В.И. Сушков

Инвариантность первого дифференциала - нелепость в курсе математического анализа

В "Примерной программе", разработанной и утверждённой в 2000 году Научно - методическим советом по математике под руководством Л.Д. Кудрявцева, в сотнях тысяч (в миллионах?) экземпляров учебников [4-6], в тысячах повторяющих их "методических" изданий ВУЗов, в лекциях сотен лекторов, на 4700 страницах 536 сайтов, словами "инвариантность первого дифференциала" и "не инвариантность старших дифференциалов" означены *неверные* и *нелепые* утверждения. Цитирую их.

ЦИТАТЫ

Насильственное переобозначение Δx на $\mathrm{d}x$ (изгнание Δx)

Цитата 0 А. Коши [1] стр. 18:

"...(2) $\mathrm{d}f(x) = hf'(x)$

В частном случае, когда $f(x) = x$ уравнение (2) дает

(3) $\mathrm{d}x = h$.

Итак дифференциал переменной независимой x, равняется постоянному количеству h. Уравнение (2) можно заменить следующим

(4) $\mathrm{d}f(x) = f'(x)\,\mathrm{d}x$ ... "

(Коши я процитировал, чтобы читатель мог сравнить его с его последователями. В советских учебниках было принято писать О. Коши. И поныне употребляют не первую букву имени Augustin, а ту букву русского алфавита, которая соответствует французскому звучанию этого имени - Огюстэн. Однако В.Я. Буняковский, первый переводчик лекций Коши, написал "А.Л. Коши" [1]. Я следую ему. - *Сушков В.И.*).

Цитата 1 Э.Гурса [2] стр. 56: "... $\mathrm{d}y = f'(x)\Delta x$. Если функция $f(x)$ равна x, то предыдущая формула обращается в $dx = \Delta x$; поэтому для большей симметричности пишут: $\mathrm{d}y = f'(x)\mathrm{d}x$ с тем условием, чтобы приращение $\mathrm{d}x$ независимого переменного рассматривалось как постоянное количество, впрочем, вполне произвольное. "

Цитата 2 Л.Д.Кудрявцев [6] стр. 160: "Для большей симметрии записи дифференциала приращение Δx обозначают $\mathrm{d}x$ и называют его дифференциалом независимого переменного. Таким образом дифференциал можно записать в виде $\mathrm{d}y = A\mathrm{d}x$ "

Поразительно, но это "обоснование" изгнания Δx из жизни дифференциалов попало в статью Г.П. Толстова "Дифференциал" в Математической энциклопедии [8].

"Инвариантность первого дифференциала"

Цитата 3 Г.М. Фихтенгольц [4] п.106: "...Мы всегда имеем право писать дифференциал y в форме... < $y'dx$ - *Сушков В.И.* >, будет ли x независимой переменной или нет; разница лишь в том, что, если за независимую переменную выбрано t, то dx означает не произвольное приращение Δx, а дифференциал x как функции от t. Это свойство и называют инвариантностью формы дифференциала."

Цитата 4 Л.Д.Кудрявцев [6] стр. 177: "...(инвариантность формы первого дифференциала относительно преобразования независимой переменной) <...> ...дифференциал имеет один и тот же вид: произведение производной по некоторой переменной на "дифференциал этой переменной" - независимо от того, является эта переменная в свою очередь функцией или нет”

"Не инвариантность старших дифференциалов"

Цитата 5 Л.Д. Кудрявцев [6] на стр. 191 пишет, что формулы $d^n f = f^{(n)}(x)\, dx^n$

“..справедливы, вообще говоря, при n > 1 (в отличие от случая n = 1) только тогда, когда *x* является независимым переменным. В случае дифференциалов высших порядков по зависимым переменным дело обстоит сложнее”.

Поясняю. Формулу $f^{(n)}(x)dx^n$ Л.Д. Кудрявцев получил, следуя [1-4], т.е. дифференцированием по *x* первого дифференциала

$$d f = f'(x)\, \Delta x \qquad (1)$$

в котором положил Δx = const и сменил обозначение Δx на dx. Без этого искусственного переобозначения он был должен писать

$$d^n f = f^{(n)}(x)\, \Delta x^n \qquad (2)$$

Смысл утверждения (цитата 5): в (2) при n > 1 якобы нельзя полагать *x функцией.*
То же самое, но с оговоркой, что формулы $d^n f = f^{(n)}(x)dx^n$ верны и в том случае, когда *x* есть линейная функция $x = at + b$ написали В.И. Смирнов [3] и Г.Е. Шилов [10]. Появляющийся при дифференцировании *x* множитель a^n они предлагают переместить в состав n-ой производной сложной функции $f(at+b)$.

Цитата 6 Шилов Г.Е. [10] стр. 285: “Мы видим, что в отличие от первого дифференциала второй (и высшие) меняются, если независимое переменное становится зависимым, т.е. функцией от нового независимого переменного.“

Количество цитат из разных книг можно довести до десятков, но для анализа хватит этих.

РАЗБОР УТВЕРЖДЕНИЙ ИЗ ЦИТАТ

***Анализ цитат* 1, 2.** *Навсегда* заменив обозначения Δx на dx, а слова "приращение аргумента" - на "дифференциал свободной переменной", авторы очень большого числа учебников и задачников создают в умах читателей неразбериху и запрещают им под Δx понимать приращение *функции*. Я наоборот, приращение любой величины (и свободной переменной, и функции от иных переменных) всегда буду обозначать добавлением к её имени большой буквы "дельта".

