1. Геометрическая оптика

Основные идеи гамильтоновой механики пришли из геометрической оптики. Поэтому мы сначала обсудим основные приницы геометрической оптики, а потом сформулируем аналогичные принципы гамильтоновой механики.

Будем рассматривать распространение света как движение частиц — фотонов. По какой траектории движется фотон? Оказывается, в большинстве случаев траектория может быть найдена, исходя из следующего принципа, сформулированного Ферма в 1650 году: *свет выбирает такую траекторию между двумя точками, движение по которой занимает минимальное время.*

Принципу Ферма предшествовали как экспериментальные, так и теоретические исследования движения света. Очень много было сделано греками. Птолемей, отказавшись от греческой традиции изучать явления природы умозрительно и дедуктивно, проводил множество экспериментов. В частности, он оставил таблицы зависимости углов преломления от углов падения. Герон Александрийский сформулировал принцип наименьшего расстояния, который применим к однородным и изотропным средам. Согласно принципу Герона, свет в таких средах должен распространяться по прямой. Но этот принцип не объяснял явление преломления. Принцип Ферма объясняет эффект преломления, а также множество других эффектов.

Например, для конструкции линз достаточно пользоваться только принципом Ферма. На закате, когда мы видим солнце у самого горизонта, оно на самом деле уже зашло, то есть прямая, соединяющая нас (наблюдателя) и солнце, пересекает земную поверхность. То, что мы все таки видим солнце, связано с изгибанием лучей. Лучи изгибаются, чтобы минимизировать свой путь через плотные слои атмосферы, в которых скорость света ниже. С изгибанием лучей связано еще и то, что видимая форма солнца на закате не круглая, а сплюснутая.

На горячем асфальте или горячем песке можно увидеть мираж. Лучи как бы отражаются от горячего потока воздуха. На самом деле, это явление можно объяснить при помощи принципа Ферма: лучам выгодно изгибаться и проходить существенное расстояние через горячий воздух, потому что скорость света в горячем воздухе больше, чем в холодном.

Постараемся формализовать принцип Ферма. Мы предполагаем, что скорость света в данной точке зависит только от точки и от направления, в котором движется свет (но не от того, скажем, где расположен источник света и какова его интенсивность). Скорость света в данной среде тем самым полностью определяется физическими свойствами среды. Среда называется *изотропной*, если скорость света зависит только от точки, но не от направления. Будем полагать для простоты, что среда изотропна. Тогда скорость света можно изобразить функцией точки.

Вообще-то, свет распространяется в трехмерном пространстве, но мы будем считать, что он распространяется на плоскости. Пусть $v(x, y)$ — скорость света в точке с координатами x и y. Тогда световая траектория минимизирует интеграл

$$\int\limits_{(x_0,y_0)}^{(x_1,y_1)} \frac{ds}{v(x,y)}$$

Этот интеграл взят по кривой, соединяющей точки (x_0, y_0) и (x_1, y_1) (эти точки зафиксированы), и называется *оптической длиной кривой.* Параметр s — натуральный параметр на кривой. Интеграл можно воспринимать как одномерный интеграл

$$\int\limits_{0}^{L} \frac{ds}{v(x(s), y(s))},$$

где L — длина кривой, а $x = x(s)$, $y = y(s)$ — параметрическое уравнение кривой. Тем самым, оптическая длина является функцией от кривой, и она определена на всех достаточно гладких кривых с фиксированными концами. Принцип Ферма утверждает, что свет движется по кривой, на которой эта функция достигает своего минимума.

Принцип Ферма нуждается в некоторых уточнениях. Например, нужно допускать не только такие траектории, вдоль которых время прохождения минимально, но и трактории, соответствующие любым критическим точкам функционала времени прохождения. Эта мадификация аналогично тому, что вместо минимума какой-либо функции от одной действительной переменной мы рассматриваем такое значение переменной, при котором производная функции равна нулю. Точки минимума этому условию удовлетворяют, но не только они.

Рассмотрим однородную и изотропную среду, то есть такую среду, в которой скорость света не зависит ни от точки, ни от направления. Тогда оптическая длина кривой пропорциональна обычной евклидовой длине кривой. Таким образом, минимизировать время

— то же самое, что минимизировать длину. Мы знаем, какие кривые в евклидовом пространстве имеют минимальную длину — это прямые линиии. Таким образом, в однородной и изотропной среде распространение света происходит вдоль прямолинейных лучей.

Правда, эти лучи могут отражаться от поверхностей, что непосредственно не следует из принципа Ферма, если понимать его буквально. Скажем, если луч света выбирает между тем, чтобы пойти по прямой или отразиться, скажем, от поверхности озера, то, согласно принципу Ферма, он должен выбрать первое. В природе же выбираются оба варианта; хотя второй вариант выбирается, так сказать, с меньшим весом, зависящим от отражающей способности поверхности.

Допустим, что луч должен отразиться от гладкой поверхности. То, как именно луч это делает, описывается принципом Ферма. Рассмотрим для простоты плоскую задачу. Луч света выходит из точки A и отражается от прямой l. Наблюдатель находится в точке B и видит отраженный свет. По какой кривой идет свет от точки A до точки B?

Это математическая задача:

Задача 1. Найдите кривую наименьшей длины, соединяющую точки A и B и имеющую хотя бы одну общую точку с прямой l. Мы предполагаем, что A и B находятся по одну и ту же сторону от прямой l (иначе задача тривиальна).

Эта задача красиво решается методом отражения, который иногда проходят в школе на геометрии. Но можно эту задачу решать и "в лоб". Для этого нужно заметить, что оптимальная кривая (если она существует) должна состоять из двух прямолинейных отрезков, один из которых идет от точки A до точки X на прямой l, а другой от точки X до точки B. Таким образом, задача свелась к задаче с единственным параметром. В качестве параметра можно взять координату точки X на прямой l. Обозначим эту координату через u. Будем писать $X(u)$ вместо X, имея в виду точку на прямой l с координатой u. Мы, конечно, предполагаем, что координата u выбрана таким образом, что расстояние между точками $X(u)$ и $X(u')$ равно $|u - u'|$.

Задача 2. Посчитайте производную по u расстояния между точками A и $X(u)$ (выразите ответ через угол между отрезком $[A, X(u)]$ и прямой l).

Пусть α — угол между отрезком $[A, X(u)]$ и прямой l. Ответ в задаче такой: $\pm\cos\alpha$. Знак зависит от того, какой именно из двух возможных углов мы принимаем за α. Теперь, чтобы найти кратчайший отраженный луч, нам нужно найти минимум следующей функции одной переменной:

$$f(u) = |AX(u)| + |X(u)B|,$$

где $|AX|$ обозначает евклидово расстояние между точками A и X. Приравняв производную функции f к нулю, получаем, что угол падения должен быть равен углу отражения.

Переходя из одной однородной изотропной среды в другую, луч преломляется. Это явление связано с тем, что скорость света меняется при переходе из одной среды в другую. Предположим, что одна среда находится в верхней полуплоскости, и скорость света в этой среде равна c_1, а вторая среда находится в нижней полуплоскости, и скорость света в этой среде равна c_2. Например, можно считать, что сверху находится воздух, а снизу вода. Пусть l — горизонтальная прямая, разделяющая две наши среды. Как и раньше, обозначим через $X(u)$ точку на прямой l, координата которой равна u.

Допустим, что луч света пытается попасть из точки A в точку B, причем первая точка находится в воздухе, а вторая — в воде. Понятно, что сначала луч по прямолинейному отрезку пойдет из точки A в какую-то точку прямой l, а потом из этой точке в точку B, опять по прямолинейному отрезку. Мы приходим к задаче минимизации следующей функции:

$$f(u) = \frac{|AX(u)|}{c_1} + \frac{|X(u)B|}{c_2}.$$

Приравнивая производную функции f к нулю, получаем следующее уравнение:

$$\frac{\cos\alpha_1}{c_1} = \frac{\cos\alpha_2}{c_2}.$$

Здесь α_1 — это угол между отрезком $[A, X(u)]$ и прямой l, а α_2 — угол между прямой l и отрезком $[X(u), B]$. Углы выбираются таким образом, что либо оба угла острые, либо оба угла тупые. Мы будем считать, что оба угла острые.

