

Лекция 1

Геометрия и физика специальной теории относительности

Специальная теория относительности основана на геометрии пространства-времени Минковского $M^4 = \mathbb{R}^{1,3}$ или, более обще $M^d = \mathbb{R}^{1,d-1}$.¹ Это пространство-время представляет собой аффинное пространство с заданной на нем плоской метрикой сигнатуры $(1, d-1)$. Чтобы разобраться в этом, начнем с векторных (линейных) пространств.

Пусть V — d -мерное линейное пространство. С каждым таким пространством связано d -мерное *двойственное пространство* V^* , то есть пространство линейных функций (форм) на пространстве V . Предположим, что на пространстве V задана невырожденная симметричная билинейная форма $g : V \otimes V \rightarrow \mathbb{R}$ (или, что то же самое, $g \in V^* \otimes V^*$), то есть билинейное симметричное отображение $g(u, v) = g(v, u)$ ($u, v \in V$), такая, что для любого вектора $u \neq 0$ существует хотя бы один вектор v , такой что $g(u, v) \neq 0$. Форму g можно также понимать как отображение $\bar{g} : V \rightarrow V^*$, сопоставляющее каждому вектору $u \in V$ линейную форму $g(u, \cdot) \in V^*$. Невырожденность формы g означает существование обратного отображения $\bar{g}^* : V^* \rightarrow V$, которое по тому же принципу задает билинейную форму $g^* : V^* \otimes V^* \rightarrow \mathbb{R}$, то есть бивектор $g^* \in V \otimes V$.

Изучим вопрос о знакоопределенности формы g . Предположим, что $g(u, u) \geq 0$ для любого $u \neq 0$. Тогда из невырожденности формы g следует, что $g(u, u) > 0$. Действительно, пусть есть такой ненулевой вектор u , что $g(u, u) = 0$. В силу невырожденности g существует такой вектор v , что $g(u, v) \neq 0$ и, по условию, $g(v, v) \geq 0$. Тогда для любого $\alpha \in \mathbb{R}$ имеем

$$0 \leq g(u + \alpha v, u + \alpha v) = 2\alpha g(u, v) + \alpha^2 g(v, v).$$

Очевидно, это неравенство не может быть соблюдено в некоторой окрестности точки $\alpha = 0$ в силу того, что $g(u, v) \neq 0$. Таким образом, любая неотрицательная форма положительно определена.

Но это значит, что если существует «нулевой» вектор u , то есть вектор, такой что $g(u, u) = 0$, то существуют и такие два вектора v_+, v_- , что $g(v_+, v_+) > 0$, а $g(v_-, v_-) < 0$. Более того, любой «нулевой» вектор u может быть представлен как линейная комбинация двух таких векторов с разными знаками формы. Пусть имеется вектор v_+ , такой что $g(v_+, v_+) = 1$, $g(u, v_+) \neq 0$. Тогда для вектора $v_- = u - g(u, v_+)v_+$ имеем $g(v_-, v_-) = -(g(u, v_+))^2$. Ни множество векторов \mathcal{C}_+ с $g(u, u) > 0$, ни множество векторов \mathcal{C}_- с $g(u, u) < 0$, ни множество векторов \mathcal{J} с $g(u, u) = 0$ не образуют векторных пространств. Но тот факт, что «нулевые» векторы раскладываются в пару «положительных» и «отрицательных» векторов, позволяет предположить, что и само векторное пространство V распадается в сумму $V = V_+ \oplus V_-$ пространств с положительно- и отрицательно-определенной формой g .

