Лекция 4

ЭКСПОНЕНТА ЛИНЕЙНОГО ДИФФЕРЕНЦИАЛЬНОГО ОПЕРАТОРА. — ФОРМУЛА ДЛЯ ЗНАЧЕНИЙ ГЛАДКИХ ФУНКЦИЙ В НОРМАЛЬНОЙ ОКРЕСТНОСТИ ЕДИНИЦЫ ГРУППЫ ЛИ. — ФОРМУЛА ДЛЯ ЗНАЧЕНИЙ ГЛАДКИХ ФУНКЦИЙ НА ПРОИЗВЕДЕНИИ ДВУХ ЭЛЕМЕНТОВ. — РЯД КЕМПБЕЛЛА — ХАУСДОРФА И МНОГОЧЛЕНЫ ДЫНКИНА. — СХОДИМОСТЬ РЯДА КЕМПБЕЛЛА — ХАУСДОРФА. — ВОССТАНОВЛЕНИЕ ГРУПУСКУЛЫ ЛИ ПО ЕЕ АЛГЕБРЕ ЛИ. — ОПЕРАЦИИ В АЛГЕБРЕ ЛИ ГРУППЫ ЛИ И ОДНОПАРАМЕТРИЧЕСКИЕ ПОДГРУППЫ. — ДИФФЕРЕНЦИАЛЫ ВНУТРЕННИХ АВТОМОРФИЗМОВ. — ДИФФЕРЕНЦИАЛ ЭКСПОНЕНЦИАЛЬНОГО ОТОБРАЖЕНИЯ. — КАНОНИЧЕСКИЕ КООРДИНАТЫ. — ЕДИНСТВЕННОСТЬ СТРУКТУРЫ ГРУППЫ ЛИ. — ГРУППЫ БЕЗ МАЛЫХ ПОДГРУПП И ПЯТАЯ ПРОБЛЕМА ГИЛЬБЕРТА.

Пусть M — произвольное гладкое многообразие. В отличие от предыдущего, мы теперь предположим, что M является *аналитическим* (класса C^{ω}) многообразием. Поскольку каждое векторное поле X на M мы можем рассматривать как линейный дифференциальный оператор, действующий на гладкие функции, по аналогии с матричным случаем можно ввести в рассмотрение линейный оператор

$$(1)\qquad e^X = E + X + \frac{X^2}{2!} + \ldots + \frac{X^n}{n!} + \ldots = \sum_{n=0}^{\infty} \frac{X^n}{n!},$$

где E — тождественный оператор, а X^n обозначает n-кратную итерацию оператора X.

Конечно, мы здесь должны определить, что следует понимать под суммой бесконечного ряда (1). В принципе для этого необходимо ввести в пространство $\mathcal{O}(M)$ топологию (скажем, задаваемую некоторой нормой). Однако для простоты мы предпочтем здесь обходный путь и будем понимать сходимость ряда (4) в «слабом» смысле. Именно, мы определим результат $e^X f$ применения оператора e^X к функции $f \in \mathcal{O}(M)$ формулой

$$(2)\qquad e^X f = f + Xf + \frac{X^2 f}{2!} + \ldots + \frac{X^n f}{n!} + \ldots = \sum_{n=0}^{\infty} \frac{X^n f}{n!},$$

где $X^n f = X(X^{n-1} f)$, и будем считать, что этот оператор применяется только к таким функциям, для которых ряд (2) (уже функциональный) имеет непустую область сходимости (которая и принимается за область определения функции $e^X f$).

Оператор e^X уже не будет дифференциальным оператором. Именно, как мы сейчас покажем, он, вообще говоря, индуцируется некоторым диффеоморфизмом Φ: $M \to M$, т. е. имеет вид $f \longmapsto f \circ \Phi$. Для этого требуется только, чтобы интегральные кривые $t \longmapsto \varphi_a(t)$ поля X были «достаточно длинны», а именно, были определены при $|t| \leqslant 1$, и чтобы область определения $W(f)$ функции f была «достаточно большой», а именно, чтобы существовали такие точки $a \in W(f)$, что $\varphi_a(t) \in W(f)$ при $|t| \leqslant 1$. Тогда функция $e^X f$ будет определена для всех таких точек a и будет выражаться формулой

$$(3)\qquad (e^X f)(a) = f(\varphi_a(1)).$$

Действительно, введем в рассмотрение функцию

$$F(t) = f(\varphi_a(t)).$$

По условию эта функция определена и аналитична при $|t| \leqslant 1$. Поэтому она разлагается в ряд Тейлора

$$F(t) = \sum_{n=0}^{\infty} \frac{F^{(n)}(0)}{n!} t^n,$$

сходящийся при $|t| \leqslant 1$. С другой стороны, так как кривая $t \mapsto \varphi_a(t)$ является интегральной кривой поля X, то

$$F'(t) = \frac{df(\varphi_a(t))}{dt} = \frac{d\varphi_a(t)}{dt} f = X_{\varphi_a(t)} f = (Xf)(\varphi_a(t)).$$

Это означает, что функцией $F(t)$ для функции Xf служит функция $F'(t)$, откуда посредством очевидной индукции выводится, что функцией $F(t)$ для функции $X^n f$ служит функция $F^{(n)}(t)$, т. е. что

$$F^{(n)}(t) = (X^n f)(\varphi_a(t)).$$

Поэтому

$$F^{(n)}(0) = (X^n f)(\varphi_a(0)) = (X^n f)(a),$$

и, значит,

$$F(t) = \sum_{n=0}^{\infty} \frac{(X^n f)(a)}{n!} t^n = \sum_{n=0}^{\infty} \frac{(tX)^n f}{n!}(a) = (e^{tX} f)(a).$$

Для завершения доказательства остается положить $t = 1$. $\square$

Мы будем применять общую формулу (3) к левоинвариантным векторным полям X на аналитической группе (или групускуле) Ли G и к функциям f, определенным и аналитическим в некоторой окрестности единицы e группы G. За точку a мы примем единицу e. Обозначив точку $\varphi_e(1)$ символом $\exp X$, мы перепишем для этого случая формулу (3) в следующем виде:

$$f(\exp X) = (e^X f)(e). \tag{4}$$

В этом виде и будем ее использовать.

Для левоинвариантного векторного поля X интегральной кривой $t \mapsto \varphi_e(t)$ является соответствующая однопараметрическая подгруппа $t \mapsto \beta(t)$. Поэтому $\exp X = \beta(1)$ и формула (4) имеет место для любых функций f, область определения которых содержит отрезок $t \mapsto \beta(t)$, $|t| \leqslant 1$, этой подгруппы.

По определению exp представляет собой такое отображение линейного пространства $\mathfrak{g} = \mathfrak{l}(G)$ в группу G, что $\exp 0 = e$. Оно, очевидно, обладает свойством естественности, т. е. для любого гомоморфизма

$\Phi\colon G \to H$ групп Ли имеет место коммутативная диаграмма

$$\begin{tikzcd} \mathfrak{l}(G) \arrow[r, "\mathfrak{l}(\Phi)"] \arrow[d, "\exp"'] & \mathfrak{l}(H) \arrow[d, "\exp"] \\ G \arrow[r, "\Phi"] & H \end{tikzcd}$$

Из теоремы о гладкой зависимости решений дифференциальных уравнений от начальных данных непосредственно вытекает, что отображение $\exp\colon \mathfrak{l}(G) \to G$ гладко.

Утверждение А. *Отображение*

$$\exp\colon \mathfrak{l}(G) \to G$$

является в точке 0 *диффеоморфизмом.*

Мы докажем это утверждение позже.

Определение 1. Окрестность $\overset{\circ}{U}$ нулевого вектора $0 \in \mathfrak{g}$ мы будем называть *нормальной*, если:

а) она обладает свойством звездности, т. е. вместе с некоторым вектором X содержит и все векторы вида tX при $|t| \leqslant 1$;

б) отображение exp диффеоморфно отображает окрестность $\overset{\circ}{U}$ на некоторую окрестность U единицы e группы G.

Окрестность U мы также будем называть *нормальной окрестностью.*

Согласно утверждению А существуют сколь угодно малые (содержащиеся в любой наперед заданной окрестности) нормальные окрестности (как вектора $0 \in \mathfrak{g}$, так и единицы $e \in G$).