Полученное соотношение между $\cos\alpha_1$ и $\cos\alpha_2$ называется законом Снелла. Этот закон получил Снелл (вариант: Снеллиус) в 1621 году. Виллеброрд Снелл — голландский математик, физик и астроном. Цитирую википедию:

Родился в Лейдене в семье профессора математики. В 1613 году стал преемником отца на должности профессора Лейденского университета. Предложил использовать метод подобия треугольников при проведении геодезических измерений, при помощи которого нашел решение задачи Потенота. В его работе “Eratosthenes Batavus” (“Голландский Эратосфен”), опубликованной в 1617 году, описывался метод триангуляции и приводились результаты измерений между городами Берген-Оп-Зомом и Алкмаром дуги меридиана 1°11’30". В 1621 году открыл закон преломления света. Однако результаты многочисленных экспериментов по оптике опубликованы не были. Позже они были обнаружены в архивах Рене Декартом, который использовал их при написании своих “Начал философии””.

Следующее неформальное рассуждение позволяет обобщить закон Снелла на случай изотропной среды, в которой скорость света (как скалярная функция точки) зависит только от y-координаты. Сначала представим себе среду, состоящую из большого числа горизонтальных слоев. Каждый слой заключен между двумя параллельными горизонтальными плоскостями (точнее, если мы думаем про плоскость, между двумя параллельными горизонтальными прямыми). Допустим, что внутри каждого слоя среда однородная. Обозначим через c_i скорость света в i-ом слое. Тогда закон Снелла говорит о том, что величина

$$\frac{\cos\alpha_i}{c_i}$$

постоянна (то есть не зависит от i). Теперь будем делать слои все тоньше и тоньше. Любая изотропная среда, в которой скорость света является гладкой функцией $v(y)$ одной только координаты y, может быть сколь угодно точно приближена средой состоящей из конечного, но очень большого числа тонких горизонтальных слоев. Переходя к пределу (мы здесь допускаем физический уровень строгости, и не обосновываем законность предельного перехода), получаем, что

$$\frac{\cos\alpha(t)}{v(y(t))} = const$$

вдоль любого светового луча (то есть величина, написанная в левой части, не зависит от t). Здесь t — любой параметр на световом луче (например, можно считать, что $x(t)$ и $y(t)$ — это координаты фотона в момент времени t). Угол $\alpha(t)$ — это угол между световым лучом и горизонтальным направлением в точке $(x(t), y(t))$. Мы будем называть полученный закон обобщенным законом Снелла.

Обобщенный закон Снелла можно переписать как дифференциальное уравнение на траекторию. Обозначим

$$\kappa = \frac{\cos \alpha(t)}{v(y(t))}.$$

Эта величина не зависит от t по обобщенному закону Снелла. Допустим, что траектория светового луча является графиком некоторой функции $y = y(x)$. Тогда можно взять x вместо параметра t на световом луче. Угол $\alpha(x)$, т.е. угол наклона светового луча (по отношению к горизонтальному направлению) связан с производной функции $y(x)$. Именно, как мы знаем, производная равна тангенсу угла наклона. Тангенс и косинус связаны следующей формулой:

$$1 + \operatorname{tg}^2 \alpha = \frac{1}{\cos^2 \alpha}.$$

Таким образом, мы получаем следующее дифференциальное уравнение на световые лучи:

$$\frac{dy}{dx} = \frac{\sqrt{1 - \kappa^2 v^2(y)}}{\kappa v(y)}.$$

Правая часть этого уравнения зависит только от y. Следовательно, мы получили уравнение с разделяющимися переменными.

2. Семинар: принцип Ферма

Одно из самых интересных применений обобщенного закона Снелла — задача о брахистохроне. Дано две точки A и B, причем точка B находится ниже точки A, но не обязательно на той же самой вертикали. Рассмотрим различные кривые, соединяющие A с B. Представим себе точку, скользящую без трения вдоль этих кривых. Брахистохрона — это такая кривая, время скольжения вдоль которой минимально. Все скольжения, которые мы рассматриваем, начинаются в точке A и заканчиваются в точке B; начальная скорость всегда нулевая (то есть мы не подталкиваем скользящие частицы).

Задача нахождения брахистохроны может быть формально сведена к задаче геометрической оптики. В самом деле, время скольжения вдоль кривой равно интегралу

$$\int \frac{ds}{v},$$

взятому вдоль этой кривой. Если окажется, что скорость v скольжения точки зависит только от точки, но не от кривой, то задача сведется к геометрической оптике.

Скорость v может быть найдена из закона сохранения энергии:

$$\frac{mv^2}{2} - mgy = 0.$$

(Мы направим ось y вниз, поэтому потенциальная энергия убывает с ростом y, отсюда знак минус. Мы также предполагаем, что точка A находится на высоте 0, поэтому потенциальная энергия равна $mv^2/2$. В начальный момент скользящая точка находится на высоте 0, и ее скорость равна нулю, поэтому полная энергия равна нулю.) Находим

$$v = \sqrt{2gy}.$$

Таким образом, v действительно зависит от точки. Более того, v зависит только от координаты y. Поэтому мы можем применить обобщенный закон Снелла.

Мы уже выяснили, что обобщенный закон Снелла переписывается как дифференциальное уравнение на световые траектории с разделяющимися переменными. При этом световые траектории ищутся как графики функций $y = y(x)$. В нашем случае, дифференциальное уравнение имеет вид

$$\frac{dy}{dx} = \frac{\sqrt{1 - 2g\kappa^2 y}}{\kappa\sqrt{2gy}}.$$

Разделяя переменные и интегрируя, получаем:

$$\int \frac{\sqrt{2gy}dy}{\sqrt{1 - 2g\kappa^2 y}} = \int \frac{dx}{\kappa}.$$

Чтобы вычислить интеграл в левой части, удобно сделать следующую замену переменной:

$$2g\kappa^2 y = \sin^2 \phi$$

(старая переменная y, новая переменная ϕ; эти переменные выражаются одна через другую из уравнения, выписанного выше). Интегрируя, получаем:

$$\frac{\phi}{2} - \frac{\sin 2\phi}{4} = g\kappa^2 x + C.$$

Заменим 2ϕ на ϕ (это просто изменение параметризации кривой), x на $4g\kappa^2 x$ и y на $4g\kappa^2 y$ (это гомотетия, т.е. изменение единицы длины). Получаем:

$$x = \phi - \sin\phi + const, \quad y = 1 - \cos\phi.$$

Как нетрудно видеть, это уравнение описывает *циклоиду*, то есть траекторию точки, прикрепленной к колесу, которое равномерно и без проскальзывания катится по горизонтальной прямой. Точнее, мы получили траекторию точки на колесе, отраженную относительно горизонтальной прямой.

Циклоида — периодическая кривая с счетным числом каспов (клювов). Брахистохрона — это часть циклоиды, начинающаяся в каспе и заканчивающаяся не позже ближайшего минимума.

Задача 3. Рассмотрим изотропную среду, занимающую верхнюю полуплоскость $y > 0$, скорость света в которой выражается формулой $v(x, y) = y$. Найдите форму световых лучей в этой среде.

Задача 4. Рассмотрим изотропную среду, занимающую внутренность единичного круга, скорость света в которой выражается формулой $v(x, y) = 1 - x^2 - y^2$. Найдите форму световых лучей в этой среде.

Обобщенный закон Снелла можно также использовать для нахождения формы цепной линии. Эту форму принимает тонкая цепочка, свободно висящая между двумя закрепленными концами. Будем считать для определенности, что концы цепочки закреплены на одной и той же высоте. Ориентируем ось y вниз, и выберем начало координат так, чтобы закрепленные концы оказались на высоте 0.

Цепочка пытается минимизировать полную потенциальную энергию. Будем считать, что масса цепочки равномерно распределена по ее длине. Тогда потенциальная энергия маленького участка цепочки длины Δs примерно равна $-y\,\Delta s$, где $-y$ — высота этого участка (мы считаем плотность равной единице). Следовательно, полная потенциальная энергия цепочки выражается интегралом

$$U = -\int y\,ds.$$

Однако цепочка не может менять длины. Поэтому интеграл потенциальной энергии нужно минимизировать при условии постоянства длины, т.е.

$$\int ds = L,$$

где L — постоянная величина.

Таким образом, мы обсуждаем задачу нахождения условного экстремума. К этой задаче можно применить принцип множителей Лагранжа. Для нахождения условного экстремума некоторой функции f (определенной, в нашем случае, на бесконечномерном пространстве кривых) при условии $g = 0$, мы делаем вид, что пытаемся найти безусловный экстремум функции $f + \lambda g$ (то есть ищем критические точки этой функции). Здесь λ — константа, называемая множителем Лагранжа. После того, как критические точки функции $f + \lambda g$ найдены, значение множителя λ для этих точек находится из условия $g = 0$.

