К СТОЛЕТИЮ ОТКРЫТИЯ ПОЛИНОМОВ БЕРНШТЕЙНА*

В. С. Виденский

ilya.viden@gmail.com

14 марта 2013 г.

Videnskii V. S. The centenary of the discovery of Bernstein polynomials. In 1912 S. Bernstein found a very elegant explicit representation for approximation polynomials $B_n f$. In our paper we discuss the history of this discovery.

В 1912 г. С. Н. Бернштейн опубликовал конструкцию своих знаменитых полиномов и при помощи соображений из теории вероятностей доказал теорему Вейерштрасса о приближении непрерывных функций. В настоящей статье сделана попытка проследить почти десятилетний путь к этому открытию.

В феврале 2012 г. в СПбГУ состоялась международная конференция, посвящённая 100-летию со дня рождения Л. В. Канторовича. В частности, я делал доклад о его работах по полиномам Бернштейна. Тогда же я заметил, что эти полиномы были открыты ровно 100 лет тому назад. Мне показалось, впрочем, что не сто́ит отвлекать внимание слушателей и смешивать обе темы, а лучше обсудить вторую из них отдельно.

Эти полиномы были введены в очень краткой заметке С. Н. Бернштейна “Доказательство теоремы Вейерштрасса, основанное на теории вероятностей”. Как известно, они определяются так:

$$B_n\big(f(t);x\big) = \sum_{k=0}^{n} f\left(\tfrac{k}{n}\right) p_{nk}(x), \tag{1}$$

$$p_{nk}(x) = C_n^k\, x^k (1-x)^{n-k}. \tag{2}$$

Опираясь на эти полиномы, С. Н. Бернштейн доказал теорему Вейерштрасса в такой форме.

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

ТЕОРЕМА 1. *Если f — функция, непрерывная на отрезке $[0,1]$, то последовательность полиномов $B_n f$ при $n \to \infty$ сходится к ней равномерно на $[0,1]$.*

Доказательство получилось простое, понять его легко. По биному Ньютона имеем
$$\sum_{k=0}^{n} p_{nk}(x) = 1; \qquad \sum_{k=0}^{n} \tfrac{k}{n}\, p_{nk}(x) = x, \tag{3}$$
Из (3) легко выводится тождество
$$\sum_{k=0}^{n} \bigl(\tfrac{k}{n} - x\bigr)^2 p_{nk}(x) = \frac{x(1-x)}{n}. \tag{4}$$
Остаётся воспользоваться неравенством Чебышева, применяемым для вывода закона больших чисел. □

Рассуждение ясно и красиво, однако, оно совершенно неожиданно — как будто с неба свалилось.

К концу марта или началу апреля 1952 г. первый том Собрания сочинений С. Н. Бернштейна был вполне готов, и наступило время подписать его к печати. В какой-то момент редактор издательства профессор Д. А. Райков, по-видимому, чтобы заполнить паузу, возникшую в деловой беседе, спросил Сергея Натановича, как ему удалось построить такие оригинальные полиномы для доказательства теоремы Вейерштрасса. Сергей Натанович на мгновение задумался, а затем сухо и иронично ответил, что в 1911 году имело место такое прогрессивное явление, как студенческие забастовки, благодаря чему у него оставалось довольно много времени для размышлений. Разумеется, мы ожидали, что Сергей Натанович войдёт в какие-нибудь интересные подробности или расскажет о каком-нибудь поразительном озарении. Но он не открыл нам никакой великой тайны, — быть может, её и не было.

Во всяком случае, С. Н. Бернштейн начал думать и публиковать статьи о приближении непрерывных функций рядами, по форме сходными с полиномами (1), ещё в 1903 г., когда приступил к решению 19-й проблемы Гильберта. Мы попытаемся восстановить в общих чертах его дорогу к полиномам, которые носят его имя и которые привели к принципиально новому доказательству теоремы Вейерштрасса.

