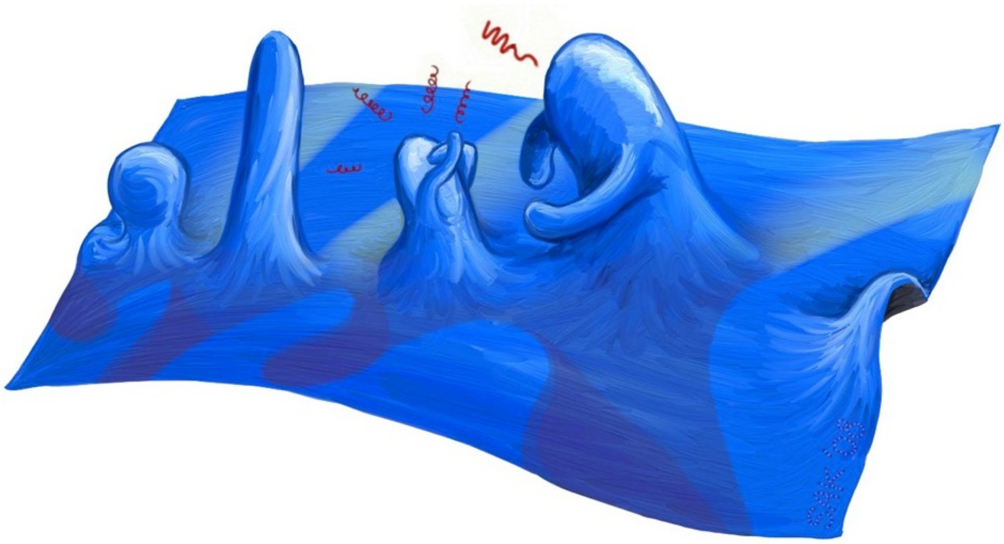
Concepts in Theoretical Physics

Part IA Mathematical Tripos



Daniel Baumann

dbaumann@damtp.cam.ac.uk

Contents

Preface **1**

1 Principle of Least Action **2**
1.1 A New Way of Looking at Things . . . 3
1.1.1 Newtonian Mechanics . . . 3
1.1.2 A Better Way . . . 3
1.1.3 Examples . . . 5
1.2 Unification of Physics . . . 7
1.3 From Classical to Quantum (and back) . . . 8
1.3.1 Sniffing Out Paths . . . 8
1.3.2 Feynman's Path Integral . . . 8
1.3.3 Seeing is Believing . . . 9

2 Quantum Mechanics **11**
2.1 The Split Personality of Electrons . . . 11
2.1.1 Particles or Waves? . . . 11
2.1.2 The Structure of Atoms . . . 12
2.2 Principles of Quantum Mechanics . . . 13
2.2.1 States are Vectors . . . 13
2.2.2 Observables are Matrices . . . 14
2.2.3 Measurements are Probabilistic . . . 15
2.2.4 Collapse of the State Vector . . . 16
2.2.5 The Uncertainty Principle . . . 17
2.2.6 Combining Systems: Entanglement . . . 17
2.3 Quantum Mechanics in Your Face . . . 19
2.3.1 The GHZ Experiment . . . 19
2.3.2 Completely Nuts, But True! . . . 21
2.3.3 Quantum Reality . . . 21

3 Statistical Mechanics **24**
3.1 More is Different . . . 25
3.2 The Distinction of Past and Future . . . 25
3.3 Entropy and The Second Law . . . 26
3.3.1 Things Always Get Worse . . . 26
3.3.2 Probability . . . 27
3.3.3 Counting . . . 28
3.3.4 Arrow of Time . . . 29
3.4 Entropy and Information . . . 31
3.4.1 Maxwell's Demon . . . 31
3.4.2 Szilard's Engine . . . 32
3.4.3 Saving the Second Law . . . 33

3.5 Entropy and Black Holes* . . . 34
3.5.1 Information Loss? . . . 34
3.5.2 Black Holes Thermodynamics . . . 34
3.5.3 Hawking Radiation . . . 35
3.5.4 Black Holes in String Theory . . . 36

4 Electrodynamics and Relativity **37**
4.1 Relativity requires Magnetism . . . 37
4.2 Magnetism requires Relativity . . . 39
4.2.1 Maxwell's Equations . . . 39
4.2.2 Let There Be Light! . . . 40
4.2.3 Racing A Light Beam . . . 42
4.2.4 My Time Is Your Space . . . 42
4.3 Relativity unifies Electricity and Magnetism . . . 43
4.3.1 Relativistic Electrodynamics . . . 43
4.3.2 A Hidden Symmetry* . . . 45
4.4 More Unification?* . . . 46
4.4.1 Kaluza-Klein Theory . . . 46
4.4.2 String Theory . . . 47

5 Particle Physics **49**
5.1 The Standard Model . . . 49
5.1.1 A New Periodic Table . . . 49
5.1.2 Four Forces Bind Them All . . . 51
5.2 From Fields to Particles . . . 51
5.3 From Symmetries to Forces . . . 52
5.4 From Virtual Particles to Real Forces . . . 54
5.4.1 Heisenberg's Uncertainty Principle . . . 54
5.4.2 Feynman Diagrams . . . 55
5.4.3 QED . . . 56
5.4.4 QCD and Confinement . . . 57
5.4.5 Why the Sun Shines* . . . 58
5.5 The Origin of Mass . . . 59
5.5.1 An Analogy . . . 59
5.5.2 Beyond Cartoons* . . . 61
5.5.3 Discovery of the Higgs . . . 62
5.6 Beyond the Standard Model* . . . 65
5.6.1 Dark Matter . . . 65
5.6.2 Supersymmetry . . . 65

6 General Relativity **66**
6.1 The Happiest Thought . . . 68
6.1.1 Fictitious Forces . . . 68
6.1.2 Equivalence Principle . . . 68
6.1.3 Black Holes . . . 71
6.2 Gravity as Geometry . . . 72

6.2.1 Spacetime . . . 72
6.2.2 Matter Tells Space How To Curve 73
6.2.3 ... Space Tells Matter How To Move. . . . 74
6.3 Gravitational Waves . . . 75
6.4 Quantum Gravity* . . . 77

7 Cosmology **78**
7.1 The Big Bang . . . 80
7.1.1 Nucleosynthesis . . . 81
7.1.2 Recombination . . . 82
7.1.3 Cosmic Microwave Background . . . 82
7.1.4 Gravitational Instability . . . 83
7.2 The Horizon Problem . . . 84
7.3 Inflation . . . 85
7.3.1 Solution of the Horizon Problem . . . 85
7.3.2 The Physics of Inflation . . . 86
7.4 From So Simple A Beginning . . . 87
7.4.1 Quantum Fluctuations . . . 87
7.4.2 CMB Anisotropies . . . 87
7.5 Breaking News: BICEP2 . . . 89
7.5.1 B-modes from Gravitational Waves . . . 89
7.5.2 Have They Been Seen? . . . 90
7.6 A Puzzle and A Mystery . . . 91
7.6.1 Dark Matter . . . 91
7.6.2 Dark Energy . . . 91

A Symmetries and Conservation Laws **94**
A.1 Conservation of Momentum . . . 94
A.2 Noether's Theorem . . . 95
A.3 Conservation of Energy . . . 96

References **99**

Preface

"If you decide you don't have to get A's, you can learn an enormous amount in college."