В нашем случае, мы должны притвориться, что пытаемся минимизировать интеграл

$$\int_\gamma (-y + \lambda) ds$$

по всем кривым γ с данной парой закрепленых концов. Экстремум этого интеграла можно найти из обобщенного закона Снелла: сам интеграл имеет вид интеграла оптической длины пути для изотропной среды, в которой скорость света записывается формулой

$$v(y) = \frac{1}{\lambda - y}.$$

Как мы видим, скорость света зависит только от y. Из обобщенного закона Снелла получаем следующее дифференциальное уравнение

$$\frac{dy}{dx} = \sqrt{\frac{(\lambda - y)^2}{\kappa^2} - 1},$$

или, разделяя переменные,

$$\frac{dy}{\sqrt{(\lambda - y)^2 - \kappa^2}} = \frac{dx}{\kappa}.$$

Введем замену переменных $\lambda - y = \kappa \operatorname{ch} \phi$ (старая переменная y, новая переменная ϕ). Интегрируя при помощи этой замены, получаем:

$$y = \lambda - \kappa \operatorname{ch} \left(\frac{x - x_0}{\kappa} \right).$$

Таким образом, с точностью до аффинной замены координат, цепная линия — это график функции гиперболический косинус.

Задача 5. Рассмотрим изотропную среду, занимающую верхнюю полуплоскость $y > 0$, скорость света в которой выражается формулой $v(x, y) = y$. Найдите форму световых лучей в этой среде.

Задача 6. Рассмотрим изотропную среду, занимающую внутренность единичного круга, скорость света в которой выражается формулой $v(x, y) = 1 - x^2 - y^2$. Найдите форму световых лучей в этой среде.

3. Лекция: принцип Гюйгенса

Мы пока занимались изучением траекторий отдельного фотона. Однако интересно также иметь описание движения пучка фотонов, или светового пятна. Обычно фотонов очень много. Настолько много, что невозможно следить за отдельными фотонами. Но можно смотреть, как распространяется световое пятно.

Предположим, что фотоны вылетают из точечного источника, и летят во всех возможных направлениях. Какова будет форма светового пятна за время t? Обозначим это световое пятно через $X(x_0, t)$, где x_0 — положение точечного источника. Более точно, $X(x_0, t)$ — это множество тех точек, в которые свет может попасть за время $\leq t$. В отличие от прошлой лекции, мы сейчас не предполагаем, что среда изотропна. Среда может быть неоднородна и неизотропна.

Принцип Гюйгенса описывает закон изменения светового пятна $X(x_0, t)$. Математически, этот принцип можно выразить следующей короткой формулой:

$$X(x_0, t) = \bigcup_{x_1 \in X(x_0, t_1)} X(x_1, t_2)$$

для всякого разложения $t = t_1 + t_2$ временного интервала t в сумму двух неотрицательных временных интервалов. Словами эта формула описывается так. Рассмотрим световое пятно за время t_1. Поместим в каждую точку этого светового пятна воображаемый источник света, и дадим всем этим источникам светить в течение времени t_2. Получится много *вторичных световых пятен*. Нужно взять объединение всех этих вторичных световых пятен, чтобы получить световое пятно за время t. Заметим, что правая часть равенства, выражающего принцип Гюйгенса, включает в себя множество $X(x_0, t_1)$. Дело в том, что x_1 обязательно принадлежит световому пятну $X(x_1, t_2)$.

Докажем принцип Гюйгенса. Пусть x взято из левой части принципа Гюйгенса, то есть $x \in X(x_0, t)$. Свет может дойти от точки x_0 до точки x за время $s \leq t$. Если $s \leq t_1$, то точка x принадлежит множеству $X(x_0, t_1)$, которое, как мы видели, является подмножеством правой части. Если же $s > t_1$, то $s = t_1 + s_2$, где $s_2 \leq t_2$. Рассмотрим траекторию фотона, идущего от x_0 к x. За время t_1, фотон доходит до некоторой точки $x_1 \in X(x_0, t_1)$, а затем за время $s_2 \leq t_2$ он доходит до точки x, откуда получаем, что $x \in X(x_1, t_2)$. Следовательно, x лежит в правой части.

Допустим теперь, что x лежит в правой части принципа Гюйгенса. Это значит, что в x можно попасть, сначала пройдя вдоль истинной световой траектории от точки x_0 до некоторой точки $x_1 \in X(x_0, t_1)$, а потом пройдя вдоль истинной световой траектории от точки x_1 до точки $x \in X(x_1, t_2)$. Мы знаем, что время движение по первой из двух траекторий не превышает t_1, а время движения по второй траектории не превышает t_2. Объединение двух рассматриваемых истинных световых траекторий может и не быть истинной световой траекторией. Однако это объединение является некоторой кривой, вдоль которой свет может пройти за время t или меньше. Значит, на прохождение света от x_0 до x по истинной световой траектории понадобится время $\leq t$ (это вытекает из принципа Ферма).

Пример. Рассмотрим однородную, но неизотропную, среду. В такой среде, все световые пятна за время t отличаются лишь параллельным переносом. Поэтому есть такая фигура $X(t)$, зависящая только от t, что

$$X(x_0, t) = x_0 + X(t)$$

для всякой точки x_0. Здесь $x_0 + X$ означает параллельный перенос множества X на вектор x_0, то есть множество точек $x_0 + x$, $x \in X$. *Сумма Минковского* двух множеств A, $B \subset \mathbb{R}^n$ определяется следующим образом:

$$A + B = \{a + b \mid a \in A,\ b \in B\}.$$

Например, если B — шар радиуса $\varepsilon > 0$ с центром в нуле, то $A + B$ есть не что иное, как ε-окрестность множества A.

Упражнение. Пусть A — это объединение двух отрезков на плоскости

$$[(0,0),(1,0)] \cup [(0,0),(0,1)].$$

Нарисуйте множества $2A$, $3A$, $4A$.

Вернемся к обсуждению однородной неизотропной среды. Мы выяснили, что такая среда характеризуется одной функцией $X(t)$, принимающей значения в подмножествах пространства $\mathbb{R}^n$. Принцип Гюйгенса переписывается как следующее функциональное уравнение:

$$X(t_1 + t_2) = X(t_1) + X(t_2)$$

для любых неотрицательных t_1 и t_2. Интересно найти все решения этого функционального уравнения при достаточно общих предположениях относительно функции $X(t)$. Естественно считать, что $X(t)$ зависит от t в некотором смысле непрерывно. Заметим, что функция $t \mapsto tA$ является решением нашего уравнения только в том случае, когда множество A выпукло (докажите).

Вернемся к рассмотрению общей (т.е., вообще говоря, неоднородной и неизотропной) среды. Следуя Гамильтону, определим функцию

$$W(x_0, x) = \inf\{t \mid x \in X(x_0, t)\}.$$

Другими словами, $W(x_0, x)$ — это минимальное время, за которое свет доходит от точки x_0 к точке x. Если принцип Ферма верен буквально, то это просто время, за которое свет доходит от x_0 до x, и оно автоматически минимально. В реальности это утверждение верно, вообще говоря, только если точки x_0 и x достаточно близки. Заметим, что множество $X(x_0, t)$ можно определить неравенством

$$W(x_0, x) \leq t.$$

Поверхность, заданная равенством $W(x_0, x) = t$ при фиксированном t, называется *световым фронтом.* Принцип Гюйгенса часто формулируют для световых фронтов, а не для световых пятен. В этом случае, нужно говорить об огибающей световых фронтов.

Индикатрисой $I(x_0)$ (данной сплошной среды) в точке x_0 называется объединение всех векторов (мы рассматриваем векторы как ориентированные отрезки; имеется в виду объединение соответствующих отрезков) $\vec{v}$, которые могут служить вектором скорости фотона, проходящего через точку x_0. В каждом направлении выходит ровно один такой вектор. Например, если среда изотропна, то индикатриса является шаром. В общем случае мы будем предполагать, что индикатриса ограничена гладкой выпуклой поверхностью.

