А. А. Кириллов

Лекции 28 декабря 1997 года и 3 января 1998 года

Метод орбит и конечные группы

Часть 1

Честно говоря, я несколько смущен тем, что назвал бы overqualified присутствующей аудитории. Потому что мне были обещаны студенты второго и отчасти первого курса. Но будем надеяться, что хотя бы записи этой лекции попадут по адресу.

Я начну с не относящейся к лекции темы, а именно, что происходит нового в математике. Это отдельная тема, которой можно было бы занять все 2 часа. Но две вещи мне все же хотелось бы упомянуть. Об одной вещи я скажу совсем коротко: это последняя статья Максима Концевича, которая закрыла тему деформационного квантования и которая сейчас бурно обсуждается во всех математических центрах. А вторая —это тривиальное доказательство так называемой космологической теоремы Конвея. Эта тема специально предназначена для студентов первого курса. Конвей —это совершенно нетривиальный математик, который занимается совершенно неожиданными вещами. Например, ему пришло в голову исследовать так называемый Audioactive operator —оператор произнесения. А именно, представьте себе, что вы, как чукча, который едет по тундре, и что видит, о том и поет песню. И вот такой человек видит число, для простоты единицу. Он видит —одна единица, и пишет: «Одна единица (11)». Тем самым получается второй член нашей последовательности. Он видит второй член последовательности —две единицы, и пишет: «Две единицы (21)». Это третий член нашей последовательности. Читая его вслух, получаем: «Одна двойка, одна единица (1211)». Следующий член: «Одна единица, одна двойка, две единицы (111221)». Так возникает последовательность

$$1,\ 11,\ 21,\ 1211,\ 111221,\ 312211, \ldots$$

Конвей заинтересовался, какие у этой последовательности асимптотические свойства. Довольно простое упражнение —доказать, что в этой последовательности не встретится никаких цифр, кроме 1, 2 и 3. Следующее нетрудное упражнение —если в первом и втором члене нет никаких цифр, кроме 1, 2 и 3, то и дальше никаких других цифр не встретится.

Оказывается, что у оператора произнесения есть ровно одна неподвижная точка: 22. Все остальные последовательности начинают расти, но весьма нерегулярным образом: иногда длина сильно увеличивается, а иногда не очень. Конвей доказал, что для всех последовательностей, состоящих из цифр 1, 2 и 3 (за исключением неподвижной точки 22), отношение длины слова a_n к длине слова a_{n-1} стремится к определенному пределу, который равен 1,301577269... Этот предел —алгебраическое число, и специалисты по динамическим системам, конечно, догадаются, что это —максимальное собственное значение некоторого оператора, но подробности я опускаю.

Это не относится к теме лекции, но я специально потратил несколько минут, чтобы все, кто хотел прийти, пришли бы и сели.

Теперь о теме сегодняшней лекции. Я собирался коротко рассказать о некоторых задачах, возникающих в связи с применением метода орбит к конечным группам. Поскольку лекция была рассчитана на способных первокурсников, то у меня принцип такой: я иногда произношу слова, которые слушатель не обязан понимать в данный момент, но может спросить у старшего товарища, или сообразить сам, или прочитать в книжке, и тем самым ликвидировать локальное непонимание. Идейных непониманий не должно быть, потому что каждые несколько минут будет сменяться, как говорит Юрий Иванович Манин, парадигма, и поэтому можно забыть всё, что было раньше, и начинать внимать сначала.

В целом речь идет о теории представлений. Я полагаю, что первокурсник, обучающийся в Независимом университете, представляет себе, что такое группа, что такое линейное пространство, что такое линейный оператор, что такое линейное представление этой группы (т. е., выражаясь более ученым языком, гомоморфизм группы в группу обратимых линейных операторов), и понимает, что полезно заниматься этой наукой —теория представлений имеет много разных приложений, о которых я сейчас говорить не буду.

Мой вклад в теорию представлений состоит в том, что я предложил так называемый метод орбит. Этот метод орбит является в какой-то степени соединением двух вещей. А именно, симплектической геометрии и теории представлений. Или, если подниматься на более высокую степень абстракции, это —соединение современной мат-физики с математикой. Дело в том, что современная мат-физика уже постепенно вытесняет математику из ее канонически устоявшихся областей и придумывает новые подходы, новые задачи, попутно решая старые. И уже примерно половина математических журналов постепенно наполняются статьями, в которых мат-физика играет ту или иную роль, иногда решающую.

Я потрачу несколько минут на то, чтобы объяснить, что такое симплектическая геометрия, хотя мне и говорили, что эта тема здесь в какой-то степени уже известна. Начну с того, что рассмотрю гладкое многообразие M. Гладкое многообразие —это то, на чем можно развивать математический анализ. А именно, это множество M, которое локально устроено как евклидово пространство. Это значит, что множество M можно покрыть открытыми областями[1] U_i так, чтобы для каждой области существовало взаимно однозначное отображение на некоторую область в евклидовом пространстве: $U_i \xrightarrow{\varphi} V_i \subset \mathbb{R}^n$. Тогда все, что мы знаем из математического анализа, а именно, правила обращения с функциями многих переменных, дифференцирование, интегрирование, подстановки, дифференциальные уравнения и прочее, можно перенести на многообразие M.

Простейший пример многообразия —это окружность S^1. На ней нельзя ввести одной координаты. Это опять же задача для первокурсников: доказать, что любая непрерывная функция на окружности принимает по крайней мере одно значение в двух разных точках, и тем самым не может служить координатой. С другой стороны, нетрудно сообразить, что на окружности можно выделить две области, на каждой из которых такая координата существует. На всей окружности, кроме верхней точки, можно ввести координату, проектируя окружность из верхней точки на горизонтальную прямую. При этом у самой верхней точки образа нет, но можно воспользоваться другой проекцией, из нижней точки. Тогда все, кроме нижней точки, можно отобразить на ту же горизонтальную прямую. Получаем две карты, которые в совокупности покрывают наше многообразие, нашу окружность. Назовем эти локальные координаты x_+ и x_-. Для окружности радиуса 1 они связаны соотношением $x_+x_- = 1$ ($x_+x_- = r^2$ для окружности радиуса r).

[1] Когда я говорю об открытых областях, я подразумеваю, что M —не просто множество, а топологическое пространство, для которого определены открытые и замкнутые множества.

Пока наш основной объект —многообразие, т. е. множество, на котором можно заниматься математическим анализом. Кроме того я сказал слово «группа». У группы можно перемножать ее элементы. Обычно группы возникают как группы преобразований чего-либо. Если мы объединим две структуры —группы и многообразия, то получится так называемая группа Ли. Это будет наш основной объект.

Группа Ли —это такой объект, который одновременно является многообразием и группой. Это значит, что на нем одновременно существуют локальные координаты и закон умножения, обладающий всеми положенными свойствами: ассоциативность, существование обратного элемента, существование единицы.

Та же окружность является прекрасным примером группы. Закон умножения проще всего ввести так. Представим себе, что окружность имеет радиус 1 и лежит не просто на вещественной плоскости, а на комплексной плоскости. Тогда ее можно задать уравнением $|z| = 1$ и ввести групповой закон в виде обычного умножения комплексных чисел: $z_1 z_2$. Если два числа имели модуль 1, то их произведение тоже имеет модуль 1.

Я забыл сказать, что две структуры —группы и многообразия — должны быть естественным образом связаны между собой. А именно, на многообразии есть понятие гладкой функции. При определении многообразия я намеренно забыл указать, что соответствия $U_i \xrightarrow{\varphi} V_i$ произвольные, но там, где возникают две системы координат, мы должны потребовать, чтобы переход от одной системы координат к другой совершался с помощью гладких функций. Гладкость означает существование нескольких (одной, двух, трех и так далее) непрерывных производных. Как правило, удобно сразу об этом не заботиться, а считать, что все функции бесконечно дифференцируемы. Так вот, согласование аксиом группы с аксиомами многообразия очень просто: групповой закон и отображение, которое каждому элементу сопоставляет обратный элемент, должны быть гладкими функциями.

Упражнение. Записать групповой закон для окружности в координатах x_+ и x_- и убедиться, что он является непрерывной функцией.

Дальше мы будем заниматься теорией представлений. Я буду рассматривать только унитарные представления. Это означает, что группа G отображается в группу $\mathrm{aut}(H)$, где H —гильбертово пространство. Гильбертово пространство —это обобщение линейного пространства в двух направлениях. А именно, размерность пространства не ограничивается и может быть бесконечной. И во-вторых, как правило, вместо вещественного поля рассматривается комплексное поле. Можно рассматривать и вещественные представления, но они сводятся к рассмотрению комплексных, а многие факты и теоремы над комплексным полем выглядят намного проще. Еще в гильбертовом пространстве есть наследуемое из конечномерного случая понятие скалярного произведения двух векторов. Оно позволяет ввести норму вектора и понятие ортогональности векторов. Под автоморфизмами H я понимаю совокупность операторов, которые сохраняют все структуры в H: структуру линейного пространства и скалярное произведение. Такие операторы называют унитарными.

Итак, каждому элементу g группы G сопоставляется некоторый унитарный оператор $U(g)$ так, чтобы выполнялось функциональное уравнение $U(g_1g_2) = U(g_1)U(g_2)$: произведение элементов группы переходит в произведение операторов. Есть довольно развитая наука о том, как обращаться с унитарными представлениями. В частности, есть естественные понятия эквивалентности представлений, разложения их в прямые суммы. Те представления, которые не допускают разложения в виде прямой суммы, называют неразложимыми. Они же оказываются в этом случае и неприводимыми, т. е. не имеют нетривиальных подпредставлений. И первая задача, возникающая при рассмотрении каждой группы, —это описание совокупности всех ее неприводимых унитарных представлений с точностью до эквивалентности. Множество классов эквивалентности неприводимых унитарных представлений группы G обозначают $\widehat{G}$. В английской мат-физике есть очень удобный термин, сокращающий это длинное название, а именно: unirrep (UNItary IRREducible REPresentation). Эти unirrep'ы и есть наш основной объект исследования, и для каждой группы хорошо бы иметь список таких unirrep'ов с указанием их свойств: как они ведут себя при ограничении на подгруппу, или при индуцировании (что это такое, я пока не буду говорить) на бóльшую группу, каковы их матричные элементы, каковы их характеры, инфинитезимальные характеры, и прочее.

Возникает столько же задач, сколько имеется групп. А поскольку группы встречаются в любой области математики и ее приложений, то и эта задача возникает везде. Но методы решения этой задачи совершенно разные для разных областей, и иногда одно решение не похоже на другое. Скажем, для компактных групп есть один набор теорем, для полупростых групп другой набор теорем, для нильпотентных третий набор теорем, и т. д.

Метод орбит, о котором я буду говорить, имеет то преимущество, что он рассматривает все группы сразу и дает универсальный ответ, как описать множество неприводимых представлений и как ответить на большинство вопросов, с ним связанных. Подробно описывать этот метод я не буду. Скажу лишь, какие понятия с ним связаны и что нужно делать. Наибольшую популярность из всех моих увлекательных дел получил так называемый User's Guide —руководство к использованию. Когда вы покупаете электрический утюг или что-то еще, к нему прилагается инструкция, как им пользоваться. Вот существует и такая инструкция, как пользоваться методом орбит. Что вам нужно —и как получить ответ: какие вещи нужно определить, какие вещи перемножить, какие сложить, чтобы получился ответ.

Я перечислю основные ингредиенты, а потом перейду к основной теме моей лекции —как применить метод орбит к конечным группам. Для начинающих математиков всегда приятнее иметь дело с чем-нибудь легко обозримым и конечным. Я, когда был первокурсником, считал, что могу решить любую задачу, в которой идет речь о конечном наборе объектов. Оказалось, что это не совсем так —есть вполне конечные задачи, которые люди до сих пор решить не могут. Некоторые такие задачи я позже еще попробую сформулировать. А сейчас я закончу описание метода орбит и скажу, каким образом его можно приложить к конечным группам.