По построению $\exp X = \beta(1)$, где β — однопараметрическая подгруппа, являющаяся интегральной кривой поля X. Но ясно, что для поля aX, где $a \in \mathbb{R}$, интегральной кривой будет кривая $t \longmapsto \beta(at)$, также, конечно, являющаяся однопараметрической подгруппой. Отсюда следует, что $\exp(aX) = \beta(a)$ или, если мы обозначим a через t, что

$$\exp(tX) = \beta(t).$$

Эта формула означает, что $\beta\colon t \to \exp(tX)$, т. е. что *однопараметрические подгруппы* β *группы* G *является образами прямых* $t \longmapsto tX$ *при отображении* exp.

В силу условия а) определения 1 отсюда следует, что условие на область определения функции f, необходимое (и достаточное) для справедливости формулы (4), заведомо выполнено, если этой областью является некоторая нормальная окрестность U точки e. При этом, поскольку любая точка из U имеет вид $\exp X$, где $X \in \mathring{U}$, *формула* (4) *задает функцию* f *на всей окрестности* U.

Имея это в виду, мы применим формулу (4) к вычислению значения $f(ab)$ функции f на произведении ab двух элементов $a = \exp X$ и $b = \exp Y$ (конечно, в предположении, что $ab \in U$).

Для данного элемента $a \in U$ формула $f_a(b) = f(ab)$ определяет (на некоторой содержащейся в U окрестности точки e, которую мы можем считать нормальной) гладкую функцию f_a. При этом, согласно формуле (4), если $b = \exp Y$, то $f_a(b) = (e^Y f_a)(e)$.

С другой стороны, так как функция f_a является не чем иным, как композицией $f \circ L_a$ левого сдвига L_a: $b \longmapsto ab$ с функцией f, то в силу левоинвариантности поля Y имеет место формула

$$Yf_a = Yf \circ L_a = (Yf)_a.$$

Но тогда $Y^n f_a = (Y^n f)_a$ для любого $n \geqslant 0$, и потому

$$e^Y f_a = (e^Y f)_a.$$

Следовательно, если $a = \exp X$, то (мы снова применяем формулу (4))

$$f_a(b) = (e^Y f_a)(e) = (e^Y f)(a) = (e^X e^Y f)(e).$$

Поскольку $f_a(b) = f(\exp X \cdot \exp Y)$, этим доказано, что

$$f(\exp X \cdot \exp Y) = (e^X e^Y f)(e) \tag{5}$$

для любых X и Y из соответствующей нормальной окрестности нуля алгебры $\mathfrak{g}$.

По определению (мы пока игнорируем вопрос о сходимости)

$$e^X e^Y = \sum_{p=0}^{\infty} \frac{1}{p!} X^p \left(\sum_{q=0}^{\infty} \frac{1}{q!} Y^q \right) = \sum_{p,\, q=0}^{\infty} \frac{1}{p!q!} X^p Y^q.$$

Подставив этот ряд в логарифмический ряд

$$\ln Z = (Z - E) - \frac{(Z-E)^2}{2} + \ldots = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} (Z-E)^k,$$

мы получим (учитывая, что операторы X и Y, вообще говоря, не коммутируют) формальный ряд

$$\ln(e^X e^Y) = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} \left(\sum_{\substack{p,\, q=0 \\ p+q>0}}^{\infty} \frac{X^p Y^q}{p!\,q!} \right)^k =$$

$$= \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} \sum \frac{X^{p_1} Y^{q_1} \ldots X^{p_k} Y^{q_k}}{p_1!\,q_1! \ldots p_k!\,q_k!},$$

где во внутренней сумме суммирование распространено на всевозможные наборы $(p_1, \ldots, p_k, q_1, \ldots, q_k)$ целых неотрицательных чисел, подчиненных условиям

$$(6) \qquad p_1 + q_1 > 0, \ \ldots, \ p_k + q_k > 0.$$

Положив

$$(7) \qquad z_n(x, y) = \sum_{k=1}^{n} \frac{(-1)^{k-1}}{k} \sum \frac{x^{p_1} y^{q_1} \ldots x^{p_k} y^{q_k}}{p_1!\,q_1! \ldots p_k!\,q_k!},$$

где во внутренней сумме показатели $p_1, \ldots, p_k, q_1, \ldots$ $\ldots, q_k$, кроме условия (6), удовлетворяют также условию

$$(8) \qquad p_1 + \ldots + p_k + q_1 + \ldots + q_k = n,$$

мы перепишем ряд $\ln(e^X e^Y)$ в следующем виде:

$$(9) \qquad \ln(e^X e^Y) = \sum_{n=1}^{\infty} z_n(X, Y).$$

Этот формальный ряд называется *рядом Кемпбелла — Хаусдорфа.*

С алгебраической точки зрения $z_n(x, y)$ является не чем иным, как многочленом от, вообще говоря, некоммутирующих переменных x и y. Рассмотрим поэтому такого рода многочлены поподробнее.

Пусть x и y — некоторые символы и $\mathbb{K}$ — произвольное поле. Обыкновенные многочлены от x и y над полем $\mathbb{K}$ получаются из x, y и элементов поля $\mathbb{K}$ применением любое число раз действий сложения и умножения. Аналогичным образом конструируются и *некоммутативные многочлены* от x и y над полем $\mathbb{K}$; единственное отличие состоит лишь в том, что их умножение не предполагается коммутативным (но по-прежнему $kf = fk$ для любого многочлена f и любого числа $k \in \mathbb{K}$). Все они образуют унитальную алгебру, которую мы будем обозначать символом $\mathbb{K}\langle x, y\rangle$. (Более формальное определение этой алгебры мы дадим в следующей лекции.)

В коммутаторной алгебре $[\mathbb{K}\langle x, y\rangle]$ алгебры $\mathbb{K}\langle x, y\rangle$ естественным образом определяются *лиевы многочлены* от x и y, получающиеся из x и y действиями сложения, умножения на элементы поля $\mathbb{K}$ и операцией Ли $a, b \longmapsto [a, b] = ab - ba$. Ясно, что они составляют подалгебру алгебры Ли $[\mathbb{K}\langle x, y\rangle]$, содержащую элементы x, y и содержащуюся в любой другой такой подалгебре (т. е. — в общеалгебраических терминах — порожденную элементами x и y). Мы будем обозначать эту подалгебру символом $\mathrm{L}\langle x, y\rangle$.

Каждый лиев многочлен $u \in \mathrm{L}\langle x, y\rangle$ является одновременно и многочленом из $\mathbb{K}\langle x, y\rangle$ (достаточно раскрыть все скобки Ли по правилу $[a, b] = ab - ba$). В этом его качестве мы будем обозначать многочлен u символом ιu. Формально, отображение

$$\iota\colon \mathrm{L}\langle x, y\rangle \to \mathbb{K}\langle x, y\rangle, \qquad u \longmapsto \iota u,$$

определяется как линейное отображение, обладающее тем свойством, что $\iota[a, b] = \iota a \cdot \iota b - \iota b \cdot \iota a$ для любых элементов $a, b \in \mathrm{L}\langle x, y\rangle$. Оно по определению инъективно.

Если поле $\mathbb{K}$ имеет характеристику 0, то формула (7) определяет в $\mathbb{K}\langle x, y\rangle$ некоторый элемент $z_n(x, y)$.

Утверждение В. *Существует такой лиев многочлен* $د_n(x, y)$, что

$$\iota د_n(x, y) = z_n(x, y).$$

Здесь символ $د$ читается «дэ» (он является одной из форм арабской буквы «даль»).

Мы докажем утверждение В позже.

Примеры.

Пусть $n = 1$. Очевидно, что

$$z_1(x, y) = x + y,$$

и потому

$$\mathfrak{Z}_1(x, y) = x + y$$

(так что скобка Ли в построении многочлена $\mathfrak{Z}_1(x, y)$ не участвует).