***Анализ цитат* 3, 4.** На самом деле дифференциал функции *не всегда* равен произведению производной на *дифференциал* ее аргумента. Он может быть равен произведению производной на *приращение* аргумента функции, которое не равно *дифференциалу* ее аргумента. Такое встречается часто - когда рассматривают суперпозицию функций $f(g(x))$ и желают использовать ее разложение по степеням Δg, а не Δx.

Существуют формула первого дифференциала (1) из определения по Лагранжу и формула дифференциала сложной функции (суперпозиции двух функций) $y = f(x(t))$

$$d f = f'(x)\, dx \qquad (3)$$

Например, если $x = \ln(t)$ то $\Delta x = \ln(t+\Delta t) - \ln(t) = \ln(1+\Delta t / t)$, $dx = \Delta t / t$

и тогда формула (1) имеет вид $df = f'(\ln(t))\ln(1+\Delta t / t)$

а формула (3) имеет вид $df = f'(\ln(t))\, \Delta t / t$

Это - *разные* первые дифференциалы функции $f(\ln(t+\Delta t))$: первый - из разложения по степеням величины $\ln(1+\Delta t / t)$, второй - из разложения по степеням Δt.

При фиксированном t и стремлении Δt к нулю эти дифференциалы - *эквивалентные* бесконечно малые, но они не равны.

Ещё пример: разложение 2/(2-sin(t)) = 1+ sin(t)/2 + $(\sin(t)/2)^2$ + ...

Здесь sin(t)/2 есть дифференциал функции 2/(2-x) в точке 0 = sin(0).

В нём sin(t) – sin(0) есть *приращение* Δx функции x = sin(t), а не её дифференциал dx = cos(0)t = t. То есть вопреки [4-6] здесь дифференциал функции 2/(2-x) в точке 0 не равен произведению производной на *дифференциал* аргумента. Следующее слагаемое - деленный на 2 второй дифференциал, в котором - то же самое.

Формула $f\,'(x)\mathrm{d}x$ не может быть универсальным заменителем формулы $f\,'(x)\Delta x$ уже потому, что вторая допускает более широкий класс функций $x(t)$. Так, в формуле $f\,'(x)\Delta x$ функция $x(t)$ может быть недифференцируемой, при этом не существуют dx и $f\,'(x)\mathrm{d}x$. Пример: x - значение кусочно- гладкой функции $x(t)$ в точке t её излома, Δx - её приращение, порождённое числами t и Δt.

Что сделали авторы [4-6] заменой $f\,'(x)\Delta x$ на $f\,'(x)\mathrm{d}x$ *навсегда*? - Они создали искусственное препятствие использованию дифференциала, когда x и Δx - *функция* $x(t)$ и её *приращение*. В частности, они этим запретили суперпозицию рядов Тэйлора.

***Анализ цитаты* 5.** Выведем формулы старших дифференциалов сложной функции не дифференцированием как в [1-6], а алгебраически. Подставим в разложение

$$f(x+\Delta x) = f(x) + f\,'(x)\,\Delta x + f\,''(x)\,\Delta x^2/2! + \dots + f^{(n)}(x)\,\Delta x^n / n! + o(\Delta x^n) \qquad (4)$$

значение *функции* $x(t)$ и разложение *приращения* Δx в точке t по степеням Δt

$$\Delta x = \mathrm{d}\,x + \mathrm{d}^2 x/2! + \mathrm{d}^3 x/3! + \dots \mathrm{d}^n x/n! + o(\Delta t^n) \qquad (5)$$

Проделаем это для n = 3. При возведении Δx в степень отнесем слагаемые высших порядков (относительно Δt) в остаток в форме Пеано.

$(\Delta x)^2 = \mathrm{d}\,x^2 + \mathrm{d}\,x\,\mathrm{d}^2x + o(\Delta t^3)$; $(\Delta x)^3 = \mathrm{d}\,x^3 + o(\Delta t^3)$; $(\Delta x)^4 = o(\Delta t^3)$ и т.д.

Группируя в (4) члены по степеням Δt (показатель степени обозначен в дифференциалах $x(t)$), после выделения факториалов в знаменатель получим

$$\Delta f = f\,'(x)\mathrm{d}x + [f\,''(x)\mathrm{d}x^2 + f\,'(x)\mathrm{d}^2x]/2! + [f\,'''(x)\mathrm{d}x^3 + 3f\,''(x)\mathrm{d}x\mathrm{d}^2x + f\,'(x)\mathrm{d}^3x]/3! + o(\Delta t^3) \qquad (6)$$

- разложение функции $f(x(t))$ по степеням Δt (отклонения от t). Его общий вид

$$\Delta f = \mathrm{d}f + \mathrm{d}^2 f/2! + \mathrm{d}^3 f/3! + \dots \mathrm{d}^n f/n! + o(\Delta t^n) \qquad (7)$$

Сравнивая (6) и (7) на основе единственности разложения получаем формулы (3), (8), (9) для дифференциалов сложной функции $f(x(t))$.

$$\mathrm{d}^2 f = f\,''(x)\,\mathrm{d}x^2 + f\,'(x)\,\mathrm{d}^2x \qquad (8)$$

$$\mathrm{d}^3 f = f\,'''(x)\,\mathrm{d}x^3 + 3f\,''(x)\,\mathrm{d}x\,\mathrm{d}^2x + f\,'(x)\,\mathrm{d}^3x \qquad (9)$$

Аналогично получают формулы более высокого порядка. Общей формулы подобно (2) здесь нет. Теперь - анализ цитаты 5 в свете *такого* вывода формул (8,9).

