx}\left[x^{1-\alpha_0} Q(x)\right] = 0 \tag{25'}$$

и замечаем, что каждый кратный корень уравнения (25) является в то же время корнем уравнения (25′) со степенью кратности на одну единицу меньшей; кроме этих корней, уравнение (25′) имеет еще не менее различных положительных корней нечетной кратности, чем уравнение (25). Таким образом, число корней уравнения (25′), взятых с их степенью кратности, не меньше числа корней уравнения (25), взятых с их степенью кратности; число же различных корней уравнения (25′) нечетной кратности не менее числа всех различных корней нечетной кратности уравнения (25), увеличенного на число различных двойных корней последнего уравнения. Из этого следует, что если мы поступим с уравнением (25′), как с уравнением (25) и т. д., то придем, наконец, к уравнению, число различных корней которого нечетной кратности будет не менее числа корней уравнения (25), взятых с их степенью кратности. Но это последнее уравнение будет того же вида

$$Q_1(x) = b_0 + b_1 x^{\beta_1} + \ldots + b_n x^{\beta_n} = 0, \tag{25''}$$

что и уравнение (25), причем будем иметь $b_i a_i > 0$, так что число перемен знака в ряде $b_0, b_1, \ldots, b_n$ то же, что и в ряде $a_0, a_1, \ldots, a_n$. Поэтому число различных корней нечетной кратности уравнения (25″) не превышает числа перемен знака в ряде $a_0, a_1, \ldots, a_n$; тем более и общее число положительных корней уравнения (25), взятых с их степенью кратности, не может превышать числа перемен знаков ряда $a_0, a_1, \ldots, a_n$.

Следствие. *Число положительных корней уравнения* (25) *не превышает* n.

33. Теорема. *Существует только одна сумма степеней*

$$R_n(x) = \sum_{i=0}^{n} B_n x^{\alpha_n},$$

наименее уклоняющаяся в промежутке AB от функции $f(x)$. При этом разность $f(x) - R_n(x)$ достигает не менее чем $n+2$ раз своего наибольшего абсолютного значения, последовательно меняя свой знак. Исключение может представляться лишь в том случае, если это наибольшее значение равно $|f(0)|$, при $\alpha_0 > 0$.

В самом деле, обозначая через $x_1, x_2, \ldots, x_k$ возрастающий ряд чисел, для которых разность $f(x) - R_n(x)$ достигает последовательно наибольшего абсолютного значения, меняя знак, находим на основании соображений, которыми мы пользовались несколько раз в главе I, что уравнения

$$\begin{aligned} b_0 x_1^{\alpha_0} + \ldots + b_n x_1^{\alpha_n} &= f(x_1) - R_n(x_1) = \pm L, \\ b_0 x_2^{\alpha_0} + \ldots + b_n x_2^{\alpha_n} &= f(x_2) - R_n(x_2) = \mp L, \\ \ldots\ldots\ldots\ldots & \ldots\ldots\ldots\ldots\ldots\ldots \\ b_0 x_k^{\alpha_0} + \ldots + b_n x_k^{\alpha_n} &= f(x_k) - R_n(x_k) = \mp (-1)^k L \end{aligned} \tag{26}$$

должны быть несовместимыми, если $R_n(x)$ представляет собой сумму указанного вида, наименее уклоняющуюся от $f(x)$ на отрезке AB. Но легко убедиться, что уравнения (26) были бы совместимы в случае $k < n+2$, ибо ни один из определителей

$$\delta_p = \begin{vmatrix} x_1^{\alpha_0} & \ldots & x_1^{\alpha_p} \\ \cdot & \cdot\ \cdot\ \cdot & \cdot \\ \cdot & \cdot\ \cdot\ \cdot & \cdot \\ x_{p+1}^{\alpha_0} & \ldots & x_{p+1}^{\alpha_p} \end{vmatrix} \equiv a_0 x_{p+1}^{\alpha_0} + \ldots + a_p x_{p+1}^{\alpha_p}$$

не может быть равен нулю: это справедливо при $p = 0$; но если $\delta_{p-1} \gtrless 0$, то и $\delta_p \gtrless 0$, так как в уравнении

$$a_0 x^{\alpha_0} + \ldots + a_p x^{\alpha_p} = 0$$

коэффициент $a_p = \delta_{p-1} \gtrless 0$, и потому это уравнение не соблюдено тождественно; но оно имеет p положительных[1] корней $x_1, \ldots, x_p$, и следовательно

$$\delta_p = a_0 x_{p+1}^{\alpha_0} + \ldots + a_p x_{p+1}^{\alpha_p} \gtrless 0.$$

Таким образом, вторая часть теоремы доказана. Допустим теперь, что кроме $R_n(x)$ существует еще сумма $R'_n(x)$, наименее уклоняющаяся от данной функции $f(x)$. В таком случае, в силу только что доказанного, разность

$$Q(x) = R_n(x) - R'_n(x)$$

в точках $x_1, x_2, \ldots, x_{n+2}$ будет последовательно менять знак или равняться нулю.

Поэтому, если $Q(x_i) \gtrless 0$, $Q(x_{i+1}) \gtrless 0, \ldots, Q(x_{i+k+1}) \gtrless 0$, то между x_i и x_{i+k+1} по крайней мере $k+1$ корней; точно так же, если $Q(x_{i+1}) = \ldots$ $\ldots = Q(x_{i+k}) = 0$, то число корней (взятых с их степенью кратности) не менее $k+1$, так как это число не менее k и, кроме того, разность между ним и $k+1$ должна быть четной. Отсюда вытекает, что общее число положительных корней уравнения

$$Q(x) = 0$$

не менее $n+1$, что невозможно на основании леммы § 32. Таким образом, существует только одна сумма $R_n(x)$, наименее уклоняющаяся от функции $f(x)$ в данном промежутке.

Примечание. Необходимо помнить, что применение доказанной теоремы в случае $\alpha_0 > 0$ и $x \geqslant 0$ законно лишь при условии $f(0) = 0$.

34. Обобщенная теорема Валле Пуссена[2]. *Отклонение* $|f(x) - R_n(x)|$ *не может в промежутке AB оставаться постоянно менее наименьшего из значений* $|f(x) - P_n(x)|$ *в $n+2$ точках, где $f(x) - P_n(x)$ последовательно меняет знак, если $P_n(x)$ — сумма того же вида, что $R_n(x)$.*

В самом деле, допустим противное. Тогда в $n+2$ точках разность

$$Q(x) = P_n(x) - R_n(x) = [f(x) - R_n(x)] - [f(x) - P_n(x)]$$

последовательно меняет знак, и, следовательно, уравнение $Q(x) = 0$ имеет по крайней мере $n+1$ положительных корней, что невозможно.

Замечание. Эта теорема была доказана Валле Пуссеном в случае многочленов, причем промежуток AB тогда может быть какой угодно;

---

[1] Если $x_1 = 0$, то коэффициент $a_0 = 0$, и потому сумма δ_p, состоящая только из p слагаемых, не имеет более $p-1$ корней.

[2] De la Vallée Poussin. Sur les polynomes d'approximation et la représentation approchée d'un angle. «Bull. Acad. Belgique», Décembre 1910.

очевидно, что данное здесь доказательство пригодно и для упомянутого случая.

35. Определения. Точки $x_1, x_2, \ldots, x_k$, где $|f(x) - R_n(x)|$ достигает наибольшего значения, мы будем называть *точками отклонения*.

Следует заметить, что расположение точек отклонения на отрезке AB может быть четырех родов. А именно: 1-го рода, когда оба конца A и B являются точками отклонения; 2-го рода, когда только A — точка отклонения; 3-го рода, когда только B — точка отклонения; 4-го рода, когда все точки отклонения находятся внутри отрезка AB. Расположение 1-го рода является вообще наиболее общим случаем. Однако если $\alpha_0 > 0$ и $A = 0$ (что большей частью будет иметь место в дальнейших приложениях), то расположения 1-го и 2-го родов будут невозможны, так как, вследствие примечания к теореме § 33, необходимо, чтобы было $f(0) = 0$; в этом случае обыкновенно представляется расположение 3-го рода.

36. Основная теорема А *. *Если сумма* $P(x, \lambda) = \sum\limits_0^n a_i x^{\alpha_i}$, *наименее уклоняющаяся на отрезке* AB *от голоморфной функции* $\lambda f(x) + (1 - \lambda)\varphi(x)$, *имеет* $n + 2$ *точки отклонения одного и того же рода при всяком* λ $(0 \leqslant \lambda \leqslant 1)$, *то коэффициенты суммы* $P(x, \lambda)$ *и абсциссы точек отклонения, так же как и наименьшее отклонение, являются голоморфными функциями параметра* λ *при условии, что во внутренних точках отклонения* $F''_{x^2} \gtreqless 0$, *где положено*

$$F(x, \lambda) = \lambda f(x) + (1 - \lambda)\varphi(x) - P(x, \lambda).$$

Достаточно будет рассмотреть, например, случай 1-го рода расположения точек отклонения; другими словами, предположим, что, при всяком λ $(0 \leqslant \lambda \leqslant 1)$, концы отрезка A и B являются точками отклонения. В таком случае для определения $P(x, \lambda)$ мы будем иметь $2n + 2$ уравнения[1]

$$\begin{aligned}
&F_x'(x_i, \lambda) = \lambda f'(x_i) + (1 - \lambda)\varphi'(x_i) - P'(x_i, \lambda) = 0 \qquad (i = 1, \ldots, n)\\
&[F(x_i, \lambda)]^2 = L^2,\\
&[F(A, \lambda)]^2 = L^2,\\
&[F(B, \lambda)]^2 = L^2
\end{aligned} \tag{27}$$

с $2n + 2$ неизвестными: внутренними точками отклонения $x_1, x_2, \ldots, x_n$ (расположенными в возрастающем порядке), коэффициентами $a_0, a_1, \ldots, a_n$ и отклонением L (с чередующимися знаками).

При всяком определенном значении $\lambda = \lambda_0$, система уравнений (27) имеет одну вполне определенную систему вещественных решений, соответствующую единственной сумме, наименее уклоняющейся от функции

---

* Обобщение этой теоремы см. «Э. П.» (стр. 40). (*Ред.*)

[1] Если $\alpha_i = i$, то промежуток AB произволен; в общем же случае предполагается, что $B > A \geqslant 0$.

$\lambda f(x) + (1 - \lambda)\varphi(x)$. Поэтому, если функциональный определитель уравнений (27) относительно неизвестных отличен от нуля, то все неизвестные будут аналитическими функциями параметра λ. Таким образом, для доказательства теоремы достаточно будет показать, что вышеупомянутый функциональный определитель не равен нулю. Но этот определитель Δ равен

$$\begin{vmatrix} +1 & \overbrace{0 \quad 0 \quad \ldots 0}^{n} & \overbrace{A^{\alpha_0} \quad A^{\alpha_1} \ldots A^{\alpha_n}}^{n+1} \\ -1 & 0 \quad 0 \quad \ldots 0 & x_1^{\alpha_0} \quad x_1^{\alpha_1} \ldots x_1^{\alpha_n} \\ \cdots & \cdots\cdots\cdots & \cdots\cdots\cdots \\ (-1)^{n+1} & 0 \quad 0 \quad \ldots 0 & B^{\alpha_0} \quad B^{\alpha_1} \ldots B^{\alpha_n} \\ 0 & F''_{x_1^2} \quad 0 \quad \ldots 0 & \ldots\ldots\ldots \\ 0 & 0 \quad F''_{x_2^2} \ldots 0 & \ldots\ldots\ldots \\ \cdots & \cdots\cdots\cdots & \cdots\cdots\cdots \\ 0 & 0 \quad 0 \quad \ldots F''_{x_n^2} & \ldots\ldots\ldots \end{vmatrix} =$$

$$= \pm F''_{x_1^2} F''_{x_2^2} \cdots F''_{x_n^2} \left[\Delta_A + \Delta_{x_1} + \cdots + \Delta_B\right],$$

где

$$\Delta_A = \begin{vmatrix} x_1^{\alpha_0}\, x_1^{\alpha_1} \ldots x_1^{\alpha_n} \\ x_2^{\alpha_0}\, x_2^{\alpha_1} \ldots x_2^{\alpha_n} \\ \ldots\ldots\ldots \\ B^{\alpha_0}\, B^{\alpha_1} \ldots B^{\alpha_n} \end{vmatrix} > 0, \qquad \Delta_{x_1} = \begin{vmatrix} A^{\alpha_0} \ldots A^{\alpha_n} \\ x_2^{\alpha_0} \ldots x_2^{\alpha_n} \\ \ldots\ldots \\ B^{\alpha_0} \ldots B^{\alpha_n} \end{vmatrix} > 0 \quad \text{и т. д.}$$

Следовательно, $\Delta_A + \Delta_{x_1} + \ldots + \Delta_B > 0$, а потому $\Delta \gtrless 0$, что и требовалось доказать.

Примечание. Можно заметить, что при доказательстве никакой роли не играло то обстоятельство, что параметр λ входит в виде $\lambda f(x) + (1 - \lambda)\varphi(x)$; все рассуждение остается в силе, если рассматриваемая функция голоморфна относительно λ. Это замечание приводит нас к другой полезной для применений формулировке основной теоремы.

37. Основная теорема В. *Если сумма* $P(x, \lambda) = \sum\limits_{i=0}^{n} a_i x^{\alpha_i}$, *наименее уклоняющаяся на отрезке* AB *от функции* $f(x) + (\lambda - 1) Q(x)$, *имеет* $n + 2$ *точки отклонения одного и того же рода, при всяком* λ $(0 \leqslant \lambda \leqslant 1)$, *и* $F''_{x^2} = f''_{x^2} + (\lambda - 1) Q''_{x^2} - P''_{x^2} \gtrless 0$ *во всех внутренних точках отклонения, то, полагая, что* $Q(x) = \sum\limits_{i=0}^{n} b_i x^{\beta_i}$ *есть сумма, наименее уклоняющаяся от* $f(x)$ *на отрезке* AB, *коэффициенты* $P(x, \lambda)$ *так же, как абсциссы точек отклонения и отклонение, суть голоморфные функции* λ, *причем* $P(x, 0) = 0$.

38. Применение основных теорем. Теоремой А следует пользоваться, если хотят определить сумму $\sum\limits_{i=0}^{n} a_i x^{\alpha_i}$, наименее уклоняющуюся от $f(x)$, зная сумму того же вида, наименее уклоняющуюся от другой данной функции $\varphi(x)$. Теорему B применяют, когда хотят определить сумму $\sum\limits_{i=0}^{n} a_i x^{\alpha_i}$, наименее уклоняющуюся от $f(x)$, зная сумму $\sum\limits_{i=0}^{n} b_i x^{\beta_i}$, составленную из других степеней x, наименее уклоняющуюся от той же функции.

Нетрудно понять общий прием пользования упомянутыми теоремами, к изложению которого мы сейчас перейдем, обратив особое внимание на вычисление функции $L(\lambda)$, представляющей наименьшее отклонение для различных значений параметра λ.

Если данная функция $f(x)$ неаналитическая, то предварительно надо будет заменить ее аналитической, достаточно мало отличающейся от данной в рассматриваемом промежутке. Таким образом, в дальнейшем мы все время предполагаем данную функцию $f(x)$ аналитической. Для применения теоремы A выбираем некоторую другую аналитическую функцию $\varphi(x)$, для которой наименее уклоняющаяся сумма того же вида $P(x) = P(x, 0) = \sum\limits_{i=0}^{n} b_i x^{\alpha_i}$ известна и, кроме того, обладающую свойством, что функция $F(x, \lambda) = \lambda f(x) + (1 - \lambda)\varphi(x) - P(x, \lambda)$ удовлетворяет условию, что во всех внутренних точках отклонения $\frac{\partial^2 F}{\partial x^2} \gtrless 0$, причем род расположения точек отклонения независим от λ.

После этого вычисляем последовательные производные $\frac{\partial P}{\partial \lambda}$, $\frac{\partial^2 P}{\partial \lambda^2}$ и т. д. для $\lambda = 0$. Многочлен или сумма степеней $P(x, \lambda)$ разлагается, таким образом, в строку Тэйлора относительно λ, представляющую голоморфную функцию при всех значениях λ $(0 \leqslant \lambda \leqslant 1)$, значение которой $P(x, 1)$, при $\lambda = 1$, равно искомой сумме, наименее уклоняющейся от функции $f(x)$. Если строка Тэйлора имеет радиус сходимости меньший единицы, то для вычисления $P(x, 1)$ можно во всяком случае применить способ суммирования Миттаг-Леффлера. Последовательные производные $\frac{\partial P}{\partial \lambda} = P_1$, $\frac{\partial^2 P}{\partial \lambda^2} = P_2$ и т. д., при $\lambda = 0$ представляющие собой суммы степеней того же вида, что и $P(x)$, последовательно вычисляются следующим образом.

Прежде всего замечаем, что в $n + 2$ точках отклонения x_i, соответствующих $\lambda = 0$ и, по предположению, известных, имеем

$$\pm L(0) = \varphi(x_i) - P(x_i, 0).$$

Затем, так как в этих точках $\frac{\partial F}{\partial x_i} = 0$ или же $\frac{dx_i}{d\lambda} = 0$, то

$$\pm \frac{dL(0)}{d\lambda} = \frac{\partial F(x_i, 0)}{\partial \lambda} = f(x_i) - \varphi(x_i) - P_1(x_i), \tag{28}$$

причем знак левой части равенства (28) для определенного i всегда тот же, что и в предыдущем равенстве.

Таким образом, для определения $\frac{dL}{d\lambda}$ и $n+1$ коэффициентов суммы P_1 имеем $n+2$ линейных уравнения с $n+2$ неизвестными; при этом определитель, составленный из коэффициентов этих уравнений,

$$\delta = \begin{vmatrix} 1 & x_0^{\alpha_0} \ldots x_0^{\alpha_n} \\ -1 & x_1^{\alpha_0} \ldots x_1^{\alpha_n} \\ \ldots & \ldots \ldots \ldots \\ \ldots & \ldots \ldots \ldots \\ (-1)^{n+1} & x_{n+1}^{\alpha_0} \ldots x_{n+1}^{\alpha_n} \end{vmatrix}$$

отличен от нуля, так что для каждого из неизвестных получается всегда одно вполне определенное значение.

Для определения $\frac{d^2L}{d\lambda^2}$ и P_2 замечаем, что если x_i представляет собой неподвижный конец отрезка AB, т. е. совпадает с A или с B, то

$$\pm \frac{d^2L}{d\lambda^2} = \frac{\partial^2 F(x_i, 0)}{\partial \lambda^2} = - P_2(x_i); \tag{29}$$

если же точка x_i внутренняя, то

$$\pm \frac{d^2L}{d\lambda^2} = \frac{\partial^2 F(x_i, 0)}{\partial \lambda^2} + 2\, \frac{\partial^2 F(x_i, 0)}{\partial \lambda\, \partial x} \frac{dx_i}{d\lambda} + \frac{\partial^2 F(x_i, 0)}{\partial x^2} \left(\frac{dx_i}{d\lambda}\right)^2;$$

и так как

$$\frac{\partial^2 F(x_i, 0)}{\partial x\, \partial \lambda} d\lambda + \frac{\partial^2 F(x_i, 0)}{\partial x^2} dx_i = 0$$

следовательно,

$$\pm \frac{d^2L}{d\lambda^2} = \frac{\partial^2 F}{\partial \lambda^2} - 2\, \frac{\left(\frac{\partial^2 F}{\partial x\, \partial \lambda}\right)^2}{\frac{\partial^2 F}{\partial x^2}} + \frac{\left(\frac{\partial^2 F}{\partial x\, \partial \lambda}\right)^2}{\frac{\partial^2 F}{\partial x^2}} = \frac{\partial^2 F}{\partial \lambda^2} - \frac{\left(\frac{\partial^2 F}{\partial x\, \partial \lambda}\right)^2}{\frac{\partial^2 F}{\partial x^2}} = -P_2 - \frac{\left(\frac{\partial^2 F}{\partial x\, \partial \lambda}\right)^2}{\frac{\partial^2 F}{\partial x^2}} \tag{29'}$$

(замечание относительно знаков то же, что в равенствах (28)).

Уравнения (29) и (29′) представляют снова систему $n+2$ линейных уравнений с $n+2$ неизвестными: коэффициентами многочлена P_2 и $\frac{d^2L}{d\lambda^2}$. При этом определителем этих уравнений служит тот же определитель δ, отличный от нуля, что и раньше.

Тем же способом можно вычислить и последующие производные; это вычисление всегда приводится к решению системы $n+2$ линейных уравнений с $n+2$ неизвестными, у которых коэффициенты при неизвестных для производных всех порядков одни и те же.

При применении теоремы В вычисления совершенно аналогичны; в частности, равенства (29) и (29′) остаются без изменений.

39. Вывод двух неравенств. В приложениях, составляющих содержание следующей главы, мы будем ограничиваться первыми двумя членами строки Тэйлора, а именно: за приближенное значение искомого отклонения $L(1)$ мы будем брать $L(0)$ или $L(0)+\frac{dL(0)}{d\lambda}$. Первым из этих значений нам придется пользоваться в различных частных случаях и в соответствующих местах будут указаны его более или менее общие свойства. Напротив, мы остановимся здесь же на втором значении, удовлетворяющем во всех случаях неравенству

$$L(1) \geqslant L(0)+\frac{dL(0)}{d\lambda}. \tag{30}$$

Очевидно, что неравенство (30) будет доказано, если будет обнаружена для всякого λ справедливость неравенства

$$\frac{d^2L(\lambda)}{d\lambda^2} \geqslant 0, \tag{31}$$

ибо

$$L(1)=L(0)+\frac{dL(0)}{d\lambda}+\frac{d^2L(\lambda)}{2d\lambda^2},$$

где $0 \leqslant \lambda \leqslant 1$.

Но неравенство (31) вытекает из формул (29) и (29′), имеющих место при всяком λ $(0 \leqslant \lambda \leqslant 1)$.

В самом деле, знак $+$ в вышеупомянутых формулах берется, когда $L=F$; знак $-$ берется, когда $L=-F$. Поэтому если бы неравенство (31) было неправильно, то во внешних точках отклонения было бы $\frac{\partial^2 F}{\partial \lambda^2} F < 0$; во внутренних же точках отклонения, где

$$F>0, \text{ т. е. } \frac{\partial^2 F}{\partial x^2}<0,$$

мы имели бы

$$\frac{d^2L}{d\lambda^2}=\frac{\partial^2 F}{\partial \lambda^2}-\frac{\left(\frac{\partial^2 F}{\partial x\,\partial \lambda}\right)^2}{\frac{\partial^2 F}{\partial x^2}}<0,$$

или

$$\frac{\partial^2 F}{\partial \lambda^2}\,\frac{\partial^2 F}{\partial x^2}-\left(\frac{\partial^2 F}{\partial x\,\partial \lambda}\right)^2>0,$$

и тем более

$$\frac{\partial^2 F}{\partial \lambda^2}\,\frac{\partial^2 F}{\partial x^2}>0;$$

а во внутренних точках отклонения, где $F<0$, т. е. $\frac{\partial^2 F}{\partial x^2}>0$, таким же образом получили бы

$$\frac{\partial^2 F}{\partial \lambda^2}-\frac{\left(\frac{\partial^2 F}{\partial x\,\partial \lambda}\right)^2}{\frac{\partial^2 F}{\partial x^2}}>0$$

и поэтому также

$$\frac{\partial^2 F}{\partial \lambda^2}\,\frac{\partial^2 F}{\partial x^2}>0.$$

Следовательно, во всех точках отклонения имело бы место неравенство

$$\frac{\partial^2 F}{\partial\lambda^2} F < 0.$$

Таким образом, сумма степеней $\frac{\partial^2 F}{\partial\lambda^2} = \sum_{i=0}^{n} c_i x^{\alpha_i}$ должна была бы иметь по крайней мере по одному корню между x_i и x_{i+1}, т. е. имела бы не менее $n+1$ положительных корней, что невозможно.

Итак, неравенство (31), а вместе с ним и неравенство (30) доказаны.

Заметим, что неравенство (30) можно получить непосредственно из теоремы § 34.

