Кривая через точки

Л.АШКИНАЗИ

Пусть в ходе эксперимента при двух значениях x_1 и x_2 некого параметра x получены два значения y_1 и y_2 функции y. Например, при двух значениях тока I, протекающего через некое сопротивление, на нем оказываются два значения напряжения U. Если бы вы знали, какая бездна вопросов начинается с этих четырех цифр, вы бы в ужасе бросили читать и убежали прочь. Но поскольку вы не убежали, то вот и задача – постройте зависимость $y(x)$.

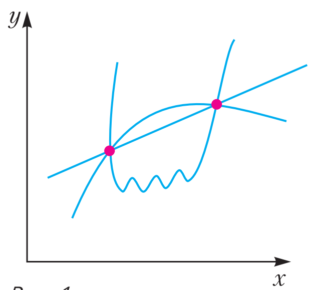

Рис. 1

В такой постановке задачу решить нельзя – у нее бесконечно много решений (рис. 1). Школьник, скорее всего, изобразит в некотором смысле простейшее решение – линейную функцию. Тут самое время поинтересоваться, чему у него будет равно $y(x=0)$ если оно окажется не равно нулю, то как бы к этому отнесся уважаемый Георг Ом (1789–1854) – с учетом того, что у нас это не какие-то x и y, а конкретно I и U (рис. 2)?

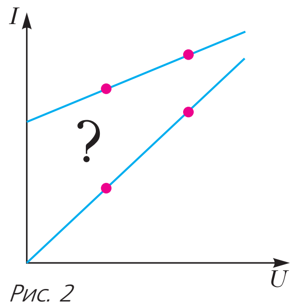

Рис. 2

Впрочем, с нами могли поступить и интереснее – изобразить одну точку, задать вопрос и посмотреть, догадаемся ли мы использовать в качестве второй точки начало координат (рис. 3).

Что же касается прямой, не проходящей через ноль, то у нас есть три выхода. Можно спросить, с какой погрешностью измерены значения. Другой вариант – сказать, что у нас нелинейное сопротивление. И, наконец, заявить, что это не простое сопротивление, даже не нелиней-

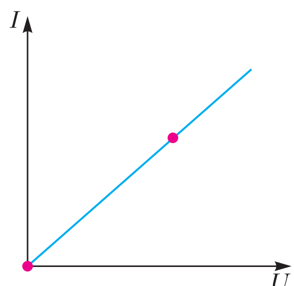

Рис. 3

DOI: https://doi.org/10.4213/kvant20230205

ное, а волшебное, т.е. линейное, но с вмонтированном внутрь него маленькой батарейкой или аккумулятором.

С последнего варианта – как самого простого – и начнем. Нарисуем схему, в которой к источнику напряжения подключено не простое и честное сопротивление, а волшебное – коробочка, внутри которой находятся последовательно соединенные обычное честное линейное сопротивление и другой источник ЭДС с напряжением U_x (рис. 4). Вы

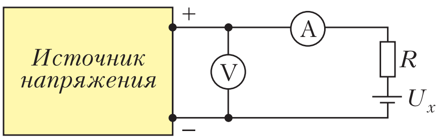

Рис. 4

можете, не решая уравнений, утверждать, что проблема решена – наклон графика зависимости $I(U)$, который определяется величиной R, остался тем же, а изменение величины U_x позволяет перемещать график параллельно самому себе. Но экзаменатор, скорее всего, захочет уравнений. Ну, так напишите их прямо сейчас и получите выражения для R и U_x через U_1, U_2, I_1, I_2 (рис. 5).

$$\begin{cases} I_1 = \dfrac{U_1 - U_x}{R} \\ I_2 = \dfrac{U_2 - U_x}{R} \\ I(U=0) = \dfrac{U_x}{R} \end{cases} \Rightarrow \quad \begin{aligned} R &= \frac{U_2 - U_1}{I_2 - I_1} \\ U_x &= \frac{I_1 U_2 - I_2 U_1}{I_1 - I_2} \end{aligned}$$

Рис. 5

Пусть экзаменатор обрадуется вашему волшебному сопротивлению.

Но тут есть опасность, и, кстати, при общении с умным экзаменатором она есть всегда. Экзаменатор, которому понравится ваша идея волшебного сопротивления, может предложить другой вариант этого чуда – не последовательно, а параллельно соединенные дополнительный источник ЭДС и обычное сопротивление. Рассмотрите этот вариант сами.