Пусть $n = 2$. При $k = 1$ внутренняя сумма в формуле (7) содержит три слагаемых $\frac{1}{2}x^2$, xy и $\frac{1}{2}y^2$, а при $k = 2$ четыре слагаемых x^2, xy, yx и y^2. Следовательно,

$$z_2(x, y) = \frac{1}{2}xy - \frac{1}{2}yx,$$

и потому

$$\mathfrak{Z}_2(x, y) = \frac{1}{2}[x, y].$$

Аналогично проверяется, что

$$z_3(x, y) = \frac{1}{12}x^2y + \frac{1}{12}yx^2 + \frac{1}{12}xy^2 + \\ + \frac{1}{12}y^2x - \frac{1}{6}xyx - \frac{1}{6}yxy.$$

Здесь уже непосредственно не видно, каков многочлен $\mathfrak{Z}_3(x, y)$ (и, вообще говоря, существует ли он). Однако после немногих проб можно убедиться, что

$$\mathfrak{Z}_3(x, y) = \frac{1}{12}[x, [x, y]] + \frac{1}{12}[y, [y, x]].$$

С ростом n вычисления стремительно усложняются. Тем не менее оказывается возможным указать для многочленов $\mathfrak{Z}_n(x, y)$ явную формулу, аналогичную формуле (7) для многочлена $z_n(x, y)$. Впервые эта формула была указана Е. Б. Дынкиным. Поэтому многочлены $\mathfrak{Z}_n(x, n)$ мы будем называть *многочленами Дынкина.*

Заметим, что *многочлен* $\mathfrak{Z}_n(x, n)$ *однороден степени* n *по* x *и* y, т. е.

$$\mathfrak{Z}_n(tx, ty) = t^n\mathfrak{Z}_n(x, y)$$

для любого t.

Поскольку для векторных полей (линейных дифференциальных операторов) операция $X, Y \mapsto [X, Y]$ также выражается формулой $[X, Y] = XY - YX$, из утверждения В следует, что *для любых операторов* $X, Y \in \mathfrak{l}(G)$

каждый оператор $z_n(X, Y)$ *принадлежит алгебре* $\mathfrak{l}(G)$. Поэтому алгебре $\mathfrak{l}(G)$ принадлежат и оператор $\ln(e^X e^Y)$, если, конечно, этот оператор имеет смысл, т. е. если ряд (9) сходится. Рассмотрим поэтому вопрос о сходимости этого ряда, или, точнее, ряда

$$(10) \qquad \mathfrak{Z}_1(X, Y) + \mathfrak{Z}_2(X, Y) + \ldots + \mathfrak{Z}_n(X, Y) + \ldots,$$

состоящего из лиевых многочленов $\mathfrak{Z}_n(X, Y)$ от X и Y.

Все члены ряда (10) лежат в конечномерном линеале $\mathfrak{g} = \mathfrak{l}(G)$, и потому исследование его сходимости можно было бы, вообще говоря, осуществить, рассматривая в $\mathfrak{g}$ произвольную (например, евклидову) норму.

Однако необходимые для этого оценки требуют довольно детальной информации о строении многочленов $\mathfrak{Z}_n(x, y)$, которую мы получим только в следующей лекции. Поэтому пока мы вынуждены идти по обходному пути и вернуться к ряду (9), строение членов которого нам известно, но при переходе к которому мы теряем преимущество конечномерности.

В соответствии с нашим общим отношением к сходимости операторных рядов мы будем понимать сходимость ряда (9) в «слабом» смысле, но несколько более сильном, чем выше. Именно, мы будем считать, что оператор $\ln(e^X e^Y)$, определенный рядом (9), применим к функции $f \in \mathcal{O}(G)$, если в области определения этой функции сходится функциональный ряд

$$(11) \qquad z_1(X, Y)f + z_2(X, Y)f + \ldots + z_n(X, Y)f + \ldots$$

Сумму g этого ряда мы и будем считать результатом применения оператора $\ln(e^X e^Y)$ к функции f.

Таким образом, в силу этого определения $W(g) = W(f)$ (тогда как раньше $W(g) \subset W(f)$).

Предполагая, что в $\mathfrak{g}$ задана некоторая норма $\| \ \|$, мы покажем сейчас, что *существует такое число* $\delta_0 > 0$, *что при* $\|X\| < \delta_0$ *и* $\|Y\| < \delta_0$ *оператор* $\ln(e^X e^Y)$ *применим к любой функции* $f \in \mathcal{O}(G)$. Поскольку каждый элемент из G, принадлежащий области определения функции f, обладает координатной окрестностью U с компактным замыканием, содержащейся в этой области, для этого достаточно доказать, что *для любой координатной окрестности* U *с компактным замыканием* $\overline{U}$ *оператор*

$\ln(e^X e^Y)$ *определен на линейном пространстве* $\mathcal{F}(\bar{U})$ *всех функций, гладких на* $\bar{U}$.

С этой целью мы заметим, что, поскольку множество $\bar{U}$ компактно, для любой функции $f \in \mathcal{F}(\bar{U})$ определено число

$$\|f\| = \max\left(|f|, \left|\frac{\partial f}{\partial x^1}\right|, \ldots, \left|\frac{\partial f}{\partial x^n}\right|\right),$$

где $x^1, \ldots, x^n$ — локальные координаты в координатной окрестности U (или, точнее, в некоторой большей окрестности, содержащей замыкание $\bar{U}$ окрестности U). Число $\|f\|$ зависит, конечно, от выбора координат $x^1, \ldots$ $\ldots, x^n$, но это обстоятельство не играет в дальнейшем никакой роли. Очевидным образом проверяется, что так построенный функционал $f \longmapsto \|f\|$ является нормой на линеале $\mathcal{F}(\bar{U})$. При этом, так как $Xf = X^i \frac{\partial f}{\partial x^i}$ и $|X^i| \leqslant$ $\leqslant \|X^i\|$ для всех i, то для каждого поля $X \in \mathfrak{g}$ с $\|X\| < \delta$ будет иметь место оценка

$$\|Xf\| \leqslant \delta \|f\|.$$

Посредством очевидной индукции отсюда выводится, что если $\|X\| < \delta$ и $\|Y\| < \delta$, то для любых показателей $p_1, \ldots, p_k, q_1, \ldots, q_k$ имеет место оценка

$$\|X^{p_1}Y^{q_1} \ldots X^{p_k}Y^{q_k} f\| \leqslant \delta^n \|f\|,$$

где $n = p_1 + \ldots + p_k + q_1 + \ldots + q_k$. Это означает, что ряд (11) мажорируется (после деления на $\|f\|$) рядом $\ln(e^X e^Y)$, в который вместо X и Y подставлено число δ, а коэффициенты заменены их абсолютными значениями, т. е. мажорируется рядом, получающимся из ряда

$$\sum_{n=1}^{\infty} \frac{(z-1)^n}{n} \tag{12}$$

подстановкой вместо z ряда

$$e^{2\delta} = \sum_{n=0}^{\infty} \frac{(2\delta)^n}{n!}.$$

Поскольку ряд (12) сходится при $|z - 1| < 1$, этот мажорирующий ряд будет сходиться при $|e^{2\delta} - 1| < 1$, т. е.

при $\delta < \frac{\ln 2}{2}$. Следовательно, при $\delta < \frac{\ln 2}{2}$ ряд (11) будет равномерно сходиться в $\bar{U}$ к некоторой гладкой функции, т. е. оператор $\ln(e^X e^Y)$ будет применим к функции f. □

Поскольку ряд (10) отличается от ряда (9) лишь обозначениями (по определению $z_n(X, Y) = \mathfrak{Z}_n(X\ Y)$ для любых полей $X, Y \in \mathfrak{g}$), этим доказано, что при $\|X\| < \delta_0$, $\|Y\| < \delta_0$, где $\delta_0 = \frac{\ln 2}{2}$, ряд (10) сходится в том же самом «слабом» смысле. Но известно, что для операторных рядов, члены которых принадлежат конечномерному линейному пространству, «слабая» сходимость совпадает с обычной «сильной» сходимостью. Поэтому *при* $\|X\| < \delta_0$ *и* $\|Y\| < \delta_0$ *ряд* (10) *сходится* (в обычном смысле).

Замечание 1. Независимо от ссылки на общую теорему о совпадении «сильной» и «слабой» сходимостей, мы можем доказать сходимость ряда (10), рассмотрев произвольную систему локальных координат $x^1, \ldots, x^n$ в точке e группы G. При $f = x^i$, $i = 1, \ldots, n$, ряд (11) превращается в ряд, составленный из компонент $z_n(X, Y)^i = \mathfrak{Z}_n(X, Y)^i$ векторных полей $\mathfrak{Z}_n(X, Y)$. Следовательно, этот ряд сходится. Поэтому сходится и ряд, составленный из их значений $\mathfrak{Z}_n(X, Y)_e^i$ в точке e, т. е. из координат $\left(\text{в базисе} \left(\frac{\partial}{\partial x^1}\right)_e, \ldots, \left(\frac{\partial}{\partial x^n}\right)_e\right)$ векторов $\mathfrak{Z}_n(X, Y)_e$. Это означает, что сходится векторный ряд

$$(13)\qquad \mathfrak{Z}_1(X, Y)_e + \mathfrak{Z}_2(X, Y)_e + \ldots + \mathfrak{Z}_n(X, Y)_e + \ldots$$

Пусть $\mathfrak{Z}(X, Y)_e$ — его сумма, а $\mathfrak{Z}(X, Y)$ — левоинвариантное векторное поле, принимающее в e значение $\mathfrak{Z}(X, Y)_e$. Для любого элемента $a \in G$ ряд

$$\mathfrak{Z}_1(X, Y)_a + \mathfrak{Z}_2(X, Y)_a + \ldots + \mathfrak{Z}_n(X, Y)_a + \ldots,$$

составленный из значений, принимаемых членами ряда (10) в точке e, получается из ряда (13) применением линейного оператора dL_a, и потому сходится к вектору $(dL_a)(\mathfrak{Z}(X, Y)_e) = \mathfrak{Z}(X, Y)_a$. Это и означает, что ряд (10) сходится к $\mathfrak{Z}(X, Y)$.

