НОВОЕ
В ЖИЗНИ,
НАУКЕ,
ТЕХНИКЕ

Серия
«Математика,
кибернетика»
№ 12, 1979 г.

Издается
ежемесячно
с 1967 г.

Ю. И. Манин,

доктор физико-математических наук,
профессор

МАТЕМАТИКА И ФИЗИКА

Издательство
«Знание»
Москва
1979

22
М23

СОДЕРЖАНИЕ

Манин Ю. И.

М23 Математика и физика. М., «Знание», 1979.

64 с. (Новое в жизни, науке, технике. Серия «Математика, кибернетика», 12. Издается ежемесячно с 1967 г.)

Связь математики и физики... Только ли в том она проявляется, что физики говорят на языке математики? Не только в этом. Брошюра рассказывает о том, как математика сопоставляет с некоторыми важнейшими физическими абстракциями (моделями) свои образы, далеко уходящие от тех представлений, которые дает непосредственный опыт и физический эксперимент.

Рассчитана брошюра на тех, кто интересуется методологией науки и знаком в известной мере и с математикой и с физикой.

20 000 22

ПРЕДИСЛОВИЕ

Есть предание о том, как один известный математик начинал читать логику второкурсникам. «Логика — это наука о законах мышления, — сообщал он. — Теперь я должен объяснить вам, что такое наука, что такое закон и что такое мышление. Что такое «о», я объяснять не буду».

Взявшись писать книжку «Математика и физика», автор понимал, что ее объема едва хватит на попытку объяснить, что такое «и» в ее названии. Две науки, бывшие единой ветвью на дереве познания, к нашему времени далеко разошлись. Одна из причин этого в том, что обе они в этом столетии активно занимались самоосознанием, т. е. своими средствами строили свои собственные модели. Физика волновало взаимоотношение мышления и действительности, а математика — мышления и формул. Оба эти отношения оказались много сложнее, чем казалось раньше, и модели, автопортреты, маски-для-себя двух дисциплин вышли очень несхожими. В результате уже со студенческой скамьи физиков и математиков учат думать по-разному. Было бы замечательно владеть обоими типами профессионального мышления, хотя бы так, как мы владеем правой и левой рукой.

Но эта книжка — партия одной руки. Автор, по образованию математик, как-то прочел студентам четыре лекции под названием «Как математик должен учить физику». В лекциях говорилось, что современная теоретическая физика — это роскошный, совершенно раблезианский полнокровный мир идей, и математик может найти в нем все, что душе угодно, кроме порядка, к которому он привык. Поэтому хороший способ настроить себя на активное изучение физики — сделать вид, что ты пытаешься, наконец, навести в ней этот самый порядок.

В книжке, выросшей из этих лекций и дальнейших размышлений, я попробовал выделить несколько крупных абстракций двух наук и сопоставить их. На самом высоком уровне такие абстракции теряют терминологичность и способны стать культурными символами времени: вспомним судьбу слов «э в о л ю ц и я», «о т н о с и т е л ь н о с т ь» или «п о д с о з н а т е л ь н о е». Здесь мы спускаемся ступенькой ниже и обсуждаем слова, еще не символы, но уже почти не термины: «множество», «симметрия», «пространство — время». (Ср. попытку М. М. Бахтина терминологически ввести последнее понятие в литературоведение в нарочито остраненной форме «хронотоп».) Часть этих слов стоит в названиях главок. У каждого читателя в сознании должны быть первоначальные образы этих понятий, образы, имеющие физическое происхождение в широком смысле слова.

Автор хотел показать, как математика сопоставляет с такими физическими абстракциями новые образы, для тренированного рассудка почти осязаемые, но далеко ушедшие от тех, которые дает прямой жизненный и физический опыт. Скажем, движение планет Солнечной системы математик представит в виде линии тока несжимаемой жидкости в 54-мерном фазовом пространстве, объем которого задается мерой Лиувилля.

Читателю может потребоваться усилие воли, чтобы увидеть в математике воспитателя образного мышления. Чаще с ней связывается представление о жесткой логике и вычислительном формализме. Но это — лишь дисциплина, линейка, которой нас учат не умирать.

Вычислительный формализм математики — мысль, экстериоризованная до такой степени, что она на время отчуждается и превращается в технологический процесс. Математический образ формируется в затяжном приживлении к человеку этой временно отторгнутой мысли. Думать — значит вычислять, волнуясь.

Безумная идея, которая ляжет в основу будущей фундаментальной физической теории, будет осознанием того, что физический смысл имеет некоторый математический образ, ранее не связывавшийся с реальностью. С этой точки зрения проблема безумной идеи — это проблема выбора, а не порождения. Не нужно понимать это слишком буквально. В шестидесятых годах (по частному поводу) было сказано, что крупнейшее открытие последних лет в физике — комплексные числа. Нечто подобное автор имеет в виду.

Я не хочу извиняться за субъективность суждений и выбора материала. О физике и математике писали Галилей, Максвелл, Эйнштейн, Пуанкаре, Фейнман, Вигнер; только надежда сказать что-то свое может оправдать новую попытку.

1. МАТЕМАТИКА С ПТИЧЬЕГО ПОЛЕТА

Математическая истинность. Вероятно, самые простые математические действия — это арифметические вычисления вроде такого:

$$\frac{0,25}{20}\cdot\frac{\sqrt{13}}{1,1}\cdot\frac{7,8\cdot 10^4}{2,04\cdot 10^5}\cdot\frac{2\cdot 0,048}{0,021+0,019}=0,038.$$

Для реалистичности этот пример списан не из школьного задачника, а из статьи Энрико Ферми «О поглощении и диффузии медленных нейтронов». Подумаем немного о смысле такого вычисления.

а) Для проверки этого равенства можно условиться, что оно относится к целым числам (возведем в квадрат, освободимся от знаменателей и будем считать все в тысячных долях единицы). Тогда наше равенство можно рассматривать как предсказание о результате некоторого «физического эксперимента», состоящего в следующем: нужно взять две группы по 48 предметов ($2\cdot 0,048$), повторить это действие 78 000 раз ($\times 7,8\cdot 10^4$) и т. п. Так в первом классе раскладывают по кучкам палочки, чтобы уяснить смысл счета, целого числа, сложения и умножения, а также смысл арифметических тождеств. Поэтому разумно представлять себе, что арифметика целых чисел есть «физика собирания предметов в кучки».

б) Все же практическое вычисление, конечно, производится иначе: оно состоит из серии некоторых стандартных преобразований левой части тождества. Мы выбираем группу символов слева, скажем $\frac{0,25}{20}$, и заменяем ее по школьным рецептам на 0,0125 и т. п. Все правила, включая правила о порядке действий, можно сформулировать заранее. Безошибочность вычисления — это его грамматическая (рецептурная) правильность; она же гарантирует «физическую истинность» результата. (Разумеется, Ферми округляет левую часть; и без вычислений ясно, что его равенство не может быть верным буквально, потому что число $\sqrt{13}$ — иррационально.)

в) Для Ферми смысл этого вычисления резюмируется следующей фразой: «группа A... является столь узкой энергетической полосой, что в процессе замедления через нее проходит только 4 % нейтронов». (4 % — это 0,038 справа.) Ясно, что к такому выводу мы не можем непосредственно прийти, как бы ни представляли себе смысл арифметического вычисления. Ни раскладывание 78 000 кучек по 96 предметов, ни деление 0,096 уголком на 0,04 сами по себе не имеют никакого отношения к нейтронам. Математическое рассуждение входит в физический текст вместе с актом его физического истолкования; именно этот акт и есть самое поразительное в современной физике.

Как бы то ни было, уже на нашем простом примере видны три аспекта математической истинности. Условно их можно обозначить как содержательную истинность, формальную правильность, или доказуемость, и адекватность физической модели.

Для математики, замкнутой в себе, существенны лишь первые два аспекта, и только двадцатый век принес понимание различия между ними. Рассмотрим такое просто формулируемое утверждение, как гипотезу Ферма. Хотя мы не знаем ни ее доказательства, ни опровержения, мы уверены, что она либо истинна, либо ложна. Эта уверенность основана на абстракции возможности произвести бесконечно много арифметических действий (или «раскладываний на кучки»), перебрав все суммы степеней пар целых чисел. Вообще, понятие об истинности (большинства) математических утверждений включает в себя представление о таких бесконечных сериях проверок. Между тем всякое математическое доказательство, т. е. рассуждение, состоящее из последовательного применения аксиом или логических правил вывода, есть существенно конечная процедура. К. Гёдель доказал в тридцатых годах, что по этой причине доказуемость значительно у́же содержательной истинности, даже когда речь идет лишь о целых числах. При этом совершенно безразлично, из каких аксиом мы исходим, лишь бы они были содержательно истинны и задавались конечным списком (или конечным числом правил их порождения). Это различие между содержательной истинностью и доказуемостью широко известно, но, кажется, его следствия поняты плохо. В литературе часто обсуждаются проблемы редукционизма: сводится ли биология или химия к физике? Ясно, что речь может идти лишь о некоторой теоретической модели явлений физики, биологии и химии, притом достаточно математизированной. Но тогда следует объяснить, что подразумевается под сводимостью — абстракция типа содержательной истинности или типа выводимости из аксиом. Продумывание обеих возможностей создает впечатление, что, говоря о сводимости законов, мы просто не понимаем, о чем говорим.

Множества. Современные представления о математической истинности связаны с развитием двух крупных концепций: теоретико-множественной математики и математики формальных языков. Математические формализмы знакомы всем. Типичный математический багаж студента может состоять из умения выполнять арифметические действия с числами в десятичной записи, преобразовывать алгебраические тождества, дифференцировать и брать некоторые интегралы. Этот язык математического анализа, практически сложившийся во времена Эйлера и Лагранжа, оказался очень удобным, эффективным в решении задач и доступным для массового изучения. Параллельно происходило развитие представлений того, о чем говорит этот язык, т. е. выяснение смысла таких понятий, как $\sqrt{-1}$, функция, дифференциал и т. п. Огромную роль при этом играли геометрические представления: комплексные числа потеряли свою таинственность лишь после того, как Арган и Гаусс предложили их последовательную интерпретацию точками евклидовой плоскости; дифференциал интерпретируется через представление о касательной и т. п. Теоретико-множественные понятия заложили универсальную базу для определения всех математических конструкций в таких «обобщенно геометрических» образах. Эти образы одновременно представляют собой вместилище смысла математических формализмов и средство отбора содержательных языковых утверждений из всего необозримого моря выводимых математических формул.

В этой книжке мне хотелось бы продемонстрировать пользу таких образов в роли посредника между математикой и физикой. Конечно, возможности их популярного изложения ограничены. На шестидесяти страницах мы не сможем объяснить их точный смысл и научить пользоваться ими для решения задач. Но, может быть, читателю станут яснее некоторые идеи математики и теоретической физики. Трудность понимания концепций квантовой теории или общей теории относительности отчасти связана с тем, что при попытках их объяснения опускается такой акт промежуточной теоретико-множественной интерпретации математических моделей. Даже в университетском образовании ему уделяется недостаточно внимания; общение физика и математика часто затруднено тем, что физик склонен переходить от формул прямо к их физическому смыслу, минуя «математический смысл». Впрочем, в последние годы положение заметно улучшается.

Хороший физик пользуется формализмом, как поэт — естественным языком. Пренебрежение ригористическими запретами оправдывается конечной апелляцией к физической истине, чего не может позволить себе математик. Выбор лагранжиана в единой теории слабых и электромагнитных взаимодействий Салама — Вейнберга, введение в него полей Хиггса, вычитание вакуумных средних и прочее колдовство, приводящее, скажем, к предсказанию нейтральных токов, оставляет математика в состоянии немого изумления.

Но вернемся к множествам.

Важнейшие множества физиков — это множества не предметов, а возможностей: конфигурационное пространство системы есть множество ее возможных мгновенных состояний, пространство — время есть множество возможных событий типа «вспышки», отмечающих точки. Физик обычно спешит ввести на этом пространстве координаты, т. е. функции с числовыми значениями. Если набор n таких функций позволяет однозначно координатизировать точки множества, то допустимо считать, что оно лежит в n-мерном вещественном числовом пространстве R^n, состоящем из векторов вида $(a_1, \ldots, a_n)$. «Фигуры», т. е. подмножества такого пространства, измерения в нем расстояний, углов, объемов и т. п., наконец, его движения или отображения в себя, — все это составляет главный арсенал геометрических образов физики. При этом важно, что размерность n может быть как угодно велика и даже бесконечна — в строгом математическом тексте определение бесконечномерности нужно вводить отдельно, но мы будем представлять себе здесь бесконечномерность как «неопределенно большую конечномерность». Если координаты принимают комплексные значения, то наши множества погружаются в C^n. Вообще же часто о координатах можно и не упоминать. Физически они иногда являются пережитком слишком упрощенных представлений о наблюдении; математически — напоминанием о времени, когда не существовало языка, на котором можно было бы содержательно обсуждать множества, не являющиеся множествами чисел или векторов.

Теоретико-множественный язык хорош тем, что он не вынуждает говорить ничего лишнего. Г. Кантор определил множество как «соединение в одно целое различимых объектов нашей интуиции или нашей мысли». Это наилучшее объяснение множества как понятия, помогающего познавать мир.

Многомерное пространство и идея линейности. Если автомобиль прошел за секунду двадцать метров, то за две секунды он, скорее всего, пройдет сорок метров. Если слабый ветер отклонил летящую пулю на три сантиметра, то вдвое более сильный отклонит ее на шесть сантиметров. Отклик на малые воздействия линейно зависит от этих малых воздействий — таков естественнонаучный принцип, лежащий в основе огромного количества математических моделей. Математик превращает этот принцип в определение дифференцируемой функции и в постулат о том, что бо́льшая часть процессов бо́льшую часть времени описывается такими функциями. Говорит ли закон упругости Гука или закон Ома что-нибудь большее, чем этот принцип линейного отклика на малые воздействия? Да, если оказывается, что законы остаются верны и для довольно больших воздействий.

Линейное пространство — это идеализация «сколь угодно больших малых воздействий». Не обязательно вводить координаты, нужно лишь помнить, что элементы линейного пространства можно складывать и умножать на числа (вещественные или комплексные — этот эпитет прибавляется к названию пространства). Исходный геометрический образ — это наше «физическое пространство» размерности три; пространства R^n, C^n с координатным сложением и умножением исчерпывают все конечномерные линейные пространства.

Размерность линейного пространства — это количество независимых линейных координатных функций на нем. Теорема о том, что от выбора самих координатных функций она не зависит, при всей ее простоте, является глубоким результатом. Она устанавливает связь между непрерывным и дискретным: целое число — размерность впервые появляется не как количество предметов или дискретных образов, но как мера величины непрерывного объекта.

Линейное отображение, или оператор, — это идеализация линейного отклика на произвольные воздействия.

Отклик может измеряться элементами того же пространства, что и воздействие, или другого; в любом случае это отображение линейного пространства в линейное пространство, переводящее сумму векторов в сумму их образов и произведение вектора на число в произведение образа на то же число.

В одномерном пространстве всякое линейное отображение в себя есть умножение на число — «коэффициент усиления». В комплексном случае геометрический образ немного сложнее: поскольку одномерное комплексное пространство устроено как вещественная плоскость, умножение на комплексное число есть комбинация вещественного растяжения и поворота. Чистые повороты, т. е. умножения на числа, по модулю равные единице, играют большую роль в квантовой механике: в их терминах формулируется закон эволюции замкнутой квантовой системы.

Важный класс линейных отображений n-мерного пространства в себя образуют растяжения вдоль n независимых направлений со своим коэффициентом вдоль каждого из них. Множество «коэффициентов растяжения» линейного оператора называется его спектром: омонимия с физическим термином отражает их глубокие связи.

В квантовой физике идея линейности приобретает фундаментальный физический смысл благодаря основному постулату о суперпозиции квантовых состояний. В классической физике и математике, кроме исходной мысли о линеаризации «чего угодно» в малом, большую роль играет замечание о том, что функции (все или непрерывные, или дифференцируемые, или интегрируемые по Риману и т. п.) на любом пространстве сами образуют линейное пространство, потому что их можно складывать друг с другом и умножать на числа. Пространства функций в большинстве случаев бесконечномерны, но возможность направленно воспитать и затем применить к ним первоначально развитую конечномерную (даже трехмерную) интуицию оказалась исключительно плодотворным открытием. В двадцатом веке этому учили нас Давид Гильберт и Стефан Банах.

Измерения в линейном пространстве. В трехмерном физическом пространстве существуют твердые тела, сохраняющие некоторую «тождественность самим себе» в больших пространственно-временных областях. Это — основа всех физических измерений. Совсем не очевидно заранее, какие из идеализированных свойств физических измерений наиболее полезны в математической теории и в приложениях.

Действительно, математические понятия, связанные с идеей классического измерения, образуют сложный идейный узор. Сразу назовем несколько образов: длина, углы, площади и скалярные произведения, движения.

Чтобы у читателя не возникло неверного впечатления, заметим, что само понятие линейного пространства не содержит ничего, позволяющего однозначно измерять что бы то ни было. У векторов нет никакой длины (правда, у пропорциональных векторов имеется естественное отношение длин), угол между векторами не имеет никакой естественной меры и т. д. Поэтому для математического оформления идеи измерения мы должны дополнительно ввести новый геометрический образ или даже несколько образов; на математическом жаргоне — снабдить пространство дополнительной структурой.