Девятнадцатая проблема Гильберта относится к дифференциальным уравнениям с частными производными второго порядка эллиптического типа, задаваемых аналитическими функциями, с двумя независимыми переменными. Иными словами, речь идёт об уравнении
$$F(r, s, t, p, q, u, x, y) = 0, \tag{5}$$

где F — аналитическая функция всех восьми переменных, и выполняется неравенство

$$4F'_r F'_t - (F'_s)^2 > 0. \tag{6}$$

Требовалось доказать, что все решения этого уравнения являются аналитическими функциями, причём предполагается, что решение уравнения u непрерывно дифференцируемо σ раз, $\sigma \geqslant 2$. Уравнение Лапласа было рассмотрено ещё Коши, а случай линейных уравнений — Пикаром. Оба они применяли метод последовательных приближений, что при переходе от декартовых координат к полярным приводило к некоторым степенным рядам. Лично Гильберт предложил молодому Бернштейну заняться проблемой в общем случае. Естественно было начать с метода последовательных приближений и дальнейшего перехода к полярным координатам. Однако оказалось, что это приводит не к степенным рядам, а к более сложным двойным рядам вида

$$f(x) = \sum_{l,m} a_{lm}\, x^l (1-x)^m.$$

Если они сходятся абсолютно и равномерно, то С. Н. Бернштейн назвал их нормальными рядами.

Число σ непрерывных производных от решения u уравнения (5), которое нужно потребовать первоначально, зависит от класса функций f, разлагаемых в нормальные ряды. Если бы оказалось, что любая непрерывная на отрезке $[0,1]$ функция f представима суммой нормального ряда, то было бы достаточно требовать от решения u, чтобы его вторые частные производные удовлетворяли условию Липшица с $\alpha > 0.5$.

Но представимость любой непрерывной функции f нормальным рядом удалось доказать не сразу. Сначала приходилось требовать от функции f большей гладкости. Приведём рассуждение для дважды непрерывно дифференцируемой функции, данное в магистерской диссертации (1908 г., Сочинения, т. 3, стр. 49). Проверим, что справедливо тождество

$$f(x) = \frac{1}{2}\int_0^1 f''(t)|x-t|\,dt + A + Bx, \qquad 0 \leqslant x \leqslant 1. \tag{7}$$

Действительно, если мы напишем

$$\frac{1}{2}\int_0^1 f''(t)|x-t|\,dt = \tfrac{1}{2}\int_0^x f''(t)(x-t)\,dt - \tfrac{1}{2}\int_x^1 f''(t)(x-t)\,dt$$

и проинтегрируем по частям, то получим

$$\frac{1}{2}\int\limits_0^1 f''(t)|x-t|\,dt = f(x) - \tfrac{1}{2}\bigl(f(1)+f(0)-f'(1)\bigr) - \tfrac{1}{2}\bigl(f'(1)+f'(0)\bigr)x.$$

Разложим функцию $|x-t|$ при $0 \leqslant x \leqslant 1$ в ряд

$$\begin{aligned}|x-t| &= \bigl(1-\bigl(1-(x-t)^2\bigr)\bigr)^{1/2} = \\ &= \bigl(1+t^2-(1-x^2+2tx)\bigr)^{1/2} = (1+t^2)^{1/2}\Bigl(1-\frac{1-x^2+2tx}{1+t^2}\Bigr)^{1/2} = \\ &= (1+t^2)^{1/2}\Bigl(1-\frac{1}{2}\,\frac{1-x^2+2tx}{1+t^2} - \sum_{k=2}^{\infty}\frac{(2k-3)!!}{k!\,2^k}\,\frac{(1-x^2+2tx)^k}{(1+t^2)^k}\Bigr).\end{aligned} \tag{8}$$

Подставляя (8) в (7), получим

$$f(x) = a + bx + cx^2 + \sum_{p,q} A_{pq}\,x^p(1-x^2)^q. \tag{9}$$

Ясно, что (9) — нормальный ряд, так как

$$(1-x^2)^q = (1-x)^q \sum_{k=0}^{q} C_q^k\, x^k.$$

Затем С. Н. Бернштейн показывает, что приведенный результат остаётся справедливым для любой непрерывно дифференцируемой (только один раз) функции.

Вернёмся, однако, к плодотворному 1911 г., как сказано в ответе Д. А. Райкову. Как раз в 1910 г. академия наук Бельгии объявила конкурс по теме, предложенной Валле Пуссеном, — о наилучшем приближении функции $f(x)=|x|$ на отрезке $[-1,+1]$ полиномами степени $\leqslant n$. Эта задача была в стиле проблем Гильберта: просто формулируется; не кажется ни слишком лёгкой, ни совершенно неприступной; можно надеяться, что её решение будет полезным также для исследования иных тем.

