Три лекции о дзета-функции

А.М. Левин

НМУ, осень 2001

Лекция 1

Я не смогу в этом мини-курсе как следует объяснить, зачем вводятся дзета-функции, так что придется сослаться на авторитеты. Этих авторитетов будет даже два: Эйлер и Риман.

Мы начнем с *дзета-функции Римана.* В полном соответствии с известным «принципом Арнольда», это название является неточным: на самом деле эту функцию ввел Эйлер еще в XVIII веке (впрочем, Риман открыл важные свойства и замечательные применения этой функции, так что дзета-функцию связывают с его именем совершенно справедливо).

Итак, дзета-функция Римана определеяется по формуле

$$\zeta(s)=\sum_{n\geqslant 1}\frac{1}{n^s}. \tag{1}$$

Для всех комплексных s с $\mathrm{Re}(s)>1$ этот ряд сходится к «хорошей» (голоморфной) функции (в самом деле, достаточно установить сходимость интеграла $\int_1^\infty dx/x^s=1/(s-1)$; нелишне отметить, что под n^s при комплексных s понимается $e^{s\ln n}$).

У ряда (1) есть важное свойство, называемое *разложимостью в эйлерово произведение:*

$$\zeta(s)=\prod_p\Bigl(1-\frac{1}{p^s}\Bigr)^{-1},$$

где p пробегает все простые числа (точнее говоря: при при $\mathrm{Re}(s)>1$ произведение в правой части сходится к $\zeta(s)$).

Наивно (в духе Эйлера) эту формулу можно обосновать так:

$$\sum\frac{1}{n^s}=\sum_{\substack{\alpha_1,\dots,\alpha_n,\dots\\ \text{почти все } \alpha_i=0}}\frac{1}{(p_1^{\alpha_1}\cdot\dots\cdot p_k^{\alpha_k}\dots)^s}=\prod_p\sum_{\alpha=0}^{\infty}\frac{1}{(p^\alpha)^s}=\prod_p\frac{1}{1-\frac{1}{p^s}}.$$

К сожалению, в XXI веке так делать нельзя: надо обосновывать законность проведенных манипуляций, а мне бы этого не хотелось. Проверьте самостоятельно, что в абсолютно сходящемся произведении абсолютно сходящихся рядов можно раскрывать скобки.

Теперь докажем еще одно свойство дзета-функции, а именно, что ее можно аналитически продолжить в полуплоскость $\mathrm{Re}(s) > 0$, где ряд $\sum_{n\geqslant 1} \frac{1}{n^s}$, очевидно, расходится (возможно, это было сделано еще до Римана).

Доказательство состоит в том, что надо более аккуратно обойтись с интегральным признаком сходимости. Именно, легко видеть, что

$$\frac{1}{n^s} - \int_n^{n+1} \frac{dx}{x}^s = \frac{1}{n^s} - \frac{1}{x_0^s},$$

где $x_0 \in (n; n+1)$; с другой стороны, из формулы Лагранжа следует, что

$$\frac{1}{n^s} - \frac{1}{x_0^s} = (x_0 - n) \cdot s \frac{1}{y_0^{s+1}},$$

где $y_0 \in (n; n+1)$; стало быть, там, где ряд (1) и интеграл сходятся, имеем

$$\left|\zeta(s) - \frac{1}{s-1}\right| \leqslant \left|\sum\Big(\frac{1}{n^s} - \int_n^{n+1} \frac{dx}{x^s}\Big)\right| \leqslant \sum \frac{|s|}{|n^{s+1}|}.$$

Ряд в правой части сходится при $\mathrm{Re}(s) > 0$, поэтому при $\mathrm{Re}(s) > 0$ дзета-функцию можно доопределить так:

$$\zeta(s) = \frac{1}{s-1} + \Big(\frac{1}{n^s} - \int_n^{n+1} \frac{dx}{x^s}\Big). \tag{2}$$

Если оценивать разность между $1/n^s$ и интегралом более точно, то можно продолжить $\zeta(s)$ и далее влево (но нам это не понадобится, так как вскоре мы другим способом получим продолжение на всю комплексную плоскость).

