О ДОКАЗАТЕЛЬСТВЕ ТЕОРЕМЫ ВЕЙЕРШТРАССА ПОСРЕДСТВОМ ПОЛИНОМОВ БЕРНШТЕЙНА*

В. С. Виденский

ilya.viden@gmail.com

25 января 2014 г.

Videnskii V. S. *On a proof of the Weierstrass theorem by Bernstein polynomials.* In 1912 S. N. Bernstein published the paper, in which he introduced his famos polynomials $B_n f$, and found a very elegant proof of the Weierstrass approximation theorem by them. In our paper we discuss the history of this investigation.

С. Н. Бернштейн при решении 19-й проблемы Гильберта об эллиптических уравнениях с частными производными пришёл к задаче о разложении любой $f \in C[0,1]$ в некоторые двойные ряды по x^k и $(1-x)^m$ (1903 г.). В дальнейшем (1912 г.), продолжая исследование этого вопроса, С. Н. Бернштейн изобрёл свои знаменитые полиномы $B_n f$, посредством которых дал новое доказательство аппроксимационной теоремы Вейерштрасса, основанное на теории вероятностей. В нашей статье излагается история этих исследований[1].

1°. Карл Вейерштрасс открыл свою фундаментальную аппроксимационную теорему в 1885 году. В году, когда ему минуло 70 лет. Он несомненно считал это даром судьбы. В математике, как и в музыке и в поэзии, крупные творческие удачи, как правило, достаются молодым, энергичным и дерзновенным.

---

*Семинар по дискретному гармоническому анализу и геометрическому моделированию «DHA & CAGD»: `http://www.dha.spb.ru/`

[1]Этот текст соответствует моему докладу, прочтенному 7 ноября 2013 года на семинаре по истории математики в ПОМИ. Статья представляет собой некоторую вариацию прошлогодней публикации [6].

В большинстве книг по теории приближения функций приводится несколько различных доказательств теоремы Вейерштрасса, данных разными авторами. Статья Вейерштрасса [1] была написана в его обычном стиле — понятно и отчетливо. Наметим её план в общих чертах. Если $f \in C[a,b]$, то она продолжается на всю вещественную ось равенствами $f(x) = f(a)$ при $x \leqslant a$, $f(x) = f(b)$ при $x \geqslant b$. Затем вводится свёртка

$$W_n(f;x) = \sqrt{\frac{2\pi}{n}} \int_{-\infty}^{+\infty} f(t) e^{-n(t-x)^2} dt. \tag{1.1}$$

При достаточно больших n частные суммы рядов

$$W_n(f;x) = \sqrt{\frac{2\pi}{n}} \sum_{k=0}^{\infty} \int_{-\infty}^{+\infty} f(t)(-1)^k \frac{n^k (t-x)^{2k}}{k!} dt \tag{1.2}$$

и есть приближающие полиномы.

В статье [1] была приведена также вторая теорема о приближении непрерывных 2π-периодических функций тригонометрическими полиномами. Её доказательство было сложнее, чем первой. Впрочем, в дальнейшем выяснилось, что обе теоремы эквивалентны — каждая из них выводится из другой заменой переменной.

В конце своей статьи Вейерштрасс выразил надежду, что молодые математики заинтересуются этой темой и найдут явные формулы для приближающих полиномов.

Среди тех, кто откликнулся на этот призыв были такие выдающиеся математики, как А. Лебег, Л. Фейер, Ш. Валле Пуссен, Э. Ландау и С. Бернштейн. Они дали ряд новых оригинальных доказательств аппроксимационной теоремы Вейерштрасса и вывели различные явные формулы для приближающих полиномов.

Интересно, что по аналогии с операторами Вейерштрасса W_n нетрудно построить аппроксимирующие операторы, которые будут полиномами, то есть искомыми явными формулами. Это заметили в 1908 г. почти одновременно и независимо Ш. Валле Пуссен и Э. Ландау. Оказалось, что вместо функции

$$e^{-n(t-x)^2} = \left(1 - (t-x)^2 + \sum_{k=0}^{\infty} (-1)^k \frac{(t-x)^{2k}}{k!}\right)^n \tag{1.3}$$