1) Формулы (8, 9), которыми [4-6] требуют заменять (2) ради избавления от подстановки *функции* $x(t)$ в (2), - именно такой подстановкой и получены. Значит, требуя отказаться от этой подстановки в задачах, авторы [4-6] требуют делать *её же*, но только в общем виде. - Нелепое требование (см. след пункт). Вот если бы они написали, что подстановку бесполезно делать в отдельных дифференциалах и полезно её делать в дифференциалах, стоящих в формуле Тэйлора...

2) В случае когда x - *функция*, в [4-6] объявлен обязательным (цитата 5) *отказ* от (2) в пользу формул старших дифференциалов сложной функции (8), (9) и т.д. - Но сами авторы [4-6] ему не следуют. Например: для вычисления некоторых пределов в задачнике Л.Д. Кудрявцева [7] нужны степенные разложения 5 порядка для суперпозиции трёх функций. Формула Тэйлора в таком случае состоит из 47 слагаемых, в которые входят значения этих трёх функций и их первых пяти производных, - она необозрима. Её слагаемые (без факториалов в знаменателях) и есть старшие дифференциалы сложной функции. Настаивая на их использовании, Л.Д. Кудрявцев настаивает на использовании этой формулы - монстра. Однако в задачнике [7], решая образцы задач, он не пользуется ею. Он подставляет разложения Маклорена с числовыми коэффициентами в таковые же, т.е. подставляет в дифференциалы (2) функцию $x(t)$ и её приращение Δx в виде степенного разложения. Таким образом в задачнике [7] Л.Д. Кудрявцев учит не обращать внимания на свои слова (цитата 5) и их обоснования в учебнике [6]. - Разве это не нелепость?

То же самое делают *все* сторонники "не инвариантности старших дифференциалов", - экстренно забывают о ней, когда создают суперпозицию двух рядов Тэйлора.

3) Благодаря *алгебраическому* выводу формул дифференциалов сложной функции (3), (8), (9) выясняется причина появления сумм в формулах (8,9). Вместо того, чтобы ясно её обрисовать, или хотя бы сформулировать в виде предупреждения о невозможности получить старший член разложения (7) из *одного* члена из (4), одинакового с ним порядка, зачем-то вообще запретили всякую подстановку в (2) функции вместо x. Да ещё и выдали *свой* запрет за свойство старших дифференциалов.

Итак, вопреки [4-6] формулы старших дифференциалов (2) верны и тогда, когда x - функция, и люди это используют 300 лет и это используют сами авторы [4-6].

Теперь прошу обратить внимание: "инвариантность дифференциала" (3) и "не инвариантность дифференциала" (2) - *одно и то же* потому, что обе они - всего лишь противоестественный *запрет* подставлять вместо x функцию в формулу дифференциала (2). Между ними разница лишь в том, чему равно n. Если n = 1 то запрет называют "инвариантностью". Если n > 1 то тот самый же запрет называют "неинвариантностью". - Разве это не нелепо?

***Анализ цитаты* 6.** Стандартное доказательство [2] неинвариантности второго дифференциала таково. Присвоив себе право называть *приращение* аргумента *дифференциалом свободной* переменной , т.е. наложив табу на "Δx - приращение *функции*", дифференцируют (3) т.е. $f\,'(x)\mathrm{d}x$ - как якобы единственного представителя понятия "дифференциал":

$$\mathrm{d}^2 f = \mathrm{d}(f'(x)\,\mathrm{d}x) = \mathrm{d}(f'(x))\,\mathrm{d}x + f'(x)\mathrm{d}^2x = f''(x)\,\mathrm{d}x^2 + f'(x)\,\mathrm{d}^2x \qquad (10)$$

После этого заявляют: полученное отличается от $f''(x)\mathrm{d}x^2$ наличием второго слагаемого. И это якобы означает, что второй дифференциал изменился.

Но в $f\,''(x)\mathrm{d}x^2$ буквами $\mathrm{d}x$ было обозначено *приращение* аргумента, а в полученном выражении $\mathrm{d}x$ - *дифференциал*. Сами себя запутали! Отличие не в одном, а в *обоих* слагаемых. Теперь осталось заметить: из того, что 3 не равно 8 и 5 не равно 8 не следует, что 3+5 не равно 8, - и доказательство рухнуло.

Но у этой "теоремы" есть дефекты намного хуже. Эти два вторых дифференциала изначально разные, потому что они из разных разложений (по степеням Δx и по степеням Δt). Потому из факта их неравенства бессмысленно делать выводы о другой причине, тоже обеспечивающей его. Теорема априори не имеет смысла.

А вот ещё дефект. Невозможность подставлять функцию $x(t)$ и её приращение в (2) при n = 2 в этой "теореме" получена после запрещения того же самого для случая n = 1, из которого и получены формулы старших дифференциалов (2). Что заложили в условие, то и получили. - Разве так принято доказывать теоремы?

Не более осмыслена и другая версия этой "теоремы", - как утверждение о не инвариантности *формы* второго дифференциала. Например, (8) на самом деле

$$\mathrm{d}^2 f = f''(x)\,\mathrm{d}x^2 + f'(x)\,\mathrm{d}^2x = f''(x(t))\,(x'(t)\Delta t)^2 + f'(x(t))\,x''(t)\,\Delta t^2 =$$
$$= [f''(x(t))\,(x'(t))^2 + f'(x(t))\,x''(t)]\,\Delta t^2 = y''(t)\,\Delta t^2\,, \quad \text{где } y = f(x(t)) \qquad (11)$$

имеет вид (2) при n = 2. Непонятно, что мешало В.И. Смирнову и Г.Е. Шилову увидеть, что старшие дифференциалы сложной функции имеют (как и полагается) структуру формул (2) не только тогда, когда x - линейная функция от t но и тогда, когда x - любая достаточное число раз дифференцируемая функция.