Теперь обратимся к более традиционному случаю – к нелинейному сопротивлению. В школьном курсе такие сопротивления упоминаются, иногда даже приводятся примеры. Это и полупроводниковый диод, и вакуумный диод, и газовый разряд. Поскольку теорий этих приборов в школьном курсе нет, то сказать, можно ли через три заданные точки, т.е. (x_1, y_1), (x_2, y_2), $(0, y(0))$, провести вольт-амперную характеристику какого-то из этих приборов, в этой статье было бы трудно. Но есть промежуточная, вполне школьная задача – вам дают вольт-амперную характеристику какого-то нелинейного сопротивления и предлагают из него и обычных линейных сопротивлений скомбинировать схему и получить вольт-амперную характеристику, проходящую через эти три заданные точки. Кстати, вы можете рассмотреть эту задачу, если в качестве нелинейного элемента вам дали идеализированный полупроводниковый диод. Подсказка: посмотрите, какие вольт-амперные характеристики вы можете получить, соединяя диод и сопротивление последовательно и параллельно.

Что касается нелинейных сопротивлений, полезно понимать еще вот что. Нелинейность сопротивления, т.е. его неподчинение закону Ома, может быть нескольких типов. Например, она может быть связана со свойствами самой среды, по которой протекает ток, – это случай полупроводникового диода. Но возможна и другая ситуация – нелинейность может быть связана тоже со свойствами среды, однако зависящими от наличия и величины протекающего по ней тока. Это случай газового разряда, вакуумного прибора, иногда – полупроводника, диэлектрика и даже просто металла! Для газового разряда ситуация такова: когда нет тока, то среда – это почти диэлектрик, т.е. почти просто газ. В вакууме, если нет тока, так вообще почти ничего нет – это вакуум! В полупроводнике и диэлектрике при некоторых условиях возможно увеличение количества свободных электронов, которые могут переносить заряд. А проще всего обстоит дело в металле, как учат нас Джоуль и Ленц, – металл нагревается током, и его сопротивление растет. Среда изменяется под воздействием тока.

Еще одна ситуация: ток может и не менять свойства среды, но его магнитное поле может сжимать его самого к оси проводника. При этом сечение, по которому протекает бульшая часть тока, уменьшается, а сопротивление (вы ведь помните, что $R = \rho L/S$) растет. Этот эффект более заметен в полупроводниках, а уж в газовом разряде вы все его видели. Именно поэтому молния – относительно тонкая, т.е. диаметр канала разряда много меньше длины молнии.

Ну и последний вариант, самый для вас загадочный. С ростом частоты ток прижимается к поверхности – это скин-эффект (англ. skin – кожа). Если кто-то из вас потрошил телевизионный кабель, то мог задуматься – почему поверхности посеребрены? Кстати, там дело не только в том, что сопротивление серебра меньше, чем у меди (обычно в книжках пишут именно об этом), ситуация хитрее. Сопротивление серебра отличается от меди незначительно, причем если серебряное покрытие нанесено гальванически, то его сопротивление может быть немного больше и вообще может оказаться не меньше, чем у меди. С медью проблема в том, что она окисляется на воздухе, ее соединения оказываются полупроводниками, а они поглощают электромагнитные волны. Кстати, на заре электроники соединения меди вообще использовали как материал для полупроводниковых диодов – так называемые купроксные выпрямители (рис. 6).

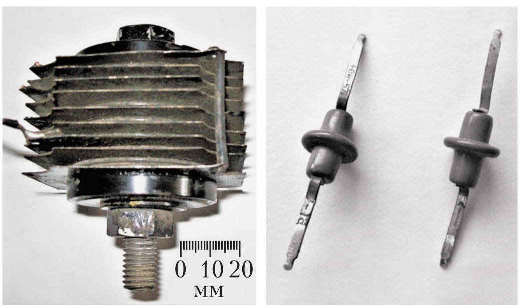

Рис. 6

Однако самый простой и самый «школьный» ответ на вопрос о двух точках и о предложении провести что-то через них, это спросить насчет погрешностей. Так что поговорим о погрешностях.

Иногда говорят, что в физике все величины определяются с какой-то погрешностью, но это не так. Есть величины, которые могут быть определены точно. Например, это некоторые величины, которые могут принимать лишь конечное количество значений. Хотя и тут все не просто – например, после приведенного значения скорости света написано просто и ясно «точно». Но это немного другой, метрологический вопрос (о нем хорошо рассказано, например, тут https://habr.com/ru/post/407675/).

Могут быть определены точно и величины, которые имют бесконечное количество значений – например, если это целые числа. Скажем, количество чего-то дискретного, например контрольных в полугодии или баллов на ЕГЭ. Правда, в физике нет чего-либо бесконечного, даже количества электронов во Вселенной, но не будем о страшном. А вот забавный вопрос – всегда ли можно точно определить величину, которая может принимать как целые числа бесконечное, но счетное множество значений? Оказывается, нет. Пусть некая величина может принимать значения 0 (ноль) и все значения $1/n$, где n – целое. У этой величины бесконечно много значений, которые она может принимать точно, но есть одно, которое она точно принять не может! Вы, наверное, уже догадались, какое, поняли, почему, и можете это доказать. Перейдем, однако, к более физическим ситуациям.