Первый из таких образов — единичная сфера пространства: множество векторов единичной длины. Если любой ненулевой вектор после умножения на подходящее число a попадает на единичную сферу, мы получаем возможность приписать ему длину — это будет $|a|^{-1}$; значит, a должно быть определено с точностью до умножения на число, по модулю равное единице. Расстояние между векторами x и y можно определить как длину их разности $|x-y|$. Если ненулевые векторы имеют ненулевую длину, выполняется неравенство треугольника $|x+y| \leqslant |x|+|y|$ и еще условие о существовании пределов последовательностей Коши, мы приходим к понятию банахова пространства. Таковы многие полезные пространства функций. Произвольная банахова сфера, однако, недостаточно симметрична, чтобы быть правильным обобщением единичной сферы в трехмерном евклидовом пространстве. Есть два разных способа добиться нужной симметрии, наложив на банахову сферу дополнительные условия: а) потребовать, чтобы некоторая $\frac{n(n-1)}{2}$-мерная непрерывная группа линейных отображений переводила ее в себя (n-размерность пространства); б) потребовать, чтобы в пространстве существовало скалярное произведение векторов (x, y) (линейная функция по обоим аргументам в вещественном случае и несколько более сложная в комплексном) такое, чтобы $|x|^2=(x, x)$ для всех x. Первый способ — обобщение идеи о том, что твердые тела можно вращать, второй — что между векторами можно измерять углы, также не меняющиеся при вращениях пары. Обе идеи тесно связаны и приводят к понятию многомерного евклидова пространства (в комплексном случае его называют гильбертовым). В подходящих координатах единичная сфера в таком пространстве задается привычным уравнением $\sum\limits_{i=1}^{n} |x_i|^2=1$. Вращения — это линейные отображения, переводящие эту сферу в себя, они образуют группу, которая обозначается $O(n)$ в вещественном случае и $U(n)$ — в комплексном. В евклидовом вещественном пространстве скалярное произведение принимает вещественные значения и является симметричным: $(x, y)=(y, x)$. В евклидовом комплексном пространстве оно принимает комплексные значения и при перемене мест векторов становится комплексно-сопряженным: $(x, y)=\overline{(y, x)}$. В обоих случаях выполняется замечательное неравенство $|(x, y)|^2 \leqslant |x|\,|y|$, так что число $\frac{(x, y)}{|x|\,|y|}$ по модулю не больше единицы. Оно вещественно для вещественных пространств и существует угол φ, для которого $\cos\varphi=\frac{(x, y)}{|x\ |y|}$; он называется углом между векторами x и y. В комплексном случае можно определить этот угол формулой $\cos\varphi=\frac{|(x, y)|}{|x|\,|y|}$. Правая часть здесь принимает только значения, лежащие между нулем и единицей, и существует еще одна замечательная физическая величина с таким свойством — это в е р о я т н о с т ь. В квантовой механике числа $\cos^2\varphi$ интерпретируются как вероятности, о чем мы подробнее расскажем ниже. В школьной геометрии векторы x, y называются ортогональными, если косинус угла между ними равен нулю, т. е. если $(x, y)=0$; эта же терминология применяется и в общем случае.

Если отказаться от тех или иных свойств евклидовости, то понятие скалярного произведения приводит к нескольким важным классам линейных геометрий. Например, в R^4 можно задать «длину» вектора $x=(x_0, x_1, x_2, x_3)$ формулой $|x|^2=x_1^2+x_2^2+x_3^2-x_0^2$. Один минус приводит к большому количеству отличий: например, имеются целые прямые, состоящие из векторов нулевой длины. Они изображают лучи света в основной модели пространства — времени специальной теории относительности — в знаменитом пространстве Минковского.

Можно отказаться от условия $(x, y)=(y, x)$ и заменить его условием $(x, y)=-(y, x)$. Всякий вектор в таком пространстве «ортогонален самому себе»! Эту геометрию, называемую симплектической, нужно долго изучать, чтобы привыкнуть к ней. Гироскоп, ориентирующий ракету, — это — посланец шестимерного симплектического мира в нашем трехмерном; там его поведение выглядит просто и естественно. Хотя симплектическая геометрия была открыта в прошлом столетии, ее роль в физике долго недооценивалась и в учебниках все еще затемняется старинным формализмом.

Но вернемся в евклидов, хотя и многомерный, мир. Последнее, что мы хотели бы обсудить, — измерение объемов. Если $e_1, \ldots, e_n$ — попарно ортогональные векторы единичной длины, то натянутый на них n-мерный кубик с ребром тоже единичной длины — множество векторов вида $x_1e_1+\ldots+x_ne_n$, где $0 \leqslant x_i \leqslant 1$. Его объем естественно считать равным единице. Сдвиг этого кубика на любой вектор не меняет объема; кубик с ребром длины a имеет объем a^n . После этого объем любой n-мерной фигуры можно определить, замостив ее большим числом маленьких кубиков и сложив их объемы. Проблемы возникнут около границы — там останется свободное пространство, при попытке замощения которого кубики начнут вылезать наружу. Если граница не очень сильно изрезана, то, делая кубики все более мелкими, мы сможем как угодно уменьшить ошибку. Это — основная идея интегрирования. Она дополняется еще следующей конструкцией: предположим, что в нашей области пространства находится нечто, «субстанция», как сказали бы в прошлом веке, которая характеризуется своей плотностью $f(x)$ вблизи точки x. Общее количество этой субстанции будет примерно выражаться суммой ее количеств во всех кубиках замощения, а количество в одном кубике — произведением объема этого кубика на значение плотности в какой-нибудь его внутренней точке. Вся сумма есть «сумма Римана», а ее предел — интеграл от функции f по объему.

В математике трудно указать более классическое и в то же время более живое понятие, чем интеграл. Каждые несколько десятилетий приносят его новые математические варианты, а физика все время требует еще. Определение интеграла Римана, которое мы привели выше, математически разумно лишь для не слишком сильно меняющихся функций f, скажем непрерывных. Но почти каждая физическая модель, как только она сменяется более детальной, обнаруживает, что функция f, казавшаяся довольно гладкой, есть результат усреднения более сложной «мелкозернистой» картины. Заряд можно измерять интегралом от его плотности, пока мы не выходим на масштабы, где носителями заряда являются электроны и ионы. Плотность заряда на точечном носителе бесконечна, а вне его равна нулю, и мы вынуждены строить аппарат для интегрирования таких функций.

Вызовом математикам остаются замечательные континуальные интегралы Фейнмана, уже превратившиеся в основной инструмент квантовой теории поля, но все еще не определившиеся как математический объект. Два обстоятельства затрудняют их понимание: интегрировать приходится по бесконечномерному пространству и притом очень сильно колеблющиеся функции. Скажем здесь несколько слов об эффектах бесконечномерности, понимая ее наивно как очень большую конечную размерность.

Из отрезка длины 1 вырежем его среднюю часть длины 0,9. Длина остатка будет, конечно, составлять 10% от длины всего отрезка. Из круга диаметром 1 вырежем концентрический круг диаметром 0,9. Площадь оставшегося кольца будет уже 19% площади круга. Из шара диаметром 1 вырежем концентрический шарик диаметром 0,9. Объем оставшегося шарового слоя будет составлять уже 27,1% объема шара: почти треть вместо одной десятой для отрезка. Объем n-мерного шара диаметром d, как нетрудно сообразить «физически», должен выражаться формулой $c(n)d^n$, где $c(n)$ — константа, от d не зависящая. Доля объема концентрического шара диаметром 0,9 d поэтому будет $(0{,}9)^n$; она стремится к нулю вместе с ростом n. Двадцатимерный арбуз радиусом 20 см с толщиной корки 1 см чуть не на две трети состоит из корки:

$$1-\left(1-\frac{1}{20}\right)^{20}\approx 1-e^{-1},\ e\approx 2{,}72.$$

Эти расчеты позволяют сформулировать геометрический образ: «объем многомерного тела почти целиком сосредоточен у его поверхности». (Интересно рассмотреть также вместо шара куб — тот же эффект проявляется в быстром росте числа его граней.)

Представим себе простейшую модель газа: N точечных атомов, движущихся в резервуаре со скоростями v_i, $i=1, \ldots, N$; каждый атом имеет массу m. Кинетическая энергия газа E равна $\sum\limits_{i=1}^{N}\frac{mv_i^2}{2}$; состояние газа, описываемое набором скоростей при фиксированной энергии E, определяет точку на $(N-1)$-мерной евклидовой сфере радиусом $\sqrt{\frac{2E}{m}}$. Для макроскопического объема газа в нормальных условиях размерность этой сферы имеет порядок $10^{23}-10^{25}$ (определяемый числом Авогадро), т. е. очень велика. Если два таких резервуара соединены так, что они могут обмениваться энергией, но не атомами, и сумма их энергий $E=E_1+E_2$ остается постоянной, то энергии E_1 и E_2 большую часть времени будут близки к таким, которые максимизируют объем пространства состояний, доступный объединенной системе. Он равен произведению объемов сфер радиусов $\sqrt{\frac{2E_1}{m}}$ и $\sqrt{\frac{2E_2}{m}}$ соответственно, первый из которых с ростом E_1 очень быстро растет, а второй очень быстро убывает. Их произведение имеет поэтому острый пик в точке, которую легко вычислить; точка отвечает условию равенства температур. «Сосредоточенность объема многомерного тела вблизи поверхности», в сущности, предопределяет существование температуры как макроскопической величины.

О каком пространстве идет речь в этом примере? О пространстве состояний физической системы, точнее, о некотором его фактор-пространстве: мы не принимаем во внимание ни положения атомов, ни направления скоростей. Одна его точка — это снова возможность. Типичное множество — это не стулья в комнате и не ученики в классе и даже не атомы в резервуаре, а возможные состояния атомов в резервуаре.

Асимптотические свойства многомерных объемов — это геометрический арсенал статистической физики. Все разнообразие мира природа конструирует из малого числа разных кирпичиков. Кирпичики одного сорта тождественны, и когда статистика описывает поведение их конгломератов, она пользуется образом точки, блуждающей в областях почти бесконечномерного фазового пространства. Макроскопические наблюдения позволяют лишь грубо указать расположение области, куда попала точка, и чем больше ее объем, тем вероятнее, что мы увидим точку именно там.

В бесконечномерии, где почти вся область — это ее граница, чтобы найти правильные способы думать и вычислять, нужны самые рафинированные орудия математического арсенала.

Нелинейность и кривизна. Так же как идея линейности экстраполирует малые приращения, идея кривизны использует такую экстраполяцию для изучения отклонения геометрического объекта (например, графика функции f) от линейного.

Малые размерности дают пищу интуиции, вырабатывающей геометрический образ кривизны. График кривой $y=ax^2$ в вещественной плоскости имеет три основные формы: «чаша» (выпуклость вниз) при $a>0$, «купол» (выпуклость вверх) при $a<0$ и горизонтальная прямая при $a=0$. Число a определяет крутизну стенок чаши или купола, а также радиус кривизны в нижней (верхней) точке, который равен $\frac{1}{2|a|}$. В физических моделях малых колебаний тяжелый шарик, катающийся по дну чаши, совершает такие же колебания, как на пружине, и эквивалентная кривизна выражается через массу грузика и жесткость пружины. В многомерии график квадратичной функции при подходящем выборе системы координат приводится к виду $y=\sum_{i=1}^{N} a_i x_i^2$. Среди чисел могут быть положительные, отрицательные и нули, они определяют количество направлений, по которым график уходит вверх, вниз или остается горизонтальным. В современной квантовой теории поля удивительно велико количество ситуаций, где эта простая модель отвечает за вид спектра масс элементарных частиц и спектра сил (констант) взаимодействий, служа первой ступенькой на долгом пути к более изощренным схемам. При этом квадратичные функции возникают в бесконечномерном пространстве, горизонтальные направления в графиках появляются из-за действия группы симметрий; когда функция не меняется при некоторых движениях пространства по себе, ее график не может быть ямой, а в лучшем случае имеет вид оврага (овраг можно сдвигать по себе вдоль дна, а яму нельзя).

Итак, первый образ нелинейности, который мы вкратце описали, это образ того, как многомерная поверхность, график функции удаляется вблизи своей точки от линейной поверхности, касающейся ее в этой точке. Представив себе касательное пространство горизонтальным, мы можем отметить в нем набор попарно ортогональных направлений, вдоль каждого из которых поверхность уходит вверх, вниз или остается горизонтальной; скорость ухода вверх или вниз измеряется радиусом кривизны; этих радиусов столько, какова размерность поверхности.

Для описания искривления мы пользуемся, стало быть, «внешним лекалом». Этот круг представлений естествен и полезен, но эйнштейновская теория тяготения и, как было понято, максвелловская теория электромагнетизма, а также, как мы начинаем понимать сейчас, теория ядерных сил и, может быть, всех взаимодействий вообще требует более тонких представлений о кривизне. Первые математические теории «внутренней кривизны» в отличие от описанной «внешней» были развиты Гауссом и Риманом.

Понятие внутренней кривизны строится сначала для области в числовом пространстве, в которой для каждой пары близких точек задано расстояние между ними. Геометрия нашей области с новым, римановым понятием расстояния должна «в бесконечно малом» быть евклидовой. Разные аспекты понятия кривизны показывают, насколько эта геометрия все же не совпадает с плоской евклидовой.

Чтобы объяснить их, удобно начать с аналога прямых в этом многообразии — это геодезические — кривые наименьшей длины, соединяющие точки пространства (например, дуги больших кругов на сфере). Длина кривой, конечно, измеряется интегралом: кривую нужно разбить на много маленьких отрезков и заменить длину каждого отрезка расстоянием между его концами.

Теперь вообразим себе маленький вектор в многообразии, движущийся вдоль геодезической кривой так, что его угол с геодезической все время остается неизменным. (На двухмерной поверхности этот рецепт определяет движение, а в многомерном случае его нужно еще уточнить.) Поскольку в малом пространство близко к евклидову, этим представлениям нетрудно придать точный смысл. Такое движение вектора называется его параллельным переносом. Можно определить и параллельный перенос вдоль любой кривой: как и для вычисления длины, ее следует заменить ломаной из коротких отрезков геодезических, и затем вектор переносить параллельно вдоль этих геодезических.

Рассмотрим маленькую замкнутую кривую в пространстве — почти плоскую петельку. Перенеся вектор вдоль нее параллельно и вернувшись в начальную точку, мы обнаружим, что вектор повернулся относительно своего начального положения на маленький угол, и этот угол пропорционален площади петельки. Сверх того, коэффициент пропорциональности зависит: а) от точки, вокруг которой расположена петелька; б) от направления двухмерной площадки, которую можно натянуть на петельку. Этот коэффициент, как функция точки и двухмерного направления в ней, называется римановым тензором кривизны. Для плоского евклидова пространства тензор кривизны тождественно равен нулю.

Понадобилось много времени, чтобы понять, какие образы в этой конструкции важнее всего, и прийти к выводу, что самой фундаментальной является идея параллельного переноса вдоль кривой. Геометрическая картина кривизны, которая наиболее актуальна для понимания, например, теории полей Янга — Миллса в современной физике, более обща, чем картина римановой кривизны. Для определения римановой кривизны мы переносили вектор вдоль кривой в пространстве. Представим себе этот вектор в виде маленького гироскопа, а кривую — в виде его мировой линии в четырехмерном пространстве — времени (подробнее об этом будет рассказано ниже, в главе о пространстве — времени). Тогда предсказание о том, как будут различаться направления осей двух гироскопов, разошедшихся в одинаковом начальном состоянии и затем соединенных для сравнения вновь в близких точках пространства — времени, есть прямое дело физики. В то же время воображаемый набор поведений всевозможных таких гироскопов есть математический образ пространства, дополненного правилами параллельного переноса касательных векторов вдоль кривых в нем.

Остается один шаг до введения общего математического понятия пространства со связностью и кривизны этой связности. Направление оси гироскопа является частным случаем представления о том, что точечная физическая система может обладать еще внутренними степенями свободы. В классической физике это идеализация, в соответствии с которой мы суммарно учитываем составные части системы, их вращения, колебания и т. п.; в квантовой физике появляются неклассические степени свободы, такие, как спин или магнитный момент электрона, не сводящиеся к воображаемому поведению «частей» электрона в пространстве — времени. Пусть вообще задана пара пространств M и E и отображение $f: E \to M$, скажем, модель пространства — времени M, в каждой точке m которого находится локализованная физическая система с пространством внутренних состояний $f^{-1}(m)$. Тогда связность на этом геометрическом объекте есть задание правила переноса системы вдоль кривых в M. Иными словами, если мы знаем отрезок мировой линии системы в M и ее начальное внутреннее состояние, то мы должны знать всю ее историю. Кривизна связности измеряет разницу конечных состояний системы, пришедших из близких начальных точек пространства — времени в близкие конечные разными путями, если сначала системы были в одинаковом состоянии.

Эти представления связывают геометрию с физикой напрямую, минуя сложные извивы гениальных догадок, ошибки, формализм и исторические случайности, сопровождавшие возникновение новых идей и постоянно переизлагаемые в учебниках.

Гравитационное поле — это связность в пространстве внутренних степеней свободы гироскопа, управляющая его эволюцией в пространстве — времени. Электромагнитное поле — связность в пространстве внутренних степеней свободы квантового электрона, управляющая его эволюцией в пространстве — времени. Поле Янга — Миллса — связность в пространстве цветовых внутренних степеней свободы кварка.

Сейчас эта геометрическая картина представляется наиболее универсальной математической схемой для классического описания идеализированного мира, в котором по очереди рассматривается небольшое число основных взаимодействий. Материя в пространстве — времени описывается сечением соответствующего расслоения $E \to M$ — указанием того, в каком состоянии эта материя находится в каждой точке в каждый момент. Поле описывается связностью в этом расслоении. Материя влияет на связность, накладывая ограничения на ее кривизну, а связность влияет на материю, заставляя ее «переноситься параллельно» вдоль мировых линий. Великие Уравнения Эйнштейна, Максвелла — Дирака и Янга — Миллса являются точным выражением этих идей.