В самом деле, заменяя в формуле (28) $\varphi(x_i)$ через $P(x_i) \pm L(0)$, находим

$$\pm\left[L(0) + \frac{dL(0)}{d\lambda}\right] = f(x_i) - P(x_i) - P_1(x_i).$$

Таким образом, приближенная сумма $P(x) + P_1(x)$, получающаяся, если в строке Тэйлора сохранить только первые два члена, отклоняется от $f(x)$ во всех $n+2$ точках x_i на $\pm\left(L(0) + \frac{dL(0)}{d\lambda}\right)$; следовательно, на основании указанной теоремы можно утверждать, что отклонение суммы того же вида, наименее уклоняющейся от $f(x)$, не менее чем $L(0) + \frac{dL(0)}{d\lambda}$, т. е.

$$L(1) \geqslant L(0) + \frac{dL(0)}{d\lambda}.$$

Примечание. Согласно терминологии Валле Пуссена (в упомянутой выше статье),

$$P(x) + P_1(x)$$

есть сумма степеней, наименее уклоняющаяся от $f(x)$ в данных $n+2$ точках x_i, причем, следовательно,

$$L(0) + \frac{dL(0)}{d\lambda}$$

есть наименьшее уклонение в этих точках.

Глава IV

ПРИБЛИЖЕННОЕ ВЫЧИСЛЕНИЕ НАИМЕНЬШЕГО УКЛОНЕНИЯ $|x|$ ОТ МНОГОЧЛЕНА ДАННОЙ СТЕПЕНИ

40. Задача. *Определить среди всех многочленов степени n, у которых коэффициент при x^p $(0 < p \leqslant n)$ равен* 1, *тот, который наименее уклоняется от нуля в промежутке* (0, 1).

Если искомый многочлен $P_n(x) = x^p - R(x)$, где $R(x) = \sum_{i=0}^{n-1} a_i x^{\alpha_i}$, причем $\alpha_i = i$, когда $i < p$, и $\alpha_i = i + 1$, когда $i \geqslant p$, то $R(x)$ есть сумма степеней указанного вида, наименее уклоняющаяся от x^p в промежутке

(0, 1). Следовательно, задача будет решена, если многочлен $P_n(x)$ будет иметь $n+1$ точек отклонения (§ 33) на отрезке [0, 1.] Но для этого достаточно взять многочлен

$$P_n(x) = \frac{\cos 2n \arccos \sqrt{x}}{A_{2p}},$$

где

$$\cos 2n \arccos \sqrt{x} = 2^{2n-1}\left[x^n - \frac{n}{2} x^{n-1} + \frac{n}{2^3} \frac{2n-3}{2!} x^{n-2} + \dots \right.$$

$$\left. \dots + (-1)^l \frac{n}{2^{2l-1}} \frac{(2n-l-1)\dots(2n-2l+1)}{l!} x^{n-l} + \dots \right],$$

и A_{2p} равен коэффициенту при x^p в многочлене $\cos 2n \arccos \sqrt{x}$ или коэффициенту при x^{2p} в $\cos 2n \arccos x$, а именно

$$A_{2p} = (-1)^{n-p} \frac{2^{2p} n (n+p-1)(n+p-2)\dots(2p+1)}{(n-p)!},$$

если $p < n-1$, $A_{2n-2} = -2^{2n-2} n$ и $A_{2n} = 2^{2n-1}$.
В самом деле, многочлен $P_n(x)$ имеет коэффициент при x^p равный единице и, кроме того, он имеет $n+1$ точек отклонения $x_i = \cos^2 \frac{i\pi}{2n}$, где $i = 0, 1, \dots, n$, на отрезке [0, 1].

Это отклонение, таким образом, равно $\frac{1}{|A_{2p}|}$; например, для $p=1$ оно равно $\frac{1}{2n^2}$; для $p=2$ оно равно $\frac{3}{2n^2(n^2-1)}$ и т. д.

41. Задача[1]. *Определить среди всех многочленов степени n, имеющих коэффициент при x^p равный единице, где $0 < p \leqslant n$, многочлен, наименее уклоняющийся от нуля в промежутке $(-1, +1)$.*

Пусть сначала p будет числом четным. В таком случае, если $x^p + Q(x)$ удовлетворяет задаче, то тем же свойством обладает и $x^p + Q(-x)$; и тем более многочлен $x^p + \frac{Q(x) + Q(-x)}{2}$ будет также наименее уклоняющимся от нуля в промежутке $(-1, +1)$; но этот последний многочлен будет составлен из одних только четных степеней. Поэтому, оставляя в стороне вопрос, будет ли это решение единственным (читатель легко убедится, что, хотя это и не вытекает непосредственно из общей теории, но и в данном случае решение будет только одно), можем ограничиться допущением, что $Q(x)$ составлен только из четных степеней.

Поэтому, полагая $x^2 = y$, мы можем привести нашу задачу к предыдущей. Следовательно, искомый многочлен будет

$$x^p + Q(x) = \frac{\cos n \arccos x}{B_p},$$

если n четное число, и

$$x^p + Q(x) = \frac{\cos (n-1) \arccos x}{B_p},$$

если n нечетное число, где B_p равен коэффициенту при x^p в числителе.

---

[1] Эта задача, как я узнал впоследствии, была уже решена при помощи других рассуждений в упомянутом выше сочинении В. Маркова.

Иными словами, *наилучшее приближение $x^p = x^{2k}$ при помощи многочлена степени $2n$ или $2n+1$ на отрезке $[-1, +1]$ то же, что наилучшее приближение x^k при помощи многочлена степени n на отрезке* $[0, 1]$.

Допустим далее, что p нечетное число, $p = 2k+1$. В таком случае, если многочлен $x^p + Q(x)$ дает решение задачи, то тем же свойством обладает и многочлен $x^p - Q(-x)$, а тем более многочлен $x^p + \frac{Q(x) - Q(-x)}{2}$ будет наименее уклоняющимся от нуля на отрезке $[-1, +1]$. Следовательно, можем ограничиться предположением, что искомый многочлен составлен из одних только нечетных степеней. Задача сводится таким образом к определению суммы нечетных степеней $x, x^3, \ldots, x^{2k-1}, x^{2k+3}, \ldots, x^n$ (или x^{n-1}, если n четное число), наименее уклоняющейся на отрезке $[0, 1]$ от $x^p = x^{2k+1}$; число этих степеней равно $\frac{n-1}{2}$, если n нечетное число, а если n четное число, оно равно $\frac{n-2}{2}$. Следовательно, задача будет решена, если сумма $x^p + Q(x)$ имеет $\frac{n+1}{2}$, а во втором случае $\frac{n}{2}$ точек отклонения на отрезке $[0, 1]$. Но этим свойством обладает

$$x^p + Q(x) = \frac{\cos n \arccos x}{B_p},$$

при n нечетном, и

$$x^p + Q(x) = \frac{\cos (n-1) \arccos x}{B_p},$$

при n четном, где B_p — коэффициент числителя при x^p.

Пусть, например, $p = 1$. Тогда

$$B_1 = (-1)^{\frac{n-1}{2}} n$$

(при n нечетном) и

$$B_1 = (-1)^{\frac{n-2}{2}} (n-1)$$

(при n четном).

Примечание. Таким образом, сумма $x + a_1x^3 + \ldots + a_n x^{2n+1}$ в промежутке $(0, 1)$ не может оставаться менее $\frac{1}{2n+1}$, при этом *сумма эта, действительно, не превышает* $\frac{1}{2n+1}$, *если она равна многочлену* $\frac{(-1)^n}{2n+1} \cos (2n+1) \arccos x$.

42. Преобразование задачи вычисления уклонения $|x|$. Ввиду того, что функция $|x|$ четная, мы заключаем, как в предыдущем параграфе, что многочлен, наименее уклоняющийся от $|x|$ на отрезке $[-1, +1]$, можно предположить состоящим только из четных степеней. Следовательно, этот многочлен есть не что иное, как сумма $\sum_{i=0}^{n} b_i x^{2i}$, наименее уклоняющаяся от x в промежутке $(0, 1)$; но вместо того, чтобы исследовать эту сумму, мы будем рассматривать сумму, составленную только из четных степеней $x^2, x^4, \dots, x^{2n}$ (без нулевой степени). Другими словами, мы будем изучать наименее уклоняющийся от $|x|$ из многочленов, равных нулю при $x=0$. Если мы обозначим через E'_{2n} наименьшее уклонение, соответствующее сумме последнего вида (без постоянного члена), а через E_{2n} наименьшее уклонение, соответствующее первоначальной сумме, то легко убедиться, что

$$E'_{2n} \geqslant E_{2n} \geqslant \frac{1}{2} E'_{2n}. \tag{32}$$

Первое из этих неравенств очевидно; второе вытекает из того, что если многочлен $P_{2n}(x)$ уклоняется на E_{2n} от $|x|$, то $P_{2n}(x) - P_{2n}(0)$ обращается в нуль при $x=0$ и не уклоняется от $|x|$ более чем на $2E_{2n}$ (нетрудно было бы убедиться, что знаки равенства в неравенствах (32) можно отбросить).

Примечание. Если многочлен $P(x)$ наименее уклоняется от $|x|$ на отрезке $[-1, +1]$, то $hP\left(\frac{x}{h}\right)$ есть многочлен, наименее уклоняющийся от $|x|$ на отрезке $[-h, +h]$; следовательно, наименьшее уклонение пропорционально длине промежутка $2h$.

43. Теорема. *Наименьшее уклонение $L(0)$ на отрезке $[0, 1]$ суммы вида $\sum_{i=0}^{n} a_i x^{\alpha_i}$ от x больше наименьшего уклонения $L(1)$ от x суммы вида $\sum_{i=0}^{n} b_i x^{\beta_i}$, если $\alpha_1 > \beta_1$, $\alpha_2 \geqslant \beta_2, \dots, \alpha_n \geqslant \beta_n$, причем вообще все $\beta_i > 1$.*

Положим сначала, что

$$\beta_1 < \alpha_1 \leqslant \beta_2 \leqslant \alpha_2 \leqslant \dots \leqslant \beta_n \leqslant \alpha_n.$$

Пусть сумма $Q(x) = B_1 x^{\alpha_1} + \dots + B_n x^{\alpha_n}$ будет наименее уклоняющейся от x на отрезке $[0, 1]$. В таком случае несомненно

$$B_1 > 0,\ B_2 < 0,\ B_3 > 0,\ \dots,$$

ибо уравнение $x - Q(x) = 0$ должно иметь по крайней мере n положительных корней.

Для применения теоремы § 37 строим функцию

$$F(x, \lambda) = x + (\lambda - 1) Q(x) - P(x, \lambda),$$

где $P(x, \lambda)$ есть сумма вида $x \sum_{i=1}^{n} A_i(\lambda) x^{\beta_i}$, наименее уклоняющаяся от $x + (\lambda - 1) Q(x)$, так что $P(x, 0) \equiv 0$, а $P(x, 1)$ есть $\sum_{i=0}^{n} b_i x^{\beta_i}$, наименее уклоняющаяся от x. Нетрудно убедиться, что в данном случае применение указанной теоремы законно.

В самом деле, коэффициенты суммы

$$\frac{\partial F}{\partial x} = 1 + (\lambda - 1) Q'(x) - P'_x(x, \lambda),$$

при $0 < \lambda \leqslant 1$, не могут иметь более, чем n, чередований знаков; поэтому уравнение $\frac{\partial F}{\partial x} = 0$ имеет не более n положительных простых корней, так что при всяком λ (учитывая, что число точек отклонения $> n$) конец отрезка 1 будет точкой отклонения и, кроме того, ни в одной из внутренних точек отклонения $\frac{\partial^2 F}{\partial x^2}$ не будет равно нулю.

Итак, вычисляем производную по параметру λ

$$\frac{\partial F}{\partial \lambda} = Q(x) - P'_{\lambda}(x, \lambda),$$

которая обращается в нуль не менее чем при n положительных значениях x (так как в точках отклонения $\frac{\partial F}{\partial \lambda} = \pm \frac{dL}{d\lambda}$ имеет чередующиеся знаки). Отсюда следует, что число чередований знаков коэффициентов не менее n, а потому первый коэффициент в $-P'_{\lambda}$ должен быть отрицательным. В таком случае $\frac{\partial F}{\partial \lambda}$ будет иметь ровно n положительных корней и, при x весьма малом, в частности в ближайшей к 0 точке отклонения, $\frac{\partial F}{\partial \lambda}$ будет иметь знак своего первого члена, т. е.

$$\frac{\partial F}{\partial \lambda} < 0.$$

Но в этой точке $F(x, \lambda) > 0$, следовательно, и

$$\frac{dL}{d\lambda} < 0,$$

откуда заключаем, что отклонение $L(\lambda)$ идет убывая, в то время как возрастает от 0 до 1. Таким образом,

$$L(1) < L(0).$$

Из правильности теоремы в только что рассмотренном случае легко заключить ее справедливость в самом общем случае. Для этого достаточно составить следующую таблицу показателей:

$$\begin{array}{llll}
\alpha_1, & \alpha_2, & \ldots, & \alpha_n, \\
\frac{(n-2)\alpha_1 + \beta_1}{n-1}, & \beta_2, \alpha_3, & \ldots, & \alpha_n, \\
\ldots\ldots\ldots & \ldots\ldots & \ldots & \ldots \\
\frac{\alpha_1 + (n-2)\beta_1}{n-1}, & \beta_2, & \ldots, \beta_{n-1}, & \alpha_n, \\
\beta_1, & \beta_2, & \ldots, & \beta_n.
\end{array}$$

Сравнивая каждый ряд показателей с предыдущим, мы видим, что они удовлетворяют условиям, только что рассмотренным нами. Поэтому, беря последовательно суммы степеней, соответствующие каждому ряду, получим тем более, что и в общем случае наименьшее уклонение x от суммы вида $\sum_{i=1}^{n} a_i x^{\alpha_i}$ больше наименьшего уклонения от суммы вида $\sum_{i=1}^{n} b_i x^{\beta_i}$.

44. Следствия. А. *Наименьшее уклонение на отрезке* [0, 1] *многочлена вида* $A_1x^2+A_2x^4+\ldots+A_nx^{2n}$ *от* x *меньше, чем* $\frac{1}{2n+1}$.

В самом деле, из § 41 мы знаем, что наименьшее уклонение на отрезке [0, 1] суммы нечетных степеней $a_1x^3+\ldots+a_nx^{2n+1}$ от x равно $\frac{1}{2n+1}$.

Б. *Наименьшее уклонение от* x *многочлена вида* $B_1x^4+\ldots+B_nx^{2n+2}$ *на отрезке* [0, 1] *больше, чем* $\frac{1}{2n+1}$.

45. Теорема. *Наименьшее уклонение* E'_{2n} *многочлена без свободного члена степени* $2n$ *от* $|x|$ *на отрезке* $[-1, +1]$, *при* $n>1$, *удовлетворяет неравенствам*[1]

$$\frac{1}{2(1+\sqrt{2})}\cdot\frac{1}{2n-1}<E'_{2n}<\frac{1}{2n+1}. \tag{33}$$

В самом деле, E'_{2n} есть в то же время наименьшее отклонение от x на отрезке [0, 1] многочлена вида $A_1x^2+A_2x^4+\ldots+A'_nx^{2n}$; следовательно, второе из неравенств равнозначно следствию А предыдущего параграфа. Для доказательства первого неравенства рассуждаем следующим образом.

По предположению

$$|x-A_1x^2-A_2x^4-\ldots-A_nx^{2n}|\leqslant E'_{2n} \tag{34}$$

на отрезке [0, 1]. Поэтому при всяком положительном значении μ будем тем более иметь на том же отрезке

$$\left|\frac{x}{1+\mu}-A_1\left(\frac{x}{1+\mu}\right)^2-\ldots\right|\leqslant E'_{2n},$$

откуда

$$|(1+\mu)x-A_1x^2-\ldots|\leqslant E'_{2n}(1+\mu)^2;$$

но, вычитая из этого неравенства неравенство (34), получим неравенство вида

$$|\mu(x-B_1x^4-\ldots-B_{n-1}x^{2n})|\leqslant E_{2n}[(1+\mu)^2+1]$$

и, наконец,

$$|x-B_1x^4-\ldots-B_{n-1}x^{2n}|\leqslant E'_{2n}\frac{(1+\mu)^2+1}{\mu}.$$

С другой стороны, из следствия Б предыдущего параграфа мы знаем, что $|x-B_1x^4-\ldots-B_{n-1}x^{2n}|$ должно (при $n>1$) становиться более, чем $\frac{1}{2n-1}$. Следовательно,

$$\frac{1}{2n-1}<E'_{2n}\frac{(1+\mu)^2+1}{\mu},$$

каково бы ни было положительное число μ.

---

[1] Случай $n=1$ непосредственно приводится к решению квадратного уравнения, из которого получается $E'_2=\frac{1}{2(1+\sqrt{2})}$.

Но

$$\frac{(1+\mu)^2+1}{\mu}$$

достигает минимума при $\mu = \sqrt{2}$; таким образом, в частности,

$$\frac{1}{2n-1} < E'_{2n}\cdot 2(1+\sqrt{2}),$$

откуда

$$E'_{2n} > \frac{1}{2(1+\sqrt{2})}\cdot\frac{1}{2n-1}.$$

Примечание. На основании неравенств (32) и (33) можем заключить, что

$$\frac{1}{2n+1} > E_{2n} > \frac{1}{4(1+\sqrt{2})}\cdot\frac{1}{2n-1}. \tag{33'}$$

46. Применение неравенства (30). Как мы видели в § 43, применение теоремы § 37 является вполне законным, если

$$F(x,\lambda) = x + (\lambda - 1)\,Q(x) - P(x,\lambda),$$

где $Q(x)$ — многочлен вида $B_1x^3 + B_2x^5 + \ldots + B_nx^{2n+1}$, наименее уклоняющийся от x в промежутке $(0, 1)$, а $P(x,\lambda)$ — многочлен вида $A_1x^2 + A_2x^4 + \ldots + A_nx^{2n}$, наименее уклоняющийся от $x + (\lambda - 1)\,Q(x)$ в том же промежутке. Мы знаем, что

$$x - Q(x) = \frac{(-1)^n}{2n+1}\cos(2n+1)\arccos x$$

и

$$L(0) = \frac{1}{2n+1},$$

а первоначальными точками отклонения служат

$$x_i = \cos\frac{i\pi}{2n+1} \qquad (i = 0,\ 1,\ \ldots,\ n).$$

Таким образом,

$$1 - Q(1) = (-1)^n L(0),$$

$$\cos\frac{\pi}{2n+1} - Q\left(\cos\frac{\pi}{2n+1}\right) = (-1)^{n-1}L(0),$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$\cos\frac{n\pi}{2n+1} - Q\left(\cos\frac{n\pi}{2n+1}\right) = L(0),$$

а уравнения, соответствующие уравнениям (28), имеют форму

$$Q(1) - P_1(1) = (-1)^n\frac{dL(0)}{d\lambda},$$

$$\ldots\ldots\ldots\ldots\ldots\ldots$$

$$Q\left(\cos\frac{n\pi}{2n+1}\right) - P_1\left(\cos\frac{n\pi}{2n+1}\right) = \frac{dL(0)}{d\lambda}.$$

Складывая каждое из равенств первой группы с соответствующим уравнением второй группы, получим

$$\begin{gathered}
1-P_1(1)=(-1)^n\left[L(0)+\frac{dL(0)}{d\lambda}\right],\\
\cos\frac{\pi}{2n+1}-P_1\left(\cos\frac{\pi}{2n+1}\right)=(-1)^{n-1}\left[L(0)+\frac{dL(0)}{d\lambda}\right],\\
\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\\
\cos\frac{n\pi}{2n+1}-P_1\left(\cos\frac{n\pi}{2n+1}\right)=L(0)+\frac{dL(0)}{d\lambda}.
\end{gathered}\tag{35}$$

Многочлен $P_1(x)$ имеет форму $A_1x^2+A_2x^4+\ldots+A_nx^{2n}$. Следовательно, уравнения (35) вполне определяют его коэффициенты, а также $\rho=L(0)+\frac{dL(0)}{d\lambda}$. Для удобства решения этих уравнений заметим, что к ним можно присоединить уравнения

$$\begin{gathered}
-\cos\frac{(n+1)\pi}{2n+1}-P_1\left(\cos\frac{(n+1)\pi}{2n+1}\right)=\rho,\\
\dots\dots\dots\dots\dots\dots\dots\dots\\
-\cos\frac{(2n+1)\pi}{2n+1}-P_1\left(\cos\frac{(2n+1)\pi}{2n+1}\right)=(-1)^n\rho.
\end{gathered}\tag{35'}$$

Таким образом, многочлен $P_1(x)$ есть многочлен степени не выше $2n+1$, который, благодаря равенствам (35) и (35′), должен в $2n+2$ точках $x_i=\cos\frac{i\pi}{2n+1}$ $(i=0,1,\ldots,2n+1)$ принимать значения $x_i-\rho(-1)^{n+i}$, если $i<n$, и $-x_i+\rho(-1)^{n+i}$, если $i>n$, которые станут определенными, если ρ выбрать так, чтобы было $P_1(0)=0$. Поэтому, применяя известную формулу для интерполирования, получим

$$P_1(x)=S(x)\left[\sum_{i=0}^{n}\frac{x_i-\rho(-1)^{n+i}}{(x-x_i)S'(x_i)}-\sum_{i=n+1}^{2n+1}\frac{x_i-\rho(-1)^{n+i}}{(x-x_i)S'(x_i)}\right]\tag{36}$$

где

$$S(x)=\sin(2n+1)\arccos x\cdot\sqrt{1-x^2}$$

— многочлен степени $2n+2$, имеющий корнями x_i $(i=0,1,\ldots,2n+1)$. Условие, что $P_1(0)=0$, приводит нас к уравнению

$$\sum_{i=0}^{n}\frac{x_i-\rho(-1)^{n+i}}{x_iS'(x_i)}-\sum_{i=n+1}^{2n+1}\frac{x_i-\rho(-1)^{n+i}}{x_iS'(x_i)}=0,$$

из которого определяем ρ. Для этого замечаем, что

$$S'(x)=-(2n+1)\cos(2n+1)\arccos x-\frac{x}{\sqrt{1-x^2}}\sin(2n+1)\arccos x,$$

откуда

$$S'(x_i)=-(2n+1)(-1)^i,\qquad\text{если } i=1,2,\ldots,2n,$$

и

$$S'(x_i)=-2(2n+1)(-1)^i,\qquad\text{если } i=0 \text{ или } 2n+1.$$

Таким образом,

$$\sum_{i=0}^{n}\frac{1}{S'(x_i)}=\frac{-1}{2n+1}\left[\frac{1}{2}-1+1+\ldots+(-1)^n\right]=\frac{-(-1)^n}{2(2n+1)}.$$

$$\sum_{i=n+1}^{2n+1}\frac{1}{S'(x_i)}=\frac{1}{2n+1}\left[\frac{1}{2}-1+1+\ldots+(-1)^n\right]=\frac{(-1)^n}{2(2n+1)}.$$

Следовательно,

$$\rho\left[\sum_{i=0}^{n}\frac{(-1)^i}{x_iS'(x_i)}-\sum_{i=n+1}^{2n+1}\frac{(-1)^i}{x_iS'(x_i)}\right]=\frac{-1}{2n+1},$$

или

$$\rho\left[1+\sum_{i=1}^{n}\frac{1}{\cos\frac{i\pi}{2n+1}}-\sum_{i=n+1}^{2n}\frac{1}{\cos\frac{i\pi}{2n+1}}\right]=1,$$

и, наконец,

$$\rho=\frac{1}{1+2\sum\limits_{i=1}^{n}\frac{1}{\cos\frac{i\pi}{2n+1}}}. \tag{37}$$

Пользуясь неравенством (30), мы получим отсюда нижнюю границу для $L(1)=E'_{2n}$, а именно,

$$E'_{2n}>\rho=\frac{1}{1+2\sum\limits_{i=1}^{n}\frac{1}{\cos\frac{i\pi}{2n+1}}}.$$

Но

$$1+\sum_{i=1}^{n}\frac{1}{\cos\frac{i\pi}{2n+1}}<\frac{1}{\cos\frac{n\pi}{2n+1}}+\int_0^n\frac{dz}{\cos\frac{z\pi}{2n+1}}=\frac{1}{\cos\frac{n\pi}{2n+1}}+$$

$$+\frac{2n+1}{\pi}\int_0^{\frac{n\pi}{2n+1}}\frac{du}{\cos u}=\frac{1}{\cos\frac{n\pi}{2n+1}}+\frac{2n+1}{\pi}\log\frac{1+\sin\frac{n\pi}{2n+1}}{\cos\frac{n\pi}{2n+1}}=$$

$$=\frac{1}{\frac{\pi}{4n+2}-\frac{1}{6}\left(\frac{\pi}{4n+2}\right)^3+\ldots}+\frac{2n+1}{\pi}\log\left(1+\sin\frac{n\pi}{2n+1}\right)-$$

$$-\frac{2n+1}{\pi}\log\left[\frac{\pi}{4n+2}-\frac{1}{6}\left(\frac{\pi}{4n+2}\right)^3+\ldots\right]=\frac{4n+2}{\pi}\cdot\frac{1}{1-\frac{\pi^2}{24(2n+1)^2}+\ldots}+$$

$$+\frac{2n+1}{\pi}\log\left(2-\frac{\pi^2}{8(2n+1)^2}+\ldots\right)+\frac{2n+1}{\pi}\log\frac{4n+2}{\pi}-$$

$$-\frac{2n+1}{\pi}\log\left(1-\frac{\pi^2}{24(2n+1)^2}+\ldots\right)=\frac{2n+1}{\pi}\log\frac{8n+4}{\pi}+\frac{4n+2}{\pi}+\varepsilon_n,$$

где ε_n стремится к нулю, когда n возрастает бесконечно, и, при всяком n, $\varepsilon_n < \frac{1}{2}$.