достаточно взять в качестве ядра свёртки не весь ряд (1.3), а только функцию

$$\big(1 - (t-x)^2\big)^n \tag{1.4}$$

и построить операторы по формуле

$$L_n(f;x) = k_n \int_a^b f(t) \big(1 - (t-x)^2\big)^n dt. \tag{1.5}$$

2°. Возможно, что кое-кто из опубликовавших новые доказательства аппроксимационной теоремы Вейерштрасса нашли их не специально, но попутно в ходе собственных исследований, так или иначе связанных с приближением функций. В частности, так случилось с С. Н. Бернштейном, который изобрёл свои знаменитые полиномы $B_n f$, работая над 19-й проблемой Гильберта о структуре решений дифференциальных уравнений с частными производными эллиптического типа. Мы остановимся на подробностях в п. 3 и в п. 5.

Для функции $f \in C[0,1]$ полиномы Бернштейна определяются так:

$$B_n\bigl(f(t);x\bigr) = \sum_{k=0}^{n} f(\tfrac{k}{n})p_{nk}(x), \quad 0 \leqslant x \leqslant 1. \tag{2.1}$$

$$p_{nk}(x) = C_n^k x^k (1-x)^{n-k}. \tag{2.2}$$

В 1912 г. в «Сообщ. Харьк. матем. об-ва» С. Н. Бернштейн [2] в очень краткой заметке «Доказательство теоремы Вейерштрасса, основанное на теории вероятностей» опубликовал такую теорему.

ТЕОРЕМА 1. *Если $f \in C[0,1]$, то последовательность полиномов $B_n f$ при $n \to \infty$ сходится к f равномерно на $[0,1]$.*

Для доказательства применяются издавна известные в теории вероятностей тождества

$$\sum_{k=0}^{n} p_{nk}(x) = \sum_{k=0}^{n} C_n^k x^k (1-x)^{n-k} = \bigl(x + (1-x)\bigr)^n = 1, \tag{2.3}$$

$$\sum_{k=0}^{n} \frac{k}{n} p_{nk}(x) = x \sum_{k=0}^{n-1} C_{n-1}^k x^k (1-x)^{n-1-k} = x, \tag{2.4}$$

и тождество, которое из них выводится дифференцированием,

$$\sum_{k=0}^{n} \left(\frac{k}{n} - x\right)^2 p_{nk}(x) = \frac{x(1-x)}{n}. \tag{2.5}$$

Действительно, пусть $0 < x < 1$. Имеем

$$p'_{nk}(x) = \frac{n}{x(1-x)} \left(\frac{k}{n} - x\right) p_{nk}(x). \tag{2.6}$$

В силу (2.3) и (2.4)

$$\sum_{k=0}^{n} \left(\frac{k}{n} - x\right) p_{nk}(x) = 0. \tag{2.7}$$

Дифференцируя (2.7), получаем (2.5).

Затем применяется неравенство Чебышёва, которое состоит в следующем: Пусть $0 < \delta < 1$; тогда, благодаря (2.5) имеем

$$\begin{gathered}\sum_{|\frac{k}{n}-x|\geqslant\delta} p_{nk}(x) \leqslant \frac{1}{\delta^2}\sum_{|\frac{k}{n}-x|\geqslant\delta}\left(\frac{k}{n}-x\right)^2 p_{nk}(x) \leqslant \\ \leqslant \frac{1}{\delta^2}\sum_{k=0}^{n}\left(\frac{k}{n}-x\right)^2 p_{nk}(x) = \frac{x(1-x)}{\delta^2 n} \leqslant \frac{1}{4n\delta^2}.\end{gathered} \tag{2.8}$$

Теорема 1 следует из (2.8). Пусть $f \in C[0,1]$; $\bigl|f(x)\bigr| \leqslant M$. Фиксируем $\varepsilon > 0$ $(0 < \varepsilon < 1)$ и выбираем $\delta(0 < \delta < 1)$ по условию, что при $|x' - x''| < \delta$ выполняется неравенство

$$\bigl|f(x') - f(x'')\bigr| < \frac{\varepsilon}{2}. \tag{2.9}$$

Покажем, что при $n > M/\varepsilon\delta^2$ справедливо неравенство

$$\Bigl|B_n\bigl(f(t);x\bigr) - f(x)\Bigr| < \varepsilon, \quad 0 \leqslant x \leqslant 1. \tag{2.10}$$