Ну и наконец, от теоремы не останется никакого следа, если отменить требование называть приращение аргумента *не приращением*, а *дифференциалом* свободной переменной.

Тогда формула $f''(x)\mathrm{d}x^2$ вернёт себе свой законный вид $f''(x)\Delta x^2$ и исчезнет причина для путаницы её с формулой $f''(x)\mathrm{d}x^2 + f'(x)\mathrm{d}^2x$ и слов о том, что дифференциал *меняется*.

Да не меняется он! При подстановке в дифференциал достаточное число раз дифференцируемой функции $x(t)$ может произойти разве лишь сужение множества его значений за счёт ограниченности множеств значений и приращений функции $x(t)$. Разве синус перестаёт быть синусом, если мы подставляем то или иное значение угла? Точно так же дифференциал (2) не перестаёт быть дифференциалом от того, что в него подставили значение функции $x(t)$ и её приращения Δx. Сама постановка вопроса неверна. Речь должна идти не об одном дифференциале, а о двух *разных* дифференциалах, из двух *разных* степенных разложений.

Для чего же в итоге нужен термин "неинвариантность"? Получается, только для того, чтобы отметить очень глубокое отличие $8\Delta t^2$ от $(3+5)\Delta t^2$.

С "теоремой" о неинвариантности второго дифференциала из [4-6] мы разобрались: её нет в природе. Но "инвариантность" первого дифференциала "доказывается" еще интереснее. В.А. Ильин и Э.Г. Позняк [5] вводят *договорённость* называть приращение "дифференциалом свободной переменной". Через несколько страниц об этой *договорённости* они пишут "В конце §2 мы установили...". Заменив в сознании читателя *договорённость* на *установленный* факт, затем пишут "мы доказали". - Ну разве это математика?! Чему мы студентов учим такими "доказательствами"?

ИСТОЧНИКИ ОШИБОК В УЧЕБНИКАХ [3-6]

<u>Первый источник ошибки</u> - самообман на почве изгнания Δx (цитаты 1,2).

Для дифференциалов (1) и (3) авторы [3-6] имеют по два варианта: либо x - независимая величина, либо зависимая. Итого четыре. Два из них совпадают: когда x - независимая (первичная) величина - формулы (1) и (3) одинаковы. Итого *три* случая.

Изгнание Δx (цитата 1,2) - это *манипуляция*: ею из поля зрения убирают один из *трёх* случаев и переводят внимание на совпадение обозначений в *двух* случаях

1) когда x - *свободная* (первичная) переменная, формулу (1) по договорённости (цитаты 1,2) пишут в виде (3);

2) когда x - *функция*, формулу (3) получают по теореме о производной сложной функции.

А дальше пишут, что дифференциал в *обоих* случаях - и когда x независимая и когда зависимая переменная, имеет вид (3). Поскольку состояний быть зависимой или независимой величиной и вправду два, создаётся *впечатление*, будто формула (3) справедлива во всех мыслимых случаях, то есть *всегда* (цитата 3).

Но случаев-то *три*, а не два! И потому *не всегда* дифференциал равен произведению производной на *дифференциал* аргумента (см. выше). Авторов [3-6] несколько оправдывает лишь то, что они (безрассудно) следовали Коши.

<u>Второй источник ошибки</u> - неверное употребление понятия "зависимая величина". У Коши [1] этих проблем не было, потому что он вводил дифференциал как предел дроби $(f(x+\mathrm{h}t)-f(x))/t$ при стремлении t к нулю и $\mathrm{h} = \mathrm{const} = \mathrm{d}x$. Т.е. у него дифференциал $f'(x)\,\Delta x$ был функцией *одного* аргумента x. У него величина x - *"свободная"*, но вот в каком смысле: это величина, через которую выражаются *все* остальные, кроме $\mathrm{d}x$. Правильнее было бы назвать её *первичной* переменной; в терминах теории многообразий это - точка карты. В [1,2,3] *свободные* (*первичные*) *переменные* те, по степеням приращений которых авторы хотят получить разложение. Деформация смысла в [4] произошла, видимо, по цепочке "несвободная" - "зависимая" - "функция". В итоге невинное замечание Гурса [2] о том, что (8) имеет не такой вид как (2) при $n = 2$, было расширено до абсурда: выбор формул (8,9) вместо (2) авторы [5, 6] подают не следствием *нашего желания* иметь разложение по степеням Δt вместо Δx, а неким свойством дифференциала (2) при $n > 1$ зависеть от того, является ли x функцией.

Если бы дифференциал (2) при n > 1, как средство обработки двух чисел x и Δx был живым существом, то по [5, 6] наше общение с ним выглядело бы так. Мы просим его обработать данные: $f(x) = \sin(x)$, точка $x = 0$, приращение $\Delta x = 0.5$. А он нас подозрительно спрашивает: "А вы откуда эти числа 0 и 0.5 взяли? *Сами* их задали, по вашей *доброй воле*, или вам их *кто-то дал*?" Мы ему: "А в чем дело?" А он нам: "А в том, что если эти числа *не свободное* проявление вашей воли, получены от какой-то функции, - я их обрабатывать не буду". - Вот до чего довели старшие дифференциалы в [5, 6].