Из того, что мы доказали, можно уже вывести некоторые следствия. Для этого возьмем от дзета-функции логарифм (считаем, что $\mathrm{Re}(s) > 1$, и выбираем ту ветвь логарифма, для которой логарифмы вещественных чисел вещественны):

$$\ln\zeta(s) = -\sum \ln(1 - p^{-s}) = \sum_p \sum_{k=1}^{\infty} \frac{1}{kp^{sk}} = \sum_p \frac{1}{p^s} + \sum_{\substack{p\\ k\geqslant 2}} \frac{1}{kp^{sk}}.$$

Второй из полученных рядов сходится при $s = 1$: в самом деле, $\frac{1}{kp^k} < \frac{1}{p^k}$, и $\sum_{k\geqslant 2} \frac{1}{kp^k} < \sum_{k\geqslant 2} \frac{1}{p^k} = 1/p(p-1)$. Поскольку в единице функция $\zeta(s)$ имеет, как мы только что выяснили, полюс, получается, что ряд, составленный из чисел, обратных к простым, расходится. Отсюда немедленно получается, что простых чисел бесконечно много.

Разумеется, это не самый простой способ доказывать бесконечность множества простых чисел, но наше рассуждение дает нам больше, чем бесконечность этого множества: из него следует, что простые числа расположены в натуральном ряду довольно густо, например, что их больше, чем полных квадратов.

Также с использованием дзета-функции, но гораздо более сложными рассуждениями, можно найти асимптотику для функции $\pi(x)$, равной, по определению, количеству простых чисел, не превосходящих x: оказывается, что $\pi(x)$ асимптотически эквивалентно $x/\ln x$.

С помощью той же техники, которую мы только что применили, доказывается и *теорема Дирихле об арифметической прогрессии,* утверждающая, что во всякой арифметической прогрессии, у которой первый член и разность — взаимно простые целые числа, содержится бесконечно много простых чисел.

До сих пор мы, можно сказать, занимались дзета-функцией не Римана, но Эйлера, поскольку рассуждения, которые мы провели, Эйлеру были в основном известны. Теперь перейдем собственно к вкладу Римана.

Риман обнаружил, что:

- Существует аналитическое продолжение дзета-функции на все $\mathbb{C}$ (как однозначной функции с единственным простым полюсом в единице);

- $\zeta(1-s)$ связана несложной формулой с $\zeta(s)$ (как говорят, «дзета-функция удовлетворяет функциональному уравнению»). Если заменить $\zeta(s)$ на просто связанную с ней функцию $\xi(s)$ (вскоре мы дадим ее точное определение), то функциональное уравнение будет иметь вид $\xi(1-s) = \xi(s)$. Кстати, функцию $\xi(s)$ тоже называют дзета-функцией.

Факт существования функционального уравнения дзета-функции представляется довольно загадочным: до сих пор непонятно, почему оно должно существовать. (Для аналогов дзета-функции над полями алгебраических функций от одной переменной над конечным полем функциональное уравнение дзета-функции тоже есть, но его смысл как раз понятен: по существу это теорема Римана-Роха для кривых.)

Перейдем к доказательству сформулированных выше результатов Римана. Доказательство будет довольно хитрым: оно напоминает решение хорошей олимпиадной задачи.

Для начала вспомним определение *гамма-функции Эйлера:*

$$\Gamma(s) = \int_0^\infty x^s e^{-x} \frac{dx}{x}, \qquad \mathrm{Re}(s) > 0$$

(запись, при которой x в знаменателе не выделен, является идеологически неправильной). Из этого определения с помощью интегрирования по частям немедленно получается тождество $\Gamma(s) = (s-1)\Gamma(s-1)$, справедливое при $\mathrm{Re}(s) > 1$. Пользуясь этим тождеством, можно распространить (а точнее — аналитически продолжить) гамма-функцию на всю комплексную плоскость, полагая $\Gamma(s) = \Gamma(s+1)/s$, и т. д.

Можно доказать, наконец, и тождество $\Gamma(s)\Gamma(1-s) = \pi/\sin \pi s$, но оно нам не понадобится.

А теперь начинаются чудеса. Умножим обе части формулы (1) на $\Gamma(s/2)$:

$$\Gamma\Big(\frac{s}{2}\Big)\zeta(s) = \Gamma\Big(\frac{s}{2}\Big) \cdot \sum_{n=1}^{\infty} \frac{1}{n^s} = \sum_{n=1}^{\infty} \int_0^{\infty} \frac{1}{n^s} x^{s/2} e^{-x} \frac{dx}{x} =$$

(сделаем замену $x = n^2 y$)

$$= \sum_{n=1}^{\infty} \int_0^{\infty} y^{s/2} e^{-n^2 y} \frac{dy}{y} = \int_0^{\infty} y^{s/2} \underbrace{\sum_{n=1}^{\infty} e^{-n^2 y}}_{\tilde{\theta}(y)} \frac{dy}{y}.$$

На самом деле в последнем равенстве есть обман: в окрестности нуля ряд для введенной нами функции $\tilde{\theta}(y)$ сходится неравномерно, так что менять местами знаки суммы и интеграла нужно осторожно.