Чтобы в этом убедиться, построим базис в V_- . Пусть v_1 — вектор с $g(v_1, v_1) < 0$. Тогда множество ортогональных ему векторов, то есть векторов u , таких что $g(v_1, u) = 0$ образует линейное пространство $V^{(1)}$. В пространстве $V^{(1)}$ форма g может быть положительно определена. Тогда $V_+ = V^{(1)}$, а $V_- = \text{span}(v_1)$. Если же нет, выберем вектор $v_2 \in V^{(1)}$ такой, что $g(v_2, v_2) < 0$ и определяем пространство $V^{(2)}$ как подпространство ортогональных ему векторов в $V^{(1)}$. Продолжаем процедуру, пока на пространстве $V^{(d-1)}$ форма g не будет положительно определена. Тогда $V_+ = V^{(d-1)}$ размерности $d_+ = d - d_-$, а $V_- = \text{span}(v_1, \dots, v_{d-1})$. Заодно мы получили ортогональный базис в V_- , который легко превратить в ортонормированный, если выбирать векторы v_i условием $g(v_i, v_i) = -1$. Мы будем говорить, что форма g имеет *сигнатуру* (d_+, d_-) . В теории относительности нас будут интересовать формы сигнатуры $(1, 3)$.² Одномерное пространство V^+ отвечает времени, а трехмерное пространство V^- — физическому пространству. Векторы из \mathcal{C}^+ (с положительной нормой) называют *временеподобными*, векторы из \mathcal{C}^- (с отрицательной нормой) — *пространственноподобными*. Векторы из \mathcal{J} (с нулевой нормой) называют *светоподобными* или *изотропными*. Само множество \mathcal{J} называют *световым конусом*, причем \mathcal{C}^+ считается внутренностью светового конуса, а \mathcal{C}^- — его внешней областью. В теории относительности светоподобный базисный вектор принято обозначать e_0 , а три пространственноподобных — e_1, e_2, e_3 . Во всех векторных и матричных формулах будем принимать, что значения индексов расположены в порядке $0, 1, 2, 3$.

¹Нас, конечно, будет интересовать прежде всего случай $d = 4$, однако там, где это не будет заметно усложнять рассуждения, мы будем рассматривать общие $d > 1$.

²Можно было бы рассматривать формы сигнатуры $(3, 1)$. В некоторых отношениях такой выбор удобнее, а в чем-то менее удобен, так что это дело вкуса.

Рассмотрим некоторый базис e_μ , $\mu = 1, \dots, d$ в пространстве V . Любой вектор $u \in V$ записывается в виде $u = u^\mu e_\mu$, где мы предполагаем суммирование по повторяющемуся индексу. Тогда числа $g_{\mu\nu} = g(e_\mu, e_\nu)$ образуют симметричную матрицу $G = (g_{\mu\nu})_{\mu,\nu=1}^d$. Замена базиса $e_\mu = V_\mu^\nu e'_\nu$ отвечает преобразованию $g_{\mu\nu} = V_\mu^\rho V_\nu^\sigma g'_{\rho\sigma}$ или $G = VG'V^T$. Тот факт, что пространство V распадается в сумму $V_+ \oplus V_-$ означает, что таким преобразованием мы можем привести матрицу G к виду

$$G = \text{diag}(\underbrace{1, \dots, 1}_{d_+}, \underbrace{-1, \dots, -1}_{d_-}) \equiv (\eta_{\mu\nu})_{\mu,\nu=1}^d \equiv H^{(d_+, d_-)}. \quad (1.1)$$

Матрицы V , сохраняющие этот вид формы G , то есть

$$H^{(d_+, d_-)} = VH^{(d_+, d_-)}V^T,$$

образуют группу псевдоортогональных матриц $O(d_+, d_-)$. Те, которые при этом имеют положительный (равный единице) определитель, образуют группу $SO(d_+, d_-)$. Группа $SO(1, 3)$ (и ее обобщение $SO(1, d-1)$ в гипотетическом пространстве-времени размерности d) называется *группой Лоренца*. В теории относительности будем писать $\eta_{00} = -\eta_{11} = -\eta_{22} = -\eta_{33} = 1$.

Бивектор g^* равен, очевидно, $g^{\mu\nu} e_\mu \otimes e_\nu$, где коэффициенты $g^{\mu\nu}$ образуют матрицу G^{-1} :