Начиная с [4] дифференциал объявили функцией *двух* аргументов x и Δx. Объявить-то объявили, а рассмотреть его как функцию двух аргументов забыли: замена переменных в дифференциале $d^n f = f^{(n)}(x)\,\Delta x^n$, - в учебниках не рассмотрена. Восполним этот пробел.

Если x - *функция* то имеем формулы замены *пары* чисел $(x,\Delta x)$ на *пару* $(t,\ \Delta t)$ (или обратно - в зависимости от вида задачи)

$$x = g(t); \quad \Delta x = g(t+\Delta t) - g(t) \tag{12}$$

Можно ли в такой ситуации в общем случае утверждать, что благодаря тому, что x - функция, между x и Δx появилась функциональная зависимость?

- Нельзя. Например: $x = \ln(t)$. При любом заданном значении $x = \ln(t)$ мы можем обеспечить для $\Delta x = \ln(t+\Delta t) - \ln(t)$ любое значение, выбрав $\Delta t = \exp(x)(\exp(\Delta x)-1)$. И наоборот, при любом значении Δx мы можем пропорционально менять t и Δt, тогда $\Delta x = \ln(1+\Delta t/t)$ сохранится, а $x = \ln(t)$ получит желаемое значение. Т.е. x и Δx *зависимы* от $(t,\Delta t)$, они *функции*, мы даём им значения с помощью чисел $(t,\Delta t)$, но они независимы *друг от друга*.

Поэтому и в том случае, когда x - функция, можно дифференцированием (1) по x получать (2), следуя лекциям Коши [1] (хотя сам он эту возможность не рассмотрел). В терминологии Коши переменную x - можно назначить *свободной*, хотя она - *функция* от иной величины t (тоже *свободной*, в терминах многообразий мы имеем две эквивалентные карты для одного многообразия - ось x и ось t). При построении разложения функции $f(x+\Delta x)$ по степеням Δx то обстоятельство, что x - функция, не играет никакой роли. Видимо, этого не осознал Коши [1]. Возможно, ему казалось, что Δx=const есть необходимое условие того, чтобы с ним обращаться как с константой при выводе (2) дифференцированием (1) по x. Так или иначе, но он не рассмотрел случай "x - функция, Δx - её приращение" в (1). А вместе с ним его упустили и его последователи [2-6, 10]. Упустили потому, что вместо замены (12) *двух* величин $(x,\ \Delta x)$ на *две* $(t,\Delta t)$ они вслед за Коши пишут про замену *одной* переменной x на функцию $x(t)$. В результате оказались совмещены в одно два разных обстоятельства: зависимость x и Δx *между собой* и зависимость x *от иных*, внешних переменных (цитата 6). У них получилось, что если величина x - значение функции, то она якобы не может иметь назначаемые нами ей значения, якобы не может быть свободной переменной.

<u>Третий источник ошибки.</u> Несомненно, Коши хотел вывести обозначение Лейбница dy/dx для производной. Для этого ему был нужен "дифференциал свободной переменной". А как только он заменил Δx на dx - сразу потерял случай "Δx - приращение функции". Повторяющие за Коши эту процедуру изгнания Δx - получают то же самое.

Теперь про вывод (10). В этом случае

$$x=g(t);\ dx = g'(t)\Delta t \tag{13}$$

Случай отличается от предыдущего тем, что мы хотим иметь разложение по степеням Δt а не dx. Право считать dx функцией вытекает не просто из того, что x - функция (как вынуждены думать читатели учебников), а ещё из того, что нам нужно варьировать t при фиксированном Δt, т.е. здесь первичны (свободны) t и Δt.

Авторы [4-6], видимо, без раздумий копировали предшественников, внеся "небольшую" поправку: первый дифференциал (1) есть функция *двух* переменных, а не одной. То что за этой поправкой следует, они не рассмотрели.

ЧТОБЫ ОГЛЯДЕТЬ СИТУАЦИЮ В ЦЕЛОМ, ВСПОМНИМ ИСТОРИЮ

Я буду опираться на [9].

В обличье, используемом всеми нами, первый дифференциал был создан в 1670-х годах Г. Лейбницем, мечтавшим видеть математику основой миропонимания [12], а алгебру - основой математики. "Не будем спорить, будем вычислять!" - вот его идеал научных рассуждений. *Обозначения* дифференциалов и интегралов - плод его трудов на *этой* ниве и вечный ему памятник. Они воплощают его желание видеть алгебру. Дифференциал был задуман им как *бесконечно малая* (в его понимании) разница (differentia) абсцисс ($\mathrm{d}x$) или ординат ($\mathrm{d}y$) двух сближаемых точек графика функции. У Лейбница производная равна $\mathrm{d}y/\mathrm{d}x$ не потому, что он так обозначает вычисленный им предел дроби $\Delta y/\Delta x$, а потому что у него и числитель и знаменатель уже являются катетами исчезающе малого треугольника, гипотенуза которого - и касательная и секущая одновременно. Грубо говоря, у Лейбница, как в современной дифференциальной геометрии, $y' = \mathrm{d}y/\mathrm{d}x$ просто потому, что ($\mathrm{d}x$, $\mathrm{d}y$) есть точка касательной прямой. Именно в этом и заключается притягательность для нас обозначений Лейбница - они нам напоминают о б.м. треугольнике и дают убежденность в праве так и рассматривать и секущую (как гипотенузу), и касательную (как ту же самую гипотенузу), и элемент длины дуги (длину гипотенузы). А в обозначении интеграла дифференциал напоминает о бесконечно узком прямоугольнике из Римановой суммы. Интересно отметить, что и сам Лейбниц не сразу пришёл к идее обозначать интеграл как сумму $f(x)\mathrm{d}x$, сначала он пытался суммировать ординаты, то есть в его обозначении интеграла сначала не было $\mathrm{d}x$.

