Ковектором называется величина B_μ, преобразующаяся по правилу

$$\widetilde{B}_\mu \mapsto B_\mu = a^\nu_\mu \widetilde{B}_\nu.$$

Ковектор — это на самом деле элемент двойственного пространства.

Величина, преобразующаяся как произведение координат, $A^{\mu\nu} \mapsto \mapsto \widetilde{A}^{\mu\nu} = a^\mu_{\mu_1} a^\nu_{\nu_1} A^{\mu_1\nu_1}$, называется 2-вектором или тензором типа $(2,0)$.

Пример. Метрика $g_{\mu\nu}$ при преобразовании координат меняется следующим образом:

$$\widetilde{g}_{\mu\nu}\widetilde{x}^\mu\widetilde{x}^\nu = \widetilde{g}_{\mu\nu} a^\mu_{\mu_1} x^{\mu_1} a^\nu_{\nu_1} x^{\nu_1} = (\widetilde{g}_{\mu\nu} a^\mu_{\mu_1} a^\nu_{\nu_1}) x^{\mu_1} x^{\nu_1} = g_{\mu_1\nu_1} x^{\mu_1} x^{\nu_1},$$

следовательно, $g_{\mu_1\nu_1} = \widetilde{g}_{\mu\nu} a^\mu_{\mu_1} a^\nu_{\nu_1}$, и следовательно, $g_{\mu\nu}$ — 2-ковектор или тензор типа $(0,2)$.

Обозначение $\varepsilon_{\mu\nu\lambda\sigma}$ закреплено за стандартным антисимметричеcким тензором ранга 4. Антисимметричность означает, что его компоненты меняют знак при перестановке любой пары индексов, кроме того, $\varepsilon_{0123} = 1$. *Рангом* называется число индексов тензора. В этих обозначениях форма объема равна $dV = \dfrac{1}{4!}\varepsilon_{\mu\nu\lambda\sigma}\, dx^\mu \wedge dx^\nu \wedge dx^\lambda \wedge dx^\sigma$. *Дуальным тензором* называется тензор, свернутый с $\varepsilon_{\mu\nu\lambda\sigma}$.

Пример. Пусть $F_{\mu\nu} = -F_{\nu\mu}$ — антисимметричный тензор. Тогда $(*F)$ по определению равен: $F^*_{\mu\nu} = \dfrac{1}{2}\varepsilon_{\mu\nu\lambda\sigma} F^{\lambda\sigma}$, где $F^{\lambda\sigma} = g^{\lambda\mu} g^{\sigma\nu} F_{\mu\nu}$ и $(g^{\lambda\mu}) = (g_{\mu\nu})^{-1}$.

Пусть (x^μ) и $(\widetilde{x}^\mu)$ — локальные координаты в точке и $\widetilde{x}^\mu = a^\mu_\nu x^\nu$. Скалярное поле — это функция от x $\varphi(x) \mapsto \widetilde{\varphi}(\widetilde{x}) = \varphi(x)$. Векторное поле — это сечение касательного расслоения $A^\mu(x) \mapsto \widetilde{A}^\mu(\widetilde{x}) = = a^\mu_\nu A^\nu(x)$.

2. Лагранжева механика

Состояние классической системы, изменение ее во времени

В евклидовой системе координат положение частицы в системе определяется декартовыми координатами x^i_α, где $\alpha = 1, \ldots, N$ — номера частиц. Таким образом, $3N$ величин задают положение частиц. В общем случае *обобщенные координаты* — это некие переменные q_i, задающие положения частиц в системе. Все возможные положения частиц или значений q_i в общем случае образуют *конфигурационное пространство*. Значения координат не определяют скоростей механической системы и не задают будущего.

Состояние задается значениями координат x^i_α и их производными по времени в данной точке — скоростями $\dot{x}^i_\alpha$. Все возможные состояния системы или значения $(q_i, \dot{q}_i)$ образуют *фазовое пространство*. Если состояние задано, то определяется будущее и прошлое системы с помощью уравнений движения (например, уравнений Ньютона).

Примеры:

1) $m\dfrac{d^2x}{dt^2} = 0$ — уравнение свободной частицы.

