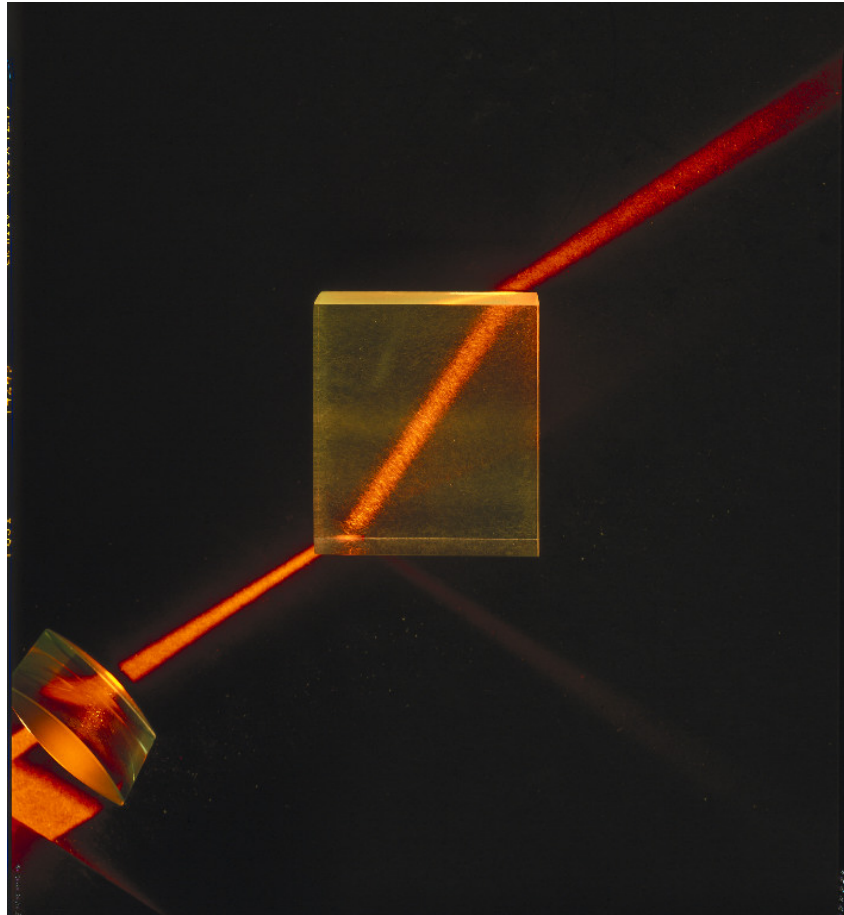
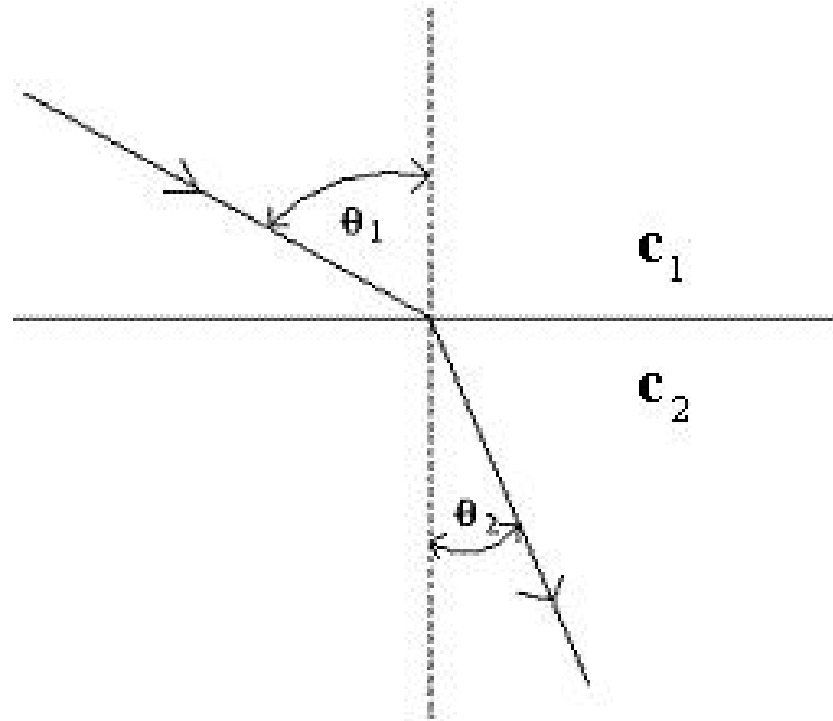