Isidor Isaac Rabi.

This course is intended to be purely for fun and entertainment. I want to show you a few of the wonderful things that lie ahead in your studies of theoretical physics. We will talk about some of the most mind-boggling and beautiful aspects of modern physics. We will discuss the physics of the very small (quantum mechanics and particle physics), the physics of the very big (general relativity and cosmology), the very fast (special relativity), and the very complex (statistical mechanics).

I will try to strike a balance between keeping things simple and staying honest. In particular, I will make an effort to avoid oversimplifying analogies. Instead I will look for ways of giving you a gist of the real deal without too much lying. This won't always possible, either because it would take too much time to present a detailed argument or because you haven't yet been exposed to the advanced mathematics used in modern physics. In that case, I will try to find words and diagrams to describe the missing equations and I ask you to focus on the big picture, rather than the technical details. You will have the opportunity to catch up on those details in future courses.

I sincerely hope that these notes will be easy to read. A few words of advice: don't obsess about details! You don't have to understand every line and every equation. You are welcome to focus on the parts that you enjoy and ignore the rest. Having said that, many of the ideas that I will present in these lectures are so beautiful, that it will be worth making some effort to understand them. If you don't understand an equation, just try to understand the words surrounding it. In fact, if you want, you can view any equations in this course simply as art and listen to my words to get a sense for their meanings.

To show you a bit more of the mathematics behind the ideas, I will include extra material as short digressions. These parts will be separated from the main text by boxes. Most of this material won't be covered live in the lectures, but is for you to read in your own time.

Sections that are marked by a '$*$' are even more non-examinable than the rest of the course.

Acknowledgements. Thanks to Prof. David Tong. I have stolen shamelessly from his notes for a previous version of the course. Whatever good jokes are in these notes are most likely his.

1 Principle of Least Action

You have all suffered through a course on Newtonian mechanics. You therefore all know how to calculate the way things move: you draw a pretty picture; you draw arrows representing forces; you add them all up; and then you use $F = ma$ to figure out where things are heading next. All of this is rather impressive—it really is the way the world works and we can use it to understand things about Nature. For example, showing how the inverse square law for gravity explains Kepler's observations of planetary motion is one of the great achievements in the history of science.

However, there's a better way to do it. This better way was found about 150 years after Newton, when classical mechanics was reformulated by some of the giants of mathematical physics—people like Lagrange, Euler and Hamilton. This new way of doing things is better for a number of reasons:

- Firstly, it's elegant. In fact, it's not just elegant: it's completely gorgeous. In a way that theoretical physics should be, and usually is, and in a way that the old Newtonian mechanics really isn't.

- Secondly, it's more powerful. It gives new methods to solve hard problems in a fairly straightforward manner. Moreover, it is the best way of exploiting the symmetries of a problem (see Appendix A). And since these days all of physics is based on symmetry principles, it is really the way to go.

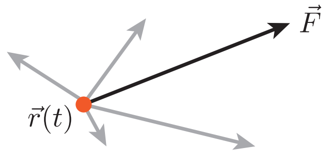
- Finally, and most importantly, it is universal. It provides a framework that can be extended to all other laws of physics, and reveals a deep relationship between classical mechanics and quantum mechanics. This is the real reason why it's so important.

In this lecture, I'll show you the key idea that leads to this new way of thinking. It's one of the most profound results in physics. But, like many profound results, it has a rubbish name. It's called the **principle of least action**.

1.1 A New Way of Looking at Things

1.1.1 Newtonian Mechanics

Let's start simple. Consider a single particle at position $\vec{r}(t)$, acted upon by a force $\vec{F}$. You might have gotten that force by adding up a bunch of different forces



Sir Isaac tells us that

$$\vec{F} = m\vec{a} = m\ddot{\vec{r}} \,. \tag{1.1.1}$$

The goal of classical mechanics is to solve this differential equation for different forces: gravity, electromagnetism, friction, etc. For conservative forces (gravity, electrostatics, but not friction), the force can be expressed in terms of a potential,

$$\vec{F} = -\vec{\nabla} V \ . \tag{1.1.2}$$

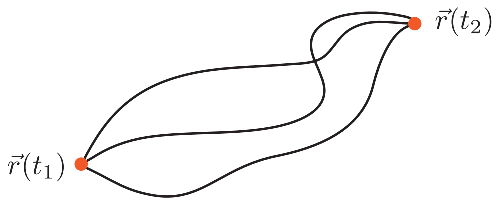
The potential $V(\vec{r})$ depends on $\vec{r}$, but not on $\dot{\vec{r}}$. Newton's equation then reads

$$m\ddot{\vec{r}} = -\vec{\nabla} V \ . \tag{1.1.3}$$

This is a second-order differential equation, whose general solution has two integration constants. Physically this means that we need to specify the initial position $\vec{r}(t_1)$ and the initial velocity $\dot{\vec{r}}(t_1)$ of the particle to figure out where it is going to end up.

1.1.2 A Better Way

Instead of specifying the initial position and velocity, let's instead choose to specify the initial and *final* positions, $\vec{r}(t_1)$ and $\vec{r}(t_2)$, and consider the possible paths that connect them:



What path does the particle actually take?

Let's do something strange: to each path $\vec{r}(t)$, we assign a number which we call the *action*,

$$S[\vec{r}(t)] = \int_{t_1}^{t_2} \mathrm{d}t \left(\tfrac{1}{2} m\dot{\vec{r}}^{\,2} - V(\vec{r}) \right) \ . \tag{1.1.4}$$

The integrand is called the *Lagrangian* $L = \frac{1}{2}m\dot{\vec{r}}^{\,2} - V(\vec{r})$. The *minus* sign is not a typo. It is crucial that the Lagrangian is the *difference* between the kinetic energy (K.E.) and the potential energy (P.E.) (while the total energy is the *sum*). Here is an astounding claim:

The true path taken by the particle is an extremum of the action.

Let us prove it: you know how to find the extremum of a function—you differentiate and set it equal to zero. But the action not a function, it is a *functional*—a function of a function. That makes it a slightly different problem. You will learn how to solve problems of this type in next year's "methods" course, when you learn about "calculus of variations". It is really not as fancy as it sounds, so let's just do it for our problem: consider a given path $\vec{r}(t)$. We ask how the action changes when we change the path slightly

$$\vec{r}(t) \to \vec{r}(t) + \delta\vec{r}(t) \ , \tag{1.1.5}$$

but in such a way that we keep the end points of the path fixed

$$\delta\vec{r}(t_1) = \delta\vec{r}(t_2) = 0 \ . \tag{1.1.6}$$

The action for the perturbed path $\vec{r} + \delta\vec{r}$ is

$$S[\vec{r} + \delta\vec{r}] = \int_{t_1}^{t_2} \mathrm{d}t \left[\tfrac{1}{2}m\left(\dot{\vec{r}}^2 + 2\dot{\vec{r}}\cdot\delta\dot{\vec{r}} + \delta\dot{\vec{r}}^2\right) - V(\vec{r} + \delta\vec{r})\right] \ . \tag{1.1.7}$$