2) $m\dfrac{d^2x}{dt^2} = -kx$ — уравнение шарика на пружине.

Лагранж придал этим уравнениям в случае потенциальных сил форму

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} = 0,$$

где $L = \dfrac{m\dot{x}^2}{2}$ в примере 1 и $L = \dfrac{m\dot{x}^2}{2} - \dfrac{kx^2}{2}$ в примере 2, причем $\dfrac{kx^2}{2}$ — потенциальная энергия.

Телеологический принцип

Почему в однородной среде свет распространяется по прямой? Можно дать следующий ответ: потому что время, затрачиваемое им в этом случае, минимально. Это подтверждается также при преломлении. В этом случае свет, добиваясь минимизации суммарного времени, предпочитает пройти больше по тому участку, где его скорость выше, чтобы сэкономить там, где его скорость меньше.

«Все к лучшему в этом лучшем из миров» — провозглашает *телеологический принцип*. Свет, и вообще любая система, идет по такому пути, чтобы сделать «нечто» наименьшим. Это «нечто» есть действие S, функционал от траектории

$$S^{\text{реальное}} = \min\ S.$$

В классической теории будущее состояние однозначно предопределено настоящим, поэтому мы не можем одновременно допускать это и верить в то, что существует свобода воли (один из важнейших принципов христианства).

В квантовой теории этого противоречия нет. Оказывается, что система идет сразу по многим траекториям. При этом амплитуда вероятности того, что система перейдет из точки A в точку B равна $\int e^{iS/\hbar}$, где интеграл берется по множеству всех траекторий, соединяющих A и B, а S — действие на данной траектории. Классические уравнения получаются, когда $\hbar \to 0$. В результате интерференции усиливаются траектории вблизи классической. Теперь будущее не однозначно определяется настоящим, и противоречия с принципом свободы воли нет.

Вариационное исчисление и принцип наименьшего действия

Вариационное исчисление занимается нахождением экстремумов функций, область определения которых — бесконечномерное пространство (пространство кривых). Такие функции называют *функционалами*.

Исторически первой серьезной задачей вариационного исчисления была задача о линии наискорейшего спуска: по какой кривой, соединяющей две данные точки, должно двигаться тело в поле силы тяжести, чтобы время спуска было наименьшим? Напишем функционал, минимум которого — нужная кривая.

Маленький кусочек траектории длины $\sqrt{dx^2 + dy^2}$ проходится за время $\sqrt{dx^2 + dy^2}/v$, где v — скорость. Скорость можно найти из закона сохранения энергии: $\dfrac{mv^2}{2} + mgy = mgh$. Поэтому $v = \sqrt{2g(h-y)}$. Итак, нам надо минимизировать функционал

$$T = \int\limits_{x_1}^{x_2} \sqrt{\frac{1 + (dy/dx)^2}{h - y}}\, dx.$$

Другая задача с заранее известным ответом: кривая наименьшей длины, соединяющая две данные точки. Длина кривой $y = f(x)$ задается формулой $\Phi(f(x)) = \int \sqrt{1+(df/dx)^2}\,dx$. Функционал можно представлять как функцию от бесконечного числа переменных: у функции $\Phi(f_1, \ldots, f_n)$ увеличивается число аргументов, и аргумент f_i постепенно превращается в $f(x)$.

Мы будем заниматься только *дифференцируемыми функционалами*, т. е. такими функционалами Φ, что $\Phi(f+h) - \Phi(f) = F(f,h) + R(f,h)$, где $F(f,h)$ зависит от h линейно, а $R(f,h) = O(h^2)$, т. е. из $|h| < \varepsilon$ и $\left|\dfrac{dh}{dx}\right| < \varepsilon$ следует, что $|R| < C\varepsilon^2$. Функционал F называют *дифференциалом*, или *первой вариацией*.

Пусть для всех функций $h(x)$ функция $\Phi(q(x)+\varepsilon h(x))$ имеет минимум при $\varepsilon = 0$. Это значит, что первая вариация $F(q,h) = \dfrac{d\Phi(q+\varepsilon h)}{d\varepsilon}$ равна нулю для всех функций $h(x)$. В этом случае функцию q называют *экстремалью* функционала Φ.