Теорема 1. *Рассмотрим световой луч, выходящий из точки* x_0 *и проходящий через точку* x*. Рассмотрим также световой фронт* $\Gamma = \{x \in \mathbb{R}^n \mid W(x_0, x) = t\}$*, проходящий через точку* x*. Пусть* $\vec{v}$ *— вектор скорости света в точке* x*, касательный к данному световому лучу. Тогда касательная гиперплоскость к индикатрисе* $I(x)$ *в точке* $\vec{v} \in \partial I(x)$ *параллельна касательной гиперплоскости к световому фронту в точке* x*.*

Эта теорема очевидна в случае, когда световой фронт является плоскостью. В общем случае, гладкий световой фронт достаточно хорошо приближается плоскостью около каждой своей точки. Мы не будем давать более формального доказательства сформулированной теоремы (чтобы было формальное доказательство, нужно еще и формулировку уточнить), но интуитивно это должно быть ясно (если не ясно, попробуйте порисовать картинки).

В заключение, обудим задачу из совсем другой области, для решения которой оказываются полезными идеи из геометрической оптики. Задача состоит в том, чтобы в данном графе найти кратчайшее расстояние между двумя вершинами. Расстояние между вершинами — это длина кратчайшего пути по ребрам из одной вершины в другую, а длина пути по ребрам — это просто количество пройденных ребер. Задача нахождения кратчайшего пути в графе имеет большое практическое значение. Например, на сайте немецкой железнодорожной компании Deutsche Bahn можно найти, как проехать от одной железнодорожной станции до другой с наименьшим числом пересадок (есть еще оптимизация по времени). При этом система распознает все станции, включая остановки местных электричек и даже остановки городского транспорта в крупных городах, при том не только в Германии, а по всей Европе. Иногда минимальное количество пересадок оказывается равным 7 (например, от остановки голландской электрички до остановки местной электички где-нибудь в Шварцвальде на юге Германии). Чтобы найти маршрут с минимальным количеством пересадок, система решает задачу о кратчайшем пути в графе. В каком именно графе? — упражнение.

Допустим, мы хотим найти кратчайшее расстояние от точки x некоторого графа до другой точки y. Пишем в точке x число 0. Во всех соседних точках, в которых еще ничего не написано, пишем 1. Для каждой точки, помеченной числом 1, проводим аналогичную процедуру: во всех соседних точках, в которых еще ничего не написано, пишем 2 и т.д. Рано или поздно, в точке y окажется некоторое число. Это число и есть расстояние от точки x до точки y. Очевидна аналогия описанного процесса с процессом распространения светового фронта.

4. Семинар: принцип Гюйгенса

Задача 7. Пусть $A \subset \mathbb{R}^n$ — подмножество, удовлетворяющее соотношению $A+A = 2A$. Докажите, что если A замкнуто или открыто, то A выпукло.

Задача 8. Рассмотрим изотропную среду, занимающую плоскость $\mathbb{R}^2$. Пусть $v(x, y)$ обозначает (скалярную) скорость света в точке (x, y). Рассмотрим функцию $W(x_0, y_0, x, y)$, равную минимальному времени, которое свет может потратить на прохождение от точки (x_0, y_0) до точки (x, y). Докажите, что вектор скорости света в точке (x, y) пропорционален вектору с координатами $\frac{\partial W}{\partial x}(x_0, y_0, x, y)$, $\frac{\partial W}{\partial y}(x_0, y_0, x, y)$. Найдите коэффициент пропорциональности.

Заметим, что вектор с координатами $\frac{\partial W}{\partial x}(x_0, y_0, x, y)$ и $\frac{\partial W}{\partial y}(x_0, y_0, x, y)$ — это вектор, перпендикулярный световому фронту в точке (x, y) (имеется в виду световой фронт, выпущенный из точки (x_0, y_0) и проходящий через точку (x, y)). Это вытекает из следующего общего утверждения: вектор $(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y})$ перпендикулярен линии уровня функции f.

Пусть $\vec{v}$ — вектор скорости света в точке (x, y), направленный вдоль луча, выходящего из (x_0, y_0). Поскольку среда изотропна, все ее индикатрисы являются кругами. В частности, касательная к индикатрисе в точке $\vec{v}$ перпендикулярна вектору $\vec{v}$. Согласно теореме 1, касательная к фронту в точке (x, y) параллельна касательной к индикатрисе в точке $\vec{v}$, то есть перпендикулярна вектору $\vec{v}$. Таким образом, вектор $\vec{v}$ перпендикулярен световому фронту. Как мы уже выяснили, вектор с координатами $\frac{\partial W}{\partial x}(x_0, y_0, x, y)$ и $\frac{\partial W}{\partial y}(x_0, y_0, x, y)$ тоже перпендикулярен световому фронту. Следовательно, эти два вектора пропорциональны.

Осталось найти коэффициент пропорциональности. Для этого достаточно посчитать длину вектора с координатами $\frac{\partial W}{\partial x}(x_0, y_0, x, y)$ и $\frac{\partial W}{\partial y}(x_0, y_0, x, y)$. Это градиент функции W, равной времени, которое свет тратит на прохождение от точки (x_0, y_0) до точки (x, y). Таким образом, он приблизительно равен $\Delta t/\Delta s$, где Δs — длина малого участка светового луча, а Δt — время, за которое фотон проходит этот малый участок (мы уже знаем, что градиент функции W направлен вдоль светового луча, выходящего из точки (x, y) — мы здесь этим пользуемся). С другой стороны, скорость света в точке (x, y) приблизительно равна $\Delta s/\Delta t$. Отсюда можно заключить, что длина вектора $(\frac{\partial W}{\partial x}, \frac{\partial W}{\partial y})$ равна $1/v(x, y)$, где $v(x, y)$ — (скалярная) скорость света в точке (x, y).

Длина вектора $\vec{v}$ равна $v(x, y)$. Таким образом, коэффициент пропорциональности равен $1/v(x, y)^2$.

Задача 9. В условиях предыдущей задачи, докажите, что функция W удовлетворяет следующему уравнению с частными производными (уравнению эйконала):

$$\left(\frac{\partial W}{\partial x}\right)^2 + \left(\frac{\partial W}{\partial y}\right)^2 = \frac{1}{v(x, y)^2}.$$

В самом деле, в левой части стоит квадрат длины градиента функции W (рассматриваемой как функция от x и y при фикмированных x_0 и y_0). Но мы уже выяснили, что длина градиента равна $1/v(x, y)$. Кстати, поэтому градиент функции W иногда называют *нормальной медлительностью светового фронта.*

Задача 10. Проверьте, что функция $f(x,y) = \sqrt{x^2+y^2}$ удовлетворяет уравнению

$$\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2 = 1$$

всюду, кроме начала координат.

Задача 11. Пусть $f(x,y)$ равно минимальному расстоянию от точки (x,y) до эллипса

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1.$$

Докажите, что для всех точек во внешности эллипса выполняется уравнение из предыдущей задачи. Найдите еще какие-нибудь решения этого уравнения с частными производными.

Рассмотрим произвольную выпуклую замкнутую кривую C. У этой кривой есть внешность и внутренность. Из каждой точки кривой C, выпустим луч, перпендикулярный этой кривой. Выпуклость кривой C гарантирует, что все эти лучи не пересекаются. Как посчитать расстояние от точки (x,y) вне кривой C до кривой C? Точка (x,y) принадлежит какому-либо лучу R. Расстояние от (x,y) до кривой C равно длине отрезка луча R от его начала (которое принадлежит кривой C) до точки (x,y). Понятно, что если мы теперь пойдем вдоль луча R, то расстояние до кривой C увеличится ровно на пройденное нами расстояние. По неравенству треугольника, как бы мы ни шли, расстояние до кривой C никогда не может увеличиться больше, чем на пройденное расстояние. Отсюда вытекает, что если $f(x,y)$ — это расстояние от точки (x,y) до кривой C, то производная функции f по направлению единичного вектора вдоль луча R равна 1, а производные по направлению всех остальных единичных векторов не больше 1. Иначе говоря, максимальная скорость приращения функции f (относительно евклидовой длины) равна 1. Но модуль градиента — это как раз максимальная скорость приращения функции. Следовательно, модуль градиента функции f равен 1. Это утверждение можно переписать как следующее дифференциальное уравнение с частными производными на функцию f:

$$\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2 = 1.$$

5. Принцип наименьшего действия

Движение любой механической системы подчиняется вариационному принципу, в чем-то похожему на принцип Ферма. Этот вариационный принцип называется *принципом наименьшего действия* и часто связывается с именем Гамильтона (хотя он был ранее известен Лагранжу). Положение механической системы можно охарактеризовать точкой конфигурационного пространства. Мы сейчас будем предполагать, что конфигурационное пространство совпадает с $\mathbb{R}^n$. При этом случай $n > 3$ является физически осмысленным, поскольку механическая система может состоять более чем из одной точки. Движение системы соответствует движению точки в конфигурационном пространстве, то есть гладкой функции $\gamma : [t_0, t_1] \to \mathbb{R}^n$. Пусть фиксированы начальный и конечный момент времени t_0 и t_1, соответственно. Принцип наименьшего действия утверждает, что истинная траектория γ такова, что интеграл

$$\int_{t_0}^{t_1} L(\gamma(t), \dot{\gamma}(t))\, dt$$

(называемый *функционалом действия*) минимален при фиксированных начальном и конечном моментах времени t_0 и t_1, соответственно, а также при фиксированных начальном и конечном положениях $x_0 = \gamma(t_0)$, $x_1 = \gamma(t_1)$. Здесь L — некоторая функция на $\mathbb{R}^n \times \mathbb{R}^n$, называемая *функцией Лагранжа*.