Действительно, учитывая (2.8) и (2.9), имеем

$$\begin{gathered}\Bigl|B_n\bigl(f(t);x\bigr) - f(x)B_n(1;x)\Bigr| \leqslant \sum_{k=0}^{n}\bigl|f(\tfrac{k}{n}) - f(x)\bigr|p_{nk}(x) = \\ = \sum_{|\frac{k}{n}-x|\geqslant\delta}\bigl|f(\tfrac{k}{n}) - f(x)\bigr|p_{nk}(x) + \sum_{|\frac{k}{n}-x|<\delta}\bigl|f(\tfrac{k}{n}) - f(x)\bigr|p_{nk}(x) \leqslant \frac{2M}{4n\delta^2} + \frac{\varepsilon}{2} < \varepsilon.\end{gathered}$$

Как мы видим, это доказательство не использует непосредственно терминологию теории вероятностей. Однако, оно использует по существу и в деталях идеи С. Н. Бернштейна. Впервые в такой форме оно было приведено в задачнике Г. Полка и Г. Сегё. Теперь эта манера изложения стала обычной.

Бесспорно, это прекрасное, простое и ясное доказательство. Вместе с тем абсолютно загадочно, как оно могло быть изобретено. В заметке [2] С. Н. Бернштейн не даёт никакого намёка на путь, который привёл его к полиномам $B_n f$.

В 1950 году С. Н. Бернштейну минуло 70 лет. По инициативе академиков П. С. Александрова, А. Н. Колмагорова и И. Г. Петровского, поддержаной президентом Академии наук С. Н. Вавиловым, Академия наук постановила издать Собрание сочинений С. Н. Бернштейна под его редакцией.

Первый том сочинений вышел в свет весной 1952 г. Заметка [2], таким образом, была напечатана вторично, спустя ровно 40 лет. Редактор этого тома профессор Д. А. Райков рискнул спросить С. Н. Бернштейна, какова история изобретения таких оригинальных полиномов $B_n f$. Но Сергей Натанович не ответил по существу, а просто отшутился. Он сказал, что частые студенческие забастовки в 1911 году предоставили достаточно времени для математических размышлений.

3°. В каждой шутке есть доля правды. Попробуем эту крупицу истины уловить в нашем случае. Мы увидим, что изобретение полиномов Бернштейна было результатом не мгновенного вдохновения, а что путь к ним был многолетним и нелёгким. Однако, разумеется, в конце пути не обошлось без вдохновения и глубокой интуиции.

Окончив в Париже университет (Сорбонна), С. Н. Бернштейн переехал в Геттинген. В поисках темы для самостоятельной научной работы он начал посещать прославленный семинар Гильберта и по личному его совету занялся 19-й проблемой. Вот её содержание. Рассматривается дифференциальное уравнение с частными производными второго порядка с двумя независимыми переменными

$$F(r,\ s,\ t,\ p,\ q,\ u,\ x,\ y) = 0, \tag{3.1}$$

где функция $u = u(x, y)$ — искомое решение уравнения. Уравнение (3.1) называется эллиптическим, если

$$4F'_r F'_t - (F'_s)^2 > 0. \tag{3.2}$$

В 19-й проблеме Гильберта предлагается доказать, что, если F аналитическая функция всех восьми переменных и если решение $u = u(x, y)$ имеет непрерывными все частные производные второго порядка, то функция u является аналитической.

Уравнение Лапласа

$$r + t = 0 \tag{3.3}$$

является частным случаем уравнения (3.1). Ещё Коши, применяя метод последовательных приближений, доказал аналитичность гармонических функций, то есть решений уравнения (3.3). Этот метод стал ведущим частным приёмом при исследовании решений общего уравнения (3.1). Сначала в 1890 г. этот способ применил Пикар к уравнению

$$r + t + ap + bq + cu = 0, \tag{3.4}$$

где $a,\ b,\ c$ аналитические функции переменных $x,\ y$, и доказал, что все решения будут аналитическими при начальных условиях, что все частные производные функции u второго порядка непрерывны.

С. Н. Бернштейн по примеру своих знаменитых предшественников начал исследование проблем Гильберта для уравнения (3.1) применением метода последовательных приближений. Повидимому, надеясь, с одной стороны, на плодотворность метода, но, с другой стороны, бесспорно не сомневаясь, что на этом пути встретятся большие трудности.