В курсе анализа доказывается, что если функция разлагается в сходящийся степенной ряд, то этот ряд единствен. Отсюда, в частности, следует, что если в ряд для e^z подставить ряд для $\ln z$, то после приведения подобных членов получится ряд z. Иными словами, равенство $e^{\ln z} = z$ имеет место не только для функций, но и для формальных рядов. Поэтому оно остается верным и при подстановке в него вместо z, скажем, ряда для $e^X e^Y$. Это означает, что при подстановке в ряд для e^z вместо z ряда (10), получается ряд для $e^X e^Y$. Поскольку все ряды здесь сходятся, этим доказано, что

$$e^{\mathfrak{Z}(X, Y)} = e^X e^Y. \tag{14}$$

Эта формула имеет место для любых элементов X, Y алгебры $\mathfrak{l}(G)$, для которых $\|X\| < \delta_0$ и $\|Y\| < \delta_0$, где δ_0 — произвольное достаточно малое число (согласно произведенному выше исследованию годится любое $\delta_0 < \frac{\ln 2}{2}$).

Вернемся теперь к формуле (5). Эта формула имеет место для любых X и Y из некоторой нормальной окрестности $\overset{\circ}{U}$ нуля алгебры $\mathfrak{g} = \mathfrak{l}(G)$, обладающих тем свойством, что $\exp X \cdot \exp Y \in U = \exp \overset{\circ}{U}$. В частности, существует такое δ_1, что при любом положительном $\delta < \delta_1$ формула (5) справедлива при $\|X\| < \delta$ и $\|Y\| < \delta$. Поэтому, предполагая, кроме того, что $\delta < \delta_0$, мы можем формулу (5) переписать в следующем виде:

$$f(\exp X \cdot \exp Y) = (e^Z f)(e), \quad \text{где} \quad Z = \mathfrak{Z}(X, Y).$$

Заметим теперь, что элемент Z алгебры $\mathfrak{l}(G)$, очевидно, непрерывно зависит от ее элементов X и Y, так что, в частности, $Z \to 0$, когда $X \to 0$ и $Y \to 0$. Поэтому при достаточно малом δ к полю Z применима формула (4), согласно которой

$$(e^Z f)(e) = f(\exp Z).$$

Таким образом, если $\|X\| < \delta$ и $\|Y\| < \delta$, где $\delta > 0$ достаточно мало, то

$$f(\exp X \cdot \exp Y) = f(\exp Z)$$

для каждой гладкой функции f, определенной в точках $\exp X \cdot \exp Y$ и $\exp Z$. В частности, это верно, когда

f является одной из координат $x^1, \ldots, x^n$ в точке e. Таким образом, все координаты $x^1, \ldots, x^n$ принимают в точках $\exp X \cdot \exp Y$ и $\exp Z$ одинаковые значения, что возможно только тогда, когда

$$\exp X \cdot \exp Y = \exp Z.$$

Резюмируя, мы получаем следующую теорему, которая была главной целью всех предыдущих рассуждений:

Теорема 1. *Единица e аналитической группы (или групускулы) Ли имеет окрестность U, обладающую следующими свойствами:*

а) *существует такое $\delta > 0$, что каждая точка окрестности U единственным образом представляется в виде* $\exp X$, *где $X \in \mathfrak{l}(G)$ и $\|X\| < \delta$;*

б) *для любых двух точек* $\exp X$ *и* $\exp Y$ *окрестности U в алгебре $\mathfrak{l}(G)$ существует такой элемент Z, что*

$$\exp X \cdot \exp Y = \exp Z; \tag{15}$$

в) *этот элемент Z является суммой $\mathfrak{Z}(X, Y)$ ряда* (10), *членами которого являются многочлены Дынкина $\mathfrak{Z}_n(X, Y)$ от X и Y.* □

Эта теорема означает, что группа (групускула) Ли G обладает частью, умножение в которой однозначно восстанавливается (в соответствии с формулой (15)) по алгебре Ли $\mathfrak{l}(G)$. Следовательно, *две (аналитические) групускулы Ли с изоморфными алгебрами Ли изоморфны* (точнее, изоморфны их ростки). В этом смысле можно сказать, что *с точностью до изоморфизма функтор Ли*

$$L\colon \text{GR-LOC} \to \text{ALG}_f\text{-LIE}$$

обратим.

Более точное утверждение мы получим в лекции 6.

С помощью теоремы 1 легко решается отложенный выше вопрос об интерпретации операций алгебры $\mathfrak{g} = \mathfrak{l}(G)$ в терминах однопараметрических подгрупп.

Если в линеале $\mathfrak{g}$ произвольно выбран базис $e_1, \ldots, e_n$, то для любой нормальной окрестности U единицы группы G композиция h диффеоморфизма $\exp^{-1}$: $U \to \overset{\circ}{U}$ с ограничением на $\overset{\circ}{U}$ соответствующего координатного изоморфизма $\mathfrak{g} \to \mathbb{R}^n$ будет диффеоморфизмом окрестности U на некоторое открытое множество пространства $\mathbb{R}^n$, т. е. пара (U, h) будет картой на группе Ли G. Соответствующие локальные координаты называются *нормальными координатами*. Таким образом, если $a = \exp X$ и $X = x^1e_1 + \ldots + x^ne_n$, то числа $x^1, \ldots, x^n$ являются нормальными координатами точки $a \in U$.

Однопараметрическую подгруппу $t \mapsto \exp(tX)$, соответствующую элементу $X \in \mathfrak{g}$ (т. е. в интерпретации X как левоинвариантного векторного поля — интегральную кривую этого поля, проходящую при $t = 0$ через точку e, а в интерпретации X как вектора касательного пространства $\mathsf{T}_e(G)$ — однопараметрическую подгруппу, имеющую при $t = 0$ касательный вектор X), мы будем обозначать символом β_X.

Предложение 1. *Для любого $X \in \mathfrak{g}$ и любого $k \in \mathbb{R}$ элемент $kX \in \mathfrak{g}$ (интерпретированный как элемент пространства $\mathsf{T}_e(G)$) является вектором, касающимся при $t = 0$ кривой*

$$t \mapsto \beta_X(kt). \tag{16}$$

Для любых $X, Y \in \mathfrak{g}$ элемент $X + Y \in \mathfrak{g}$ является (в той же интерпретации) вектором, касающимся при $t = 0$ кривой

$$t \mapsto \beta_X(t)\beta_Y(t), \tag{17}$$

а элемент $[X, Y] \in \mathfrak{g}$ — вектором, касающимся при $t = 0$ кривой

$$t \mapsto \beta_X(\sqrt{t})\beta_Y(\sqrt{t})\beta_X(\sqrt{t})^{-1}\beta_Y(\sqrt{t})^{-1}. \tag{18}$$

Доказательство. Первое утверждение очевидно, поскольку кривая (16), т. е. кривая $t \mapsto \exp(ktX)$ является не чем иным, как однопараметрической подгруппой β_{kX}.

Для доказательства второго утверждения заметим, что так как $\mathfrak{Z}_n(tX, tY) = t^n\mathfrak{Z}_n(X, Y)$ и $\mathfrak{Z}_1(X, Y) = X + Y$, то

$$\mathfrak{Z}(tX, tY) = t(X + Y) + O(t^2),$$

где символ $O(t^2)$ обозначает члены, имеющие по t степень $\geqslant 2$. Поэтому кривая (17) имеет вид

$$t \mapsto \exp(t(X + Y) + O(t^2))$$

и, значит, в нормальных координатах (определенных произвольным базисом алгебры Ли $\mathfrak{g}$) задается функциями

$$x^i(t) = t(X^i + Y^i) + O(t^2).$$

Следовательно, ее касательный вектор при $t = 0$ имеет координаты

$$\left.\frac{dx^i(t)}{dt}\right|_{t=0} = X^i + Y^i$$

и потому совпадает с вектором $X + Y$.