Принцип наименьшего действия нуждается в некоторых уточнениях, аналогичных тем, которые нужно было ввести для принципа Ферма. В частности, следует говорить не о минимуме функционала действия, а о критической точке этого функционала.

Существенным отличием принципа наименьшего действия от принципа Ферма является то, что минимизация интеграла происходит при фиксированных начальном и конечном моментах времени. Впрочем, можно от времени полностью избавиться и сформулировать вариационный принцип, вполне аналогичный принципу Ферма, для нахождения формы траекторий в конфигурационном пространстве. Когда форма траекторий найдена, параметризация находится из закона сохранения энергии. Этот вариационный принцип мы сформулируем позже.

Аргументы функции Лагранжа мы будем обозначать через $x \in \mathbb{R}^n$ и $v \in \mathbb{R}^n$. При этом x имеет смысл положения (координаты), а v имеет смысл скорости. В физических убебниках пишут $L(x, \dot{x})$ вместо $L(x, v)$. Но при этом $\dot{x}$ воспринимается как независимый набор переменных, а не как вектор скорости конкретной траектории.

Теорема 2 (уравнения Эйлера–Лагранжа)**.** *Пусть* $\gamma : [t_0, t_1] \to \mathbb{R}^n$ *— достаточно гладкая траектория, минимизирующая функционал действия (мы не обсуждаем сейчас вопрос существования такой траектории). Тогда*

$$\frac{d}{dt}\frac{\partial L}{\partial v}(\gamma(t), \dot\gamma(t)) = \frac{\partial L}{\partial x}(\gamma(t), \dot\gamma(t)).$$

Доказательство. Рассмотрим однопараметрическое семейство путей (=параметризованных кривых), соединяющих точки x_0 и x_1 и запараметризованных отрезком $[t_0, t_1]$. Мы будем писать $t \mapsto \gamma(t, s)$ для обозначения пути с параметром s. Таким образом, параметр t — это параметр вдоль пути (имеющий физический смысл времени), а s — параметр, отвечающий за вариацию самого пути.

Допустим, что функция $\gamma(t, s)$ достаточное число раз дифференцируема как функция двух переменных. Дифференцирование по t будем обозначать точкой, а дифференцирование по s символом δ. Имеем

$$0 = \delta \int_{t_0}^{t_1} L(\gamma(t,s), \dot\gamma(t,s))\, dt = \int_{t_0}^{t_1} \left(\frac{\partial L}{\partial x}\delta\gamma + \frac{\partial L}{\partial v}\delta\dot\gamma\right) dt =$$

$$= \int_{t_0}^{t_1} \left(\frac{\partial L}{\partial x}\delta\gamma - \left(\frac{d}{dt}\frac{\partial L}{\partial v}\right)\delta\gamma\right) dt + L \cdot \delta\gamma\,\Big|_{t_0}^{t_1} =$$

$$= \int_{t_0}^{t_1} \left(\frac{\partial L}{\partial x} - \left(\frac{d}{dt}\frac{\partial L}{\partial v}\right)\right)\delta\gamma\, dt.$$

Первое равенство верно только для такого параметра s, для которого функционал действия на траектории $t \mapsto \gamma(t, s)$ принимает минимальное значение (мы можем, например, считать, что это минимальное значение достигается при $s = 0$). Равенство следует из того, что производная дифференцируемой функции в точке минимума должна быть равна нулю. Второе равенство — это дифференцирование под знаком интеграла. Третье равенство — интегрирование по частям. В четвертом равенстве используется то, что $\gamma(s, t_0) = x_0$ не зависит от s, и что $\gamma(s, t_1) = x_1$ тоже независит от s, а значит, $\delta\gamma = 0$ при $t = t_0$ и при $t = t_1$.

Наконец, заметим, что если $t \mapsto \gamma(t)$ — это достаточно гладкая траектория, минимизирующая функционал действия, а $t \mapsto \alpha(t)$ — любое достаточно гладкое отображение из $[t_0, t_1]$ в $\mathbb{R}^n$, то можно рассмотреть семейство траекторий

$$\gamma(t, s) = \gamma(t) + s\alpha(t),$$

для которого $\delta\gamma = \alpha$ при $s = 0$. Таким образом, мы доказали, что интеграл

$$\int_{t_0}^{t_1} \left(\frac{\partial L}{\partial x} - \left(\frac{d}{dt} \frac{\partial L}{\partial v} \right) \right) \alpha \, dt$$

равен нулю для любой достаточно гладкой функции $\alpha : [t_0, t_1] \to \mathbb{R}^n$. Отсюда следует, что подынтегральное выражение тождественно равно нулю, то есть

$$\frac{\partial L}{\partial x} - \frac{d}{dt} \frac{\partial L}{\partial v} = 0,$$

если вместо x и v (еще до того, как дифференцировать по t) подставить $\gamma(t)$ и $\dot{\gamma}(t)$. $\square$

Пусть $\gamma : [t_0, t_1] \to \mathbb{R}^n$ — гладкая траектория, которая минимизирует функционал действия среди всех траекторий, подчиняющихся условиям $\gamma(t_0) = x_0$, $\gamma(t_1) = x_1$. Будем называть такую траекторию *оптимальной*. Принцип наименьшего действия в том виде, в каком он сформулирован выше, говорит о том, что механическая система всегда движется по оптимальной траектории. Однако, в реальности это не всегда так. *Истинной траекторией* называется траектория, удовлетворяющая уравнениям Эйлера–Лагранжа. Механические системы могут двигаться по истинным, но не обязательно оптимальным траекториям. Уточненный принцип наименьшего действия гласит: движение механической системы подчиняется уравнениям Эйлера–Лагранжа, то есть представляет собой истинную траекторию.

Пример. Допустим, что лагранжиан механической системы не зависит от координат, а зависит только от скоростей (этот случай аналогичен рассмотрению оптических свойств однородной, но неизотропной среды). В этом случае, лагранжиан имеет вид $T(v)$. Мы предположим, что отображение $\frac{\partial T}{\partial v} : \mathbb{R}^n \to \mathbb{R}^n$ (то есть первый дифференциал функции T; формально говоря, это отображение из пространства $\mathbb{R}^n$ в двойственное пространство $\mathbb{R}^{n*}$) обратимо. Уравнения Эйлера–Лагранжа имеют вид

$$\frac{d}{dt} \frac{\partial T}{\partial v} = 0.$$

Это значит, что $\frac{\partial T}{\partial v}$ постоянно вдоль истинных траекторий, то есть для любой истинной траектории γ, имеем $\frac{\partial T}{\partial v}(\dot{\gamma}(t)) = const$. Отсюда и из обратимости отображения $\frac{\partial T}{\partial v}$ вытекает, что $\dot{\gamma}(t)$ не зависит от t, то есть истинное движение равномерное и прямолинейное.

Пример. Пусть $L(x, v)$ — квадратичная форма от $v \in \mathbb{R}^n$, коэффициенты которой зависят от $x \in \mathbb{R}^n$ (достаточно гладко). По теореме Эйлера об однородных функциях,

$$v \cdot \frac{\partial L}{\partial v} = 2L.$$

Вместо x и v подставим $x(t)$ и $v(t)$ — положение и скорость в момент времени t при движении вдоль истинной траектории. Теперь продифференцируем обе части равенства по t:

$$\dot{v} \cdot \frac{\partial L}{\partial v} + v \cdot \frac{d}{dt}\frac{\partial L}{\partial v} = \frac{\partial L}{\partial x} \cdot v + \frac{\partial L}{\partial v} \cdot \dot{v} + \frac{dL}{dt}.$$

Мы здесь сделали хитро: в правой части стояло $2L$, мы одно L продифференцировали по правилу сложной функции, а у другого L оставили полную производную по времени. Пользуясь уравнениями Эйлера–Лагранжа и уничтожая одинаковые члены из правой и левой частей, получаем $\frac{d}{dt}L = 0$. Это означает, что величина $L(x(t), v(t))$ постоянна вдоль истинных траекторий, то есть не зависит от времени.