В лагранжевой динамике мы будем рассматривать только функционалы вида $A(q(t)) = \displaystyle\int_{t_0}^{t_1} L(q(t), \dot q(t), t)\,dt$, где L — обыкновенная функция (а не функционал) от нескольких переменных. Такой функционал является дифференцируемым. Действительно,

$$F(x,h) = \frac{d}{d\varepsilon}A(x(t)+\varepsilon h(t)) = \int_{t_0}^{t_1} \frac{d}{d\varepsilon}L(x(t)+\varepsilon h(t), \dot x(t)+\varepsilon \dot h(t))\,dt =$$

$$= \int_{t_0}^{t_1} \left(\frac{\partial L}{\partial x}(x,\dot x)\cdot h(t) + \frac{\partial L}{\partial \dot x}(x,\dot x)\cdot \frac{d}{dt}h(t)\right)\,dt =$$

$$= \int_{t_0}^{t_1} \left(\frac{\partial L}{\partial x}(x,\dot x)\cdot h(t) + \frac{d}{dt}\left(\frac{\partial L}{\partial \dot x}(x,\dot x)h(t)\right) - \left(\frac{d}{dt}\frac{\partial L}{\partial \dot x}\right)\dot h(t)\right)\,dt =$$

$$= \int_{t_0}^{t_1} \left(\frac{\partial L}{\partial x} - \frac{d}{dt}\frac{\partial L}{\partial \dot x}\right)\cdot h(t)dt + \left(\frac{\partial L}{\partial \dot x}(x,\dot x)h(t)\right)\Bigg|_{t_0}^{t_1}.$$

Теорема. Экстремали такого функционала подчиняются уравнению Эйлера—Лагранжа

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{x}} - \frac{\partial L}{\partial x} = 0.$$

Доказательство. Так как мы оставляем неподвижными концы $x(t)$, то $h(t_0) = 0$ и $h(t_1) = 0$. Поэтому для всех h должно выполняться следующее равенство:

$$\int_{t_0}^{t_1} \left(\frac{\partial L}{\partial x} - \frac{d}{dt}\frac{\partial L}{\partial \dot{x}}\right) h(t)\, dt = 0. \tag{1.1}$$

Но если не выполнено уравнение Эйлера—Лагранжа, т. е. если существует t_2 такое, что

$$\left(\frac{d}{dt}\frac{\partial L}{\partial \dot{x}} - \frac{\partial L}{\partial x}\right)(t_2) \neq 0,$$

мы можем взять функцию h, почти везде равную 0, а в окрестности точки t_2 принимающую очень большие значения. Для такой h интеграл (1.1) не будет равен 0. Противоречие.

Уравнения Лагранжа

Все это нужно нам потому, что теперь мы можем доказать принцип наименьшего действия и переформулировать уравнения Ньютона в виде уравнений Эйлера—Лагранжа. Действительно, уравнения движения Ньютона для потенциальных сил имеют вид

$$\frac{d}{dt} m\dot{x}^i_\alpha = F^i_\alpha = -\frac{\partial U}{\partial x^i_\alpha}.$$

Введем функцию $L = T - U$ (лагранжиан), где

$$T = \sum_\alpha \sum_i \frac{m(\dot{x}^i_\alpha)^2}{2}.$$

(α — номер частицы). Элементарно проверяется, что уравнения Ньютона можно переписать в виде уравнений Эйлера—Лагранжа. Следовательно, это уравнение задает экстремаль функционала действия

$$A(x^i_\alpha(t), \dot{x}^i_\alpha(t)) = \int\limits_{t_0}^{t_1} L(x^i_\alpha(t), \dot{x}^i_\alpha(t))\, dt.$$

Итак, принцип наименьшего действия доказан.

В уравнениях в форме Эйлера—Лагранжа можно делать произвольные замены координат $x^i_\alpha \to q_\alpha$, и от этого вид этих уравнений не изменится. В самом деле, экстремальность траектории не зависит от того, в каких координатах мы работаем. Уравнения Ньютона не обладают таким свойством. Величины q_α называют обобщенными координатами, а величины $\dot{q}_\alpha$ — обобщенными скоростями. Координаты и скорости задают состояние системы.