Но даже не записывая Уравнений, мы сказали очень многое. Открытие того, что основные физические поля суть связности, не было ни столь драматичным, ни столь точно датированным, как открытие этих Уравнений. В теории тяготения, например, основным понятием для Эйнштейна была не связность, а (псевдо) риманова метрика в пространстве — времени. Что электромагнитное поле является связностью, впервые предположил Герман Вейль, но в доквантовой физике он не смог указать, на каком расслоении эта связность определяет параллельные переносы, решив, что поле меняет длины отрезков, прошедших по разным путям в пространстве — времени. На неправдоподобность этого указал Эйнштейн, а правильное расслоение, на котором действует связность Максвелла, открыл Дирак. Но все равно осознание физических особенностей поля связности затянулось так надолго, что лишь в шестидесятых годах Ааронов и Бом предложили эксперимент, который показывает истинно «связностную» природу поля Максвелла. Для этого следует разделить на две части электронный пучок и пустить эти две части в обход цилиндрической области, внутри которой заключен магнитный поток, после чего наблюдать интерференционную картину на экране. По их предположению при включении и выключении магнитного поля картина должна меняться, хотя пучки, обходя область магнитного потока, проходят в области, где напряженности электромагнитного поля нулевые. Таким образом, разделенные и вновь соединенные на экране пучки будут «чувствовать» кривизну связности на расслоении Дирака, возникающую от включения поля в области, которую они обходят. Интерференционная картина на экране отражает именно разность углов поворота фаз в спиновом пространстве степеней свободы электрона, появляющуюся из-за того, что электрон может прийти в точку экрана разными путями в обход магнитного поля. (Эксперимент был реально проведен и подтвердил эти предсказания.)

Некоторые новинки. «Выставка» важнейших геометрических образов, по которой мы торопливо провели читателя, далеко не исчерпывается показанными экспонатами. Число таких образов пополняется. Из тех, которые начали привлекать внимание физиков и математиков в последнее время, можно, например, назвать «катастрофы», «суперсимметрии » и «солитоны».

Термин «катастрофы» ввел французский математик Рене Том для передачи интуитивных представлений, связанных с математическими схемами описания таких явлений, как разрывы, скачки, углы, поверхности раздела между однородными фазами, биологическая дифференциация тканей и т. п. Их польза в естественнонаучных моделях пока остается под вопросом и стала даже предметом горячих споров в газетах. Историку науки предоставляется счастливый случай наблюдать попытки установления новой парадигмы в смысле Томаса Куна и размышлять над социальными аспектами процесса установления научного общественного мнения.

«Суперсимметрии», изучаемые в супергеометрии, начинают входить в науку с меньшим шумом, хотя, возможно, окажут больше влияния на дальнейшее развитие физики и геометрии. Формально говоря, супергеометрия предлагает рассматривать на пространстве, скажем R^n, не только обычные функции, но также антикоммутирующие, т. е. удовлетворяющие условию $fg = -gf$, откуда, в частности, следует, что $f^2 = 0$. Так как ненулевых чисел с нулевым квадратом нет, такая функция не может принимать числовые значения; ее естественные области значений — так называемые грассмановы алгебры, введенные в прошлом веке замечательным математиком и санскритологом Грассманом. В физике грассмановы алгебры появились лишь после возникновения квантовой теории и понятия о спине; оказалось, что адекватное описание коллектива тождественных частиц с полуцелым спином, например электронов, требует введения антикоммутирующих переменных.

Образ солитона возник в результате открытия некоторых специальных решений уравнений, описывающих волны в разных средах, например на воде. Классические волновые уравнения линейны, т. е. сумма их решений и произведение решения на число также являются решениями. Иными словами, это уравнения описывают волны, не взаимодействующие между собой. Учет таких свойств реальных сред, как дисперсия (нелинейная связь между частотой и длиной элементарной волны) и нелинейная зависимость скорости волны от ее амплитуды приводит к гораздо более сложной картине взаимодействия волновых возмущений, чем простое их сложение. Поэтому крайнее удивление вызвало открытие в шестидесятых годах группой американских физиков и математиков эффекта нелинейного сложения некоторых уединенных возбуждений, описываемых уравнением Кортевега-де Фриза (эти уединенные возбуждения и были сначала названы солитонами). Высота солитонной волны пропорциональна ее скорости. Поэтому можно попытаться проследить судьбу суммы двух далеко разнесенных в начальный момент солитонов, из которых больший движется в сторону меньшего и потому обязательно догонит его. Общее ожидание состояло в том, что после «столкновения» волновая картина разрушится, но машинный эксперимент показал, что ничего подобного не происходит: после периода взаимодействия больший солитон «проходит сквозь меньший», и оба начинают расходиться, сохранив свою форму. Точная математическая теория явления была построена вскоре после этого — она подтвердила сохранение индивидуальности солитонов после взаимодействия, сколько бы их ни было вначале. После этого число нелинейных волновых уравнений, обнаруживающих аналогичные свойства, росло линейно со временем, а число публикаций, посвященных им, росло экспоненциально. Высказываются надежды, что солитоноподобные возбуждения полей являются адекватным классическим образом элементарных частиц: на новом идейном уровне возрождается столетней давности идея Ранкина и Томсона о том, что атомы суть «вихревые кольца основной жидкости». Дело в том, что уравнения для связностей Янга — Миллса в отличие от уравнений Максвелла нелинейны.

Небольшая историческая справка о первооткрывателях солитона, содержащая нравоучительные детали. Дидерик Иоханнес Кортевег родился в 1848 году и умер в 1941 году в Голландии. Он был известным ученым, и его памяти посвящено несколько некрологов. Ни один из некрологов даже не упоминает работы, в которой был открыт солитон. Сама эта работа представляет собой, в сущности, отрывок из диссертации Густава де Фриза, выполненной под руководством Кортевега и защищенной 1 декабря 1894 года. Де Фриз был гимназическим учителем, и о нем почти ничего не известно.

Множества, формулы и расщепленный мозг. Каково соотношение между математическим текстом и его содержанием в широком смысле слова (множественностью его потенциальных содержаний)? Мы пытались показать, что между уравнениями, скажем, Максвелла и их прямым истолкованием в терминах физических понятий должен быть построен промежуточный теоретико-множественный образ, интерпретация - посредник, функционально подобный языку-посреднику в современных лингвистических моделях машинного перевода. На самом деле внимательный анализ научного мышления позволил бы обнаружить целую иерархию языков - посредников, участвующих в потенциальном объяснении таких понятий, как «число», «фотон» или «время». Однако большинство этих объяснений существует в непроявленном, незаконченном и зыбком облике, часто специфичном для индивидуального сознания, поддающемся коммуникации лишь в той мере, в какой удается использовать средства естественного языка. Естественный язык играет огромную роль в открытии, обсуждении и хранении научных знаний, но очень плохо приспособлен к точной передаче содержания этих знаний и той их обработке, которая составляет важную часть научного мышления. У него иные функции и иные достоинства.

Язык современной, теоретико-множественной математики может осуществлять роль такого языка-посредника благодаря его уникальной способности одновременно формировать геометрические, пространственные, кинематические образы и максимально точную запись их математического содержания в формализме. Канторовское определение множества, которое приведено выше, с долей иронии называли «наивным», сравнивая его с определением точки по Евклиду как «места без длины и ширины». Эта критика связана с непониманием того, что фундаментальные понятия математики, в данной системе не сводимые к более элементарным, обязательно должны вводиться двумя способами: содержательным («наивным») и формальным. Целью содержательного определения — создание первоначального, еще не вполне оформленного образа, настройка разных индивидуальных сознаний на один лад, как камертоном. Формальное же определение вводит, собственно говоря, не понятие, а термин, не образ «множества» в структуру сознания, а слово «множество» в структуру допустимых языковых текстов о множествах, которые описываются правилами их порождения примерно так же, как инструкции по АЛГОЛУ описывают правила составления программы. В пределе идеализации вся математика может предстать как потенциальная совокупность грамматически правильных текстов на формальном языке.

В этом образе есть странная и для многих притягательная эстетика уродливости. Возник он в работах мыслителей, задумывавшихся над тем, как согласовать веру в абсолютную истинность математических принципов с абстракциями бесконечных множеств, бесконечных процессов проверок и т. п., через которые эта истинность вводится. Исходная гипотеза Давида Гильберта состояла в том, что эти абстракции, строго говоря, не нуждаются в такой «почти физической» и потому сомнительной интерпретации и что их можно считать чисто языковыми фактами. «Бесконечность» — это слово, а не явление, помогающее каким-то образом узнать истины о конечных вещах. Мы уже упоминали, что позже Гёдель показал, что такое языковое понятие «доказуемой истины» является несравненно более узким, чем абстракция истины, вводимой через идеи бесконечных проверок.

Внешние, естественнонаучные, прикладные, в широком смысле слова, аспекты математического знания при их гносеологическом анализе позволяют понять кое-что о математическом творчестве и диалектике его взаимоотношении с гёделевским запретом. Принятие интерпретации формализма, физической в том или ином смысле этого слова, вера в адекватность этой интерпретации и знание каких-то черт поведения физического явления позволяет указать или постулировать математические истины, не доступные «чистой интуиции». Это — источник расширения самой базы математического знания.

В более частном плане соотношение между математическим символизмом, неформальным мышлением и познанием природы в последние годы стало возможно рассматривать с точки зрения новых данных о структуре и функциях центральной нервной системы.

Мозг состоит из двух полушарий, левого и правого, которые перекрестно связаны с правой и левой половинами тела. Нейронные связи между полушариями проходят через мозолистое тело и комиссуры. В нейрохирургической практике известен метод лечения, в частности, тяжелых эпилептических припадков, состоящий в рассечении мозолистого тела и комиссур, что прерывает прямые связи между полушариями. После такой операции у больных наблюдается необычная картина «двух сознаний». По лаконичной формулировке американского нейропсихолога К. Прибрама, результаты исследования таких больных, а также больных с различными поражениями левого и правого полушарий, можно резюмировать следующим образом: «У правшей левое полушарие обрабатывает информацию во многом подобно цифровой вычислительной машине, тогда как правое полушарие функционирует скорее по принципам оптических и голографических систем обработки данных». В частности, левое полушарие содержит генетически заданные механизмы усвоения естественного языка и, более общо, символизма, логики, «рацио»; правое, молчаливое полушарие ведает образами, целостным восприятием, интуицией. Функционирование человеческого сознания в норме постоянно обнаруживает это сочетание двух компонент, одна из которых может проявляться заметнее другой, и открытие их физиологических носителей проливает свет на природу и типологию математических интеллектов и даже школ в проблеме оснований математики. Можно строить догадки о том, что два великих интеллекта, стоявших у колыбели современной математики, — Ньютон и Лейбниц — принадлежали соответственно к правополушарному и левополушарному типам. Ньютону мы обязаны созданием математического анализа и первыми фундаментальными результатами математической физики — закон всемирного тяготения, вывод из него законов Кеплера, теория приливов. Лейбниц же ввел обозначения, в частности $\int y dx$, которыми мы пользуемся и поныне. По словам историка математики Д. Стройка, он был «одним из самых плодовитых изобретателей математических символов» и даже свою версию математического анализа изобрел в результате поисков универсального языка. (Интересно, что Ньютон, также не избежавший этого поветрия времени, не создал формализма анализа и доказательства излагал геометрически.)

Принимая современное представление о функциональной асимметрии мозга, можно высказать предположение о том, что язык теории множеств позволяет кратчайшим путем достичь сбалансированной активности правого и левого полушарий работающего математика, чем и объясняется его замечательная эффективность.

Я хотел бы в заключение привести слова И. А. Соколянского, посвятившего жизнь воспитанию слепоглухонемых детей. Они содержатся в письме к Вяч. Вс. Иванову, из книги которого «Чет и нечет. Асимметрия мозга и знаковых систем» (М., «Советское радио», 1978) почерпнута и следующая информация.

Если у слепоглухонемого ребенка не поражены отделы центральной нервной системы, ведающие наглядным восприятием внешнего мира, то его можно научить языку, даже звуковому, и обеспечить полное развитие его личности. Этот процесс происходит в несколько этапов. Сначала ребенок поддерживает постоянный контакт с матерью или воспитательницей, держится за руку или юбку, ходит по дому, ощупывает предметы ее действий, и на этой основе вырабатывает язык жестов, в той или иной мере имитирующий действия и свойства предметов. В норме это функция правого полушария. Открытие Соколянского состояло в том, что на следующем этапе можно научить ребенка перекодировать язык жестов в пальцевую азбуку, так что жест-иероглиф замещается жестом-словом. Символ перевода — специальный жест, подобный математическому знаку равенства, — две вытянутые параллельно ладони. Смысл этого перекодирования состоит в том, что информация передается в левое полушарие, которое, будучи предрасположено к научению дискретному и символическому языку (не обязательно звуковому!), начинает развивать эту функцию практически с той же скоростью, что и у здорового ребенка, — овладение языком происходит за два-три года. Синтаксис такого левополушарного языка отличен от «синтаксиса мира», запечатлеваемого в правом полушарии, и тождествен синтаксису словесного естественного языка. Семантика же его, видимо, более ограничена или, во всяком случае, неадекватна семантике зрячего и слышащего. Как объяснить, что значит «звезда», тому, кто никогда не увидит звезд? Соколянский дает замечательный ответ: «Словесная речь, как бы ею ни овладели безъязычные, сама по себе не может обеспечить слепоглухонемому полноценное умственное развитие в такой степени, чтобы он мог отразить внешний физический мир так, как это доступно нормальному человеку. Истинная картина этого мира может быть раскрыта только математически развитым мышлением...»

Что такое звезда, спрашивают и те, кто видит звезды, потому что видеть глазами — это еще очень мало.

2. ФИЗИЧЕСКИЕ ВЕЛИЧИНЫ, РАЗМЕРНОСТИ И КОНСТАНТЫ: ОТКУДА В ФИЗИКЕ БЕРУТСЯ ЧИСЛА

«Главная цель физических теорий — найти число, и притом с достаточной точностью!» (Р. Фейнман).

Это преувеличение. Главная цель физических теорий — понимание. Способность теории найти число — полезный критерий правильности понимания.

Числа в физике — чаще всего значения физических величин, описывающие состояния физических систем. Величины — это родовое имя для таких абстракций, как расстояние, время, энергия, действие, вероятность, заряд и т. п. В свою очередь, состояние системы характеризуется значениями на нем достаточно полного набора физических величин, а систему естественнее всего описывать заданием множества возможных ее состояний. Выйти из этого логического круга, ограничиваясь чисто словесными описаниями, нельзя. Он может быть разорван в двух местах — операционально, когда мы объясняем, как измерить массу Земли или электрона, и математически, когда мы предлагаем теоретическую модель системы или класса систем и объявляем, что масса m — это, скажем, коэффициент в формуле Ньютона $F=ma$.

Содержательная, хотя и простая математика, связанная с физическими величинами, начинается с напоминания о том, что значения физической величины (точнее, скалярной вещественной величины) можно отождествлять с числами, вообще говоря, только после выбора единицы измерения и начала отсчета (нуля). Разумно не вносить этого произвола как можно дольше — некоторые из самых фундаментальных физических законов гласят, что у определенных физических величин имеются естественные единицы. Разберемся в этом подробнее.

Спектр скалярной величины. Назовем спектором величины множество всех значений, которые она может принимать (на состояниях данной системы, определенного класса систем, «всех» систем — это следует уточнять по мере необходимости). Основной математический постулат, который можно считать определением скалярной величины в теоретических моделях, состоит в том, что спектр всегда является подмножеством одномерного аффинного пространства над вещественными числами. Иными словами, он лежит на прямой, где не отмечены нуль и единица; если две такие точки отметить, спектр превратится в множество вещественных чисел. Вся соль в том, что иногда эти точки можно отметить не как попало, а пользуясь самим спектром. Вот основные примеры.

а) Скорость. Наименьшую (относительную) скорость естественно назвать нулем. Вторая отмеченная точка на спектре скоростей — это c, скорость света. Общепринятый (после создания специальной теории относительности) постулат о спектре скоростей состоит в том, что он заполняет отрезок от нуля до c. Тогда естественно объявить c единицей скорости и считать, что все скорости заполняют отрезок [0, 1]: в более обычных обозначениях так ведут себя отношения v/c. В обыденной жизни мы редко встречаемся со скоростями, большими 10^{-6} по этой шкале (скорость звука).

б) Действие. Это, может быть, самая важная величина во всей теоретической физике, и мы посвятим ей отдельную главку. Она принимает значения не на мгновенных состояниях, а на отрезках истории физической системы. В классической физике она определяет физически возможные отрезки истории — на них действие принимает наименьшие допустимые значения. Естественный нуль на спектре действия — это действие «бесконечно короткой» истории системы. Верхней границы спектра действия мы не знаем. Можно представить себе космологическую модель, где этой границей будет действие Вселенной на всем отрезке ее истории от Большого Взрыва до Большого Коллапса, если последний предсказывается моделью.

Тем не менее вторая отмеченная точка на спектре действия известна: это знаменитая постоянная Планка h. В человеческих масштабах она крайне мелка — действие ручки, написавшей слово «действие», имеет порядок 10^{29}—10^{30} h. Прагматически говоря, h указывает, когда следует пользоваться квантовомеханическими, а не классическими моделями: в тех случаях, когда нас интересуют такие подробности истории системы, на которых действие меняется всего на несколько h. (Впрочем, это условие не необходимо и не достаточно.) Выбирая h в качестве единицы действия, мы можем считать, что спектр действия есть полупрямая [0, ∞), а спектр приращений действия — вся вещественная прямая.