Следовательно, при всяком n,

$$E'_{2n} > \rho > \frac{\pi}{(4n+2)\left[2+\log\frac{8n+4}{\pi}\right]}. \tag{38}$$

Неравенство (38), как мы видим, дает значительно менее близкую к E'_{2n} нижнюю границу, чем неравенство (33).

47. Замена приближенного многочлена $P_1(x)$ другим многочленом. Вместо того чтобы продолжать систематическое применение общего метода, рассмотрим многочлен $R(x)$ степени $2n$, определяемый условиями, что он равен $|x|$ в точках $x_k = \cos\frac{\left(k+\frac{1}{2}\right)\pi}{2n}$ $(k = 0, 1, \ldots, 2n-1)$, где $T(x) = \cos 2n \arccos x = 0$, и, кроме того, равен нулю при $x = 0$.

Замечаем, что

$$T'(x) = \frac{2n \sin 2n \arccos x}{\sqrt{1-x^2}}.$$

Поэтому

$$T'(x_k) = (-1)^k \frac{2n}{\sin\frac{\left(k+\frac{1}{2}\right)\pi}{2n}}.$$

Следовательно,

$$R(x) = \frac{xT(x)}{2n}\left[\sum_{k=0}^{n-1}\frac{(-1)^k \sin\frac{\left(k+\frac{1}{2}\right)\pi}{2n}}{x-\cos\frac{\left(k+\frac{1}{2}\right)\pi}{2n}} - \sum_{k=n}^{2n-1}\frac{(-1)^k \sin\frac{\left(k+\frac{1}{2}\right)\pi}{2n}}{x-\cos\frac{\left(k+\frac{1}{2}\right)\pi}{2n}}\right]. \tag{39}$$

Но, с другой стороны,

$$x = \frac{xT(x)}{2n}\left[\sum_{k=0}^{n-1}\frac{(-1)^k \sin\frac{\left(k+\frac{1}{2}\right)\pi}{2n}}{x-\cos\frac{\left(k+\frac{1}{2}\right)\pi}{2n}} + \sum_{k=n}^{2n-1}\frac{(-1)^k \sin\frac{\left(k+\frac{1}{2}\right)\pi}{2n}}{x-\cos\frac{\left(k+\frac{1}{2}\right)\pi}{2n}}\right],$$

откуда

$$x - R(x) = \frac{xT(x)}{n}\sum_{k=n}^{2n-1}\frac{(-1)^k \sin\left(k+\frac{1}{2}\right)\frac{\pi}{2n}}{x-\cos\left(k+\frac{1}{2}\right)\frac{\pi}{2n}} =$$

$$= \frac{-xT(x)}{n}\sum_{k=0}^{n-1}\frac{(-1)^k \sin\left(k+\frac{1}{2}\right)\frac{\pi}{2n}}{x+\cos\left(k+\frac{1}{2}\right)\frac{\pi}{2n}}; \tag{40}$$

и так как многочлен $R(x)$ представляет собой сумму четных степеней, то $|x| - R(x)$ как при положительных, так и при отрицательных значениях x равняется разности $x - R(x)$, взятой только для *положительных* значений x.

Преобразуем сумму

$$H(x) = -\sum_{k=0}^{n-1} \frac{(-1)^k \sin\left(k+\frac{1}{2}\right)\frac{\pi}{2n}}{x+\cos\left(k+\frac{1}{2}\right)\frac{\pi}{2n}} =$$

$$= -\sum_{k=1,3,\dots,n-1} \left\{ \frac{\sin\left(k-\frac{1}{2}\right)\frac{\pi}{2n}\left[x+\cos\left(k+\frac{1}{2}\right)\frac{\pi}{2n}\right]}{\left[x+\cos\left(k-\frac{1}{2}\right)\frac{\pi}{2n}\right]\left[x+\cos\left(k+\frac{1}{2}\right)\frac{\pi}{2n}\right]} - \right.$$

$$\left. - \frac{\sin\left(k+\frac{1}{2}\right)\frac{\pi}{2n}\left[x+\cos\left(k-\frac{1}{2}\right)\frac{\pi}{2n}\right]}{\left[x+\cos\left(k-\frac{1}{2}\right)\frac{\pi}{2n}\right]\left[x+\cos\left(k+\frac{1}{2}\right)\frac{\pi}{2n}\right]} \right\} =$$

$$= \sum_{k=1,\ 3,\dots,\ n-1} \frac{2x\sin\frac{\pi}{4n}\cos\frac{k\pi}{2n}+\sin\frac{\pi}{2n}}{\left[x+\cos\left(k-\frac{1}{2}\right)\frac{\pi}{2n}\right]\left[x+\cos\left(k+\frac{1}{2}\right)\frac{\pi}{2n}\right]}, \tag{41}$$

полагая для определенности n четным.

Теперь легко убедиться, что для всякого *определенного* положительного значения x

$$\lim_{n\to\infty} xH(x) = \frac{1}{2}. \tag{42}$$

Действительно,

$$\lim_{n\to\infty} xH(x) = \lim_{n\to\infty} \sum_{k=1,\ 3,\dots,n-1} \frac{\pi}{2n} \frac{x^2\cos\frac{k\pi}{2n}+x}{\left(x+\cos\frac{k\pi}{2n}\right)^2} = \frac{x}{2}\int_0^{\frac{\pi}{2}} \frac{x\cos\alpha+1}{(x+\cos\alpha)^2}\,d\alpha = \frac{1}{2}.$$

Таким образом,

$$|x| - R(x) = \frac{\cos 2n \arccos x}{2n} + \frac{\varepsilon_n(x)\cos 2n \arccos x}{2n}, \tag{43}$$

причем $\varepsilon_n(0) = -1$ и $\lim\limits_{n\to\infty} \varepsilon_n(x) = 0$, если $|x| > 0$.

48. Определение нижней границы E'_{2n}. Многочленом $R(x)$ можно воспользоваться для определения нижней границы E'_{2n} при помощи обобщенной теоремы Валле Пуссена.

Для этого покажем сначала[1], что при всяком $x > 0$,

$$H(x) > \frac{n}{2n+1}\left[\frac{1}{x+\frac{\pi}{4n}} - \frac{1}{x+\frac{2n+1}{4n}\pi}\right]. \tag{44}$$

---

[1] Мы предполагаем $n \geqslant 2$. Случай, когда $n = 1$, не представляет никаких трудностей, как это уже было замечено ранее.

В самом деле,

$$H(x) > \sin\frac{\pi}{2n} \sum_{1,\,3,\ldots,n-1} \frac{1}{\left[x+\cos\left(k-\frac{1}{2}\right)\frac{\pi}{2n}\right]\left[x+\cos\left(k+\frac{1}{2}\right)\frac{\pi}{2n}\right]} =$$

$$= \sin\frac{\pi}{2n} \sum_{1,\,3,\ldots,\,n-1} \frac{1}{\left[x+\sin\left(k-\frac{1}{2}\right)\frac{\pi}{2n}\right]\left[x+\sin\left(k+\frac{1}{2}\right)\frac{\pi}{2n}\right]} >$$

$$> \sin\frac{\pi}{2n} \sum_{1,\,3,\ldots,\,n-1} \frac{1}{\left[x+\frac{(2k-1)\,\pi}{4n}\right]\left[x+\frac{(2k+1)\,\pi}{4n}\right]} >$$

$$> \frac{\pi}{2n+1} \sum_{1,\,3,\ldots,\,n-1} \frac{1}{\left[x+\frac{(2k-1)\,\pi}{4n}\right]\left[x+\frac{(2k+1)\,\pi}{4n}\right]} >$$

$$> \frac{\pi}{2\,(2n+1)} \sum_{1,\,2,\ldots,\,n} \frac{1}{\left[x+\frac{2k-1}{4n}\pi\right]\left[x+\frac{2k+1}{4n}\pi\right]} =$$

$$= \frac{n}{2n+1}\left[\frac{1}{x+\frac{\pi}{4n}} - \frac{1}{x+\frac{2n+1}{4n}\pi}\right].$$

Отсюда заключаем без труда, что, при $x \geqslant \frac{\pi}{8n}$,

$$xH(x) > \frac{n}{2n+1}\left(\frac{1}{3} - \frac{1}{4n+3}\right);$$

а потому, как бы мало ни было ε, можно взять n достаточно большим, чтобы иметь

$$xH(x) > \frac{1-\varepsilon}{6}.$$

Поэтому разность

$$x - R(x) = \frac{xH(x)\,T(x)}{n},$$

в точках

$$Z_i = \cos\frac{i\pi}{2n} \qquad (i = 0, 1, \ldots, n-1)$$

последовательно меняя знак, становится по абсолютному значению больше $\frac{1-\varepsilon}{6n}$ и, наконец, при новой перемене знака, в точке $\frac{\pi}{8n}$ превышает

$$\frac{1-\varepsilon}{6n}\,T\left(\frac{\pi}{8n}\right).$$

Применяя обобщенную теорему Валле Пуссена, заключаем, что

$$E'_{2n} > \frac{1-\varepsilon}{6n}\left|T\left(\frac{\pi}{8n}\right)\right| = \frac{1-\varepsilon'}{6n}\left|\cos 2n \arccos \sin\frac{\pi}{8n}\right| =$$

$$= \frac{1-\varepsilon'}{6n}\left|\cos 2n\left(\frac{\pi}{2} - \frac{\pi}{8n}\right)\right| = \frac{1-\varepsilon'}{6n}\cos\frac{\pi}{4},$$

или, полагая n достаточно большим, находим

$$E'_{2n} > \frac{1}{3\sqrt{2}} \cdot \frac{1}{2n}. \tag{45}$$

Примечание. Легко было бы проверить, что $E'_{2n} > \frac{1}{6} \cdot \frac{1}{2n}$ для всякого n; но это неравенство менее точно, чем неравенство (33), которое получено было уже выше другим способом (§ 45).

В прилагаемом ниже добавлении* к этой главе речь будет итти о приближенном вычислении E_{2n}. Что же касается E'_{2n}, то, пользуясь более точным вычислением $xH(x)$, для весьма больших значений n можно получить, пользуясь тем же многочленом $R(x)$,

$$E'_{2n} > \frac{0{,}34}{2n}.$$

Глава V

РАЗЛИЧНЫЕ ПРИЛОЖЕНИЯ ОСНОВНЫХ ТЕОРЕМ. ОБОБЩЕНИЯ ТЕОРЕМЫ ВЕЙЕРШТРАССА

49 [58]. **Теорема.** *Если производные $(n+1)$-го порядка двух функций $f(x)$ и $\varphi(x)$ удовлетворяют в промежутке AB неравенствам*

$$0 < f^{(n+1)}(x) < \varphi^{(n+1)}(x),$$

то наименьшие уклонения $E_n[f(x)]$ и $E_n[\varphi(x)]$ рассматриваемых функций от многочленов степени n на отрезке AB удовлетворяют неравенству

$$E_n[f(x)] < E_n[\varphi(x)].$$

В самом деле, составляя функцию

$$F(x, \lambda) = \lambda f(x) + (1-\lambda)\varphi(x) - P(x, \lambda),$$

где $P(x, \lambda)$ — многочлен степени n, наименее уклоняющийся от $\lambda f(x) + (1-\lambda)\varphi(x)$ на отрезке AB, мы видим, что при всяком λ расположение точек уклонения будет первого рода, и во внутренних точках уклонения $F''_{x^2} \gtrless 0$, так как на всем отрезке

$$\frac{\partial^{n+1} F(x, \lambda)}{\partial x^{n+1}} = \lambda f^{(n+1)}(x) + (1-\lambda)\varphi^{(n+1)}(x) > 0,$$

и, следовательно, уравнение $F'_x = 0$ имеет не более n корней.

Таким образом, мы вправе применять теорему § 36, замечая вместе с тем, что в последней точке уклонения $F > 0$, т. е. $F = L$. Но уравнение

$$\frac{\partial F}{\partial \lambda} = f - \varphi - P'_\lambda = 0$$

имеет $n+1$ корней, так как

$$\frac{\partial^{n+2} F}{\partial \lambda\, \partial x^{n+1}} = f^{(n+1)} - \varphi^{(n+1)} < 0;$$

---

* В настоящем издании это добавление не воспроизводится; см. [3.1]. — (*Ред.*).

при этом в последней точке отклонения

$$\frac{\partial F}{\partial \lambda} = \frac{\partial L}{\partial \lambda} < 0.$$

Следовательно,

$$L(1) = E_n[f(x)] < L(0) = E_n[\varphi(x)],$$

что и требовалось доказать.

50 [59]. Следствия. А. *Если в промежутке AB*

$$0 < \psi^{(n+1)}(x) < f^{(n+1)}(x) < \varphi^{(n+1)}(x),$$

то

$$E_n[\psi(x)] < E_n[f(x)] < E_n[\varphi(x)].$$

Б. *Если* в промежутке AB*

$$|f^{(n+1)}(x)| < \varphi^{(n+1)}(x),$$

то

$$E_n[f(x)] < 2E_n[\varphi(x)].$$

В самом деле,

$$0 < \varphi^{(n+1)}(x) \pm f^{(n+1)}(x) < 2\varphi^{(n+1)}(x);$$

поэтому

$$E_n[\varphi + f] < 2E_n[\varphi]; \qquad E_n[\varphi - f] < 2E_n[\varphi]$$

и, следовательно, тем более

$$E_n[f] = E_n\left[\frac{f + \varphi + f - \varphi}{2}\right] < 2E_n[\varphi].$$

В. *Если в промежутке AB, длина которого равна 2h,*

$$0 < N < f^{(n+1)}(x) < M,$$

то

$$\frac{2N}{(n+1)!}\left(\frac{h}{2}\right)^{n+1} < E_n[f(x)] < \frac{2M}{(n+1)!}\left(\frac{h}{2}\right)^{n+1}.$$

В самом деле, на основании следствия А,

$$E_n\left[\frac{Nx^{n+1}}{(n+1)!}\right] < E_n[f(x)] < E_n\left[\frac{Mx^{n+1}}{(n+1)!}\right],$$

а потому, замечая, что

$$E_n[Ax^{n+1}] = 2A\left(\frac{h}{2}\right)^{n+1},$$

получаем

$$\frac{2N}{(n+1)!}\left(\frac{h}{2}\right)^{n+1} < E_n[f(x)] < \frac{2M}{(n+1)!}\left(\frac{h}{2}\right)^{n+1}.$$

---

* Следствие Б было автором впоследствии уточнено («Э. П.», глава I, § 10), а именно, вместо неравенства $E_n[f(x)] < 2E_n[\varphi(x)]$ дано более сильное неравенство $E_n[f(x)] < E_n[\varphi(x)]$. Поэтому множитель 2 оказывается лишним и в следствиях Г и Д. (*Ред.*).

Г. *Если в промежутке AB длины* $2h$

$$|f^{(n+1)}(x)|<M,$$

то

$$E_n[f(x)]<\frac{4M}{(n+1)!}\left(\frac{h}{2}\right)^{n+1}.$$

Это вытекает из следствия Б.

Д. *Если в промежутке AB*

$$f^{(n+1)}(x)>k\,|f^{(n+2)}(x)|,$$

то

$$2E_n[f(x)]>kE_n[f'(x)];$$

если же

$$|f^{(n+1)}(x)|<kf^{(n+2)}(x),$$

то

$$E_n[f(x)]<2kE_n[f'(x)].$$

Это вытекает из следствия Б.

Е. *Если в промежутке AB* $(x\geqslant 0)$

$$f^{(n+1)}(x)>0,\qquad f^{(n+2)}(x)>0,$$

то

$$E_n[f(x)]<\frac{1}{n+1}E_n[xf'(x)].$$

В самом деле, полагая

$$\varphi(x)=\frac{xf'(x)}{n+1},$$

находим

$$\varphi^{(n+1)}(x)=\frac{xf^{(n+2)}(x)}{n+1}+f^{(n+1)}(x)>f^{(n+1)}(x)>0.$$

50 bis [59 bis]. Примеры. Предыдущие результаты, получаемые при помощи общего метода, если ограничиваться только первым членом соответствующей строки Тэйлора, в некоторых случаях дают довольно тесные границы для наилучшего приближения E_n.

Рассмотрим, например, наилучшее на отрезке $[a, b]$ приближение $E_n(e^x)$ функции e^x при помощи многочлена степени n. Применяя следствие В, находим немедленно

$$\frac{2e^a}{(n+1)!}\left(\frac{b-a}{4}\right)^{n+1}<E_n(e^x)<\frac{2e^b}{(n+1)!}\left(\frac{b-a}{4}\right)^{n+1}.$$

В частности, на отрезке $[-1, +1]$

$$\frac{e^{-1}}{(n+1)!}\left(\frac{1}{2}\right)^n<E_n(e^x)<\frac{e}{(n+1)!}\left(\frac{1}{2}\right)^n.$$

Рассмотрим еще наилучшее приближение функции $\sin x$ на отрезке $[-h, +h]$, где $h<\frac{\pi}{2}$, при помощи многочленов степени $2m$ или $2m-1$ (нетрудно видеть, что так как $\sin x$ есть нечетная функция, многочлены, наименее уклоняющиеся от $\sin x$ на отрезке $[-h, +h]$, будут также нечетными функциями). На основании того же следствия В, получим

$$\frac{2\cos h}{(2m+1)!}\left(\frac{h}{2}\right)^{2m+1} < E_{2m-1}(\sin x) < \frac{2}{(2m+1)!}\left(\frac{h}{2}\right)^{2m+1};$$

например, если $h = \frac{\pi}{3}$, то

$$\frac{1}{(2m+1)!}\left(\frac{\pi}{6}\right)^{2m+1} < E_{2m-1}(\sin x) < \frac{2}{(2m+1)!}\left(\frac{\pi}{6}\right)^{2m+1}.$$

Рассмотрим, наконец, наилучшее приближение $E_n\left(\frac{1}{b+x}\right)^{\alpha}$, где $b>0$ и $\alpha>0$, на отрезке $[0, 1]$.

Полагая $f(x) = \left(\frac{1}{b+x}\right)^{\alpha}$ и $\varphi(x) = \left(\frac{1}{b+x}\right)^{\alpha+h}$, находим

$$f^{(n+1)}(x) = (-1)^{n+1}\alpha(\alpha+1)\ldots(\alpha+n)\left(\frac{1}{b+x}\right)^{\alpha+n+1},$$

$$\varphi^{(n+1)}(x) = (-1)^{n+1}(\alpha+h)(\alpha+h+1)\ldots(\alpha+h+n)\left(\frac{1}{b+x}\right)^{\alpha+h+n+1},$$

откуда

$$f^{(n+1)}(x) = \frac{(\alpha+h)\ldots(\alpha+h+n)}{\alpha\ldots(\alpha+n)}(b+x)^h\varphi^{(n+1)}(x).$$

Поэтому, применяя следствие А, получим

$$\frac{(\alpha+h)\ldots(\alpha+h+n)}{\alpha\ldots(\alpha+n)}b^hE_n\left(\frac{1}{b+x}\right)^{\alpha+h} < E_n\left(\frac{1}{b+x}\right)^{\alpha},$$

$$E_n\left(\frac{1}{b+x}\right)^{\alpha} < \frac{(\alpha+h)\ldots(\alpha+h+n)}{\alpha\ldots(\alpha+n)}(b+1)^hE_n\left(\frac{1}{b+x}\right)^{\alpha+h},$$

полагая $h>0$.

В частности, если $h=1$, то

$$\frac{\alpha+n+1}{\alpha}bE_n\left(\frac{1}{b+x}\right)^{\alpha+1} < E_n\left(\frac{1}{b+x}\right)^{\alpha} < \frac{\alpha+n+1}{\alpha}(b+1)E_n\left(\frac{1}{b+x}\right)^{\alpha+1}.$$

Упражнение. Показать, при помощи следствия В, что на отрезке $[0, 1]$

$$E_n(x^{n+1+h}) < 2\frac{(n+1+h)\ldots(1+h)}{(n+1)!}\left(\frac{1}{4}\right)^{n+1},$$

если $h>0$.

51 [61]. Преобразование строк Тэйлора в ряды тригонометрических многочленов*.

Из тождества

$$(\cos t)^m = \left(\frac{e^{it}+e^{-it}}{2}\right)^m = \frac{1}{2^{m-1}}[\cos mt + m\cos(m-2)t + \ldots$$

$$\ldots + \frac{m(m-1)\ldots(m-l+1)}{l!}\cos(m-2l)t + \ldots]$$

---

* Так как в первоначальном тексте § 52 [60] предшествовал § 51 [61], то в настоящем издании в эти параграфы внесены незначительные редакционные изменения. (*Автор.*)

выводим, полагая $x = \cos t$ и $T_n(x) = \cos n \arccos x$,

$$x^m = \frac{1}{2^{m-1}} \left[T_m(x) + mT_{m-2}(x) + \frac{m(m-1)}{2} T_{m-4}(x) + \ldots \right]; \tag{46}$$

следовательно,

$$\begin{aligned} f(x) &= a_0 + a_1 x + \ldots + a_n x^n + \ldots = \\ &= a_0 + \frac{2}{2^2} a_2 + \frac{4 \cdot 3}{2^4 \cdot 2!} a_4 + \ldots + \frac{2l(2l-1)\ldots(l+1)}{2^{2l} l!} a_{2l} + \ldots \end{aligned}$$

$$\ldots + T_1(x) \left[a_1 + \frac{3}{2^2} a_3 + \frac{5 \cdot 4}{2^4 \cdot 2!} a_5 + \ldots + \frac{(2l+1)\ldots(l+2)}{2^{2l} l!} a_{2l+1} + \ldots \right] +$$

$$+ T_2(x) \left[\frac{a_2}{2} + \frac{4}{2^3} a_4 + \frac{6 \cdot 5}{2^5 \cdot 2!} a_6 + \ldots + \frac{(2l+2)\ldots(l+3)}{2^{2l+1} l!} a_{2l+2} + \ldots \right] +$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\begin{aligned} &+ T_n(x) \left[\frac{a_n}{2^{n-1}} + \frac{n+2}{2^{n+1}} a_{n+2} + \frac{(n+4)(n+3)}{2^{n+3} 2!} a_{n+4} + \ldots \right. \\ &\qquad \left. \ldots + \frac{(2l+n)\ldots(l+n+1)}{2^{2l+n-1} l!} a_{2l+n} + \ldots \right] + \end{aligned} \tag{47}$$

$$+ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

Отбрасывая первый член в правой части равенства (46), получаем многочлен степени $m-2$, наименее уклоняющийся от x^m на отрезке $[-1, +1]$.