The Variational Principle, from Fermat to Feynman



Nick Sheridan

1662: Fermat points out that light always follows the path between two points that takes least time. This explains the laws of reflection and refraction.



If the speed of light in the upper medium is c_1 , and that in the lower region is c_2 , then the angle of incidence is related to the angle of refraction by **Snell's Law**:

$$\frac{\sin \theta_1}{\sin \theta_2} = \frac{c_1}{c_2}.$$

Note that if $c_1 > c_2$ then $\theta_1 > \theta_2$, as we might expect.

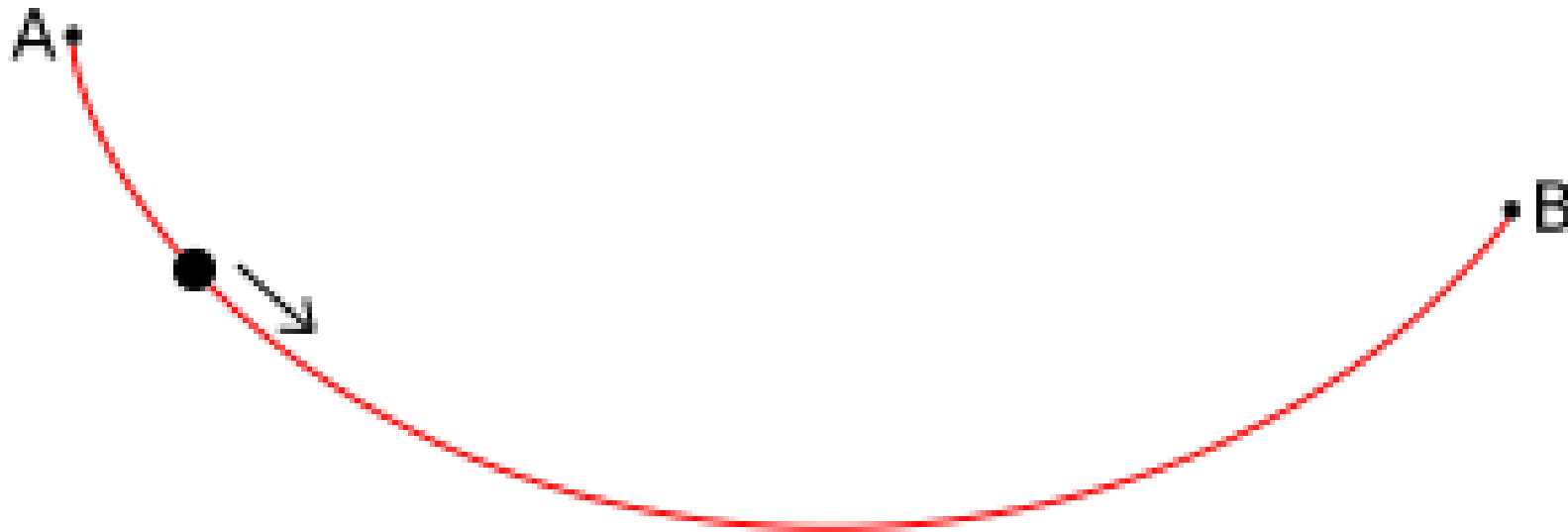
1696: Johann Bernoulli unleashes a broadside on Newton in the Calculus Wars:

I, Johann Bernoulli, address the most brilliant mathematicians in the world. Following the example set by Pascal, Fermat, etc., I hope to gain the gratitude of the whole scientific community by placing before the finest mathematicians of our time a problem which will test their methods and the strength of their intellect. If someone communicates to me the solution of the proposed problem, I shall publicly declare him worthy of praise... There are few who are likely to solve our excellent problems, aye, fewer even among the very mathematicians who boast that they have wonderfully extended its bounds by means of the golden theorems which (they thought) were known to no one, but which in fact had long previously been published by others.