Аналогично, так как

$$\begin{aligned}
(\exp X)\cdot(\exp Y)\cdot(\exp X)^{-1}\cdot(\exp Y)^{-1} &= \\
&= \exp X\cdot\exp Y\cdot\exp(-X)\cdot\exp(-Y) = \\
&= \exp(\mathfrak{I}(X, Y))\exp(\mathfrak{I}(-X, -Y)) = \\
&= \exp\mathfrak{I}(\mathfrak{I}(X, Y), \mathfrak{I}(-X, -Y))
\end{aligned}$$

и

$$\begin{aligned}
\mathfrak{I}(\mathfrak{I}(X, Y), \mathfrak{I}(-X, -Y)) &= \mathfrak{I}(X, Y) + \mathfrak{I}(-X, -Y) + \\
&+ \frac{1}{2}[\mathfrak{I}(X, Y), \mathfrak{I}(-X, -Y)] + \ldots = \\
&= \left\{(X + Y) + \frac{1}{2}[X, Y] + \ldots\right\} + \\
&+ \left\{(-X - Y) + \frac{1}{2}[-X, -Y] + \ldots\right\} + \\
+ \frac{1}{2}[X + Y + \ldots, -X - Y + \ldots] + \ldots &= \frac{1}{2}[X, Y] + \\
&+ \frac{1}{2}[X, Y] + \frac{1}{2}[X, -Y] + \frac{1}{2}[Y, -X] + \ldots = \\
&= [X, Y] + \ldots,
\end{aligned}$$

то кривая (18) имеет вид

$$\begin{aligned}
t \mapsto \exp\left([\sqrt{t}\,X, \sqrt{t}\,Y] + O(t^{3/2})\right) &= \\
&= \exp\left(t[X, Y] + O(t^{3/2})\right),
\end{aligned}$$

и потому в нормальных координатах задается функциями

$$x^i(t) = t[X, Y]^i + O(t^{3/2}), \qquad i = 1, \ldots, n. \tag{19}$$

Следовательно,

$$\left.\frac{dx^i(t)}{dx}\right|_{t=0} = [X, Y]^i,$$

и, значит, касательным вектором к кривой (18) при $t=0$ является вектор $[X, Y]$. $\square$

Замечание 2. Обратим внимание, что в то время как кривая (16) является однопараметрической подгруппой, кривые (16) и (18) однопараметрическими подгруппами не являются. Более того, кривая (18) определена только при $t \geqslant 0$, так что говорить о ее касательном векторе при $t=0$ мы не имеем права. Поэтому мы должны дать специальное ad hoc определение касательного вектора кривой (18) при $t=0$. Мы будем считать этим вектором предел при $t \to 0$ касательных векторов кривой (18) при $t > 0$. Из формул (19) следует, что этот предел существует и равен $[X, Y]$.

Замечание 3. Полезно также иметь в виду, что в формулах (17) и (18) однопараметрические подгруппы β_X и β_Y можно заменить произвольными кривыми α_X: $t \mapsto \alpha_X(t)$ и α_Y: $t \mapsto \alpha_Y(t)$, проходящими при $t=0$ через точку e и имеющими в этой точке касательные векторы X и Y. Действительно, при достаточно малом $|t|$ в $\mathfrak{g}$ будут определены векторы $\mathring{\alpha}_X(t) = \exp^{-1}\alpha_X(t)$ и $\mathring{\alpha}_Y(t) = \exp^{-1}\alpha_Y(t)$, и для этих векторов будут иметь место равенства

$$\mathring{\alpha}_X(t) = tX + O(t^2), \qquad \mathring{\alpha}_Y(t) = tY + O(t^2)$$

и, значит, равенство

$$\mathfrak{Z}(\mathring{\alpha}_X(t), \mathring{\alpha}_Y(t)) = t(X+Y) + O(t^2).$$

Поэтому касательным вектором, скажем, кривой

$$t \mapsto \alpha_X(t)\,\alpha_Y(t) = \exp \mathfrak{Z}(\mathring{\alpha}_X(t), \mathring{\alpha}_Y(t))$$

будет при $t=0$ вектор $X+Y$.

Теорема 1 позволяет также вычислить дифференциал произвольного внутреннего автоморфизма.

Для каждого элемента a группы Ли G дифференциал $(d\Phi_a)_e = \mathfrak{l}(\Phi_a)$ соответствующего внутреннего автоморфизма Φ_a: $x \mapsto axa^{-1}$, $x \in G$, обозначается символом $\mathrm{Ad}(a)$. Он представляет собой линейное обратимое отображение

$$\mathrm{Ad}(a)\colon \mathfrak{g} \to \mathfrak{g}$$

линеала $\mathfrak{g} = \mathfrak{l}(G)$ на себя. Так как $\Phi_a = L_a \circ R_{a^{-1}}$, то (в интерпретации $\mathfrak{g} = \mathsf{T}_e(G)$)

$$\mathrm{Ad}(a) = (dL_a^{-1})_{a^{-1}} \circ (dR_{a^{-1}})_e.$$

Ясно, что отображение

$$\mathrm{Ad}\colon a \mapsto \mathrm{Ad}(a)$$

является гомоморфизмом группы Ли G в группу Ли $\mathrm{Aut}\,\mathfrak{g}$ всех невырожденных линейных операторов линеала $\mathfrak{g}$. Этот гомоморфизм называется *присоединенным представлением* группы Ли G.

Дифференциал $(d\,\mathrm{Ad})_e = \mathfrak{l}(\mathrm{Ad})$ гомоморфизма Ad в точке e представляет собой линейное отображеиие линеала $\mathfrak{g} = \mathfrak{l}(G)$ в линеал $\mathrm{End}\,\mathfrak{g} = \mathfrak{l}(\mathrm{Aut}\,\mathfrak{g})$ всех линейных операторов линеала $\mathfrak{g}$. С другой стороны, из лекции 3 мы знаем, что для любой алгебры Ли $\mathfrak{g}$ имеется естественное линейное отображение $\mathrm{ad}\colon \mathfrak{g} \to \mathrm{End}\,\mathfrak{g}$, действующее по формуле

$$\mathrm{ad}\,X\colon Y \mapsto [X, Y], \qquad X, Y \in \mathfrak{g}.$$

Предложение 2. *Имеет место равенство*

$$\mathfrak{l}(\mathrm{Ad}) = \mathrm{ad}.$$

Доказательство (ср. с доказательством предложения 1). Так как

$$\begin{aligned}\mathfrak{H}(\mathfrak{H}(X, Y), -X) &= \mathfrak{H}\left(X + Y + \frac{1}{2}[X, Y] + \ldots, -X\right) = \\ &= Y + \frac{1}{2}[X, Y] + \frac{1}{2}[Y, -X] + \ldots = \\ &= Y + [X, Y] + \ldots,\end{aligned}$$

где многоточие обозначает члены не менее чем третьей степени по X и Y, то

$$\begin{aligned}(\exp X)(\exp Y)(\exp X)^{-1} &= \exp \mathfrak{H}(\mathfrak{H}(X, Y), -X) = \\ &= \exp(Y + [X, Y] + \ldots),\end{aligned}$$

и потому

$$\Phi_{\beta_X(s)}\left(\beta_Y(t)\right) = (\exp(sX))(\exp(tY))(\exp(sX))^{-1} = \\ = \exp(tY + st[X, Y] + \ldots),$$

где многоточие обозначает члены не менее чем третьей степени по s и t. Следовательно, нормальные координаты вектора

$$\left(d\Phi_{\beta_X(s)}\right)_e Y = \left(d\Phi_{\beta_X(s)}\right)\left(\frac{d\beta_Y(t)}{dt}\bigg|_{t=0}\right) = \\ = \left(\frac{d}{dt}\Phi_{\beta_X(s)}\left(\beta_Y(t)\right)\right)\bigg|_{t=0}$$