Таким образом, точка h на спектре действия «не видна» в отличие, скажем, от c, которая является правым концом своего спектра. Это очень странно. Впрочем, есть два контекста, в которых h проявляется.

Один из них связан со спектром спина — внутреннего момента количества движения элементарных частиц. Спин имеет ту же размерность, что и действие, и состоит из целых кратных $\frac{\hbar}{2} = \frac{h}{4\pi}$. Не означает ли это, что спин есть истинно фундаментальная величина, а действие — лишь пережиток классической физики?

Второй контекст — это знаменитое соотношение неопределенностей Гейзенберга. Квантовые модели определяют разбиение системы классических величин на пары сопряженных: координата — проекция импульса, энергия — время. Размерность произведения сопряженных величин есть размерность действия. Принцип неопределенности в словесной формулировке утверждает, что оба члена пары сопряженных величин не могут одновременно принимать точного значения ни на каком состоянии систем. Произведение неточностей ограничено снизу величиной $\frac{\hbar}{2}$. Применяя этот принцип к энергии и времени, мы получаем формально соотношение $\Delta E \Delta t \geqslant \frac{\hbar}{2}$, содержательный смысл которого многократно обсуждался в физической литературе. С нашей точки зрения, оно означает, что представление о классическом отрезке истории системы, на котором действие меняется меньше чем на $\frac{\hbar}{2}$, лишено смысла. Позже мы подробнее обсудим трудный вопрос о сравнительном смысле одноименных классических и квантовых величин.

в) Масса. По Ньютону, значения инертной массы можно приписать стабильным материальным телам. Наименьшие объекты, к которым ньютоновское понятие массы еще применимо без принципиальных оговорок, — электрон и протон. Они приводят к двум точкам на спектре масс (кроме нуля): m_e и m_p. Характерная масса человеческих масштабов определяется с помощью числа Авогадро — $6{,}02 \cdot 10^{23}\ m_p$. Отношение $m_p/m_e \approx 1840$ является первым истинно фундаментальным числом, которое мы до сих пор встретили, в отличие от точек спектра, которые числами, строго говоря, не являются.

Теория, которая его объяснит, наверное, будет важной теорией.

Другие элементарные частицы определяют другие точки на спектре масс; измеряя их в единицах m_e или m_p, мы получаем кучу чисел, нуждающихся в теоретическом объяснении.

г) Гравитационная постоянная. Если две точечные массы m_1 и m_2 находятся на расстоянии r друг от друга и притягиваются с силой F, обусловленной только ньютоновским гравитационным взаимодействием, то величина $\frac{Fr^2}{m_1 m_2}$ не зависит от m_1 и m_2. Она была открыта Ньютоном и обозначается G.

Последний пример идейно сложнее предыдущих: для введения G мы должны явно апеллировать к «физическому закону». Кроме того, мы получили точку нового спектра — спектра констант связи фундаментальных взаимодействий, к которым относятся еще электромагнитное, сильное и слабое взаимодействия.

В этом месте пора ввести следующую крупную группу физических абстракций.

Физический закон, размерность и подобие. Для нужд этого пункта под «физическим законом» будем понимать содержание таких формул, как $F=ma$, $F=G\frac{m_1 m_2}{r^2}$ (Ньютон), $E=h\nu$ (Планк), $E=mc^2$ (Эйнштейн) и т. п. Физическая теория, скажем, механика Ньютона или электромагнитная теория Максвелла, с математической стороны включает в себя указание следующих данных: а) основные величины теории; б) основные связывающие их законы. Кроме того, с операциональной стороны, должны быть описаны: в) физические ситуации, в которых можно применять теорию; г) принципы сопоставления теоретических высказываний с измерениями и наблюдениями.

Мы занимаемся лишь первой частью. Зная величины и связывающие их законы, мы можем построить фундаментальную математическую характеристику теории — ее группу размерностей D. На математическом языке это абелева группа, которую можно задать образующими и соотношениями: образующие — это физические величины теории, а соотношения определяются условием, чтобы все законы теории были однородными. Класс величины в группе D называется размерностью этой величины. Можно выбрать основные величины, которые в группе D составят независимую систему образующих; размерности остальных величин теории будут выражаться через них в виде формальных одночленов. Единицы основных величин определят единицы остальных. (Все это — сжатое изложение принципов, лежащих за такими школьными обозначениями, как, например см/с².) Мы отметим несколько обстоятельств, в которых явное введение группы помогает разобраться в существе дела.

Группа размерностей ньютоновской механики. Она порождена размерностями длины L, времени T и массы M.

Закон $F=ma$ показывает, что сила имеет в этой группе размерность MLT^{-2}, энергия (сила $\times$ длина) — ML^2T^{-2}, а действие (энергия $\times$ время) — ML^2T^{-1}.

Прогресс физики постоянно сопровождается двумя противоположными процессами: увеличением группы размерностей D в силу открытия величин новой природы (электромагнетизм после Ньютона; новые квантовые величины, такие, как «странность», «очарование», в наши времена) и уменьшением этой группы в силу открытия новых законов, которые дают соотношения между прежде независимыми размерностями.

Чтобы понять этот второй процесс, вернемся к ньютоновской гравитационной постоянной G. Размерность ее есть по предыдущим правилам сила $\times$ (длина) ^{2} $\times$ (масса)$^{-2}=M^{-1}L^3T^{-2}$. Ее числовое значение, таким образом, зависит от выбора единиц массы, длины и времени. Постоянна же она в том смысле, что после выбора таких единиц ее числовое значение, полученное по формуле $Fr^2(m_1m_2)^{-1}$, где F, r, m_1, m_2 измеряются в разных экспериментах типа эксперимента Этвеша или вычисляются по данным астрономических наблюдений, не зависит от переменных величин этих экспериментов: r, m_1, m_2.

После установления этого физического факта мы можем использовать его для построения уменьшенной группы размерностей D' теории «механика Ньютона» + «гравитация Ньютона». Эта уменьшенная группа математически является фактор-группой D по подгруппе, порожденной всеми степенями $M^{-1}L^3T^2$. В качестве основных размерностей в D' можно выбрать любую пару (ML), (MT) или (LT), а оставшуюся размерность выразить через эту пару и размерность G. Соответственно число основных единиц, отвечающих D', уменьшается до двух, если выбрать G в качестве единицы измерения размерности $M^{-1}L^3T^2$. На этом примере также виден физический смысл отмеченных точек спектров: это точки, воспроизводимые в серии экспериментов некоторого типа, изолирующих определенные взаимодействия, системы определенного сорта и т. д.

Масштабная инвариантность. Группу подобия, или масштабной инвариантности, D^* данной теории с математической точки зрения можно определить как состоящую из *характеров* группы размерностей D, т. е. из отображений χ группы D в положительные вещественные числа со свойством мультипликативности: $\chi(d_1d_2)=\chi(d_1)\chi(d_2)$ для всех $d_1, d_2 \in D$. Эта группа имеет прямой физический смысл:

она показывает, в какой пропорции можно увеличивать (или уменьшать) разные характеристики явления, не выводя его за пределы применимости теории. Если все законы теории известны, D^* вычисляется тривиально. Польза D^* состоит в том, что иногда ее можно угадать из физических соображений до того, как становится известным точный вид этих законов. Тогда оказывается, что D^* несет о них важную информацию. Известный пример классического открытия, сделанного таким способом, — закон Вина $\varepsilon(\nu, T)=\nu^3 F(\nu/T)$ для испускательной способности абсолютно черного тела как функции частоты и температуры. Он отвечает характеру $\chi_a([\nu])=a$, $\chi_a([\varepsilon])=a^3$, $\chi_a([T])=a$ в группе D^*, где $[\varepsilon]$, $[\nu]$, $[T]$ — соответствующие размерности; a — любое вещественное число. Можно упомянуть еще соображения Галилея о размерах животных и многочисленные приложения теории подобия в гидро- и аэродинамических расчетах. На уровне фундаментальных теорий группа D^* является простейшим примером групп симметрий, которые в физике элементарных частиц и в квантовой теории поля все чаще выступают в роли самостоятельных физических законов высшего уровня, накладывающих жесткие ограничения на вид законов следующего уровня, например лагранжианов. Важнейшие из этих групп — некоммутативные и комплексные, как группы унитарных вращений $U(n)$, потому что в квантовой механике основные величины лежат в многомерных комплексных пространствах, а не одномерных вещественных. На это уже другая история.

Планковские единицы и проблема единой физической теории. В ньютоновской физике нет других естественных единиц, кроме G. Скорость света c может быть объявлена естественной единицей лишь внутри новой теории, постулирующей ее особую роль как верхнего предела скоростей распространения материальных тел (недостижимого) или сигналов (достижимого), как инварианта относительно смены инерциальной системы координат, и т. п. Подобным же образом планковская единица действия $\hbar$ стала гербом новой физической теории — квантовой механики.

Однако c и $\hbar$, так же как G, имеют вполне определенные размерности в ньютоновской группе D: LT^{-1} для c и ML^2T^{-1} для $\hbar$. Выбрав G, c и $\hbar$ в качестве основных единиц соответствующих размерностей, мы обнаруживаем, что имеется естественный масштаб, делающий значения всех вообще физических величин, выразимых в D, вещественными числами, т. е. имеются естественные единицы всего на свете! В самом деле, размерности $M^{-1}L^3T^{-2}(G)$, $LT^{-1}(c)$ и $ML^2T^{-1}(\hbar)$ порождают всю группу D (если уж быть совсем точным, то они порождают подгруппу индекса два).

В частности, естественные единицы длины, времени и массы — знаменитые единицы Планка — суть:

$$L^*=(\hbar G/c^3)^{1/2}=1{,}616\cdot10^{-33}\text{ см};$$
$$T^*=(\hbar G/c^5)^{1/2}=5{,}391\cdot10^{-44}\text{ с};$$
$$M^*=(\hbar c/G)^{1/2}=2{,}177\cdot10^{-5}\text{ г}.$$

У читателя должен возникнуть вопрос — почему же мы ничего не измеряем в планковских единицах? Прагматический ответ: потому что они определяют совершенно несуразные масштабы. Боровский радиус равен $3{,}9\cdot10^{-11}$ см: L^* меньше него на 22 порядка! Во столько же раз T^* меньше времени, за которое свет проходит боровский радиус. С другой стороны, M^* — это масса вполне макроскопической пылинки, содержащей примерно 10^{19} протонов. Планковская единица плотности M^*/L^{*3} равна $5\cdot10^{93}$ г/см, ничего отдаленно подобного этому ни в каких условиях мы не можем даже вообразить.

Более содержательное замечание состоит в том, что у нас на самом деле нет единой физической теории, в которой бы фигурировали одновременно G, c и $\hbar$. Уже теории, соединяющие эти константы попарно, являются крупнейшими достижениями двадцатого века: (G, c) — это общая теория относительности; $(c, \hbar)$ — это релятивистская квантовая теория поля, сравнительно завершенная лишь для электромагнитных взаимодействий. К фрагментам будущей $(G, c, \hbar)$-теории относятся расчеты квантового рождения частиц в сильных классических гравитационных полях, в частности, вблизи черных дыр малой массы (С. Хоукинг). Пока полностью квантовой $(G, c, \hbar)$-теории не существует, планковские единицы остаются отдаленными пограничными столбами обширной неисследованной территории.

Можно посмотреть на это странное несоответствие порядков величин естественных единиц друг с другом и с привычными единицами с другой точки зрения. В гипотетических фундаментальных уравнениях единой теории — «всеобщей теории всего» (Станислав Лем) — разные члены (теперь безразмерные числа!) будут принимать (благодаря этому несоответствию) очень резко отличающиеся по величине значения в зависимости от масштабов области пространства (времени, импульсов, энергий), в которой помещаются изучаемые нами явления. Самые маленькие члены можно будет отбросить с ничтожной ошибкой, придя к одной из приближенных теорий. Так и происходит на границах применимости известных ныне моделей, когда мы описываем мир классически в человеческих масштабах (считая $v/c = 0$ и $\hbar/S = 0$, где v — типичные скорости; S — типичные действия) или не учитываем гравитацию в микромасштабах, полагая $G = 0$.

На самом деле это содержащее долю истины рассуждение крайне наивно. Настоящая смена теории не есть смена уравнений — это смена математических структур, и лишь фрагменты конкурирующих теорий, часто не самые важные идейно, допускают сравнение друг с другом на ограниченном круге явлений реальности. «Гравитационный потенциал» Ньютона и «кривизна метрики Эйнштейна» описывают разные миры на разных языках.

Кроме того, жизнь — может быть, самое интересное физическое явление — вышита на ажурной канве игры неустойчивостей, когда несколько квантов энергии могут иметь огромную информационную ценность, а отбрасывание малых членов в уравнениях означает смерть.

Классификация физических констант. Подведем некоторые итоги. Справочник «Таблицы физических величин» (М., Атомиздат, 1976) содержит 1005 страниц текста и многие миллионы чисел; как в них разобраться? Эти величины делятся по крайней мере на четыре типа.

а) Естественные единицы измерения, или физически отмеченные точки спектров. Это — не числа, а такие величины, как G, c, h, m_e, e (заряд электрона). Это — размерные характеристики некоторых явлений, поддающихся воспроизведению многократно, с высокой степенью точности. Это — отображение того, что природа тиражирует элементарные ситуации огромными сериями. Размышления над тождественностью подобных кирпичиков мироздания приводили иногда к таким глубоким физическим идеям, как статистики Бозе — Эйнштейна и Ферми — Дирака. Фантастическая мысль Уилера, что все электроны тождественны потому, что представляют собой мгновенные сечения запутанной в клубок мировой линии одного электрона, привела Фейнмана к изящному упрощению диаграммной техники вычислений в квантовой теории поля.

б) Истинные, или безразмерные, константы. Это — отношения нескольких отмеченных точек на спектре величины одной размерности, например, отношения масс электрических частиц: мы уже упоминали m_p/m_e. Отождествление разных размерностей при учете нового закона, т. е. редукция группы размерностей, приводит к объединению прежде разных спектров и к необходимости объяснять новые числа.

Например, размерности m_e, c и $\hbar$ порождают группу Ньютона и потому приводят к столь же естественным атомным единицам размерностей M, L, T, как и единицы Планка. Поэтому их отношения к планковским единицам нуждаются в теоретическом объяснении. Но, как мы говорили, это невозможно, пока отсутствует (G, c, $\hbar$)-теория. Однако и в (m_e, c, $\hbar$)-теории — квантовой электродинамике — имеется безразмерная величина, значению которой современная квантовая электродинамика в некотором смысле слова обязана своим существованием. Поместим два электрона на расстоянии $\hbar/m_e c$ (так называемая комптоновская длина волны электрона) и измерим отношение энергии их электростатического отталкивания к энергии $m_e c^2$, эквивалентной массе покоя электрона. Получится число $\alpha = 7{,}2972 \cdot 10^{-3} \approx 1/137$. Это — знаменитая постоянная тонкой структуры.

Квантовая электродинамика описывает, в частности, процессы, в которых не сохраняется число частиц: вакуум рождает электрон-позитронные пары, они аннигилируют. Из-за того что энергия рождения (не меньшая, чем $2m_e c^2$) в сотни раз больше энергии характерного кулоновского взаимодействия (благодаря значению α), удается провести эффективную схему вычислений, в которой эти радиационные поправки не отбрасываются начисто, но и не «портят жизнь» теоретика безнадежно.

Теоретического объяснения величины α не существует.

У математиков есть свои замечательные спектры: спектры выделенных линейных операторов — генераторов простых групп Ли в неприводимых представлениях, объемы фундаментальных областей, размерности пространств гомологий и когомологий и т. п. Простор для фантазии, отождествляющей спектры математиков и спектры физиков, открыт — нужны скорее принципы, ограничивающие выбор. Но вернемся к константам.

Следующий их тип, занимающий много места в таблицах, это:

в) Коэффициенты пересчета из одних масштабов в другие, например, из атомных в «человеческие». К ним относятся: уже упомянутое число Авогадро $N_0 = 6{,}02 \times 10^{23}$ — по существу, один грамм, выраженный в единицах «масса протона», хотя традиционное определение немного другое, а также такие вещи, как световой год в километрах. Наиболее отвратительны для математика здесь, конечно, коэффициенты перехода от одних физически бессмысленных единиц к другим, столь же бессмысленным: от локтей к футам или от Реомюра к Фаренгейту. По-человечески это иногда самые главные числа; как мудро заметил Винни-Пух: «Не знаю, сколько в нем литров, и метров, и килограмм, но тигры, когда они прыгают, огромными кажутся нам».

г) «Диффузные спектры». Это — характеристика материалов (не элементов или чистых соединений, а обыкновенных технологических марок стали, алюминия, меди), астрономические данные (масса Солнца, диаметр Галактики...) и многие в том же роде. Природа производит камни, планеты, звезды и Галактики, не заботясь об их одинаковости, в отличие от электронов, но все же их характеристики меняются лишь в достаточно определенных пределах. Теоретические объяснения этих «разрешенных зон», когда они известны, бывают замечательно интересными и поучительными.

Серию таких объяснений собрал В. Вайскопф в прекрасной статье «Современная физика в элементарном изложении» («Успехи физических наук», 1971, т. 103, вып. 1, с. 155—179).