Интегралы движения и симметрии

Принцип наименьшего действия позволяет связывать интегралы движения с симметриями.

Интегралы движения — это функции состояния I, сохраняющиеся при движении, т. е. $\dfrac{d}{dt}I(q, \dot{q}) = 0$.

Примеры первых интегралов:

1. Свободная частица: $m\dfrac{d^2x}{dt^2} = 0$. Интегралы: $P = m\dot{x}$, $E = \dfrac{m(\dot{x})^2}{2}$.

2. Пусть $\dfrac{\partial}{\partial t}L = 0$. Рассмотрим величину $\dot{q}^i\dfrac{\partial L}{\partial \dot{q}^i}$ (здесь, как и в дальнейшем, подразумевается суммирование по повторяющимся индексам):

$$\frac{d}{dt}\left(\dot{q}^i\frac{\partial L}{\partial \dot{q}^i}\right) = \frac{d^2q^i}{dt^2}\frac{\partial L}{\partial \dot{q}^i} + \dot{q}^i\frac{d}{dt}\frac{\partial L}{\partial \dot{q}^i} = \frac{d^2q^i}{dt^2}\frac{\partial L}{\partial \dot{q}^i} + \dot{q}^i \cdot \frac{\partial L}{\partial q^i} = \frac{d}{dt}L(q(t), \dot{q}(t)).$$

Следовательно,

$$\frac{d}{dt}\left(\dot{q}^i\frac{\partial L}{\partial\dot{q}^i}-L\right)=0$$

и мы получили первый интеграл — энергию

$$E=\dot{q}^i\frac{\partial L}{\partial\dot{q}^i}-L=\mathrm{const}\,.$$

Симметриями называют однопараметрические группы преобразований конфигурационного пространства

$$q^i\to f^i(q,\varepsilon),\quad f^i(q,0)=q^i,$$

сохраняющие лагранжиан. В принципе можно рассматривать и более сложные симметрии.

Теорема (Нётер). Каждой симметрии соответствует первый интеграл

$$I(q,\dot{q})=\frac{\partial L}{\partial\dot{q}^i}\cdot h^i(q),\ \text{ где }\ h^i(q)=\frac{\partial f^i}{\partial\varepsilon}|_{\varepsilon=0}.$$

Доказательство.

$$\frac{d}{dt}\left(\frac{\partial L}{\partial\dot{q}^i}\cdot h^i\right)=\frac{d}{dt}\left(\frac{\partial L}{\partial\dot{q}^i}\right)\cdot h^i+\frac{\partial L}{\partial\dot{q}^i}\cdot\dot{h}^i=\frac{\partial L}{\partial q^i}\cdot h^i+\frac{\partial L}{\partial\dot{q}^i}\cdot\dot{h}^i=$$
$$=\frac{\partial L}{\partial q^i}\cdot\frac{\partial f^i}{\partial\varepsilon}+\frac{\partial L}{\partial\dot{q}^i}\cdot\frac{\partial\dot{f}^i}{\partial\varepsilon}\Big|_{\varepsilon=0}=\frac{dL(f(q,\varepsilon),\dot{f}(q,\varepsilon))}{d\varepsilon}\Big|_{\varepsilon=0}=0,$$

так как каждая симметрия сохраняет лагранжиан.

Пусть у нас есть несколько тел и их потенциальная энергия u зависит только от разностей $\vec{x}_a-\vec{x}_b$. Тогда преобразование $\vec{x}_a\to\vec{x}_a+$ $+\vec{\varepsilon}$ — симметрия, а

$$p^k=\sum_a\frac{\partial L}{\partial\dot{x}_a^k}=\sum_a m_a\dot{x}_a^k$$

является импульсом (первым интегралом, соответствующим этой симметрии).

Пусть имеется несколько тел, и U является функцией от расстояний $|\vec{x}_a - \vec{x}_b|$. Тогда кроме параллельных переносов есть еще один набор симметрий — повороты:

$$x_a^i \to x_a^i + \omega^k \varepsilon^i_{kl} x_a^l.$$

Этой симметрии соответствует интеграл движения

$$M^k = \sum_a \varepsilon^k_{il} \cdot \frac{\partial L}{\partial \dot{x}_a^i} \cdot x_a^l = \varepsilon^k_{il} \sum_a p_a^i \cdot x_a^l = [\,\vec{p}, \vec{x}\,]^k,$$

называющийся моментом количества движения.