Заметим, что

$$\tilde{\theta}(y) \leqslant \sum_{n=1}^{\infty} e^{-ny} = \frac{e^{-y}}{1 - e^{-y}}.$$

Поэтому в окрестности бесконечности $\tilde{\theta}$ убывает экспоненциально, стало быть, сходимость ряда для $\tilde{\theta}$ равномерна, и наша перестановка суммы и интеграла допустима. Вблизи нуля $\tilde{\theta}$ растет не быстрее $1/y$ (на самом деле, как мы увидим позже, она растет как $1/y^{\frac{1}{2}}$) и при $\mathrm{Re}\, s > 2$ ряд $\sum y^{s-1} e^{-n^2 y}$ сходится на замкнутом луче $[0, \infty)$ и стало быть, сходится равномерно в окрестности нуля, и перестановка суммирования и интегрирования также допустима.

Разобьем интеграл в сумму двух:

$$\int_0^{\infty} y^{s/2} \sum_{n=1}^{\infty} e^{-n^2 y} \frac{dy}{y} = \int_0^{T} y^{s/2} \sum_{n=1}^{\infty} e^{-n^2 y} \frac{dy}{y} + \int_T^{\infty} y^{s/2} \sum_{n=1}^{\infty} e^{-n^2 y} \frac{dy}{y} \quad (3)$$

(число T подберем потом). Интеграл от T до ∞ сходится при всех s и определяет голоморфную функцию от s.

Для того, чтобы разобраться с первым слагаемым нам потребуется еще один трюк.

Отступление: формула суммирования Пуассона

Пусть f — гладкая и быстро убывающая (последнее означает, что и сама функция, и все ее производные убывает на бесконечности быстрее, чем $1/|x|^N$, для любого N) функция на $\mathbb{R}$. Рассмотрим ее преобразование Фурье $\hat{f}(p) = \int_{-\infty}^{\infty} f(x)e^{2\pi ipx}dx$. Формула Пуассона гласит, что при этих условиях

$$\sum_{n\in\mathbb{Z}} f(n) = \sum_{m\in\mathbb{Z}} \hat{f}(m).$$

Доказательство формулы Пуассона. Положим $F(x) = \sum_{n\in\mathbb{Z}} f(x+n)$. Функция F бесконечно дифференцируема и периодична, так что ее ряд Фурье к ней сходится, и можно записать разложение $F(x) = \sum_{m\in\mathbb{Z}} a_m e^{-2\pi imx}$. При этом

$$\begin{aligned} a_n &= \int_0^1 F(x)e^{2\pi imx}dx = \int_0^1 \sum_n f(x+n)e^{2\pi imx}dx = \\ &= \sum_n \int_0^1 f(x+n)e^{2\pi imx}dx = \sum_n \int_n^{n+1} f(x)e^{2\pi imx}dx = \\ &= \int_{-\infty}^{\infty} f(x)e^{2\pi imx}dx = \hat{f}(m). \end{aligned}$$

Отсюда получаем $F(0) = \sum_m \hat{f}(m)$, и формула суммирования Пуассона доказана.

Применим формулу суммирования Пуассона к функции

$$\theta(y) = 1 + 2\tilde{\theta}(y) = \sum_{n\in\mathbb{Z}} e^{-n^2y}.$$

Для этого положим $f(x) = e^{-x^2y}$, тогда $\theta(y) = \sum_{n\in\mathbb{Z}} f(n)$. Как известно, $\hat{f}(p) = \sqrt{\pi/y}e^{-(\pi^2/y)p^2}$. Поэтому формула Пуассона в нашем случае превращается в такое тождество:

$$\theta(y) = \sqrt{\frac{\pi}{y}} \cdot \theta\Big(\frac{\pi^2}{y}\Big). \tag{4}$$

Завершение доказательства

Вернемся к выводу функционального уравнения. Для начала слегка преобразуем первое слагаемое в формуле (3):

$$\begin{aligned} \int_0^T y^{s/2} \sum_{n=1}^{\infty} e^{-n^2y}\frac{dy}{y} &= \frac{1}{2}\Big(\sum_{n\in\mathbb{Z}} \int_0^T y^{s/2}e^{-n^2y}\frac{dy}{y}\Big) - \frac{1}{2}\int_0^T y^{s/2}\frac{dy}{y} = \\ &= \frac{1}{2}\Big(\sum_{n\in\mathbb{Z}} \int_0^T y^{s/2}e^{-n^2y}\frac{dy}{y}\Big) - \frac{T^{s/2}}{s} \end{aligned}$$