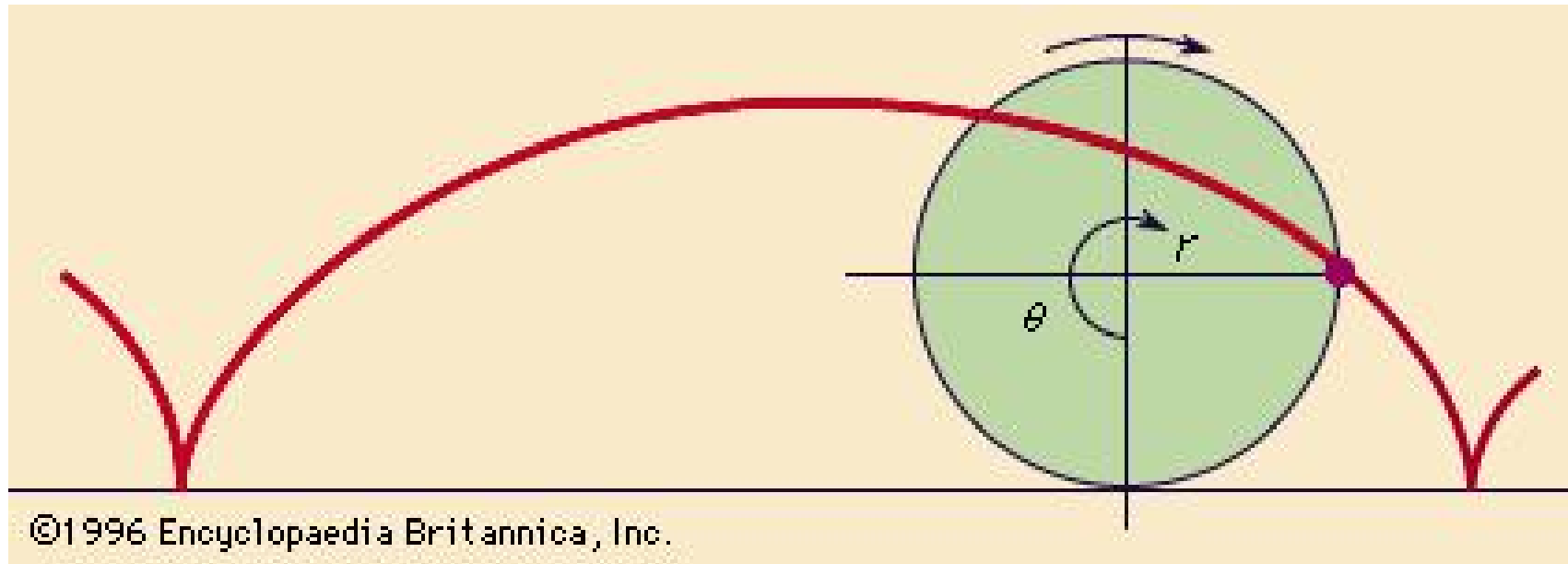
What was this guy's problem?

Problem: We are given points A and B in space. We want to build a ramp between them so that the time taken for a ball to roll from A to B is as short as possible. What shape should the ramp be?

Answer: A **brachistochrone**, or **cycloid**, or **tautochrone**! These are all names for the same curve.



It is the same as the shape traced out by a point on the circumference of a circle as it rolls along the ground:



It is also the shape of an ideal pendulum, i.e., one whose period does not depend on its amplitude.

Solutions were submitted by Newton (anonymously), Leibniz, Jacob Bernoulli and l'Hôpital.

The problem is made tricky by the fact that we must find the minimum time over all possible paths between two points in space: a vast and ugly set. In the effort to deal with this problem, the **calculus of variations** was born.