Следующим нашим объектом будет алгебра Ли $\mathfrak{g} = Lie(G)$. Буква $\mathfrak{g}$, маленькая готическая буква g, —теперь уже общепринятое обозначение для алгебры Ли, связанной с группой Ли G. Теория групп Ли и теория представлений претерпели долгие изменения, и у каждой эпохи были свои обозначения. Сейчас уже более или менее устоялся обычай обозначать группу Ли большой латинской буквой, а алгебру Ли —маленькой готической. Хотя есть отсталые ученые, которые используют устаревшие обозначения, и наоборот —передовые ученые, которые выдумывают свои новые обозначения. Но я буду придерживаться наиболее общепринятых обозначений.

Алгебра Ли —это то, что остается от группы, если мы рассматриваем инфинитезимальную окрестность единицы. А именно, возьмем касательное пространство T_eG к группе Ли G в единице e, и будем интерпретировать векторы этого касательного пространства как точки, бесконечно близкие к единице. Что остается от группового закона, если мы рассматриваем только точки, очень близкие к единице? Если мы введем локальные координаты с началом координат в единице, то каждый элемент группы изобразится вектором. Возьмем два таких вектора $\vec{x} = (x_1, \ldots, x_n)$ и $\vec{y} = (y_1, \ldots, y_n)$ и рассмотрим их произведение как элементов группы Ли $\vec{x} \cdot \vec{y} = \vec{z}$. Координаты полученного вектора $\vec{z}$

будут функциями от $\vec{x}$ и $\vec{y}$:

$$z_k = \varphi_k(\vec{x}, \vec{y}).$$

Оказывается, что если учесть ассоциативность умножения и то, что умножение любого элемента на единицу дает тот же самый элемент, то функция φ_k становится совсем не произвольной, а при подходящем выборе локальных координат имеет следующий замечательный вид:

$$\varphi_k(\vec{x}, \vec{y}) = \vec{x} + \vec{y} + [\vec{x}, \vec{y}] + \dots$$

(Многоточием обозначены члены третьего порядка и более высоких порядков.) Эта формула означает, что на всякой группе Ли групповой закон в первом приближении коммутативен, а во втором приближении задается билинейным антисимметричным выражением $[\cdot, \cdot]$, которое удовлетворяет тождеству Якоби. Эта билинейная операция превращает касательное пространство к группе в единице в так называемую алгебру Ли. С каждой группой Ли связана алгебра Ли.

Если в алгебре Ли выбрать базис $X_1, \dots, X_n$, то для того чтобы задать структуру алгебры Ли, достаточно задать произведение базисных векторов: $[X_i, X_j] = c_{ij}^k X_k$. Здесь я использую стандартное эйнштейновское правило, что если в некотором выражении один индекс встречается сверху и снизу, то по этому индексу подразумевается суммирование. Таким образом, алгебра Ли —это просто набор структурных констант c_{ij}^k. И великое открытие Софуса Ли состояло в том, что в этом наборе структурных констант заключена вся информация о группе Ли. Всё, что вы захотите о ней узнать, можно, в принципе, извлечь из этого набора структурных констант. Чуть позже я скажу, как это сказывается на тех задачах, которыми мы будем заниматься.

Геометрически нам понадобится не само пространство $\mathfrak{g}$, а двойственное к нему пространство, которое обозначается $\mathfrak{g}^*$. Это —пространство линейных функционалов на пространстве $\mathfrak{g}$. Если в исходном пространстве был базис $X_1, \dots, X_n$, то можно ввести двойственные функционалы $F^1, \dots, F^n$ так, что каждый функционал на своем векторе равен 1, на прочих базисных векторах равен 0, а дальше распространяется по линейности.

Пространство $\mathfrak{g}^*$, в отличие алгебры Ли $\mathfrak{g}$, никакого умножения в себе не содержит —функционалы перемножать нельзя. Правда, можно их насильно заставить умножаться, и тогда мы получим очень интересную конструкцию —так называемую биалгебру. Она связана с квантовыми группами, алгебрами Дринфельда и прочим. Это отдельная наука, которой я здесь касаться не буду. Нас интересует это пространство как линейное пространство.

И последний ингредиент из общей теории групп и алгебр Ли, который мне здесь понадобится, —это действие группы Ли G на алгебре Ли $\mathfrak{g}$ и двойственное действие на пространстве $\mathfrak{g}^*$. Это действие можно определять многими эквивалентными способами. Проще всего убедиться, что такое действие существует, можно так. Для любого $x \in G$ определим преобразование $A(x)\colon g \mapsto xgx^{-1}$. Это преобразование является автоморфизмом группы. Кроме того, это —гладкое преобразование многообразия G, оставляющее неподвижной точку e. Поэтому есть производное отображение, которое действует в касательном пространстве. Оно обозначается $\operatorname{Ad} x$ и переводит элемент X касательного пространства в другой элемент касательного пространства, который я тоже условно обозначу xXx^{-1}. Такое обозначение оправданно, потому что подавляющее большинство групп Ли можно реализовать как подгруппы в группе матриц (если рассматривать и бесконечные матрицы, то так реализуется любая группа Ли; если же рассматривать только матрицы конечного порядка, то есть исключения). А для группы матриц эту формулу можно понимать буквально. Большинство физиков так и делает, потому что для них любая группа состоит из матриц.

Мы описали действие в самой алгебре Ли, а нам нужно действие в сопряженном пространстве. Обозначу $K(x)$ действие элемента x в сопряженном пространстве $\mathfrak{g}^*$ (K —от слова «коприсоединенное»: представление $\operatorname{Ad} x$ называют присоединенным, а представление $K(x)$ — коприсоединенным). $K(x)$ действует на функционал F по следующему правилу:

$$\langle K(x)F, X\rangle = \langle F, \operatorname{Ad} x^{-1}X\rangle.$$

Матрица коприсоединенного представления отличается от матрицы присоединенного представления тем, что она заменяется на обратную матрицу и транспонируется. Казалось бы, разница небольшая. Но оказывается, что присоединенное представление и коприсоединенное представление выглядят совершенно по-разному. Грубо говоря, коприсоединенное представление имеет гораздо больше свойств, чем присоединенное.

Одно из геометрических свойств, которое сразу бросается в глаза, заключается в том, что все коприсоединенные орбиты —симплектические многообразия, причем действие группы сохраняет симплектическую структуру. Выражение «коприсоединенные орбиты» —это жаргон; полностью надо говорить «орбиты коприсоединенного представления». Орбита получается следующим образом: берем функционал F и применяем к нему $K(x)$ для всех $x \in G$.

Теперь еще одно краткое отступление о симплектических многообразиях. Это понятие пришло в математику из механики. Оно является, в каком-то смысле, нечетным аналогом риманова многообразия. Риманово многообразие, может быть, даже ближе начинающим математикам. Это —многообразие, на котором можно измерять длины, т. е. задана положительно определенная квадратичная форма в касательном пространстве. В симплектическом многообразии задается не квадратичная форма в касательном пространстве, а кососимметрическая билинейная форма, т. е. антисимметричное скалярное произведение. Это, грубо говоря, —нечетный аналог понятия метрики. Вообще, сейчас люди склоняются к тому, что у каждого понятия должен быть свой нечетный аналог, и только рассматривая одновременно сам объект вместе с его нечетным аналогом, мы получаем правильное понимание того или иного явления.

Более точная аналогия между римановыми многообразиями и симплектическими многообразиями получается, если учесть, что у римановых многообразий самое главное качество —это кривизна. Бывают плоские пространства, а бывают неплоские, и эта неплоскость выражается в том, что метрика имеет кривизну. Если метрика заменяется на кососимметрическую квадратичную форму, то для нее тоже есть понятие кривизны. Так вот, симплектические многообразия —это нечетные аналоги плоских римановых многообразий, т. е. кривизна у них равна нулю. Координатное выражение этого таково: если мы локально запишем кососимметрическую форму в виде $\omega = \omega_{ij} dx_i \wedge dx_j$, то условие плоскости запишется в виде

$$d\omega = \frac{\partial \omega_{ij}}{\partial x_k} dx_i \wedge dx_j \wedge dx_k = 0.$$

Известно, что всякую плоскую метрику можно в подходящей системе координат заменить на постоянную метрику. Точно так же симплектическую структуру можно в подходящей системе координат заменить на постоянную симплектическую структуру, т. е. выбрать локальную систему координат так, что коэффициенты ω_{ij} —постоянные числа. Локально симплектическая геометрия никаких инвариантов не имеет. А глобальные инварианты она имеет, и это очень интересная наука —симплектическая топология. С ней каждый может познакомиться по недавно вышедшему собранию сочинений Арнольда. Кстати сказать, оно интересно и во многих других отношениях.

Каждая орбита коприсоединенного представления каноническим образом является симплектическим многообразием, и симплектическая форма ω, которая на ней существует, инвариантна относительно действия нашей группы. Это, казалось бы, очень сложно формулируемое свойство имеет одно яркое и наглядное следствие: все коприсоединенные орбиты обязательно имеют четную размерность.

Самый простой пример такой. Если в качестве исходной группы взять группу вращений трехмерного пространства, то соответствующая алгебра Ли —это обычное трехмерное пространство, присоединенное представление —это так называемое тавтологическое представление (каждое вращение трехмерного пространства и действует именно как вращение трехмерного пространства), а орбиты —это двумерные сферы или само начало координат. Как видите, встречаются только размерности 0 и 2.

Основной принцип метода орбит состоит в том, что следующие два множества не то, чтобы совпадают, но в каком-то смысле близки друг к другу и отвечают друг за друга. А именно, множество $\widehat{G}$ (классы эквивалентности унитарных неприводимых представлений, сокращенно unirreps) и множество коприсоединенных орбит $\mathfrak{g}/G$:

$$\widehat{G} \cong \mathfrak{g}/G.$$

Написанное соответствие выглядит как нечто очень универсальное: оно должно быть справедливо для любой группы Ли. Поэтому естественно воспринимать его как некий общий принцип, который справедлив всегда, когда имеют смысл обе части равенства. Например, группу Ли можно понимать в несколько обобщенном смысле. Можно считать, что это не классическое гладкое многообразие с вещественными координатами, а, скажем, комплексное многообразие, или многообразие над другим полем (алгебраическое многообразие). Можно также считать, что это бесконечномерное многообразие. Можно, наконец, считать, что это квантовая группа, которая группой вообще не является, но понятие алгебры Ли для нее есть и понятие коприсоединенного представления тоже есть. И самый интересный для меня сегодня случай, это случай, когда наша группа является алгебраическим многообразием над конечным полем. Тогда она сама как группа является конечной группой. И все теоремы и гипотезы, о которых я буду говорить, относятся к очень простому и обозримому объекту —конечным группам. Замечательно, что некоторые утверждения, которые получаются применением метода орбит, оказываются известными верными теоремами, другие оказываются хотя и верными, но новыми и неизвестными теоремами, третьи оказываются неверными, и, наконец, четвертые остаются проблемами. В основном о проблемах я и хотел поговорить.

Общий принцип метода орбит состоит в том, что он —линейное, или, как еще говорят, квазиклассическое, приближение к настоящей теории представлений. А именно, вместо нелинейного многообразия (группы Ли) рассматривается линейное многообразие (ее алгебра Ли) и закон умножения рассматривается только с точностью до квадратичных членов, что является первым шагом приближения. Поэтому чем ближе группа Ли к алгебре Ли, тем лучше работает этот метод.