We can Taylor expand the potential

$$V(\vec{r} + \delta\vec{r}) = V(\vec{r}) + \vec{\nabla}V\cdot\delta\vec{r} + \mathcal{O}(\delta\vec{r}^2) \ . \tag{1.1.8}$$

Since $\delta\vec{r}$ is infinitesimally small, we can ignore all terms of quadratic order and higher, $\mathcal{O}(\delta\vec{r}^2, \delta\dot{\vec{r}}^2)$. The difference between the action for the perturbed path and the unperturbed path then is

$$\delta S \equiv S[\vec{r} + \delta\vec{r}] - S[\vec{r}] \ = \int_{t_1}^{t_2} \mathrm{d}t \left[m\dot{\vec{r}}\cdot\delta\dot{\vec{r}} - \vec{\nabla}V\cdot\delta\vec{r}\right] \tag{1.1.9}$$

$$= \int_{t_1}^{t_2} \mathrm{d}t \left[-m\ddot{\vec{r}} - \vec{\nabla}V\right]\cdot\delta\vec{r} + \left[m\dot{\vec{r}}\cdot\delta\vec{r}\right]_{t_1}^{t_2} \ . \tag{1.1.10}$$

To go from the first line to the second line we have integrated by parts. This picks up a term that is evaluated at the boundaries t_1 and t_2. However, this term vanishes since the end points are fixed, i.e. $\delta\vec{r}(t_1) = \delta\vec{r}(t_2) = 0$. Hence, we get

$$\delta S \ = \int_{t_1}^{t_2} \mathrm{d}t \left[-m\ddot{\vec{r}} - \vec{\nabla}V\right]\cdot\delta\vec{r} \ . \tag{1.1.11}$$

The condition that the path we started with is an extremum of the action is

$$\delta S = 0 \ . \tag{1.1.12}$$

This should hold for *all* changes $\delta\vec{r}(t)$ that we could make to the path. The only way this can be true is if the expression in $[\cdots]$ is zero. This means

$$m\ddot{\vec{r}} = -\vec{\nabla}V \ . \tag{1.1.13}$$

But this is just Newton's equation (1.1.3)! Requiring that the action is an extremum is equivalent to requiring that the path obeys Newton's equation. It's magical.

Comments:

- From the Lagrangian $L(\vec{r}, \dot{\vec{r}}, t)$ we can define a generalized momentum and a generalized force,

$$\vec{\mathsf{P}} \equiv \frac{\partial L}{\partial \dot{\vec{r}}} \qquad \text{and} \qquad \vec{\mathsf{F}} \equiv \frac{\partial L}{\partial \vec{r}} \ . \tag{1.1.14}$$

  In our simple example, $L = \frac{1}{2} m \dot{\vec{r}}^2 - V(\vec{r})$, these definitions reduce to the usual ones: $\vec{\mathsf{P}} = m\dot{\vec{r}}$ and $\vec{\mathsf{F}} = -\vec{\nabla} V$. (For more complicated examples, $\vec{\mathsf{P}}$ and $\vec{\mathsf{F}}$ can be less recognizable.) Newton's equation can then be written as

$$\frac{d\vec{\mathsf{P}}}{dt} = \vec{\mathsf{F}} \quad \Leftrightarrow \quad \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\vec{r}}}\right) = \frac{\partial L}{\partial \vec{r}} \ . \tag{1.1.15}$$

  This is called the *Euler-Lagrange equation.*

- The principle of least action easily generalizes to systems with more than one particle. The total Lagrangian is simply the sum of the Lagrangians for each particle

$$L = \sum_{i=1}^{N} \frac{1}{2} m_i \dot{\vec{r}}_i^2 - V(\{\vec{r}_i\}) \ . \tag{1.1.16}$$

  Each particle obeys its own equation of motion

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\vec{r}}_i}\right) = \frac{\partial L}{\partial \vec{r}_i} \ . \tag{1.1.17}$$

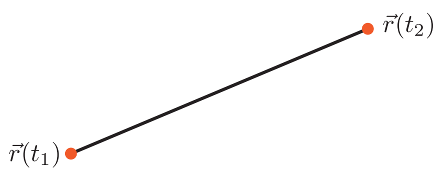
1.1.3 Examples

To get a bit more intuition for this strange way of doing mechanics, let us discuss two simple examples:

- Consider the Lagrangian of a free particle

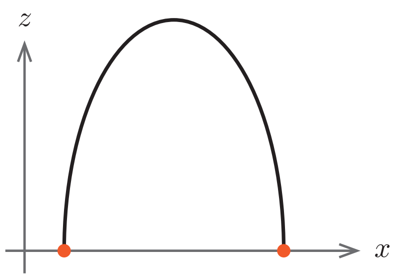
$$L = \tfrac{1}{2} m \dot{\vec{r}}^2 \ . \tag{1.1.18}$$

  In this case the least action principle implies that we want to minimize the kinetic energy over a fixed time. It is reasonable to expect that the particle must take the most direct route, which is a straight line:



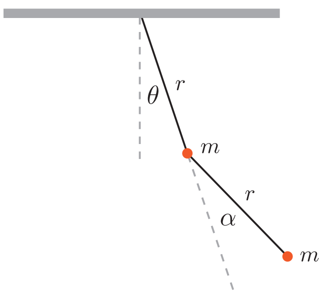
  But, do we slow down to begin with, and then speed up? Or, do we accelerate like mad at the beginning and slow down near the end? Or, do we go at a uniform speed? The average speed is, of course, fixed to be the total distance over the total time. So, if you do anything but go at a uniform speed, then sometimes you are going too fast and sometimes you are going too slow. But, the mean of the square of something that deviates around an average, is always greater than the square of the mean. (If this isn't obvious, please think about it for a little bit.) Hence, to minimize the integrated kinetic energy—i.e. the action—the particle should go at a uniform speed. In the absence of a potential, this is of course what we should get.

- As a slightly more sophisticated example, consider a particle in a uniform gravitational field. The Lagrangian is
$$L = \tfrac{1}{2}m\dot{x}^2 + \tfrac{1}{2}m\dot{z}^2 - mgz \ . \tag{1.1.19}$$
Imagine that the particle starts and ends at the same height $z_0 = z(t_1) = z(t_2)$, but moves horizontally from $x_1 = x(t_1)$ to $x_2 = x(t_2)$. This time we don't want to go in a straight line. Instead, we can minimize the difference between K.E. and P.E. if we go up, where the P.E. is bigger. But we don't want to go too high either, since that requires a lot of K.E. to begin with. To strike the right balance, the particle takes a parabola:



At this point, you could complain: what is the big deal? I could easily do all of this with $F = ma$. While this is true for the simple examples that I gave so far, once the examples become more complicated, $F = ma$ requires some ingenuity, while the least action approach stays completely fool-proof. No matter how complicated the setup, you just count all kinetic energies, subtract all potential energies and integrate over time. No messing around with vectors. You will see many examples of the power of the least action approach in future years. I promise you that you will fall in love with this way of doing physics!