Примеры систем

Пример 1. Свободная нерелятивистская частица:

$$L = \frac{m(\dot{x}^i)^2}{2}.$$

Приведем некоторые симметрии и соответствующие им интегралы движения.

(а) Параллельные переносы $\vec{x} \to \vec{x} + \vec{\varepsilon}$, $p^k = m\dot{x}^k = \mathrm{const}$.
(б) Повороты $x^i \to x^i + \omega^k \varepsilon^{ikl} x^l$, $M = [\,\vec{p}, \vec{x}\,] = \mathrm{const}$.
(в) Перенос во времени, $E = \dfrac{m(\dot{x}^i)^2}{2} = \mathrm{const}$.

Пример 2. Рассмотрим две частицы в трехмерном пространстве, причем силы зависят только от расстояния:

$$L = \frac{m_1(\dot{x}_1)^2}{2} + \frac{m_2(\dot{x}_2)^2}{2} - u(|\vec{x}_1 - \vec{x}_2|).$$

Введем координаты

$$\vec{X} = \frac{m_1\vec{x}_1 + m_2\vec{x}_2}{m_1 + m_2}, \quad \vec{x} = \vec{x}_1 - \vec{x}_2.$$

Пусть $M = m_1 + m_2$, величина $m = \dfrac{m_1 m_2}{m_1 + m_2}$ называется приведенной массой. В случае $m_1 \ll m_2$ имеем $M \approx m_2$, $m \approx m_1$.

Ясно, что $L = \frac{M\dot{X}^2}{2} + \frac{m\dot{x}^2}{2} - U(|\vec{x}|) = L_1 + L_2$, где $L_1 = \frac{M\dot{X}^2}{2}$, $L_2 = \frac{m\dot{x}^2}{2} - u(|\vec{x}|)$.

Мы видим, что такое движение распадается на два — движение свободной частицы и движение частицы в радиальном силовом поле. Исследуем последнее движение.

В этом случае сохраняются энергия $E = \frac{m\dot{x}^2}{2} + U(|\vec{x}|)$ и момент количества движения относительно $\vec{x} = 0$, т. е. величина $\vec{M} = [\vec{x}, \vec{p}]$. Так как $\vec{x} \perp \vec{M}$ и $\vec{M} = \mathrm{const}$, движение происходит в плоскости. Введя в ней координаты r и φ, получаем

$$L = \frac{m}{2}(\dot{r}^2 + r^2\dot{\varphi}^2) - U(r).$$

Из конкретного вида L следует, что $\frac{\partial L}{\partial \varphi} = 0$. Такая координата называется *циклической*. Поэтому

$$p_\varphi = \frac{\partial L}{\partial \dot{\varphi}} = mr^2\dot{\varphi} = M = \mathrm{const}\,.$$

$\frac{1}{2}r \cdot r\dot{\varphi}$ — площадь сектора, пройденная за единицу времени. Итак, секториальная скорость постоянна. Это *второй закон Кеплера*.

$$E = \frac{m(\dot{r}^2 + r^2\dot{\varphi}^2)}{2} + U(r) = \frac{m\dot{r}^2}{2} + \frac{M^2}{2mr^2} + U(r) = \frac{m\dot{r}^2}{2} + U_{\text{ef}}(r),$$

где $U_{\text{ef}}(r) = \frac{M^2}{2mr^2} + U(r)$ называется *эффективной потенциальной энергией*. Поэтому

$$\int dt = \int \frac{dr}{\sqrt{2(E-U) - M^2/mr}},$$

$$\Delta\varphi = \int \frac{(M/r^2)\,dr}{\sqrt{2(E-U) - M^2/mr^2}}.$$

Расстояние до центра колеблется между $r_{\min}$ и $r_{\max}$ (r может быть только таким, что $U_{\text{ef}}(r) < E$). Но, вообще говоря, траектории при таком движении незамкнуты, $\Delta\varphi \neq 2\pi n$.

В случае $U(r) = kr^2$ орбиты замкнуты.