Вот пример физического рассуждения из этой статьи, в котором свои роли играют все наши главные герои: «высота гор определяется фундаментальными физическими постоянными». Имеется в виду вот что: самая высокая вершина Земли Джомолунгма (Эверест) имеет высоту около 10 км; почему нет более высоких гор? Оказывается, даже без учета геологических механизмов выветривания и разрушения высота горы ограничена несколькими десятками километров из-за конкретных размеров Земли и значений фундаментальных констант. Аргументы Вайскопфа таковы: гора слишком большой высоты не сможет существовать из-за ожижения своей нижней части под давлением верхней. Подсчет высоты, при которой давление еще не достаточно для ожижения, дает оценку

$$\gamma \frac{\alpha a_0}{\alpha_G} \cdot \frac{1}{N^{1/3}} \cdot \frac{1}{A^{5/3}} \approx 40 \text{ км},$$

где $\gamma = 0{,}02$ — характеристика теплоты плавления (вполне оцениваемая через фундаментальные константы); α — постоянная тонкой структуры, $\alpha_G = Gm_p^2/\hbar c$; $N \approx 3 \cdot 10^{51}$ — число протонов и нейтронов в составе Земли; $A \approx 60$ — средний атомный вес вещества горы. Только число N здесь не фундаментально. Но и его место на диффузном спектре масс планет ограничено фундаментальными постоянными. Вайскопф с помощью совсем грубых оценок показывает, что N не может превосходить примерно 10^{53}, иначе вещество планеты не сможет существовать в виде неионизированных атомов. Наконец, оценка N снизу получается, если потребовать, чтобы высота гор на планете была не больше ее радиуса, т. е. чтобы планета была в основном круглая, иначе и о горах нельзя говорить! Эта оценка приводит к величине больших астероидов.

3. КАПЛЯ МОЛОКА, ИЛИ НАБЛЮДАТЕЛЬ, НАБЛЮДЕНИЕ, НАБЛЮДАЕМОЕ И НЕНАБЛЮДАЕМОЕ

«...Что наблюдалось бы, если не глазами во лбу, то очами умственными, когда орел, несомый силой ветра, выпустит из своих когтей камень?» (Г. Галилей).

Глазами во лбу мои сверстники наблюдали, как летит бомба, когда открывается замок бомбодержателя, на фоне дымного неба, на экранах кинохроник и на тысячах детских рисунков; я сам их рисовал. Попробуем забыть об этом и посмотрим на мир очами умственными, как учил простодушного Симпличио наш вечный современник Галилео Галилей.

Изолированная система. Среди всех абстракций классической физики одной из главных является идея изолированной, или замкнутой, системы. Эта часть Вселенной, эволюция которой в течение некоторого периода существования определяется лишь внутренними законами. Внешний мир или не взаимодействует с системой вовсе, или в некоторых моделях это взаимодействие учитывается суммарно как эффект связей, внешнего поля, термостата (таким образом, мы пользуемся словами «изолированная» «замкнутая» шире, чем общепринято; изолированность относится, скорее, к математической модели). Петли обратной связи нет или она искусственно перерезана. Мир разбирается на детали, узлы и сборки, как в заводских спецификациях. И в самом деле, это идеология не только Человека Размышляющего, но Человека Делающего. Винтики и шестеренки большой машины мира, когда их поведение понято, могут быть собраны и соединены в новом порядке. Так, появляется лук, ткацкий станок или большая интегральная схема.

Для математика изолированная система — это: а) ее фазовое пространство, т. е. множество мгновенных состояний движения системы; б) множество кривых в фазовом пространстве, изображающих возможные истории системы, проходимые ею с течением времени последовательности состояний. Первое — кинематика, второе — динамика. Важно отличать состояние системы от состояния движения: первое традиционно задается координатами, второе — координатами и скоростями; зная лишь координаты, мы не можем предсказать дальнейшее движение системы, но зная координаты и скорости — можем. Предположение о том, что замкнутую систему можно описать хоть каким-то фазовым пространством и системой кривых в нем (иногда все вместе называют фазовым портретом),— это математическое содержание классического принципа детерминизма.

Один из знаменитых парадоксов Зенона Элейского можно истолковать как первую догадку о роли фазового пространства: стрела летящая и стрела неподвижная в каждый момент времени находятся там, где они находятся; чем же отличается полет от неподвижности? Ответ: видимое место стрелы есть лишь проекция на пространство положений ее «истинного места» в пространстве пар (положение, вектор скорости).

Классическая замкнутая система изолирована от всего внешнего мира, значит, и от внешнего наблюдателя. Она изолирована от воздействий, которые на нее может оказать наблюдатель. Наблюдение — не воздействие. Наблюдение — это важнейший мысленный эксперимент, который можно произвести над системой и цель которого состоит в первую очередь в локализации системы в ее фазовом пространстве. Можно сказать и наоборот: фазовое пространство есть множество возможных результатов мгновенных полных наблюдений. Полное наблюдение позволяет вычислить полную эволюцию классической системы; существование полных наблюдений — это другая форма постулата детерминизма. Эволюция — это набор результатов наблюдений во все моменты времени. Идея мысленного наблюдения без воздействия подкрепляется рассмотрением разных способов наблюдения, более приближенных к реальности, где воздействие входит в схему, но может быть сделано сколь угодно малым или полностью учтено в расчетах, т. е. контролируемо. Эти рассмотрения, по существу, состоят в том, что изолированная система S включается как часть в большую изолированную систему (S, T). Наблюдению отвечает акт слабого взаимодействия между S и T, почти не нарушающий эволюции S (может быть, включенный ненадолго и тут же выключенный). Принципиально важно здесь вот что: к объединению (S, T) все равно применяется абстракция мысленного наблюдения, уже не влияющего на эволюцию объединенной системы. Кроме того, предполагается, что S может стать частью (S, T), не потеряв своей индивидуальности, ненадолго, обратимо.

Это очень естественный постулат для человека, главное средство наблюдения которого — видение. Электромагнитные взаимодействия столь слабы, что в масштабах от космических до человеческих взгляд на систему ничуть не действует на нее.

Умственные очи должны видеть в фазовом пространстве механики, в пространстве элементарных событий теории вероятностей, в кривом четырехмерном пространстве — времени общей теории относительности, в комплексном бесконечномерном проективном пространстве квантовой теории. Чтобы понимать видимое глазами во лбу, мы должны знать, что оно есть лишь проекция на сетчатку бесконечномерного мира. Образ платоновской пещеры кажется мне лучшей метафорой структуры современного научного знания: мы в самом деле видим лишь тени, ибо тень — лучшая метафора проекции.

Человеку психологически очень трудно выйти за пределы привычных пространственных трех измерений. Но мы вредим себе, пытаясь описать квантовые внутренние степени свободы неловкими словами вроде «значение проекции спина на ось z» — вектор спина находится в совсем другом пространстве, чем ось z. Стоит вспомнить, что и трехмерность мира вошла в создание после огромных усилий — ее научили нас видеть художники Возрождения. Уччелло на десять лет удалился от дел, чтобы посвятить себя изучению перспективы. Современная математика среди прочего — это суровый тренаж многомерной перспективы по унифицированной программе. Если верить нейропсихологам, левая и правая части мозга при этом ведут себя, как слепой и его безногий поводырь, которого первый несет на своих плечах.

В классике наблюдатель, в общем, представлен системой координат в основных пространствах теории. Единица измерения определяет координату в спектре измеряемой величины. Когда эти единицы выбраны, координатные функции, т. е. наблюдаемые величины, отождествляют пространство положений, фазовое пространство или их части с подмножествами R^n и C^n математиков. Теория с наблюдаемыми величинами хороша, поскольку она одновременно описывает и идеи, и их наблюдаемые «тени». Теория с наблюдаемыми величинами плоха, поскольку может оказаться проще, поучительнее, вернее как можно раньше явно отделить наблюдаемое от наблюдателя и изучать их соотношение как отдельный объект исследования. Цвет по Ньютону и Эйлеру — это спектральный состав светового излучения в диапазоне длин волн около полумикрона; цвет по Гёте — это то, что мы видим. Поразительно, насколько эти два представления не поддаются прямому сравнению — их связывает лишь сложная и нетривиальная физиологическая теория цветового зрения. Гуманитарий Гёте не мог допустить отречения от наблюдателя, ибо вся его система ценностей не способна существовать без идеи человеческого участия как мерила вещей. Многое можно сказать в пользу этой точки зрения. Многое можно и возразить; часто лучший способ узнать себя — отвернуться от себя. Ньютоновская теория цвета — и все последующие физические теории — призваны объяснить, что такое свет безотносительно к тому, что его можно видеть. Для Гёте свет — это главным образом то, что можно видеть. И опять, как всегда, оказывается, что видимое нужно объяснять через невидимое.

«Все движения, замечающиеся у небесной тверди, принадлежат не ей самой, а Земле» (Коперник, 1515). После этой фразы, сдвинувшей Землю, все теории, основанные лишь на «замечающемся», стали архаизмом еще до своего рождения.

Классический наблюдатель живет в мире человеческих масштабов, и концепция классического наблюдения претерпевает естественные изменения при переходе к масштабам космологии или микромира. Расстояния, времена, энергии и действия астрономических явлений столь велики, что гипотезу о невлиянии наблюдателя хочется принять без дальнейших обсуждений. Другие проблемы наблюдения выступают на передний план; две из них можно кратко суммировать в виде вопросов. Можно ли рассматривать Вселенную как замкнутую систему? Как относиться к теории, описывающей явления, которые не могут наблюдаться из-за их разрушительного влияния на наблюдателя или из-за того, что какие-то области пространства — времени от него принципиально изолированы? (Звездные температуры, массы, давления, гравитационные поля черных дыр, условия Большого Взрыва.)

Самые принципы описания замкнутых систем основаны на гипотезе их воспроизводимости — фазовое пространство системы реализует идею осуществимости разных состояний и разных путей эволюции. Как совместить эту идею с единственностью эволюции, данной нам в наблюдениях системы? Ответ, конечно, связан с представлением о локальном взаимодействии «частей мира» между собой и о существенной одинаковости законов физики, действующих в разных частях. В самые простые и самые фундаментальные модели Вселенной (модель Фридмана, модель Эйнштейна — де Ситтера) заложена идея однородности, проявляющейся в существовании большой группы симметрий математической модели. Во всех моделях космологии на первый план выступает аспект представления о замкнутой системе, который затемнен в описании более привычных примеров — степень огрубления деталей. В космологической модели мира не остается и следов обыденности. Но вопреки этому или благодаря этому статья в «Успехах физических наук» может начинаться фразой: «Мы были бы счастливы, если бы Лебедь X-1 оказался черной дырой» («Успехи физических наук», 1978, т. 126, вып. 3, с. 515). Мы знаем кое-что о мире потому, что мы счастливы познавать его.

Принципы квантового описания. Итак, идеальный наблюдатель макромира не может его изменить, но даже идеальный наблюдатель микромира не может его изменить. Это объясняется в бесчисленных изложениях квантовой механики, но, кажется, мы понимаем это очень плохо. Квантовая механика не просто научила нас новым математическим моделям явлений, она явила образец нового соотношения между описанием и явлением. В частности, целый ряд характеристик этих моделей на естественном языке приходится объяснять, привлекая идею «ненаблюдаемости». Смысл этого слова меняет оттенки как Протей: ненаблюдаемы фаза пси-функции, виртуальный фотон, цвет кварка, разница между тождественными частицами и многое другое.

Попробуем взглянуть на геометрию квантовой механики умственными очами.

Фазовое пространство. Фазовое пространство замкнутой квантовой системы есть множество лучей (одномерных подпространств) в комплексном линейном пространстве $\mathcal{H}$, в котором задано также скалярное произведение. В этом постулате выражены: а) принцип линейной суперпозиции; б) принцип «ненаблюдаемости фазы». Вместо целой прямой в $\mathcal{H}$, описывающей состояние системы, обычно рассматривают один вектор, лежащий в этой прямой. Он определен только с точностью до умножения на комплексное число. Если даже нормировать его условием, чтобы его длина была равна единице, все еще останется произвол в выборе множителя $e^{i\theta}$. Это θ и есть «ненаблюдаемая фаза».

Фазовые кривые. Чтобы описать их, мы должны объяснить, как каждый луч в $\mathcal{H}$ меняется со временем t, изображая эволюцию замкнутой системы. Стандартное описание таково: а) в $\mathcal{H}$ имеется N попарно ортогональных лучей, которые вообще не меняются: они соответствуют стационарным состояниям системы. (Здесь N — размерность $\mathcal{H}$; как и в главе 1, мы для простоты рассматриваем лишь конечномерный случай.); б) каждому из стационарных состояний ψ_j отвечает величина E_j, имеющая размерность энергии, энергетический уровень соответствующего стационарного состояния. Если в нулевой момент времени система находилась в состоянии $\Sigma\, a_j\psi_j$, то через время t она будет находиться в состоянии $\psi\,(t) = \Sigma\, a_j\psi_j e^{E_j t/i\hbar}$. Заметим, что $E_j t$ имеет размерность действия и, естественно, измеряется единицей Планка $\hbar$. Поскольку $e^{Et/i\hbar} = \cos\frac{Et}{\hbar} - i\sin\frac{Et}{\hbar}$, каждое слагаемое здесь периодично по времени, в сущности, описывает движение по окружности со своей угловой скоростью. Их сумма, таким образом, изображает вращение вокруг N осей с разными скоростями. Траектории таких двухмерных движений — это известные фигуры Лиссажу. Другой образ из давней истории науки — эпициклы Птолемея, также приводившие к сумме круговых движений. Любая координата вектора $\psi(t)$ испытывает со временем частые и исключительно нерегулярные колебания; график уже такой простой функции, как $\sum_{n=1}^{10} \cos(n^2 t)$ выглядит как сейсмограмма. Удобно записывать $\psi(t)$ в виде $e^{-iS(t)}\psi(0)$, где $S(t)$ — линейный оператор «действие за время t».

Наблюдение: печки, фильтры и квантовые скачки. Классическая идеализация наблюдателя, способного фиксировать мгновенное положение системы на ее фазовой кривой, заменяется радикально новой системой понятий. Назовем их сначала не общеупотребительными словами, чтобы не создавать иллюзий. Сильно идеализированные предположения о связи описанной схемы с реальностью состоят в том, что для каждого состояния $\psi \in \mathcal{H}$ можно сделать физический прибор («печку») A_ψ, производящий систему в состоянии ψ. Сверх того, для каждого состояния $\chi \in \mathcal{H}$ можно сделать прибор («фильтр») B_χ, на вход которого подаются системы в состоянии ψ, а на выходе обнаруживаются они же в состоянии χ или не обнаруживается ничего («система через фильтр не проходит»). Третий основной (после принципа суперпозиции и закона эволюции) постулат квантовой механики состоит в следующем: *система, приготовленная в состоянии ψ и сразу же после этого пропущенная через фильтр B_χ, пройдет через него и окажется в состоянии χ с вероятностью, равной квадрату косинуса угла между лучами ψ и χ в $\mathcal{H}$.*

Если между приготовлением системы в состоянии ψ и ее пропусканием через фильтр B_χ прошло время t, то вероятность будет равна квадрату косинуса угла между $e^{-iS(t)}\psi$ и χ. Пока с системой ничего не делают, она движется по своей фазовой кривой. Но как только ее подают на фильтр, пропускающий лишь системы в состоянии χ, вектор ее состояния скачком меняется — он либо доворачивается на угол между ψ и χ, и система проходит через фильтр, либо фильтр ее задерживает. Система, прошедшая через фильтр B_χ, не несет никаких следов памяти о состоянии, с которым она вошла в фильтр, — χ может получиться из чего угодно.

Если ψ, χ имеют единичную длину, то «вероятность перехода» от ψ к χ обозначается $|\langle\chi|\psi\rangle|^2$, а само скалярное произведение $\langle\chi|\psi\rangle$ называется амплитудой перехода. Поскольку фазы χ и ψ не определены, не определен и аргумент комплексного числа $\langle\chi|\psi\rangle$ — однозначный смысл имеют лишь разности аргументов, скажем, $\langle\chi_1|\psi\rangle$ и $\langle\chi_2|\psi\rangle$. Квадрат модуля суммы двух комплексных чисел зависит не только от самих чисел, но и от угла между ними, т. е. разности их аргументов. Это — «интерференция амплитуд».

Взаимодействие системы ψ с фильтром $B_\varkappa$ — частный случай того, что в квантовой механике называют наблюдением, или измерением. Более общая схема получается, если представить себе, что систему ψ подают на набор фильтров $B_{\varkappa_1}, \ldots, B_{\varkappa_n}$, где $\chi_1, \ldots, \chi_n$ — некоторая полная совокупность ортогональных базисных векторов; эти фильтры следует представлять себе расположенными «параллельно», так что через какой-нибудь из них система пройдет и окажется в состоянии χ_j. С таким набором фильтров связывают представление о некоторой физической величине B, которая в состояниях $\chi_1, \ldots, \chi_n$ принимает значения $b_1, \ldots, b_n$ соответственно, и говорят, что акт измерения, или наблюдения, приводит к значению b_j величины B на состоянии ψ, если ψ прошла через фильтр $B_{\varkappa_j}$. Математическим представителем системы фильтров $(B_{\varkappa_j})$ или величины B принято считать линейный оператор $\mathcal{H} \to \mathcal{H}$, который переводит вектор x_j в вектор $b_j x_j$ для всех j. Все такие линейные операторы, осуществляющие растяжение $\mathcal{H}$ по N взаимно ортогональным направлением с вещественными коэффициентами, называют наблюдаемыми.