Exericise.—You still don't believe me that Lagrange wins over Newton? Then look at the following example:



Derive the equations of motion of the two masses of the double pendulum.
Hint: Show first that the Lagrangian is

$$L = mr^2\left[\dot{\theta}^2 + \frac{1}{2}(\dot{\theta}+\dot{\alpha})^2 + \dot{\theta}(\dot{\theta}+\dot{\alpha})\cos\alpha\right] - mgr\left[2\cos\theta + \cos(\theta-\alpha)\right] \ .$$

Try doing the same the Newtonian way.

Another advantage of the Lagrangian method is that it reveals a deep connection between symmetries and conservation laws. I describe this in Appendix A. You should read that on you own.

1.2 Unification of Physics

The Lagrangian method has taken over all of physics, not just mechanics. All fundamental laws of physics can be expressed in terms of a least action principle. This is true for electromagnetism, special and general relativity, particle physics, and even more speculative pursuits that go beyond known laws of physics such as string theory.

To really explain this requires many concepts that you don't know yet. Once you learn these things, you will be able to write all of physics on a single line. For example, (nearly) every experiment ever performed can be explained by the Lagrangian of the Standard Model

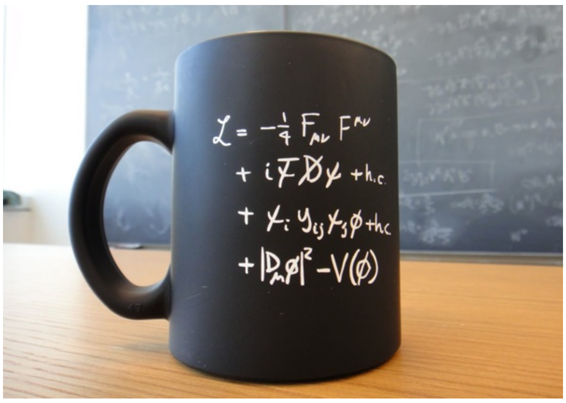
$$\mathcal{L} \sim \underbrace{R}_{\text{Einstein}} - \underbrace{\overbrace{\tfrac{1}{4}F_{\mu\nu}F^{\mu\nu}}^{\text{Maxwell}}}_{\text{Yang-Mills}} + \underbrace{i\bar{\psi}\gamma^{\mu}D_{\mu}\psi}_{\text{Dirac}} + \underbrace{|D_{\mu}h|^2 - V(|h|)}_{\text{Higgs}} + \underbrace{h\bar{\psi}\psi}_{\text{Yukawa}} \quad . \tag{1.2.20}$$

Don't worry if you don't understand many of the symbols. You are not supposed to. View this equation like art.

Let me at least tell you what each of the terms stands for:

- The first two terms characterize all fundamental forces in Nature: The term 'Einstein' describes gravity. Black holes and the expansion of the universe follow from it (see Lectures 6 and 7).
- The next term, 'Maxwell', describes electric and magnetic forces (which, as we will see in Lecture 4, are just different manifestations of a single (electromagnetic) force). A generalization of this, 'Yang-Mills', encodes the strong and the weak nuclear forces (we will see what these are in Lecture 5).
- The next term, 'Dirac', describes all matter particles (collectively called fermions)—things like electrons, quarks and neutrinos. Without the final two terms, 'Higgs' and 'Yukawa', these matter particles would be massless.

The Lagrangian (1.2.20) is a compact way to describe all of fundamental physics. You can print it on a T-shirt or write it on a coffee mug:



1.3 From Classical to Quantum (and back)

1.3.1 Sniffing Out Paths

As we have seen, the principle of least action gives a very different way of looking at things:

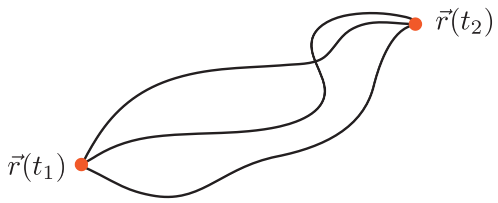
- In the Newtonian approach, the intuition for how particles move goes something like this: at each moment in time, the particle thinks "where do I go now?". It looks around, sees the potential, differentiates it and says "ah-ha, I go this way." Then, an infinitesimal moment later, it does it all again.
- The Lagrangian approach suggests a rather different viewpoint: Now the particle is taking the path which minimizes the action. How does it know this is the right path? It is sniffing around, checking out all paths, before it decides: "I think I'll go this way".

At some level, this philosophical pondering is meaningless. After all, we proved that the two ways of doing things are completely equivalent. This changes when we go beyond classical mechanics and discuss **quantum mechanics**. There we find that the particle really does sniff out every possible path!

1.3.2 Feynman's Path Integral

> *"Thirty-one years ago, Dick Feynman told me about his "sum over histories" version of quantum mechanics. "The electron does anything it likes," he said. "It just goes in any direction at any speed, . . . however it likes, and then you add up the amplitudes and it gives you the wavefunction." I said to him, "You're crazy." But he wasn't."*
>
> Freeman Dyson

I have been lying to you. It is not true that a free particle only takes one path (the straight line) from $\vec{r}(t_1)$ to $\vec{r}(t_2)$. Instead, according to quantum mechanics it takes *all possible paths*:



It can even do crazy things, like go backwards in time or first go to the Moon. According to quantum mechanics all these things can happen. But they happen *probabilistically*. At the deepest level, Nature is probabilistic and things happen by random chance. This is the key insight of quantum mechanics (see Lecture 2).

The probability that a particle starting at $\vec{r}(t_1)$ will end up at $\vec{r}(t_2)$ is expressed in terms of an *amplitude* $\mathcal{A}$, which is a complex number that can be thought of as the square root of the probability

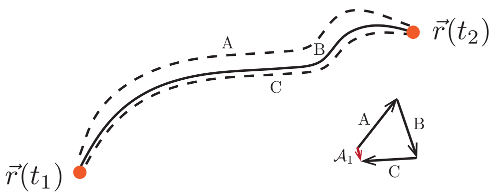
$$\text{Prob} = |\mathcal{A}|^2 \, . \tag{1.3.21}$$

To compute the amplitude, you must sum over all path that the particle can take, weighted by a *phase*

$$\mathcal{A} = \sum_{\text{paths}} e^{iS/\hbar} \, . \tag{1.3.22}$$

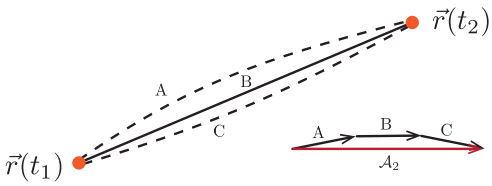
Here, the phase for each path is just the value of the action for the path in units of Planck's constant $\hbar = 1.05 \times 10^{-34}$J·s. Recall that complex numbers can be represented as little arrows in a two-dimensional xy-plane. The length of the arrow represents the magnitude of the complex number, the phase represents its angle relative to say the x-axis. Hence, in (1.3.22) each path can be represented by a little arrow with phase (angle) given by $S/\hbar$. Summing all arrows and then squaring gives the probability. This is Feynman's *path integral approach to quantum mechanics*. Sometimes it is called *sum over histories*.