$$g^{\mu\lambda} g_{\lambda\nu} = \delta_\nu^\mu. \quad (1.2)$$

Использование индексных обозначений компактно, но имеет один недостаток: мы заменяем реальные инвариантные объекты (тензоры) их компонентами. В то же время безындексная запись математически корректна, обща, но громоздка. Поэтому мы часто будем пользоваться *методом формальных индексов*. Рассмотрим тензорное произведение $V \otimes V \otimes V \otimes \dots$ и перенумеруем входящие в него пространства натуральными числами по порядку. Будем соответствующим числом жирным шрифтом указывать, в каком пространстве расположен или на какое пространство действует объект. Например, u^k — это вектор u , расположенный в k -й тензорной компоненте, а ω_k — 1-форма, действующая на k -ю компоненту. Тогда $\omega_1 u^1 = u^1 \omega_1 = \omega u = \omega(u)$. Здесь от формального индекса, вроде бы, нет никакой пользы. Но вот $g_{12} u^1 v^2 = g(u, v)$ и $g_{12} u^1 = (\bar{g}(u))_2$ записать с формальными индексами уже удобней. И все четыре объекта $g, \bar{g}, g^*, \bar{g}^*$ можно теперь обозначать одной буквой g . Действительно,

$$g^{13} g_{32} = \delta_2^1 \quad (1.3)$$

формально похоже на (1.2), но выражает инвариантный факт, что $\bar{g}^* = \bar{g}^{-1}$. На общих многообразиях запись с формальными индексами будет значительно упрощать вывод явных формул в компонентах, не мешая пониманию инвариантного смысла выражений.

Нам часто удобно будет отождествлять вектор $u \in V$ с формой $\bar{g}(u) \in V^*$ и, наоборот, форму $\omega \in V^*$ с вектором $\bar{g}^*(\omega) \in V$. Чтобы подчеркнуть разницу мы будем писать верхний формальный индекс у объекта из V и нижний у объекта из V^* . Если не будет необходимости указывать конкретную «цифру», но нужно будет подчеркнуть, какому пространству принадлежит объект, мы будем вместо индекса ставить жирную точку, например, $u^\bullet = u$, $u_\bullet = \bar{g}(u)$ (то есть $u_1 = g_{12} u^2$).

Теперь перейдем к аффинным пространствам. Аффинное пространство отличается от линейного только отсутствием выделенного нуля. То есть это множество точек, такое что каждой паре точек A, B отвечает вектор в линейном пространстве, обозначаемый $\overrightarrow{AB} = -\overrightarrow{BA}$ а для любых трех точек A, B, C выполняется правило треугольника $\overrightarrow{AC} = \overrightarrow{AB} + \overrightarrow{BC}$. Форма g определяет *метрику* на аффинном пространстве, то есть расстояние $|AB|$ между точками A и B задается формулой

$$|AB|^2 = g(\overrightarrow{AB}, \overrightarrow{AB}).$$

Понятно, что определенное таким образом расстояние может быть как вещественным, так и мнимым. В теории относительности величина $|AB|$ называется *собственным временем*, отвечающим отрезку $[AB]$. Точки в аффинном пространстве-времени называются *событиями*.

В аффинном пространстве точки можно задать координатами. Если мы выделим «начальную» точку O и базис $\{e_\mu\}$ в пространстве V , то координаты x_A^μ точки A задаются уравнением

$$\overrightarrow{OA} = x_A^\mu e_\mu.$$

С одной стороны, это отождествляет аффинное пространство с пространством \mathbb{R}^d , которое, в случае метрики $g = \eta$ обозначается как $\mathbb{R}^{d+,d-}$. С другой стороны, это вводит на аффинном пространстве структуру многообразия. Понятно, что аффинное пространство покрывается одной картой, но специальная теория относительности имеет смысл на многообразиях с разной топологией. Главное, чтобы матрицы перехода между картами с (псевдо)декартовыми координатами задавались бы псевдоортогональными матрицами. Обозначения и понятия, связанные со структурой многообразия мы разберем в следующей лекции, посвященной более общей геометрии.

В теории относительности мы иногда будем пользоваться обозначениями $x^0 = t, x^1 = x, x^2 = y, x^3 = z$. Скорость света мы будем всюду принимать равной единице. Соответственно, квадрат расстояния (собственного времени) между событиями в специальной теории относительности равен

$$|AB|^2 = (t_A - t_B)^2 - (x_A - x_B)^2 - (y_A - y_B)^2 - (z_A - z_B)^2. \quad (1.4)$$