Отбрасывая еще один член, получаем многочлен, приближающий x^m, степени $m-4$, и т. д. При этом легко оценить соответствующую погрешность, так как $|T_n(x)| \leqslant 1$. Откуда и заключаем, что

$$\begin{gathered} E_n(x^{n+3}) = E_{n-1}(x^{n+3}) < \frac{n+3}{2^{n+2}} \left[1 + \frac{1}{n+3} \right], \\ E_n(x^{n+5}) = E_{n-1}(x^{n+5}) < \frac{(n+4)(n+5)}{2^{n+4} 2!} \left[1 + \frac{2}{n+4} + \frac{2 \cdot 1}{(n+4)(n+5)} \right] \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ E_n(x^{n+2k+1}) = E_{n-1}(x^{n+2k+1}) < \\ < \frac{(n+2k+1)\ldots(n+k+2)}{2^{n+2k} k!} \left[1 + \frac{k}{n+k+2} + \frac{k(k-1)}{(n+k+2)(n+k+3)} + \ldots \right]. \end{gathered} \tag{48}$$

Полагая

$$A_{n+1} = \frac{1}{2^n} \left[a_{n+1} + \frac{n+3}{2^2} a_{n+3} + \frac{(n+4)(n+5)}{2^4 \cdot 2!} a_{n+5} + \ldots \right],$$

замечаем, что остаток, получаемый, если отбросить в разложении (47) члены степени выше n, не более, чем

$$|A_{n+1}| + |A_{n+2}| + \ldots,$$

поэтому

$$E_n[f(x)] < |A_{n+1}| + |A_{n+2}| + \ldots \tag{49}$$

52 [60]. Применение теоремы Валле Пуссена или неравенства (30).

Мы можем получить нижнюю границу $E_n[f(x)]$ на отрезке $[-1,+1]$, применяя неравенство (30), т. е. беря первые два члена строки Тэйлора, представляющей многочлен степени n, наименее уклоняющийся от $\lambda f(x)+(1-\lambda)\varphi(x)$, где $\varphi(x)=x^{n+1}$.

На основании примечания к § 39 заключаем, что эти первые два члена строки Тэйлора представляют вместе с тем многочлен $Q(x)$, наименее уклоняющийся от $f(x)$ в $n+2$ точках x_i, где разность $|\varphi(x)-P_n(x)|$ достигает максимума (через $P_n(x)$ обозначен многочлен степени n, наименее уклоняющийся от $\varphi(x)$). В данном случае $\varphi(x)=x^{n+1}$, поэтому $x_i=\cos\frac{i\pi}{n+1}$.

Пусть

$$f(x)=a_0+a_1x+\ldots+a_nx^n+\ldots$$

имеет радиус сходимости $R>1$.

Многочлен $Q(x)$ степени n удовлетворяет $n+2$ уравнениям

$$Q(x_i)=f(x_i)+(-1)^i\rho,$$

причем $|\rho|$, как мы видели, является нижней границей для $E_n[f(x)]$. Применяя формулу интерполирования Лагранжа, находим

$$Q(x)=S(x)\sum_{i=0}^{n+1}\frac{f(x_i)+(-1)^i\rho}{(x-x_i)S'(x_i)},$$

где $S(x)=\sqrt{1-x^2}\sin(n+1)\arccos x$.

Так как степень $Q(x)$ не выше n, то

$$\sum_{i=0}^{n+1}\frac{f(x_i)+(-1)^i\rho}{S'(x_i)}=0,$$

откуда, вследствие[1] равенства $\sum\limits_{i=0}^{n+1}\frac{(-1)^i}{S'(x_i)}=\pm 1$, получаем

$$\rho=\pm\sum_{i=0}^{n+1}\frac{f(x_i)}{S'(x_i)}=\pm\frac{1}{n+1}\left[\frac{1}{2}f(1)-f(x_1)+\ldots+(-1)^nf(x_n)+\frac{1}{2}(-1)^{n+1}f(-1)\right].$$

Чтобы преобразовать правую часть последнего равенства, разложим функцию $f(x)$ в ряд по тригонометрическим многочленам

$$f(x)=\sum_{k=0}^{\infty}A_kT_k(x).$$

Принимая во внимание, что

$$\frac{1}{2}-\cos k\frac{\pi}{n+1}+\cos k\frac{2\pi}{n+1}-\ldots+(-1)^n\cos k\frac{n\pi}{n+1}+$$
$$+(-1)^{n+1}\frac{1}{2}\cos k\frac{(n+1)\pi}{n+1}=\begin{cases}0 & \text{при } k\not\equiv 0 \pmod{(n+1)},\\ 0 & \text{при } k=2s(n+1),\\ n+1 & \text{при } k=(2s+1)(n+1),\end{cases}$$

---

[1] См. § 46.

получаем

$$\rho = |A_{n+1} + A_{3(n+1)} + A_{5(n+1)} + \dots|.$$

Таким образом, на отрезке $[-1, +1]$

$$E_n[f(x)] > |A_{n+1} + A_{3(n+1)} + \dots|. \tag{50}$$

В частности, на отрезке $[-1, +1]$ имеем

$$\begin{aligned} &E_n[x^{n+3}] = E_{n-1}[x^{n+3}] > \frac{n+3}{2^{n+2}}, \\ &E_n[x^{n+5}] = E_{n-1}[x^{n+5}] > \frac{(n+4)(n+5)}{2^{n+4}2!}, \\ &\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ &E_n[x^{n+2k+1}] = E_{n-1}[x^{n+2k+1}] > \frac{(n+k+2)\dots(n+2k+1)}{2^{n+2k}k!}. \end{aligned} \tag{51}$$

Сопоставление неравенств (48) с неравенствами (51) показывает, что каково бы ни было определенное целое число k, на отрезке $[-1, +1]$

$$\lim_{n\to\infty} \frac{E_n[x^{n+2k+1}]\,2^{n+2k}k!}{(n+k+2)(n+k+3)\dots(n+2k+1)} = 1. \tag{52}$$

Из неравенств (49) и (50) получаем окончательно [3.6]

$$|A_{n+1} + A_{3(n+1)} + A_{5(n+1)} + \dots| < E_n[f(x)] < |A_{n+1}| + |A_{n+2}| + \dots \tag{53}$$

53. Общие соображения о наилучшем приближении функции, заданной своим разложением в ряд Тэйлора*. Пусть

$$f(x) = a_0 + a_1x + a_2x^2 + \dots$$

функция, заданная своим разложением в ряд Тэйлора, радиус сходимости которого $R_1 > 1$. Мы увидим в третьей части, что наилучшее приближение произвольной функции $E_n[f(x)]$ больше, чем $\frac{k}{\log n} I_n[f]$, где k — постоянная, а $I_n[f(x)]$ — приближение, которое получается, если в разложении функции $f(x)$ по тригонометрическим многочленам остановиться на многочлене n-й степени[1]. В данном случае наиболее простой способ получения этого разложения состоит в применении формулы (46) к последовательным членам разложения $f(x)$, как это было сделано в § 51. Далее мы увидим также, что если B есть особая точка функции $f(x)$, лежащая на наименьшем из эллипсов с фокусами в точках -1 и $+1$, то $I_n[f(x)]$, а следовательно также и $E_n[f(x)]$, будет удовлетворять при всех достаточно больших n неравенствам вида

$$E_n[f(x)] < \frac{1}{R'^n},$$

каково бы ни было произвольное число R' такое, что $R > R'$, где R — полусумма осей конфокального эллипса, проходящего через точку B;

---

* § 53 отсутствовал в диссертации и взят из мемуара «О» (§ 46). (*Автор.*)

[1] См. теорему § 63.

кроме того, по теореме § 29, для бесконечного множества значений n имеют место неравенства $E_n\ [f(x)] > \frac{1}{R''}$, каково бы ни было $R'' > R$. Если, с другой стороны, мы вспомним, что остаток $r_n\,[f(x)]$, соответствующий ряду Тэйлора, удовлетворяет неравенствам вида

$$\frac{k}{R_1''^{n}} < r_n[f(x)] < \frac{k}{R_1'^{n}},$$

где $R_1'' > R_1 > R_1'$, то для нас станет ясно, что невозможно указать сколько-нибудь точного общего соотношения между $E_n\,[f(x)]$ и $r_n\,[f(x)]$ или между $E_n\,[f(x)]$ и R_1. Чтобы получить более точные результаты, нужно знать прежде всего положение и, кроме того, характер особой точки B. Этот случай представится, в частности, если все коэффициенты разложения $f(x)$ одного знака или попеременно его меняют (к тому же второй случай сводится к первому), так как тогда точка B будет лежать на действительной оси, и мы будем иметь $R = R_1 + \sqrt{R_1^2 - 1}$. Мы ограничимся рассмотрением случаев: 1) когда $f(x) = \frac{1}{R_1 - x}$, 2) когда $f(x)$ — целая функция.

54 [62]. Следствия. А. *Если* $\frac{|A_{n+1}| + |A_{n+2}| + \ldots}{A_n}$ *стремится к нулю при* $n \to \infty$, *то на отрезке* $[-1, +1]$

$$\lim_{n\to\infty} \frac{E_n\,[f(x)]}{A_{n+1}} = 1.$$

Б. *На отрезке* $[-h, +h]$

$$\lim_{n\to\infty} \frac{E_n\,(e^x)\cdot 2^n\,(n+1)!}{h^{n+1}} = 1.$$

В самом деле, $E_n\,(e^x)$ на отрезке $[-h, +h]$ равно $E_n\,(e^{hx})$ на отрезке $[-1, +1]$. Но

$$e^{hx} = \sum \frac{h^n x^n}{n!},$$

поэтому

$$A_{n+1} = \frac{h^{n+1}}{2^n\,(n+1)!}\left[1 + \frac{1}{n+2}\left(\frac{h}{2}\right)^2 + \frac{1}{1\cdot 2\cdot (n+2)\,(n+3)}\left(\frac{h}{2}\right)^4 + \ldots\right] =$$

$$= \frac{h^{n+1}}{2^n\,(n+1)!}\,[1 + \varepsilon_n],$$

где ε_n стремится к нулю при $n \to \infty$, следовательно, $\frac{|A_{n+2}| + |A_{n+3}| + \cdots}{A_{n+1}}$ также стремится к нулю.

В. *На отрезке* $[-h, +h]$

$$\lim_{k\to\infty} \frac{E_{2k}\,(\sin x)\cdot 2^{2k}\,(2k+1)!}{h^{2k+1}} = \lim_{k\to\infty} \frac{E_{2k-1}\,(\sin x)\cdot 2^{2k}\,(2k+1)!}{h^{2k+1}} = 1,$$

и

$$\lim_{k\to\infty}\frac{E_{2k}(\cos x)\cdot 2^{2k+1}(2k+2)!}{h^{2k+2}}=\lim_{k\to\infty}\frac{E_{2k+1}(\cos x)\cdot 2^{2k+1}(2k+2)!}{h^{2k+2}}=1.$$

Доказательство подобно предыдущему[1].

Г. *Если* [3.7]

$$\lim_{n\to\infty}\frac{a_n}{R^n}=1,$$

то на отрезке $[-1, +1]$ *при* n *достаточно большом*

$$\frac{2RB^n}{\sqrt{1-R^2}\left(1+\sqrt{1-R^2}\right)\left(1-B^{2n+2}\right)}<E_n[f(x)]<\frac{2RB^n}{\sqrt{1-R^2}\left(1-R+\sqrt{1-R^2}\right)},$$

где

$$B=\frac{R}{1+\sqrt{1-R^2}}.$$

Для простоты письма положим $a_n=R^n$ (что соответствует $f(x)=\frac{1}{1-Rx}$). В таком случае

$$A_{n+1}=\frac{R^{n+1}}{2^n}\left[1+(n+3)\left(\frac{R}{2}\right)^2+\frac{(n+4)(n+5)}{2!}\left(\frac{R}{2}\right)^4+\right.$$

$$\left.+\frac{(n+5)(n+6)(n+7)}{3!}\left(\frac{R}{2}\right)^6+\ldots\right]=\frac{R^{n+1}}{2^n}F\left(\frac{n}{2}+1,\ \frac{n}{2}+\frac{3}{2},\ n+2,\ R^2\right);$$

замечая, что

$$F\left(\frac{n}{2}+1,\ \frac{n}{2}+\frac{3}{2},\ n+2,\ R^2\right)=\frac{2^{n+1}}{\sqrt{1-R^2}\left(1+\sqrt{1-R^2}\right)^{n+1}},$$

получаем разложение в ряд по тригонометрическим многочленам для $f(x)=\frac{1}{1-Rx}$ в таком виде:

$$\frac{1}{1-Rx}=\sum_{n=0}^{\infty}\frac{2R^n}{\sqrt{1-R^2}\left(1+\sqrt{1-R^2}\right)^n}T_n(x).$$

---

[1] Назовем функцию $Q_n(x)$ асимптотическим выражением многочлена $P_n(x)$ степени n, наименее уклоняющегося от данной функции $f(x)$, если уклонение $E'_n f$ функции $f(x)$ от $Q_n(x)$ удовлетворяет условию, что $\lim\limits_{n\to\infty}\frac{E'_n f}{E_n f}=1$. Таким образом, асимптотическими выражениями многочленов, наименее уклоняющихся от рассматриваемых здесь функций, во всех случаях служат многочлены $Q_n(x)$ степени n, получаемые при помощи преобразования, указанного в § 51, т. е. $Q_n(x)$ является n-й частной суммой ряда, полученного при помощи разложения $f(x)$ в ряд по тригонометрическим многочленам.

Следовательно,

$$\sum_{k=n+1}^{\infty} A_k = \frac{2RB^n}{\sqrt{1-R^2}\,(1-R+\sqrt{1-R^2})},$$

$$\sum_{k=0}^{\infty} A_{(2k+1)(n+1)} = \frac{2RB^n}{\sqrt{1-R^2}\,(1+\sqrt{1-R^2})\,(1-B^{2n+2})}.$$

Отсюда, применяя неравенство (53), получаем

$$\frac{2RB^n}{\sqrt{1-R^2}\,(1+\sqrt{1-R^2})\,(1-B^{2n+2})} < E_n[f(x)] < \frac{2RB^n}{\sqrt{1-R^2}\,(1-R+\sqrt{1-R^2})}.$$

Интересно сравнить полученный результат с теоремой § 29.

Не останавливаясь на этом, перейдем к рассмотрению неаналитических функций.

55 [63]. Теорема Вейерштрасса. Выведем некоторые следствия из неравенства *

$$E_{2n}[|x|] < \frac{0,286}{2n}, \qquad (54)$$

имеющего место на отрезке $[-1, +1]$ для достаточно больших значений n.

Хорошо известно, что из того, что

$$\lim_{n\to\infty} E_n[|x|] = 0,$$

вытекает теорема Вейерштрасса, т. е.

$$\lim_{n\to\infty} E_n[f(x)] = 0$$

для какой угодно непрерывной функции[1]. Я хочу заметить только, что при помощи формул, указанных мной в 1905 г.[2], из неравенства (54) можно вывести в некоторых случаях довольно точную верхнюю границу для $E_{2n}[f(x)]$.

Пусть $f(x)$ будет непрерывная на отрезке $[0, 1]$ функция, и пусть $y = f_n(x)$ будет уравнением ломаной линии, имеющей вершинами точки на линии $y = f(x)$, с абсциссами $x_k = \frac{k}{n}$ $(x = 0, 1, \ldots, n)$.

Упомянутые мною формулы заключаются в том, что

$$f_n(x) = \sum_{k=1}^{n-1} A_k |x - x_k| + A + Bx,$$

---

* См. [12], § 26, неравенство (36). Вместо неравенства $E_{2n}[|x|] < \frac{0,286}{2n}$ в первоначальном тексте стояло неравенство $E_{2n}[|x|] < \frac{0,32}{2n}$, которое было доказано в опущенном здесь добавлении к главе IV. (*Автор.*)

[1] Небесполезно обратить внимание на то, что непрерывность функции $f(x)$ есть условие, необходимое и достаточное для того, чтобы $\lim E_n[f(x)] = 0$.

[2] См. статью «Об интерполировании» [1].

где

$$A_k = \frac{n}{2}\left[f\left(\frac{k+1}{n}\right) + f\left(\frac{k-1}{n}\right) - 2f\left(\frac{k}{n}\right)\right],$$

$$A = \frac{1}{2}\left[f(0) + nf\left(\frac{n-1}{n}\right) - (n-1)f(1)\right],$$

$$B = \frac{n}{2}\left[f\left(\frac{1}{n}\right) + f(1) - f(0) - f\left(\frac{n-1}{n}\right)\right].$$

Заменяя $|x - x_k|$ приближенными многочленами степени p, получаем приближенный многочлен $f_{n,p}(x)$ степени p для $f_n(x)$ и заключаем, что, при p достаточно большом, ошибка $|f_{n,p}(x) - f_n(x)|$ и, тем более, $E_p[f_n(x)]$ будет удовлетворять неравенству

$$E_p[f_n(x)] < \frac{0,286}{p} \sum_{k=1}^{n-1} |A_k|. \tag{54'}$$

Ограничимся только рассмотрением случая, когда функция $f(x)$ удовлетворяет условию Дини-Липшица, а именно пусть

$$|f(x+h) - f(x)| < \frac{\delta(h)}{|\log h|},$$

где $\delta(h)$ стремится к нулю вместе с h.

В таком случае, очевидно,

$$|f(x) - f_n(x)| < \frac{2\delta\left(\frac{1}{n}\right)}{\log n}$$

и, с другой стороны,

$$\frac{1}{p}\sum_{k=1}^{n-1} |A_k| < \frac{n^2}{p} \cdot \frac{\delta\left(\frac{1}{n}\right)}{\log n},$$

так как $|A_k| < \frac{n\delta\left(\frac{1}{n}\right)}{\log n}$. Поэтому, полагая $p = n^2$, находим

$$E_p[f(x)] < |f(x) - f_{n,p}(x)| < 4,572\frac{\delta\left(\frac{1}{\sqrt{p}}\right)}{\log p}. \tag{55}$$

Аналогичное неравенство дал Лебег в цитированной выше работе из «Annales de Toulouse». Заметим, что в случае существования обобщенного условия Дини-Липшица неравенство (55) соблюдается не для всех, но для бесчисленного множества значений p. Следовательно, принимая во внимание результат § 27, находим, что *условие необходимое и достаточное, чтобы функция $f(x)$ удовлетворяла обыкновенному условию Дини-Липшица, заключается в том, чтобы* $\lim_{n\to\infty} E_n[f(x)]\log n = 0$; *условие необходимое и достаточное, чтобы функция $f(x)$ удовлетворяла обобщенному условию Дини-Липшица, заключается в том, чтобы, при бесчисленном множестве значений* $n > n_0$, $\lim_{n\to\infty} E_n[f(x)]\log n = 0$.

56 [64]. **Первое обобщение теоремы Вейерштрасса.** *Если дан бесконечный ряд чисел*

$$\alpha_1 < \alpha_2 < \dots < \alpha_n < \dots,$$

обладающий тем свойством, что $H < \alpha_i < K$, где H и K — два независимых от i положительных числа, то для всякой непрерывной на отрезке $[0, 1]$ *функции $f(x)$ можно составить сумму $\sum\limits_{i=1}^{n} A_i x^{\alpha_i}$ так, чтобы на всем отрезке было*

$$\left| f(x) - f(0) - \sum_{i=1}^{n} A_i x^{\alpha_i} \right| < \varepsilon,$$

как бы мало ни было число ε.

(Указанным свойством обладают, например, числа $\alpha_i = 1 - \frac{1}{2^i}$).

Наша теорема будет, очевидно, доказана, если мы покажем, что она справедлива для $f(x) = x^p$, где p — произвольное целое число, большее единицы.

Для этого замечаем сначала, что, на основании рассуждения, совершенно подобного доказательству теоремы § 43, можно утверждать, что *наилучшее приближение x^p на отрезке* $[0, 1]$ *при помощи суммы вида $\sum\limits_{i=1}^{n} A_i x^{\alpha_i}$ всегда меньше наилучшего приближения при помощи суммы вида $\sum\limits_{i=1}^{n} B_i x^{\beta_i}$, если $p > \alpha_i > \beta_i > 0$.*

С другой стороны, полагая в неравенствах (48) $x^2 = y$, выводим из них, что на отрезке $[0, 1]$

$$E_{m-1}(x^{m+k}) < \frac{(2m+2k)\dots(2m+k+1)}{2^{2m+2k-1}k!}\left[1 + \frac{k}{2m+k+1} + \right.$$
$$\left. + \frac{k(k-1)}{(2m+k+1)(2m+k+2)} + \dots\right];$$

обозначая через $E'_n(x^p)$ наилучшее приближение x^p на отрезке $[0, 1]$ при помощи суммы $\sum\limits_{i=1}^{n} A_i x^i$, имеем

$$E'_n(x^p) < \frac{2p\dots(p+n+2)}{2^{2p-2}(p-n-1)!}\left[1 + \frac{p-n-1}{p+n+2} + \frac{(p-n-1)(p-n-2)}{(p+n+2)(p+n+3)} + \dots\right] =$$
$$= I_{n+1} + I_{n+2} + \dots + I_p,$$

где

$$I_s = \frac{2p\dots(p+s+1)}{2^{2p-2}(p-s)!} = I_0 \frac{p(p-1)\dots(p-s+1)}{(p+1)\dots(p+s)}.$$

Поэтому

$$\begin{aligned}\log I_s &= \log I_0 + \log \frac{p(p-1)\ldots(p-s+1)}{(p+1)\ldots(p+s)} = \\ &= \log I_0 + \left[\log\left(1-\frac{1}{p}\right) - \log\left(1+\frac{1}{p}\right)\right] + \ldots \\ &\ldots + \left[\log\left(1-\frac{s-1}{p}\right) - \log\left(1+\frac{s-1}{p}\right)\right] - \log\left(1+\frac{s}{p}\right) < \\ &< \log I_0 - \frac{2}{p} - \frac{4}{p} - \ldots - \frac{2(s-1)}{p} - \frac{s}{p} + \frac{s^2}{2p^2} = \\ &= \log I_0 - \frac{s^2}{p} + \frac{s^2}{2p^2}.\end{aligned}$$

Отсюда

$$I_s < I_0 e^{-\frac{s^2}{p}+\frac{s^2}{2p^2}} < \frac{4}{\sqrt{p\pi}}\, e^{-\frac{s^2}{p}+\frac{s^2}{2p^2}},$$

так как

$$I_0 = \frac{(2p)!}{2^{2p-2}(p!)^2} = \frac{1\cdot 3\ldots(2p-1)}{2\cdot 4\ldots 2p}\cdot 4 < \frac{4}{\sqrt{p\pi}}.$$

Но, при $p > 1$, $s > 0$,

$$e^{-\frac{s^2}{p}+\frac{s^2}{2p^2}} < \frac{1}{2}\left[e^{-\frac{\left(s-\frac{1}{2}\right)^2}{p}} + e^{-\frac{s^2}{p}}\right]. \tag{56}$$

Действительно, это неравенство равнозначно неравенству

$$e^{\frac{s^2}{2p^2}} < \frac{1}{2}\left[e^{\frac{s}{p}-\frac{1}{4p}} + 1\right],$$

или, полагая $u = \frac{s}{p}$, $\alpha = \frac{1}{4p}$, — неравенству

$$f(u) = 2e^{\frac{u^2}{2}-u} - e^{-u} < e^{-\alpha},$$

справедливость которого нужно, следовательно, доказать при предположении, что $\alpha \leqslant \frac{1}{8}$, $1 \geqslant u \geqslant 4\alpha$. Но нетрудно видеть, что при рассматриваемых значениях u, $f''(u) > 0$; поэтому наибольшее значение $f(u)$ будет равно $f(1)$ или $f(4\alpha)$, так что достаточно заметить, что, при $\alpha \leqslant \frac{1}{8}$,

$$f(1) = 2e^{-1/2} - e^{-1} < e^{-\alpha}; \qquad f(4\alpha) = 2e^{8\alpha^2-4\alpha} - e^{-4\alpha} < e^{-\alpha}.$$

Из неравенства (56) заключаем, что

$$e^{-\frac{s^2}{p}+\frac{s^2}{2p^2}} < \int\limits_{s-1}^{s} e^{-\frac{z^2}{p}}\,dz,$$

и потому

$$I_s < \frac{4}{\sqrt{p\pi}} \int\limits_{s-1}^{s} e^{-z^2/p}\,dz = \frac{4}{\sqrt{\pi}} \int\limits_{\frac{s-1}{\sqrt{p}}}^{\frac{s}{\sqrt{p}}} e^{-z^2}\,dz.$$

Следовательно [1], наконец,

$$E'_n(x^p) < \frac{4}{\sqrt{\pi}} \int\limits_{\frac{n}{\sqrt{p}}}^{\infty} e^{-z^2}dz. \tag{57}$$

Таким образом, $E'_n(x^p)$ *стремится к нулю, если* $\frac{n}{\sqrt{p}}$ *возрастает бесконечно*. Поэтому, в частности, $E'_n(x^{pn})$ стремится к нулю, если, при данном p, n возрастает бесконечно. Но, полагая $x^n = y$, мы видим, что $E'_n(x^{pn})$ есть вместе с тем наилучшее приближение функции x^p при помощи суммы $\sum\limits_{i=1}^{n} B_i x^{in}$ на том же отрезке $[0, 1]$. Следовательно, благодаря замечанию, сделанному в начале доказательства, приближение x^p при помощи суммы вида $\sum\limits_{i=1}^{n} A_i x^{\alpha_i}$ стремится к нулю вместе с $1/n$, так как (введя, если понадобится, переменную x^k вместо x) всегда можно предположить, что $1 \leqslant H < \alpha_i$, что и требовалось доказать.