The way to think about this is as follows: when a particle moves, it really does take all possible paths. However, away from the classical path (i.e. far from the extremum of the action), the action varies wildly between neighbouring paths, and the sum of different paths therefore averages to zero:



i.e. far from the classical path, the little arrows representing the complex amplitudes of each path point in random directions and cancel each other out.

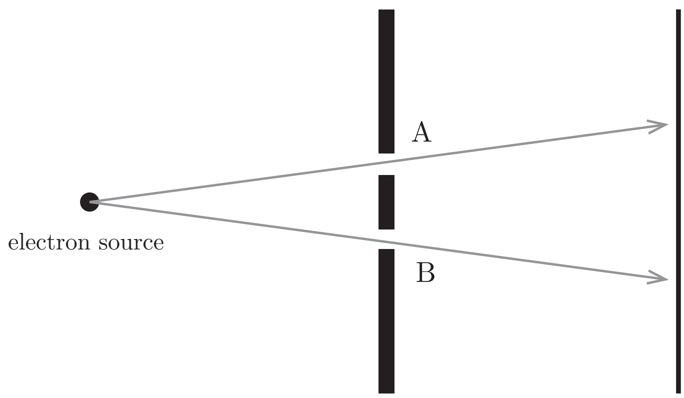
Only near the classical path (i.e. near the extremum of the action) do the phases of neighbouring paths reinforce each other:



i.e. near the classical path, the little arrows representing the complex amplitudes of each path point in the same direction and therefore add. This is how classical physics emerges from quantum physics. It is totally bizarre, but it is really the way the world works.

1.3.3 Seeing is Believing

You don't believe me that quantum particles really take all possible paths? Let me prove it to you. Consider an experiment that fires electrons at a screen with two slits, "one at a time":



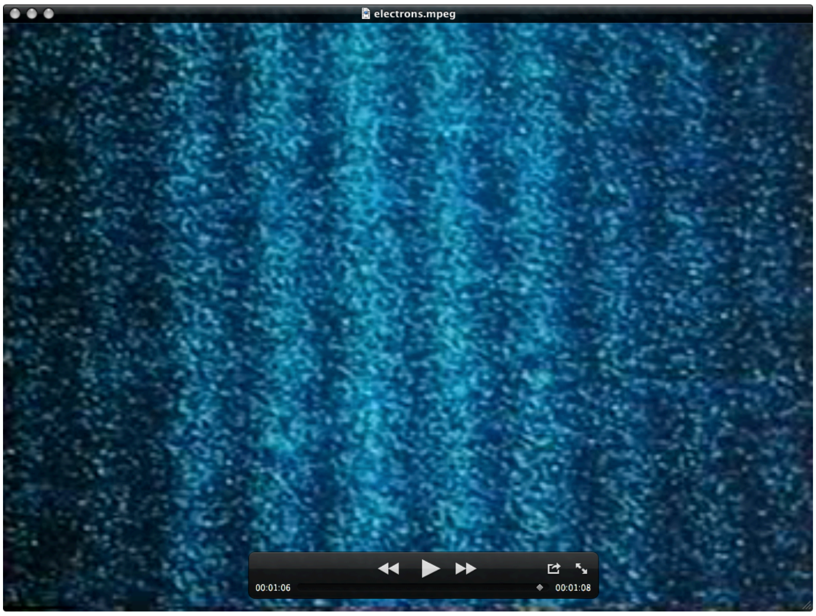
In Newtonian physics, each electron either takes path A OR B. We then expect the image behind the two slits just to be the sum of electrons going through slit A or B, i.e. we expect the electrons to build up a pair of stripes—one corresponding to electrons going through A, and one corresponding to electrons going through B.

Now watch this video of the actual experiment:

www.damtp.cam.ac.uk/user/db275/electrons.mpeg

We see that the electrons are recorded one by one:

Slowly a pattern develops, until finally we see this:



We don't just see two strips, but a series of strips. In quantum physics, each electron seems to take the paths A AND B simultaneously, and interferes with itself! (Since the electrons are fired at the slits one by one, there is nobody else around to interfere with.) It may seem crazy, but it is really the way Nature ticks. You will learn more about this in various courses on **quantum mechanics** over the next few years. I will tell you a little bit more about the strange quantum world in Lecture 2.

2 Quantum Mechanics

"If quantum mechanics hasn't profoundly shocked you, you haven't understood it."

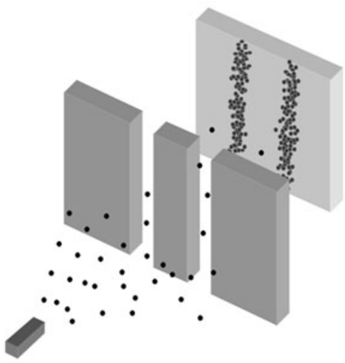
Niels Bohr

Today, I will tell you more about **quantum mechanics**—what weird thing it is and why it is so weird. The importance of quantum mechanics cannot be overstated. It explains the structure of atoms, it describes how elementary particles interact (Lecture 5) and it may even be the ultimate origin of all structure in the universe (Lecture 7). And yet it is so mind-boggling strange and bizarre that it is very difficult to come to terms with. To the point where it suggests that, ultimately, our brains may not be best equipped to understand what's going on! On the other hand, if we just "shut up and calculate", quantum mechanics simply works. In fact, it is arguably the most successful scientific theory ever developed. In this lecture, I will try to give you a glimpse of the wonderful world of quantum mechanics. This will be, by far, the most challenging lecture of the course. Are you ready?

2.1 The Split Personality of Electrons

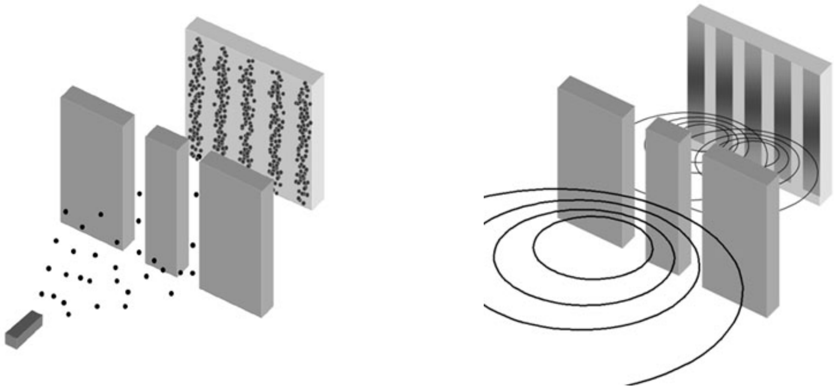
2.1.1 Particles or Waves?

Electrons (and other elementary particles) behave like nothing you have ever seen before. I have demonstrated this to you at the end of the last lecture when we looked at an experiment in which electrons were fired at a screen with two slits. In a world governed by classical (i.e. non-quantum) mechanics, there is a clear expectation for what should happen: electrons should behave like bullets. The bullets go through either one slit or the other, and we expect them just to pile up behind the two slits:



We would predict that the pattern behind the two slits simply to be the sum of the patterns for each slit considered separately: if half the bullets were fired with only the left slit open and then half were fired with just the right slit open, the result would be the same.

