Глава I

ПРОИЗВОДНАЯ

§ 1. ТЕОРЕМА ЛЕБЕГА О ПРОИЗВОДНОЙ МОНОТОННОЙ ФУНКЦИИ

1. Пример непрерывной функции, не имеющей производной. В классическом анализе обычно рассматриваются функции, имеющие производные, даже непрерывные до некоторого порядка. Иногда в отдельных точках производные могут не существовать или претерпевать разрывы. Но до начала нынешнего столетия лишь изредка задавались вопросом, всегда ли обладают производными функции того или иного класса, например непрерывные или монотонные, и каковы множества тех точек, в которых эти производные могут не существовать. В этом направлении были получены лишь некоторые почти очевидные результаты, как, например, существование левой и правой производных у выпуклой функции, откуда следовала дифференцируемость такой функции при всех x, кроме, может быть, счетного множества значений.

Историю этого вопроса принято начинать с 1806 г., когда Ампер в статье [1], посвященной „теории производных функций“, безуспешно пытался установить дифференцируемость „произвольной“ функции всюду, за исключением некоторых „исключительных и изолированных“ значений аргумента. Впрочем, если иметь в виду эволюцию понятия функции, то возникает уверенность — хотя оригинальный текст не дает на этот счет определенных указаний,— что утверждение Ампера относилось лишь к кусочно-монотонным функциям.

Последующие десятилетия, сколь плодотворны они ни были для развития анализа, не продвинули решения этой проблемы. Лишь Вейерштрасс положил конец попыткам доказать дифференцируемость непрерывных функций сколько-нибудь общего типа, построив пример *непрерывной функции, нигде не имеющей производной* [1]). Построением других, более простых функций такого рода и выяснением их свойств в связи с рядом других задач занимались почти все выдающиеся математики второй половины прошлого века, работавшие в области анализа; занимаются этим и в наши дни. Здесь мы приведем пример, по-видимому, наиболее элементарный, принадлежащий ван дер Вардену [1]; идея его основана на том, что последовательность, состоящая из целых чисел, сходится только тогда, когда все ее члены, начиная с некоторого, совпадают.

[1]) Опубликоваио Дю Буа-Реймоиом [1].

Будем говорить, как обычно, что функция $f(x)$ имеет в точке x производную, равную $f'(x)$, если отношенне

$$\frac{f(x+h)-f(x)}{h}$$

стремнтся к определенному конечному пределу $f'(x)$, когда $h \to 0$ и $x+h$ пробегает значения, при которых $f(x+h)$ определена. Обозначив $\{x\}$ расстояние от x до ближайшего целого чнсла, рассмотрим функцию

$$f(x)=\sum_{n=0}^{\infty}\frac{\{10^n x\}}{10^n}. \tag{1}$$

Члены этого ряда положительны, непрерывны и не превосходят 10^{-n}, поэтому функцня $f(x)$ непрерывна. Попытаемся теперь вычислить производную функции $f(x)$ в какой-нибудь точке x.

Для этого заметим, что можно ограничиться рассмотреннем ннтервала $0 \leqslant x < 1$, н запншем x в внде

$$x=0,a_1a_2\ldots a_n\ldots,$$

причем в случае конечной десятичной дробн условнмся дополнять ее нулями. Дробь

$$0,a_{n+1}a_{n+2}\ldots$$

может быть $\leqslant \frac{1}{2}$ или $> \frac{1}{2}$. В первом случае

$$\{10^n x\}=0,a_{n+1}a_{n+2}\ldots,$$

во втором —

$$\{10^n x\}=1-0,a_{n+1}a_{n+2}\ \ldots.$$

Положив $h_m=-10^{-m}$ тогда, когда a_m равно 4 илн 9, и $h_m=10^{-m}$ при любом другом a_m, рассмотрнм отношение