равны

$$\frac{d}{dt}(tY^i + st[X, Y]^i + \ldots)|_{t=0} = Y^i + s[X, Y]^i + \ldots,$$

где последнее многоточие обозначает члены не меньше чем второй степени по s, и значит,

$$\left(d\Phi_{\beta_X(s)}\right)_e Y = Y + s[X, Y] + \ldots$$

Поскольку

$$(d\,\mathrm{Ad})_e X = (d\,\mathrm{Ad})_e\left(\frac{d\beta_X(s)}{ds}\bigg|_{s=0}\right) = \frac{d}{ds}\mathrm{Ad}\,(\beta_X(s))|_{s=0} = \\ = \lim_{s\to 0}\frac{\mathrm{Ad}\,(\beta_X(s)) - \mathrm{Ad}\,(e)}{s} = \lim_{s\to 0}\frac{\left(d\Phi_{\beta_X(s)}\right)_e - E}{s},$$

отсюда следует, что

$$((d\,\mathrm{Ad})_e X)\,Y = \lim_{s\to 0}\frac{\left(d\Phi_{\beta_X(s)}\right)_e Y - Y}{s} = \\ = \lim_{s\to 0}([X, Y] + \ldots) = [X, Y] = \\ = (\mathrm{ad}\,X)\,Y. \quad \square$$

Следствие. *Для любого элемента $X \in \mathfrak{g}$ имеет место равенство*

$$\mathrm{Ad}\,(\exp X) = e^{\mathrm{ad}\,X}.$$

Доказательство. Формулы

$$t \longmapsto \operatorname{Ad}(t \exp X) \quad \text{и} \quad t \longmapsto e^{t \operatorname{ad} X}$$

задают однопараметрические подгруппы группы Ли $\operatorname{Aut}\mathfrak{g}$, имеющие при $t = 0$ один и тот же касательный вектор

$$(d \operatorname{Ad})_e X = \operatorname{ad} X,$$

и потому совпадающие при всех t. $\square$

Еще один пример применения теоремы 1 мы получим, рассмотрев в алгебре Ли $\mathfrak{g} = \mathfrak{l}(G)$ произвольную гладкую кривую $t \longmapsto X(t)$. Для любого $s \in \mathbb{R}$ отображение $t \longmapsto \exp(sX(t))$ представляет собой некоторую кривую в группе G, проходящую при $t = 0$ через точку $a(s) = \exp(sX)$, где $X = X(0)$. Пусть

$$A(s) = \frac{d}{dt} \exp(sX(t))\,|_{t=0}$$

— касательный вектор этой кривой в точке $a(s)$. Перенеся посредством дифференциала $(dR_{a(s)^{-1}})_{a(s)}$ этот вектор в точку e, мы получим в $\mathfrak{g} = \mathsf{T}_e(G)$ вектор

$$B(s) = (dR_{a(s)^{-1}})_{a(s)} A(s).$$

Отображение $s \longmapsto B(s)$ представляет собой гладкую кривую в $\mathfrak{g}$, и потому для любого s определен ее касательный вектор $B'(s)$. Оказывается, что

$$B'(s) = \operatorname{Ad}(a(s))\,Y, \tag{20}$$

где $Y = X'(0)$. Действительно, по определению действия дифференциала гладкого отображения на касательные векторы кривых

$$B(s) = \frac{d}{dt}\left(R_{a(s)^{-1}}(\exp(sX(t)))\right)|_{t=0} =$$

$$= \frac{d}{dt}(\exp(sX(t)))\exp(-sX)\,|_{t=0}$$

и, следовательно, согласно замечанию 3,

$$B(s+\Delta s)-B(s)=\frac{d}{dt}((\exp(sX(t))\exp(-sX))^{-1}\times$$
$$\times(\exp((s+\Delta s)X(t))\exp(-(s+\Delta s)X))|_{t=0}=$$
$$=\frac{d}{dt}(\exp(sX)\exp(-sX(t))\times$$
$$\times\exp((s+\Delta s)X(t))\exp(-(s+\Delta s)X))|_{t=0}=$$
$$=\frac{d}{dt}\left(a(s)\exp(\Delta sX(t))a(s+\Delta s)^{-1}\right)=$$
$$=\left(dL_{a(s)}\circ dR_{a(s+\Delta s)^{-1}}\right)\frac{d}{dt}(\exp(\Delta sX(t)))|_{t=0}.$$

Поэтому

$$B'(s)=\lim_{\Delta s\to 0}\frac{B(s+\Delta s)-B(s)}{\Delta s}=$$
$$=\left(dL_{a(s)}\circ dR_{a(s)^{-1}}\right)\lim_{\Delta s\to 0}\frac{\frac{d}{dt}(\exp(\Delta sX(t)))|_{t=0}}{\Delta s}=$$
$$=\operatorname{Ad}(a(s))\lim_{\Delta s\to 0}\frac{\frac{d}{dt}(\exp(\Delta s\,X(t)))|_{t=0}}{\Delta s}$$

и, значит, для доказательства равенства (20) достаточно доказать, что

$$\lim_{\Delta s\to 0}\frac{\frac{d}{dt}(\exp(\Delta sX(t)))|_{t=0}}{\Delta s}=Y.$$

Но это очевидно, поскольку в нормальных координатах, соответствующих произвольному базису $e_1, \ldots, e_n$ линеала $\mathfrak{g}$, точка $\exp(\Delta sX(t))$ имеет координаты $\Delta sX^i(t)$, где $X^i(t)$ — координаты вектора $X(t)$ в базисе $e_1, \ldots, e_n$, и, значит, вектор

$$\lim_{\Delta s\to 0}\frac{\frac{d}{dt}(\exp(\Delta sX(t))|_{t=t_0}}{\Delta s}$$

имеет в базисе $e_1, \ldots, e_n$ координаты $\frac{dX^i(0)}{dt}$, т. е. те же координаты, что и вектор $X'(0)=Y$. $\square$

Участвующий в формуле (20) линейный оператор $\operatorname{Ad}(a(s))$ мы можем переписать в следующем виде:

$$\operatorname{Ad}(a(s))=\operatorname{Ad}(\exp(sX))=e^{s\operatorname{ad}X}=$$
$$=E+s\operatorname{ad}X+\ldots+s^n\frac{(\operatorname{ad}X)^n}{n!}+\ldots$$

Интегрируя это операторное тождество по s от 0 до 1, мы получим тождество

$$\int_0^1 \mathrm{Ad}\,(a\,(s))\,ds = \frac{e^{\mathrm{ad}\,X} - E}{\mathrm{ad}\,X},$$

где под $\frac{e^{\mathrm{ad}\,X} - E}{\mathrm{ad}\,X}$ понимается сумма операторного ряда

$$E + \frac{\mathrm{ad}\,X}{2!} + \ldots + \frac{(\mathrm{ad}\,X)^n}{(n+1)!} + \ldots,$$

получающегося из степенного ряда для функции $\frac{e^z - 1}{z}$ подстановкой вместо z оператора $\mathrm{ad}\,X$. Для вектора $B\,(1)$ отсюда в силу формулы (20) следует, что

$$B\,(1) = \int_0^1 B'\,(s)\,ds = \frac{e^{\mathrm{ad}} - E}{\mathrm{ad}\,X}\,Y.$$

Поскольку по определению

$$\begin{aligned} B\,(1) &= \left(dR_{a\,(1)^{-1}}\right)_{a\,(1)} \frac{d}{dt} \exp\,(X\,(t)\,|_{t=0} = \\ &= (dR_{\exp\,(-X)})_{\exp X} \frac{d}{dt} \exp X\,(t)\,|_{t=0} = \\ &= \frac{d}{dt}\,(\exp X\,(t) \exp\,(-X))\,|_{t=0}, \end{aligned}$$

этим доказано, что

$$\frac{d}{dt}\,(\exp X\,(t) \exp\,(-X))\,|_{t=0} = \frac{e^{\mathrm{ad}\,X} - E}{\mathrm{ad}\,X}\,Y.$$

Переходя к нормальным координатам, мы немедленно получаем отсюда, что для вектора $Z\,(t) \in \mathfrak{g}$, удовлетворяющего соотношению $\exp X\,(t) \exp\,(-X) = \exp Z\,(t)$, справедливо равенство

$$Z'\,(0) = \frac{e^{\mathrm{ad}\,X} - E}{\mathrm{ad}\,X}\,Y,$$

откуда следует, что

$$Z\,(t) = t\,\frac{e^{\mathrm{ad}\,X} - E}{\mathrm{ad}\,X}\,Y + O\,(t^2).$$

Возвращаясь к $\exp Z(t)$ и полагая $X(t) = X + tY$, мы видим, что нами доказано

Предложение 3. *Для любых элементов $X, Y \in \mathfrak{g}$ имеет место равенство*

$$\exp(X + tY)\exp(-X) = \exp\left(t\frac{e^{\operatorname{ad} X} - E}{\operatorname{ad} X}Y + O(t^2)\right). \quad \square$$