Между группой Ли и алгеброй Ли существует некоторая связь, а именно, имеется естественное отображение из алгебры Ли в группу Ли. Если у вас есть касательный вектор, т. е. направление движения из единицы, то можно двигаться в этом направлении наиболее естественным образом —так, чтобы траектория движения была однопараметрической подгруппой. Параметр на этой кривой можно выбрать так, чтобы выполнялись следующие два свойства:

1) $g_t g_s = g_{t+s}$ (это можно сделать в силу общей теоремы: любая одномерная группа локально есть аддитивная группа прямой);

2) $\dot{g}_0 = X$, где X —заданный вектор.

Система обыкновенных дифференциальных уравнений имеет единственное решение с таким начальным условием, и g_1 по определению обозначается $\exp X$. Это обозначение не случайно: для матричной группы Ли $\exp X = \sum_{k\geq 0} X^k/k!$, т. е. мы получаем настоящую экспоненту.

Отображение exp определено на всей алгебре Ли. Но образ этого отображения покрывает не всю группу Ли. Поэтому обратное отображение, которое естественно было бы назвать логарифмом, не везде определено, не везде однозначно, и вообще обладает гораздо худшими свойствами, чем прямое отображение exp.

Метод орбит проще всего работает в том случае, когда отображение exp взаимно однозначно. Самый замечательный пример, когда это отображение взаимно однозначно, я и хочу сегодня рассмотреть в качестве основного объекта.

Обозначим через G_n алгебраическую группу верхнетреугольных матриц. Я сделаю краткое отступление об алгебраических многообразиях. Это очень важное понятие, и даже не столь важно само понятие, как идеология, с ним связанная. На заре этой науки алгебраическим многообразием называлось множество совместных решений алгебраической системы уравнений. Со временем было понято, что это неправильная точка зрения. Например, если вы рассматриваете вещественные числа, то уравнения $x^2 + y^2 = -1$ и $x^2 + y^2 = -3$ эквивалентны, потому что ни то ни другое решений не имеет. В то же время ясно, что это немножко разные уравнения, они должны описывать разные множества. И действительно, если мы рассмотрим комплексные решения, то решения этих двух уравнений окажутся разными. Или рассмотрим еще другое уравнение $x + y = x + y + 2$. Оно тоже не имеет решений, поэтому должно рассматриваться как эквивалентное двум первым уравнениям. Но ясно, что никакого преобразования, переводящего их друг в друга, придумать нельзя. Это разные объекты. Если первые уравнения можно спасти, рассматривая комплексные решения, то это уравнение не имеет и комплексных решений. Но можно, например, рассматривать решения в поле вычетов по модулю 2. Тогда это уравнение задает всю плоскость —любая пара (x, y) является решением.

Системы уравнений имеют разные множества решений, если их рассматривать над разными полями, или, более общо, над разными алгебрами. Постепенно люди пришли к такому пониманию: система уравнений —это функтор из категории алгебр в категорию множеств. А именно, если у вас есть какая-то система уравнений с коэффициентами из какого-то поля k и если вы зададите какую-то алгебру A над этим полем, то множество решений над этой алгеброй будет каким-то множеством X_A. Поэтому мы получаем соответствие $A \mapsto X_A$, и это соответствие есть функтор. Что такое функтор, я объяснять не буду. Каждый должен спросить у умного соседа.

Это уже правильное понимание того, что такое алгебраическое многообразие. Какие нужно рассматривать алгебры —это отдельный вопрос. Иногда достаточно рассматривать комплексные числа, иногда достаточно рассматривать более узкие образования. А иногда полезно рассматривать также некоммутативные алгебры A. Это уже начала так называемой некоммутативной алгебраической геометрии.

Алгебраическая группа —это одновременно группа и алгебраическое многообразие. Только надо помнить, что алгебраическое многообразие не является множеством. Оно становится множеством лишь после того, как вы выберете некоторую алгебру A. Так и алгебраическая группа не является группой в обычном смысле этого слова по той простой причине, что она не является множеством. А группа, как учат в теории групп, является множеством.

Алгебраическая группа —это нечто, что становится группой после того, как вы выберете алгебру A и рассмотрите множество решений соответствующей системы уравнений над алгеброй A. Это, как вы понимаете, не одна группа, а целое семейство групп, зависящих от выбора алгебры.

После этой подготовки я напишу пример алгебраической группы. Группа G_n будет алгебраическим подмногообразием в множестве всех матриц: $G_n \subset \mathrm{Mat}_n$. И будет она задаваться такой системой уравнений: $x_{ij} = 0$ при $i > j$ и $x_{ij} = 1$ при $i = j$. Можно рассматривать решения этой системы над любым полем и над любой алгеброй. Я утверждаю, что G_n —алгебраическая группа. Что это такое, я не буду аккуратно определять. Проверю лишь, что как только мы рассматриваем решения над какой-то алгеброй, то в результате получается группа. Пусть для простоты $n = 3$. Тогда множество решений этой системы над заданной алгеброй A выглядит так:

$$\begin{pmatrix} 1 & a_{12} & a_{13} \\ 0 & 1 & a_{23} \\ 0 & 0 & 1 \end{pmatrix},$$

где $a_{ij} \in A$. Такие матрицы называют строго верхнетреугольными (строго —потому что на диагонали стоят единицы). Для любой алгебры A множество строго верхнетреугольных матриц является группой относительно обычного матричного умножения.

Теперь уже я могу сформулировать основной предмет изучения сегодняшней лекции. А именно, что может сказать метод орбит про теорию представлений этой группы в случае, когда $A = \mathbb{F}_q$ —конечное поле из q элементов. Здесь $q = 2, 3, 4, 5, 7, 8, 9, 11, 13, \ldots$ Не существует полей, состоящих из 6, 10, 12 элементов. Самый интересный контрольный пример —поле из 4 элементов: существует поле из 4 элементов, но оно не является полем вычетов по модулю 4.

Интересные задачи здесь возникают, когда начинаешь заниматься так называемой асимптотической теорией представлений. Это значит, что у нас есть не одна конечная группа, а целая бесконечная серия конечных групп, и мы интересуемся асимптотическими свойствами. Например: нельзя ли эти асимптотические свойства проинтерпретировать как какие-то результаты по теории представлений какого-то идеального объекта, находящегося в бесконечности.

Для группы $G_n(\mathbb{F}_q) \subset \mathrm{Mat}_n(\mathbb{F}_q)$ есть два параметра: n и q. Они играют совершенно разную роль, хотя и есть любопытные соображения, по которым есть связь между этими двумя параметрами. Можно действовать двумя способами. Во-первых, можно фиксировать n и менять q. Это означает, что мы рассматриваем одну и ту же алгебраическую группу над разными полями. Оказывается, что метод орбит к этому прекрасно приспособлен и дает универсальное описание представлений всех этих групп единообразным образом с некоторыми дополнительными поправками, зависящими от n. А вот если мы фиксируем q и n устремляем к бесконечности, то эта задача намного интереснее, и даже самые простые вопросы пока еще не имеют ответов. Вот именно эту задачу я и хочу обсудить: как устроена теория представлений группы верхнетреугольных матриц очень большого порядка с элементами из фиксированного конечного поля. Что касается поля, то здесь во многих случаях от поля мало что зависит, и результат примерно одинаков для всех полей. Поэтому проще всего ограничиться простейшим полем, а именно полем $\mathbb{F}_2 = \{0, 1\}$, состоящим из двух элементов.

Поле $\mathbb{F}_2 = \{0, 1\}$ является простейшим полем в силу аксиомы, утверждающей, что $0 \neq 1$. Эта и только эта аксиома запрещает полю состоять из одного элемента. И в некоторых случаях имеет смысл говорить о поле, состоящем из одного элемента. Тогда эта аксиома нарушается, но других препятствий нет. И любопытно, что некоторые формулы имеют смысл и в том случае, когда q обращается в 1. Есть даже такая метаматематическая философия, которая говорит, что случай $q = 1$ в какой-то степени соответствует бесконечному полю, а именно, комплексному полю $\mathbb{C}$. Более того, $q = -1$ соответствует полю $\mathbb{R}$. Но это отдельная тема, о которой я не хочу сейчас говорить из-за недостатка времени. Желающие могут посмотреть рассуждения на эту тему в «Сочинениях» Арнольда.

В дальнейшем я часто буду ограничиваться простейшим полем из двух элементов, хотя интересные вопросы будут и для общих полей.

Мы начнем с вопроса о том, как описать коприсоединенную орбиту. Уж если мы верим, что метод орбит работает, то тогда неприводимые представления соответствуют коприсоединенным орбитам, поэтому хорошо бы научиться как-то описывать коприсоединенные орбиты. И тут возникает следующий вопрос. В алгебраической геометрии есть понятие элемента общего положения. И орбиты общего положения описываются обычно легко, во всяком случае, для треугольной группы они описываются легко. А дальше идут вырожденные орбиты, и чем вырожденнее класс орбит, тем труднее их классифицировать. Этот вопрос пока еще не решен до конца. Более того, он не решен ни для какого поля. Для вещественного поля этот вопрос возник очень давно, лет 30 тому назад. Уже тогда такая задача была известна. Уже тогда было непонятно, как классифицировать представления треугольной группы с вещественными элементами. После появления метода орбит стало ясно, что это то же самое, что классифицировать коприсоединенные орбиты. Задача, казалось бы, существенно упростилась: вместо бесконечномерных представлений в гильбертовом пространстве возник вполне конечный объект —конечномерные треугольные матрицы и конечномерная группа, действующая на них указанным образом. Тем не менее, классификация орбит не поддается легкому решению. И даже есть соображения, почему тривиального решения здесь и не должно быть. С другой стороны, те, кто занимался этой классификацией, рано или поздно приходили к убеждению, что если такая классификация и получится, то от поля она не будет сильно зависеть. Для комплексного поля, для вещественного, для конечного —классификация примерно одна и та же. Неизвестно только, как она устроена. Поэтому случай простейшего поля вызывает наибольший интерес.

Теперь я напишу несколько простых формул, показывающих, как устроено коприсоединенное представление для треугольной группы. Сама группа G_n состоит из элементов

$$g = \begin{pmatrix} 1 & * & * & * & * \\ 0 & 1 & * & * & * \\ 0 & 0 & 1 & * & * \\ 0 & 0 & 0 & 1 & * \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

(Для примера здесь $n = 5$.) На диагонали стоят единицы, под диагональю —нули, а над диагональю —любые элементы.

Алгебра Ли —это касательное пространство. Хотя наше множество G_n теперь не многообразие, и формального касательного пространства у него нет, но для любого алгебраического многообразия можно определить касательное пространство. В нашем случае касательное пространство состоит из таких же элементов, только с нулями на главной диагонали:

$$\mathfrak{g}_n \ni X = \begin{pmatrix} 0 & * & * & * & * \\ 0 & 0 & * & * & * \\ 0 & 0 & 0 & * & * \\ 0 & 0 & 0 & 0 & * \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Элементы двойственного пространства удобно записывать в виде нижнетреугольных матриц:

$$\mathfrak{g}_n^* \ni F = \begin{pmatrix} 0 & . & . & . & . \\ * & 0 & . & . & . \\ * & * & 0 & . & . \\ * & * & * & 0 & . \\ * & * & * & * & 0 \end{pmatrix}.$$

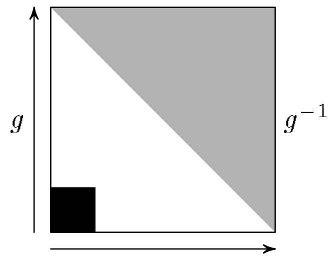
На диагонали стоят нули, под диагональю стоят какие-то элементы, а над диагональю мне не хочется ставить нули —я поставил точки. Смысл в этом вот какой. Если некоторое пространство является подпространством в пространстве матриц, то двойственное пространство будет факторпространством в двойственном пространстве матриц. Переход к двойственному пространству —это контравариантный функтор, который меняет подпространства и факторпространства. Пространство матриц само себе двойственно. След $\mathrm{tr}(F \cdot X)$ —билинейная функция, зависящая от матриц F и X. При фиксированном F она будет линейным функционалом от X, а при фиксированном X она будет линейным функционалом от F. При этом каждая линейная функция от матрицы X записывается в таком виде для подходящей матрицы F. Так что пространство $\mathrm{Mat}_n(\mathbb{F}_q)$ само себе двойственно. Более того, эта двойственность, что особенно важно, инвариантна относительно сопряжения: если заменить X и F на xXx^{-1} и xFx^{-1}, то результат не изменится. Поэтому двойственность выдерживает присоединенное действие на X и одновременно коприсоединенное действие на F. Но если X пробегает не все матрицы, а только часть —верхнетреугольные матрицы, тогда в роли F нужно рассматривать факторпространство всех матриц по тем матрицам, которые аннулируют верхнетреугольные матрицы. Точки над главной диагональю означают, что там стоят любые числа, на которые мы не обращаем внимания.