$$\frac{f(x+h_m)-f(x)}{h_m}; \tag{2}$$

согласно (1), его можно записать в виде

$$10^m\sum_{n=0}^{\infty}\pm\frac{\{10^n(x\pm 10^{-m})\}-\{10^n x\}}{10^n}.$$

Числнтели напнсанных дробей обращаются в нуль при $n \geqslant m$ н приводятся к $\pm 10^{n-m}$ при $n \leqslant m$. Следовательно, выражение (2) принимает целые значения, положительные, отрицательные или равные нулю, но во всяком случае той же четностн, что $m-1$. Таким образом, последовательность отношений (2), состоящая из целых чисел, поочередно четных н нечетных, не может сходиться.

2. Теорема Лебега о производной монотонной функции. Множества меры нуль. Перейдем к рассмотрению класса монотонных функций. Следующая теорема, одна из самых замечательных и важных в анализе, принадлежит Лебегу.

Теорема. *Любая монотонная функция $f(x)$ имеет определенную конечную производную всюду, кроме, может быть, некоторого множества значений x меры нуль* (иначе говоря, *почти всюду*).

Прежде чем дать необходимые пояснения, заметим, что Лебег доказал эту теорему *при дополнительном предположении, что $f(x)$ непрерывна*. Доказательство приведено в первом издании его книги [2] (1904 г.) в конце последней главы, как итог всей теории. Однако в формулировку теоремы не входят ни понятие интеграла, ни понятие меры. Действительно, понятие множества меры нуль не связано существенным образом с общей теорией меры, и основные свойства таких множеств могут быть установлены в нескольких словах.

Множеством меры нуль называется, согласно Лебегу, любое множество чисел x, могущее быть покрытым конечным числом или счетной последовательностью интервалов, сумма длин которых сколь угодно мала. Из этого определения сразу следует, что всякое подмножество такого множества само является множеством меры нуль. Множеством меры нуль будет также объединение конечного числа или счетной последовательности множеств меры нуль; для того чтобы в этом убедиться, достаточно покрыть заданные множества системами интервалов с суммой длин, меньшей $\varepsilon\cdot 2^{-n}$; тогда совокупность всех таких интервалов покроет объединение заданных множеств, а сумма их длин будет меньше ε. В частности, любое конечное или счетное числовое множество представляет собой множество меры нуль.

Иногда бывает полезно следующим образом видоизменить наше определение: E есть множество меры нуль, если E можно покрыть счетной последовательностью интервалов с конечной суммой длин таким образом, что любая точка множества E окажется внутри бесконечной системы этих интервалов. Такое определение эквивалентно первому. В самом деле, если все точки множества E принадлежат бесконечной системе интервалов с конечной суммой длин, то эту сумму можно сделать сколь угодно малой, удалив конечное число интервалов из заданной системы. Обратно, когда E есть множество меры нуль в смысле первого определения, E можно покрыть последовательно системами интервалов с суммами длин $< 2^{-n}$ и, если это нужно, растянуть интервалы вправо и влево так, чтобы они стали вдвое длиннее; объединение таких систем будет удовлетворять всем требованиям, содержащимся в нашем втором определении.

Термин „почти всюду“, употребляемый по отношению к некоторому свойству, означает, что это последнее имеет место всюду, за исключением, быть может, некоторого множества меры нуль.

Прежде чем доказать сформулированную выше теорему Лебега, покажем, что эта теорема в известном смысле дает окончательное решение вопроса о дифференцируемости монотонной функции и поэтому не может быть усилена. Пусть дано какое-нибудь множество E меры нуль. Построим возрастающую функцию $f(x)$, не имеющую конечной производной в точках множества E (говоря точнее, в этих точках она будет иметь бесконечную производную). Для этой цели заключим E в такую систему интервалов, о которой идет речь во втором определении множеств меры нуль, и зададим $f(x)$ как сумму длин тех интервалов и тех частей интервалов выбранной системы, которые окажутся левее точки x; такая функция будет, очевидно, обладать требуемым свойством.