Теперь сделаем в интегралах в последней формуле замену $y = \pi^2/u$. Выражение примет вид

$$-\frac{T^{s/2}}{s} + \frac{\pi^s}{2}\Big(\sum_{n\in\mathbb{Z}} \int_{\pi^2/T}^{\infty} u^{-s/2}e^{-n^2\pi^2/u}\frac{du}{u}\Big). \tag{5}$$

Заметим, что, по формуле (4), имеем

$$\sum_{n\in\mathbb{Z}} u^{-s/2}e^{-n^2\pi^2/u} = \sum_{n\in\mathbb{Z}} \sqrt{\frac{u}{\pi}}u^{-s/2}e^{-n^2u}.$$

Стало быть, выражение (5) равно

$$-\frac{T^{s/2}}{s}+\frac{\pi^{s-\frac12}}{2}\Big(\int_{\pi^2/T}^{\infty} u^{\frac{1-s}{2}}\theta(u)\frac{du}{u}\Big) = \\ = -\frac{T^{s/2}}{s}+\frac{\pi^{s-\frac12}}{s-1}\Big(\frac{\pi^2}{T}\Big)^{\frac{1-s}{2}}+\frac{\pi^{s-\frac12}}{2}\int_{\pi^2/T}^{\infty} u^{\frac{1-s}{2}}\tilde\theta(u)\frac{du}{u},$$

где второй член есть интеграл от слагаемого θ-ряда, отвечающего $m = 0$.

Теперь положим $T=\pi$, определим обещанную функцию $\xi(s)$ по формуле

$$\xi(s) = \pi^{-s/2}\Gamma\Big(\frac{s}{2}\Big)\zeta(s)$$

и подставим в (3); получим формулу

$$\xi(s) = -\frac{1}{s}+\frac{1}{s-1}+\pi^{-s/2}\int_{\pi}^{\infty} y^{\frac{s}{2}}\tilde\theta(y)\frac{dy}{y}+\pi^{(s-1)/2}\int_{\pi}^{\infty} u^{\frac{1-s}{2}}\tilde\theta(u)\frac{du}{u}.$$

Отсюда видно и тождество $\xi(1-s)=\xi(s)$, то есть функциональное уравнение, и аналитическое продолжение $\xi(s)$ (а значит, и $\zeta(s)$) на всю комплексную плоскость.

Итак, доказательство Римана завершено. Теперь несколько замечаний.

Во-первых, можно написать равенство

$$\xi(s) = \pi^{-s/2}\Gamma\Big(\frac{s}{2}\Big)\prod_p (1-p^{-s})^{-1}.$$

Значит, можно предполагать, что в каком-то высшем смысле множитель $\pi^{-s/2}\Gamma(s/2)$ имеет ту же природу, что и $(1-p^{-s})^{-1}$. Дж. Тэйт в свое время показал, что так оно и есть (при этом $\pi^{-s/2}\Gamma(s/2)$ соответствует полю $\mathbb{R}$, а $(1-p^{-s})^{-1}$ — полю p-адических чисел $\mathbb{Q}_p$).

Во-вторых, давайте обсудим, как устроена функция $\xi(s)$. В силу функционального уравнения, при $\mathrm{Re}(s) > 1$ она устроена так же, как при $\mathrm{Re}(s) < 0$. Далее, при $\mathrm{Re}(s) > 1$ функция $\zeta(s)$ не имеет нулей (поскольку к ней сходится бесконечное произведение), стало быть, при $\mathrm{Re}(s) > 1$ (а также, в силу функционального уравнения, при $\mathrm{Re}(s) < 0$) у функции $\xi(s)$ нулей нет, и все ее нули сконцентрированы в полосе $0 < \mathrm{Re}(s) < 1$.

Знаменитая *гипотеза Римана* гласит, что все нули функции $\xi(s)$ расположены на прямой $\mathrm{Re}(s) = 1/2$. Она подтверждается многими численными экспериментами, но как ее доказывать, совершенно непонятно.

Упражнение. Определим *дзета-функцию Эпштейна* по формуле

$$\zeta_{\mathbf{1}_k}(s) = \sum_{\substack{(n_1,\ldots,n_k)\in\mathbb{Z}^k\\(n_1,\ldots,n_k)\neq(0,\ldots,0)}} \frac{1}{(n_1^2+\ldots+n_k^2)^{s/2}}.$$

Выведите для нее функциональное уравнение.