Примечание. Отрезок $[0, 1]$ может быть заменен произвольным отрезком AB на положительной оси; и, кроме того, нетрудно убедиться, что если отрезок AB не доходит до 0, то условие, чтобы $H > 0$, $K > 0$, может быть отброшено.

57 [65]. Второе обобщение теоремы Вейерштрасса. *Если показатели α_n возрастают бесконечно вместе с n, то наилучшее приближение непрерывной функции $f(x)$ на отрезке $[0, 1]$ при помощи $\sum A_i x^{\alpha_i}$ стремится к нулю, если $\frac{\alpha_n}{n \log n}$ стремится к нулю; напротив, наилучшее приближение не может стремиться к нулю, если есть такое число ε, что $\alpha_n \geqslant n(\log n)^{2+\varepsilon}$ или $\alpha_n \geqslant n(\log n)^2(\log\log n)^{1+\varepsilon}$ и т. д.*

Займемся сначала доказательством первой части теоремы.

Достаточно будет рассмотреть случай, когда $f(x) = x^p$, где p — произвольное целое положительное число, если брать только те α_i, которые больше p, и, тем более, достаточно будет доказать, что, как бы мало ни было число δ, возможно на всем отрезке $[0, 1]$ удовлетворить неравенству

$$\left| x - \sum_{i=i_0+1}^{i_0+n} A_i x^{\alpha_i - p + 1} \right| < \delta, \tag{58}$$

---

[1] Указанное здесь вычисление аналогично тому, которое я сделал в заметке «Sur le calcul approché des probabilités par la formule de Laplace» («Сообщ. Х. М. О», т. XII, № 3), и приводит к следующему результату для теории вероятностей: если вероятность события равна $\frac{1}{2}$, то, при $2p$ $(p > 1)$ *испытаниях, вероятность, что число m появлений события удовлетворяет неравенству $|m - p| \leqslant z_0\sqrt{p}$, более, чем* $\Phi(z_0) = \frac{2}{\sqrt{\pi}} \int\limits_0^{z_0} e^{-z^2}dz$.

ибо, если это неравенство имеет место, то, конечно,

$$\left|x^p - \sum_{i=i_0+1}^{i_0+n} A_i x^{\alpha_i}\right| < \delta x^{p-1} \leqslant \delta.$$

Пусть

$$\alpha_n = \varepsilon_n n \log(n+1);$$

в таком случае, по предположению, как бы мало ни было число γ, можно указать достаточно большое число n_0, чтобы при $n \geqslant n_0$, иметь $\varepsilon_n < \gamma$.

На основании теоремы § 43 неравенство (58) может быть осуществлено, если известно, что

$$\left|x - \sum_{h=1}^{n} B_h x^{\beta_h}\right| < \delta,$$

где

$$\beta_h > \alpha_{i_0+h} - p + 1.$$

Положим $\beta_h = kh$; тогда

$$\left|x - \sum_{h=1}^{n} B_h x^{kh}\right| = \left|y^{1/k} - \sum_{h=1}^{n} B_h y^h\right|.$$

Мы увидим в следующей главе (и это вытекает также из примечания в) к теореме § 16), что эта разность может быть сделана менее $\frac{b}{n^{1/k}}$, где b — не зависящая от n и k постоянная. Таким образом,

$$\delta < \frac{b}{n^{1/k}},$$

если

$$k > \frac{\alpha_{i_0+h} - p + 1}{h} = \frac{\varepsilon_{i_0+h}(i_0+h)\log(i_0+h) - p + 1}{h}. \tag{59}$$

Для значений h, которые меньше, чем i_0, и меньше, чем $n_0 - i_0$, неравенству (59) можно удовлетворить, взяв для k некоторое вполне определенное число k_0; для остальных же значений h неравенство будет соблюдено, если взять

$$k = 2\gamma \log 2n.$$

Можно предположить n настолько большим, что $2\gamma \log 2n > k_0$. Следовательно,

$$\delta < \frac{b}{n^{\frac{1}{2\gamma \log 2n}}} = \frac{b}{e^{\frac{\log n}{2\gamma \log 2n}}} < be^{-\frac{1}{4\gamma}};$$

поэтому δ может быть сделано сколь угодно малой, и первая часть теоремы доказана.

Для доказательства второй части теоремы замечаем, что наилучшее приближение x на отрезке $[0, 1]$ при помощи суммы[1] $\sum_{i=1}^{n} A_i x^{\alpha_i}$ (где $\alpha_i > 1$),

---

[1] Если бы одно из чисел α_i было равно 1, то вместо наилучшего приближения x можно было бы рассматривать наилучшее приближение x^p, где $p \gtrless \alpha_i$.

β_n удовлетворяет, при всяком положительном значении μ, неравенству

$$\beta_n > \beta_{n-1} \frac{(1+\mu)^{\alpha_n-1}-1}{(1+\mu)^{\alpha_n}+1}. \tag{60}$$

Действительно, из

$$\left| x + A_1 x^{\alpha_1} + \ldots + A_n x^{\alpha_n} \right| < \beta_n$$

заключаем, что и

$$\left| \frac{x}{1+\mu} + A_1 \left(\frac{x}{1+\mu} \right)^{\alpha_1} + \ldots + A_n \left(\frac{x}{1+\mu} \right)^{\alpha_n} \right| < \beta_n;$$

а потому

$$\left| x(1+\mu)^{\alpha_n-1} + \ldots + A_n x^{\alpha_n} \right| < \beta_n (1+\mu)^{\alpha_n},$$

откуда

$$\left| x \left[(1+\mu)^{\alpha_n-1} - 1 \right] + \ldots + A'_{n-1} x^{\alpha_{n-1}} \right| < \beta_n \left[(1+\mu)^{\alpha_n} + 1 \right]$$

и

$$\left| x + B_1 x^{\alpha_1} + \ldots + B_{n-1} x^{\alpha_{n-1}} \right| < \beta_n \frac{(1+\mu)^{\alpha_n}+1}{(1+\mu)^{\alpha_n-1}-1},$$

следовательно,

$$\beta_{n-1} < \beta_n \frac{(1+\mu)^{\alpha_n}+1}{(1+\mu)^{\alpha_n-1}-1},$$

или

$$\beta_n > \beta_{n-1} \frac{(1+\mu)^{\alpha_n-1}-1}{(1+\mu)^{\alpha_n}+1}.$$

Из неравенства (60) получаем немедленно

$$\beta_n > \beta_{n_0} \cdot \prod_{i=n_0+1}^{n} \frac{(1+\delta_i)^{\alpha_i-1}-1}{(1+\delta_i)^{\alpha_i}+1}, \tag{61}$$

где δ_i — какие угодно положительные числа. Достаточно теперь будет показать, что при соответствующем выборе чисел δ_i произведение, стоящее во второй части неравенства, *не* стремится к нулю, при $n \to \infty$, если $\alpha_n \geqslant n (\log n)^{2+\varepsilon}$ или $\alpha_n \geqslant n (\log n)^2 (\log \log n)^{1+\varepsilon}$ и т. д.

Но

$$\frac{(1+\delta_i)^{\alpha_i-1}-1}{(1+\delta_i)^{\alpha_i}+1} = \frac{1}{1+\delta_i} \frac{1 - \dfrac{1}{(1+\delta_i)^{\alpha_i-1}}}{1 + \dfrac{1}{(1+\delta_i)^{\alpha_i}}}.$$

Поэтому рассматриваемое произведение не может стремиться к нулю, если оба ряда

$$\sum \delta_i, \quad \sum \frac{1}{(1+\delta_i)^{\alpha_i}}$$

будут сходящимися. Для сходимости первого ряда достаточно взять

$$\delta_n = \frac{2}{n(\log n)^{1+\varepsilon}}, \quad \text{или} \quad \delta_n = \frac{2}{n \log n (\log \log n)^{1+\varepsilon}} \quad \text{и т. д.};$$

возьмем, например, первое из этих значений. В таком случае и ряд

$$\sum \frac{1}{(1+\delta_i)^{\alpha_i}}$$

будет сходящимся, если $\alpha_i \geqslant i (\log i)^{2+\varepsilon}$.

В самом деле общий член этого ряда меньше, чем

$$\frac{1}{\left[1 + \frac{2}{i(\log i)^{1+\varepsilon}}\right]^{i(\log i)^{2+\varepsilon}}},$$

т. е., при i достаточно большом, меньше, чем

$$\frac{1}{e^{2\log i}} = \frac{1}{i^2};$$

а потому ряд $\sum \frac{1}{(1+\delta_i)^{\alpha_i}}$ сходящийся, и, следовательно, вторая часть теоремы доказана.

Примечание. Отрезок $[0, 1]$ может быть заменен произвольным отрезком AB положительной оси.

Добавление к главе V

РАЗЛОЖЕНИЕ ПРОИЗВОЛЬНЫХ ФУНКЦИЙ В НОРМАЛЬНЫЕ РЯДЫ

58 [66]. Нормальные ряды. Нормальным рядом на отрезке $[0,1]$ называется ряд вида

$$\sum_{p=0}^{\infty} \sum_{q=0}^{\infty} A_{p,q} x^p (1-x)^q,$$

абсолютно и равномерно сходящийся на этом отрезке. В моем сочинении «Исследование и интегрирование дифференциальных уравнений с частными производными 2-го порядка эллиптического типа»[1] дано (в главе II) разложение в нормальный ряд, пригодное для всякой функции, имеющей непрерывную производную на отрезке $[0, 1]$. Естественно задать себе вопрос, может ли совершенно произвольная непрерывная функция быть разложена в нормальный ряд.

Ответ на этот вопрос, как мы увидим далее, оказывается утвердительным. А именно, мы укажем прием для преобразования произвольного, равномерно сходящегося ряда многочленов в нормальный ряд. С этой целью разрешим предварительно следующую алгебраическую задачу.

---

[1] «Сообщ. Харьк. матем. об-ва», серия 2, т. 11 (1908), стр. 1—96.

Задача. *Преобразовать многочлен*

$$P(x) = a_n x^n + a_{n-1} x^{n-1} + \ldots + a_1 x + a_0$$

в выражение

$$P(x) = \sum_{p=0}^{p+q=m} \sum_{q=0} A_{p,q} x^p (1-x)^q,$$

где $m \geqslant n$, *так, чтобы максимум суммы*

$$\sum_{p=0}^{p+q=m} \sum_{q=0} |A_{p,q}| x^p (1-x)^q$$

на отрезке $[0, 1]$ *был возможно мал.*

Ввиду того, что коэффициенты $A_{p,q}$ и число их ограничены, задача, очевидно, имеет решение, т. е. можно выбрать эти коэффициенты так, чтобы максимум суммы

$$\sum_{p=0}^{p+q=m} \sum_{q=0} |A_{p,q}| x^p (1-x)^q$$

достигал своего низшего предела; этому минимальному значению максимума мы для краткости дадим название *нормального максимума степени* m данного многочлена на отрезке $[0, 1]$.

Весьма замечательно, что поставленная задача разрешается совершенно элементарно, причем обнаруживается интересный факт, что *нормальный максимум степени* m *любого многочлена* $P(x)$ *имеет пределом, при* $m \to \infty$, *максимум* $|P(x)|$ *на данном отрезке.* Искомое решение вытекает из простого замечания: допустим, что задача решена, и пусть выражение

$$P(x) = \sum_{p=0}^{p+q=m} \sum_{q=0} a_{p,q} x^p (1-x)^q$$

есть одно из возможных решений. Я говорю, что если среди членов $a_{p,q} x^p (1-x)^q$ есть такие, степень которых $p + q = m - k$, где $k > 0$, то решением задачи будет служить и то выражение, которое получится от замены $a_{p,q} x^p (1-x)^q$ суммой членов степени m,

$$a_{p,q} x^p (1-x)^q [x + (1-x)]^k =$$
$$= a_{p,q} [x^{p+k} (1-x)^q + k x^{p+k-1} (1-x)^{q+1} + \ldots + x^p (1-x)^{q+k}].$$

В самом деле,

$$|a_{p,q}| x^p (1-x)^q =$$
$$= |a_{p,q}| x^{q+k} (1-x)^q + |k a_{p,q}| x^{p+k-1} (1-x)^{q+1} + \ldots + |a_{p,q}| x^p (1-x)^{q+k},$$

поэтому сумма модулей преобразованного выражения не может превысить суммы модулей данного выражения.

Отсюда следует, что среди решений задачи всегда есть одно решение, в котором сумма показателей $p + q = m$. Другими словами, задача будет решена, если представим $P(x)$ в виде

$$P(x) = A_m x^m + A_{m-1} x^{m-1} (1-x) + \ldots + A_0 (1-x)^m.$$

Остается вычислить коэффициенты A_i так, чтобы иметь тождественно

$$A_m x^m + A_{m-1} x^{m-1} (1-x) + \ldots + A_0 (1-x)^m = a_n x^n + a_{n-1} x^{n-1} + \ldots + a_0,$$

откуда находим для определения $m+1$ коэффициентов $m+1$ уравнений

$$\begin{aligned} &A_0 = a_0, \\ &A_1 - mA_0 = a_1, \\ &\ldots\ldots\ldots \\ &A_k - (m-k+1) A_{k-1} + \ldots + (-1)^k \frac{m(m-1)\ldots(m-k+1)}{1\cdot 2 \ldots k} A_0 = a_k, \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ &A_m - A_{m-1} + \ldots + (-1)^m A_0 = a_m, \end{aligned} \tag{62}$$

где $a_k = 0$, если $k > n$.

Решение уравнений (62) не представляет труда и дает немедленно

$$\begin{aligned} &A_0 = a_0, \\ &A_1 = a_1 + ma_0, \\ &A_2 = a_2 + (m-1) a_1 + \frac{m(m-1)}{2} a_0, \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ &A_k = a_k + C^1_{m-k+1} a_{k-1} + \ldots + C^k_m a_0, \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ &A_m = a_m + a_{m-1} + \ldots + a_0, \end{aligned} \tag{63}$$

где

$$C^k_m = \frac{m(m-1)\ldots(m-k+1)}{1\cdot 2 \ldots k}.$$

Итак, поставленная задача решена; *нормальный максимум степени m данного многочлена равен максимуму суммы*

$$\sum_{k=0}^{m} |A_k| x^k (1-x)^{m-k}$$

где коэффициенты A_k определяются формулами (63).

59 [67]. Исследование величины нормального максимума. Формулу, определяющую A_k, можно преобразовать следующим образом:

$$A_k = C^k_m \left[a_0 + \frac{C^{k-1}_{m-1}}{C^k_m} a_1 + \frac{C^{k-2}_{m-2}}{C^k_m} a_2 + \ldots \right] = C^k_m \left[a_0 + \frac{k}{m} a_1 + \frac{k(k-1)}{m(m-1)} a_2 + \ldots \right] =$$

$$= C^k_m \left[a_0 + \frac{k}{m} a_1 + \left(\frac{k}{m}\right)^2 a_2 \frac{1-\frac{1}{k}}{1-\frac{1}{m}} + \ldots + \left(\frac{k}{m}\right)^n a_n \frac{\left(1-\frac{1}{k}\right)\left(1-\frac{2}{k}\right)\ldots\left(1-\frac{n-1}{k}\right)}{\left(1-\frac{1}{m}\right)\left(1-\frac{2}{m}\right)\ldots\left(1-\frac{n-1}{m}\right)} \right].$$

Из полученной формулы видно, что, при бесконечном возрастании m,

$$\lim_{m\to\infty} \frac{A_k}{C^k_m} = \lim_{m\to\infty} P\left(\frac{k}{m}\right). \tag{64}$$

Действительно, если k есть определенное число, то все члены суммы, состоящей из данного числа $n+1$ слагаемых,

$$\frac{A_k}{C_m^k} = a_0 + \frac{k}{m}\, a_1 + \left(\frac{k}{m}\right)^2 a_2 \frac{1-\frac{1}{k}}{1-\frac{1}{m}} + \cdots,$$

кроме a_0, стремится к нулю, поэтому

$$\lim_{m\to\infty} \frac{A_k}{C_m^k} = a_0 = P(0) = \lim_{m\to\infty} P\left(\frac{k}{m}\right).$$

Если же k также возрастает бесконечно, то

$$\lim_{m\to\infty} \frac{A_k}{C_m^k} = \lim_{m\to\infty}\left[a_0 + a_1\frac{k}{m} + a_2\left(\frac{k}{m}\right)^2 + \ldots + a_n\left(\frac{k}{m}\right)^n\right] = \lim_{m\to\infty} P\left(\frac{k}{m}\right).$$

Следует прибавить, что разность

$$\delta_k = \frac{A_k}{C_m^k} - P\left(\frac{k}{m}\right)$$

равномерно стремится к нулю при бесконечном возрастании m.

В самом деле,

$$\delta_k = \left(\frac{k}{m}\right)^2 a_2\left[\frac{1-\frac{1}{k}}{1-\frac{1}{m}} - 1\right] + \ldots + \left(\frac{k}{m}\right)^n a_n\left[\frac{\left(1-\frac{1}{k}\right)\left(1-\frac{2}{k}\right)\cdots\left(1-\frac{n-1}{k}\right)}{\left(1-\frac{1}{m}\right)\left(1-\frac{2}{m}\right)\cdots\left(1-\frac{n-1}{m}\right)} - 1\right],$$

поэтому

$$|\delta_k| < \left(\frac{k}{m}\right)^2 |a_2|\left[1-\left(1-\frac{1}{k}\right)\right] + \ldots + \left(\frac{k}{m}\right)^n |a_n|\left[1-\left(1-\frac{1}{k}\right)\cdots\left(1-\frac{n-1}{k}\right)\right] <$$

$$< \left(\frac{k}{m}\right)^2 |a_2|\frac{1}{k} + \ldots + \left(\frac{k}{m}\right)^n |a_n|\frac{(n-1)^2}{k} < \frac{B}{m},$$

где

$$B = |a_2| + 4|a_3| + \ldots + (n-1)^2|a_n|.$$

Итак

$$A_k = C_m^k\left[P\left(\frac{k}{m}\right) + \delta_k\right], \tag{64'}$$

где

$$|\delta_k| < \frac{B}{m}.$$

Следовательно,

$$\sum_{k=0}^{m} |A_k|\, x^k(1-x)^{m-k} = \sum_{k=0}^{m}\left|P\left(\frac{k}{m}\right) + \delta_k\right| C_m^k x^k (1-x)^{m-k} <$$

$$< \left(M + \frac{B}{m}\right)\sum_{k=0}^{m} C_m^k x^k (1-x)^{m-k} = \left(M + \frac{B}{m}\right)[x + (1-x)]^m = M + \frac{B}{m},$$

обозначая через M максимум многочлена $P(x)$ на отрезке $[0, 1]$. Таким образом, обозначая через M_m нормальный максимум степени m многочлена $P(x)$ на отрезке $[0, 1]$, имеем

$$M_m < M + \frac{B}{m}. \tag{65}$$

Следствие. *Если многочлен $P(x)$ положителен на отрезке $[0, 1]$, то, при m достаточно большом, все коэффициенты A_k положительны.*

60 [68]. **Теорема.** *Всякая непрерывная на отрезке $[0, 1]$ функция разлагается в нормальный ряд на отрезке.*

В самом деле, на основании теоремы Вейерштрасса, всякую непрерывную функцию $f(x)$ можно представить в виде равномерного сходящегося ряда многочленов

$$f(x) = Q_0(x) + Q_1(x) + \dots + Q_s(x) + \dots \tag{66}$$

Написанный ряд можно будет преобразовать в нормальный ряд следующим образом: соединяя вместе, если это понадобится, по нескольку членов, ряд (66) преобразуем в ряд

$$f(x) = P_0(x) + P_1(x) + \dots + P_s(x) + \dots, \tag{66'}$$

в котором все многочлены $P_s(x)$ (при $s > 0$) удовлетворяют условию

$$|P_s(x)| < \frac{1}{2^s}.$$

После этого представим все многочлены $P_s(x)$ в виде

$$P_s(x) = \sum_{k=0}^{m} A_k^{(s)} x^k (1-x)^{m-k}.$$

Полагая m достаточно большим, чтобы нормальный максимум $M_m^{(s)}$ многочлена P_s не превышал более чем в два раза его обыкновенного максимума, получим

$$\sum_{k=0}^{m} |A_k^{(s)}| x^k (1-x)^{m-k} < \frac{1}{2^{s-1}}.$$

Делая то же преобразование для всех s, мы, очевидно, преобразуем ряд $f(x)$ в нормальный ряд; что и требовалось доказать.

Следствие. *Для всякой непрерывной функции имеет место равенство*[1]

$$f(x) = \lim_{m \to \infty} \sum_{k=0}^{m} f\left(\frac{k}{m}\right) C_m^k x^k (1-x)^{m-k}.$$

---

[1] Эта формула выведена мною при помощи теории вероятностей в заметке «Démonstration du théorème de Weierstrass fondée sur le calcul des probabilités», помещенной в «Сообщ. Харьк. матем. об-ва», т. 13, № 1, 1912 [4].

В самом деле, если $f(x) = P(x)$ есть многочлен, то на основании равенства (64')

$$\left| P(x) - \sum_{k=0}^{m} P\left(\frac{k}{m}\right) C_m^k x^k (1-x)^{m-k} \right| < \frac{B}{m}. \qquad (67)$$

Если же $f(x)$ есть произвольная функция (66'), то, полагая

$$P_0 + P_1 + \ldots + P_s = P,$$

имеем

$$|f - P| < \frac{1}{2^s}; \qquad (68)$$

поэтому, применяя к многочлену $P(x)$ неравенство (67), заключаем, что

$$\left| f(x) - \sum_{k=0}^{m} f\left(\frac{k}{m}\right) C_m^k x^k (1-x)^{m-k} \right| < \frac{1}{2^{s-1}} + \frac{B}{m}.$$

Таким образом, как бы мало ни было число α, выбираем s достаточно большим, чтобы

$$\frac{1}{2^{s-1}} < \frac{\alpha}{2};$$

после выбора s многочлен P и коэффициент B будут определены, и, следовательно, выбирая m достаточно большим, найдем

$$\left| f(x) - \sum_{k=0}^{m} f\left(\frac{k}{m}\right) C_m^k x^k (1-x)^{m-k} \right| < \alpha,$$

т. е.

$$f(x) = \lim_{m \to \infty} \sum_{k=0}^{m} f\left(\frac{k}{m}\right) C_m^k x^k (1-x)^{m-k},$$

что и требовалось доказать.