В школьной физике, да и вообще в метрологии, когда речь заходит о точности, пишут, например, что напряжение равно 10 ± 1 В. Это означает, что некая величина лежит между 9,9 В и 10,1 В, причем равновероятно в любом месте этого отрезка, и не может лежать вне него. В реальной жизни это не так. Случайные погрешности обычно подчиняются так называемому нормальному распределению – чем ближе к краям указанного в конкретном случае отрезка, тем они более редки, вне него, они, естественно, еще более редки, но вовсе не исчезают. Тем не менее, пределы погрешностей во многих случаях (а в технике и инженерии почти всегда) указывают именно так – для простоты. Соответственно, так сделано и в экзаменационных задачах. Поэтому возникает вопрос. Как провести прямую через две точки, если указаны не только точки (x_1, y_1) и (x_2, y_2), но и погрешности, причем их величины Δx_1, Δy_1, Δx_2, Δy_2 все могут быть разными?

Ответ прост – нарисуйте на клетчатой бумаге эти точки с их погрешностями, и вы увидите ответ (рис. 7). Но при попытке его записать, возникнет небольшая проблема. Описание множества прямых вам хотелось бы получить в виде $y = ax + b$ и пределов на a и b, т.е. $a_1 < a < a_2$ и $b_1 < b < b_2$, но таком

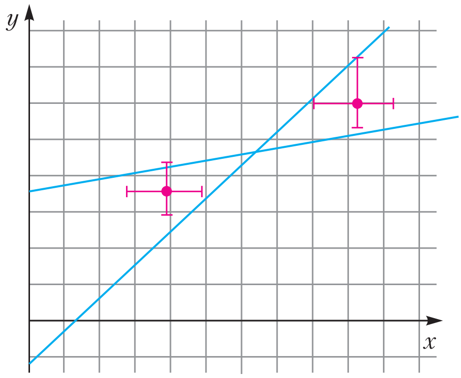

Рис. 7

простым ответом вы не обойдетесь. Пределы окажутся связаны, т.е. на плоскости (a, b) область допустимых значений a и b окажется не прямоугольной.

Впрочем, для двух точек вы выкрутитесь, записав семейство прямых в некотором виде, который подскажет вам ваш собственный рисунок. А вот если точек несколько, хотя бы три, этот фокус не пройдет. Это уже тема для небольшого исследования и, может быть, написания компьютерной программы. Поэтому остановимся, переведем дух и посмотрим на вопрос с несколько другой, более фундаментальной стороны. Почему мы вообще взялись проводить прямую? В учебной задаче – потому, что задача так сформулирована. А в реальной жизни?

В реальной жизни физик, увидев, как кто-то проводит что-то (не важно, что) через две точки, улыбнется и посоветует получить хотя бы еще десяток или лучше несколько десятков точек. Но эксперимент может быть сложен и дорог и количество экспериментальных данных так или иначе ограниченным, а главное – все-таки, какую кривую проводить? Самая естественная ситуация – когда есть какая-то теория и надо ее проверить либо есть несколько теорий и надо выбрать из них ту, которая лучше согласуется с экспериментом. Тут все ясно – проводим ту кривую, которую предсказывает теория, или те кривые, которые предсказывают разные теории, и смотрим, согласуются ли они с экспериментом какая согласуется лучше.

Распространенная ситуация – когда в теории есть некоторый параметр, значения которого мы не знаем. Но имея экспериментальные данные, мы можем подобрать такое значение этого параметра, что точки лягут на соответствующую кривую. Впрочем, бывает и так, что ни при каком значении они не ложатся. Тогда надо либо искать ошибку в измерениях (впрочем, это всегда не вредно сделать), либо идти к теоретикам за другой теорией. Впрочем, можно и самому о ней подумать.

А если теории, на которую точки ложатся хорошо, все-таки нет и данные надежны? Хотя мы говорим о науке, но бывают ситуации, когда нам теория, собственно говоря, и не нужна, нам только нужен простой и компактный способ хранения данных. Про обработку данных мы поговорим немного позже, но для хранения и использования есть простой, но не очень физичный способ. Берем любую функцию (как правило, берут полином) и подбираем коэффициенты так, чтобы кривая прошла во всем исследованном диапазоне через наши экспериментальные точки, точнее – через поля ошибок, которые около этих точек показаны. А если нас интересует именно физический смысл?

Бывает, что теория, упрощенная и приближенная, но хоть какая-то, есть. Тогда смотрят, как наши точка отклоняются от этой примитивной теории. Например, может оказаться, что согласование хорошее при малых значениях аргумента и плохое при больших, или наоборот, или расхождения возникают именно в какой-то узкой области. Все это может оказаться важным для уточнения теории. Еще интереснее, если отклонения сами зависят от какого-то параметра, например от времени. Что вы скажите, если отклонения от закона Ома будут возникать при больших токах и нарастать со временем? Ясное дело, прямая дорога в объятия Джоуля и Ленца.