Можно считать, что $L(x, v)$ — это энергия точки, движущейся свободно со скоростью v в искривленном пространстве. То, что энергия зависит от x, связано с искривлением пространства. Однако важен тот факт, что кинетическая энергия зависит квадратично от скорости. Свободное движение характеризуется отсутствием потенциальной энергии. В плоском (неискривленном) пространстве кинетическая энергия частицы имеет вид

$$L(x, v) = \frac{m|v|^2}{2},$$

где $m > 0$ — это масса частицы. Можно всегда считать, что масса равна единице. Число $|v|$ — это длина модуля скорости, измеренная относительно стандартной евклидовой метрики.

Допустим теперь, что для каждого $x \in \mathbb{R}^n$, квадратичная форма $v \mapsto L(x, v)$ обладает следующим свойством: $L(x, v) > 0$ для всех ненулевых векторов v. Тогда $L(x, v)$ является квадратом модуля вектора v относительно некоторой евклидовой метрики. Но только эта евклидова метрика зависит от точки x. Такие метрики (евклидовы метрики, зависящие от точки) называют римановыми метриками. Таким образом, наш лагранжиан является римановой метрикой. В смысле этой метрики, квадрат длины вектора v, отложенного от точки x, равен $L(x, v)$. Нетрудно посчитать длину кривой $\gamma : [t_0, t_1] \to \mathbb{R}^n$ в метрике L:

$$Length(\gamma) = \int\limits_{t_0}^{t_1} \sqrt{L(\gamma(t), \dot\gamma(t))}\, dt.$$

Таким оборазом, длина кривой тоже имеет вид действия, но это действие посчитано не для лагранжиана L, а для лагранжиана $\sqrt{L}$.

Заметим, что длина кривой не зависит от того, как эта кривая запараметризована. В частности, если $\gamma : [t_0, t_1] \to \mathbb{R}^n$ является оптимальной для лагранжиана $\sqrt{L}$, то есть является кривой наименьшей длины, соединяющей точки $x_0 = \gamma(t_0)$ и $x_1 = \gamma(t_1)$, то оптимальной будет также любая другая кривая, полученная из γ заменой параметра, то есть кривая вида $\tilde\gamma : [s_0, s_1] \to \mathbb{R}^n$, где $\tilde\gamma(s) = \gamma(h(s))$ для некоторого гомеоморфизма $h : [s_0, s_1] \to [t_0, t_1]$. Однако на кратчайшей кривой γ есть естественный параметр — так называемый натуральный параметр — относительно которого скорость постоянна, то есть $L(\gamma(t), \dot\gamma(t))$ не зависит от t. Оказывается, что истинные траектории лагранжиана L — это истинные траектории лагранжиана $\sqrt{L}$, запареметризованные натуральным параметром.

Теорема 3. *Истинные траектории лагранжиана L являются также истинными траекториями лаграндэкиана $\sqrt{L}$. Обратно, если истинная траектория лагранжиана $\sqrt{L}$ обладает тем свойством, что $L = const$ вдоль этой траектории (то есть при подстановке истинных координат и скоростей вместо аргументов функции L получается величина, не зависящая от времени), то эта траектория является также истинной траекторией лагранжиана L.*

Доказательство. Имеем:

$$\frac{\partial\sqrt{L}}{\partial x} = \frac{1}{2\sqrt{L}}\frac{\partial L}{\partial x}, \quad \frac{\partial\sqrt{L}}{\partial \dot x} = \frac{1}{2\sqrt{L}}\frac{\partial L}{\partial \dot x},$$

Запишем уравнения Эйлера–Лагранжа для лагранжиана $\sqrt{L}$:

$$\frac{d}{dt}\left(\frac{\partial\sqrt{L}}{\partial \dot x}\right) = \frac{1}{2\sqrt{L}}\left(\frac{d}{dt}\frac{\partial L}{\partial \dot x}\right) - \frac{1}{4L}\frac{\partial L}{\partial \dot x}\frac{dL}{dt} = \frac{1}{2\sqrt{L}}\frac{\partial L}{\partial x}.$$

Мы сначала должны проверить, что эти уравнения выполняются для истинной траектории лагранжиана L. Для такой траектории, мы имеем $\frac{d}{dt}L = 0$ (см. выше), а также уравнения Эйлера–Лагранжа:

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}}\right) = \frac{\partial L}{\partial x}.$$

Подставляя и то, и другое в уравнения Эйлера–Лагранжа для $\sqrt{L}$, получаем тождество. Доказательство обратного утверждения аналогично. □

6. Семинар: принцип наименьшего действия

Задача 12. Напишите систему дифференциальных уравнений в полярных координатах, которой удовлетворяют все прямые. Точнее, пусть $x(t)$, $y(t)$ координаты точки на прямой, то есть $x(t) = a + bt$, $y(t) = c + dt$ для некоторых констант a, b, c, $d \in \mathbb{R}$, таких, что $b^2 + d^2 \neq 0$. Обозначим через $r(t)$, $\phi(t)$ соответствующие полярные координаты. Напишите дифференциальные уравнения на функции $r(t)$, $\phi(t)$.

В декартовых координатах, дифференциальные уравнения прямой имеют вид

$$\ddot{x} = 0, \quad \ddot{y} = 0.$$

Это уравнения Эйлера–Лагранжа для такого лагранжиана:

$$L(x, y, \dot{x}, \dot{y}) = (\dot{x})^2 + (\dot{y})^2.$$

Запишем этот же лагранжиан в полярных координатах:

$$L = (r\dot{\phi})^2 + (\dot{r})^2.$$

Теперь осталось выписать уравнения Эйлера–Лагранжа для нового лагранжиана. Несложные вычисления дают такой результат:

$$\ddot{r} = r\dot{\phi}^2, \quad \ddot{\phi} = -\frac{\dot{r}\dot{\phi}}{r}.$$

Задача 13. Рассмотрим мыльную пленку, натянутую на некоторый контур. Предположим, что пленка однозначно проецируется на некоторую область D на плоскости xy. Таким образом, форма пленки совпадает с графиком некоторой функции $u(x, y)$. Напишите уравнение с частными производными на функцию u.

Как известно, мыльная пленка минимизирует площадь поверхности. Площадь пленки выражается следующим интегралом:

$$\iint_D \sqrt{1 + u_x^2 + u_y^2}\; dx \wedge dy.$$

Для того, чтобы в этом убедиться, достаточно посчитать площадь маленького параллелограмма со сторонами $(dx, 0, u_x\, dx)$ и

$(0, dy, u_y\, dy)$. Этот параллелограмм приближает участок пленки над прямоугольником на плоскости xy со сторонами $(dx, 0)$ и $(0, dy)$. Площадь параллелограмма с данными сторонами можно посчитать как модуль векторного произведения:

$$dS = \det \begin{vmatrix} e_x & e_y & e_z \\ 1 & 0 & u_x \\ 0 & 1 & u_y \end{vmatrix} dx\, dy = \sqrt{1 + u_x^2 + u_y^2}\, dx\, dy.$$

Здесь e_x, e_y и e_z — единичные векторы, направленные вдоль координатных осей. Таким образом, площадь выражается как интеграл по области D от функции $\sqrt{1 + u_x^2 + u_y^2}$.

Рассмотрим более общую задачу минимизации интеграла

$$\iint_D L(u, u_x, u_y)\, dx \wedge dy,$$

где L — некоторая гладкая функция от трех переменных, а u — неизвестная функция, определенная на замыкании области D и такая, что ограничение функции u на границу области D совпадает с заданной функцией (это граничное условие соответствует тому, что граница пленки закреплена на некотором контуре). Мы будем предполагать, что D имеет гладкую границу.

Мы теперь выведем дифференциальное уравнение на оптимальную функцию u, аналогичное уравнению Эйлера–Лагранжа. Рассмотрим семейство функций

$$u^t = u + tv,$$

где v — произвольным образом выбранная гладкая функция на $\overline{D}$.