Часть третья

РАЗЛОЖЕНИЕ НЕПРЕРЫВНЫХ ФУНКЦИЙ В РЯДЫ ТРИГОНОМЕТРИЧЕСКИХ МНОГОЧЛЕНОВ

Глава VI

О ПРИБЛИЖЕНИИ, ОСУЩЕСТВЛЯЕМОМ ПОСРЕДСТВОМ РАЗЛОЖЕНИЯ ФУНКЦИИ В РЯД ТРИГОНОМЕТРИЧЕСКИХ МНОГОЧЛЕНОВ

61 [69]. **Средняя квадратичная ошибка.** Отыскание многочлена данной степени, наименее уклоняющегося от некоторой функции $f(x)$, представляет, как это видно из предшествующих глав, задачу чрезвычайной трудности. Поэтому интересно выяснить, какую выгоду для решения этой задачи можно извлечь из решения другой аналогичной, но несравненно

более легкой задачи отыскания многочлена $R_n(x)$ степени n по условию чтобы *средняя квадратичная ошибка*

$$\int_a^b p(x)[f(x)-R_n(x)]^2\,dx$$

(при данном *весе* $p(x) \geqslant 0$) была возможно малой. Полагая, для определенности, $a=-1$, $b=1$, мы ограничимся рассмотрением случая[1], когда $p(x)=\frac{1}{\sqrt{1-x^2}}$.

Но

$$\delta_n^2=\int_{-1}^{+1}[f(x)-R_n(x)]^2\frac{dx}{\sqrt{1-x^2}}=\int_0^{\pi}[f(\cos\theta)-R_n(\cos\theta)]^2\,d\theta, \tag{69}$$

и, замечая (§ 10), что

$$R_n(\cos\theta)=A_0+A_1\cos\theta+\ldots+A_n\cos n\theta,$$

находим условия, необходимые и достаточные для минимума δ:

$$\begin{aligned} A_0&=\frac{1}{\pi}\int_0^{\pi}f(\cos\theta)\,d\theta=\frac{1}{2\pi}\int_0^{2\pi}f(\cos\theta)\,d\theta,\\ A_p&=\frac{2}{\pi}\int_0^{\pi}f(\cos\theta)\cos p\theta\,d\theta=\frac{1}{\pi}\int_0^{2\pi}f(\cos\theta)\cos p\theta\,d\theta. \end{aligned} \tag{70}$$

Формулы (70) дают не что иное, как хорошо известные коэффициенты Фурье[2] разложения функции $\varphi(\theta)=f(\cos\theta)$ в тригонометрический ряд. Эти же коэффициенты мы находим и для разложения $f(x)$ в ряд тригонометрических многочленов $T_n(x)=\cos n \arccos x$,

$$f(x)=A_0+A_1T_1(x)+\ldots+A_nT_n(x)+\ldots; \tag{71}$$

а многочлен

$$R_n(x)=A_0+A_1T_1(x)+\ldots+A_nT_n(x), \tag{71'}$$

обращающий в минимум среднюю квадратичную ошибку, получается, если в разложении (71) отбросить члены степени выше n.

В главе V (§ 51) мы уже рассматривали приближенные многочлены $R_n(x)$ и видели, что в некоторых редких случаях они дают асимптоти-

---

[1] Обобщение результатов, которые будут получены в этом случае, не представляет серьезных трудностей. См.: Haar. Orthogonale Funktionensysteme, «Math. Ann.», Bd. 69, 1910; В. А. Стеклов. Sur la théorie de fermeture des systèmes de fonctions orthogonales, «Зап. Акад. Наук», 1911.

[2] Коэффициенты при синусах равны нулю.

ческие выражения многочленов, наименее уклоняющихся от данной функции. Во многих случаях, как будет показано дальше,

$$1 < \frac{I_n[f(x)]}{E_n[f(x)]} < k, \tag{72}$$

где k — не зависящая от n постоянная, а $I_n[f(x)]$ есть максимум разности $|f(x) - R_n(x)|$. Но уже один тот факт, что существуют непрерывные функции, которые не могут быть разложены в сходящийся тригонометрический ряд, показывает, что неравенство (72) не всегда имеет место, так как возможно, что $E_n[f(x)]$ стремится к нулю, между тем как $I_n[f(x)]$ возрастает бесконечно. Исследование условий, каким должна удовлетворять функция $f(x)$, чтобы неравенство (72) было соблюдено, является, таким образом, непосредственным продолжением классической теории разложения функций в тригонометрический ряд.

62 [70]. Некоторые следствия из теоремы Рисса. Прежде чем перейти к изучению наименьшего уклонения с новой точки зрения, на которую мы становимся в этой главе, сделаем несколько замечаний о минимуме средней квадратичной ошибки, не имеющих прямого отношения к дальнейшему. Напомню сначала теорему Фридриха Рисса[1]: *для того, чтобы функция $\varphi(\theta)$ была квадратично интегрируема (т. е. чтобы интеграл $\int\limits_a^b \varphi^2(\theta)\,d\theta$, при $0 \leqslant a < b \leqslant 2\pi$, существовал в смысле Лебега[2]), необходимо и достаточно, чтобы ряд $\sum\limits_{p=0}^{\infty} A_p^2$, где через A_p обозначены коэффициенты Фурье (70) функции $\varphi(\theta)$, был сходящимся; при этом*

$$\frac{1}{2\pi}\int\limits_0^{2\pi} \varphi^2(\theta)\,d\theta = A_0^2 + \frac{1}{2}\sum_{p=1}^{\infty} A_p^2.$$

Применяя теорему Рисса к функции

$$-\varphi'(\theta) = f'(\cos\theta)\sin\theta = f'(x)\sqrt{1-x^2},$$

у которой коэффициенты Фурье равны pA_p, находим, что условие, необходимое и достаточное для того, чтобы интеграл

$$\int\limits_a^b [f'(x)]^2(1-x^2)\,dx$$

существовал (в смысле Лебега), при $-1 \leqslant a < b \leqslant 1$, заключается в

---

[1] Fr. Riesz. Untersuchungen über Systeme integrierbarer Funktionen, «Math. Ann.», Bd. 69.

[2] Lebesgue. Leçons sur l'intégration et la recherche des fonctions primitives. [Есть русский перевод: А. Лебег. Интегрирование и отыскание примитивных функций. ГТТИ, 1934. (*Ред.*)]

том, чтобы ряд $\sum\limits_{p=1}^{\infty} p^2 A_p^2$ был сходящимся (коэффициенты A_p даны формулами (70)), т. е. чтобы сумма $\beta_{p_0} = \sum\limits_{p=p_0+1}^{\infty} p^2 A_p^2$ стремилась к нулю с возрастанием p_0.

Но

$$\delta_{p_0}^2 = \frac{\pi}{2} \sum_{p=p_0+1}^{\infty} A_p^2, \tag{69'}$$

поэтому

$$(p_0+1)^2 \delta_{p_0}^2 < \pi\beta_{p_0} < (p_0+1)^2 \delta_{p_0}^2 + \sum_{p=p_0+1}^{\infty} (2p+1)\,\delta_p^2.$$

Таким образом, полагая $\delta_p = \frac{\varepsilon_p}{p+1}$, видим, что *для существования интеграла* $\int\limits_a^b [f'(x)]^2 (1-x^2)\,dx$ *необходимо, чтобы* $\varepsilon_p = \delta_p(p+1)$ *стремилось к нулю, и достаточно, чтобы ряд* $\sum\limits_{p=1}^{\infty} \frac{\varepsilon_p^2}{p+1} = \sum\limits_{p=1}^{\infty} (p+1)\,\delta_p^2$ *был сходящимся.* Последнее условие соблюдается, если $\varepsilon_p \leqslant \frac{1}{(\log p)^{\frac{1}{2}+\alpha}}$ или $\leqslant \frac{1}{(\log p)^{\frac{1}{2}} (\log\log p)^{\frac{1}{2}+\alpha}}$ и т. д. Аналогичные результаты можно получить и для последующих производных; не останавливаясь на этом, заметим только, что величина минимума средней квадратичной ошибки δ_n^2 также тесно связана с *интегрально-дифференциальными* свойствами функции *на всем промежутке,* как наименьшее уклонение $E_n[f(x)]$ связано с *дифференциальными* свойствами функции в каждой отдельной точке (глава II).

Примечание. Из равенств (69) и (69') видно, что $\delta_n < E_n[f(x)] \cdot \sqrt{\pi}$; поэтому*

$$\sqrt{\frac{1}{2}(A_{n+1}^2 + A_{n+2}^2 + \ldots)} < E_n[f(x)] < |A_{n+1}| + |A_{n+2}| + \ldots \tag{53'}$$

63 [71]. Теорема. *Для всякой непрерывной фукции* $f(x)$ *имеет место неравенство (при обозначениях* § 61)

$$\frac{I_n[f(x)]}{E_n[f(x)]} < k_1 \log(n+1), \tag{73}$$

где k_1 *— не зависящая от* n *и от функции* $f(x)$ *постоянная.*

Эта теорема вытекает из аналогичной теоремы, доказанной Лебегом в цитированной уже ранее работе «Sur les intégrales singulières»[1], отли-

---

* В первоначальном тексте в левой части неравенства (53') отсутствовал множитель $\sqrt{\frac{1}{2}}$. (*Автор.*)

[1] «Ann. de Toulouse», т. I (1909 г.) См. также упомянутую выше работу Джексона. В работе «Lebesguesche Konstanten und divergente Fourierreihe», «Journ. reine und angew. Math.», Bd. 138, Л. Фейер производит вычисление, из которого вытекает, что коэффициент k_1 в формуле (73) имеет пределом $8/\pi^2$ при $n \to \infty$.

чающейся от нашей теоремы тем, что у него I_n есть максимум разности $\left| f(x) - \sum_{p=0}^{n} A_p \cos px + B_p \sin px \right|$, где A_p и B_p — коэффициенты Фурье, а E_n — наилучшее приближение $f(x)$ при помощи тригонометрической суммы n-го порядка. Таким образом, считая теорему Лебега для тригонометрических сумм доказанной, мы получим неравенство (73), если, как в § 61, сделаем подстановку $x = \cos\theta$.

64 [72]. Следствия. 1) Лебег выводит из своей теоремы и из того, что наилучшее приближение E_n функций, удовлетворяющих условию Дини-Липшица, меньше, чем $\varepsilon_n \log(n+1)$, где $\lim\limits_{n\to\infty} \varepsilon_n = 0$, что эти функции разлагаются в сходящиеся тригонометрические ряды. Мы можем, следовательно, также утверждать на основании неравенств (73) и (55), что *всякая функция, удовлетворяющая условию Дини-Липшица, разлагается в сходящийся ряд тригонометрических многочленов.* Заметим, кроме того, что, вследствие замечания, заканчивающего § 55, *функция, удовлетворяющая обобщенному условию Дини-Липшица, разлагается в ряд тригонометрических многочленов, который можно сделать сходящимся простой группировкой членов.*

2) Теорема § 63 показывает нам, что, если вообще порядок убывания E_n не равен I_n, тем не менее последний всегда определяет порядок E_n, с точностью до множителя $\log(n+1)$. Укажем, например, верхнюю и нижнюю границу для $E_{2n}[|x|]$ в промежутке $(-1, +1)$. Для этого разлагаем $|x|$ в строку тригонометрических многочленов. Применяя формулы (70), находим

$$A_0 = \frac{1}{\pi}\int_0^{\pi} |\cos\theta|\, d\theta = \frac{2}{\pi}; \qquad A_{2k+1} = 0,$$

$$A_{2k} = \frac{2}{\pi}\int_0^{\pi} |\cos\theta| \cos 2k\theta\, d\theta = \frac{4}{\pi}\int_0^{\frac{\pi}{2}} \cos\theta \cos 2k\theta\, d\theta =$$

$$= \frac{2}{\pi}\int_0^{\frac{\pi}{2}} [\cos(2k+1)\theta + \cos(2k-1)\theta]\, d\theta = \frac{4}{\pi} \cdot \frac{(-1)^{k+1}}{(2k-1)(2k+1)}.$$

Следовательно,

$$|x| = \frac{4}{\pi}\left[\frac{1}{2} + \frac{T_2}{1\cdot 3} - \frac{T_4}{3\cdot 5} + \frac{T_6}{5\cdot 7} - \ldots\right]; \tag{74}$$

поэтому

$$I_{2n} = \frac{4}{\pi}\left[\frac{1}{(2n+1)(2n+3)} + \frac{1}{(2n+3)(2n+5)} + \ldots\right] = \frac{2}{\pi(2n+1)}. \tag{75}$$

Таким образом, на основании теоремы § 63

$$\frac{k_1}{(2n+1)\log(2n+1)} < E_{2n} < \frac{2}{\pi(2n+1)}.$$

Первая часть этого неравенства[1], разумеется, несравненно менее

---

[1] Она имеется и в упомянутой работе Джексона, который, независимо от меня, получил ее аналогичным образом.

удовлетворительна, чем результаты, найденные нами ранее; но вторая часть неравенства дает довольно точную верхнюю границу $E_{2n} < \frac{0,637}{2n}$. Другими словами, приближение $|x|$, которое дает столь простое разложение (74), лишь незначительно хуже наилучшего приближения; а именно, припоминая неравенство (54) § 55, имеем (по крайней мере для весьма больших значений n)

$$2,22 < \frac{I_{2n}|x|}{E_{2n}|x|} < 2,36. \tag{76}$$

65 [73]. **Теорема.** *Если функция $f(x)$ удовлетворяет условию Липшица степени $\alpha < 1$, то*

$$E_n[f(x)] < \frac{k}{n^\alpha}, \tag{77}$$

где k — не зависящий от n коэффициент; при этом многочлены степени n, осуществляющие приближение k/n^α, получаются посредством применения способа суммирования Фейера к разложению рассматриваемой функции в ряд тригонометрических многочленов. (То же самое mutatis mutandis *имеет место и для тригонометрических сумм.)*

В самом деле, полагая $x = \cos\theta$ и обозначая через

$$S_n = A_0 + A_1T_1(x) + \ldots + A_nT_n(x) = A_0 + A_1\cos\theta + \ldots + A_n\cos n\theta$$

сумму $n+1$ членов разложения $f(x) = f(\cos\theta) = \varphi(\theta)$, мы получим приближенную сумму Фейера $(n-1)$-го порядка

$$\sigma_n = \frac{s_0 + s_1 + \ldots + s_{n-1}}{n};$$

и при этом остаток R_n равен[1]

$$R_n = \sigma_n - \varphi(\theta) = \frac{1}{n\pi}\int_0^{\frac{\pi}{2}}\left(\frac{\sin nt}{\sin t}\right)^2[\varphi(\theta+2t) + \varphi(\theta-2t) - 2\varphi(\theta)]\,dt.$$

По предположению,

$$|f(x+h) - f(x)| < Nh^\alpha,$$

где N — данное число, а следовательно, и

$$|\varphi(\theta+2t) - \varphi(\theta)| < N(2t)^\alpha = Mt^\alpha.$$

Поэтому

$$|R_n| < \frac{2M}{n\pi}\int_0^{\frac{\pi}{2}}\left(\frac{\sin nt}{\sin t}\right)^2 t^\alpha dt <$$

$$< \frac{2M}{n\pi}\left[\int_0^{\frac{1}{n}} n^2t^\alpha dt + \int_{\frac{1}{n}}^{\frac{\pi}{2}} \frac{t^\alpha dt}{\sin^2 t}\right] < \frac{2M}{\pi n^\alpha}\left[\frac{1}{1+\alpha} + \frac{\pi^2}{4}\frac{1}{1-\alpha}\right].$$

---

[1] Lebesgue. Leçons sur les séries trigonométriques (стр. 94).

Таким образом, при $\alpha < 1$,

$$|f(x) - \sigma_n| < \frac{k}{n^\alpha},$$

где k — не зависящий от n коэффициент, что и требовалось доказать.

Примечание. Из доказательства видно, что вывод не нарушится, если даже N не постоянная величина, а возрастает бесконечно при $x = \pm 1$. С тем обстоятельством, что одна и та же особенность функции внутри отрезка и на концах его не одинаково влияет на приближение функции при помощи многочленов, мы уже встречались в главе II. Не останавливаясь на подробном исследовании этого вопроса, укажем лишь один простой пример, на котором отчетливо видна сущность этой разницы: из доказанной теоремы вытекает, что $E_{2n}|x|^\alpha < \frac{k}{(2n)^\alpha}$, где $\alpha < 1$; при этом ясно, что многочлен степени $2n$, наименее уклоняющийся от $|x|^\alpha$, не содержит нечетных степеней x; поэтому, полагая $x^2 = y$, мы видим, что наименьшее уклонение $E'_n(y^{\alpha/2})$ на отрезке $[0, 1]$ также удовлетворяет неравенству $E'_n(y^{\alpha/2}) = E_{2n}|x|^\alpha < \frac{k}{(2n)^\alpha}$. Другими словами, условие Липшица степени α внутри отрезка имеет существенно то же значение для наименьшего уклонения, что условие Липшица степени $\alpha/2$ в концах отрезка.

66 [74]. Результаты Джексона [1]. Нетрудно заметить, что остаток, получаемый при применении способа Фейера в случае, когда $\alpha = 1$, не подчиняется закону, выраженному предшествующей теоремой: в этом случае можно утверждать только, что

$$|R_n| < \frac{k \log n}{n}.$$

Джексон, независимо от меня, при помощи другого метода, получил более законченный результат, а именно, он показал, что при $\alpha = 1$

$$E_n[f(x)] < \frac{k}{n}. \tag{77'}$$

Кроме того, он доказал еще, что, если $f(x)$ имеет производную p-го порядка, удовлетворяющую условию Липшица степени $\alpha \leqslant 1$, то [3.8]

$$E_n[f(x)] < \frac{k}{n^{p+\alpha}}, \tag{78}$$

где k — не зависящая от n постоянная.

Следствие. *Если функция $f(x)$ удовлетворяет условию Липшица степени $\alpha \leqslant 1$, то*

$$I_n[f(x)] < \frac{k_2 \log n}{n^\alpha}. \tag{79}$$

---

[1] D. Jackson. Über die Genauigkeit der Annäherung stetiger Funktionen durch ganze rationale Funktionen. Этот параграф, разумеется, не мог войти в первоначальную (французскую) редакцию моего сочинения, как и все ссылки на работу Джексона.

Это вытекает из неравенств (77) и (77′) благодаря неравенству (73).

Примечание. Этот результат, для тригонометрических сумм, был получен непосредственно Лебегом[1], который показал также, что верхняя граница $I_n[f(x)]$ не может быть понижена, если о функции $f(x)$ ничего более не известно. Отсюда следует, что и верхняя граница $E_n[f(x)]$, найденная Джексоном и мной, также не может быть понижена, если взять неопределенную функцию, удовлетворяющую данному условию Липшица. Если принять неравенство (78), то из него точно так же можно получить, что

$$I_n[f(x)] < \frac{k \log n}{n^{p+\alpha}} \qquad (79')$$

для функций, имеющих p-ю производную, удовлетворяющую условию Липшица степени α.

Но я воспроизведу с небольшим упрощением свой первоначальный вывод неравенства (79′), который представляет, быть может, некоторый принципиальный интерес.

67 [75]. Доказательство неравенства (79′). Заметим прежде всего, что условие, что $\frac{d^p f(x)}{dx^p}$ удовлетворяют условию Липшица степени α, влечет за собой существование условия Липшица степени α для $\frac{d^p \varphi(\theta)}{d\theta^p}$.

Рассмотрим сперва четные значения $p = 2\mu$. Пусть

$$\varphi(\theta) = f(\cos\theta) = A_0 + A_1 \cos\theta + \ldots + A_n \cos n\theta + \ldots ;$$

в таком случае,

$$\frac{d^p \varphi(\theta)}{d\theta^p} = \pm [A_1 \cos\theta + \ldots + n^p A_n \cos n\theta + \ldots].$$

Полагая

$$\rho_n = (n+1)^p A_{n+1} \cos(n+1)\theta + (n+2)^p A_{n+2} \cos(n+2)\theta + \ldots,$$

мы заключаем из неравенства (79), что

$$|\rho_n| < \frac{k \log n}{n^\alpha}.$$

А потому, на основании известной леммы Абеля,

$$|R_n| = |A_{n+1} \cos(n+1)\theta + A_{n+2} \cos(n+2)\theta + \ldots| < \frac{|\rho_n|}{(n+1)^p} < \frac{k \log n}{n^{p+\alpha}}.$$

Для рассмотрения случая, когда $p = 2\mu - 1$ нечетное число, выведем предварительно следующее неравенство, справедливое для всякого значения $s > 1$: если

$$|R_n| = |A_{n+1} \cos(n+1)\theta + A_{n+2} \cos(n+2)\theta + \ldots| < \frac{k \log n}{n^s}, \qquad (80)$$

---

[1] Lebesgue. Sur la représentation trigonométrique approchée des fonctions satisfaisant à une condition de Lipschitz. «Bull. Soc. Math. de France», 1910.

то

$$|R'_n| = |(n+1)A_{n+1}\sin(n+1)\theta +$$
$$+ (n+2)A_{n+2}\sin(n+2)\theta + \ldots| < \frac{2^{s+1}k\log n}{(2^{s-1}-1)^2 n^{s-1}}\,.$$

В самом деле, из (80) вытекает, что

$$|A_{n+1}\cos(n+1)\theta + \ldots + A_{2n}\cos 2n\theta| < \frac{2k\log n}{n^s}\,,$$

а потому, вследствие § 10,

$$|(n+1)A_{n+1}\sin(n+1)\theta + \ldots + 2nA_{2n}\sin 2n\theta| < \frac{4k\log n}{n^{s-1}}\,.$$

Следовательно,

$$|R'_n| < \frac{4k}{n^{s-1}}\left[\log n + \frac{\log 2n}{2^{s-1}} + \frac{\log 4n}{4^{s-1}} + \ldots\right] =$$
$$= \frac{4k\log n}{n^{s-1}}\cdot\frac{2^{s-1}}{2^{s-1}-1} + \frac{4k\log 2}{n^{s-1}}\cdot\frac{2^{s-1}}{(2^{s-1}-1)^2} < \frac{2^{s+1}k\log n}{(2^{s-1}-1)^2 n^{s-1}}\,. \qquad (81)$$

Само собой понятно, что то же самое неравенство мы получим и в том случае, когда R_n состоит из синусов.

После этого, берем функцию

$$\Phi(\theta) = \int_0^\theta \varphi(\theta)\,d\theta,$$

где попрежнему $\varphi(\theta) = f(\cos\theta)$.

В таком случае остаток R_n тригонометрического разложения функции $\Phi(\theta)$, имеющей производную четного порядка $p+1 = 2\mu$, удовлетворяет неравенству

$$|R_n| < \frac{k\log n}{n^{p+1+\alpha}}\,,$$

а следовательно, остаток $|R'_n|$ в разложении $\varphi(\theta)$, вследствие (81), будет менее, чем

$$\frac{2^{p+2}k\log n}{(2^p-1)^2 n^{p+\alpha}}\,;$$

таким образом, неравенство (79′) справедливо для всякого p.

Следствия. а) *Если функция $f(x)$ в промежутке $(-1, +1)$ имеет производные всех порядков, то ее разложение в ряд тригонометрических многочленов равномерно сходится, так же как и ряды, получаемые от дифференцирования рассматриваемого разложения какое угодно число раз.*

б) *Если функция $f(x)$ имеет производные всех порядков в промежутке $(-1, +1)$, то*

$$\lim_{n\to\infty} n^p I_n[f(x)] = \lim_{n\to\infty} n^p E_n[f(x)] = 0,$$

при всяком p (теорема § 22).

68 [76]. **Теорема**[1]. *Если модуль аналитической функции $f(x)$ менее M внутри эллипса E, имеющего фокусами точки -1, $+1$ и полусумму осей, равную $1/\rho$, то*

$$E_n[f(x)] < I_n[f(x)] < \frac{2M\rho^{n+1}}{1-\rho}$$

на отрезке $[-1, +1]$.