Иногда можно сказать, что эта задача похожа на такие-то или что этот процесс похож на такие-то, и тогда, наверное, решением будет комбинация экспонент, или синусоид, или... Но такие рассуждения требуют наличия опыта – и вообще накопленного в физике, и имеющегося у того, кто смотрит на точки и размышляет, что с ними делать. Кроме того, опыт может и подвести, всякое бывает. Как сформулировал Айзек Азимов, «Самая волнующая фраза, какую можно услышать в науке, – вовсе не “эврика”, а “вот это забавно”».

Теперь обратимся к обработке данных и получению самих этих точек, о которых шла речь. Начнем с понятий интерполяция и экстраполяция. Вам наверняка придется с ними встретиться в учебе и работе, так почему бы не сейчас? Итак, вы делаете некий эксперимент и получаете, среди прочего, нашу знаменитую четверку – x_1, y_1, x_2, y_2. Других данных у вас нет и получить вы их не можете, а вам потребовалось значение y для некого x, причем такого, что $x_1 < x < x_2$. Так вот, нахождение y называется интерполяцией. Понятно, что найти его вы не можете, однако все это как-то делают. Простейший вариант – провести прямую линию от точки (x_1, y_1) до точки (x_2, y_2) и… вы все поняли. Способ так и называется – линейная интерполяция. Существуют и другие варианты – например, можно интерполировать куском полинома, оговорив не только сами точки, но и производные в этих точках.

С точки зрения физики, нам важно знать хоть что-то о всей функции $y(x)$ – ну, скажем, о максимальном значении ее производной, грубо говоря, насколько она гладкая. Ибо если этой функции свойственны короткие выбросы длиной Δx, а у нас $x_2 - x_1 \gg \Delta x$, то в нашей линейной интерполяции нет смысла – на интервале от x_1 до x_2 она успеет напрыгаться. Иными словами, и тут, для того чтобы действовать разумно, нам нужно что-то знать об этом классе задач вообще и об этой функции, с которой мы общаемся, в частности. Например, мы можем видеть ее поведение на больших интервалах, во многих точках. И рассуждать примерно так – ну, раз она там гладкая, так, наверное, и здесь тоже гладкая. Но мы можем и ошибиться. В физике, да и в любой науке, это бывает.

Экстраполяция – это почти то же самое, но x лежит не между x_1 и x_2, а снаружи. Это значит, что мы что-то исследуем, добрались до каких-то значений аргумента, дальше техника не позволяет, а очень хочется. Вот тут надо быть особенно осторожным – по психологическим причинам. Раньше у нас были по бокам все-таки x_1 и x_2, а теперь с одного бока свобода! По опыту автора, самая распространенная ошибка – экстраполяция в область низких температур данных, полученных при высоких температурах. Дело в том, что с ростом температуры множество процессов ускоряется (об этом рассказывается, например, в статье А.Митрофанова «Время жизни шипучей таблетки в стакане воды» в «Кванте» №9 за 2021 г.), и их можно изучить за разумное время. Но экстраполировать данные, полученные при высоких температурах, в область низких температур не всегда удается. Уравнение-то соблюдается, но при изменении температуры начинает доминировать другой процесс, с другой энергией активации. Так что и здесь необходимо более общее знание, более глубокое представление о процессах.

Ну, и в заключение обратимся к тому, с чего начали, – к самим точкам. Много ли их надо? Сколько раз надо производить измерения? На этот вопрос нет общего и надежного ответа. Иногда говорят, что любую величину нужно измерять три раза. На самом деле это не так, необходимое количество измерений зависит от того, что мы измеряем и что мы об этом уже знаем, зачем мы измеряем, какая точность и надежность нам нужна, а также от стоимости и трудоемкости измерений. Например, если вы проверяете батарейку, вынутую из фонарика, который стал плохо светить, и прибор показывает напряжение заметно меньше нормы, вряд ли вы будете повторять замер и уж точно не будете делать три замера. А вот если он покажет норму, вы удивитесь и наверняка проверите (а потом проверите напряжение при работе на нагрузку – подумайте, почему). Многократные измерения проделывают, если заметны случайные ошибки и вы хотите усреднением поднять точность. Но в этом случае может быть полезнее подумать, откуда они взялись и не лежат ли рядом с ними ошибки систематические, от которых не спасет усреднение. А вот понимание, что, как и почему вы измеряете, может и спасти.

Так что удачных измерений, надежных интерполяций, осторожных экстраполяций и осмысленных проведений кривых через точки!