С. Н. Бернштейн с большой скоростью — к лету 1911 года — дал полное решение поставленной проблемы. Для этой цели он создал общую теорию наилучшего приближения непрерывных функций многочленами, приведя во взаимодействие идеи и методы Вейерштрасса и Чебышева. Работа была премирована бельгийской академией наук.

На Международном конгрессе математиков в Кембридже в 1912 г. С. Н. Бернштейн в своём докладе сказал: "Пример задачи о наилучшем приближении $f(x)=|x|$, предложенный Валле Пуссеном, даёт ещё одно подтверждение того факта, что хорошо поставленный частный вопрос способен быть отправной точкой для далеко идущих теорий".

Эти исследования С. Н. Бернштейна составили также его докторскую диссертацию, защищённую в Харьковском университете (Сочинения, т. 1, № 3). В неё было включено Добавление, в котором доказывалось, что любая непрерывная на $[0,1]$ функция f представима нормальным рядом. В основу была положена простая идея. Так как система функций $\{x^k(1-x)^{n-k}\}_{k=0}^n$ линейно независима, то любой многочлен

$$P_m(x)=\sum_{\nu=0}^m a_\nu x^\nu$$

степени $m\leqslant n$ можно однозначно представить в виде

$$P_m(x)=\sum_{\nu=0}^m a_\nu\, x^\nu\bigl((1-x)+x\bigr)^{n-\nu}=\sum_{k=0}^n A_{nk}\, x^k(1-x)^{n-k}. \tag{10}$$

У С. Н. Бернштейна явно вычислены A_{nk} через a_ν, но проще рассматривать отдельно

$$x^\nu=x^\nu\bigl((1-x)+x\bigr)^{n-\nu}=\sum_{k=\nu}^n C_{n-\nu}^{k-\nu}\, x^k(1-x)^{n-k}. \tag{11}$$

Для применения к (10) удобно формулу (11) переписать по полиномам p_{nk} (2):

$$x^\nu=\sum_{k=\nu}^n \frac{C_{n-\nu}^{k-\nu}}{C_n^k}\, p_{nk}(x). \tag{12}$$

Мы имеем

$$\frac{C_{n-\nu}^{k-\nu}}{C_n^k}=\frac{k(k-1)\cdots(k-\nu+1)}{n(n-1)\cdots(n-\nu+1)}=\bigl(\tfrac{k}{n}\bigr)^\nu\,\frac{\bigl(1-\frac{1}{k}\bigr)\cdots\bigl(1-\frac{\nu-1}{k}\bigr)}{\bigl(1-\frac{1}{n}\bigr)\cdots\bigl(1-\frac{\nu-1}{n}\bigr)}. \tag{13}$$

Учитывая формулы (12) и (13) и применяя обозначения, принятые в наше время, можем написать приближающий полином для x^ν так:

$$B_n(t^\nu;x)=\sum_{k=0}^n \bigl(\tfrac{k}{n}\bigr)^\nu p_{nk}(x). \tag{14}$$

При помощи (13) и достаточно тонких вычислений С. Н. Бернштейн получает такую оценку разности

$$\bigl|x^\nu-B_n(t^\nu;x)\bigr|<\frac{K_\nu}{n},\qquad \nu\leqslant m,$$

где K_ν зависит только от ν. Отсюда для многочлена P_m (10) имеем

$$B\bigl(P_m(t);x\bigr)=\sum_{k=0}^{n}P_m\bigl(\tfrac{k}{n}\bigr)\,p_{nk}(x),$$

$$\bigl|B_n(P_m(t);x)-P_m(x)\bigr|<\frac{K_m}{n},\qquad m\leqslant n. \tag{15}$$

Заметим, между прочим, что неравенство (15) можно получить почти без вычислений и без применения неравенства Чебышева. Воспользуемся для этого одним тождеством, которое я недавно вывел для других целей. $B_n(1;x)$ и $B_n(t;x)$ даются формулами (3). Покажем, что при $\nu\geqslant 1$ справедливо тождество

$$B_n(t^{\nu+1};x)=B_n(t^\nu;x)-\bigl(1-\tfrac{1}{n}\bigr)^\nu(1-x)B_{n-1}(t^\nu;x). \tag{$*$}$$