Для бесконечно-малых расстояний собственное время обозначается ds и равно

$$ds^2 = dt^2 - dx^2 - dy^2 - dz^2. \quad (1.5)$$

Обсудим группу Лоренца $SO(1,3)$ и ее физическую интерпретацию. Рассмотрим две системы отсчета K и K' и связанные с ними системы пространственных координат x^i и x'^i ($i = 1, 2, 3$). Пространственными вращениями $SO(3) \subset SO(1,3)$ мы можем добиться того, чтобы базисные векторы e_i и e'_i совпали. Более того, мы можем сделать так, чтобы скорость системы отсчета K' в системе отсчета K была направлена по одной из осей, например вдоль e_1 . Тогда преобразование (*преобразование Лоренца*)

$$t' = \frac{t - vx}{\sqrt{1 - v^2}}, \quad x' = \frac{x - vt}{\sqrt{1 - v^2}}, \quad y' = y, \quad z' = z. \quad (1.6)$$

принадлежит группе Лоренца. Три вращения и преобразование Лоренца порождают всю группу Лоренца $SO(1,3)$. Преобразования Лоренца, также как и все остальные факты релятивистской кинематики, полностью выводятся из следующего постулата: *Квадрат собственного времени $|AB|^2$ для двух событий A и B , определенный формулой (1.4), не зависит от системы отсчета.*

Полезно ввести *быстроту* θ уравнением $v = \text{th } \theta$. Тогда преобразование Лоренца принимает вид

$$t' = t \text{ch } \theta - x \text{sh } \theta, \quad x' = -t \text{sh } \theta + x \text{ch } \theta, \quad y' = y, \quad z' = z, \quad (1.7)$$

напоминающий вращение в евклидовом пространстве. Легко проверить, что оно прямо переходит во вращение на комплексный угол $-i\theta$ после *викова поворота* $t = -i\tau$ времени в комплексной плоскости.

Рассмотрим теперь динамику специальной теории относительности. Нас будет интересовать поведение частиц в геометрии Минковского.

1. Свободная частица. Движение частицы описывается *мировой линией*, то есть линией $x^\bullet(t) = (t, \mathbf{r}(t)) = te_0 + x^i(t)e_i$, где $\mathbf{r}(t) = (x(t), y(t), z(t)) = x^i(t)e_i$ — закон движения частицы. Свободная частица согласно первому закону Ньютона движется равномерно и прямолинейно. То есть, если имеются два события A и B , то мировая линия свободной частицы, в момент t_A находящейся в точке $\mathbf{r}_A = (x_A, y_A, z_A)$, а в момент t_B — в точке $\mathbf{r}_B = (x_B, y_B, z_B)$, будет отрезком $[AB]$. Такая линия минимизирует действие

$$S[x] = -m \int_A^B ds = -m \int_{t_A}^{t_B} dt \sqrt{1 - \mathbf{v}^2}. \quad (1.8)$$

В нерелятивистском пределе $v \ll 1$ действие сводится к интегралу от кинетической энергии

$$S[x] \simeq -m \int_{t_A}^{t_B} dt \left(1 - \frac{\mathbf{v}^2}{2} \right) = \text{const} + \int_{t_A}^{t_B} dt \frac{m\mathbf{v}^2}{2}. \quad (1.9)$$

Уравнение движения, отвечающее действию (1.8), очевидно, имеет вид

$$\dot{\mathbf{v}} = 0. \quad (1.10)$$

Тем не менее, полезно найти импульсы и гамильтониан системы:

$$\mathbf{p} = \frac{\partial L}{\partial \mathbf{v}} = \frac{m\mathbf{v}}{\sqrt{1 - \mathbf{v}^2}}, \quad H(\mathbf{r}, \mathbf{p}) = \mathbf{v}\mathbf{p} - L = \frac{m}{\sqrt{1 - \mathbf{v}^2}} = \sqrt{m^2 + \mathbf{p}^2}. \quad (1.11)$$

Таким образом, в гамильтоновой форме уравнение движения имеет вид

$$\dot{\mathbf{p}} = 0, \quad \dot{\mathbf{r}} = \frac{\mathbf{p}}{E}, \quad E = \sqrt{m^2 + \mathbf{p}^2}. \quad (1.12)$$

Энергия E и импульс \mathbf{p} образуют 4-импульс

$$p^\bullet = (E, \mathbf{p}) = mu^\bullet, \quad (1.13)$$

где u — 4-скорость

$$u^\bullet = \frac{dx^\bullet}{ds} = \left(\frac{1}{\sqrt{1 - \mathbf{v}^2}}, \frac{\mathbf{v}}{\sqrt{1 - \mathbf{v}^2}} \right). \quad (1.14)$$