Приведем в качестве иллюстрации идеализированное описание эксперимента Штерна — Герлаха по квантовому измерению момента количества движения (спина) ионов серебра. Гильбертово пространство $\mathcal{H}$, соответствующее спиновым степеням свободы этой системы, двухмерно. Серебро испаряется в электрической печке; ионы коллимируются небольшим отверстием в экране, и получившийся пучок пропускается между полюсами магнита, создающего неоднородное магнитное поле. В пучке ионы находятся во всевозможных спиновых состояниях, но проходя через магнитное поле имеют тенденцию «сваливаться» в одно из двух стационарных состояний χ_+, χ_- в этом поле, которые по традиции называются состояниями со спином «вверх» и «вниз», ели магнитное поле вертикально. На выходе из области поля эти состояния из-за неоднородности поля оказываются пространственно разделенными — пучок делится пополам. Таким образом, магнитное поле действует как совокупность фильтров.

Итак, квантовое «наблюдение», по сути дела, не имеет ничего общего с классическим наблюдением: а) акт «наблюдения» почти неизбежно выбивает систему с ее фазовой траектории; б) акт «наблюдения» позволяет зарегистрировать в лучшем случае новое положение системы на фазовой кривой, но не то, на котором она находилась к моменту наблюдения, память о чем теряется; в) новое положение системы лишь статистически определяется старым; наконец, г) среди квантовых «наблюдаемых» имеются (и в действительности играют основную роль) физические величины, которым не отвечают никакие классические наблюдаемые.

Сопоставление между значениями квантовых и классических наблюдаемых может быть лишь очень непрямым. Например, квантовой наблюдаемой B можно поставить в соответствие ее среднее значение $\hat{B}_\psi$ на состоянии ψ (в смысле статистического усреднения). Оно оказывается равным $<\psi|B|\psi>$ (если $|\psi| = 1$). (Читатель может принять эту запись просто за новое обозначение.) Среднее значение $\Delta\hat{B}_\psi = \sqrt{[(B-\hat{B}_\psi{}^2)]_{\hat{\psi}}}$ тогда измеряет разброс значений B относительно среднего значения $\hat{B}_\psi$ на состоянии ψ. Пусть B, C — две наблюдаемые величины, $[B,C] = \frac{1}{i}(BC - CB)$. Можно показать, что «теорема Пифагора» в $\mathcal{H}$ приводит к неравенству

$$\Delta\hat{B}_\psi \cdot \Delta\hat{C}_\psi \geqslant \frac{1}{2}\left|[B,C]_{\hat{\psi}}\right|,$$

которое является математическим выражением соотношения неопределенностей Гейзенберга. Оно чаще всего применяется к парам наблюдаемых B, C, для которых $[B, C]$ сводится к умножению на $\hbar$. Тогда неравенство принимает более привычный вид: $\Delta\hat{B} \cdot \Delta\hat{C} \geqslant \frac{\hbar}{2}$ все равно на каком ψ.

Заметим, что если $\mathcal{H}$ конечномерно, таких пар наблюдаемых нет; соотношение неопределенности обычно применяется к квантовым аналогам пар классических наблюдаемых (координата и проекция импульса на соответствующую ось, энергия и время).

Объединение квантовых систем. Уже говоря о классическом наблюдении, мы отметили, что попытка детального описания подразумевает включение наблюдаемой системы S в бо́льшую систему (S, T). Поэтому следует подозревать, что необычные свойства квантовых наблюдений удается лучше понять, разобравшись в принципах квантового описания объединенной системы.

Относящийся к этому постулат квантовой механики состоит в том, что пространство состояния $\mathcal{H}_{S,T}$ объединенной системы есть некоторое подпространство тензорного произведения $\mathcal{H}_S \otimes \mathcal{H}_T$ (если S и T бесконечномерны, то это произведение нужно пополнить; эти тонкости мы опускаем). Какое именно подпространство $\mathcal{H}_S \otimes \mathcal{H}_T$ нужно взять, решается на основе дальнейших постулатов. Пока рассмотрим случай $\mathcal{H}_{(S,T)} = \mathcal{H}_S \otimes \mathcal{H}_T$. Уже сама формулировка математической модели показывает возможность совершенно неклассических связей между «частями» S и T объединенной системы. В самом деле, оказывается, что для подавляющего большинства состояний (S, T) нельзя сказать, в каком состоянии находятся S и T «по отдельности», так что представление о частях оказывается имеющим очень ограниченный смысл. Действительно, в $\mathcal{H}_S \otimes \mathcal{H}_T$ имеются разложимые состояния $\psi_S \otimes \psi_T$, где $\psi_S \in \mathcal{H}_S$, $\psi_T \in \mathcal{H}_T$. Когда система (S, T) находится в одном из таких разложимых состояний, мы имеем основания говорить что она состоит из S в состоянии ψ_S и T в состоянии ψ_T. Но уже для состояния $\psi_S \otimes \psi_T + \psi_S \otimes \otimes \psi'_T$ такое утверждение несостоятельно. Между тем принцип суперпозиции позволяет строить такие состояния $\sum_i \psi_S^{(i)} \otimes \psi_T^{(i)}$ в большом количестве. Множество разложимых состояний имеет размерность $m + n$, а всех — размерность mn, где mn — размерности $\mathcal{H}_S$ и $\mathcal{H}_T$ соответственно, почти все состояния (S, T) неразложимы. В подавляющем большинстве состояний (S, T) подсистемы S и T существуют лишь «виртуально».

Эта неклассическая связь между частями объединенной системы часто не может объясняться на основе классических представлений о том, что связь частей системы осуществляется через обмен энергией между ними. Действительно, имеется фундаментальный случай объединения двух тождественных систем S и T, когда в фазовом пространстве объединенной системы вообще нет ни одного разложимого состояния. Пусть S — фермионная элементарная частица, скажем, электрон, T — другая такая же частица. Тогда $\mathcal{H}_{S,T}$ есть собственное подпространство $\mathcal{H}_S \otimes \mathcal{H}_T = \mathcal{H}_S \otimes \mathcal{H}_S$, состоящее из векторов, меняющих знак при перестановке S и T,— это линейные комбинации векторов $\psi_1 \otimes \psi_2 - \psi_2 \otimes \psi_1$. Легко убедиться, что каждое состояние системы двух электронов неразложимо. Векторов $\psi \otimes \psi$, в частности, в фазовом пространстве нет; в популярном изложении говорят, что два электрона не могут находиться в одинаковом состоянии; это основа существования стабильных атомов и, в конечном счете, материального мира, окружающего человека. Однако почти невозможно объяснить словами «квазиразложимые» состояния $\psi_1 \otimes \psi_2 - \psi_2 \otimes \psi_1$. Говорят, что один из электронов находится в состоянии ψ_1, а другой — в состоянии ψ_2, но нельзя сказать, «который» из них в каком состоянии. Наконец, естественный язык оказывается уже в безвыходном положении, когда нужно объяснить разницу между объединением двух тождественных фермионов и двух тождественных бозонов, где фазовое пространство $\mathcal{H}_{(S,S)}$ состоит из симметричных относительно перестановок векторов в $\mathcal{H}_S \otimes \mathcal{H}_S$, т. е. объяснить, что «различными, но неразличимыми» две системы могут быть двумя разными способами (уже и один способ причинял массу хлопот натурфилософии).

Здесь уместно сделать отступление об «естественном языке». В действительности наши представления о «классической» и «неклассической» физике очень тесно связаны с представлением о том, что можно и что нельзя адекватно выразить простыми словами. Положение дел здесь очень нетривиально. Не только популяризатор, но и работающий физик часто стремится объяснить новое явление, закон или принцип «на пальцах». Нужно лишь отдавать себе отчет в том, каково место такого объяснения. Оно призвано: а) назвать и быть способным вызвать из памяти соответствующий фрагмент точной теории с математическими формулами, структурами и т. п., подобно тому как действует код команды в алгольной программе, включая процесс выполнения этой команды, который и составляет ее смысл; б) включить процесс порождения ассоциаций, т. е. помочь обнаружить, что нечто похоже на нечто другое; в) создать в мозгу структуру интуитивных представлений о предмете, значение которой состоит не в замене точного знания о нем, а в формировании ценностных принципов и возможности быстрых оценок —

что искать дальше, в каком направлении думать, что правдоподобно и что неправдоподобно. (В частности, в этом польза популяризации для ученых другой специальности.) Мы должны подчеркнуть еще раз следующую точку зрения: семантикой словесного описания какого-то фрагмента физики является, в общем, не соответствующий комплекс явлений природы, а соответствующий фрагмент теории, семантика которой, в свою очередь, эксплицируется через другие фрагменты теории, операциональные предписания и т. п. Тем не менее побуждение интерпретировать непосредственно языковые выражения может оказаться исключительно плодотворным. Так были открыты кварки: когда выяснилось, что пространство некоторых внутренних степеней свободы нуклона разлагается в тензорное произведение трех подпространств, возник соблазн рассматривать эти три подпространства как внутренние степени свободы трех новых частиц, из которых состоит нуклон. Эти частицы и суть кварки u, d, s (они «открыты, но не обнаружены в свободном состоянии»).

Возвращаясь к проблеме квантовых наблюдений, мы приходим к выводу, что неклассичность их математической модели связана в первую очередь с тем, что она является огрублением гораздо более сложной модели, призванной описывать взаимодействие системы с другой системой — «прибором». Во время первых дискуссий о смысле математического аппарата квантовой механики особенно подчеркивалось то обстоятельство, что прибор макроскопичен и нет никакой надежды на полную квантовую теорию процесса его взаимодействия с системой. Это и вынуждает заменять его линейным оператором наблюдаемой. Итак, логика математического описания приводит к следующим выводам, которые в совокупности почти противоречивы. В той мере, в какой абстракция замкнутой квантовой системы правомерна, для ее описания мы нуждаемся лишь в одной «наблюдаемой» — операторе энергии. Однако с ней не следует связывать представлений об измерении энергии, ибо акт «измерения» требует расширения системы. Локализация системы в ее фазовом пространстве может быть произведена и с помощью «измерения» других наблюдаемых, но они суть огрубленные модели недоступного для полного описания объединения (система + прибор + оператор энергии системы/прибора). После взаимодействия с прибором система может потерять свою индивидуальность, и представление о том, что она начинает новую жизнь в точке новой фазовой кривой в своем пространстве, может потерять всякий смысл. Наконец, поскольку и до взаимодействия с прибором система была частью чего-то, скорее всего она ни в какой момент не имеет индивидуальности, нужной для адекватности модели. Кажется, нет меньшей замкнутой системы, чем весь Мир.

После всего этого следует считать чудом, что наши модели успешно описывают хоть что-нибудь. На самом деле они успешно описывают очень многое: мы наблюдаем то, что предсказали, и понимаем то, что наблюдаем. Однако этот последний акт наблюдения и понимания всегда ускользает от физического описания.

В конце шестидесятых годов были сконструированы оптические затворы для фотокамер, позволяющие получить выдержку в десять пикосекунд. За это фантастически короткое время световой луч в воде проходит расстояние всего 2,2 мм и можно получить фотографию короткого лазерного импульса в флаконе, пока он движется внутри него. В лаборатории «Белл телефон», где были сделаны такие фотографии, в воду добавляли каплю молока, чтобы усилить рассеяние света и сделать след импульса ярче. Эта капля молока — символ человеческого участия в мире, где нельзя быть только наблюдателем.

4. ПРОСТРАНСТВО — ВРЕМЯ КАК ФИЗИЧЕСКАЯ СИСТЕМА

«Что такое время, пространство, место и движение, я не объясняю, ибо это известно всем» (И. Ньютон).

«Время и пространство — это категории нашего мышления, а не условия нашего существования» (А. Эйнштейн).

Мы ощущаем себя локализованными в пространстве и длящимися во времени, и почти все схемы современной физики в конечном счете описывают события, происходящие на арене пространства — времени. Однако со времени создания общей теории относительности и квантовой механики все усиливается тенденция рассматривать пространство — время как особую физическую систему. Принципы ее описания все еще остаются классическими. Трудности квантовой теории поля, видимо, указывают на то, что эти принципы вступают в противоречие с универсальными квантовыми законами.

Главный математический образ пространства — времени — дифференцируемое четырехмерное пространственно-временнóе многообразие, для краткости Мир. Одна точка Мира — идеализация очень краткого и «маленького» события, вроде вспышки, излучения или поглощения фотона атомом. Сверх того, точка Мира — это событие потенциальное; точка Мира «готова принять» событие, но «существует» и помимо него. Образ сосредоточенного в малой области пространства, но длящегося события, такого, как жизнь наблюдателя, звезды, галактики,— это линия в пространстве — времени, мировая линия события или его история. Исключительно важно научиться представлять себе Мир Становящийся как Мир Ставший, т. е. всю историю Вселенной или ее большой части как завершенный четырехмерный образ, нечто вроде «дао» древнекитайской философии. Введение временнóй динамики — это следующий шаг, который осуществляется так. В Мире между двумя близкими точками $x^\alpha = (x^0, x^1, x^2, x^3)$ и $x^\alpha + dx^\alpha = (x^0 + dx^0, \ldots, x^3 + dx^3)$ определено пространственно-временнóе расстояние. Его квадрат $ds^2 = \Sigma g_{\alpha\beta} dx^\alpha dx^\beta$ — квадратичная форма от разностей координат близких точек. Здесь x^α — произвольная локальная система координат. Классический наблюдатель со своей малой лабораторией, состоящей из линеек с делениями и часов, может установить локальную систему координат, в которой $x_0 = ct$; x_1, x_2, x_3 — прямоугольные координаты в физическом пространстве наблюдателя, t — показания часов. Метрика вблизи него будет близка к метрике Минковского $dx_0^2 - (dx_1^2 + dx_2^2 + dx_3^2)$. Если представить себе, что Мир заселен такими наблюдателями, области действия координат которых его покрывают, то координаты событий должны пересчитываться от одного наблюдателя к другому. Но пространственно-временнóй интервал между двумя близкими точками, вычисленный разными наблюдателями, будет одним и тем же. Скорость света c употребляется для пересчета временны́х единиц в пространственные, в них время и измеряется. Мировая линия наблюдателя есть его собственная река времени: атомные часы наблюдателя отсчитывают значения интеграла $\int \sqrt{ds^2}$ вдоль этой мировой линии, т. е. ее длину. Никакого физически осмысленного «общего времени» Вселенной нет; правда, его иногда можно ввести в специальных моделях Мира. Нет ничего удивительного в том, что две кривые в Мире с общим началом и концом могут иметь разную длину — это так уже на евклидовой плоскости. Поэтому не удивительно, что два наблюдателя, сверившие свои часы и расставшиеся, при новой встрече обнаружат, что их часы разошлись. Менее привычно, что если две точки пространства — времени вообще можно соединить мировой линией наблюдателя, то среди таких линий есть самая длинная, но нет самой короткой (на евклидовой плоскости верно как раз обратное). Это специальное свойство метрики Минковского, связанное с тем, что она не является положительно определенной: квадрат интервала между разными точками может быть положительным, отрицательным и нулем.

Наблюдатели с самыми длинными мировыми линиями, т. е. самым быстрым течением собственного времени, называются инерциальными. С точки зрения общей теории относительности, они свободно падают в поле тяготения. Их мировые линии называются времениподобными геодезическими.

Второй важный класс линий в Мире — траектории частиц, летящих со скоростью света, вроде нейтрино. Вдоль них пространственно-временно́й интервал тождественно обращается в нуль — «время останавливается», что и составляет их геометрическое определение. Геометрия таких светоподобных геодезических определяет, что может наблюдать наблюдатель и, более общо, какие события в Мире могут влиять на другие события. Именно в ее терминах точно формулируется, в частности, постулат о том, что никакие сигналы не могут распространяться быстрее света.

Мир Минковского. Простейшим и важнейшим конкретным примером Мира является плоский Мир Минковского. Он хорошо имитирует любой другой Мир локально. В этом Мире имеется инерциальный наблюдатель с системой координат (x^{α}), покрывающей весь Мир, в которой метрика тождественно равна $(dx^0)^2 - ((dx^1)^2 + (dx^2)^2 + dx^3)^2)$. Фиксируем начало отсчета — точку на мировой линии этого наблюдателя. Мир Минковского $\mathcal{M}$ превратится в линейное пространство в выбранной системе координат, и эта структура линейного пространства от инерциального наблюдателя на самом деле не зависит, если не обращать внимания на сдвиг начала отсчета. Поэтому понятия прямой, плоскости, трехмерного подпространства в $\mathcal{M}$ имеют абсолютный смысл. Преобразования $\mathcal{M}$ (линейные), сохраняющие метрику Минковского, образуют группу Пуанкаре, а часть из них, оставляющая на месте некоторое начало координат, образуют группу Лоренца. Это — основные группы симметрий всей физики, точнее, физических законов: ни точки Мира Минковского, ни система координат инерциальных наблюдателей ничем не предпочтительны одни перед другими, и все координатные формулировки одного закона должны быть эквивалентными.

Множество точек, отстоящих на нулевое расстояние от начала отсчета P, образует световой конус C_P с уравнением $(x^0)^2 - (x^1)^2 - (x^3)^2 = 0$. Проходящие через P времениподобные геодезические — это прямые, лежащие внутри C_P, а светоподобные геодезические — прямые, лежащие на самом C_P,— образующие этого конуса. Конус состоит из двух пол — приходящей и уходящей. Время по времениподобной геодезической течет по направлению из приходящей полы в уходящую. Это различие между полами C_P непрерывно зависит от P, в чем и выражается существование единого направления времени во всем Мире Минковского при отсутствии единого времени.