Теперь коприсоединенное действие можно записать следующим образом:

$$K(x)F = [xFx^{-1}]_{\text{нижняя часть}}.$$

Матрицу F при этом тоже можно себе представлять как нижнюю часть какой-то матрицы: нижняя часть преобразованной матрицы зависит только от нижней части исходной матрицы.

После этой предварительной подготовки я могу выписать явные формулы, как выглядят орбиты действия коприсоединенного представления. Матрицу F мы умножаем на верхнетреугольную матрицу g слева и на верхнетреугольную матрицу g^{-1} справа:



Как известно, умножение слева на треугольную матрицу g указанного вида сводится к тому, что последняя строка остается неизменной, к предпоследней строке добавляется последняя с каким-то коэффициентом, к предпредпоследней —линейная комбинация двух предыдущих строк, и т. д. Левая стрелочка показывает, что действие происходит снизу вверх. Точно так же, когда умножаем на матрицу g^{-1} справа, неизменным остается первый столбец. Ко второму столбцу добавляется первый, умноженный на какой-то коэффициент, и т. д. Нижняя стрелочка показывает, в какую сторону происходит действие. Видно, что влияние распространяется из левого нижнего угла вправо и вверх. В частности, сам левый нижний элемент остается неизменным. Это — пример так называемого инварианта. Точно так же минор, составленный из двух последних строк и двух первых столбцов, тоже остается неизменным. К его второму столбцу добавляется первый с каким-то коэффициентом, но определитель при этом не меняется. То же самое и для строк. Левые нижние миноры, которые я обозначу $\Delta_1, \Delta_2, \ldots, \Delta_{[n/2]}$, являются инвариантами: при коприсоединенном действии эти величины не меняются. Оказывается, что если мы приравняем эти инварианты константам, то система уравнений

$$\Delta_1 = c_1,\ \Delta_2 = c_2,\ \ldots,\ \Delta_{[n/2]} = c_{[n/2]}$$

задает инвариантное относительно коприсоединенного действия множество, которое при общих (типичных) значениях констант состоит ровно из одной орбиты. Что здесь означает «типичные»? Например, достаточно, чтобы все эти константы, кроме последней для нечетного n, были отличны от нуля. Если же какие-то константы обращаются в нуль, то эта система уравнений по-прежнему задает инвариантное множество, но оно может распадаться на более мелкие орбиты. Картина распадения большого инвариантного множества на мелкие орбиты усложняется по мере увеличения количества инвариантов, обращающихся в нуль, т. е. по мере того как падает размерность соответствующей коприсоединенной орбиты. Структура напоминает ветвящееся дерево. Имеется инвариантное множество общего положения, как правило, это просто одна орбита. Некоторые исключительные множества оказываются не орбитами, и там возникают дополнительные инварианты, которые являются инвариантами только на этом множестве: вообще говоря, они не инварианты, но на данном множестве инварианты. Происходит дальнейшее разбиение, где опять же типичные слои являются орбитами, а исключительные слои допускают дополнительные инварианты третьего уровня, и т. д. Этот ветвящийся процесс распространяется все дальше и дальше, пока он не доходит до орбит нулевой размерности, т. е. до точек. И пока еще не придуман адекватный аппарат для описания всего этого процесса.

Теперь я вступаю в новую область, которая связана с двумя вещами. Первая —это так называемая экспериментальная математика. Человек сидит и что-то считает и получает какой-то результат. Потом он еще что-то считает и получает другой результат. А когда получено несколько результатов, он думает, нельзя ли их объединить какой-нибудь теорией. Обычно успех зависит от того, сколько предварительных вычислений было сделано. В прошлом веке люди сидели годами, исписывали килограммы бумаги. Сейчас люди работают на компьютере, и не годами, а месяцами, потому что надо писать работы, надо участвовать в конкурсах на замещение должности, и т. д. Жизнь ускоряется, времени на обдумывание остается все меньше и меньше. Но тем не менее, экспериментальная математика процветает. Может быть, отчасти, потому, что есть люди, которые уже не озабочены своим положением и могут думать неограниченное количество времени. По крайней мере, Конвей и Кокстер —два замечательных математика, которые настолько прочно сидят на своих местах, что им не нужно думать о вещах насущных, и они могут думать о смысле жизни и о смысле математики. И они оба придумали в последнее время несколько совершенно неожиданных вещей, которые можно придумать, только если у тебя есть свободное время и ты можешь его тратить в неограниченном количестве. Последнее изобретение Кокстера очень пригодно для математической олимпиады.

Экспериментальная математика —это одна часть. Вторая —это наука, которая еще, может быть, даже и не существует как наука, но направление такое уже есть. Оно называется частично полностью интегрируемые системы. Интегрируемые системы —это такие механические системы, где решение, грубо говоря, можно выписать в явном виде. Если речь идет не о классических, а квантовых системах, то имеются в виду матрицы, собственные значения которых и собственные векторы можно выписать явно. Таких матриц немного, и обычно это связано с какой-то симметрией. Если есть симметрия, то это облегчает нахождение собственных значений и собственных векторов. Но лет 10 тому назад было обнаружено, что есть такие матрицы, для которых удается написать несколько первых собственных значений и несколько первых собственных векторов, а дальше не получается. И не потому не получается, что человек туп, а потому, что природа так устроена: первые значения хороши, а дальше наступает какой-то хаос, где уже сказать ничего нельзя.

Что-то похожее наблюдается и в той науке, о которой я рассказываю. Внешне это проявляется следующим образом. Применим экспериментальную математику, взяв самое простое поле из двух элементов и проделав тупые вычисления. Возьмем матрицы порядка 1, 2, 3, и т. д., и вычислим там количество присоединенных орбит. Получается последовательность, начало которой я сейчас выпишу. Чтобы она была убедительна, проделаем здесь вычисление до тех пределов, пока это легко сделать.

Начнем с матриц порядка 1. Группа Ли состоит и одной матрицы с элементом 1, алгебра Ли состоит из одной матрицы с элементом 0. Эта матрица и является единственной орбитой.

Для матриц порядка 2 алгебра Ли состоит из матриц

$$\begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}.$$

Двойственное пространство состоит из матриц

$$\begin{pmatrix} \cdot & \cdot \\ x & \cdot \end{pmatrix}.$$

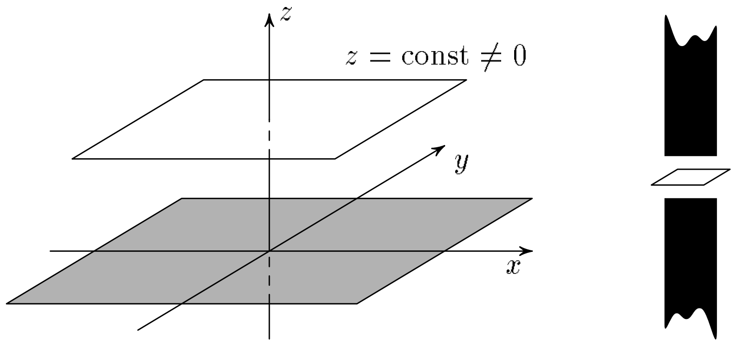
Над полем из двух элементов получаем 2 орбиты, потому что группа в данном случае коммутативна и ее действие тривиально, а само x принимает 2 значения.

Для третьего порядка нужно уже немножко считать. Алгебра Ли состоит из матриц

$$\begin{pmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{pmatrix}.$$

Двойственное пространство состоит из матриц

$$\begin{pmatrix} \cdot & \cdot & \cdot \\ x & \cdot & \cdot \\ z & y & \cdot \end{pmatrix}.$$



Преобразование выглядит так: $(x, y, z) \mapsto (x + az, y - bz, z)$.

Я нарисую, как выглядит множество орбит. Нас интересует поле из двух элементов, но картинка годится для любого поля, поэтому я нарисую картинку над полем вещественных чисел. Орбиты выглядят вот как: z является инвариантом (это тот самый инвариант Δ_1, который обсуждался в общем случае). Поэтому каждая плоскость $z = \mathrm{const}$ есть инвариантное множество. Если $z \neq 0$, то за счет подбора a и b мы можем сделать x и y произвольными. Это означает, что вся плоскость состоит из одной орбиты. Это соответствует общему геометрическому замечанию о том, что все орбиты должны быть четномерными. Если же $z = 0$, то все добавки тоже нулевые и все точки неподвижные. Это означает, что координатная плоскость $z = 0$ разбивается на орбиты, состоящие из отдельных точек. Размерность точки тоже четна.

Наглядно множество орбит можно изобразить так. Нарисуем толстую ось z, состоящую из толстых точек —они соответствуют орбитам. Одну точку $z = 0$ нужно вырезать и вместо нее вклеить целую плоскость, состоящую из тощих точек. То, что получится, и есть пространство орбит вместе с его естественной топологией: если мы стремимся к нулю по толстой оси z, то в качестве предела получаются все тощие точки. Такая топология типична для пространства орбит. Она же, кстати сказать, присутствует и в теории представлений. У множества неприводимых представлений тоже есть своя естественная топология, и оказывается, что она совпадает с топологией пространства орбит.

Теперь вернемся к конечному полю. Количество толстых орбит равно $q - 1$, а количество тощих орбит равно q^2. Всего будет $q - 1 + q^2$ орбит. Замечу, что мы получили многочлен от q. Это факт общий, для любого n количество орбит будет многочленом от q.

Продолжая такие вычисления, получаем табличку

n	-1	0	1	2	3	4	5
$O(n)$	1	1	1	2	5	16	61

Я дополнил табличку двумя начальными членами: числом орбит для $n = 0$ и для $n = -1$. Дело в том, что такая последовательность появляется и из совсем других соображений, и тогда она как раз начинается с трех единиц.

Теперь я сделаю отступление о частично точно решаемых задачах. Оказывается, что написанная выше последовательность ровно до того места, до которого она выписана, очень хорошо описывается другой последовательностью, весьма замечательной и встречающейся во многих других ситуациях. Неприятность заключается в том, что у хорошо известной последовательности на следующем месте должно стоять 272, а количество орбит при $n = 6$ равно 275. Формулу для последовательности K_n, о которой идет речь, несложно написать:

$$\sum_{n\geq 0} K_n \frac{x^n}{n!} = \operatorname{tg} x + \sec x.$$

Количество орбит начинает хорошо, но потом неожиданно отклоняется от этой хорошей последовательности.