Допустим, что u минимизирует интеграл от $L(u, u_x, u_y)$ по области D. Тогда

$$\frac{d}{dt} \iint_D L(u^t, u_x^t, u_y^t)\, dx \wedge dy \,|_{t=0} = 0.$$

С другой стороны, мы можем продифференцировать под знаком интеграла, то есть

$$\iint_D \left(\frac{\partial L}{\partial u} v + \frac{\partial L}{\partial u_x} v_x + \frac{\partial L}{\partial u_y} v_y \right) dx \wedge dy = 0. \tag{1}$$

Рассмотрим дифференциальную форму

$$\omega = v \left(\frac{\partial L}{\partial u_x} dy - \frac{\partial L}{\partial u_y} dx \right).$$

Дифференциаль этой формы равен

$$d\omega = \left(\frac{\partial L}{\partial u_x} v_x + \frac{\partial L}{\partial u_y} v_y + v \left(\frac{\partial}{\partial x} \frac{\partial L}{\partial u_x} + \frac{\partial}{\partial y} \frac{\partial L}{\partial u_y} \right) \right) dx \wedge dy.$$

По теореме Стокса, интеграл от формы $d\omega$ по оболасти D равен нулю (поскольку форма ω обращается в 0 на границе области D). Следовательно, форму $d\omega$ можно прибавить к подынтегральному выражению, и при этом интеграл (по области D) не изменится. Отнимем форму $d\omega$ от подынтегрального выражения в равенстве (1). Получим следующее равенство:

$$\iint_D \left(\frac{\partial L}{\partial u} - \left(\frac{\partial}{\partial x} \frac{\partial L}{\partial u_x} + \frac{\partial}{\partial y} \frac{\partial L}{\partial u_y} \right) \right) v \, dx \wedge dy = 0.$$

Поскольку интеграл обращается в 0 при любом выборе достаточно гладкой функции v, подынтегральное выражение обязано быть тождественно равным нулю. Следовательно, получаем такое уравнение:

$$\frac{\partial L}{\partial u} = \frac{\partial}{\partial x} \frac{\partial L}{\partial u_x} + \frac{\partial}{\partial y} \frac{\partial L}{\partial u_y}.$$

Это аналог уравнений Эйлера–Лагранжа для функций двух переменных. Вернемся к нашему примеру: $L = \sqrt{1 + u_x^2 + u_y^2}$. Для такого лагранжиана, уравнение Эйлера–Лагнранжа записывается следующим образом:

$$u_{xx} + u_{yy} + u_{xx} u_y^2 + u_{yy} u_x^2 - 2 u_x u_y u_{xy} = 0.$$

Это нелинейное уравнение. Если считать, что u_x и u_y малы, то это уравнение приближается уравнением Лапласа:

$$u_{xx} + u_{yy} = 0.$$

Задача 14. Найдите закон движения частицы с лагранжианом

$$L(x, y, \dot{x}, \dot{y}) = \frac{(\dot{x})^2 + (\dot{y})^2}{2} + y.$$

Задача 15. Рассмотрим лагранжиан, определенный в верхней полуплоскости формулой

$$L(x, y, \dot{x}, \dot{y}) = \frac{(\dot{x})^2 + (\dot{y})^2}{y}$$

Докажите, что истинные траектории — это полуокружности с центром на горизонтальной оси.

7. Лекция: фунция действия и гамильтониан

Основная заслуга Гамильтона в теоретической механике состоит в том, что он ввел *функцию действия*

$$S(x_0, t_0, x_1, t_1) = \inf_{\gamma} \int_{t_0}^{t_1} L(\gamma(t), \dot{\gamma}(t))\, dt.$$

Если траектория γ оптимальная, то функция действия совпадает с функционалом действия, посчитанным для этой конкретной траектории γ. В дальнейшем, при рассмотрении функции действия мы всегда будем предполагать, что оптимальная траектория существует.

Пример. Вернемся к примеру, в котором функция Лагранжа имеет вид $L(x, v) = T(v)$, причем отображение $\frac{\partial T}{\partial v} : \mathbb{R}^n \to \mathbb{R}^{n*}$ обратимо. В нашем примере, мы можем явно посчитать функцию действия. Предположим, что оптимальные траектории совпадают с истинными (это так для строго выпуклых функций T, но мы этого не доказываем). Тогда оптимальные траектории имеют вид

$$\gamma(t) = x_0 + (t - t_0)\frac{x_1 - x_0}{t_1 - t_0}.$$

В частности, скорость

$$\dot{\gamma}(t) = \frac{x_1 - x_0}{t_1 - t_0}$$

не звисит от времени. Подставим истинные траектории в функционал действия:

$$S(x_0, t_0, x_1, t_1) = \int_{t_0}^{t_1} T\left(\frac{x_1 - x_0}{t_1 - t_0}\right) dt =$$

$$= (t_1 - t_0) T\left(\frac{x_1 - x_0}{t_1 - t_0}\right).$$

Последнее равенство — это интегрирование постоянной функции по отрезку $[t_0, t_1]$. Заменяя x_1 на x и t_1 на t, получаем формулу

$$S(x_0, t_0, x, t) = (t - t_0) T\left(\frac{x - x_0}{t - t_0}\right).$$

Теперь мы будем считать, что t_0 и x_0 фиксированы, а x и t меняются.

Посчитаем частные производные функции действия по переменным x и t:

$$\frac{\partial S}{\partial x} = \frac{\partial T}{\partial v}, \quad \frac{\partial S}{\partial t} = T - \frac{\partial T}{\partial v} \cdot v.$$

Здесь вместо аргумента v у функции T следует подставить значение истинной скорости

$$v = \frac{x_1 - x_0}{t_1 - t_0}.$$

Во втором неравенстве $\frac{\partial T}{\partial v} \cdot v$ — это скалярное произведение вектора $\frac{\partial T}{\partial v}$ на вектор v (более инвариантно: линейный функционал $\frac{\partial T}{\partial v}$, примененный к вектору v). Вычисление, которое мы проделали, очень важно. Оно показывает, что производные $\frac{\partial S}{\partial x}$ и $\frac{\partial S}{\partial t}$ зависят (кроме как от x и t) только от вектора истинной скорости v, но не от начальных данных x_0, t_0. Это верно и в общем случае, для лагранжианов, зависящих от координат и скоростей, а не только от скоростей. Мы пока не будем доказывать это утверждение, но используем его в качестве мотивировки для определения гамильтониана. Неформальное определение такое: *функция Гамильтона*, или *гамильтониан*, определяется как функция $-\frac{\partial S}{\partial t}$, выраженная через координату x и *импульс* $p = \frac{\partial S}{\partial x} = \frac{\partial L}{\partial v}$. Функция гамильтона обозначается обычно через H, а ее значение в точке (x, p) — через $H(x, p)$.

Знак минус в определении гамильтониана вызывал в свое время протест у большинства математиков, и привел к ожесточенному спору между математиками и физиками. Физики настояли на своем. Причина отрицательного знака состоит в том, что гамильтониан измеряет механическую энергию, а при поднятии камня его энергия увеличивается, а не уменьшается. По поводу обозначений велся еще один спор, в котором победили, наоборот, математики. Математики обозначали координаты буквой q (от латинского qualitas), а импульсы буквой p (от латинского potentia). Физики делали наоборот: координаты обозначали буквой p, а импульсы буквой q. В настоящее время, математические обозначения являются общепринятыми. Подчеркнем еще раз, что в физических задачах $H(x, p)$ — это механическая энергия, выраженная через координаты и импульсы.

Дадим формальное определение функции Гамильтона. Мы не будем сейчас доказывать, что формула $-\frac{\partial S}{\partial t} = p \cdot v - L$ верна для любых лагранжианов, а не только для тех, которые не зависят от x. Поэтому определим H как правую часть этой формулы:

$$H(x, p) = p \cdot v(p) - L(x, v(p)).$$

Здесь $v(p)$ найдено из уравнения $\frac{\partial L}{\partial v} = p$ (мы, как и раньше, предполагаем, что это уравнение имеет единственное решение). Формально говоря, p является ковектором, то есть линейным функционалом на $\mathbb{R}^n$, то есть элементом двойственного пространства $\mathbb{R}^{n*}$. Запись

$p \cdot v$ означает результат применения функционала p к вектору v. При фиксированной системе координат, можно (но не нужно) думать про p как про вектор, а про $p \cdot v$ как про скалярное произведение векторов p и v. Наше формальное определение гамильтониана совпадает с общепринятым.