Обозначения Лейбница обеспечивают преемственность в развитии нашего мышления от школьной геометрии прямых, треугольников и прямоугольников к кривым линиям и фигурам, разбитым на мелкие части, обслуживаемые геометрией прямых линий.

У Лейбница как и у Ньютона, бесконечно малые имеют не тот смысл, как в наших учебниках анализа, не тот, который нам дал Коши. У Лейбница это чрезвычайно малые (исчезающе малые) *числа*, равные нулю в суммах, но не равные нулю в числителях и знаменателях дробей. Такое их свойство противоречит правилам арифметики, хотя и порождено *практикой* приближённых вычислений. Лейбниц в задаче о построении касательной к графику функции довёл особенности приближённых вычислений до, так сказать, предельного состояния, и каждый, знакомый с азами философии, скажет, что это вызывает качественное изменение. Указанное отличие от правил арифметики и есть это качественное следствие, плата за идеализацию. Сегодня его можно видеть в теории гипервещественных чисел, получаемых добавлением к вещественным числам некоего особого числа $\omega > 0$, обладающего свойством быть меньше любого положительного вещественного числа. На этой основе идея Лейбница воплощена Дэвисом в 1960-х годах в "нестандартном анализе". Последний был опробован в преподавании примерно в трети американских колледжей и показал свою бесполезность. - А чего ещё следовало ожидать от методики, отрывающей анализ бесконечно малых от его источника - приближённых вычислений, - и возрождающей в умах молодёжи те же сомнения и неясности, которые терзали математиков в первой половине XVIII века? Логическая непротиворечивость и понятность - не одно и то же!

Противоречие исчисления бесконечно малых обычной арифметике вызвало острую критику со стороны учёных XVIII века. Епископ Беркли в обширной публикации высмеял материалистов, пользовавшихся в новом исчислении идеальными, нематериальными средствами. Но и без критики сторонники исчисления бесконечно малых стремились избавить основания исчисления от неясностей и противоречий.

Другое понимание дифференциала, свободное от Лейбницевских бесконечно малых, пришло из теории степенных разложений. В 1715 году Тэйлор выяснил, во что превращается интерполяционный полином Ньютона, когда узлы интерполяции неограниченно сближаются, а их количество неограниченно растёт. В XVIII веке выкристаллизовался аппарат производных и благодаря этому пришло понимание исключительной важности формулы Тэйлора (4), как раз и выражающей собой идею превращения анализа бесконечно малых в алгебру (многочленов). Именно по этой причине имя Тэйлора навсегда осталось в науке, хотя формула частной суммы ряда Тэйлора встречается в существенно более ранних записках Ньютона.

В те времена всё больше исследователей приходило к мысли, что числители членов формулы Тэйлора и есть дифференциалы, но освобождённые от экстравагантного свойства быть и равными и неравными нулю. Эйлер в 1755 году использовал для членов степенного разложения обозначения дифференциалов. Арбогаст в 1772 году прямо назвал члены разложения дифференциалами. Лагранж в 1789 году ввёл производную как коэффициент главной, линейной по приращению аргумента, части приращения функции (т.е. определение, которым пользуемся все мы). Для коэффициентов степенного разложения он ввёл нумерацию штрихами и назвал их "производными функциями" (function dérivée) за то, что каждая из них выводилась из предыдущий (в итоге - из f) одним и тем же способом - как множитель в главной линейной части приращения предыдущей. Всё это он делал в надежде изгнать бесконечно малые и пределы вообще. Первые он изгнал, а вторые упустил (сходимость ряда).

Итак, понятия дифференциалов (2) и производных появились во второй половине XVII века, но оформились только к концу XVIII века в теории степенных рядов. В результате возникла двойственность, унаследованная и нами:

с одной стороны - удобные обозначения Лейбница для дифференциалов, производных и интегралов, апеллирующие к малости приращений и дифференциалов, к привычной нам евклидовой геометрии и к нашим привычкам работы с дробями;

с другой стороны - Лагранжево понимание дифференциалов, как функций двух переменных, не обязанных быть малыми членов Тэйлоровского степенного разложения.

Всякий раз, когда мы пишем интеграл, или обозначаем производную как отношение $\mathrm{d}y/\mathrm{d}x$, или пишем правило дифференцирования сложной функции в виде сокращения дроби $(\mathrm{d}y/\mathrm{d}x)(\mathrm{d}x/\mathrm{d}t)=\mathrm{d}y/\mathrm{d}t$, - мы используем Лейбницевское понимание дифференциала как *исчезающе малой* величины вместе с упомянутым противоречием её использования правилам арифметики. Идея малости дифференциала проникает в сознание современников, знакомых хотя бы с интегралами: им *кажется*, что дифференциал обязан быть бесконечно малой. Таково свойство обозначений Лейбница!

Всякий раз, когда мы пользуемся степенными разложениями, мы используем идею Лагранжа. Всякий раз, когда мы рассматриваем точки касательной, далёкие от точки касания, мы наделяем $\mathrm{d}y$ и $\mathrm{d}x$ свойствами дифференциала Лагранжа.

Мы не можем отказаться от одного в пользу другого *везде*! А совместить Лагранжев и Лейбницев подход в рамках теории вещественных чисел невозможно. Ведь для этого надо распространить теорему о пределе дроби на случай 0/0, это позволяют гипервещественные числа, упомянутые выше, бесполезные для преподавания дифференциального исчисления.