Lagrange saw the potential of this new calculus of variations. In **1756**, he communicated to Euler his new formulation of Newton's classical mechanics.

Newton's Law: Let $\mathbf{x}(t)$ be the position of a particle of mass m at time t , moving in a potential $V(\mathbf{x})$. The particle moves according to the following differential equation:

$$m \frac{d^2 \mathbf{x}}{dt^2} + \nabla V = 0.$$

Lagrange's formulation: We call the quantity

$$L = \frac{1}{2} m \left| \frac{d\mathbf{x}}{dt} \right|^2 - V(\mathbf{x}(t))$$

the **Lagrangian**. Given a path P to be followed by the particle, defined by $\mathbf{x} : [t_1, t_2] \rightarrow \mathbb{R}^3$, we define the **action** of this path to be

$$S(P) = \int_{t_1}^{t_2} L dt.$$

A particle always moves on the path between two points that **minimizes the action among all paths between those points**.

Suppose that the path P , defined by $\mathbf{x} : [t_1, t_2] \rightarrow \mathbb{R}^3$, minimizes the action. Let us consider a small variation in the path, \bar{P} , defined by $\bar{\mathbf{x}}(t) = \mathbf{x}(t) + \epsilon \mathbf{h}(t)$ where ϵ is some small number and $\mathbf{h}(t_1) = \mathbf{h}(t_2) = \mathbf{0}$. Then

$$\begin{aligned}
 S(\bar{P}) &= \int_{t_1}^{t_2} \left[\frac{1}{2} m \left(\frac{d\mathbf{x}}{dt} + \epsilon \frac{d\mathbf{h}}{dt} \right) \cdot \left(\frac{d\mathbf{x}}{dt} + \epsilon \frac{d\mathbf{h}}{dt} \right) - V(\mathbf{x}(t) + \epsilon \mathbf{h}(t)) \right] dt \\
 &= \int_{t_1}^{t_2} \left[\frac{1}{2} m \left(\left(\frac{d\mathbf{x}}{dt} \right) \cdot \left(\frac{d\mathbf{x}}{dt} \right) + 2 \frac{d\mathbf{x}}{dt} \cdot \epsilon \frac{d\mathbf{h}}{dt} \right) - V(\mathbf{x}(t)) - \nabla V \cdot \epsilon \mathbf{h}(t) + O(\epsilon^2) \right] dt \\
 &= S(P) + \epsilon \int_{t_1}^{t_2} \left[m \frac{d\mathbf{x}}{dt} \cdot \frac{d\mathbf{h}}{dt} - \nabla V \cdot \mathbf{h}(t) \right] dt + O(\epsilon^2)
 \end{aligned}$$

As P minimizes the action, we must have $S(\bar{P}) \geq S(P)$ for all ϵ . This is only possible if the term in ϵ above is 0. Thus

$$\int_{t_1}^{t_2} m \frac{d\mathbf{x}}{dt} \cdot \frac{d\mathbf{h}}{dt} - \nabla V \cdot \mathbf{h}(t) dt = 0,$$

no matter which \mathbf{h} we choose.

We integrate the first term by parts:

$$\left[m \frac{d\mathbf{x}}{dt} \cdot \mathbf{h} \right]_{t_1}^{t_2} - \int_{t_1}^{t_2} m \frac{d^2\mathbf{x}}{dt^2} \cdot \mathbf{h} + \nabla V \cdot \mathbf{h} dt = 0.$$