В самом деле, согласно формулам (70),

$$A_p = \frac{1}{\pi}\int\limits_0^{2\pi} f(\cos\theta)\cos p\theta\, d\theta,$$

или, полагая $z = e^{i\theta}$,

$$A_p = \frac{1}{2\pi i}\int\limits_C f\left(\frac{z^2+1}{2z}\right)\frac{z^p + z^{-p}}{z}dz,$$

причем последний интеграл взят по окружности C радиуса, равного единице. В то время как комплексная переменная x описывает эллипс E, комплексная переменная z описывает либо окружность C_1 радиуса ρ, либо окружность C_2 радиуса $1/\rho$, так как

$$x = \cos\theta = \frac{1}{2}\left(z + \frac{1}{z}\right).$$

Но $f(x)$, по предположению, остается голоморфной внутри эллипса E; поэтому $f\left(\frac{z^2+1}{2z}\right)$ также голоморфна между окружностями C_1 и C_2. Следовательно,

$$\left|\int\limits_C f\left(\frac{z^2+1}{2z}\right)z^{p-1}dz\right| = \left|\int\limits_{C_1} f\left(\frac{z^2+1}{2z}\right)z^{p-1}dz\right| < 2\pi M\rho^p$$

и

$$\left|\int\limits_C f\left(\frac{z^2+1}{2z}\right)\frac{dz}{z^{p+1}}\right| = \left|\int\limits_{C_2} f\left(\frac{z^2+1}{2z}\right)\frac{dz}{z^{p+1}}\right| < 2\pi M\rho^p,$$

откуда

$$|A_p| < 2M\rho^p.$$

И, наконец,

$$I_n[f(x)] < [|A_{n+1}| + |A_{n+2}| + \ldots] < \frac{2M\rho^{n+1}}{1-\rho},$$

что и требовалось доказать.

Примечание. В предшествующей теореме, так же, как и в условиях теорем §§ 22 и 29, наименьшее уклонение E_n может быть заменено минимумом средней квадратичной ошибки δ_n.

69 [77]. Различные следствия и приложения.

A) *Если функция $f(x)$ в промежутке $(-1, +1)$ имеет производную порядка k, полное изменение которой ограничено, то*

$$I_n[f(x)] < \frac{h'}{n^k},$$

где h' — не зависящая от n постоянная.

---

[1] См. теорему § 29.

В самом деле, согласно формуле (70),

$$A_p = \frac{1}{\pi}\int_0^{2\pi} f(\cos\theta)\cos p\theta\, d\theta = \frac{1}{\pi p^k}\int_0^{2\pi}\frac{d^k f(\cos\theta)}{d\theta^k}\cos\left(p\theta - \frac{k\pi}{2}\right)d\theta,$$

а потому

$$|A_p| < \frac{h}{p^{k+1}},$$

где h — не зависящий от p коэффициент; следовательно,

$$I_n[f(x)] < h\left[\frac{1}{(n+1)^{k+1}} + \frac{1}{(n+2)^{k+1}} + \ldots\right] < \frac{h'}{n^k}.$$

Б) *Если линия* $y = f(x)$ *имеет одну или несколько точек излома, а между точками излома угловой коэффициент касательной удовлетворяет какому-нибудь условию Липшица, то*

$$\frac{a}{n} < E_n[f(x)] < I_n[f(x)] < \frac{b}{n}, \qquad (72')$$

где a *и* b *— два независящих от* n *числа.*

В самом деле, пусть x_0 и x_1 будут абсциссы точек излома. В таком случае,

$$f(x) = M|x - x_0| + N|x - x_1| + \varphi(x),$$

где M и N — постоянные коэффициенты, а $\varphi(x)$ удовлетворяет условию Липшица на всем промежутке. Поэтому

$$E_n[f(x)] < I_n[f(x)] < MI_n[|x - x_0|] + NI_n[|x - x_1|] + I_n[\varphi(x)] < \frac{b}{n}.$$

С другой стороны, ясно, что наименьшее уклонение E_n на всем отрезке не меньше, чем наименьшее уклонение E_n' на части его, содержащей лишь одну точку излома; следовательно,

$$E_n[f(x)] > E_n'[f(x)] > ME_n'[|x - x_0|] - E_n[\varphi(x)] > \frac{a}{n}.$$

В) *Если*

$$\left|\sum_{p=n}^{\infty} A_p \cos p\theta\right| < \frac{\lambda_n}{n^\alpha},$$

причем

$$\frac{\lambda_n}{n^\varepsilon} \geqslant \frac{\lambda_{n+1}}{(n+1)^\varepsilon},$$

где $\varepsilon < \alpha$, *то*

$$\left|\sum_{p=n}^{\infty} A_p \sin p\theta\right| < \frac{2^{2+\alpha}}{2^\alpha - 2^\varepsilon}\,\frac{\lambda_n}{n^\alpha}.$$

В самом деле, применяя лемму Абеля, замечаем, что

$$\left|\sum_{p=n}^{\infty}\frac{A_p}{p}\cos p\theta\right| < \frac{\lambda_n}{n^{1+\alpha}}.$$

В таком случае,

$$\left|\sum_{p=n}^{2n-1}\frac{A_p}{p}\cos p\theta\right|<\frac{2\lambda_n}{n^{1+\alpha}},$$

а потому, вследствие § 10,

$$\left|\sum_{p=n}^{2n-1}A_p\sin p\theta\right|<\frac{4\lambda_n}{n^{\alpha}},$$

откуда

$$\left|\sum_{p=n}^{\infty}A_p\sin p\theta\right|<\frac{4}{n^{\alpha}}\left[\lambda_n+\frac{\lambda_{2n}}{2^{\alpha}}+\frac{\lambda_{4n}}{4^{\alpha}}+\ldots\right]<\frac{4\lambda_n}{n^{\alpha}}\cdot\frac{1}{1-2^{\varepsilon-\alpha}}=\frac{2^{2+\alpha}}{2^{\alpha}-2^{\varepsilon}}\cdot\frac{\lambda_n}{n^{\alpha}}.$$

Например, если $\lambda_n=\log n$ или $\lambda_n=1$, то ε сколь угодно близко к нулю для весьма больших значений n, так что из неравенства

$$\left|\sum_{p=n}^{\infty}A_p\cos p\theta\right|<\frac{\log n}{n}$$

вытекает

$$\left|\sum_{p=n}^{\infty}A_p\sin p\theta\right|<\frac{8\log n}{n};$$

а из

$$\left|\sum_{p=n}^{\infty}A_p\cos p\theta\right|<\frac{1}{n}$$

вытекает

$$\left|\sum_{p=n}^{\infty}A_p\sin p\theta\right|<\frac{8}{n}.$$

Само собой понятно, что косинусы и синусы могут быть взаимно перемещены. Этот результат заслуживает внимания потому, что вообще из сходимости ряда косинусов нельзя вывести сходимости ряда синусов и обратно. Например, сумма

$$\sum_{p=1}^{\infty}\frac{\sin px}{p}=\frac{\pi}{2}-\frac{x}{2}$$

конечна, а между тем не только $\sum_{p=1}^{\infty}\frac{\cos px}{p}$, но и $\sum_{p=2}^{\infty}\frac{\cos px}{p\log p}$ возрастает бесконечно.

Относительно медленно сходящихся рядов при помощи предыдущего рассуждения нетрудно показать, что если [1]

$$\left|\sum_{p=n}^{\infty}A_p\cos p\theta\right|<\varepsilon_n,$$

---

[1] Для определенности мы рассматриваем все время все значения θ; но аналогичные неравенства могут быть даны, если вместо всех значений θ брать в данном неравенстве $-\pi<a\leqslant\theta\leqslant b<\pi$, а в том, которое из него вытекает, предполагать $a<a'\leqslant\theta\leqslant b'<b$.

где числа ε_n идут не возрастая, то

$$\left|\sum_{p=n}^{\infty} A_p \sin p\theta\right| < 4(\varepsilon_n + \varepsilon_{2n} + \ldots);$$

таким образом, из сходимости ряда косинусов можно вывести сходимость ряда синусов, когда ряд $\varepsilon_n + \varepsilon_{2n} + \ldots$ сходится, т. е., например, если $\varepsilon_n < \frac{1}{(\log n)^{1+\alpha}}$.

При помощи тех же рассуждений можно показать, что если ряд $\sum_{n=0}^{\infty} A_n \cos nx$ не ограничен, то не ограничены также и ряды

$$\sum_{n=1}^{\infty} A_n (\log n)^{1+\varepsilon} \sin nx; \quad \sum_{n=1}^{\infty} A_n \log n\,(\log\log n)^{1+\varepsilon} \sin nx \qquad (\varepsilon > 0)$$

и т. д.

Отсюда следует, в частности, что ряды

$$\sum_{n=1}^{\infty} (\log n)^{\varepsilon} \frac{\sin nx}{n} \quad \text{и} \quad \sum_{n=1}^{\infty} (\log\log n)^{\varepsilon} \frac{\sin nx}{n}$$

не могут быть ограничены [3.9].

Глава VII

О НЕКОТОРЫХ СВОЙСТВАХ ФУНКЦИЙ ДВУХ ПЕРЕМЕННЫХ [3.10]

70 [78]. **Введение.** В настоящее время еще весьма мало изучен вопрос, какова зависимость между свойствами функции $f(x, y)$, рассматриваемой как функция двух переменных, и свойствами той же функции, рассматриваемой как функция одного только x и одного только y. Некоторые простые примеры, вроде функции $z = \frac{xy}{x^2 + y^2}$, дали повод преувеличить трудность этого вопроса. Действительно, функция z вещественной переменной x голоморфна при всяком определенном значении вещественного параметра y, и точно так же функция z голоморфна относительно y при всяком x, а между тем та же функция z, рассматриваемая как функция x и y одновременно, при $x = y = 0$ не только не голоморфна, но не стремится ни к какому пределу.

Пользуясь соотношениями между приближением функции посредством многочленов или тригонометрических сумм и ее дифференциальной природой, можно однако указать ряд теорем, которые во многих случаях позволяют свести исследование функции двух (или n) переменных к исследованию двух (или n) функций одной переменной.

Рассмотрим сначала периодические функции. Пусть

$$f(x, y) = \sum_{p=0}^{\infty} \sum_{q=0}^{\infty} [A_{pq} \cos px \cos qy + B_{pq} \cos px \sin qy +$$
$$+ C_{pq} \sin px \cos qy + D_{pq} \sin px \sin qy]$$

разлагается в тригонометрический ряд, абсолютно и равномерно сходящийся. Назовем *обобщенным модулем* функции $f(x, y)$ сумму*

$$\boxed{f(x, y)} = \sum_{p=0}^{\infty} \sum_{q=0}^{\infty} [|A_{pq}| + |B_{pq}| + |C_{pq}| + |D_{pq}|].$$

Приняв это, легко доказать следующую теорему.

71[79]. Теорема. *Если $f(x, y)$ обладает частной производной k-го порядка по x и частной производной k-го порядка по y и если их обобщенные модули удовлетворяют неравенствам*

$$\boxed{\frac{\partial^k f(x, y)}{\partial x^k}} < M \quad \text{и} \quad \boxed{\frac{\partial^k f(x, y)}{\partial y^k}} < M,$$

то существуют все частные производные порядка k от $f(x, y)$, которые будут ограничены и непрерывны, а их обобщенные модули удовлетворяют неравенству

$$\boxed{\frac{\partial^k f(x, y)}{\partial x^l \partial y^{(k-l)}}} < 2M \qquad (l = 1, 2, \ldots, k-1).$$

Действительно, по предположению,

$$\boxed{\frac{\partial^k f}{\partial x^k}} = \sum_{p=0}^{\infty} \sum_{q=0}^{\infty} p^k [|A_{pq}| + |B_{pq}| + |C_{pq}| + |D_{pq}|] < M,$$

$$\boxed{\frac{\partial^k f}{\partial y^k}} = \sum_{p=0}^{\infty} \sum_{q=0}^{\infty} q^k [|A_{pq}| + |B_{pq}| + |C_{pq}| + |D_{pq}|] < M.$$

Если ряд

$$\sum_{p=0}^{\infty} \sum_{q=0}^{\infty} p^{k_1} q^{k_2} \Big[A_{pq} \cos\left(px + \frac{k_1\pi}{2}\right) \cos\left(qy + \frac{k_2\pi}{2}\right) +$$

$$+ B_{pq} \cos\left(px + \frac{k_1\pi}{2}\right) \sin\left(qy + \frac{k_2\pi}{2}\right) + C_{pq} \sin\left(px + \frac{k_1\pi}{2}\right) \cos\left(qy + \frac{k_2\pi}{2}\right)$$

$$+ D_{pq} \sin\left(px + \frac{k_1\pi}{2}\right) \sin\left(qy + \frac{k_2\pi}{2}\right)\Big]$$

сходится абсолютно и равномерно, то он представляет, очевидно, $\frac{\partial^{k_1+k_2} f}{\partial x^{k_1} \partial y^{k_2}}$. Но, если $k_1 + k_2 = k$, то $p^k + q^k > p^{k_1} q^{k_2}$, так как $\left(\frac{p}{q}\right)^{k_2} + \left(\frac{q}{p}\right)^{k_1} > 1$, откуда следует, что

$$\sum_{p=0}^{\infty} \sum_{q=0}^{\infty} p^{k_1} q^{k_2} [|A_{pq}| + |B_{pq}| + |C_{pq}| + |D_{pq}|] < 2M,$$

чем доказательство завершается.

---

* В первоначальном тексте отсутствовало определение обобщенного модуля, которое приводится здесь по мемуару «О». Формулировка теоремы § 71 перередактирована с учетом этого определения. (*Ред.*)

72 [80]. Теорема. *Если периодическая относительно* (u, v) *функция* $\varphi(u, v)$ *имеет квадратично интегрируемые вторые частные производные* $\frac{\partial^2\varphi}{\partial u^2}$ *и* $\frac{\partial^2\varphi}{\partial v^2}$, *т. е.*

$$\int_0^{2\pi}\int_0^{2\pi}\left(\frac{\partial^2\varphi}{\partial u^2}\right)^2 du\,dv < M, \qquad \int_0^{2\pi}\int_0^{2\pi}\left(\frac{\partial^2\varphi}{\partial v^2}\right)^2 du\,dv < M,$$

то она имеет также квадратично интегрируемую частную производную $\frac{\partial^2\varphi}{\partial u\,\partial v}$, *а именно*

$$\int_0^{2\pi}\int_0^{2\pi}\left(\frac{\partial^2\varphi}{\partial u\,\partial v}\right)^2 du\,dv = \int_0^{2\pi}\int_0^{2\pi}\frac{\partial^2\varphi}{\partial u^2}\,\frac{\partial^2\varphi}{\partial v^2}\,du\,dv < M.$$

В самом деле, полагая

$$\begin{aligned} A_{pq} &= \frac{1}{\pi^2}\int_0^{2\pi}\int_0^{2\pi}\varphi(u, v)\cos pu\cos qv\,du\,dv, \\ B_{pq} &= \frac{1}{\pi^2}\int_0^{2\pi}\int_0^{2\pi}\varphi(u, v)\cos pu\sin qv\,du\,dv, \end{aligned} \tag{82}$$

и т. д., получаем

$$\int_0^{2\pi}\int_0^{2\pi}\left(\frac{\partial^2\varphi}{\partial u^2}\right)^2 du\,dv = \pi^2\sum_{p=0,\ q=0}^{p=\infty,\ q=\infty} p^4\left[A_{pq}^2 + B_{pq}^2 + \ldots\right] < M,$$

$$\int_0^{2\pi}\int_0^{2\pi}\left(\frac{\partial^2\varphi}{\partial v^2}\right)^2 du\,dv = \pi^2\sum_{p=0,\ q=0}^{p=\infty,\ q=\infty} q^4\left[A_{pq}^2 + B_{pq}^2 + \ldots\right] < M.$$

На основании теоремы Рисса, для того чтобы $\frac{\partial^2\varphi}{\partial u\,\partial v}$ была квадратично интегрируема, необходимо и достаточно, чтобы ряд

$$S = \pi^2\sum_{p=0,\ q=0}^{p=\infty,\ q=\infty} p^2q^2\left[A_{pq}^2 + B_{pq}^2 + \ldots\right]$$

был сходящимся, и тогда

$$\int_0^{2\pi}\int_0^{2\pi}\left(\frac{\partial^2\varphi}{\partial u\,\partial v}\right)^2 du\,dv = S.$$

Но

$$S = \int_0^{2\pi}\int_0^{2\pi}\frac{\partial^2\varphi}{\partial u^2}\,\frac{\partial^2\varphi}{\partial v^2}\,du\,dv < M.$$

73 [81]. Теорема. *Если периодическая функция* $\varphi(u, v)$, *рассматриваемая как функция* u, *имеет частную производную* $\frac{\partial^l\varphi}{\partial u^l}$, *удовлетворяю-*

щую определенному условию Липшица степени α, и точно так же, рассматриваемая как функция v, имеет производную $\frac{\partial^l \varphi}{\partial v^l}$, удовлетворяющую условию Липшица степени α, то функция $\varphi(u, v)$ имеет все частные производные порядка l, и эти последние также удовлетворяют условиям Липшица какой угодно степени $\alpha_1 < \alpha$ (относительно обеих переменных).

Пусть

$$\varphi(u, v) = \sum_{p=0,\ q=0}^{p=\infty,\ q=\infty} A_{pq} \cos pu \cos qv,$$

где, для сокращения письма, мы записываем только член, составленный из косинусов.

Припоминая значение коэффициентов A_{pq} (82), находим

$$S_n = \sum_{p=0,\ q=0}^{p=n,\ q=n} A_{pq} \cos pu \cos qv =$$

$$= \frac{1}{\pi^2} \int_0^{\frac{\pi}{2}} \int_0^{\frac{\pi}{2}} \frac{\sin(2n+1)t}{\sin t} \frac{\sin(2n+1)\theta}{\sin \theta} [\varphi(u+2t,\ v+2\theta) +$$

$$+ \varphi(u-2t,\ v+2\theta) + \varphi(u+2t,\ v-2\theta) + \varphi(u-2t, v-2\theta)]\, dt\, d\theta,$$

откуда

$$R_n = \varphi(u,\ v) - S_n = \frac{-1}{\pi^2} \int_0^{\frac{\pi}{2}} \int_0^{\frac{\pi}{2}} \frac{\sin(2n+1)t}{\sin t} \frac{\sin(2n+1)\theta}{\sin \theta} \times$$

$$\times \{[\varphi(u+2t,\ v+2\theta) + \varphi(u-2t,\ v+2\theta) - 2\varphi(u,\ v+2\theta)] +$$

$$+ [\varphi(u+2t,\ v-2\theta) + \varphi(u-2t,\ v-2\theta) - 2\varphi(u,\ v-2\theta)] +$$

$$+ 2[\varphi(u,\ v+2\theta) + \varphi(u,\ v-2\theta) - 2\varphi(u,\ v)]\}du\, dv. \tag{83}$$

Но

$$\varphi(u,\ v+2\theta) = \sum_{p=0}^{\infty} a_p(v+2\theta) \cos pu,$$

$$\varphi(u,\ v-2\theta) = \sum_{p=0}^{\infty} a_p(v-2\theta) \cos pu, \qquad \varphi(u,\ v) = \sum_{p=0}^{\infty} b_p(u) \cos pv,$$

где

$$a_p(z) = \frac{1}{\pi} \int_0^{2\pi} \varphi(u,\ z) \cos pu\, du, \qquad b_p(u) = \frac{1}{\pi} \int_0^{2\pi} \varphi(u,\ v) \cos pv\, dv;$$

поэтому

$$\rho_n(u,\ v+2\theta) = \sum_{p=n+1}^{\infty} a_p(v+2\theta) \cos pu =$$

$$= -\frac{1}{\pi} \int_0^{\frac{\pi}{2}} \frac{\sin(2n+1)t}{\sin t} [\varphi(u+2t,\ v+2\theta) + \varphi(u-2t,\ v+2\theta) - 2\varphi(u,\ v+2\theta)]dt,$$

$$\rho_n(u, v - 2\theta) = \sum_{p=n+1}^{\infty} a_p(v - 2\theta)\cos pu =$$

$$= -\frac{1}{\pi}\int_0^{\frac{\pi}{2}} \frac{\sin(2n+1)t}{\sin t}[\varphi(u + 2t, v - 2\theta) + \varphi(u - 2t, v - 2\theta) - 2\varphi(u, v - 2\theta)]dt,$$

$$\rho'_n(u, v) = \sum_{p=n+1}^{\infty} b_p(u)\cos pv =$$

$$= -\frac{1}{\pi}\int_0^{\frac{\pi}{2}} \frac{\sin(2n+1)\theta}{\sin\theta}[\varphi(u, v + 2\theta) + \varphi(u, v - 2\theta) - 2\varphi(u, v)]\,d\theta.$$

Следовательно, на основании неравенства (79′),

$$|\rho_n(u, v + 2\theta)| < \frac{k\log n}{n^{l+\alpha}}; \qquad |\rho_n(u, v - 2\theta)| < \frac{k\log n}{n^{l+\alpha}},$$

$$|\rho'_n(u, v)| < \frac{k\log n}{n^{l+\alpha}}.$$

А потому

$$|R_n| < \frac{4k\log n}{\pi n^{l+\alpha}}\int_0^{\frac{\pi}{2}}\left|\frac{\sin(2n+1)t}{\sin t}\right|dt.$$

Последний интеграл вычислен Фейером [1]; но нам достаточно заметить, что

$$\int_0^{\frac{\pi}{2}}\left|\frac{\sin(2n+1)t}{\sin t}\right|dt < \int_0^{\frac{1}{2n+1}}(2n+1)\,dt + \int_{\frac{1}{2n+1}}^{\frac{\pi}{2}}\frac{dt}{\sin t} < 1 + \frac{\pi}{2}\log(2n+1).$$

Следовательно (для достаточно больших n),

$$|R_n| < \frac{2k\log^2 n}{n^{l+\alpha}},$$

и при всяком $\alpha_1 < \alpha$ можно выбрать k_1 так, чтобы

$$|R_n| < \frac{k_1}{n^{l+\alpha_1}(\log n)^2}.$$

Но в таком случае, *применяя результаты главы* II (§§ 15—17), *убеждаемся в существовании всех частных производных* $\frac{\partial^l\varphi}{\partial u^i\partial v^{l-i}}$ *и в том, что они удовлетворяют условию Липшица степени* α_1. Что и требовалось доказать.

Примечание. В частности, если функция $\varphi(u, v)$ удовлетворяет условию Липшица степени α по отношению к каждой переменной в от-

---

[1] См. сноску к § 63.

дельности, то она удовлетворяет также условию Липшица степени α_1 относительно обеих переменных.

74 [82]. **Следствия.** А. *Если функция $f(x, y)$ (непериодическая), рассматриваемая как функция одного только x и одного только y, имеет внутри некоторого контура C производную порядка l, удовлетворяющую условию Липшица степени α, то функция $f(x, y)$ имеет все частные производные порядка l, и эти последние во всякой области S внутри контура C удовлетворяют условиям Липшица любой степени $\alpha_1 < \alpha$.*

В самом деле, всю область S можно поместить внутри нескольких квадратов C_1, стороны которых не выходят из контура C. Для определенности положим, что прямые, на которых расположены стороны квадрата C_1, имеют уравнениями: $x = \pm 1$, $y = \pm 1$. В таком случае, полагая $x = \cos u$, $y = \cos v$,

$$f(x, y) = f(\cos u, \cos v) = \varphi(u, v)$$

есть периодическая функция u, v, которая удовлетворяет условиям только что доказанной теоремы. А потому частные производные $\frac{\partial^l \varphi}{\partial u^i \partial v^{l-i}}$ существуют и удовлетворяют условиям Липшица степени α_1.

Но

$$\frac{\partial^l f}{\partial x^i \partial y^{l-i}} = \sum_{h+k \leqslant l} A_{hk} \frac{\partial^{h+k} \varphi}{\partial u^h \partial v^k},$$

где все коэффициенты A_{hk} суть вполне определенные функции x, y, которые голоморфны внутри квадрата C_1 (на сторонах квадрата они делаются бесконечными). Следовательно, внутри S все частные производные $\frac{\partial^l f}{\partial x^i \partial y^{l-i}}$ существуют и удовлетворяют условию Липшица степени α_1.