Действительно, препятствия незамедлительно появились. Переход от декартовых координат к полярным координатам в случае уравнений (3.3) и (3.4) приводил к степенным рядам, то есть к вопросам известным и хорошо развитым. А в общем случае уравнения (3.1) та же процедура привела не к степенным рядам, а двойным рядам

$$f(x) = \sum_{l,m} a_{lm} x^l (1-x)^m, \quad 0 \leqslant x \leqslant 1. \tag{3.5}$$

С. Н. Бернштейн называл ряд (3.5) нормальным, если он сходился абсолютно и равномерно. В исследованиях С. Н. Бернштейна нормальные ряды играют существенную роль в решении 19-й проблемы Гильберта.

Было естественно требовать в качестве начальных условий чтобы решение u уравнения (3.1) имело непрерывными все вторые частные производные. Для этого достаточно было доказать, что любая $f \in C[0,1]$ раскладывается в нормальный ряд. Но долгие годы это установить не удавалось, и приходилось считать функцию u более гладкой.

В п. 5 будет доказано при помощи полиномов Бернштейна, что любая $f \in C[0,1]$ раскладывается в нормальный ряд.

Среди частных сумм ряда (3.5) имеются такие

$$\sum_{l+m=n} a_{lm} x^l (1-x)^m. \tag{3.6}$$

С некоторой натяжкой их можно было бы считать дальними предшественниками полиномов Бернштейна.

С. Н. Бернштейн в 1908 г. в харьковской магистерской диссертации [3] доказал, что дважды непрерывно дифференцируемая функция f разлагается в нормальный ряд. Для этого он применил такую формулу

$$f(x) = \frac{1}{2}\int_0^1 f''(t)|x-t|dt + A + Bx, \quad 0 \leqslant x \leqslant 1, \tag{3.7}$$

которая легко проверяется интегрированием по частям. Затем $|x-t|$ представляется в виде

$$|x-t| = \Big(1 - \big(1-(x-t)^2\big)\Big)^{\frac{1}{2}} = (1+t^2)^{\frac{1}{2}} \left(1 - \frac{1-x^2+2xt}{1+t^2}\right)^{\frac{1}{2}}, \tag{3.8}$$

раскладывается в степенной ряд и подставляется в интеграл (3.7). То, что почленное интегрирование даёт нормальный ряд, требует ещё оценки полученных коэффициентов.

4°. Великие математики П. Л. Чебышёв (1821–1894) и К. Вейерштрасс (1815–1897) были современниками. Каждый из них внёс выдающийся вклад в проблему приближения непрерывных на отрезке функций. Чебышёв выдвинул и развил идею наилучшего приближения данной функции полиномами фиксированной степени n (неявно предполагая, что результат будет принципиально лучше, если n выбрать достаточно большим). А Вейерштрасс установил общий факт, что любая $f \in C[a, b]$ может быть равномерно приближена как угодно хорошо полиномом достаточно высокой степени. Интересно, что ни Вейерштрасс, ни Чебышёв не сделали попытки привести обе идеи в соприкосновение.

Первый шаг к сближению обоих кругов мысли сделал в 1908 г. Ш. Валле Пуссен. Он исследовал наилучшее приближение функции $|x|$ на отрезке $[-1, 1]$ полиномами степени n. Обозначим это наилучшее приближение через $E_n|x|$. Он получил неравенства

$$\frac{A}{n \ln^3 n} < E_n|x| < \frac{B}{n}, \tag{4.1}$$

в которых нижняя и верхняя границы разных порядков. Год спустя, А. Лебег улучшил оценку снизу:

$$\frac{C}{n \ln n} < E_n|x|. \tag{4.2}$$

По инициативе Валле Пуссена в 1910 г. Бельгийская академия наук объявила конкурс на тему о наилучшем приближении функции $|x|$.

С. Н. Бернштейн со свойственной ему сосредоточенностью включился в конкурс. Он получил неравенства

$$\frac{\alpha}{n} < E_n|x| < \frac{\beta}{n} \tag{4.3}$$

и, кроме того, доказал, что существует предел

$$\lim_{n\to\infty} nE_n|x| = \mu, \quad 0,278 < \mu < 0,282. \tag{4.4}$$

В процессе этих исследований он глубоко объединил идеи и методы Чебышёва и Вейерштрасса и построил теорию, которую назвал конструктивной теорией функций, где приведена классификация непрерывных функций f по скорости стремления $E_n f$ к нулю при $n \to \infty$.