Заметим, что гамильтониан строится исключительно по лагранжиану, то есть есть конкретное преобразование, определенное на функциях от $x \in \mathbb{R}^n$ и $v \in \mathbb{R}^n$, результат применения которого к лагранжиану дает гамильтониан. Это преобразование называется *преобразованием Лежандра.* Допустим, что при фиксированном x, функция $v \mapsto L(x, v)$ *строго выпукла* (функция $f : \mathbb{R}^n \to \mathbb{R}$ называется строго выпуклой, если для любых двух точек $a, b \in \mathbb{R}^n$ весь отрезок с концами в $(a, f(a))$ и $(b, f(b))$ лежит выше графика функции f, за исключением концов) и два раза дифференцируема. Поскольку производная (=первый дифференциал) в точке максимума равна нулю, получаем уравнение $\frac{\partial L}{\partial v}(x, v(p)) = p$, которое в нашем случае должно однозначно разрешиться относительно $v(p)$ при всех p из некоторого открытого множества (см. задачу на следующем семинаре). Вообще говоря, это открытое множество не совпадает с $\mathbb{R}^n$, поэтому преобразование Лежандра будет не всюду определено. Преобразование Лежандра строго выпуклой функции L можно определить следующей формулой:

$$H(x, p) = \max_v (p \cdot v - L(x, v)).$$

Точка x в данном случае выступает просто как параметр. Поэтому можно говорить про преобразования Лежандра строго выпуклых функций от v. Преобразование Лежандра функции f обычно обозначается через $\hat{f}$. Перепишем определение:

$$\hat{f}(p) = \max_v (p \cdot v - f(v)).$$

Из определения преобразования Лежандра вытекает следующее *неравенство Юнга:*

$$\hat{f}(p) + f(v) \geq p \cdot v,$$

справедливое для любых $v \in \mathbb{R}^n$ и $p \in \mathbb{R}^{n*}$. На семинаре мы разберем примеры использования этого общего неравенства.

Теорема 4 (уравнения Гамильтона)**.** *Пусть $x(t)$, $p(t)$ — положение и импульс системы в момент времени t (разумеется, мы предполагаем, что система движется по истинной траектории). Тогда*

$$\dot{x}(t) = \frac{\partial H}{\partial p}(x(t), p(t)), \quad \dot{p}(t) = -\frac{\partial H}{\partial x}(x(t), p(t)).$$

Доказательство. Продифференцируем равенство

$$H(x,p) = p \cdot v(p) - L(x, v(p))$$

по p. Получаем:

$$\frac{\partial H}{\partial p} = v(p) + p\frac{\partial v(p)}{\partial p} - \frac{\partial L}{\partial v}(x, v(p))\frac{\partial v(p)}{\partial p} = v(p).$$

Во втором равенстве, мы воспользовались тем, что $\frac{\partial L}{\partial v}(x, v(p)) = p$. На истинной траектории, имеем $v(p) = \dot{x}(t)$. Первое равенство доказано. Заметим, что для доказательства этого равенства мы не пользовались уравнением Эйлера–Лагранжа. Следовательно, это равенство справедливо для любой траектории, не обязательно даже истинной.

Теперь продифференцируем то же равенство по x:

$$\frac{\partial H}{\partial x} = -\frac{\partial L}{\partial x}(x, v(p)).$$

Теперь воспользуемся уравнениями Эйлера–Лагранжа:

$$-\frac{\partial L}{\partial x}(x(t), v(p(t))) = -\frac{\partial L}{\partial x}(x(t), \dot{x}(t)) = -\frac{d}{dt}p(t).$$

Мы здесь должны вспомнить, что $p(t) = \frac{\partial L}{\partial v}(x(t), \dot{x}(t))$ □

Систему уравнений Гамильтона называют еще иногда системой *канонических уравнений.*

Теорема 5 (закон сохранения энергии)**.** *Пусть $x(t)$ и $p(t)$ — положение и импульс системы в момент времени t. Тогда $H(x(t), p(t))$ не зависит от t (другими словами, механическая энергия сохраняется).*

Доказательство. Действительно, продифференцируем функцию $H(x(t), p(t))$ по t:

$$\frac{d}{dt}H(x(t), p(t)) = \frac{\partial H}{\partial x}\dot{x} + \frac{\partial H}{\partial p}\dot{p} =$$

$$= \frac{\partial H}{\partial x}\left(\frac{\partial H}{\partial p}\right) + \frac{\partial H}{\partial p}\left(-\frac{\partial H}{\partial x}\right) = 0.$$

Во втором равенстве мы воспользовались каноническими уравнениями Гамильтона. □

Можно рассматривать гамильтонианы, зависящие явно от времени. Для таких гамильтонианов, закон сохранения энергии не имеет места. Он заменяется формулой $\frac{d}{dt}H(x(t), p(t), t) = \frac{\partial H}{\partial t}(x(t), p(t), t)$. В правой части имеется в виду дифференцирование только по третьему аргументу t.

8. Семинар: преобразование Лежандра и уравнения Гамильтона

Задача 16. Найдите преобразование Лежандра для следующих функций:

(a): $f(u) = e^u$
(b): $f(u) = \frac{u^\alpha}{\alpha}$

Ответ:
(*a*) $\hat{f}(p) = p(\log p - 1)$.
(*b*) $\hat{f}(p) = \frac{p^\beta}{\beta}$, где $\frac{1}{\alpha} + \frac{1}{\beta} = 1$.

Задача 17. Пусть $f : \mathbb{R}^n \to \mathbb{R}$ — дважды дифференцируемая строго выпуклая функция. Докажите, что дифференциалы функции f в различных точках пространства $\mathbb{R}^n$ не могут совпадать.

Рассмотрим множество $D \subset \mathbb{R}^{n*}$, состоящее из всех функционалов вида $d_x f$ (дифференциал функции f в точке $x \in \mathbb{R}^n$). Можно доказать, что это множество открыто. Множество D совпадает с областью определения функции $\hat{f}$.

Задача 18. Докажите, что для любого $u \in \mathbb{R}^n$ и для любого $p \in D$, выполнено неравенство

$$f(u) + \hat{f}(p) \geq p \cdot u.$$

Это неравенство называется неравенством Юнга. Если положить $f(u) = u^\alpha/\alpha$, где $\alpha > 1$, то получим классическое неравенство Юнга (мы пользуемся уже выполненным вычислением преобразования Лежандра для функции f):

$$\frac{u^\alpha}{\alpha} + \frac{p^\beta}{\beta} \geq p \cdot u, \quad \frac{1}{\alpha} + \frac{1}{\beta} = 1.$$

Заметим, что неравенство

$$f(u) + \hat{f}(p) \geq p \cdot u.$$

симметрично относительно замены u на p и f на $\hat{f}$. Отсюда можно вывести следующее:

Задача 19. Предположим, что преобразование Лежандра $\hat{f}$ от строго выпуклой достаточное число раз дифференцируемой функции f определено на всем $\mathbb{R}^n$. Докажите, что функция $\hat{f}$ тоже строго выпукла, и преобразование Лежанждра от функции $\hat{f}$ совпадает с функцией f.

Задача 20. Пусть $U \subset \mathbb{R}^n$ — открытое множество, такое, что $\lambda U \subseteq U$ для всякого $\lambda > 0$. Функция $f : U \to \mathbb{R}$ называется *положительно-однородной степени* k, если

$$f(\lambda x_1, \lambda x_2, \ldots, \lambda x_n) = \lambda^k f(x_1, \ldots, x_n)$$

для всякой точки $(x_1, \ldots, x_n) \in \mathbb{R}^n$ и всякого числа $\lambda > 0$. Докажите теорему *Эйлера об однородных функциях*: если f — дифференцируемая положительно-однородная функция степени k, то

$$\frac{\partial f}{\partial x} \cdot x = \sum_{i=1}^{n} x_i \frac{\partial f}{\partial x_i} = kf.$$

Из теоремы Эйлера об однородных функциях вытекает следующее.

Задача 21. Предположим, что лагранжиан механической системы имеет вид

$$L(x, \dot{x}) = T(x, \dot{x}) - U(x),$$

где T — квадратичная форма от $\dot{x}$, коэффициенты которой могут зависеть (гладко) от x, а U — гладкая функция от x. (Практически все лагранжианы, встречающиеся в классической механике, имеют именно такой вид). Тогда соответствующий гамильтониан равен $T + U$, причем квадратичная форма T записана в терминах импульсов p, а не скоростей.