3. Доказательство теоремы Лебега. Существование почти всюду производной у монотонной функции мы докажем, не обращаясь к теории интеграла. Первые доказательства такого рода принадлежали Фаберу [1], а также Г. и У. Юнгам [1].

Для простоты предположим сначала, что рассматриваемая функция не только монотонна, но и непрерывна, а затем укажем те изменения, впрочем почти очевидные, которые нужно внести в доказательство для того, чтобы освободиться от условия непрерывности.

Доказательство будет основано на следующем вспомогательном предложении.

Лемма[1]). *Пусть $g(x)$ — непрерывная функция, определенная на интервале $a \leqslant x \leqslant b$, и пусть E — множество внутренних точек x этого интервала, обладающих тем свойством, что правее x найдется точка ξ, в которой $g(\xi) > g(x)$. Тогда E либо пусто, либо является открытым множеством, т. е. оно разбивается на конечное или счетное множество непересекающихся открытых интервалов (a_k, b_k), причем для всех них*

$$g(a_k) \leqslant g(b_k).$$

Для доказательства заметим сначала, что E — открытое множество, так как если $\xi > x_0$ и $g(\xi) > g(x_0)$, то, в силу непрерывности функции $g(x)$, неравенства $\xi > x$, $g(\xi) > g(x)$ выполняются для всех x из некоторой окрестности точки x_0. Пусть теперь (a_k, b_k) — какой-нибудь из открытых интервалов, составляющих E. Точка b_k этому последнему не принадлежит; мы покажем, что для любой точки x интервала (a_k, b_k) выполняется неравенство $g(x) \leqslant g(b_k)$, и, перейдя к пределу при $x \to a_k$, получим требуемое неравенство.

---

[1]) Рисс [18].

Пусть x_1 — точка, лежащая между x и b_k и ближайшая к b_k из тех, в которых выполняется неравенство $g(x_1) \geqslant g(x)$; покажем, что x_1 совпадает с b_k. Если бы этого не было, то ни одна из точек ξ_1, соответствующих x_1, согласно определению множества E, не совпадала бы с b_k, и так как сама b_k множеству E не принадлежит, то мы пришли бы к противоречию: $g(x_1) < g(\xi_1) \leqslant \leqslant g(b_k) < g(x_1)$.

Читатель без труда сможет убедиться в том, что в действительности $g(a_k) = g(b_k)$, за исключением, может быть, того случая, когда $a_k = a$, но для дальнейшего это нам не понадобится.

Итак, пусть $f(x)$ — непрерывная монотонная функция на интервале $a \leqslant x \leqslant b$; для определенности предположим, что $f(x)$ не убывает. Рассмотрим производные числа этой функции. Как известно, правыми *производными числами* Λ_d и λ_d, верхним и нижним, называются соответственно верхний и нижний пределы отношения $\frac{f(x+h)-f(x)}{h}$ при $h > 0$, $h \to 0$; аналогично определяются левые производные числа Λ_g и λ_g. Для тех и других допускаются бесконечные значения. Определенная конечная производная существует в тех точках x, в которых все четыре производных числа имеют одно и то же конечное значение.

Для доказательства теоремы Лебега достаточно установить почти всюду неравенства:

$$1)\ \Lambda_d < \infty; \quad 2)\ \Lambda_d \leqslant \lambda_g.$$

В самом деле, применив 2) к функции $-f(-x)$, мы получим, что почти всюду

$$\Lambda_g \leqslant \lambda_d,$$

откуда, в силу 1) и 2), будут следовать соотношения

$$\Lambda_d \leqslant \lambda_g \leqslant \Lambda_g \leqslant \lambda_d \leqslant \Lambda_d < \infty,$$

означающие существование конечной производной почти всюду.