Работа была удостоена премии Бельгийской академии наук. С. Н. Бернштейн был приглашён сделать часовой доклад на Международном конгрессе в Кембридже в 1912 г. В этом докладе [4] С. Н. Бернштейн так выразил своё отношение к роли удачно выдвинутых конкретных проблем в развитии математики: «Пример задачи о наилучшем приближении функции $|x|$, предложенный Валле Пуссеном, даёт ещё одно подтверждение того факта, что хорошо поставленный частный вопрос способен быть отправной точкой для далеко идущих теорий».

5°. Исследования С. Н. Бернштейна по теории приближения функций вновь привлекли его внимание к вопросу, возможно ли любую $f \in C[0,1]$ разложить в нормальный ряд. Теперь попытка решить эту задачу удалась блестяще. Были изобретены полиномы Бернштейна (2.1) и установлены их аппроксимационные свойства, что давало положительный ответ [5] (Добавление к главе 5 харьковской докторской диссертации).

К построению полиномов Бернштейна привела простая и плодотворная мысль. Так как система функций $\left\{x^k(1-x)^{n-k}\right\}_{k=0}^n$ линейно независима, то она образует базис в пространстве полиномов степени n, и любой полином

$$P_m(x) = \sum_{\nu=0}^{m} \alpha_\nu x^\nu, \quad m \leqslant n, \tag{5.1}$$

можно однозначно представить в виде

$$P_m(x) = \sum_{\nu=0}^{m} \alpha_\nu x^\nu \big((1-x)+x\big)^{n-\nu} = \sum_{k=0}^{n} A_{nk} x^k (1-x)^{n-k}. \tag{5.2}$$

Напишем для x^ν отдельно формулу (5.2):

$$\begin{aligned} x^\nu = x^\nu\big((1-x)+x\big)^{n-\nu} &= \sum_{k=\nu}^{n} C_{n-\nu}^{k-\nu} x^k (1-x)^{n-k} = \\ = \sum_{k=\nu}^{n} \frac{C_{n-\nu}^{k-\nu}}{C_n^k} C_n^k x^k (1-x)^{n-k} &= \sum_{k=\nu}^{n} \frac{C_{n-\nu}^{k-\nu}}{C_n^k} p_{nk}(x), \end{aligned} \tag{5.3}$$

где p_{nk} — полином (2.2). Легко вычислить, что

$$\frac{C_{n-\nu}^{k-\nu}}{C_n^k} = \frac{k(k-1)\dots(k-\nu+1)}{n(n-1)\dots(n-\nu+1)} = (\tfrac{k}{n})^\nu \frac{(1-\frac{1}{k})\dots(1-\frac{\nu-1}{k})}{(1-\frac{1}{n})\dots(1-\frac{\nu-1}{n})} = (\tfrac{k}{n})^\nu \alpha_{nk}^{(\nu)}, \tag{5.4}$$

$$\lim_{k\to\infty} \alpha_{nk}^{(\nu)} = 1. \tag{5.5}$$

Учитывая формулы (5.3)–(5.5), С. Н. Бернштейн по интуиции построил полиномы

$$B_n(t^\nu; x) = \sum_{k=1}^{n} \left(\frac{k}{n}\right)^\nu p_{nk}(x). \tag{5.6}$$

С. Н. Бернштейн путём тонких вычислений оценивает почленно разности между x^ν и $B_n t^\nu$ и получает оценки

$$\begin{gathered} \left|B_n(t^\nu; x) - x^\nu\right| \leqslant \frac{K_\nu}{n}, \quad \nu \leqslant n, \\ \Big|B_n\big(P_m(t); x\big) - P_m(x)\Big| \leqslant \frac{K_m}{n}, \quad m \leqslant n, \end{gathered} \tag{5.7}$$

где K_ν — константы, которые зависят только от ν. Теперь для вывода (5.7) можно обойтись без трудных вычислений. Достаточно воспользоваться выведенной мною недавно рекурентной формулой

$$B_n(t^{\nu+1};x) = B_n(t^\nu;x) - \left(1-\tfrac{1}{n}\right)^\nu (1-x)B_{n-1}(t^\nu;x)$$

и применить индукцию.