Для каждого элемента $X \in \mathfrak{g}$ дифференциал $(d\exp)_X$ в точке X гладкого отображения $\exp\colon \mathfrak{g} \to G$ в силу отождествления $\mathsf{T}_X(\mathfrak{g}) = \mathfrak{g}$ представляет собой линейное отображение $\mathfrak{g} \to T_a(G)$, где $a = \exp X$. Поэтому его композиция с отображением $(dR_a)_e^{-1}\colon \mathsf{T}_a(G) = \mathsf{T}_e(G) = \mathfrak{g}$ будет отображением из $\mathfrak{g}$ в $\mathfrak{g}$

Следствие 1. *Имеет место формула*

$$(dR_a)_e^{-1} \circ (d\exp)_X = \frac{e^{\operatorname{ad} X} - E}{\operatorname{ad} X}, \quad a = \exp X.$$

Доказательство. Пусть $Y \in \mathfrak{g}$. Так как $\left.\frac{d(X + tY)}{dt}\right|_{t=0} = Y$, то

$$((dR_a^{-1})_e \circ (d\exp)_X)Y = \frac{d}{dt}(\exp(X + tY)\exp(-X))|_{t=0} =$$

$$= \frac{d}{dt}\left(t\frac{e^{\operatorname{ad} X} - E}{\operatorname{ad} X}Y + O(t^2)\right)\Big|_{t=0} =$$

$$= \frac{e^{\operatorname{ad} X} - E}{\operatorname{ad} X}Y. \quad \square$$

Следствие 2. *Отображение* $\exp\colon \mathfrak{g} \to G$ *тогда и только тогда является диффеоморфизмом в точке $X \in \mathfrak{g}$, когда ни один характеристический корень оператора* $\operatorname{ad} X$ *не имеет вида $2m\pi i$.*

Доказательство. Отображение $\exp$ тогда и только тогда является диффеоморфизмом в точке X, когда его дифференциал $(d\exp)_X$ в этой точке является изоморфизмом, а оператор $\operatorname{ad} X$ тогда и только тогда имеет характеристические корни вида $2m\pi i$, когда оператор $e^{\operatorname{ad} X} - E$, а значит, и оператор $\frac{e^{\operatorname{ad} X} - E}{\operatorname{ad} X}$ вырожден. $\square$

Подчеркнем, что все эти результаты нами доказаны лишь по модулю утверждений А и В. Поэтому нашей ближайшей целью должно быть доказательство этих утверждений.

Мы докажем сначала утверждение А и притом в несколько более общем виде.

Пусть линейное пространство $\mathfrak{g} = \mathfrak{l}(G)$ разложено в прямую сумму

$$\mathfrak{g} = \mathscr{A} \oplus \mathscr{B} \tag{21}$$

двух подпространств $\mathscr{A}$ и $\mathscr{B}$. Определим отображение

$$\Phi\colon \mathfrak{g} \to G, \tag{22}$$

полагая для любого $X \in \mathfrak{g}$

$$\Phi(X) = \exp A \cdot \exp B,$$

где $A \in \mathscr{A}$ и $B \in \mathscr{B}$ — компоненты вектора $X \in \mathfrak{g}$ в разложении (21). Очевидно, что это отображение гладко и переводит нуль $0 \in \mathfrak{l}(G)$ в единицу $e \in G$. Найдем дифференциал

$$(d\Phi)_0\colon \mathsf{T}_0(\mathfrak{g}) \to \mathsf{T}_e(G)$$

этого отображения в точке 0.

Пусть

$$l\colon \mathfrak{g} \to \mathsf{T}_0(\mathfrak{g})$$

— естественный изоморфизм, переводящий вектор $X \in \mathfrak{g}$ в вектор, касающийся в точке 0 кривой $t \mapsto tX$. Отображение Φ переводит эту кривую в кривую

$$t \mapsto \exp tA \cdot \exp tB = \exp \mathfrak{J}(tA,\ tB), \tag{23}$$

и, следовательно, его дифференциал $(d\Phi)_0$ переводит вектор $l(X)$ в вектор, касающийся кривой (23) в точке e. Это означает, что для любой функции $f \in \mathcal{O}_e(G)$ имеет место формула

$$[((d\Phi)_0 \circ l)(X)]\, f = \left.\frac{df\left(\exp \mathfrak{J}(tA,\ tB)\right)}{dt}\right|_{t=0}$$

и, значит, (см. формулу (4)), формула

$$[((d\Phi)_0 \circ l)(X)]\, f = \left.\frac{d\left(e^{\mathfrak{J}(tA,\ tB)}f\right)(e)}{dt}\right|_{t=0}.$$

Но

$$e^{\mathfrak{I}(tA,\,tB)}f = (E + \mathfrak{I}(tA,\,tB) + O(t^2))f = \\ = f + t(A+B)f + O(t^2),$$

и потому

$$\left.\frac{d\left(e^{\mathfrak{I}(tA,\,tB)}f\right)(e)}{dt}\right|_{t=0} = ((A+B)f)(e) = (Xf)(e) = X_e f.$$

Следовательно,

$$((d\Phi)_0 \circ l)(X) = X_e,$$

т. е.

$$(d\Phi)_0 \circ l = i,$$

где i — изоморфизм $X \to X_e$ линеала $\mathfrak{g} = \mathrm{l}(G)$ на линеал $\mathrm{T}_e(G)$.

Поскольку i и l являются изоморфизмами, отсюда следует, что отображение $(d\Phi)_0$ также является изоморфизмом. В силу стандартной теоремы об этальных (локально диффеоморфных) отображениях этим доказано следующее предложение:

Предложение 4. *Отображение* (22) *является диффеоморфизмом в точке* $0 \in \mathfrak{g}$. $\square$

При $\mathfrak{g} \in \mathscr{A}$ (и $\mathscr{B} = 0$) мы получаем утверждение А, которое тем самым полностью доказано.

Согласно предложению 4 точка $0 \in \mathfrak{g}$ обладает сколь угодно малой звездной окрестностью $\mathring{U}$, на которой отображение (22) является ее диффеоморфизмом на некоторую окрестность U единицы $e \in G$.

Определение 2. Обладающие этим свойством окрестности $\mathring{U}$ и U называются *каноническими окрестностями* (точек $0 \in \mathfrak{g}$ и $e \in G$ соответственно), отвечающими данному прямому разложению (21).

Выбрав в подпространствах $\mathscr{A}$ и $\mathscr{B}$ произвольные базисы, мы получим некоторый базис пространства $\mathfrak{g}$. Композиция h диффеоморфизма $\Phi^{-1}\colon U \to \mathring{U}$ с ограничением на $\mathring{U}$ соответствующего координатного изоморфизма $\mathfrak{g} \to \mathbb{R}^n$ будет диффеоморфизмом окрестности U на некоторое открытое множество пространства $\mathbb{R}^n$, т. е. пара (U, h) будет некоторой картой на G.

Карты такого вида называются *каноническими картами*, отвечающими разложению (21), а соответствующие локальные координаты $x^1, \ldots, x^n$ — *каноническими координатами.*

Ясно, что все эти определения (вместе с предложением 4) автоматически переносятся на случай, когда задано разложение

$$\mathfrak{g} = \mathscr{A}_1 \oplus \ldots \oplus \mathscr{A}_m \tag{24}$$

линеала $\mathfrak{g}$ в прямую сумму любого числа подпространств.

Если в разложении (21) $\mathscr{A} = \mathfrak{g}$ и $\mathscr{B} = 0$, т. е. если в разложении (24) $m = 1$, то канонические окрестности совпадают с нормальными окрестностями в смысле определения 1, а канонические координаты — с нормальными координатами.

Другой крайний случай возникает при $m = n$, когда все пространства $\mathscr{A}_1, \ldots, \mathscr{A}_m$ одномерны (т. е. когда разложение (24) определяется выбором в $\mathfrak{g}$ некоторого базиса). Соответствующие канонические координаты называются *каноническими координатами второго рода* (тогда как *каноническими координатами первого рода* называются нормальные координаты).

Заметим, что канонические координаты первого и второго рода задаются произвольным базисом линеала $\mathfrak{g}$.