Чтобы доказать утверждение 1), т. е. то, что точки x, в которых $\Lambda_d = \infty$, образуют множество E_∞ меры нуль, заметим, что, каково бы ни было число C, множество E_∞ содержится в множестве E_C тех точек x, в которых $\Lambda_d > C$. Если же $\Lambda_d > C$, то существует точка $\xi > x$, такая, что

$$\frac{f(\xi)-f(x)}{\xi - x} > C,$$

т. е. если положить $g(x) = f(x) - Cx$, то $g(\xi) > g(x)$. Таким образом, множество E_C заключено в интервалах (a_k, b_k), о которых шла речь в предыдущей лемме, а, согласно этой последней,

$$f(b_k) - Cb_k \geqslant f(a_k) - Ca_k,$$

откуда следует, что

$$C(b_k - a_k) \leqslant f(b_k) - f(a_k).$$

Суммируя, получим

$$C\sum(b_k-a_k)\leqslant\sum[f(b_k)-f(a_k)]\leqslant f(b)-f(a).$$

Отсюда вытекает, что при достаточно большом C сумма длин соответствующих интервалов $(a_k,\ b_k)$ становится сколь угодно малой. Следовательно, E_∞ представляет собой множество меры нуль.

Утверждение 2) можно получить последовательным применением подобных рассуждений. Пусть $c<C$ — какие-нибудь положительные числа. Возьмем функцию $g(x)=f(-x)+cx$, и пусть Σ_1 — система интервалов $(a_k,\ b_k)$, построенных для этой функции так, как указано в предыдущей лемме (точнее, система интервалов, симметричных им относительно начала); рассуждая так же, как и выше, мы придем к заключению, что Σ_1 содержит все те точки x, в которых $\lambda_g<c$. Пусть, далее, Σ_2 — система интервалов $(a_{kl},\ b_{kl})$, представляющих собой пересечения интервалов $(a_k,\ b_k)$ со всевозможными интервалами, построенными, в согласии с предыдущей леммой, для функции $g(x)=f(x)-Cx$. Тогда для всех интервалов $(a_k,\ b_k)$ и $(a_{kl},\ b_{kl})$ получим неравенства

$$f(b_k)-f(a_k)\leqslant c(b_k-a_k),\qquad C(b_{kl}-a_{kl})\leqslant f(b_{kl})-f(a_{kl}),$$

из которых будут следовать соотношения

$$C|\Sigma_2|\leqslant V_2\leqslant V_1\leqslant c|\Sigma_1|,$$

где $|\Sigma_1|$ и $|\Sigma_2|$ означают суммы длин интервалов, образующих соответствующие системы, а V_1 и V_2 — суммы изменений функции $f(x)$ на таких интервалах. Числа $|\Sigma_1|$ и $|\Sigma_2|$ оказываются связанными неравенством

$$|\Sigma_2|\leqslant\frac{c}{C}|\Sigma_1|.$$

Повторяя это построение, мы получаем последовательность систем интервалов $\Sigma_1,\ \Sigma_2,\ \ldots$, каждая из которых содержится в предыдущих, причем

$$|\Sigma_{2n}|\leqslant\frac{c}{C}|\Sigma_{2n-1}|.$$

Отсюда следует, что

$$|\Sigma_{2n}|\leqslant\left(\frac{c}{C}\right)^n|\Sigma_1|\to 0.$$

Точки x, в которых $\Lambda_d>C$ и одновременно $\lambda_g<c$, заключены во всех системах интервалов Σ_n, и, следовательно, образуют множество E_{cC} меры нуль. Всякая точка x, в которой $\Lambda_d>\lambda_g$, содержится в одном из таких множеств E_{cC}, причем соответствующие числа c и C можно считать рациональными, так как между любыми двумя действительными числами найдется пара рациональных чисел. Таким образом, если образовать множества E_{cC} для всевозможных пар рациональных положительных чисел, то объединение E^* таких множеств будет содержать все точки x, в которых $\Lambda_d > \lambda_g$. Но существует всего лишь счетное множество пар рациональных чисел, следовательно, E^*, будучи объединением счетной системы множеств меры нуль, само является множеством меры нуль. Вместе с тем множеством меры нуль окажется содержащееся в E^* множество тех x, в которых $\Lambda_d > \lambda_g$.