В таком состояние дело находилось уже довольно давно, первый мой доклад на эту тему был уже года два тому назад. Тогда были известны эти две последовательности и было известно их расхождение в шестом члене. Оказывается, что если произвести более детальные исследования, а именно, рассмотреть не только поле из двух элементов, но и поле из q элементов, то можно придумать q-аналог последовательности K_n. Членами этой последовательности будут многочлены от q. Они тоже, как и числа K_n, задаются так называемым треугольником Эйлера-Бернулли. Он опять же имеется в «Сочинениях» Арнольда. Треугольник Эйлера-Бернулли имеет замечательный q-аналог, элементы которого являются многочленами от q. На самом деле там есть еще вспомогательный параметр t, поэтому получаются многочлены даже от двух переменных: они превращаются в многочлены от q при $t = 1$. Так вот, оказалось, что несовпадение, если рассматривать его над полем с произвольным q, делится на $(t - q)(t - q^2)$. В частности, если параметр t принимает исключительные значения q и q^2, то невязка пропадает.

Объясню теперь это всё подробно. Мы занимаемся разбиением множества $\mathfrak{g}_n^*(\mathbb{F}_q)$ на орбиты (здесь n —порядок треугольных матриц). Это множество является линейным пространством над полем $\mathbb{F}_q$, размерность которого равна числу ненулевых элементов в треугольных матрицах. Поэтому $|\mathfrak{g}_n^*(\mathbb{F}_q)| = q^{n(n-1)/2}$. Это конечное множество распадается на орбиты. Каждая орбита Ω четномерна (над конечным полем это имеет вполне определенный смысл: орбита Ω состоит из q^{2r} точек). Мы можем классифицировать орбиты по их размерности. Это довольно насильственная классификация, но какой-то смысл она имеет. Пусть $a_n^r(q)$ —число $2r$-мерных орбит в $\mathfrak{g}_n^*(\mathbb{F}_q)$. Для последовательности $a_n^r(q)$, зависящей от параметров n и r, естественно ввести производящую функцию

$$P_n(t,q) = \sum a_n^r(q)t^r.$$

Теперь у нас есть последовательность многочленов, которые зависят от двух переменных: t и q. Эта последовательность многочленов может быть включена в обобщенный треугольник Эйлера-Бернулли вплоть до $n = 5$. При $n = 6$ между ними получается разница, но эта разница делится на $(t - q)(t - q^2)$. При $t = q$ или q^2 эта разница пропадет. Для $t = q^2$ это не удивительно. Грубо говоря, t классифицирует орбиты по размерности, и когда мы полагаем $t = q^2$, это означает, что мы заменяет каждую орбиту размерности $2r$ группой из q^{2r} точек. Это —просто число точек в орбите, а общее число точек легче считать, чем число орбит: оно известно. Поэтому делимость на $t - q^2$ доказывается совсем просто. А вот когда $t = q$, мы получаем квадратный корень из числа точек на орбите. Согласно методу орбит это число имеет истолкование: оно равно размерности неприводимого представления. Значит, мы получаем не число орбит, а сумму размерностей неприводимых представлений. Эта сумма размерностей ведет себя уже лучше, чем число орбит. Например, она дальше совпадает с предсказаниями теории. Оказывается, что если положить $t = q$, то совпадение с теорией продолжается дальше, и проверено оно вплоть до $n = 11$. Это потребовало большой вычислительной работы, на которую я лично не способен, но я нашел себе соавтора, Аню Мельникову из института Вейцмана, которая смогла это сделать. Она проделала все эти вычисления и убедилась, что до $n = 11$ обе формулы совпадают. Мы уже было обрадовались, но пришло известие, что для $n = 13$ формула не может быть верна. Дело в том, что из нее вытекает некое следствие, которое я сейчас приведу, а это следствие опровергается неким примером.

Это не совсем противоречащий пример —это противоречащий пример к гораздо более общему утверждению, я бы даже сказал более нахальному утверждению, которое состоит в том, что метод орбит работает *буквально* для рассматриваемой последовательности групп. Тогда, в частности, из метода орбит вытекает явная формула для характеров неприводимых представлений. А если эта формула верна, то все характеры неприводимых представлений вещественны (для поля из двух элементов). А уже из этого, как знает каждый специалист по теории групп, вытекает, что любой элемент группы сопряжен своему обратному: $g \sim g^{-1}$. Классы сопряженности различаются характерами, а характер на элементе g —это сопряженное значение характера на элементе g^{-1}. Для вещественных чисел они совпадают.

Возникает задача, которую можно сформулировать для первокурсника. Имеется треугольная матрица с элементами из $\mathbb{F}_2$. Спрашивается: сопряжена ли она своей обратной матрице внутри треугольной группы или не сопряжена? Если бы речь шла обо всей группе матриц, то ответ положителен. Дело в том, что треугольная матрица приводится к набору жордановых блоков, а жордановы блоки у самой матрицы и у ее обратной одни и те же. Но внутри треугольной группы ответ не такой очевидный. Для матриц порядка 13 с помощью компьютера показано, что имеется ровно один противоречащий пример. Поэтому неверна та формула для характеров, которая получается совсем тупым применением метода орбит. Это жаль, потому что если бы такое применение было верно, то можно было бы доказать совпадение всех многочленов.

На самом деле, я немножко поспешил. Из того, что имеется противоречащий пример, еще не вытекает, что при $t = q$ многочлены различны. Верно лишь то, что то доказательство, которое я имел в виду, не годится. Сами же многочлены настолько красивы, что, может быть, они и продолжают годиться.

Часть 2

Эта лекция хотя и будет продолжением первой, но ее можно слушать независимо. Я надеюсь рассказать более конкретные вещи. Я хочу рассказать о задачах, связанных с представлениями конечных групп, которые возникают из метода орбит. Хотя сам метод орбит относится к бесконечным (и даже бесконечномерным) группам, оказывается, что правильное его истолкование применимо и к конечным группам тоже.

Начну я с формулировки некоторой гипотезы, для понимания которой не нужно вообще ничего. Это будет в стиле тех занятий с молодыми математиками, которые ведутся с молодыми математиками здесь в Независимом университете, и будет также продолжением традиции математических кружков. Я буду рассказывать про некоторую замечательную последовательность многочленов. Я уверен, что она замечательная, и одним из оправданий этого является то, что эта последовательность имеет по крайней мере 5 разных определений, про которые не доказано, что эти определения совпадают.

Мы будем рассматривать последовательности многочленов, которые обозначаются $A_n(q)$, $B_n(q)$, $C_n(q)$, $D_n(q)$ и $Y_n(q)$. Эти многочлены не имеют отношения к простым алгебрам Ли серий A, B, C, D. Тем, кто был в прошлый раз, я напомню, что мы рассматривали теорию представлений. Для конечной группы G одной из теоретико-представленческих характеристик является набор размерностей неприводимых комплексных представлений этой группы. Сам набор неприводимых комплексных представлений группы G, рассматриваемых с точностью до эквивалентности, я буду обозначать $\widehat{G}$. Элемент множества $\widehat{G}$ я буду обозначать λ, а то представление, класс эквивалентности которого обозначается λ, будет обозначаться π_λ. Таким образом, π_λ —это гомоморфизм $G \to \mathrm{GL}(d(\lambda), \mathbb{C})$. Группа конечна, поэтому ее любое конечномерное представление эквивалентно унитарному представлению, и можно считать, что $\pi_\lambda\colon G \to U(d(\lambda)) \subset \mathrm{GL}(d(\lambda), \mathbb{C})$.

Набор целых положительных чисел $d(\lambda)$ обладает несколькими замечательными свойствами. Например, все они являются делителями порядка группы. Кроме того, имеется знаменитое тождество Бернсайда

$$\sum d^2(\lambda) = \#G.$$

(Здесь $\#G$ обозначает число элементов группы G). А если взять сумму нулевых степеней, то получим

$$\sum d^0(\lambda) = \#\widehat{G}.$$

Хотелось бы эти два тождества проинтерполировать и вставить между ними сумму самих $d(\lambda)$: $\sum d(\lambda) = ?$ В действительности не известно, обладает ли эта сумма какими-либо хорошими свойствами. Но есть частный случай, когда эта сумма все-таки обладает хорошими свойствами. Это тот случай, когда все представления (которые априори комплексные) являются вещественными.

Я сделаю отступление и напомню тем, кто знает, и сообщу тем, кто не знает, что комплексные неприводимые представления группы разбиваются на три типа:

вещественный тип: представление π_λ эквивалентно вещественному, т. е. при подходящем выборе базиса все операторы представления записываются матрицами с вещественными коэффициентами;

комплексный тип: комплексно сопряженное представление $\overline{\pi}_\lambda$ не эквивалентно представлению π_λ (для каждого представления можно построить комплексно сопряженное представление, выбрав базис и заменив в каждой матрице представления каждый элемент на комплексно сопряженный: комплексное сопряжение является автоморфизмом комплексного поля, поэтому все тождества сохраняются и представление остается представлением);

кватернионный тип: представление π_λ не эквивалентно вещественному представлению, но $\overline{\pi}_\lambda \sim \pi_\lambda$.

Кватернионный тип, можно сказать, самый интересный тип представлений. Он возникает тогда, когда комплексная размерность представления четна и это четномерное комплексное пространство получается сужением поля скаляров из кватернионного пространства вдвое меньшей размерности. При этом сами операторы представления можно записывать кватернионными матрицами.

Самый известный пример —группа $SU(2)$, которая сама по себе является группой кватернионов, по модулю равных 1. Эта группа бесконечная, но у нее имеется много конечных подгрупп. Тавтологическое представление группы $SU(2)$ как одномерных кватернионных матриц дает пример кватернионного представления.

Как известно, каждый кватернион можно изобразить матрицей порядка 2 с комплексными элементами, поэтому каждое кватернионное представление размерности n можно при желании рассматривать как комплексное представление размерности $2n$. Это будет представление кватернионного типа.

Назовем индексом представления λ число

$$\operatorname{ind}\lambda = \begin{cases} 1, & \text{если представление } \lambda \text{ вещественного типа;} \\ 0, & \text{если представление } \lambda \text{ комплексного типа;} \\ -1, & \text{если представление } \lambda \text{ кватернионного типа.} \end{cases}$$

Существует замечательная формула Германа Вейля, позволяющая этот индекс сосчитать:

$$\operatorname{ind}\lambda = \frac{1}{\#G}\sum_{g\in G} \operatorname{tr}\pi_\lambda(g^2).$$

В качестве задачи я предлагаю доказать это равенство самостоятельно. Сделать это непросто. Даже не все специалисты по теории представлений, которых я спрашивал, могли это равенство сразу доказать, если они его раньше не знали. Поэтому здесь необходимо указание. Я не знаю, существует ли много разных доказательств. То доказательство, которое я сам придумал, мне очень нравится, поэтому я хочу его пропагандировать.

Рассмотрим пространство $\mathbb{C}[G]$ всех комплекснозначных функций на нашей группе. Для любой конечной группы существует преобразование Фурье, которое отображает это пространство в пространство матричных функций на двойственном объекте, на $\widehat{G}$. Обозначим это пространство $\mathbb{C}\mathrm{Mat}[\widehat{G}]$. Это соответствие взаимно однозначное, как и полагается преобразованию Фурье. А точная его запись такова: функции $f \in \mathbb{C}[G]$ соответствует функция

$$\tilde{f}(\lambda) = \sum_{g \in G} f(g)\pi_\lambda(g).$$

Для каждого класса λ представление π_λ сопоставляет элементу g матрицу $\pi_\lambda(g)$ порядка $d(\lambda)$. Можно проверить, что если взять не конечную группу, а окружность или прямую, то эта процедура дает, соответственно, ряд Фурье или интеграл Фурье.

Из взаимной однозначности преобразования Фурье следует, в частности, совпадение размерностей пространств $\mathbb{C}[G]$ и $\mathbb{C}\mathrm{Mat}[\widehat{G}]$. Это, кстати сказать, и есть тождество Бернсайда. Преобразование Фурье является не только изоморфизмом пространств, но и изоморфизмом алгебр: свертка функций переходит в произведение матриц.