But recall that $\mathbf{h}(t_1) = \mathbf{h}(t_2) = 0$, thus

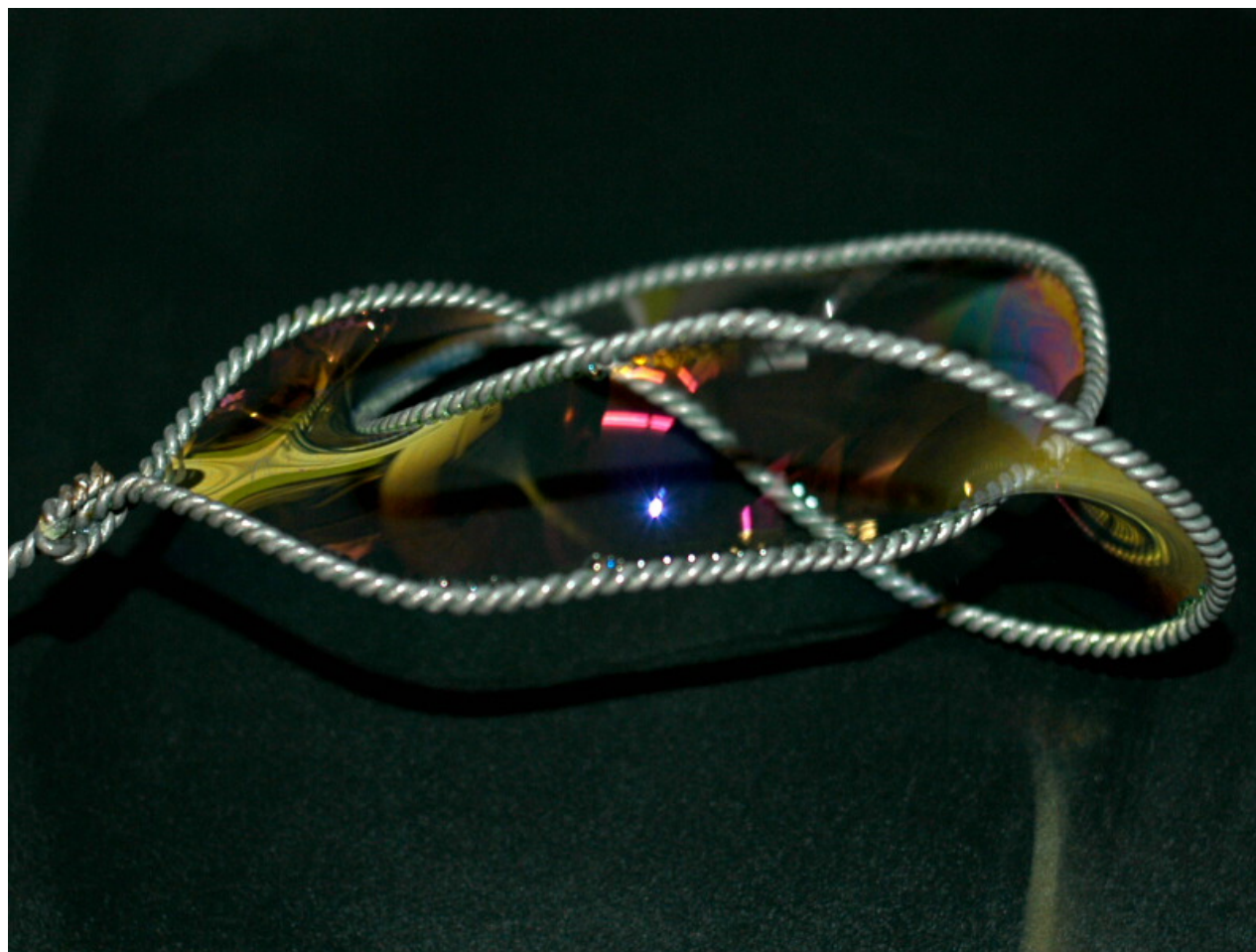
$$\int_{t_1}^{t_2} \left(m \frac{d^2\mathbf{x}}{dt^2} + \nabla V \right) \cdot \mathbf{h} dt = 0$$

no matter what \mathbf{h} is. This is only possible if

$$m \frac{d^2\mathbf{x}}{dt^2} + \nabla V = 0.$$

This is Newton's law.

The calculus of variations can also be used to show that a surface with given boundary has minimal area only if its mean curvature is zero at each point. The most well-known example of a minimal surface is a soap bubble:



Lagrange proposed in **1760** that for every circle embedded in space there is a minimal surface with that circle as boundary, a problem that became known as Plateau's problem, and was solved by Jesse Douglas and Tibor Radó in 1927.