С. Н. Бернштейн из неравенств (5.7), считая аппроксимационную теорему Вейерштрасса известной, извлёк два выжных следствия.

ТЕОРЕМА 2. *Любая функция $f \in C[0,1]$ разлагается в нормальный ряд.*

Действительно, для фиксированной функции $f \in C[0,1]$ по теореме Вейерштрасса существует последовательность полиномов P_{m_s} степени m_s, такая, что

$$\left|f(x) - P_{m_s}(x)\right| < \frac{1}{2^{s+2}}, \quad 0 \leqslant x \leqslant 1, \tag{5.8}$$

а по (5.7) существует последовательность полиномов B_{n_s}, такая, что

$$\left|B_{n_s}(P_{m_s};x) - P_{m_s}(x)\right| < \frac{1}{2^{s+2}}. \tag{5.9}$$

Тогда

$$\left|B_{n_s}(P_{m_s};x) - f(x)\right| < \frac{1}{2^{s+1}},$$
$$\left|B_{n_{s+1}}(P_{m_{s+1}};x) - B_{n_s}(P_{m_s};x)\right| < \frac{1}{2^s},$$

и следовательно функция f разлагается в нормальный ряд

$$f(x) = \sum \bigl(B_{n_{s+1}}(P_{m_{s+1}};x) - B_{n_s}(P_{m_s};x)\bigr). \tag{5.10}$$

ТЕОРЕМА 3. *Из теоремы Вейерштрасса и неравенств* (5.7) *следует, что*

$$\lim_{n\to\infty} B_n\bigl(f(t),x\bigr) = f(x), \quad 0 \leqslant x \leqslant 1. \tag{5.11}$$

Из (5.8) и (5.9) имеем

$$\begin{aligned}\left|B_n(f;x) - f(x)\right| \leqslant \left|B_n(f;x) - B_n(P_{m_s};x)\right| + \left|B_n(P_{m_s};x) - P_{m_s}(x)\right| + \\ + \left|P_{m_s}(x) - f(x)\right|,\end{aligned}$$

$$\left|B_n(f;x) - B_n(P_{m_s};x)\right| \leqslant \sum_{k=0}^{n} \left|f(\tfrac{k}{n}) - P_{m_s}(\tfrac{k}{n})\right| p_{nk}(x) < \frac{1}{2^{s+2}},$$

$$\left|B_n(P_{m_s};x) - P_{m_s}(x)\right| < \frac{K_{m_s}}{n}, \quad m_s < n.$$

И (5.11) доказано.

Изложенное здесь Добавление С. Н. Бернштейн кончает фразой: «Формула (5.11) выведена мною при помощи теории вероятностей в заметке [2], 1912 г». Указаний по истории вопроса и литературных ссылок в заметке [2] нет.

Полиномы Бернштейна не сразу привлекли внимание математиков. Но в дальнейшем их стали активно исследовать, обобщать и применять. Методы заметки [2] стали основой для построения других аппроксимирующих линейных положительных операторов. Список трудов по этим темам превышает 1500 названий.

ЛИТЕРАТУРА

1. Weierstrass K. *Uber die analytische Darstellbarkeit sogenannter willkurlicher Functionen einer Argumente* (1885). Werke, Bd. 3.

2. Бернштейн С. Н. *Доказательство теоремы Вейерштрасса, основанное на теории вероятностей* (1912). Сочинения, Т. 1, статья № 4.

3. Бернштейн С. Н. *Исследование и интегрирование дифференциальных уравнений с частными производными второго порядка эллиптического типа* (1908). Сочинения, Т. 3, статья № 9, стр. 48–50.

4. Бернштейн С. Н. *О новых исследованиях, относящихся к наилучшему приближению непрерывных функций многочленами* (1912). Сочинения, Т. 1, статья № 6.

5. Бернштейн С. Н. *О наилучшем приближении непрерывных функций посредством многочленов данной степени* (1912). Сочинения, Т. 1, статья № 3, стр. 79–84.

6. Виденский В. С. *К столетию открытия полиномов Бернштейна* // Некоторые актуальные проблемы современной математики и математического образования. Герценовские чтения–2005. СПб. РГПУ. 2013. С. 5–11.