But, I showed you what happens in the actual experiment: The electrons, like bullets, strike the target one at a time. We slowly see a pattern building up—yet, it is *not* the pattern that we would expect if electrons behaved like bullets. Instead the pattern we get for electrons looks more like the pattern we would get for waves:



With waves, if the slits were opened one at a time, the pattern would resemble that for particles: two distinct peaks. But when the two slits are open at the same time, the waves pass through both slits at once and interfere with each other: where they are in phase they reinforce each other; where they are out of phase they cancel each other out. Electrons seem to do the same. However, if each electron passes individually through one slit, what does it "interfere" with? Although each electron arrives at the target at a single place and time, it seems that each has passed through both slits at once. Remember that in Lecture 1 we claimed that electrons "sniff out all possible paths." In the double slit experiment we see this in action.

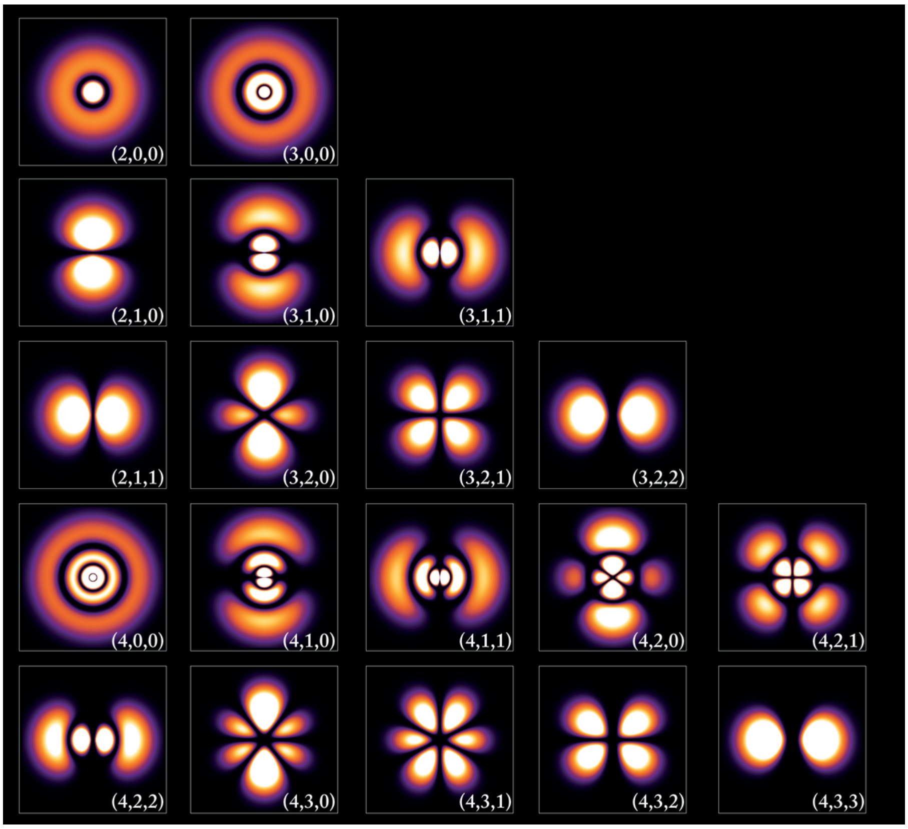
2.1.2 The Structure of Atoms

This behaviour of electrons determines the structure of atoms:

In classical physics, we have a simple (and wrong!) mental picture of the hydrogen atom:

The electric force keeps an electron in orbit around a proton, just like the gravitational force keeps the Earth in orbit around the Sun. According to this view of the atom, the electron follows a single classical path with no restriction on the size or eccentricity of the orbit (which only depend on the initial conditions).

In quantum mechanics, the answer is very different. The electron "sniffs out" many possible orbits and its position smears out into 'quantum orbits':



Moreover, there are only a finite number of allowed quantum orbits, with specific energies. The orbits and their energies are 'quantized'. In this way, quantum mechanics explains the periodic table of elements and therefore underlies all of chemistry. Having agreed that this is important stuff, let's see some of the mathematical details.

2.2 Principles of Quantum Mechanics

Quantum mechanics doesn't use any fancy mathematics. Everything that we will need was described in your course on '*Vectors and Matrices*'. In this section, we will use some of that knowledge to introduce the basic postulates of quantum mechanics. We don't have time to waste, so let's jump right in.

2.2.1 States are Vectors

The *state* of a system consists of all the information that we have to specify about a system in order to determine its future evolution. For example, in classical mechanics we need to list the positions and momenta of all particles. Now, in quantum mechanics,

$$\boxed{\textbf{states } \textit{are} \textbf{ vectors}}\,.$$

We will use $|\Psi\rangle$ for the state vector.[1] As a simple example consider a system that only has two possible states: 0/1, up/down, on/off, left/right, dead/alive, etc. Such a system is also called a

[1] The notation $|\Psi\rangle$ goes back to Dirac. It is called a *ket-vector* or just *ket*. There are also *bra-vectors*, $\langle\Phi|$, which are the transposes (or Hermitian conjugates) of the ket-vectors. The inner product of a bra and a ket, $\langle\Phi|\Psi\rangle$, is a *bra-ket*, or bracket.

bit. The two states of the bit can be represented by the following two-component vectors

$$|\uparrow\rangle \equiv \begin{pmatrix} 1 \\ 0 \end{pmatrix} , \tag{2.2.1}$$

$$|\downarrow\rangle \equiv \begin{pmatrix} 0 \\ 1 \end{pmatrix} . \tag{2.2.2}$$

To have a physical example in mind, you may think of the spin of an electron relative to say the z-axis.[2] In fact, this will be our go-to example for the rest of this lecture.

In the classical world, the system has to be either in one state or the other. However, in the quantum world, the state can be in a superposition of both states at the same time, i.e. we can add the state vectors

$$|\Psi\rangle \;=\; \alpha\,|\uparrow\rangle + \beta\,|\downarrow\rangle = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} , \tag{2.2.3}$$

where α and β are complex numbers satisfying $|\alpha|^2 + |\beta|^2 = 1$. Such a state is called a *qubit* (quantum bit). The possibility of a *superposition of states* is the origin of all the weirdness of quantum mechanics.

More generally, the state vector can be the sum of many basis states

$$|\Psi\rangle \;=\; \sum_i \alpha_i\,|i\rangle , \tag{2.2.4}$$

where each basis state $|i\rangle$ is weighted by a complete number α_i and $\sum_i |\alpha_i|^2 = 1$. The spectrum of states can also be continuous and the corresponding vector space infinite-dimensional. For example, this is the case for a particle with position x. In quantum mechanics, the state of the particle is a superposition of all possible locations

$$|\Psi\rangle \;=\; \int \mathrm{d}x\,\alpha(x)|x\rangle , \tag{2.2.5}$$

and we call the function $\alpha(x)$ the *wavefunction*.

2.2.2 Observables are Matrices

You know that matrices are things that can act on vectors to produce other vectors. You also know that observables are things you can measure. In quantum mechanics,

observables *are (Hermitian)* **matrices** .