Величины

$$J^{\mu\nu} = x^\mu p^\nu - x^\nu p^\mu \quad (1.15)$$

тоже сохраняются и дают пространственно-временной момент импульса. Его чисто пространственные компоненты³ J^{ik} ($k = 1, \dots, 3$) дают обычный момент импульса:

$$J_i = \epsilon_{ijk} J^{jk}, \quad (1.16)$$

где ϵ_{ijk} — полностью антисимметричный тензор с $\epsilon_{123} = 1$. Временные компоненты J^{0i} не порождают независимых сохраняющихся величин.

2. Частица в электромагнитном поле. Электромагнитное поле описывается 4-потенциалом $A_\bullet = (\varphi, -\mathbf{A})$ или, более точно, 1-формой $A(x) = A_\mu(x) dx^\mu$. Движение частицы описывается действием

$$S[x] = \int_A^B (-m ds - eA) = \int_A^B (-m ds - eA_\mu dx^\mu) = \int_{t_A}^{t_B} dt \left(-m\sqrt{1 - \mathbf{v}^2} + e\mathbf{A}\mathbf{v} - e\varphi \right). \quad (1.17)$$

Уравнение движения удобнее написать в гамильтоновой форме. Имеем для импульсов

$$\mathbf{P} = \frac{\partial L}{\partial \mathbf{v}} = \frac{m\mathbf{v}}{\sqrt{1 - \mathbf{v}^2}} + e\mathbf{A} = \mathbf{p} + e\mathbf{A} \quad (1.18)$$

и для гамильтониана

$$H(\mathbf{r}, \mathbf{P}) = \mathbf{v}\mathbf{P} - L = \sqrt{m^2 + (\mathbf{P} - e\mathbf{A}(\mathbf{r}))^2} + e\varphi(\mathbf{r}). \quad (1.19)$$

Уравнение движения

$$\dot{\mathbf{P}} = -\frac{\partial H}{\partial \mathbf{r}}, \quad \dot{\mathbf{r}} = \frac{\partial H}{\partial \mathbf{P}} \quad (1.20)$$

нетрудно привести к стандартному виду

$$\dot{\mathbf{p}} = e\mathbf{E} + e\mathbf{v} \times \mathbf{H}, \quad \mathbf{E} = -\dot{\mathbf{A}} - \nabla\varphi, \quad \mathbf{H} = \nabla \times \mathbf{A}. \quad (1.21)$$

В более инвариантном виде уравнение движения можно записать как

$$m \frac{du^\mu}{ds} = e F^\mu{}_\nu u^\nu, \quad (1.22)$$

где тензор электромагнитного поля определяется как

$$F = \frac{1}{2} F_{\mu\nu} dx^\mu \wedge dx^\nu = dA, \quad (1.23)$$

или

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu, \quad \partial_\mu = \frac{\partial}{\partial x^\mu}. \quad (1.24)$$

Этот тензор выражается через компоненты напряженности поля

$$F_{\bullet\bullet} = \begin{pmatrix} 0 & E_x & E_y & E_z \\ -E_x & 0 & -H_z & H_y \\ -E_y & H_z & 0 & -H_x \\ -E_z & -H_y & H_x & 0 \end{pmatrix}. \quad (1.25)$$

Обратите внимание, что в левой части (1.22) дифференцирование выполняется по лоренц-инвариантному собственному времени.

³Всюду греческие буквы из середины алфавита $\kappa, \lambda, \mu, \nu, \dots$ будут использоваться для пространственно-временных индексов, пробегающих значения $0, 1, \dots, d-1$. Латинские буквы из середины алфавита i, j, k, l, \dots будут использоваться для пространственных индексов и пробегать значения $1, \dots, d-1$.

Задачи

1. Покажите, что прямой мировой линии отвечает именно минимум (а не максимум) действия (1.8), то есть максимум собственного времени $s = \int_A^B ds$. Приведите примеры мировых линий, отвечающих *наименьшему* собственному времени. Чему равно это время?