Точка P и времениподобный касательный единичный вектор в ней — это модель «мгновенного наблюдателя» в Мире Минковского. Вектор указывает направление его личного времени. Ортогональное к этому вектору подпространство в $\mathscr{M}$ — это модель трехмерного физического пространства мгновенного наблюдателя. Его метрика (с обратным знаком) получается ограничением метрики Минковского. У двух разных мгновенных наблюдений, даже находящихся в одной точке Мира $\mathscr{M}$, разные и физические пространства. Они пересекают мировую полосу, скажем линейки, под разными углами. Такое пересечение есть в некотором приближении мгновенный наблюдаемый образ линейки. Он может иметь поэтому для разных наблюдателей разную длину—пространственные и временны́е координаты могут перетекать друг в друга.

Как следует представлять себе наблюдение удаленного объекта, скажем звезды, в Мире Минковского? Пусть наблюдатель движется по мировой линии P, а звезда — по своей мировой линии S. Вообразим себе приходящую полу светового конуса $C^+_{P_0}$ точки P_0, движущегося вместе с $P_0 \in P$. Она «заметает» за собой некоторую часть $\mathscr{M}$ — область Мира, которую наблюдатель мог наблюдать. В конкретной точке P_0 наблюдатель видит звезду в точке пересечения S и $C^+_{P_0}$ посредством светового луча, соединяющего $S \cap C^+_{P_0}$ и P_0. Но мы должны еще понять, как узнать видимое положение звезды на небосводе наблюдателя. Дело в том, что его небосвод «лежит в его физическом пространстве E_{P_0}», а не в пространстве Минковского $\mathscr{M}$, и положение звезды моделируется лучом в E_{P_0}. Чтобы получить этот луч, мы должны спроецировать луч в $\mathscr{M}$ — полупрямую с концом P_0, проходящую через $S \cap C^+_{P_0}$, в физическое пространство E_{P_0}. Эта проекция — ортогональная, но, конечно, по отношению к метрике Минковского.

Таким образом, удобно различать «абсолютный небосвод» в точке P_0 — базу приходящей полы светового конуса, и небосвод мгновенного наблюдателя в этой точке — проекцию абсолютного небосвода в физическое пространство этого наблюдателя. В классической космографии небосвод вполне можно представлять себе как хрустальную сферу неопределенного радиуса; между точками небосвода определены угловые расстояния, и геометрия небосвода совпадает с геометрией твердой сферы. Но для другого наблюдателя угловые расстояния между звездами будут иными; летя с очень большой скоростью в направлении созвездия Ориона, мы увидим, что оно сожмется в овчинку, а противоположная небесная полусфера растянется (астрономы называют это аберрацией). Таким образом, математическая структура «абсолютного небосвода» не совпадает со структурой евклидовой сферы: угловые расстояния на ней не имеют смысла, не зависящего от наблюдателя. Подробное исследование показывает, что естественная структура абсолютного небосвода — это комплексная сфера Римана: плоскость комплексных чисел, дополненная бесконечно удаленной точкой, причем различие между конечными и бесконечной точками забыто. Более точно сфера Римана — это множество одномерных векторных подпространств в двухмерном комплексном векторном пространстве или комплексная проективная прямая CP^1. В частности, естественные координаты звезд на небе — это комплексные числа. Выберем три опорные звезды и припишем им координаты $0, 1, \infty$. Тогда имеется несложная процедура, позволяющая по результатам наблюдений поставить в соответствие любой четвертой звезде комплексное число z, и оно получится одним и тем же, какой бы наблюдатель в данной точке Мира не измерил положение звезды.

Если координаты $0, 1, \infty$ приписываются не опорным звездам, а точкам неба, в которые направлены три ортогональные оси пространственной системы координат наблюдателя, тогда, конечно, комплексные координаты z и z' одной и той же звезды для разных наблюдателей в данной точке Мира могут быть разными. Но они обязательно связаны дробнолинейным соотношением вида $z' = \frac{az+b}{cz+d}$, где a, b, c, d — комплексные числа, не зависящие от звезды, связанные условием $ad - bc = 1$. Матрица $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ с определителем единица является необычным представителем преобразования Лоренца, связывающего две инерциальные системы координат в одной точке Мира. В современной физике это представление группы Лоренца, однако гораздо более фундаментально, чем обычные матрицы пересчета систем координат.

Искривленный Мир. Более общие модели Мира отличаются от мира Минковского в нескольких отношениях. Во-первых, даже локально в системе координат инерциального наблюдателя метрика не может быть приведена к форме $(dx^0)^2 - (dx^1)^2 - (dx^2)^2 - (dx^3)^2$. Во-вторых, может вообще не существовать глобальной системы координат. В-третьих, метрика Мира и история материи и полей в в Мире не являются независимыми — кривизна метрики определяется материей, и, в свою очередь, метрика накладывает сильные ограничения на возможные истории материи; эти связи — суть уравнения Эйнштейна. Геометрически типичные образы искривления создаются при исследовании возможного поведения близких времениподобных геодезических (локальный аспект) и световых конусов (глобальный аспект). Попытаемся дать словесное описание эффектов очень сильного искривления, приводящего к понятию черной дыры. Представим себе мировую линию S точечной массы m. С ней связана характерная длина пространственно-временнóго интервала $2\,Gm/c^2$ — так называемый радиус Шварцшильда. Точки Мира, лежащие на времениподобном расстоянии от S, не большем радиуса Шварцшильда, образуют трубу Шварцшильда $\bar{S}$ вокруг S. Далеко вне ее Мир почти плоский (если отвлечься от влияния остальной материи). Но внутри нее Мир настолько искривлен, что для любой точки $P_0 \in \bar{S}$ уходящая пола светового конуса C_{P_0} целиком лежит внутри трубы Шварцшильда: каждый световой луч на границе $\bar{S}$ выглядит как спираль на поверхности цилиндра. Поле тяготения массы m не выпускает фотоны, испущенные внутри $\bar{S}$. Приходящая пола $C^{+}_{P_0}$, однако, не обязана лежать в $\bar{S}$ — труба Шварцшильда может поглощать внешнее излучение.

Рассмотрим теперь более реалистическую модель, когда масса m сама сосредоточена в конечной области, радиус которой может быть больше радиуса Шварцшильда. Например, для Земли он имеет значение около 1 см, а для Солнца — около 3 км. В этом случае труба Шварцшильда не имеет особого физического смысла. Но в процессе эволюции звезда достаточно большой массы может под влиянием собственного тяготения сколлапсировать, так что в какой-то точке K_0 мировой линии ее центра радиус звезды сравняется с радиусом Шварцшильда и затем станет убывать. Отрезок трубы Шварцшильда после точки K_0 будет пространственно-временной областью, не доступной внешнему наблюдению. Четырехмерная картина того, что увидит внешний наблюдатель, будет примерно такой. Приходящая пола светового конуса наблюдателя в любой точке наблюдения пересекается с мировой трубой звезды, но это пересечение всегда происходит раньше точки K_0. Иначе говоря, локальное время вдоль того конца луча, который воспринимает наблюдатель, будет стремиться к бесконечности, тогда как локальное время вдоль звезды, испускающей луч, будет стремиться к конечной величине, отвечающей точке K_0 (напомним, что локальное время — это длина соответствующей мировой линии). «По часам наблюдателя коллапс длится бесконечно долго».

Еще раз повторим, как соотносится с четырехмерной картиной ее видимый образ на небосводе наблюдателя. В искривленном Мире к и без того сложному процессу перевода добавляется лишний шаг — построение плоского Мира, касательного к искривленному в точке, где находится мгновенный наблюдатель. Чтобы получить на своем небосводе точку видимого образа звезды S, наблюдатель должен: а) построить в кривом четырехмерии приходящую световую геодезическую, соединяющую его положение с мировой линией звезды; б) провести к этой геодезической касательную полупрямую в плоском Мире, касающемся кривого Мира в точке, где находится наблюдатель; в) в плоском же Мире провести касательную к мировой линии наблюдателя; г) построить в этом плоском Мире мгновенное физическое пространство наблюдателя; д) спроецировать на него касательную к лучу от звезды.

Так движутся «тени идей» на стене Пещеры.

Спиноры, твисторы и комплексный Мир. Если согласиться с идеей, что точка пространства—времени есть идеализация классического образа «мельчайшего события», то мы неизбежно придем к необходимости рассматривать и другие геометрические модели по мере возрастания знаний о том, какими характеристиками обладают такие события. Скажем, акт поглощения фотона далеко не полностью характеризуется указанием, в какой точке Мира он произошел, — нужно указать энергию и поляризацию фотона. Положение электрона на своей мировой линии также еще не определяет полностью его состояние — нужно указать направление его спина.

Хотя и поляризация, и спин являются квантовомеханическими внутренними степенями свободы, замечательно, что их геометрическое описание автоматически и очень естественно встроено в геометрию Мира. Именно, значение поляризации или спина в точке Мира — это луч в двухмерном комплексном пространстве, или точка на сфере Римана CP^1. Оказывается, что эту сферу Римана можно совершенно однозначно отождествить с абсолютным небосводом этой точки, который, как мы объясняли выше, также является сферой Римана. Поэтому для каждого Мира $\mathcal{M}$ можно рассмотреть расширенный Мир $\overline{\mathcal{M}}$, одна точка которого является парой (точка $\mathcal{M}$, точка абсолютного небосвода над этой точкой $\mathcal{M}$). Имеется естественное отображение $\overline{\mathcal{M}} \to \mathcal{M}$, превращающее $\overline{\mathcal{M}}$ в расслоение над $\mathcal{M}$ со слоем CP^1. Если заменить здесь каждое CP^1 на двухмерное комплексное пространство, из которого CP^1 получается как пространство лучей, мы придем к знаменитому спинорному пространству Дирака.

Мировую линию частицы со спином, скажем электрона, естественно представлять себе как линию именно в $\overline{\mathcal{M}}$, а не в $\mathcal{M}$. Лучи света тоже, естественно, поднимаются в $\overline{\mathcal{M}}$: в каждой своей точке такой луч определяет точку на соответствующем небосводе, куда он направлен; множество таких пар в $\overline{\mathcal{M}}$ и есть образ луча в $\overline{\mathcal{M}}$. Р. Пенроуз предложил рассмотреть замечательное пространство H, которое получается, если в $\overline{\mathcal{M}}$ каждый луч стянуть в одну точку. Математически H называется фактор-пространством $\overline{\mathcal{M}}$ по отношению эквивалентности, определенному лучами в $\overline{\mathcal{M}}$. Чтобы понять устройство H, разберем случай, когда $\mathcal{M}$ — плоский мир Минковского. Тогда в $\overline{\mathcal{M}}$ никакая точка не лежит на пересечении лучей (в отличие от $\mathcal{M}$) и каждый луч с каждым небосводом либо совсем не пересекается, либо пересекается в одной точке. Поэтому каждый небосвод CP^1 просто вкладывается в H без самопересечений, а некоторые небосводы попарно не пересекаются. Простейшее пространство, куда можно уложить много проективных комплексных прямых, — это проективная комплексная плоскость CP^2. Но она не может быть кандидатом на роль H, ибо любые две прямые в CP^2 пересекаются. На самом деле H лежит в CP^3 — проективном комплексном трехмерном пространстве, или пространстве проективных твисторов. Правда, небеса над точками мира Минковского — это не все прямые в CP^3, а лишь их часть, лежащая на пятимерной гиперповерхности (всё CP^3 шестимерно). Очень полезно ввести дополнительные «идеальные» точки плоского мира $\mathcal{M}$, небеса над которыми отвечают недостающим прямым в CP^3. Получится комплексное компактное пространство—время Пенроуза $C\mathcal{M}$: в нем координаты точек могут быть любыми комплексными числами и, сверх того, имеется еще целый комплексный световой конус, «лежащий на бесконечности».

Может ли такая абстрактная конструкция иметь какое-либо отношение к физике? По-видимому, ответ должен быть утвердительным. Один из аргументов состоит в открытой сравнительно недавно аналогии между квантовой теорией поля и статистической физикой. Если в основных формулах квантовой теории поля чисто мнимую координату ict заменить вещественной, то, грубо говоря, они перейдут в основные формулы статистической физики (роль ict играет обратная температура). Такая замена геометрически означает переход от мира Минковского с неопределенной метрикой к трехмерному миру Евклида с метрикой «сумма квадратов». Этот мир благополучно помещается в $C\mathcal{M}$, нужно лишь развернуть временну́ю ось в $\mathcal{M}$ на 90°. Все остальные точки $C\mathcal{M}$ получаются в результате интерполяции между мирами Евклида и Минковского, и, видимо, описания многих важных событий допускают аналитическое продолжение с $\mathcal{M}$ на $C\mathcal{M}$ или на часть $C\mathcal{M}$. Стандартный пример—сопоставление «туннельного перехода» квантовой механики с классической эволюцией системы в мнимом времени.

Сам «рай Пенроуза» $H = CP^3$ (пространство, где помещаются все небеса, но ничего не осталось от пространства—времени, естественно назвать раем), как оказалось совсем недавно, очень полезен для изучения уравнений Максвелла и их обобщений — уравнений Янга-Миллса, которые, как теперь предполагается, описывают глюонные поля, связывающие кварки в нуклоне. Имеются глубокие физические основания считать, что мир, заполненный лишь излучением (или частицами, летящими с околосветовыми скоростями, почти вдоль световых конусов), должен лучше описываться в терминах геометрии H, чем уже привычного нам вещественного четырехмерия. К пространству—времени нас привязывает масса, она мешает нам лететь со световой скоростью, когда время останавливается, а пространство теряет смысл. В мире света нет ни точек, ни мгновений; сотканные из света существа жили бы «нигде» и «никогда», лишь поэзия и математика способны говорить о таких вещах содержательно. Одна точка CP^3 есть вся история жизни свободного фотона— самое маленькое «событие», которое может произойти со светом.

Пространство — время, гравитация и квантовая механика. Самое важное, что следует усвоить при обдумывании взаимоотношений между нашими моделями классического Мира и квантовомеханическими принципами описания материи, состоит в том, что мы очень плохо понимаем эти взаимоотношения. Основные принципы описаний взаимно несогласованы.

Вот один из примеров расхождений. Пытаясь отметить точку Мира на мировой линии частицы массы m, мы не можем сделать это с меньшей ошибкой, чем радиус Шварцшильда $2Gm/c^2$ этой массы. Поэтому, пытаясь увеличить точность, мы должны пользоваться частицами как можно меньшей массы. С другой стороны, как заметили Ландау и Пайерлс, из соотношения неопределенности для пары (координата, импульс) и ограниченности всех скоростей скоростью света следует, что неопределенность положения не может быть меньше $\frac{\hbar}{mc}$, т. е. увеличение точности требует использования частиц как можно большей массы! Оба предела сравниваются при массе, которая находится из равенства $\frac{\hbar}{mc} = \frac{2G}{c^2}$ т. е. как раз массе Планка. Этот общий предел есть планковская длина. Значит, планковская единица длины (напомним, что ее порядок равен 10^{-33} см) указывает границы, за пределами которых заведомо неприменимо представление о пространственно-временной области, способной быть носителем элементарного события. События, изучаемые на современных ускорителях, происходят в гораздо больших областях, но все же предел локализации в рамках современных теорий указывается четко.

Квантовые принципы еще многими способами мешают представлять точку Мира как элементарное событие. Свободная квантовая частица с фиксированным четырехмерным импульсом k не локализована нигде — она «равномерно размазана по Миру», ее пси-функция есть плоская волна де Бройля e^{ikx}. Две тождественные квантовые частицы должны описываться пси-функцией, зависящей от двух точек Мира: симметричной для бозонов и антисимметричной для фермионов. Но это буквально означает, что в одной точке Мира вообще ничего не может «происходить», точнее говоря, через пси-функции материи и полей одна точка неразрывно связана со всеми остальными.

Образ вещественного четырехмерного Мира с метрикой Минковского в будущей теории может оказаться чем-то вроде квазиклассического приближения к бесконечномерному комплексному квантовому Миру. Например, в геометрической оптике, являющейся приближением к волновой, есть понятие каустики — множества точек, где интенсивность излучения в этом приближении бесконечна. Заманчиво представлять себе четырехмерный Мир как своего рода каустическое многообразие квантового волнового бесконечномерия. Наши затруднения с бесконечной плотностью энергии вакуума прекрасно разрешились бы в этой схеме.

Группа Лоренца является странной группой с вещественной точки зрения, но если заменить ее на SL (, C) — группу комплексных (2×2) матриц, мы получаем очень естественный объект — группу симметрий простейшего мыслимого пространства состояний квантовой системы. Не значит ли это, что спиновые степени свободы являются более фундаментальными, чем пространственно-временные? В группе SL (2, C) таинственное для нас разделение Мира на пространство и время содержится неявно, и потому его существование «объясняется» на основе принципов, не предполагающих такого разделения заранее. Более того, как мы видели, Мир без массы можно получить из SL (2, C) (или ее обобщения SL (4, C), и не вводя пространства—времени. Точки нашего четырехмерного Мира, или, лучше, его мелкие области, отмечены событиями, которые происходят с массивной материей. Возможно, и масса, и пространство—время есть результат спонтанного нарушения симметрии основных законов.

Такую теорию трудно придумать. Мы все еще пытаемся квантовать классическую Вселенную как атом водорода, вместо того чтобы пытаться получить ее образ как предел квантового описания. Может быть, первая квантовая модель Мира, скажем, вблизи Большого Взрыва, будет совсем простой математически, и лишь привычки инертного ума мешают нам угадать ее сейчас. Хотелось бы дожить до времени, когда такая модель будет предложена и принята.