Рассмотрим в пространстве $\mathbb{C}[G]$ линейный оператор V, который переводит функцию f в функцию $\check{f}$, где по определению $\check{f}(g) = f(g^{-1})$. Вычислим след этого оператора двумя разными способами. А именно, можно вычислить след оператора в исходном пространстве $\mathbb{C}[G]$, выбрав там естественный базис $\{\chi_g\}$: функция χ_g равна 1 на элементе g и равна 0 на всех остальных элементах. След —это сумма диагональных элементов, поэтому он равен количеству тех элементов группы, которые равны своему обратному элементу: только эти базисные элементы переходят сами в себя, а все прочие базисные элементы переходят в другие базисные элементы и дают только недиагональные элементы матрицы. Равенство $g = g^{-1}$ можно записать в виде $g^2 = 1$. Преобразования, квадрат которых тождествен, называют инволюциями. Итак, если $\mathrm{Inv}(G)$ —множество всех инволюций в группе G, то $\mathrm{tr}\, V = \# \,\mathrm{Inv}(G)$.

По-другому вычислить след оператора V можно, сделав преобразование Фурье. Тогда элементами пространства будут уже не функции на группе, а наборы матриц. Количество этих матриц равно количеству неприводимых представлений, а их порядки равны размерностям неприводимых представлений. Функции f соответствует набор матриц $\tilde{f}(\lambda_1), \ldots, \tilde{f}(\lambda_n)$ порядков $d_1 = d(\lambda_1), \ldots, d_k$. Выясним, как оператор V действует на эти матрицы.

Если представление вещественного типа, то оператор V переводит матрицу в транспонированную матрицу: $A \mapsto A^t = {}^T A$. Легко указать, какой вклад в след этого оператора дает каждое вещественное представление. Недиагональные элементы переходят в другие недиагональные элементы и не дают вклада, а диагональные элементы дают вклад, равный 1. Поэтому каждое вещественное представление дает вклад, равный своей размерности.

С представлениями комплексного типа разобраться еще проще: они не дают вообще никакого вклада. Представления $\overline{\pi}_\lambda$ и $\pi_{\overline{\lambda}}$ не эквивалентны. Преобразование V просто переставляет значения в разных точках: представления λ и $\overline{\lambda}$ не эквивалентны, это две разные точки пространства $\widehat{G}$. Преобразование V переставляет совершенно разные элементы матрицы, и это не дает никакого вклада в след.

С представлениями кватернионного типа разобраться сложнее. Это нужно считать, нужно знать, как кватернионы после комплексификации дают матрицы второго порядка, и что с ними происходит. Я скажу без доказательства результат. Каждое комплексное представление кватернионного типа дает вклад, равный своей размерности со знаком минус.

Сравнивая всё это с определением индекса, мы видим, что в целом получается следующий результат:

$$\#\operatorname{Inv}(G) = \operatorname{tr}(V) = \sum_{\lambda \in \widehat{G}} \operatorname{ind}(\lambda) d(\lambda).$$

Число инволюций в группе выражается в виде знакопеременной суммы размерностей неприводимых представлений. Особенно просто эта формула выглядит в случае, когда все представления группы вещественны. Тогда индекс тождественно равен 1, и получается просто сумма размерностей неприводимых представлений. Таким образом, если все неприводимые представления вещественны, то

$$\#\operatorname{Inv}(G) = \sum_{\lambda \in \widehat{G}} d(\lambda).$$

Всё это было присказкой к определению первой последовательности многочленов. Сегодняшняя моя лекция будет состоять из определений пяти последовательностей многочленов и формулировки гипотезы о том, что эти пять последовательностей совпадают. Я уже почти готов к тому, чтобы определить первую последовательность многочленов, но все-таки почти. Нужны еще некоторые сведения о конечных полях. На прошлой лекции я ввел обозначение $\mathbb{F}_q$ для конечного поля из q элементов, где $q = p^k$ и p —простое число. В дальнейшем мы будем рассматривать многочлены от q. И довольно странно было бы, что буква, изображающая независимую переменную, могла бы принимать только отдельные выбранные значения. Есть разные попытки интерпретировать значения этих многочленов для $q \neq p^k$. Но это задача не только не решенная, но даже и не сформулированная. Поэтому я пока про нее говорить не буду. Хотя, как отмечал Арнольд в своей книжке, самые интересные задачи —это те задачи, которые еще даже и не сформулированы. Вот это —пример не сформулированной задачи.

Рассмотрим группу $G_n(\mathbb{F}_q)$, состоящую из верхнетреугольных матриц g порядка n, у которых на диагонали стоят единицы, а над диагональю стоят элементы поля $\mathbb{F}_q$. Порядок этой группы равен $q^{n(n-1)/2}$, потому что каждый из $n(n-1)/2$ наддиагональных элементов может принимать ровно q различных значений.

Вообще говоря, для этой группы возможны представления всех трех типов —вещественные, комплексные и кватернионные. Но если $q = 2^l$, то при малых n все представления этой группы вещественны, а при $n = 13$ есть пример не вещественного представления. В первом приближении мы закроем глаза на этот факт и будем считать, что все представления вещественны (в случае, когда $q = 2^l$). Тогда равенство

$$\#\operatorname{Inv}(G) = \sum_{\lambda \in \widehat{G}} d(\lambda)$$

дает выражение для числа инволюций в группе через размерности неприводимых представлений. А для поля из четного числа элементов инволюции очень легко вычислить. А именно, запишем матрицу g в виде $g = 1_n + X$, где 1_n —единичная матрица порядка n. Тогда $g^2 = 1_n + 2X + X^2 = 1_n + X^2$, так как в поле из четного числа элементов $2X = 0$. Поэтому уравнение $g^2 = 1_n$ эквивалентно уравнению $X^2 = 0$.

Поставим такую задачу: сколько существует верхнетреугольных матриц X с элементами из поля $\mathbb{F}_q$, для которых $X^2 = 0$.

Теперь можно дать определение первой последовательности многочленов. Определим $A_n(q)$ как число решений уравнения $X^2 = 0$ в верхнетреугольных матрицах порядка n с элементами из поля $\mathbb{F}_q$. Здесь имеются в виду не только четные q, а и все другие q, для которых существует поле из q элементов.

Задача. Доказать, что $A_n(q)$ —многочлен от q.

Это —хорошая комбинаторная задача. А именно, у вас есть некоторый конечный объект, и нужно вычислить количество точек в этом объекте. Это и есть основная задача комбинаторики. Комбинаторика в этом смысле не является вполне наукой, потому что для каждой задачи изобретается своя собственная теория. Но в этом смысле многие другие науки тоже науками не являются. Когда-то Юрий Иванович Манин объяснял мне, что алгебраическая геометрия не является наукой, потому что она состоит из набора решенных задач, а каждая задача требует своего собственного метода. Есть, конечно, какие-то общие понятия и общие приемы, но, как правило, они действуют в двух или самое лучшее в трех случаях, а так, чтобы больше, то такое редко бывает. Наиболее интересные результаты в алгебраической геометрии относятся к индивидуальным приемам. Вот так и комбинаторика. Она тоже умеет решать несколько конкретных задач. Слово «несколько» может здесь означать несколько тысяч, но тем самым имеется несколько тысяч конкретных задач, которые умеет решать комбинаторика.

Как же решить нашу конкретную комбинаторную задачу? Естественно попытаться получить рекуррентное выражение $A_{n+1}(q)$ через $A_n(q)$. Но я этого делать не умею, и, насколько мне известно, этого никто пока делать не умеет. Тогда остается один из немногих стандартных приемов, а именно, давайте усложним задачу, тогда решить ее будет легче. Разобьем множество всех решений на типы по рангу матрицы:

$$A_n^r(q) = \{X \in A_n(q) \mid \operatorname{rk} X = r\}.$$

Казалось бы, задача усложнилась, потому что вместо одного числа мы должны вычислить сразу много чисел. С другой стороны, как и во многих других случаях, такое усложнение приводит к упрощению, потому что для чисел $A_n^r(q)$ уже существует рекуррентное соотношение.

Это рекуррентное соотношение получается следующим образом. Рассмотрим матрицу $\mathcal{X} = \begin{pmatrix} X & x \\ 0 & 0 \end{pmatrix}$. Ясно, что $\mathcal{X}^2 = \begin{pmatrix} X^2 & Xx \\ 0 & 0 \end{pmatrix}$. Поэтому $\mathcal{X}^2 = 0$ тогда и только тогда, когда $X^2 = 0$ и $Xx = 0$. Число решений уравнения $X^2 = 0$ мы, по предположению, уже знаем: оно было получено на предыдущем шаге. А число решений уравнения $Xx = 0$ зависит от ранга матрицы X. Поэтому в общем случае число решений этого уравнения не известно. Но если известно, что $\operatorname{rk} X = r$, то число решений тоже известно. Более того, зная кое-что из линейной алгебры, нетрудно догадаться, каким будет ранг матрицы $\mathcal{X}$. Он может либо не измениться, либо увеличиться на 1. Что именно происходит, зависит от соотношения между X и x, которое легко контролируется. На этом я заканчиваю свои объяснения и выписываю рекуррентное соотношение:

$$A_{n+1}^{r+1}(q) = q^{r+1} A_n^{r+1}(q) + (q^{n-r} - q^r) A_n^r(q).$$

Теперь можно забыть о том, что q принимает лишь весьма специальные значения, и пользоваться рекуррентным соотношением. Тогда оно определяет многочлены A_n^r, которые можно по очереди вычислять, если знать самый первый многочлен A_0^0. А он, естественно, равен 1. Получается таблица многочленов совершенно общего вида, у которых все коэффициенты отличны от нуля и про которые ничего хорошего сказать нельзя. Но если рассмотреть $\sum\limits_r A_n^r(q) = A_n(q)$, то происходит магическое сокращение, почти все члены взаимно уничтожаются. В качестве доказательства я нарисую таблицу многочленов A_n:

$$\begin{aligned}
A_0 &= 1,\\
A_1 &= 1,\\
A_2 &= q,\\
A_3 &= 2q^2 - q,\\
A_4 &= 2q^4 - q^2,\\
A_5 &= 5q^6 - 4q^5,\\
A_6 &= 5q^9 - 5q^7 + q^5,\\
A_7 &= 14q^{12} - 14q^{11} + q^7,\\
A_8 &= 14q^{16} - 20q^{14} + 7q^{12},\\
A_9 &= 42q^{20} - 48q^{19} + 8q^{15} - q^{12},\\
A_{10} &= 42q^{25} - 75q^{23} + 35q^{21} - q^{15},\\
A_{11} &= 132q^{30} - 165q^{29} + 44q^{25} - 10q^{22}.
\end{aligned}$$

Я напомню смысл $A_2(q)$ —это число решений уравнения $X^2 = 0$ в верхнетреугольных матрицах второго порядка. Равенство $X^2 = 0$ выполняется для всех верхнетреугольных матриц второго порядка (с нулями на диагонали). Это и соответствует тому, что $A_2(q) = q$.

В табличке сначала идут 3 многочлена с одним мономом, затем идут 3 многочлена с двумя мономами, затем 3 многочлена с тремя мономами, затем 3 многочлена с четырьмя мономами. Примерно два дня, или даже три, я смотрел на эту таблицу многочленов (она была выписана до размерности 20) и нашел в ней много свойств, которые позволяли продолжать эту таблицу неограниченно. И только потом догадался, какой правильный ответ.