Пример 3. Частица в кулоновском поле. Это частный случай предыдущего примера. $U(r) = -\frac{\alpha}{r}$ — кулоновский потенциал. В частности, это случай гравитации.

$$L = \frac{m\dot{r}^2}{2} + \alpha/r.$$

В этом случае есть скрытая симметрия с группой $\mathrm{O}(4)$ и дополнительные три интеграла движения

$$\vec{K} = [\vec{v}, \vec{M}] + \frac{\alpha\vec{r}}{r} = mv^2\vec{r} - m\vec{v}(\vec{v}, \vec{r}) + \frac{\alpha\vec{r}}{r},$$
$$\vec{K} = m\vec{r}(\dot{r}^2 + r^2\dot{\varphi}^2) - m\vec{v}r\dot{r} + \frac{\alpha\vec{r}}{r}.$$

Существование этого интеграла показывает, что орбита замкнута. Действительно, в момент, когда $\dot{r} = 0$, имеем $\vec{r} \parallel \vec{K}$, т. е. перигелий не смещается.

Докажем третий закон Кеплера. Уравнение движения планеты имеет вид

$$m\frac{d^2\vec{r}}{dt^2} = \frac{\alpha\vec{r}}{r^3}.$$

Если мы заменим t на μt, а $\vec{r}$ на $\lambda\vec{r}$, причем $\mu^2 = \lambda^3$, уравнение сохраняется. Из этого следует, что $T^2 = \gamma r^3$.

Пример 4. Осциллятор.

$$L = \frac{m\dot{x}^2}{2} - \frac{kx^2}{2}.$$

Такому уравнению удовлетворяют шарики на пружине, маятники. Имеется симметрия $x \to \lambda x$. Из нее следует независимость периода колебаний такого маятника от амплитуды.

Пример 5. Системы связанных осцилляторов.

$$L = \sum \frac{m_i\dot{x}_i^2}{2} - \sum \frac{k(x_i - x_{i+1})^2}{2}.$$

Пример 6. Релятивистская частица.

В этом случае траектория лежит в четырехмерном пространстве-времени $(ct, x^1, x^2, x^3) = x^\nu$. Траектории должны переходить в траектории при преобразованиях Лоренца. Для этого достаточно, чтобы действие было лоренц-инвариантно. Мы можем выбрать на кривой в качестве параметра t. Тогда простейшее лоренц-инвариантное действие имеет вид

$$S = m\int\sqrt{c^2\,dt^2 - d\vec{x}^2} = mc^2\int\sqrt{1-(dx^i/dt)^2/c^2}\,dt.$$

Легко доказать, что экстремали этого действия (на самом деле на этих траекториях достигается минимум) — прямые.

Итак, лагранжиан $L = -mc^2\sqrt{1-(\dot{x}/c)^2}$. Уравнения Эйлера–Лагранжа имеют вид

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{x}^i} = \frac{\partial L}{\partial x^i}.$$

Так как $\dfrac{\partial L}{\partial x^i} = 0$, импульс $\vec{P} = \dfrac{\partial L}{\partial \dot{x}^i} = \dfrac{m\vec{v}}{\sqrt{1-v^2/c^2}} = \mathrm{const}.$

Энергия имеет вид $E = \dot{x}^i p_i - L = \dfrac{mc^2}{\sqrt{1-v^2/c^2}}$. При малых v энергия равна $mc^2 + \dfrac{m\vec{v}^2}{2}$. Первое слагаемое в этой сумме $E = mc^2$ можно интерпретировать как внутреннюю энергию. А вектор $p^i = (E/c, \vec{p})$ является 4-вектором, т. е. величина $(p^i)^2 = m^2$ сохраняется при преобразованиях Лоренца.

Структура пространства-времени — это вопрос не геометрии, а физики. Фактически, это вопрос о группе инвариантности уравнений движения. Его нельзя решить без эксперимента.

3. Уравнения теории поля

Для возникновения теории поля имеются две причины. Первая причина состоит в том, что мы можем рассматривать системы с бесконечным числом степеней свободы. Система с конечным числом степеней свободы описывается обобщенными координатами $q_i(t)$ и обобщенными скоростями $\dot{q}_i(t)$. Если число координат становится континуальным,