В XIX веке этого не знали. Естественное желание избавить обозначения Лейбница от противоречий, - как часть процесса систематизации накопленных аналитических знаний, - привело в 1820-х годах к появлению лекций А. Коши [1]. Последователям Коши *казалось* (возможно до сих пор *кажется* некоторым членам Научно - методического совета по математике), что он совместил дифференциал по Лагранжу с дифференциалом по Лейбницу. Как у него получилась видимость такого совмещения - я написал выше (12). Есть задачи, в которых между Лагранжевым и Лейбницевским подходами не возникает противоречий и в них действительно можно положить, что $\Delta x= \mathrm{d}x$. Но сделать такое навсегда означает сузить круг приложений дифференциального исчисления.

У нас нет иного выхода, кроме как рассказывать всё это студентам, ничего не скрывая.

ЗАКЛЮЧЕНИЕ

1) "Инвариантность" первого дифференциала в смысле [4-6], т.е. как замена (1) на (3) *навсегда,* - не существует. Такая замена допустима лишь в рамках некоторых задач, например - в задачах внутренней геометрии многообразий (в частности в ОТО), при условии, что карта многообразия не меняется. Провозглашение "инвариантности" в смысле [4-6] сужает круг задач, обслуживаемых дифференциальным исчислением.

2) "Неинвариантность" старших дифференциалов (2) в смысле [4-6] тоже не существует: вопреки [4-6] в них можно заменять x функцией, а Δx - её приращением и это делают 300 лет. Выбор между (2) и формулами дифференциалов сложной функции (8,9) обусловлен не неким свойством "не инвариантности" формулы (2) при n>1, а нашим желанием иметь разложение по степеням Δx или Δt в случае $x = x(t)$.

3) Все формулы дифференциалов - и первого и старших и для сложной функции - инвариантны относительно того, является ли x функцией или нет. В учебниках [4-6] вопрос о связи между собой *разных* дифференциалов, членов (одинакового порядка) разложений функции по степеням разных величин, выдают за вопрос о свойствах *одного* дифференциала. Причина в том, что первый дифференциал в них дают по Лагранжу, а старшие - по Коши. При таком подходе старшие дифференциалы теряют связь со степенным разложением и кажутся лишь итогом формального дифференцирования.

4) Нельзя путать свойство величины быть функцией с невозможностью свободно задавать её значения. Эту путаницу насаждают в [4-6].

5) Формулы дифференциалов сложной функции (3,8,9) *не являются* более общими, чем (2), потому что (2) применимы к более широкому классу функций $x(t)$.

6) Методика [1-6] приучает к мысли, что показатель в обозначении дифференциала - количество *дифференцирований*. А вот главное, - то, что это показатель *степени* Δx, - оказывается в тени. Этим между анализом и алгеброй (которой он должен быть по замыслу Лейбница) в сознании студентов авторами [4-6] воздвигнута стена.

7) Насилия над Δx в [4-6] проистекли из унаследованного от [1-3] (вместе со схемой изложения) желания обосновать Лейбницевское обозначение производной $y' = \mathrm{d}y/\mathrm{d}x$. Лагранжево определение дифференциала даёт другое обозначение $y' = \mathrm{d}y/\Delta x$. Сказанное не означает, что первая формула должна уступить место второй. Сказанное означает, что всегда надо помнить, из какого разложения взят дифференциал функции. Так, в этих двух *верных* формулах разнятся не только знаменатели, но и числители.

8) Пользоваться Лейбницевскими обозначениями производных и интегралов мы будем, но мы должны сказать студентам об их условности. Мы должны объяснить, что радикально исправить обозначения вряд ли удастся, потому что в обозначение производной или дифференциала невозможно вместить всю информацию о породившем их полиноме Тэйлора. Проще написать сам полином. Надо не пытаться вслед за Коши из определения Лагранжа выводить обозначения Лейбница, невыводимые в R. - Надо рассказать о мечте Готфрида Лейбница и о 300-летней традиции его обозначений, хоть и противоречивых, но удобных в классических задачах.

9) Мы обязаны объяснять студентам, что одних лишь правил действий с производными недостаточно для корректных действий с дифференциалами. Ещё нужно помнить, из какого разложения взят дифференциал. Все действия с ним оправданы лишь тогда, когда они есть часть действий с разложениями Тэйлора. Методика Коши [1], кладущая производную в основание дифференциального исчисления, не объясняет этого, потому она рождает грубые ошибки и беспомощность.

10) Мы должны вернуть Δx в учебники и задачники, ликвидировать жуткую неразбериху, порождаемую заменой обозначения Δx на $\mathrm{d}x$ *навсегда*. В учебники и в Примерную программу надо ввести вопрос о замене *двух* переменных в дифференциалах (2) и (3,8,9)

11) Вместо "инвариантности" мы должны рассказывать, что (3) является воплощением гомоморфизма "суперпозиции функций отвечает суперпозиция их производных отображений": если $f = f(x(t))$ и $\Delta f = \mathrm{A}\Delta x + \mathrm{o}(\Delta x)$ и $\Delta x = \mathrm{B}\Delta t + \mathrm{o}(\Delta t)$, то $\Delta f = \mathrm{AB}\Delta t + \mathrm{o}(\Delta t)$. Отсюда - шаг до понимания, что (4) есть $f(\mathrm{x}+\Delta x) = \exp(\mathrm{d}) f(x)$. "Инвариантность первого дифференциала" в смысле [4-6] не только не существует, но и не нужна, т.к. на самом деле вместо неё работает указанный гомоморфизм.