Старшие коэффициенты многочленов —числа Каталана. Все знают, как записывается треугольник Паскаля. Давайте сделаем такой же треугольник, но только поставим зеркало и запретим за это зеркало заходить:

$$
\begin{array}{ccccccc}
1 & & & & & & \\
 & 1 & & & & & \\
1 & & 1 & & & & \\
 & 2 & & 1 & & & \\
2 & & 3 & & 1 & & \\
 & 5 & & 4 & & 1 & \\
5 & & 9 & & 5 & & 1
\end{array}
$$

В первом столбце стоят как раз числа Каталана. Но зеркало мешает. Хорошо бы его убрать, а закон сохранить. Как хорошо известно из курса физики, зеркало заменяется отражением. В зеркальном мире еще нужно поменять знаки. Тогда получается треугольник, который образуется в точности по тому же правилу, что и треугольник Паскаля. На самом зеркале всюду стоят нули:

$$
\begin{array}{ccccccccccccccc}
 & & & & & & -1 & & 1 & & & & & & \\
 & & & & & -1 & & 0 & & 1 & & & & & \\
 & & & & -1 & & -1 & & 1 & & 1 & & & & \\
 & & & -1 & & -2 & & 0 & & 2 & & 1 & & & \\
 & & -1 & & -3 & & -2 & & 2 & & 3 & & 1 & & \\
 & -1 & & -4 & & -5 & & 0 & & 5 & & 4 & & 1 & \\
-1 & & -5 & & -9 & & -5 & & 5 & & 9 & & 5 & & 1
\end{array}
$$

Легко сообразить, что это просто разность двух треугольников Паскаля: один треугольник Паскаля растет из 1, а другой —из -1.

Теперь я предлагаю вам продолжить треугольник Каталана и сравнить его с табличкой многочленов. Это прекрасный пример экспериментальной математики, когда можно обнаружить формулу. Но это еще не значит, что ее можно доказать. Доказана она была сравнительно недавно, и это потребовало усилий двух крупных комбинаториков. Они использовали для этого разработанную ими ранее теорию. Сейчас имеется книжка под названием «$A = B$». Содержание этой книжки состоит в построении некоего алгоритма: на входе пишется формула $A = B$, а на выходе пишется доказательство этой формулы при условии, что формула верная. Один из авторов этой книги является моим партнером по университету в Филадельфии. Когда я ему сообщил, что у меня есть формула, которую я не могу доказать, он ответил, что у нас есть книжка, в которой доказана эта формула. Я предложил им засунуть мою формулу в их книгу. Они засунули. Алгоритм сразу не сработал, но они поднатужились и через некоторое время довели доказательство до конца.

Если считать, что в первой строке треугольника Каталана стоят числа $c_{1,-1} = -1$ и $c_{-1,1} = 1$, то для чисел $c_{k,l}$ можно написать рекуррентную формулу

$$c_{k,l} = c_{k-1,l-1} + c_{k-1,l+1}.$$

Чтобы рисовать треугольную таблицу, нужно иметь два индекса, причем хорошо бы, чтобы эти индексы имели одинаковую четность.

Введем теперь новую последовательность многочленов, которую я в честь Каталана обозначаю C_n:

$$C_n(q) = \sum_{s \equiv n+1 (2) \equiv (-1)^n (3)} c_{n+1,s}\, q^{\frac{n^2}{4} + \frac{1-s^2}{12}}.$$

Это —экспериментальное правило, какие числа из треугольника Каталана нужно взять, чтобы получить коэффициенты полиномов A_n.

Теорема 1. $A_n(q) = C_n(q)$.

Эта теорема доказана с помощью книжки «$A = B$».

Мы уже выполнили значительную часть программы —определили две последовательности многочленов. Про многочлены C_n мы знаем явную формулу, но не знаем никакой другой интерпретации. Про многочлены A_n мы знаем интерпретацию: если q —степень двойки, то $A_n(q)$ —это число решений уравнения $X^2 = 0$ в верхнетреугольных матрицах над полем $\mathbb{F}_q$. Но мы не знаем для этих многочленов явной формулы. И вот теперь есть теорема, которая дает и явную формулу и интерпретацию. Я, кстати сказать, не очень удовлетворен доказательством этой теоремы, хотя бы потому, что я не могу ни слова понять в этом доказательстве. Поэтому если кто-нибудь придумает более прозрачное доказательство, то это будет вполне содержательный шаг в науке, и мы с удовольствием это где-нибудь опубликуем.

Наряду с треугольником Паскаля есть гораздо более хитрый треугольник, о котором я лично узнал из лекции Арнольда, прочитанной

4 года назад. А из новой книжки Арнольда я узнал, что этот треугольник известен уже более ста лет. Но кому известно и где известно, не сказано. Косвенно можно сделать вывод, что не Арнольд это придумал, потому что ему ста лет еще нет. Это —так называемый треугольник Эйлера-Бернулли. Он строится способом, похожим на построение треугольника Паскаля, но с помощью челночного движения. Правило такое: нужно двигаться по нечетным строкам слева направо, по четным строкам справа налево, и каждый раз нужно начинать с нуля. А во всем остальном нужно следовать правилам треугольника Паскаля.

Нулевая строка состоит из одного элемента 1. Первая строка состоит из двух элементов 0 и 1. Вторую строку начинаем уже не слева, а справа. Ставим 0 и движемся влево: берем сумму предыдущего числа той же строки и предыдущего числа строки сверху, т. е. числа строго справа и числа справа вверху. Затем начинаем строку слева. Сначала всегда ставим 0, а затем берем суммы пар полученных ранее чисел:

$$
\begin{array}{ccccccccccccccc}
 & & & & & & 1 & & & & & & & & \\
 & & & & & 0 & & 1 & & & & & & & \\
 & & & & 1 & & 1 & & 0 & & & & & & \\
 & & & 0 & & 1 & & 2 & & 2 & & & & & \\
 & & 5 & & 5 & & 4 & & 2 & & 0 & & & & \\
 & 0 & & 5 & & 10 & & 14 & & 16 & & 16 & & & \\
61 & & 61 & & 56 & & 46 & & 32 & & 16 & & 0 & & \\
 & & & & & & & & & & & & & 272 &
\end{array}
$$

В левой части треугольника через раз по определению идут нули. А оставшиеся числа $(1, 1, 5, 61, \ldots)$ известны в науке. Они называются числами Эйлера. В любом справочнике по специальным функциям есть таблица чисел Эйлера. Они обозначаются E_n и определяются соотношением

$$\sum_{n\geq 0} E_n \frac{x^n}{n!} = \frac{1}{\cos x}.$$

Кажется, первоначальное определение Эйлера было как раз такое.

Если взять правую диагональ, то мы получим другие числа, а именно, нормированные числа Бернулли. Сами числа Бернулли, как известно, рациональные. А у нас получаются так называемые нормированные числа Бернулли, которые являются целыми. Нормированное число Бернулли $\widetilde{B}_n$ получается из обычного числа Бернулли B_n по формуле

$\widetilde{B}_n = B_n \frac{2^{2n}(2^{2n}-1)}{2n}$. Нормированные числа Бернулли определяются соотношением

$$\sum_{n>0} \widetilde{B}_n \frac{x^n}{n!} = \operatorname{tg} x.$$

Это факт, который обычно скрывают от студентов мехмата, что не только синус и косинус имеют простые разложения в ряд Тэйлора, но и $1/\cos x$ и тангенс тоже имеют разложения в ряд Тэйлора, которые можно описать в явном виде (через числа Эйлера и через числа Бернулли).

Отдельно числа Эйлера и числа Бернулли считать трудно, а весь треугольник вместе считать легко.

Можно написать производящую функцию для всего треугольника. Это тоже хорошее упражнение.

Тот факт, который я в эту науку внес, состоит в том, что правильное понимание треугольника Эйлера-Бернулли состоит в том, что нужно заменить числа, туда входящие, многочленами. Более того, многочленами от двух переменных q и t. Заменять числа многочленами —хорошо известный принцип математики. Его иногда называют принципом q-аналога. Многие математические величины, которые первоначально кажутся целыми числами, на самом деле нужно интерпретировать как многочлены. Один из способов это объяснить такой. Целое число в математике встречается чаще всего как размерность некоторого пространства. С другой стороны, пространства часто бывают градуированными, т. е. представляются в виде суммы подпространств, занумерованных целыми числами. Тогда, чтобы учесть градуировку, нужно писать не глобальную размерность всего пространства, а писать отдельно размерность каждой однородной компоненты. А это уже можно кодировать многочленами: свободный член означает размерность пространства с градуировкой 0, коэффициент при q —размерность пространства с градуировкой 1, и т. д. Поэтому каждый раз, когда у вас есть градуированное пространство, вместо его размерности можно рассматривать градуированную размерность, т. е. многочлен.

Хорошо известный пример —это $n!$. Правильной интерпретацией числа $n!$ является многочлен $\prod\limits_{k=1}^{n} \dfrac{1-q^k}{1-q}$. Степень этого многочлена равна $n(n-1)/2$. При $q=1$ этот многочлен равен $n!$. (Если рассматривать дроби $\dfrac{1-q^k}{1-q}$, то нужно вычислять предел при q стремящемся к 1.)

В отличие от самого $n!$ выражение $\prod_{k=1}^{n} \frac{1-q^k}{1-q}$ имеет глубокий хорошо известный геометрический смысл —это размерность группы гомологий пространства

флагов. Группы гомологий градуированы по размерности, поэтому вместо числа возникает многочлен, который задается этой формулой. Пространство флагов —это одно из самых замечательных многообразий, которые сейчас имеются. Про него очень много всего известно. Это в какой-то степени объясняет, почему $n!$ встречается в разных формулах. Почему, например, x^n любит, чтобы его делили на $(n!)$? Потому, что пространство флагов существует. Я не буду сейчас это объяснять. Есть масса книг, в которых это объясняется.

Я объясню еще один пример замены чисел на многочлены. Давайте возьмем другой пример —число сочетаний

$$C_n^k = \binom{n}{k} = \frac{n!}{k!(n-k)!}.$$

Это выражение имеет q-аналог. Например, каждый факториал можно заменить приведенным выше многочленом. На самом деле q-аналог биномиального коэффициента был придуман раньше, чем q-аналог числа $n!$. Это было придумано Гауссом следующим образом. Что такое число сочетаний из n по k? Это число k-точечных подмножеств в n-точечном множестве. Но это объект плохой —конечные множества никакой структуры не имеют. Это слишком слабый объект, чтобы им занимались математики. Давайте возьмем объект поинтереснее, например, линейные пространства. Но если взять вещественные или комплексные пространства, то объект будет бесконечным, сосчитать число точек будет невозможно. Выход такой —возьмем пространство над конечным полем. Тогда в нем будет уже конечное число точек. Что является правильным аналогом q-точечного подмножества в n-точечном множестве? Этим аналогом будет q-мерное подпространство в n-мерном пространстве над полем $\mathbb{F}_q$. Число q-мерных подпространств в n-мерном пространстве над полем $\mathbb{F}_q$ является многочленом от q. При $q = 1$ этот многочлен равен биномиальному коэффициенту. В пределе можно понимать так: n-мерное линейное пространство над полем из одного элемента —это просто n точек. Это одна из возможных интерпретаций предела $q \to 1$, но далеко не единственная.

Существует еще по крайней мере два или три разных способа заменять число на многочлен. Этот многочлен по традиции считается многочленом от q. Но в разных подходах буква q играет совершенно разную роль. Только что q было числом элементов конечного поля. В примере с пространством флагов число q было размерностью группы гомологий. Бывают и другие ситуации. Иногда q —это корень из единицы определенной степени. Иногда —малый параметр. Пожалуй, самый замечательный факт состоит в том, что в разных подходах возникают одни и те же выражения. Вот это уже область экспериментальной математики. Я объясняю это тем, что число разумных выражений невелико, а природа необъятна. Поэтому она обязана пользоваться одним и тем же выражением много раз, просто потому, что не хватает разумных выражений для всех природных явлений.