5. ДЕЙСТВИЕ И СИММЕТРИЯ

«Физика там, где есть Действие» (неизвестный автор).

Шарик, катающийся в желобе под действием собственного веса,— простейшая физическая система. Состояние покоя — его простейшая история. Шарик может покоиться лишь в тех точках желоба, где касательная к желобу горизонтальна, иначе он скатится под уклон. Зададим положение шарика горизонтальной координатой x и обозначим через $V(x)$ его потенциальную энергию (пропорциональную высоте желоба) в точке x. Точки x, в которых шарик может покоиться,— это решения уравнения $\frac{dV}{dx} = 0$ или $dV = 0$; приращение функции при малом удалении от такой точки имеет высший порядок малости по сравнению с приращением координаты. В этих точках V стационарна.

Второй пример — мыльная пленка, натянутая, скажем, между двумя проволочными окружностями. Равновесная форма покоящейся пленки определяется тем, что при малых изгибах ее энергия поверхностного натяжения V меняется на величину высшего порядка малости по сравнению с некоторой естественной мерой величины изгиба. Энергия поверхностного натяжения V пропорциональна площади пленки, так что форма покоящейся пленки есть состояние, в котором площадь стационарна. Функция (или функционал) V в этом случае определена не на числах x, а на всевозможных поверхностях, натянутых на заданный контур. Вместо дифференциала dV принято писать «первую вариацию» δV.

Пусть, например, проволочные обручи радиусов r и R находятся в параллельных плоскостях на расстоянии l, и линия, соединяющая их центры, перпендикулярна этим плоскостям. Она является осью симметрии контура, и потому следует ожидать, что равновесная форма пленки будет поверхностью вращения кривой $q(x)$ с условиями $q(0) = r$, $q(l) = R$. Примем это; тогда энергия V пропорциональна площади $2\pi \int_0^l q(x)\left(1 + \left(\frac{dq}{dx}\right)^2\right)^{\frac{1}{2}} dx$ и уравнение $\delta V = 0$ можно переписать: $-q\frac{d^2q}{dx^2} + \left(\frac{dq}{dx}\right)^2 + 1 = 0$ и затем решить его с поставленными граничными условиями.

Такой переход к формализму и полезен и опасен. Мы постулировали, что симметрия граничных условий ведет к симметрии равновесной формы. Вот совершенно аналогичный пример, когда это, очевидно, не так. Нагрузим упругий вертикальный стержень сжимающей силой; при некоторой критической величине нагрузки он изогнется. Направление изгиба в плоскости, перпендикулярной стержню, ничем не выделяется в первоначальной осесимметричной картине, но выделено, когда изгиб произошел. В таких ситуациях физики говорят о спонтанном нарушении симметрии: явление, подчиняющееся некоторым законам, менее симметрично, чем сами эти законы. Еще в нашей задаче мы забыли о решении, состоящем из двух плоских пленок, натянутых на каждый обруч в отдельности. Это напоминание о том, что, исследуя функционал V на таком бесконечномерном многообразии, как пространство поверхностей, нужно попробовать разобраться заранее в его геометрическом устройстве. Такими задачами занимается топология. Лишь в последние годы доставляемые ею геометрические образы стали применяться в физике, например, в квантовой теории поля появилось представление о «топологическом заряде». К сожалению, нам некогда этим заниматься. Математически обычно полезно считать основным уравнением $\delta V = 0$, пусть с не до конца определенной областью существования функции V, подлежащей уточнению в ходе решения задачи. График V — это бесконечномерный желоб, по которому движется, ища покоя, наша система.

Физически этот образ оправдан поразительно универсальным принципом, который можно сформулировать так: развитие во времени фундаментальных классических систем есть их равновесие в пространстве — времени. Более точно, кинематика системы определяется описанием множества ее виртуальных историй — «пленок» в доступных ей пространственно-временны́х областях. На этих виртуальных историях определен функционал S размерности действия. Динамика системы описывается условием $\delta S = 0$. Переход от размерности энергии (V) к размерности действия (S) связан, конечно, с добавлением временной координаты. Действие первично; энергия есть всего лишь его производная по времени. В следующей фундаментальной теории действие останется, тогда как энергия станет квазиклассической величиной.

Виртуальная история φ системы в пространственно-временной области U классически определяется сечением расслоения степеней свободы системы над этой областью. Если область представлена в виде объединения двух своих непересекающихся частей U_1, U_2, а история φ есть объединение историй φ_1 и φ_2, то действие φ равно сумме действий φ_1 и φ_2. Этому постулату удовлетворяют почти все используемые в классической физике функционалы действия. Стало быть, полное действие истории φ в области U можно записать в виде суммы действий φ по многим маленьким областям, покрывающим U в пределе в виде интеграла $S(\varphi)=\int\limits_U L\,(\varphi(x,y,z,t))\,dx\,dy\,dz\,dt$, где (φ) — набор внутренних координат системы и их производных по пространству и времени. Функцию L (или ее интеграл по пространственным координатам) называют лагранжианом системы: это плотность действия. Если мы поместим в область пространства — времени две системы с лагранжианами L_1, L_2, то лагранжиан объединенной системы имеет вид $L_1+L_2+L_{12}$, где третий член есть «плотность взаимодействия». Две системы не взаимодействуют, если этот член равен нулю.

Само пространство — время («вакуум») вносит в лагранжиан член, пропорциональный его кривизне. Поэтому пространство — время можно рассматривать на тех же основаниях, что и системы, включающие массивную материю или электромагнитное поле.

Роль действия в квантовой физике чрезвычайно прояснил Ричард Фейнман, основываясь на более ранней работе Дирака. Его идеи за отсутствием фундаментальной квантовой теории мы вынуждены сейчас формулировать как рецепт «квантования», т. е. перехода от классического описания некоторой физической системы к ее квантовому описанию. Согласно этому рецепту следует представлять себе, что в квантовое описание истории системы вносит свой вклад каждая классическая история φ, но со своим комплексным весом (фазовым множителем) $e^{iS(\varphi)}$ (действие, конечно, измеряется в единицах $\hbar$). Поясним это подробнее. Фиксируем классическое поведение системы на границе области U, скажем, φ_1 и φ_2 в моменты времени t_1 и t_2. Квантовая теория ставит в соответствие этим условиям — «обручам для пленки» — не классическую историю развития от φ_1 до φ_2, а комплексное число $G\,(\varphi_1,\ t_1,\ \varphi_2,\ t_2)$ — амплитуду вероятности перехода из состояния $(\varphi_1,\ t_1)$ в состояние $(\varphi_2,\ t_2)$, квадрат модуля которой в принципе наблюдаем или входит в другие наблюдаемые величины. Предписание Фейнмана состоит в том, что эта амплитуда есть $\int e^{-iS(\varphi)}\, D\varphi$, где интеграл берется по бесконечномерному множеству классических историй, соединяющих (φ_1, t_1) и (φ_2, t_2); $D\varphi$ вместо $d\varphi$ служит напоминанием об этой бесконечномерности: это не дифференциал, а «элемент объема»!

В предыстории интегрального исчисления важное место занимает замечательный труд Кеплера «Стереометрия винных бочек». Интегралы, выражающие объемы тел вращения, полезных в народном хозяйстве, были вычислены в этой работе до появления общего определения интеграла. Математическая теория великолепных интегралов Фейнмана, которые физики пишут в огромных количествах, все еще не далеко ушла от стереометрии винных бочек. С точки зрения математика, каждое такое вычисление есть заодно определение того, что «вычисляется», либо построение текста в формальном языке, грамматика которого заранее не описана. В процессе таких вычислений физик спокойно делит и умножает на бесконечности (точнее, на нечто, что если бы оно было определено, вероятно, оказалось бы бесконечным); суммирует бесконечные ряды бесконечностей, предполагая при этом, что два-три члена ряда дают хорошее приближение ко всему ряду, и вообще живет в царстве свободы, нарушая все «моральные нормы».

Едва ли можно будет построить последовательную математическую теорию интегралов Фейнмана без прогресса в понимании физики. Сама идея «квантования» принадлежит не физике, а истории и психологии науки — содержательный смысл может иметь лишь «деквантование», т. е. переход от квантового описания к классическому, когда он разумен, но никак не наоборот. Классические поля, входящие в лагранжианы слабых и сильных взаимодействий, являются физическими фантомами: мы не знаем их смысла, помимо вторичного квантования, и неправдоподобно, чтобы они описывали виртуальные классические истории чего бы то ни было. (Считается, что с квантованием электромагнетизма дело обстоит лучше.)

Конечномерные квантовые модели позволяют предположить, какие черты фейнмановской формулировки существенны, а какие являются атавизмом. Как было объяснено в третьей главке, оператор эволюции замкнутой локализованной квантовой системы за ее локальное время t имеет вид $e^{-iS(t)}$, где $S(t)$ — на этот раз оператор размерности действия. Представляя себе разные мировые линии системы с разными локальными временами, мы убеждаемся, что квантовое действие есть связность в пространстве внутренних степеней свободы системы, определяющая физически допустимые истории как параллельные переносы. Рецепты вторичного квантования — это примитивное оформление представления о том, что из-за виртуального рождения частиц уже у вакуума пространство внутренних степеней свободы «в одной точке» бесконечномерно. Дальнейшее понимание блокируется, пока мы не отказались от идеи пространства — времени как основы всей физики.

«Симметрия обозначает тот вид согласованности отдельных частей, который объединяет их в единое целое. Красота тесно связана с симметрией» (Г. Вейль).

Кроме этой цитаты, я не повторю ничего из написанного Германом Вейлем в его замечательной книге «Симметрия» (М., «Наука», 1968). Ее нужно прочесть каждому, кто хочет пройти хотя бы часть дороги от восприятия симметрии как чувственной данности (цветы, орнаменты, кристаллы) до ее понимания как глубочайшей физико-математической идеи.

Универсальный математический образ симметрии — это группа G и ее действие на множестве X, например, группа S_n всех перестановок чисел $(1, \ldots, n)$. Действие есть отображение $G \times X \to X$, ставящее в соответствие паре (элемент группы g, точка множества x) элемент множества gx (образ x под действием g). Все элементы вида gx при переменном g составляют орбиту x под действием группы. Сама группа G никогда не задана как физический объект — мы можем вообразить твердое тело как чувственную данность, но множество всех его вращений есть идея, находящаяся на следующей ступени абстракции. Омонимичность слов «действие» в контекстах «действие группы» и «действие отрезка виртуальной истории» в основных европейских языках есть случайное следствие смутного исходного представления об «изменении как результате делания», но в выражении для квантового оператора эволюции $U(t)=e^{-iS(t)}$ эта омонимия неожиданно приобрела глубокий смысл.

Отделение понятия абстрактной группы от понятия ее действия на множестве было одним из великих достижений математической мысли и оказалось очень важным для физики. Вообразим себе атом водорода в виде электрона, движущегося в центральном кулоновском поле около неподвижного ядра. Евклидова группа вращений вокруг ядра $SO(3)$ действует на комплексном линейном пространстве квантовых состояний электрона. Оказывается, что все бесконечномерное пространство возможных связанных состояний разбивается на сумму конечномерных подпространств, на которых $SO(3)$ действует по отдельности. Эти действия — неприводимые линейные представления группы и суть возможные стационарные состояния атома водорода. Их точное математическое описание объясняет спектр, квантовые числа и т. п. Аналогично группа Пуанкаре — полная группа симметрий Мира Минковского — действует на пространстве квантовых состояний воображаемой одинокой элементарной частицы в Мире, где кроме нее ничего нет. Как показал Юджин Вигнер, экспериментальная классификация элементарных частиц по их массе и спину вкладывается в классификацию (бесконечномерных) неприводимых представлений группы Пуанкаре. Это дало повод к шутке, что мир двадцатого века состоит не из четырех стихий — огня, воздуха, земли и воды, а из неприводимых представлений некоторой группы. Теоретическая физика последних десятилетий усиленно занималась поисками группы симметрий фундаментальных взаимодействий: их законы (лангражианы) выступают как вторичный объект в математическом описании.

В физических изложениях, подчеркивающих скорее формализм, чем его теоретико-множественную интерпретацию, бывает иногда затемнено описание того, на чем действует группа симметрий. Вот два крайних случая, ведущих, по существу, к разной физике: а) группа симметрий G действует на фазовом пространстве системы и переводит в себя ее фазовый портрет; б) фазовое пространство X системы само представляется в виде множества орбит некоторого другого пространства Y под действием группы G.

К случаю а) принадлежит описание рассмотренного выше атома водорода. Дискретные инварианты неприводимого представления — это квантовые числа соответствующего состояния, но сам вектор состояния в пространстве представления еще не определен однозначно, как в примере с изгибом стержня. Симметрия снова спонтанно нарушена. В схемах описания фундаментальных взаимодействий часто постулируют схему нарушения симметрии более слабым взаимодействием. Нарушение симметрии — это тоже многозначный термин. О спонтанном нарушении можно говорить, когда фазовый портрет системы симметричен, но не симметричны его отдельные траектории. Учет нового взаимодействия нарушает симметрию всего фазового портрета, ибо, вообще говоря, изменяет его. Но если это изменение невелико, то его можно учитывать приближенно, рассматривая как малое возмущение исходной симметричной картины.

К случаю б) относятся так называемые калибровочные теории. Классическое поле материи в такой теории (объект, подлежащий вторичному квантованию) является не сечением расслоения внутренних степеней свободы, как мы говорили раньше, а целой орбитой таких сечений под действием калибровочной группы преобразований. Над одной точкой пространства — времени такая группа представляется некоторыми вращениями пространства внутренних степеней свободы, но эти вращения могут независимо и непрерывно меняться с изменением точки. Условие, чтобы лагранжиан теории был инвариантен относительно таких преобразований, накладывает на него очень жесткие требования, что резко ограничивает выбор лагранжиана. В суперсимметричных теориях, о которых мы вкратце говорили в первой главе, калибровочная группа может быть еще более общим объектом, перемешивающим, в частности, бозонные и фермионные поля. С калибровочными и суперсимметричными теориями сейчас связаны основные надежды на построение единой теории фундаментальных взаимодействий.

В заключение я хотел бы сказать несколько слов о теории чисел — высоко развитой и обладающей изумительной красотой математической дисциплине, которая до сих пор не нашла никаких глубоких естественнонаучных приложений. Один из главных объектов ее изучения — простые числа: целые положительные числа, не имеющие целых делителей, кроме себя и единицы. Еще в «Началах» Евклида содержится теорема о том, что простых чисел бесконечно много (зная их конечную систему $p_1, \ldots, p_n$, мы можем построить еще одно простое число как наименьший, отличный от единицы делитель числа $p_1, \ldots p_n + 1$). Это — замечательное рассуждение при всей его краткости. Вот более свежий образец теоремы о простых числах. Обозначим через $\tau(n)$ (n-е число Рамануджана) n-й коэффициент ряда, который получится после формального разложения бесконечного произведения $x \prod_{m=1}^{\infty} (1 - x^m)^{24}$ Если

p — простое число, то $|\tau(p)| < 2p^{\frac{11}{2}}$. Доказать это здесь никак невозможно; по оценке его автора, П. Делиня, чтобы изложить это доказательство, считая известным все, что знает студент третьего курса мехмата, понадобилось бы около двух тысяч страниц печатного текста. Вероятно, по отношению длины доказательства к длине формулировки эта теорема занимает рекордное место во всей современной математике. Разумеется, вместе с ней мы поняли еще много интересных вещей — для доказательства была создана большая новая теория («l-адические когомологии») и пришлось пользоваться двумя-тремя старыми (группы Ли, автоморфные функции...).

Замечательно, что самые глубокие идеи теории чисел обнаруживают далекоидущее сходство с идеями современной теоретической физики. Подобно квантовой механике, теория чисел доставляет совершенно не очевидные образы соотношения между непрерывным и дискретным (техника рядов Дирихле и тригонометрических сумм, p-адические числа, неархимедов анализ) и подчеркивает роль скрытых симметрий (теория полей классов, описывающая связь между простыми числами и группами Галуа полей алгебраических чисел). Хочется надеяться, что это сходство не случайно, и мы уже слышим новые слова о мире, в котором живем, но только не понимаем пока их смысла.

ЛИТЕРАТУРА

Вигнер Е. Этюды о симметрии. М., «Мир», 1971.
Компанеец А. С. Симметрия микро- и макромира. М., «Наука», 1978.
Математика в современном мире. М., «Мир», 1969.
Фейнман Р. Характер физических законов. М., «Мир», 1968.
Эйнштейн А. Физика и реальность. М, «Наука», 1965.

Юрий Иванович МАНИН
МАТЕМАТИКА И ФИЗИКА

Главный отраслевой редактор В. П. Демьянов.
Редактор Г. Г. Карвовский. Мл. редактор Т. Г. Иншакова. Обложка художника Л. П. Ромасенко. Худож. редактор М. А. Бабичева. Техн. редактор А. М. Красавина. Корректор С. П. Ткаченко

Т-21206. Индекс заказа 94312.
Сдано в набор 10.10.79 г. Подписано к печати 14.11.79 г. Формат бумаги 84×108¹/₃₂. Бумага типографская № 3. Бум. л. 1. Печ. л. 2. Усл. печ. л. 3,36. Уч.-изд. л. 3,83. Тираж 38 240 экз. Издательство «Знание», 101835, ГСП, Москва, Центр, проезд Серова, д. 4. Заказ 2432.
Цена 11 коп.
Чеховский полиграфический комбинат Союзполиграфпрома
Государственного комитета СССР по делам издательств,
полиграфии и книжной торговли
г. Чехов Московской области