12) Самое раннее употребление термина "инвариантность первого дифференциала " в учебниках анализа, найденное мной, - у Г.М. Фихтенгольца и Куранта (конец 1940-х). Кто у кого заимствовал - судите сами. Этот термин характерен для тензорного анализа, для ОТО. Видимо, он был введён - искусственно - в курс дифференциального исчисления в кампанию прославления А. Эйнштейна. [*]

13) Рассказ о дифференциалах в [4-6] неудовлетворителен. В случаях включения вопроса "инвариантность первого дифференциала" в экзамен мы можем говорить о принуждении студентов следовать [4-6] в теме "дифференциалы", запоминать и признавать математикой запутанные в клубок нелепости. Этот вопрос должен быть изъят из Примерной программы и курсов лекций.

14) 18 лет назад я пытался создать курс XXI века [11]. Испробовал в лекциях идею Пеано вводить производные как коэффициенты степенного разложения. Но уже тогда систему образования разъела толерантность.

15) В 1990-х годах многие преподаватели математики ВУЗов России, чувствуя, что "здесь что-то не так", не сговариваясь, предали забвению "инвариантность первого дифференциала" и "не инвариантность старших дифференциалов". Но в 2000 году "реформаторы" образования вновь вытащили на свет "инвариантность первого дифференциала" и под руководством Л.Д. Кудрявцева включили её в "Примерную программу дисциплины Математика". Поскольку эти понятия лекторы берут из самых массовых учебников [4-6], которым ещё в СССР создали репутацию эталонных, - автоматически ВУЗам России оказалась навязана и методика изложения дифференциалов и формулы Тэйлора из этих учебников. Странная ситуация: огромная страна в обучении молодёжи дифференциальному исчислению оказалась вынуждена следовать *личным привычкам* и даже *заблуждениям* четверых человек из Научно-методического совета.

16) 60 лет студентам морочат голову "инвариантностью" в теме "дифференциалы", - и не видно тому конца. Почему? - А руководители озабочены личным благополучием и потому заняты выгодным делом - уничтожением образования под разговоры о "реформах" и "инновациях". Под давлением "реформаторов" курс математики сжат до профанации, до абсурда! За 14 занятий одного семестра нас обязали научить студентов функциям нескольких переменных, рядам, обыкновенным дифференциальным уравнениям и двойным интегралам. А некоторые факультеты требуют сюда еще и ТФКП уместить! Система управления и воспитания кадров сгнила и работает только на саму себя. Множатся фиктивные диссертации. Наверх всплывают отнюдь не сливки. Методическая работа убита. Нет связи между качеством преподавания и зарплатой. Растёт доля фиктивных положительных оценок, местами она зашкалила за 75%. Госвузы уже торгуют дипломами как коммерческие. Не то чтобы нет стимулов к повышению качества - есть мощнейшие стимулы к его снижению. Старые преподаватели без учёной степени, обладающие неоценимым опытом и моральными принципами, ввергнуты в нищету (зарплата старшего преподавателя - 3196 рублей 58 копеек при официальном прожиточном минимуме 4100 рублей и средней по Петербургу 15000 рублей). По выражению одной мамаши (сынок - банкир и наркоман), они нас "выдавливают как пасту из тюбика" из наших квартир, из наших городов, из образования, из жизни. Ничего иного они не умеют. Это их способ жить.

---

* Примечание: уже после публикации статьи доцент И.В. Бурков указал мне, что первое упоминание "инвариантности первого дифференциала", возможно, принадлежит Поссе.

ЛИТЕРАТУРА

1. Дифференціальное и интегральное изчисленіе. Краткое изложеніе уроковъ, преподаваемыхъ въ Королевской политехнической школђ Г. *А.Л. Коши*. (Перевод В.Я.Буняковского). - СПб, 1831 г.

2. Гурса Э. Курс математического анализа. Том 1. - М.: ОНТИ НКТП СССР, 1936. (Издание 3, тираж 7000).

3. Смирнов В.И. Курс высшей математики. Том 1. - М.: Наука, 1974 (Издание 23, тираж 64000).

4. Фихтенгольц Г.М. Курс дифференциального и интегрального исчисления. Том 1. - М.: Наука, 1966. (Издание 4, тираж 75000)

5. Ильин В.А., Позняк Э.Г. Основы математического анализа. Часть 1. - М.: Наука, 1971. (Издание 3, тираж 97000)

6. Л.Д. Кудрявцев Курс математического анализа 1-й том. - М.: "Высшая школа", 1981. (Тираж 80000).

7. Кудрявцев Л.Д., Кутасов А.Д., Чехлов В.И., Шабунин М.И. Сборник задач по математическому анализу. Предел. Непрерывность. Дифференцируемость. - М.: Наука, 1984.

8. Математическая энциклопедия. Том 2. - М.: "Советская энциклопедия", 1979.

9. История математики с древнейших времен до начала XIX столетия. Под ред. А.П. Юшкевича. Том 3. Математика XVIII столетия. - М.: Наука, 1972.

10. Шилов Г.Е. Математический анализ. Функции одного переменного. Часть 1 и 2. - М.: Наука, 1969 (тираж 75000).

11. Сушков. В.И. Общий курс высшей математики начала XXI века. - Математика в ВУЗе №3, 2002 г. Интернет-журнал на сайте СПбГПУ. - http://www.spbstu.ru/public/m_v/N_003/Sushkov/Cours.XXI/plan.html

12. Избранные отрывки из математических сочинений Лейбница (перевод А.П.Юшкевича)// УМН, 1948, том III, вып. 1 (23), стр.165 - 173