2. В случае системы нескольких свободных частиц момент импульса равен сумме их моментов:

$$J^{\mu\nu} = \sum_s (x_s^\mu p_s^\nu - x_s^\nu p_s^\mu).$$

Покажите, что сохранение компонент J^{0i} эквивалентно тому, что центр инерции системы

$$\mathbf{R} = \frac{\sum_s E_s \mathbf{r}_s}{\sum_s E_s}$$

движется с постоянной скоростью.

3. Покажите, что при калибровочном преобразовании

$$A \rightarrow A + d\chi,$$

где $\chi(x)$ — произвольное скалярное поле, действие (1.17) преобразуется как

$$S \rightarrow S + e(\chi(x_A) - \chi(x_B)).$$

Объясните, почему отсюда следует, что уравнение движения частицы не меняется при калибровочных преобразованиях.

4. Выведите уравнение (1.22).

5*. Рассмотрите частицу во внешнем скалярном поле, которое описывается зависящей от точки массой $m(x)$ в действии (1.17). Напишите гамильтониан и уравнения движения такой частицы. Покажите, что если $m(x) = m_0 + U(x)$, $U(x) \ll m_0$, то в нерелятивистском пределе эти уравнения описывают частицу во внешнем потенциальном поле $U(x)$.

Семинар 1

Преобразования Лоренца, светоподобные координаты, диаграмма Пенроуза

Для простоты мы будем рассматривать двумерное пространство-время.

1. Графическое представление преобразований Лоренца в двух измерениях (Рис. 1.1).
2. Координаты Риндлера и равноускоренное движение. Введем координаты τ и ξ уравнениями:

$$t = \xi \operatorname{sh} \tau, \quad x = \xi \operatorname{ch} \tau.$$

Метрика в этих координатах принимает вид

$$ds^2 = \xi^2 d\tau^2 - d\xi^2.$$

Координаты Риндлера определены в во внешней области светового конуса: $|x| < |t|$ (Рис. 1.2). Линии постоянной координаты ξ отвечают мировым линиям наблюдателей с инвариантным ускорением ξ^{-1} , причем собственное время такой частицы равно $\xi\tau + \text{const}$.

Важной особенностью риндлеровских наблюдателей является наличие горизонта событий. События вне правого клина для них ненаблюдаемы. Другое полезное наблюдение: твердый стержень не может ускоряться так, чтобы все его точки ускорялись с одним и тем же ускорением. Задний конец должен иметь большее ускорение («парадокс» Белла).

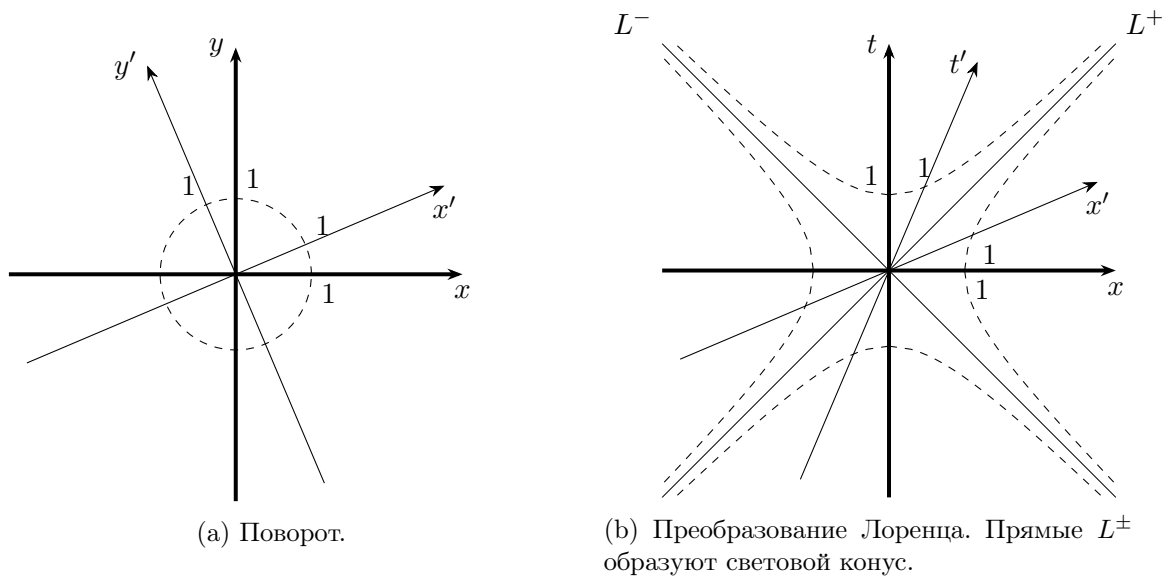


Рис. 1.1. Поворот и преобразование Лоренца в двумерном пространстве-времени. Пунктирные линии содержат точки, удаленные на единичное (по модулю) расстояние от начала координат.