Напишем явно

$$B_n(t^{\nu+1};x)=\sum_{k=1}^{n}\bigl(\tfrac{k}{n}\bigr)^{\nu+1}C_n^k\,x^k(1-x)^{n-k}$$

и преобразуем общий член, стоящий под знаком суммы

$$\begin{aligned}\bigl(\tfrac{k}{n}\bigr)^{\nu+1}C_n^k\,x^k(1-x)^{n-k}&=\bigl(\tfrac{k}{n}\bigr)^{\nu}\bigl(1-(1-\tfrac{k}{n})\bigr)C_n^k\,x^k(1-x)^{n-k}=\\ =\bigl(\tfrac{k}{n}\bigr)^{\nu}C_n^k\,x^k(1-x)^{n-k}&-\bigl(1-\tfrac{1}{n}\bigr)^{\nu}(1-x)\bigl(\tfrac{k}{n}\bigr)^{\nu-1}C_{n-1}^k\,x^k(1-x)^{n-1-k}.\end{aligned}$$

Равенство ($*$) доказано. При $\nu=1$, используя (3), получим

$$B_n(t^2;x)=x^2+\frac{x(1-x)}{n}. \tag{$**$}$$

Предположим по индукции, что имеет место

$$B_n(t^\nu;x)=x^\nu+O\Bigl(\frac{x(1-x)}{n}\Bigr)\quad\text{при}\quad \nu=1,2,\,\dots\,,m,$$

где константа в O зависит только от m, и примéним ($*$). Тогда

$$\begin{aligned}B_n(t^{m+1};x)&=x^{m+1}+\bigl(1-\bigl(1-\tfrac{1}{n}\bigr)^m\bigr)\,x^m(1-x)+\\ &+O\Bigl(\frac{x(1-x)}{n}\Bigr)=x^{m+1}+O\Bigl(\frac{x(1-x)}{n}\Bigr).\end{aligned}$$

Напомним, что теорема о приближении непрерывных функций полиномами самим Вейерштрассом была доказана как теорема существования. Используя её в такой форме и применяя неравенство (15), С. Н. Бернштейн доказал два важных следствия.

ТЕОРЕМА 2. *Любая непрерывная на отрезке* $[0,1]$ *функция* f *разлагается в нормальный ряд.*

Доказательство. В самом деле, по теореме Вейерштрасса для любого натурального s существует полином P_{m_s} $(m_s > m_{s-1})$ такой, что

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

а по неравенству (15) существует $n_s > n_{s-1}$ такое, что

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

Значит, f разлагается в нормальный ряд

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

ТЕОРЕМА 3. *Для любой непрерывной на отрезке* $[0,1]$ *функции* f *последовательность* $B_n(f;x)$ *сходится к ней равномерно.*

Для *доказательства* теоремы 3 применяется теорема Вейерштрасса, так что теорема 3 слабее теоремы 1. Обозначим через (P_m) последовательность полиномов, которая сходится к f. По теореме Вейерштрасса и неравенству (15) для фиксированного $\varepsilon > 0$ и достаточно больших n и m имеем

$$\left|P_m(x) - f(x)\right| < \frac{\varepsilon}{3}, \qquad \left|B_n(P_m; x) - P_m(x)\right| < \frac{\varepsilon}{3},$$
$$\left|B_n(f,x) - f(x)\right| < B\bigl(\left|f - P_m\right|; x\bigr) + \left|B_n(P_m; x) - P_m(x)\right| + \left|P_m(x) - f(x)\right| < \varepsilon.$$

Таким образом,

$$\lim_{n\to\infty} B_n(f;x) = f(x). \tag{16}$$

В тексте С. Н. Бернштейном сказано, что “формула (16) выведена мною при помощи теории вероятностей в заметке, помещённой в «Сообщениях Харьковского математического общества, т. 13, № 1, 1912 г.»”.

Мне кажется несомненным, что С. Н. Бернштейн доказал теорему 1 после теоремы 3, но следует признать, что он этого нигде не пишет!

ЛИТЕРАТУРА

1. С. Н. Бернштейн. *Собрание сочинений.* М.: АН СССР. Т. 1, № 3, с. 79–84; № 4. Т. 2, №№ 57, 64, 65, 81. Т. 3, № 1, № 9, с. 49–53.