We have a different matrix M for each kind of measurement we can make on a system: energy, position, momentum, spin, etc. The matrices corresponding to the spin of an electron along the x-axis, y-axis and $z$axis are

$$\mathsf{X} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} , \quad \mathsf{Y} = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} , \quad \mathsf{Z} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} . \tag{2.2.6}$$

[2]Do *not* try to visualize the electron as a spinning top. Particle spin is a purely quantum phenomenon. In quantum mechanics, the spin can either be up, $|\uparrow\rangle$, or down, $|\downarrow\rangle$. To show that the electron can only have these two spin states is hard.

The matrices "act" on the state vectors simply by matrix multiplication. In general, when a matrix acts on a vector, it will change the direction of the vector. However, there are certain vectors whose direction are the same after the matrix multiplication. These special vectors are called *eigenvectors*. In our notation, this means

$$\mathsf{M}\,|i\rangle = m_i\,|i\rangle\ . \tag{2.2.7}$$

where $|i\rangle$ are the eigenvectors and m_i the corresponding eigenvalues. The vectors (2.2.1) and (2.2.2) are eigenvectors of Z with eigenvalues $+1$ and -1, i.e.

$$\mathsf{Z}\,|\uparrow\rangle = +1\,|\uparrow\rangle \qquad \text{and} \qquad \mathsf{Z}\,|\downarrow\rangle = -1\,|\downarrow\rangle\ . \tag{2.2.8}$$

They are not eigenvectors of X and Y, e.g. $\mathsf{X}\,|\uparrow\rangle = |\downarrow\rangle$.

In quantum mechanics,

measurements *are* **eigenvalues** ,

i.e. the possible outcomes of a measurement are the eigenvalues m_i of the matrix M associated with the observable. Since the measurement matrices are Hermitian, the eigenvalues are always real. You can easily check that the eigenvalues of all matrices in (2.2.6) are $+1$ and -1. Measuring the spin of an electron along any axis will therefore always give $+1$ or -1.

In general, the outcome of a measurement will be probabilistic (see below). An important exception is that

measurements of M *lead to definite values if the states are* **eigenvectors** *of* M.

In other words, if the system is prepared in an eigenvector of M, i.e. $|\Psi\rangle = |i\rangle$, you will always measure the corresponding eigenvalue m_i, with certainty! This is important, so let me say it again: Although we will see that quantum mechanics is all about probabilities and uncertainties, if the system is prepared in an eigenvector of a particular observable and you measure that observable, you will always just get the eigenvalue. For example, if we prepare a system in the state $|\uparrow\rangle$ and decide to measure its spin along the z-axis, we will always get $+1$.

A special matrix H (called the *Hamiltonian*; see Appendix A) represents a measurement of energy. The possible outcomes of the corresponding measurement are the energy eigenvalues E_i, and (2.2.7) becomes the famous *Schrödinger equation*,

$$\mathsf{H}|i\rangle = E_i|i\rangle\ . \tag{2.2.9}$$

For a given system, you need to figure out the matrix H, then solve its eigenvalues. This is usually hard! Sometimes, like for the hydrogen atom it can be done exactly. Most of the time (e.g. for all other atoms) it is done in some approximation scheme. Solving the eigenvalue equation (2.2.9) explains the discrete energy levels of atoms.

2.2.3 Measurements are Probabilistic

We have seen that if the system is in an eigenstate $|i\rangle$ of M, then we always measure the corresponding eigenvalue m_i. However,

if the state is not an eigenvector of the observable M*,*
then the outcomes of measurements of M *will be* **probabilistic***.*

The measurement could give any one of the eigenvalues m_i. Each with some probability. We can expand any arbitrary state $|\Psi\rangle$ in a basis of eigenvectors of M,

$$|\Psi\rangle = \sum_i \alpha_i |i\rangle \ , \tag{2.2.10}$$

where α_i are complex constants. The probability of the measurement giving the eigenvalue m_i is

$$\text{Prob}(m_i) = |\alpha_i|^2 \ , \tag{2.2.11}$$

i.e. the probability of measuring m_i is simply the square of the expansion coefficient α_i. The sum of all probabilities has to add up to 100%, so we have

$$\sum_i \text{Prob}(m_i) = \sum_i |\alpha_i|^2 = 1 \ . \tag{2.2.12}$$

Let's see what this means for the spinning electron. If the system is prepared in the qubit state (2.2.3) then

$$\text{Prob}(\uparrow) = |\alpha|^2 \ , \tag{2.2.13}$$
$$\text{Prob}(\downarrow) = |\beta|^2 \ . \tag{2.2.14}$$

Since we are guaranteed to measure either spin-up or spin-down, we have

$$\text{Prob}(\uparrow) + \text{Prob}(\downarrow) = |\alpha|^2 + |\beta|^2 = 1 \ . \tag{2.2.15}$$

2.2.4 Collapse of the State Vector

Finally,

the state vector **collapses** *after the measurement* :

$$|\Psi\rangle \xrightarrow{m_i} |i\rangle \ ,$$

i.e. if the eigenvalue m_i has been measured, then the state of the system after the measurement is the corresponding eigenvector $|i\rangle$. If we now were to repeat the measurement we would get the same value m_i with certainty. But, if we then perform a measurement of a different observable, corresponding to a matrix N, the outcome will again be probabilistic, unless $|i\rangle$ is also an eigenvector of N.

Again, the spinning electron makes this clear. If the system is prepared in the state (2.2.3), then we measure spin-up or spin-down with probabilities $|\alpha|^2$ and $|\beta|^2$, respectively. Once the measurement is made, the state vector $|\Psi\rangle$ will collapse into one of the two states, $|\uparrow\rangle$ or $|\downarrow\rangle$, depending on which is measured. Every subsequent measurement of spin along the z-axis will give the same value. However, if we then decide to measure a different quantity, say the spin along the x-axis, the result will be probabilistic again. I warned you that this is strange stuff.

Let's say you make a measurement of the spin along the z-axis and find $+1$. After that the state will be

$$|\Psi\rangle = |\!\uparrow\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} . \tag{2.2.16}$$

Check for yourself that the eigenvectors of the matrix X in (2.2.6) are

$$|\!\rightarrow\rangle \equiv \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad , \quad |\!\leftarrow\rangle \equiv \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} . \tag{2.2.17}$$

The state in (2.2.16) can therefore be written as a superposition of the eigenstates of X:

$$\begin{aligned} |\Psi\rangle = |\!\uparrow\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} &= \frac{1}{2} \left[\begin{pmatrix} 1 \\ 1 \end{pmatrix} + \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right] \\ &= \frac{1}{\sqrt{2}} \Big[|\!\rightarrow\rangle + |\!\leftarrow\rangle \Big] . \end{aligned} \tag{2.2.18}$$

Measuring X next will therefore give $+1$ or -1 with equal probability. Say it is -1. The state has then collapsed onto $|\!\leftarrow\rangle$. It is not an eigenstate of Z anymore. Instead, $|\!\leftarrow\rangle = \frac{1}{\sqrt{2}} \big[|\!\uparrow\rangle + |\!\downarrow\rangle \big]$. The next measurement of the spin along the z-axis would again be probabilistic, giving up or down with equal probabilities. And so on ...