Таким образом, теорема доказана для случая *непрерывной* монотонной функции. Для того чтобы распространить эту теорему на *разрывные* функции, достаточно заметить, что и нашу лемму можно соответствующим образом обобщить. В самом деле, для наших целей достаточно предположить, что в любой точке x существуют пределы $g(x-0)$ и $g(x+0)$, так как этому условию удовлетворяет монотонная функция $f(x)$, а вместе с ней и функции $f(x)-Cx$, $f(-x)+cx$. Наибольшее из чисел $g(x-0)$, $g(x)$ и $g(x+0)$ мы обозначим $G(x)$, причем условимся считать, что $g(a-0)=g(a)$ и $g(b+0)=g(b)$. Тогда те внутренние точки x интервала (a, b), если они существуют, для которых найдутся $\xi > x$, такие, что $g(\xi) > G(x)$, образуют открытое множество, и для открытых интервалов (a_k, b_k), на которые это множество распадается, будет выполняться неравенство $g(a_k+0) \leqslant G(b_k)$.

Изменения, которые придется произвести в доказательстве нашей леммы в такой обобщенной формулировке, а также в доказательстве основной теоремы применительно к разрывным монотонным функциям, настолько очевидны, что мы позволим себе опустить подробности. Мы хотим, однако, подчеркнуть, что введение функции $G(x)$ никак не отразится на наших рассуждениях, когда они относятся к точкам непрерывности; что же касается точек разрыва, то они образуют счетное множество[1]), следовательно, множество меры нуль, и его можно просто исключить из рассмотрения.

4. Функции с ограниченным изменением. Теперь мы распространим полученный результат на еще более обширный класс функций, а именно на функции с ограниченным изменением. Эти функции играют важную роль в различных областях анализа, в частности в теории рядов Фурье, в вопросах, относящихся к спрямлению кривых, и, разумеется, в теории интегрирования. К понятию функции с ограниченным изменением можно прийти,

---

[1]) Достаточно заметить, что, каково бы ни было целое положительное число k, число точек x, в которых

$$|g(x+0)-g(x-0)| > \frac{1}{k},$$

может быть лишь конечным. В самом деле, будь их множество бесконечно, из него можно было бы выделить сходящуюся последовательность, возрастающую или убывающую. В точке ξ, служащей пределом такой последовательности, по крайней мере один из пределов $g(\xi-0)$ или $g(\xi+0)$ не мог бы существовать, что противоречит нашему предположению.

заметив, что если функции $f_1(x)$ и $f_2(x)$ дифференцируемы почти всюду, то и их разность $f(x)=f_1(x)-f_2(x)$ обладает тем же свойством, и если $f_1(x)$ и $f_2(x)$ — функции неубывающие, то для любого разбиения интервала (a, b) на частичные интервалы (x_{k-1}, x_k) $(k=1, 2, \ldots, n;\ x_0=a,\ x_n=b)$ выполняется неравенство

$$\sum_1^n |f(x_k)-f(x_{k-1})| \leqslant f_1(b)-f_1(a)+f_2(b)-f_2(a).$$

Функция $f(x)$, непрерывная или нет, называется *функцией с ограниченным изменением*, если суммы

$$\Sigma_{ab}=\sum_1^n |f(x_k)-f(x_{k-1})| \tag{3}$$

не превосходят некоторого фиксированного числа, не зависящего от выбора разбиения. Верхняя грань таких сумм называется *полным изменением функции $f(x)$ на интервале* (a, b) и обозначается $T(a, b)$.