После всего этого отступления я хочу сказать, каким образом нужно превратить треугольник Эйлера-Бернулли в треугольник, элементами которого являются многочлены от q. Аккуратное определение требует многочлена от двух переменных q и t. Но из тех соображений, о которых я кратко сказал на предыдущей лекции, нужно положить $t = q$ и тогда получается многочлен от одной переменной. Сейчас я напишу правило, по которому строится q-аналог треугольника Эйлера-Бернулли. Небольшая модификация правила образования треугольника Эйлера-Бернулли приводит к правилу

$$b_{k,l} = q^{-1} b_{k-1,l+1} + (q^{l+1} - q^l) b_{l,k-1}$$

для $k > 0$. Кроме того, $b_{0,l} = q^l b_{l-1,0}$ и $b_{0,0} = 1$.

Набор этих правил позволяет восстановить весь треугольник. Обратите внимание, что в рекуррентной формуле у последнего члена индексы k и l поменялись местами. Это —следствие челночного прохода: каждое число зависит от предыдущих чисел либо слева, либо справа. От этого k и l меняются местами.

Здесь я хочу сформулировать две нерешенные задачи.

1. Рекуррентное соотношение есть, треугольник можно написать, но формула для него не известна.

2. Не известна даже асимптотика роста членов треугольника.

Вторая задача, пожалуй, наиболее интересна: насколько быстро растут члены треугольника при фиксированном q.

Третья последовательность многочленов определяется так: $B_n(q) = b_{n-1,0}(q)$.

Теперь из пяти последовательностей многочленов мы определили три. Остались самый интересные многочлены D_n и Y_n. Чтобы заинтриговать слушателей, я сразу напишу для D_n формулу:

$$D_n(q) = \zeta_{G_n(\mathbb{F}_q)}(-1), \quad \text{где } q = 2^l.$$

Здесь $G_n(\mathbb{F}_q)$ —группа строго верхнетреугольных матриц порядка n с элементами из $\mathbb{F}_q$.

Я напомню определение классической дзета-функции Римана: $\zeta(s) = \sum\limits_{n\geq 1} n^{-s}$. Эта функция очень знаменита. Последняя нерешенная великая задача в математике —нахождение нулей дзета-функции. По известной гипотезе Римана они лежат на прямой $\Re s = 1/2$. Поэтому многие люди пытались обобщить дзета-функцию, чтобы примазаться к славе Римана. Специалисту по теории представлений обобщить дзета-функцию ничего не стоит, потому что натуральные числа —это размерности неприводимых представлений простейшей некоммутативной компактной группы, а именно $SU(2)$. Эта группа имеет ровно одно неприводимое представление в каждой размерности (с точностью до эквивалентности). Поэтому если у вас имеется любая группа G, у которой неприводимые представления конечномерные (например, конечная группа или компактная группа), то вы можете определить дзета-функцию, связанную с этой группой:

$$\zeta_G(s) = \sum_{\lambda\in\widehat{G}} d(\lambda)^{-s}.$$

Как ни странно, эта функция до сих пор не привлекала к себе большого внимания. Может быть, из-за того, что здесь содержательные примеры либо сводятся к обычной дзета-функции, либо про них нельзя сказать ничего содержательного. Хотя, конечно, некоторые известные тождества можно переписать как тождества для значений дзета-функции. Например, тождество Бернсайда, которое я выписывал, означает, что для конечной группы G выполняется равенство

$$\zeta_G(-2) = \#G.$$

В начале лекции я обсуждал, что для конечных групп еще имеет смысл

$$\zeta_G(0) = \#\widehat{G}.$$

Кроме того, иногда имеет смысл

$$\zeta_G(-1) = \sum d(\lambda).$$

А именно, в некоторых случаях выполняется равенство $\zeta_G(-1) = \#\operatorname{Inv} G$. В определении многочлена D_n участвует как раз $\zeta_G(-1)$.

При нашем определении совершенно не ясно, почему D_n —многочлен.

Последнее, что я хочу сказать, это определение пятой последовательности многочленов Y_n. Эти многочлены связаны с описанием коприсоединенных орбит для группы верхнетреугольных матриц над конечным полем. Что такое коприсоединенная орбита я объяснял на прошлой лекции. Это некоторая траектория, описываемая точкой под действием группы. Если бы поле было вещественным, то эти траектории были бы довольно простыми алгебраическими многообразиями. А есть замечательная теорема, которая не доказана в полной общности, что если у вас есть хорошее алгебраическое многообразие, т. е. множество, задаваемое алгебраическими уравнениями, и если эти уравнения имеют смысл над конечным полем (например, если коэффициенты уравнений —целые числа), то тогда число решений этих уравнений над полем $\mathbb{F}_q$ является многочленом от q.

Например, уравнение $x^2 + y^2 = 1$ над полем $\mathbb{F}_q$ имеет конечное число решений, и это число решений является многочленом от q. Это —очень красивая задача. Я очень рекомендую тем, кто хочет поближе познакомиться с конечными полями, попробовать эту задачу решить.

Есть еще гипотеза, которая для многообразий размерности 1 была высказана Андре Вейлем, одним из самых замечательных математиков нашего века, и доказана сравнительно недавно другим замечательным математиком, Пьером Делинем. Гипотеза состоит в том, что число решений является многочленом, коэффициенты которого выражаются через структуру множества решений над комплексным полем.

То же самое уравнение $x^2 + y^2 = 1$ над комплексным полем задает множество, гомеоморфное $\mathbb{C} \setminus \{0\}$. Естественная компактификация этого множества —двумерная сфера $S^2 \simeq \mathbb{CP}^1$.

К сожалению, в полном объеме эта теорема не доказана. Не известно, какие алгебраические многообразия обладают таким замечательным свойством, что число точек этого алгебраического многообразия над полем $\mathbb{F}_q$ является многочленом от q. Не известно также, какой смысл имеют коэффициенты этого многочлена. Но мнемоническое правило известно. Эти коэффициенты должны соответствовать размерностям групп гомологий этого многообразия над вещественным или комплексным полем. Точнее говоря, $q = 1$ отвечает комплексному полю, а $q = -1$ отвечает вещественному полю. Неформальное объяснение этому примерно такое. Многообразие в каком-то смысле склеивается из клеток, которые устроены как обычные аффинные пространства. Над вещественным полем это дает вклад в гомологии размерности n, равный $(-1)^n$ (здесь n —размерность клетки). Над комплексным полем клетка дает вклад в гомологии размерности $2n$, всегда равный 1. А над конечным полем клетка размерности n дает вклад, равный q^n. Поэтому если многообразие допускает клеточное разбиение, согласованное с его структурой алгебраического многообразия (например, окружность допускает разбиение на точку и прямую), то мы получаем доказательство. Это «доказательство» очень простое, убедительное, но неправильное, потому что слишком мало многообразий допускает такое клеточное разбиение. А сам факт верен и в гораздо более общей ситуации, когда разбиения на клетки не существует.

В интересующем меня случае я не знаю, есть ли теорема, обеспечивающая это утверждение, или нет такой теоремы. И те специалисты, которых я об этом спрашивал, тоже не могут сказать сразу. Так вот, последний многочлен $Y_n(q)$ описывает число коприсоединенных орбит для нашей группы. Это число можно интерпретировать как число точек на некотором алгебраическом многообразии над конечным полем. И если к этому многообразию применимы те или иные теоремы, которые в этой области известны, то тогда Y_n будет многочленом от q.

Многочлен Y_n имеет интерпретацию в духе современной квантовой статистической физики. В этой теории запас слов довольно бедный. Нужно знать такие слова: состояние, гибсовское распределение, энергия и статсумма. Больше ничего знать не нужно. Система состояний —это нечто, когда рисуется решетка, система узлов, или что-то, где имеются клеточки, ребра, вершины, какое-либо двумерное образование или, в крайнем случае, трехмерное. Состояния обычно получаются, если каждому элементу системы приписать какое-либо конкретное значение, чаще всего 0 или 1.

В моем случае системой можно назвать строго верхнетреугольную матрицу, а состояние системы заключается в том, что в каждую клеточку над диагональю можно поставить элемент конечного поля. Еще интересно рассмотреть составную систему, в которую входят верхнетреугольные матрицы X и Y и нижнетреугольная матрица F. Количество состояний этой системы равно $q^{3n(n-1)/2}$. Это число довольно большое. Для первого интересного примера $q = 2$ и $n = 13$. Для него имеется 2^{228} состояний.

Когда у вас есть система и заданы состояния, нужно придумать некоторую функцию, которая была бы явно лучше всех остальных функций. Обычно, если такую функцию вы придумали, то вы называете ее энергией и обозначаете H. Потом берете сумму по всем состояниям s:

$$\sum_s e^{-\beta H(s)}.$$

Здесь β —параметр, который обычно интерпретируют как величину, обратно пропорциональную температуре: $\beta = 1/kT$, где T —абсолютная температура в градусах Кельвина, а k —постоянная Больцмана. Сосчитав такую сумму, вы получаете функцию от β. При всей неуправляемости этого процесса, про такую функцию от β иногда можно что-нибудь сказать. Например, иногда можно доказать, что эта функция гладкая. Тогда с точки зрения физики система неинтересная —у нее нет фазовых переходов. А иногда эта функция не является гладкой. Тогда у системы есть фазовые переходы. В очень редких случаях эту функцию можно сосчитать явно и увидеть фазовые переходы невооруженным глазом: ответ задается одной формулой на одном отрезке и другой формулой на другом отрезке.

Вернемся к нашему примеру. Какая наиболее интересная функция может быть построена из трех матриц, две из которых верхнетреугольные, а одна нижнетреугольная? Я утверждаю, что наиболее интересная функция —это $\mathrm{tr}(F\cdot[X,Y])$. Значения этой функции лежат в конечном поле. Но надо придать смысл не самой функции, а ее экспоненте. Что такое экспонента, здесь хорошо известно. Экспонента —это функция, удовлетворяющая функциональному уравнению: сложение переходит в умножение. Для конечного поля —это так называемый аддитивный характер χ. Получаем сумму

$$\sum_{X,Y,F} \chi(\mathrm{tr}(F\cdot[X,Y])).$$

Есть предположение, что эта сумма равна $q^{n(n-1)/2}Y_n(q)$. Это означает, что функция очень сильно осциллирует и почти все члены взаимоуничтожаются. Поэтому считать сумму необычайно трудно. Обычно для вычисления таких сумм используют метод Монте-Карло, а для сильно осциллирующих функций он непригоден.

Теперь я могу сформулировать общую гипотезу: все 5 последовательностей многочленов на самом деле совпадают.

Что по этому поводу известно? Результаты таковы:

1) $A_n = C_n$ для всех n. Я повторяю, что доказательство этого утверждения имеется и опубликовано.[2] Модный современный журнал, в котором оно опубликовано, физического тела не имеет, он не на бумаге

[2] S. B. Ekhad and D. Zeilberger, The Electronic J. Combinatorics 3 (1996), #R2.

печатается, а существует только в виде файлов. Но любой подписчик может иметь его на своем компьютере. При желании можно распечатать доказательство, что я и проделал, но, надо сказать, ничего из него не понял. Так что если кто-нибудь мне объяснит, как это доказать, не обращаясь к той статье, то будет очень приятно.

2) Утверждение о том, что $A_n = B_n$, проверено экспериментально для $n \leq 26$. На мой взгляд, это не оставляет сомнений в том, что это верно всегда. Но науке известны разные примеры подобных совпадений. Например, многочлен $n^2 - n + 41$ для первых сорока значений n дает простое число.

3) Равенство $A_n = B_n = C_n = D_n = Y_n$ экспериментально проверено для $n \leq 11$.

4) По-видимому, $A_n \neq D_n$ для $n = 13$. Это объяснялось на предыдущей лекции. Недавно выяснилось, что существует треугольная матрица порядка 13 над полем из двух элементов, которая не сопряжена своей обратной. Структура ее жордановых блоков такова: 4, 5, 4.

На этом я хочу закончить лекцию. Но если кто-нибудь заинтересуется этим и готов это обсуждать, то я всегда доступен по электронной почте:

`kirillov@math.upenn.edu`