Б. *Если функция $f(x, y)$ внутри контура S, не имеющего острых углов* [1], *рассматриваемая как функция одного только x и одного только y, имеет ограниченные производные каждого порядка, то она имеет также внутри области S ограниченные частные производные любого порядка.*

Из предыдущего следствия вытекает непосредственно существование и ограниченность всех производных внутри всякой области S_1, расположенной внутри S. Чтобы показать, что производные ограничены во всякой точке M контура S, строим квадрат C_1, не выходящий из S и имеющий одну из вершин в точке M. Для определенности можно предположить снова, что квадрат C_1 составлен прямыми $x = \pm 1$, $y = \pm 1$. Разлагая функцию $f(x, y)$ в ряд тригонометрических многочленов внутри C_1 и отбрасывая члены степени выше n относительно x или

---

[1] Из доказательства будет видно, что это условие вводится для того, чтобы внутри S можно было поместить квадрат, имеющий вершину в любой точке контура S; но, заменяя прямоугольные координаты косоугольными, можно квадрат заменить ромбом; таким образом, существенно только, чтобы контур S не имел точек возврата.

относительно y, находим, на основании формул (79′) и (83), что для достаточно больших значений n ошибка

$$|R_n| < \frac{1}{n^p},$$

каково бы ни было число p. А потому наше утверждение есть прямое следствие из теоремы § 22.

75 [83]. **Теорема.** *Пусть $f(x, y)$ будет некоторая функция двух вещественных переменных (x, y), данная внутри прямоугольника C_1, образованного прямыми $x = \pm h$, $y = \pm k$. Если, при всяком вещественном x_0 $(-h \leqslant x_0 \leqslant h)$, функция $f(x_0, y)$ голоморфна относительно y и $|f(x_0, y)| < M$, когда комплексная переменная y находится внутри эллипса E, имеющего фокусы $-k$, $+k$ и полусумму осей k/ρ; и, при всяком вещественном y_0 $(-k \leqslant y_0 \leqslant k)$, функция $f(x, y_0)$ голоморфна относительно x и $|f(x, y_0)| < M$, когда комплексная переменная x находится внутри эллипса E_1, имеющего фокусы $-h$, $+h$ и полусумму осей h/ρ_1,— то функция двух переменных $f(x, y)$ голоморфна, и $|f(x, y)| < \frac{4M}{(1-\lambda)^2}$ $(\lambda < 1)$ в то время, как комплексная переменная y находится внутри эллипса E', гомофокального с E и имеющего полусумму осей k/R, а комплексная переменная x находится внутри эллипса E_1', гомофокального с E_1 и имеющего полусумму осей h/R_1, причем*

$$\frac{\log \lambda R_1}{\log \rho_1} + \frac{\log \lambda R}{\log \rho} = 1.$$

В самом деле, полагая $x = h \cos u$, $y = k \cos v$ и раскладывая функцию

$$f(x, y) = \sum_{p=0,\ q=0}^{p=\infty,\ q=\infty} A_{pq} T_p\left(\frac{x}{h}\right) T_q\left(\frac{y}{k}\right)$$

в ряд тригонометрических многочленов, мы выводим из формул (82), при помощи рассуждений § 68, что

$$|A_{pq}| < 4M\rho_1^p; \qquad |A_{pq}| < 4M\rho^q.$$

А потому, на основании неравенства (9), заключаем, что если y находится внутри эллипса E', а x находится внутри эллипса E_1', то

$$|f(x, y)| < \sum_{p=0,\ q=0}^{p=\infty,\ q=\infty} \frac{4M\rho_1^{\frac{ap}{a+b}} \rho^{\frac{bq}{a+b}}}{R_1^p R^q},$$

каковы бы ни были положительные числа a и b.

Полагая

$$\frac{\rho_1^{\frac{a}{a+b}}}{R_1} = \frac{\rho^{\frac{b}{a+b}}}{R} = \lambda,$$

получим

$$|f(x, y)| < 4M \sum_{p=0,\ q=0}^{p=\infty,\ q=\infty} \lambda^p \lambda^q = \frac{4M}{(1-\lambda)^2};$$

при этом, очевидно,

$$\frac{a}{a+b} = \frac{\log \lambda R_1}{\log \rho_1}; \qquad \frac{b}{a+b} = \frac{\log \lambda R}{\log \rho},$$

откуда

$$\frac{\log \lambda R_1}{\log \rho_1} + \frac{\log \lambda R}{\log \rho} = 1. \tag{84}$$

Следствие. *Если $\rho = \rho_1$, то*

$$\lambda^2 = \frac{\rho}{RR_1}, \tag{84'}$$

что вытекает из формулы (84), в которой полагаем $\rho = \rho_1$.

76 [84]. **Применение к уравнениям с частными производными.** Результаты предшествующих параграфов находятся в тесной связи с теорией уравнений с частными производными, и было бы интересно вывести из них систематически свойства уравнений эллиптического типа. Я ограничусь только двумя замечаниями.

1) *Уравнение эллиптического типа*

$$Ar + 2Bs + Ct = 0 \qquad (AC - B^2 > 0),$$

где A, B, C — какие угодно функции x, y, z, p, q, не имеет иных решений, периодических относительно x, y, обладающих конечными производными первых двух порядков[1]*, кроме постоянной величины.*

В самом деле, из теоремы § 72 мы знаем, что

$$\iint s^2 dx\, dy = \iint rt dx\, dy = -\iint \frac{t\,(Ct + 2Bs)}{A}\, dx\, dy,$$

откуда

$$\iint \frac{As^2 + 2Bst + Ct^2}{A}\, dx\, dy = 0,$$

а потому

$$t = s = r = 0;$$

следовательно, z есть постоянная величина.

2) *Если производные функции z, до порядка k включительно, удовлетворяют в некоторой области S какому-нибудь условию Липшица и, кроме того, функция z удовлетворяет двум уравнениям*

$$\begin{aligned} \frac{\partial^{k+1} z}{\partial x^{k+1}} &= f\left(x,\ y,\ z,\ \frac{\partial z}{\partial x},\ \ldots,\ \frac{\partial^k z}{\partial y^k}\right), \\ \frac{\partial^{k+1} z}{\partial y^{k+1}} &= \varphi\left(x,\ y,\ z,\ \frac{\partial z}{\partial x},\ \ldots,\ \frac{\partial^k z}{\partial y^k}\right), \end{aligned} \tag{85}$$

---

[1] Это вытекает также из обобщенной теоремы Лиувилля, указанной мною в «Comptes rendus» за 10 октября 1910 г.

где f *и* φ *имеют конечные производные всех порядков при конечных значениях переменных, то функция* z *имеет также конечные производные всех порядков во всякой области* S_1 *внутри* S.

Действительно, из уравнений (85) выводим непосредственно, что $\frac{\partial^{k+1}z}{\partial x^{k+1}}$ и $\frac{\partial^{k+1}z}{\partial y^{k+1}}$ существуют и удовлетворяют условию Липшица. Поэтому, на основании следствия A § 74, тем же свойством обладают все производные порядка $k+1$ во всякой области S_1' внутри S. Дифференцируя первое уравнение относительно x, а второе относительно y, мы можем то же рассуждение применить к производным $(k+2)$-го порядка; последовательное дифференцирование, оказывающееся возможным, приводит таким образом к доказательству высказанного утверждения.

АВТОРСКИЕ КОММЕНТАРИИ

3. О НАИЛУЧШЕМ ПРИБЛИЖЕНИИ НЕПРЕРЫВНЫХ ФУНКЦИЙ ПОСРЕДСТВОМ МНОГОЧЛЕНОВ ДАННОЙ СТЕПЕНИ

3.1. (К Введению.) Здесь печатается только второе из упомянутых добавлений, так как первое покрывается мемуаром из «Acta Math.» 1912 г., который печатается в этом же томе [12]. В связи с этим, начиная с главы V, нумерация формул и параграфов соответствующим образом изменена. После нового номера параграфа в квадратных скобках стоит его прежний номер.

3.2. (К Введению.) Приведем здесь перевод Введения к французскому мемуару «Sur l'ordre de la meilleure approximation des fonctions continues par des polynomes de degré donné» (Extrait des Mémoires publiés par la Classe des sciences de l'Académie royale de Belgique. Collection in 4°, Deuxiéme série, tome IV, 1912).

«Главная цель этой работы заключается в том, чтобы решить следующий вопрос, поставленный Ш. Валле Пуссеном[1]: возможно ли представить ординату некоторой ломаной линии посредством многочлена степени n с приближением лучшим, чем порядка $1/n$? Что указанный порядок приближения в самом деле может быть достигнут, установлено Ш. Валле Пуссеном в цитированном мемуаре.

Первый шаг в направлении решения проблемы недавно был сделан самим Ш. Валле Пуссеном в мемуаре[2], вышедшем в свет в то время, когда настоящая работа редактировалась. В этом последнем мемуаре уважаемый лувэнский профессор излагает общий метод нахождения приближающих многочленов и, в частности, устанавливает, что наилучшее приближение ломаной линии многочленом степени n имеет нижнюю границу порядка $\frac{1}{n(\log n)^3}$.

Настоящий мемуар дает полный ответ на поставленный вопрос:

Наилучшее приближение функции $|x|$ *в промежутке* $(-1, +1)$ *посредством многочлена степени* $2n > 0$ *заключено между* $\frac{\sqrt{2}-1}{4(2n-1)}$ и $\frac{2}{\pi(2n+1)}$.

Поиски решения проблемы Ш. Валле Пуссена естественно привели меня к постановке ряда других аналогичных вопросов и к общему исследованию зависимости между наилучшим приближением функции и ее дифференциальными свойствами. В настоящем мемуаре[3] изложена совокупность результатов, полученных мною в

---

[1] Sur la convergence des formules d'interpolation entre ordonnées équidistantes. «Bull. Acad. sci. Belgique», 1908, стр. 403.

[2] Sur les polynomes d'approximation et la représentation approchée d'un angle. Там же, 1910, стр. 808.

[3] Необходимо также назвать мемуар Д. Джексона, появившийся уже после посылки этой работы Бельгийской академии наук: D. Jackson. Über die Genauigkeit der Annäherung stetiger Funktionen durch ganze rationale Funktionen. Göttingen, Preisschrift und Innaugural-Dissertation, 1911. Он касается тех же вопросов и содержит также многочисленные библиографические указания.

этом направлении (некоторые из них, без доказательств, приведены мною в заметке «Sur l'approximation des fonctions continues par des polynomes», «Comptes rendus», t. 152, 27. II. 1911 [2])».

3.3. (К § 9.) В первоначальной редакции вместо отрывка, воспроизводящего соответствующее место из монографии «L. S.» (стр. 42), начинающегося словами: «Я утверждаю. . .» и кончающегося «в которой $S^2(\gamma) = L^2$. Этим наше утверждение доказано», стояло кратко: «подобно предыдущему находим, что $|\gamma| > |\beta|$, $|\delta| > |\beta|$ и $\gamma\beta > 0$, $\delta\beta > 0$» (причем в последних неравенствах $\gamma\beta > 0$ и $\delta\beta > 0$, вместо γ и δ, должны были стоять их вещественные части, как это сделано в нынешней редакции).

3.4. (К § 10.) Вместо неравенства $|f_n'(t)| \leqslant nL$ в первоначальной редакции в общем случае было установлено лишь, что $|f_n'(t)| < 2nL$. Приведенный здесь вывод, показывающий, что общее неравенство является элементарным следствием того же неравенства, доказанного для суммы синусов, сообщенный мне Э. Ландау вскоре после появления диссертации [3], впервые был опубликован в «L. S.», стр. 39.

3.5. (К главе II.) Применяя аналогичный прием, основанный на теореме § 6, Валле Пуссен («Leçons sur l'approximation des fonctions d'une variable réelle», 1919) несколько усилил некоторые результаты главы II. Принципиальное значение имеет данное им уточнение теоремы § 12 для случая p нецелого; а именно, если (для всех $n > 0$)

$$E_n[f(x)] < \frac{A}{n^p}, \tag{I}$$

где $k < p < k+1$ (k — целое число), то $f(x)$ не только имеет производную k-го порядка, удовлетворяющую внутри промежутка $(-1, +1)$ условию Липшица любого порядка $\alpha < p-k = b < 1$ (согласно утверждению теоремы § 12), но удовлетворяет также условию Липшица порядка $\alpha = p-k = b < 1$

$$|f^{(k)}(x_2) - f^{(k)}(x_1)| = O(|x_2 - x_1|^b) \qquad (-1 < x_1 < x_2 < 1). \tag{II}$$

В § 65 [73] настоящей работы установлено, что из условия Липшица порядка b

$$|f(x_2) - f(x_1)| = O(|x_2 - x_1|^b) \qquad (b < 1) \tag{II'}$$

вытекает

$$E_n[f(x)] < \frac{A}{n^b}, \tag{I'}$$

и одновременно Джексоном, независимо от меня, был получен общий результат, что (II) всегда влечет за собой (I). Таким образом, упомянутое выше уточнение теоремы § 12 означает, что в случае $b < 1$ свойства (I) и (II) (которые при $k = 0$ обращаются в (I') и (II') эквивалентны *.

Этот важный факт в то время не был обнаружен мною, так как меня интересовал главным образом случай *целого* p, для которого, как я показал (§§ 18—20), такой эквивалентности *не существует*.

В связи с основной задачей, которая была исходным пунктом исследования, меня особенно интересовал вопрос о том, существует ли однозначная зависимость между скоростью убывания $E_n[f(x)]$ и непрерывностью производной $f'(x)$ (подобно тому как непрерывность самой функции $f(x)$ равнозначна условию $E_n[f(x)] \to 0$). Несмотря на тесную связь (§ 18), ответ на этом вопрос оказался отрицательным и формулирован в конце § 20. На это обстоятельство особое внимание обращено также в докладе [6]: «существуют предельные случаи, когда природа непрерывности производной (которая не выражается никаким условием Липшица) так мало отличается от некоторой формы разрывности, что посредством рассмотрения наилучших

---

* См. примечания на стр. 29 и 90. Постоянная в условиях Липшица (II) и (II') при приближении точек x_1 и x_2 к концам отрезка $[-1, +1]$ неограниченно возрастает, согласно неравенствам § 12. При этом, как видно из §§ 17 и 65, соответствующая оговорка является излишней при рассмотрении наилучшего приближения $E_n^*[f(x)]$ периодической функции $f(x)$ при помощи тригонометрических сумм.

приближений $E_n[f(x)]$ (для всех n) невозможно решить, является ли эта производная непрерывной или нет. Я полагаю, что изучение этих критических случаев... могло бы способствовать более глубокому уяснению самого понятия непрерывности».

В недавней работе Зигмунда («Duke Math. Journ.», 12, 1945), после которой появились мои заметки (**245***) и (**250***), дополняющие результаты Зигмунда, идея, намеченная в приведенных словах доклада, в соединении с понятием обобщенных условий Липшица, введенным в § 24 работы [**3**], получила глубоко плодотворное развитие. Следует также отметить, что С. Б. Стечкин в своей кандидатской диссертации (см. «Докл. АН СССР», т. 65, стр. 135—137) весьма остроумно развил идеи §§ 24—26 в другом направлении, благодаря чему ему удалось существенно обобщить теорему § 25 и получить посредством рассмотрений, аналогичных проведенным в § 26, принципиально новое доказательство существования постоянных C_1 и C_2, для которых при всех n соблюдается неравенство

$$\frac{C_1}{n} < E_n[|x|] < \frac{C_2}{n}.$$

3.6. (К § 52 [60].) В первоначальном тексте (как диссертации, так и мемуара «О») в левой части неравенства (53) вследствие погрешности в вычислениях, которую отметил и исправил в своей книге Валле Пуссен, стояло $|A_{n+1}|$. Вместо примененного в настоящем издании способа определения левой части неравенства $\rho = \pm \sum\limits_{i=0}^{n+1} \frac{f(x_i)}{S'(x_i)}$ в основу вычислений была положена формула

$$\sum_{i=0}^{n+1} \frac{f(x_i)}{S'(x_i)} = \frac{1}{2\pi i} \int\limits_C \frac{f(x)}{S(x)}\,dx,$$

где C — контур, окружающий отрезок $[-1, +1]$.

Однако, нахождение нижней грани $E_n[f(x)]$, если известно только значение A_{n+1}, не лишено интереса. Нетрудно показать (при помощи других соображений), что всегда имеет место неравенство

$$\frac{\pi}{4}|A_{n+1}| \leqslant E_n[f(x)].$$

Для этого замечаем, что $E_n[f(x)]$ есть наименьшее возможное значение максимума $|f(x)|$ при $-1 \leqslant x \leqslant 1$, если в разложении $f(x) = \sum\limits_{k=0}^{\infty} A_k T_k(x)$ по многочленам $T_k(x)$ дан лишь один определенный коэффициент $A_{n+1} = \frac{2}{\pi}\int\limits_0^{\pi} f(\cos\theta)\cos(n+1)\theta\,d\theta$.

Как показано на стр. 30 монографии «Э. П.», наименьшее значение $\max\limits_{-1\leqslant x\leqslant 1}|f(x)|$ соответствует разрывной функции

$$f(x) = A_{n+1}\left[T_{n+1}(x) - \frac{T_{3(n+1)}(x)}{3} + \ldots + (-1)^k \frac{T_{(2k+1)(n+1)}(x)}{2k+1} + \ldots\right] =$$
$$= A_{n+1}\frac{\pi}{4}\operatorname{sign}[\cos(n+1)\arccos x],$$

для которой $\max\limits_{-1\leqslant x\leqslant 1}|f(x)| = \frac{\pi}{4}|A_{n+1}|$. Отсюда видно, что неравенство

$$\frac{\pi}{4}|A_{n+1}| \leqslant E_n[f(x)]$$

справедливо и не может быть улучшено.

3.7. (К § 54 [62].) В первоначальном тексте в формулировке следствия Г стояло

$$\frac{R^{n+1}}{2^n} F\left(\frac{n}{2}+1, \frac{n}{2}+\frac{3}{2}, n+2, R^2\right) < E_n[f(x)] <$$

$$< \frac{R^{n+1}}{2^n(1-R)} F\left(\frac{n}{2}+1, \frac{n}{2}+\frac{3}{2}, n+2, R^2\right),$$

так как, во-первых, не была использована формула (см. сноску на стр. 128).

$$F\left(\frac{n}{2}+1, \frac{n}{2}+\frac{3}{2}, n+2, R^2\right) = \frac{2^{n+1}}{\sqrt{1-R^2}\,(1+\sqrt{1-R^2})^{n+1}},$$

а, во-вторых, в левой части неравенства (53) стояло A_{n+1} вместо $\sum\limits_{k=0}^{\infty} A_{(2k+1)(n+1)}$.

3.8. (К § 66 [74].) Общее неравенство (78) легко получается из (77) и (77′) $(0 < p \leqslant 1)$ интегрированием. Согласно (77) и (77′), если

$$|f'(x+h) - f'(x)| < Ah^{\alpha} \qquad (0 < \alpha \leqslant 1),$$

то

$$f'(x) = S_n(x) + \varphi'(x),$$

где $|\varphi'(x)| < \frac{kA}{n^{\alpha}}$ и $S_n(x)$ — тригонометрическая сумма n-го порядка. Но в этом случае

$$|\varphi(x+h) - \varphi(x)| < \frac{Akh}{n^{\alpha}}$$

и, следовательно,

$$E_n^*[f(x)] = E_n^*[\varphi(x)] < \frac{Ak^2}{n^{1+\alpha}},$$

где $E_n^*[f(x)]$ означает наилучшее приближение при помощи тригонометрических сумм n-го порядка. Повторяя то же рассуждение, можно получить неравенство (78), если условию Липшица удовлетворяет производная любого данного порядка $f^{(p)}(x)$.

В дальнейшем нам еще придется вернуться к вопросу о нижней грани значения постоянных в неравенствах Джексона. Сейчас отметим только, что в случае $\alpha = 1$ в неравенстве (77′) можно положить* $k = \frac{\pi}{2}$, т. е.

$$E_n[f(x)] < \frac{\pi A}{2(n+1)}, \tag{I}$$

когда

$$|f(x+h) - f(x)| \leqslant Ah, \tag{II}$$

причем, как бы мало ни было $\varepsilon > 0$, неравенство (77′) может оказаться неверным для $n > n(\varepsilon)$, если положить в нем $k = \frac{\pi}{2} - \varepsilon$; для периодических функций неравенство $E_n^*[f(x)] \leqslant \frac{\pi A}{2(n+1)}$, также верное при условии (II), не может быть улучшено ни при каком n (см. Н. И. Ахиезер. Лекции по теории аппроксимации, § 87, ОГИЗ, М. — Л., 1947).

При помощи проведенного выше рассуждения мы получаем, таким образом, общие неравенства

$$E_{n+1}[f(x)] < \frac{\pi}{2(n+1)} E_n[f'(x)]; \qquad E_n^*[f(x)] \leqslant \frac{\pi}{2(n+1)} E_n^*[f'(x)], \tag{III}$$

первое из которых верно для любой функции, а второе — для периодической функции.

---

* С. М. Никольский, «Докл. АН СССР», т. 52 (1946).

3.9. (К § 69 [77].) Во французской редакции «О» содержится также формулировка следующего предложения:

Если вещественная часть функции, голоморфной внутри круга C, удовлетворяет условию Липшица степени α на окружности C, то мнимая часть удовлетворяет условию Липшица любой степени $\alpha_1 < \alpha$ на этой окружности.

При этом отмечается, что остаток в разложении в ряд Фурье мнимой части $\rho_n(\theta) = \sum\limits_{p=n+1}^{\infty} B_p \cos p\theta - A_p \sin p\theta$, так же как и остаток вещественной части $R_n(\theta) = \sum\limits_{p=n+1}^{\infty} A_p \cos p\theta + B_p \sin p\theta$, удовлетворяет неравенству вида

$$|\rho_n(\theta)| < \frac{k \log n}{n^{\alpha}}, \qquad |R_n(\theta)| < \frac{k \log n}{n^{\alpha}}.$$

3.10. (К главе VII.) Метод исследования связи между свойствами различных частных производных функций нескольких переменных, примененный в этой главе, значительно усовершенствован в моей заметке (250*), опубликованной в 1948 г. («Докл. АН СССР» т. 59, № 8).

10. ОБ ОДНОМ СВОЙСТВЕ МНОГОЧЛЕНОВ

10.1. Это неравенство $M < Le^{\rho}$, вытекающее из написанного выше равенства

$$P(\xi) = L \cos \zeta = \frac{L}{2}(e^{i\zeta} + e^{-i\zeta}),$$

где $\zeta = c \pm i\rho$, может быть заменено вытекающим отсюда же более сильным неравенством

$$M = |P(\xi)| \leqslant \frac{L}{2}[e^{\rho} + e^{-\rho}].$$

Поэтому, не изменяя ничего в дальнейшем рассуждении, вследствие которого было получено $\rho < nb$, находим вместо (9) более сильное неравенство

$$M < \frac{L}{2}[e^{nb} + e^{-nb}] = \frac{L}{2}\left[R^n + \frac{1}{R^n}\right] \tag{9'}$$

(вместо этого неравенства имеем равенство только в случае тригонометрического многочлена, когда $\beta = \gamma = \delta$, т. е. $\rho = nb$). Неравенство (9') вытекает также из соответствующего неравенства, относящегося к целым функциям конечной степени (Duffin a. Schaeffer. «Bull. Am. Math. Soc.», 1938).

11. НЕСКОЛЬКО ЗАМЕЧАНИЙ К НЕРАВЕНСТВУ ВЛАДИМИРА МАРКОВА

11.1. В 1938 г. мною дано простое доказательство неравенства В. А. Маркова. См. статью «О теореме В. А. Маркова» («Труды Лнгр. индустр. ин-та», № 5, разд. физ.-матем. наук, вып. 1, стр. 8—13 (199*), которая войдет во второй том настоящего собрания сочинений.

В том же году A. S. Schaeffer и R. J. Duffin («Bull. Am. Math. Soc.», v. 44, 1938, стр. 289—297) дали другое доказательство неравенства В. А. Маркова при помощи следующего интересного обобщения теоремы § 2 моей диссертации [3]. Если многочлен $P_n(x)$ степени $\leqslant n$ удовлетворяет на отрезке $[-1, +1]$ неравенству $|P_n(x)| \leqslant 1$, то

$$|P_n^{(k)}(x)| \leqslant |T_n^{(k)}(x) + iS_n^{(k)}(x)| \quad (-1 \leqslant x \leqslant 1,\ k = 1, 2, \ldots, n),$$