Более общим образом, пусть Q — положительно определенная квадратичная форма:

$$Q(n_1,\ldots,n_k) = \sum a_{ij}n_in_j.$$

Для нее дзета-функция Эпштейна определяется по формуле

$$\zeta_Q(s) = \sum_{\substack{(n_1,\ldots,n_k)\in\mathbb{Z}^k\\(n_1,\ldots,n_k)\neq(0,\ldots,0)}} \frac{1}{Q(n_1,\ldots,n_k)^{s/2}}.$$

Выведите функциональное уравнение и для нее.
Указание. Начните с того, что умножьте на $\Gamma(s/2)$.

Случай произвольного числового поля

Теперь перейдем к более интересным вещам. Пусть F — конечное расширение поля $\mathbb{Q}$; в нем, как известно, есть подкольцо целых $\mathcal{O}_F \subset F$. Как написать для этого кольца что-то вроде дзета-функции Римана?

Первое, что приходит в голову, — это ряд

$$\sum_{a\in\mathcal{O}_F\setminus\{0\}} \frac{1}{|N(a)|^s}$$

(напомним, что $N(a)$ (норма элемента a) — это определеитель оператора $a\colon F\to F$ умножения на a, где F рассматривается как векторное пространство над $\mathbb{Q}$).

К сожалению, так делать нельзя, потому что этот ряд может чудовищно расходиться. Дело в том, что чисел $a\in\mathcal{O}_F$, у которых $|N(a)| = 1$, может быть бесконечно много. Рассмотрим, например, поле $F = \mathbb{Q}(\sqrt{3})$; тогда $\mathcal{O}_F = \mathbb{Z}[\sqrt{3}]$, $N(x+\sqrt{3}y) = x^2-3y^2$, и у уравнения $x^2-3y^2 = 1$ есть бесконечно много решений: в самом деле, $N(2+\sqrt{3}) = 1$, и, возводя это число в степень, получаем бесконечно много других чисел с нормой 1. (Вообще, элемент кольца $\mathcal{O}_F$ обратим тогда и только тогда, когда его норма равна ±1.)

Чтобы исправить положение, попробуем отфакторизовать кольцо $\mathcal{O}_F$ по обратимым элементам (собственно говоря, при определении дзета-функции Римана мы так и делали: суммировали не по всем целым числам, а только по положительным). Получится ряд

$$\sum_{a \in \mathcal{O}_F/\mathcal{O}_F^*} \frac{1}{|N(a)|^s}.$$

Этот ряд уже сходится (его сумма называется *частичной дзета-функцией*); недостаток этой функции в том, что, вообще говоря, ее нельзя разложить в эйлеровское произведение, так как нельзя гарантировать, что в кольце $\mathcal{O}_F$ есть однозначное разложение на простые множители.

Чтобы исправить этот недостаток частичной дзета-функции, заметим, что $\mathcal{O}_F/\mathcal{O}_F^*$ — множество *главных* идеалов кольца $\mathcal{O}_F$; так вот, возьмем сумму по множеству всех идеалов кольца $\mathcal{O}_F$, а не только главных:

$$\zeta_F(s) = \sum_{\mathfrak{a} \text{ — идеал } \mathcal{O}_F} \frac{1}{(N\mathfrak{a})^s}.$$

Здесь $N\mathfrak{a}$ (норма идеала $\mathfrak{a}$) — число элементов в факторкольце $\mathcal{O}_F/\mathfrak{a}$. Заметим, кстати, что для $a \in \mathcal{O}_F$ имеется равенство $N(\,(a)\,) = |N(a)|$ (доказательство: если $M\colon \mathbb{Z}^n \to \mathbb{Z}^n$ — эндоморфизм, то $|\det M|$ равен порядку группы $\mathbb{Z}^n/M(\mathbb{Z}^n)$).

Функция $\zeta_F(s)$ называется *дзета-функцией Дедекинда;* поскольку идеалы в $\mathcal{O}_F$ однозначно разлагаются в произведение простых идеалов, имеет место разложение в эйлеровское произведение:

$$\zeta_F(s) = \sum_{\mathfrak{a} \text{ — идеал } \mathcal{O}_F} \frac{1}{(N\mathfrak{a})^s} = \prod_{\mathfrak{p} \text{ простой}} (1 - (N\mathfrak{p})^{-s})^{-1}.$$

На следующих лекциях мы построим для дзета-функции Дедекинда функциональное уравнение и аналитическое продолжение.

Заметим, кстати, что терминология и тут непоследовательна: уж если открытая Эйлером дзета-функция называется дзета-функцией Римана потому, что Риман нашел для нее функциональное уравнение, то и дзета-функцию Дедекинда логично было бы назвать в честь того, кто доказал функциональное уравнение для нее, а это был не Дедекинд, а Гекке.