As physics developed, it was found that all physical theories could be formulated using least-action principles:

Maxwell's electromagnetism (1860):

$$\mathcal{L} = \frac{\epsilon_0}{2} |\mathbf{E}|^2 - \frac{1}{2\mu_0} |\mathbf{B}|^2$$

where \mathbf{E} and \mathbf{B} are the electric and magnetic fields respectively, and the rest are constants.

Einstein's General Relativity (1915):

$$\mathcal{L} = kR$$

where R is the scalar curvature of spacetime and k is a constant (this Lagrangian was found by Hilbert).

Dirac's theory of the electron and subsequently quantum field theory were also most naturally formulated in terms of a Lagrangian and least-action principle.

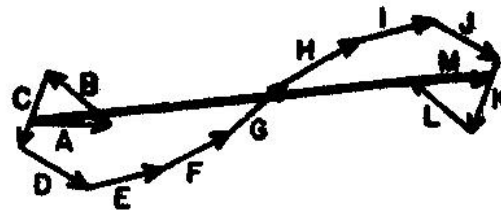
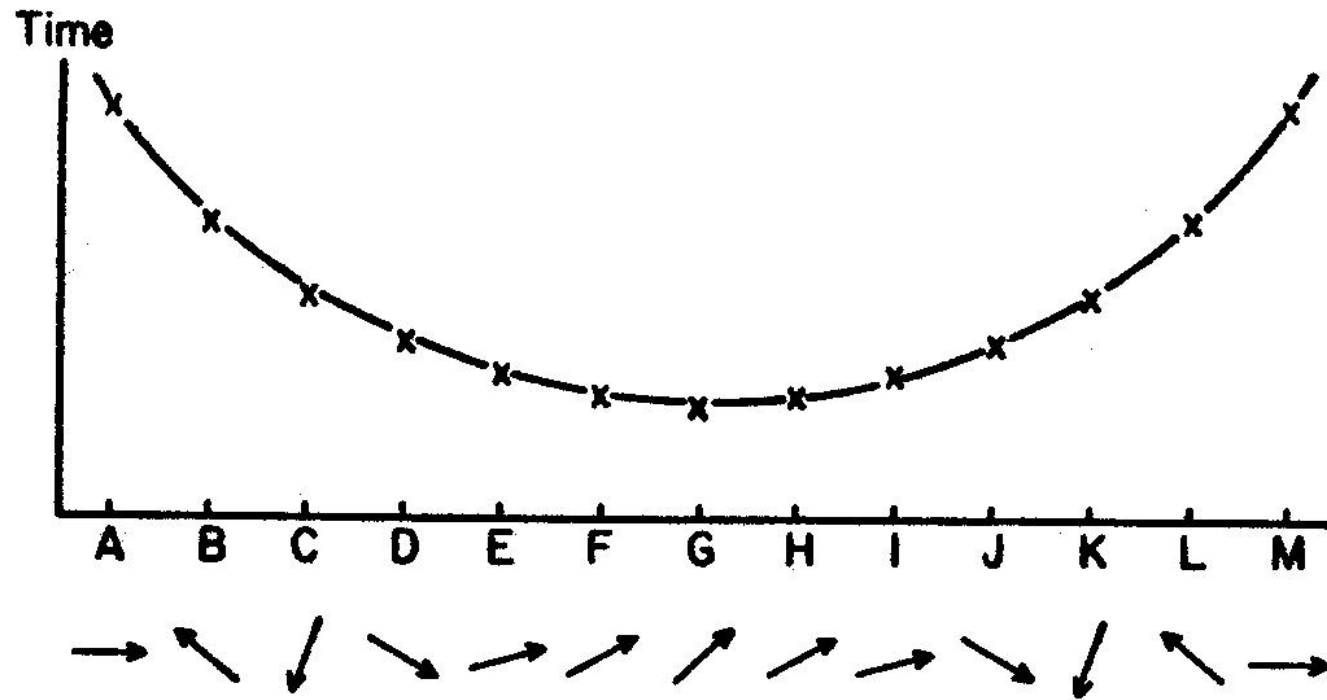
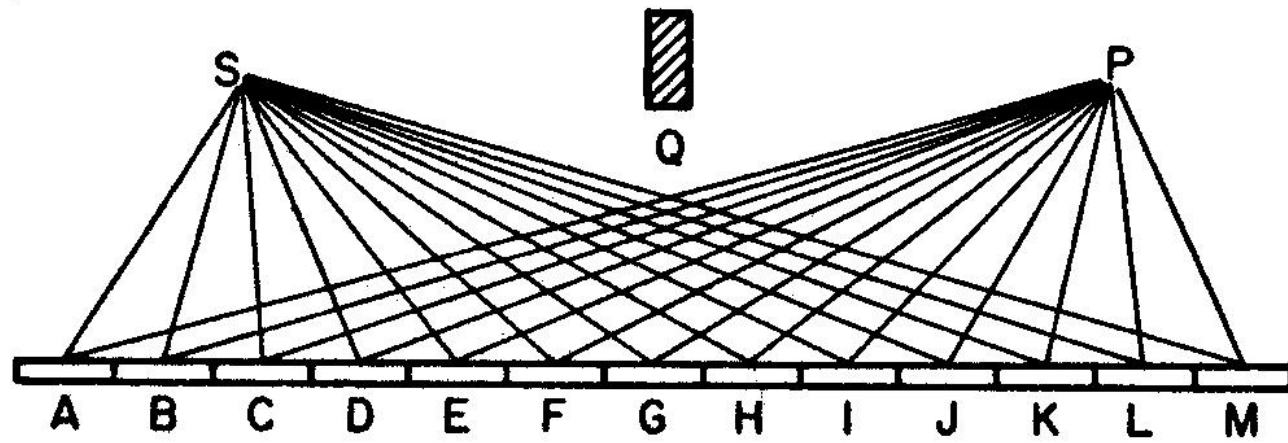
In **1948**, the physicist Richard Feynman, building on work of Dirac, showed how the magical least-action principle emerged from a more fundamental idea. Consider the example of a particle that starts at point A at time t . Feynman's proposal was:

- With every possible path P that the particle might have taken, associate an **amplitude** $e^{iS(P)/\hbar}$ where \hbar is some constant.
- For each point B , sum these amplitudes over all paths that end at B . This gives us some complex number at each point B in space.
- The modulus squared of this complex number is the probability density of finding the particle at that point.

In some sense, the particle takes **every** possible path!

What will happen when we add up all these complex numbers? For the most part they will cancel each other out, except for the contributions coming from paths close to a path of least action. Thus, these are the only paths that really matter, and this is why particles always appear to follow them.

We will show how this happens for the simple case of the reflection of light. The action of a path in this case is the time taken for the light to travel along the path. The following diagram comes from Feynman's book 'QED: The Strange Theory of Light and Matter'.



Theorem: (The stationary phase formula) Suppose $f : (a, b) \rightarrow \mathbb{R}$ has a unique critical point c in (a, b) , with $f''(c) > 0$, and $g : (a, b) \rightarrow \mathbb{R}$ vanishes with all derivatives at a, b . Then

$$\int_a^b g(x) e^{if(x)/\hbar} dx = \hbar^{1/2} g(c) e^{if(c)/\hbar} I(\hbar)$$

where $I(\hbar)$ extends to a smooth function on $[0, \infty)$ and

$$I(0) = \sqrt{2\pi} e^{i\pi/4} \frac{g(c)}{\sqrt{f''(c)}}.$$

This is a very simple example – it is helpful if the paths can be parametrized by a real number. But how do we define an “integral over all paths between two points”? No-one has worked out how to do this in a mathematically satisfactory way yet.

If we modify Feynman's rules by replacing the amplitude $e^{iS(P)/\hbar}$ with $e^{-S(P)/k}$ then the same model describes Brownian motion. It is possible to define the integral over paths properly in this case.

Is it possible to define Feynman's path integral rigorously, and can all physical rules (in particular general relativity) be explained by path integral ideas?