С помощью канонических координат легко доказывается следующее важное предложение:

***Предложение* 5.** *Любой непрерывный гомоморфизм $\Phi\colon H \to G$ групп Ли (т. е. их гомоморфизм как топологических групп) является гладким отображением (их гомоморфизмом как групп Ли).*

Доказательство. Пусть $a \in H$ и $g \in \mathcal{O}_{\Phi(a)}(G)$. Нам нужно доказать, что функция $g \circ \Phi$, определенная в окрестности точки a, является гладкой функцией в этой точке, т. е. принадлежит $\mathcal{O}_a(H)$. Но, поскольку Φ является гомоморфизмом, справедливо равенство

$$\Phi = L_{\Phi a}^{-1} \circ \Phi \circ L_a.$$

Поэтому, так как отображение L_a и функция $f = g \circ L_{\Phi a}^{-1}$ гладки, то нам достаточно доказать, что для любой функции $f \in \mathcal{O}_e(G)$ функция $f \circ \Phi$ гладка в окрестности единицы $e \in H$, т. е. что отображение Φ гладко в e.

Рассмотрим сначала случай, когда H является аддитивной группой $\mathbb{R}$ вещественных чисел. Пусть U — нормальная (каноническая первого рода) окрестность единицы e в группе G, и пусть $x^1, \ldots, x^n$ — соответствующие нормальные координаты. Пусть, далее, $\varepsilon > 0$ — такое число, что $\Phi(t) \in U$, когда $|t| < \varepsilon$. Если $0 < t < \varepsilon$ и $0 < |r| < s$, где r и s — целые числа, то

$$\Phi\left(\frac{r}{s}t\right) = \Phi\left(\frac{t}{s}\right)^r \quad \text{и} \quad \Phi(t) = \Phi\left(\frac{t}{s}\right)^s.$$

С другой стороны, поскольку координаты $x^1, \ldots, x^n$ являются нормальными, то для любого $i = 1, \ldots, n$, любого элемента $a \in U$ и любого s такого, что $a^s \in U$, имеет место равенство

$$x^i(a^s) = sx^i(a)$$

(достаточно заметить, что если $a = \exp A$, то $a^s = \exp(sA)$). Следовательно,

$$(x^i \circ \Phi)\left(\frac{r}{s}t\right) = r \cdot (x^i \circ \Phi)\left(\frac{t}{s}\right)$$

и

$$(x^i \circ \Phi)(t) = s \cdot (x^i \circ \Phi)\left(\frac{t}{s}\right).$$

Поэтому

$$(x^i \circ \Phi)\left(\frac{r}{s}t\right) = \frac{r}{s} \cdot (x^i \circ \Phi)(t).$$

В силу непрерывности отображения Φ аналогичное равенство

$$(x^i \circ \Phi)(\alpha t) = \alpha \cdot (x^i \circ \Phi)(t)$$

справедливо и при любом вещественном α, $|\alpha| < 1$. Это означает, что функции $x^i \circ \Phi$ линейны. Таким образом, отображение Φ задается в локальных координатах линейными и, следовательно, аналитическими функциями. Поэтому оно гладко.

Пусть теперь группа H произвольна. Выберем в ее алгебре Ли $\mathfrak{h}$ произвольный базис $Y_1, \ldots, Y_n$. Тогда для любого $i = 1, \ldots, n$ отображение $t \mapsto \Phi(\exp tY_i)$ будет непрерывным, а следовательно, по уже доказанному и гладким гомоморфизмом $\mathbb{R} \to G$, т. е. будет однопараметрической подгруппой группы G. Поэтому в алгебре Ли $\mathfrak{g}$ группы G существуют такие элементы $X_1, \ldots, X_n$, что

$$\Phi(\exp tY_i) = \exp tX_i$$

для любого $i = 1, \ldots, n$. Поскольку Φ является гомоморфизмом, отсюда следует, что для любых чисел $t^1, \ldots$ $\ldots, t^n \in \mathbb{R}$ справедливо равенство

$$\Phi(\exp t^1 Y_1 \ldots \exp t^n Y_n) = \exp t^1 X_1 \ldots \exp t^n X_n.$$

Ясно, что элемент $\exp t^1 X_1 \ldots \exp t^n X_n$ группы G гладко зависит от $t^1, \ldots, t^n$, т. е. его локальные координаты $x^1, \ldots, x^n$ (в произвольной карте) являются гладкими функциями $x^1(t), \ldots, x^n(t)$ от $t = (t^1, \ldots, t^n)$. Но, согласно определению, числа $t^1, \ldots, t^n$ являются (в некоторой окрестности единицы группы H) не чем иным, как каноническими координатами второго рода, определенными данным базисом алгебры $\mathfrak{h}$. Поэтому функции $x^1(t), \ldots, x^n(t)$ представляют собой функции, задающие отображение Φ в локальных координатах $t^1, \ldots, t^n$ и $x^1, \ldots, x^n$. Поскольку эти функции гладки, отображение Φ гладко. □

Следствие. *Если две группы Ли изоморфны как топологические группы, то они изоморфны и как группы Ли.* □

В частности, отсюда следует, что *если на топологической группе можно ввести согласованную с топологией гладкость так, чтобы она стала группой Ли, то это можно сделать только одним способом.*

Это означает, что функтор игнорирования

$$\text{GR-DIFF} \to \text{GR-TOP}$$

переводит различные группы Ли в различные топологические группы. Поэтому мы можем считать, что этот функтор осуществляет в л о ж е н и е категории GR-DIFF в категорию GR-TOP. Иными словами, категорию GR-DIFF групп Ли мы можем считать подкатегорией категории GR-TOP всех топологических групп. Тогда предложение 5 будет означать, что эта подкатегория полна.

Теперь естественно возникает вопрос, а нельзя ли подкатегорию GR-DIFF охарактеризовать в н у т р и категории GR-TOP общетопологическими условиями без привлечения гладкостей, т. е. нельзя ли внутри категории GR-TOP охарактеризовать топологические группы G, допускающие структуру группы Ли?

Одно необходимое условие очевидно: топологическая группа G, допускающая структуру группы Ли, обязательно должна быть хаусдорфова и локально компактна. Чтобы сформулировать более тонкое необходимое условие, мы введем следующее определение:

Определение 3. Топологическая группа G называется группой *без малых подгрупп*, если ее единица e обладает окрестностью, не содержащей никаких подгрупп $H \neq \{e\}$.

Оказывается, что это свойство необходимо для того, чтобы в G можно было ввести структуру группы Ли.

Предложение 6. *Каждая группа Ли G (точнее, топологическая группа G_{top}, получающаяся из группы G игнорированием гладкости) является группой без малых подгрупп.*

Доказательство. Введем на линейном пространстве $\mathsf{T}_e(G) = \mathfrak{g}$ произвольную евклидову метрику. Тогда для достаточно малого $\delta > 0$ шар радиуса δ с центром в точке 0 будет нормальной окрестностью этой точки, и, следовательно, его образ при отображении exp будет нормальной окрестностью единицы в группе G. Пусть U — нормальная окрестность, аналогичным образом строящаяся по числу $\delta/2$. Ясно, что для любого отличного от нуля вектора $A \in \mathfrak{g}$, длина которого меньше $\delta/2$, существует такое целое число m, что длина вектора mA больше $\delta/2$ и меньше δ. Это означает, что для любого отличного от единицы элемента $a = \exp A$ окрестности U существует такое m, что $a^m = \exp mA$ не принадлежит U. Поэтому окрестность U не может содержать никакой подгруппы $H \neq \{e\}$. □

Замечательно, что вместе с условием локальной компактности условие отсутствия малых подгрупп уже достаточно для того, чтобы на хаусдорфовой группе G можно было ввести структуру группы Ли (по доказанному, единственную).

Теорема (Глисон и Ямабе). *Топологическая хаусдорфова группа тогда и только тогда является группой Ли, когда она локально компактна и не имеет малых подгрупп.*

Доказательство теоремы Глисона — Ямабе слишком сложно, чтобы мы могли его здесь изложить.

Другое необходимое условие лиевости топологической группы состоит в ее *локальной евклидовости*, т. е. в том, чтобы она была топологическим многообразием. Вопрос о том, является ли это необходимое условие достаточным, составляет содержание так называемой пятой проблемы Гильберта (в ее современной формулировке). На основе тщательного исследования строения топологических групп с малыми подгруппами было установлено (Монтгомери, Циппин, Ивасава, Глисон, Ямабе), что *никакая локально евклидова группа малых подгрупп иметь не может*. В комбинации с теоремой Глисона — Ямабе это немедленно дает положительное решение проблемы Гильберта: *любая локально евклидова группа является группой Ли.*

Подробности см. в [2] и [6].