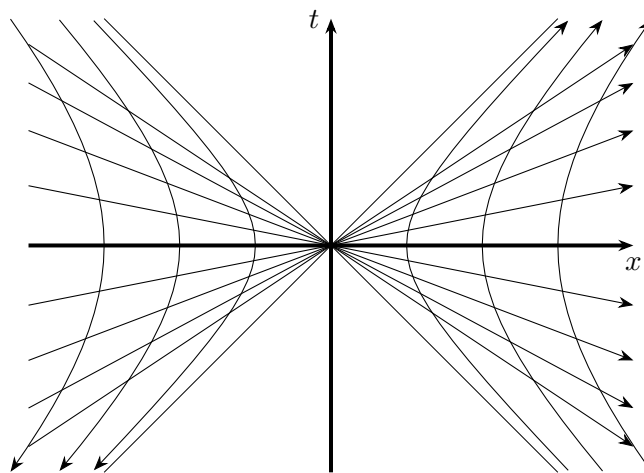


Рис. 1.2. Координаты Риндлера. Гиперболы отвечают линиям постоянной координаты $\xi = -3, -2, \dots, 3$, а прямые — линиям постоянного времени $\tau = -0.8, -0.6, \dots, 0.8$.

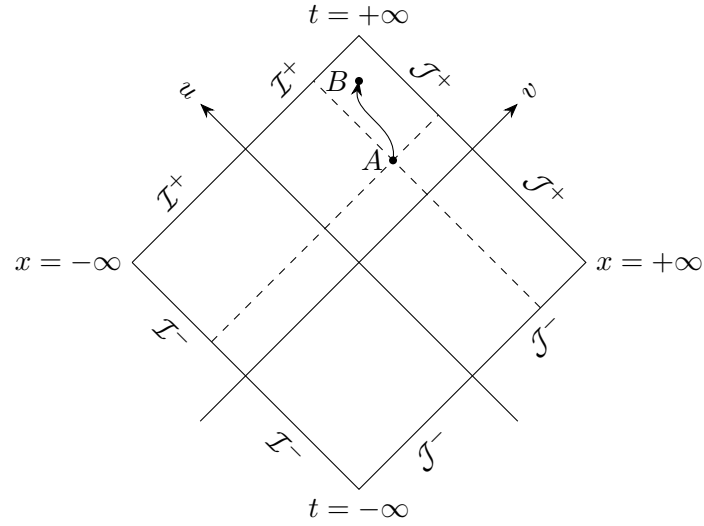


Рис. 1.3. Диаграмма Пенроуза для пространства Минковского. Граница \mathcal{I}^+ , \mathcal{I}^- отвечает бесконечному будущему, а граница \mathcal{J}^+ , \mathcal{J}^- — бесконечному прошлому.

3. Координаты светового конуса и причинная диаграмма (диаграмма Пенроуза) пространства Минковского.

Введем координаты

$$u = t - x, \quad v = t + x.$$

В этих координатах метрика имеет вид

$$ds^2 = du dv.$$

В отличие от риндлеровских координат эти координаты глобально определены на всем двумерном пространстве-времени. Используя их можно «сжать» пространство-время, чтобы изобразить его на конечном листе:

$$u = 2 \operatorname{tg} u', \quad v = 2 \operatorname{tg} v', \quad ds^2 = \frac{4 du' dv'}{\cos^2 u' \cos^2 v'}.$$

Тогда пространство-время примет вид, изображенный на Рис. 1.3. Диаграмма Пенроуза описывает причинную структуру пространства-времени. Световой конус любого события на этой диаграмме изображается двумя прямыми, наклоненными под углов 45° к горизонту.