Полное изменение представляет собой „аддитивную функцию интервала". Это означает, что если точка c заключена между a и b, то $f(x)$ имеет ограниченное изменение на (a, b) тогда и только тогда, когда она имеет ограниченное изменение на (a, c) и на (c, b), и при этом

$$T(a, b)=T(a, c)+T(c, b).$$

Для доказательства достаточно заметить, что суммы Σ_{ab} могут только возрастать при добавлении новых точек деления, и поэтому можно рассматривать лишь такие разбиения, при которых существует x_k, равное c; при этом $\Sigma_{ab}=\Sigma_{ac}+\Sigma_{cb}$ и, взяв верхние грани, мы получим требуемое равенство.

Мы только что показали, что разность двух неубывающих функций имеет ограниченное изменение. Жордану принадлежит обратная

Теорема. *Всякая функция с ограниченным изменением представляется в виде разности двух неубывающих функций.*

Доказательство ее очень просто. Введем функцию $T(x)=T(a, x)$, равную полному изменению функции $f(x)$, вычисленному для произвольного интервала (a, x); по аналогии с неопределенным интегралом назовем эту функцию *неопределенным полным изменением* функции $f(x)$. Тогда $T(x)$ и $T(x)-f(x)$ будут неубывающими функциями, и мы получим представление

$$f(x)=T(x)-[T(x)-f(x)].$$

В самом деле, если $x<\xi$, то

$$T(a, \xi)=T(a, x)+T(x, \xi),$$

откуда следует, что

$$T(\xi)-T(x)=T(x, \xi)\geqslant 0.$$

Для того чтобы показать, что

$$T(x)-f(x)\leqslant T(\xi)-f(\xi)$$

или, что то же самое,

$$f(\xi)-f(x)\leqslant T(\xi)-T(x)=T(x,\ \xi),$$

заметим, что $|f(\xi)-f(x)|$ представляет собой частный случай суммы $\Sigma_{x\xi}$ (когда между x и ξ нет точек деления); поэтому

$$|f(\xi)-f(x)|\leqslant T(x,\ \xi).$$

Это неравенство дает возможность представить $f(x)$ еще иначе, в виде разности двух монотонных функций:

$$f(x)=P(x)-N(x),$$

где функции $P(x)$ и $N(x)$, определяемые с точностью до постоянного слагаемого равенствами

$$P(x)=\frac{1}{2}[T(x)+f(x)],\qquad N(x)=\frac{1}{2}[T(x)-f(x)],$$

называются соответственно *положительным* и *отрицательным изменением* $f(x)$ на интервале $(a,\ x)$; мы будем называть их также *неопределенным положительным изменением* и *неопределенным отрицательным изменением* функции $f(x)$.

Наконец, так как $f(x)=P(x)-N(x)$ дифференцируема всюду, где дифференцируемы одновременно и $P(x)$ и $N(x)$, а объединение двух множеств меры нуль является множеством меры нуль, то мы получаем следующий окончательный результат:

Теорема Лебега. *Любая функция с ограниченным изменением почти всюду имеет конечную производную.*

§ 2. НЕКОТОРЫЕ СЛЕДСТВИЯ ТЕОРЕМЫ ЛЕБЕГА

5. Теорема Фубини о почленном дифференцировании ряда с монотонными членами. Теперь мы рассмотрим некоторые следствия, вытекающие более или менее непосредственно из только что доказанной основной теоремы. Начнем с теоремы Фубини, касающейся почленного дифференцирования ряда с монотонными членами.

Теорема Фубини[1]). *Пусть $f_1(x),\ f_2(x),\ \ldots$ — неубывающие (невозрастающие) функции на интервале $a\leqslant x\leqslant b$ и ряд*

$$f_1(x)+f_2(x)+\ldots=s(x) \tag{4}$$

сходится на этом интервале. Тогда для всех x, за исключением,

[1]) Фубини [2]; см. также Тонелли [1], Райхман и Сакс [1].