2.2.5 The Uncertainty Principle

What we just said has an important implication. Most matrices have different eigenvectors. (This is the case if the matrices don't commute, i.e. $\mathsf{MN} \neq \mathsf{NM}$.) So if you're in a state that is an eigenvector of one matrix, it is unlikely to be an eigenvector of a different matrix. If one type of measurement is certain, another type is uncertain.

For example, you can easily check that the matrices X, Y and Z in (2.2.6) do not commute with each other. You can also check that they have distinct eigenvectors. No matter what the state $|\Psi\rangle$ is, we can therefore never simultaneously know the spins along several different axes. (You may know the spin along the z-axis, if the system is in an eigenstate of Z, but then the spin along x and y is completely unknown.)

The most famous application of the uncertainty principle is to the position and momentum of a particle. In classical mechanics, we can know both with certainty. In fact, we need to know them in order to predict the future evolution of the particle. However, in quantum mechanics, if we know the position x of a particle, its momentum p becomes completely uncertain. And vice versa. This is quantified in *Heisenberg's uncertainty relation*

$$\Delta x \Delta p \geq \frac{\hbar}{2} , \tag{2.2.19}$$

where $\hbar \approx 10^{-34}\,\mathrm{J \cdot s}$ is Planck's constant. The reason we don't notice these uncertainties in everyday life is because $\hbar$ is so tiny. Similar uncertainty relations exist for other observables.

2.2.6 Combining Systems: Entanglement

Things get interesting in quantum mechanics when we combine systems. For concreteness, we will consider states made from a combination of two or more qubits. For example, we could have two electrons each with spin-up or spin-down. We will label the electrons A and B. If A is in the state up, $|\uparrow\rangle_{\rm A}$, and B is in the state down, $|\downarrow\rangle_{\rm B}$, then the combined state is

$$|\uparrow\rangle_{\rm A} \otimes |\downarrow\rangle_{\rm B} \;\equiv\; |\uparrow\downarrow\rangle \ . \tag{2.2.20}$$

The left-hand side is called a tensor product. The simplified notation on the right-hand side should be clear: the first arrow denotes the spin of A, the second arrow that of B. In total, there are four possible combined states:

$$|\uparrow\uparrow\rangle \quad , \quad |\uparrow\downarrow\rangle \quad , \quad |\downarrow\uparrow\rangle \quad , \quad |\downarrow\downarrow\rangle \quad . \tag{2.2.21}$$

The state vector $|\Psi\rangle$ may be a superposition of these four states. For example, imagine that the system is prepared in the state

$$|\Psi\rangle = \tfrac{1}{\sqrt{2}}|\uparrow\downarrow\rangle + \tfrac{1}{\sqrt{2}}|\downarrow\uparrow\rangle \ . \tag{2.2.22}$$

This is called an *entangled state* since it cannot be separated into the product of states for the individual electrons. Not all superpositions of the states (2.2.21) are entangled states. An example of a *product state*[3] is

$$|\Psi\rangle = \tfrac{1}{\sqrt{2}}|\uparrow\downarrow\rangle + \tfrac{1}{\sqrt{2}}|\downarrow\downarrow\rangle \ , \tag{2.2.25}$$

which can be written as

$$|\Psi\rangle = \Big(\tfrac{1}{\sqrt{2}}|\uparrow\rangle_{\rm A} + \tfrac{1}{\sqrt{2}}|\downarrow\rangle_{\rm A}\Big) \otimes |\downarrow\rangle_{\rm B} \ . \tag{2.2.26}$$

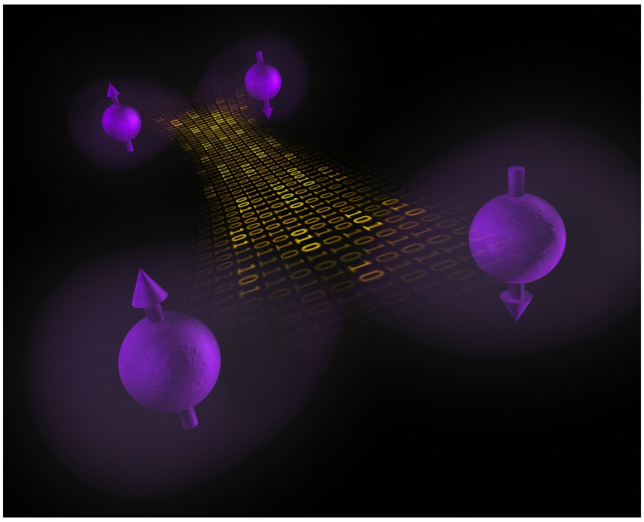
The main feature of a product state is that each subsystem behaves independently of the other. If we do an experiment on B, the result is exactly the same as if A did not exist. The same is true for A of course. In contrast, in an entangled state measurements of A and B are not independent.

The state (2.2.22) can arise if a particle without spin decays into two electrons. Since angular momentum is conserved, the spins of the two electrons have to be anti-aligned. While in classical physics the system then has to be in the state $|\uparrow\downarrow\rangle$ *or* $|\downarrow\uparrow\rangle$, in quantum physics it can be in the state $|\uparrow\downarrow\rangle$ *and* $|\downarrow\uparrow\rangle$. The two electrons are then separated by a large distance (as large as you wish). Say one of the electrons stays here, while the other is brought to the end of the universe. We then measure the spin of the electron that stayed here. There is a 50% chance that the result will be spin-up, and a 50% chance that it will be spin-down. However, once the spin of the electron is measured, the spin of the far-away electron is determined *instantly*. By itself this *correlation* of the measurements is not a problem. The same happens in classical physics. Imagine a pair of gloves. Each glove is put into a box and the two boxes are then separated by a large distance. If you open one of the boxes and find a left-hand glove, you know immediately that the far-way glove is a right-hand glove. No problem. However, entangled states like (2.2.22) are *superpositions* of correlated states. As we will see in the next section, such states can have correlations that are purely quantum. Quantum gloves are actually left- and right-handed (and everything in between) up until the moment you observe them. Moreover, the left-handed glove

[3] In general, the set of product states takes the form

$$\begin{aligned} |\Psi\rangle \;&=\; \alpha_\uparrow\beta_\uparrow|\uparrow\uparrow\rangle + \alpha_\uparrow\beta_\downarrow|\uparrow\downarrow\rangle + \alpha_\downarrow\beta_\uparrow|\downarrow\uparrow\rangle + \alpha_\downarrow\beta_\downarrow|\downarrow\downarrow\rangle \ , &(2.2.23)\\ &=\; \big(\alpha_\uparrow|\uparrow\rangle_{\rm A} + \alpha_\downarrow|\downarrow\rangle_{\rm A}\big) \otimes \big(\beta_\uparrow|\uparrow\rangle_{\rm B} + \beta_\downarrow|\downarrow\rangle_{\rm B}\big) \ . &(2.2.24)\end{aligned}$$

doesnt become left until you observe the right-handed one — at which moment both instantly gain a definite handedness. This really upset Einstein. He called it "spooky action at a distance".



2.3 Quantum Mechanics in Your Face

Entanglement is arguably the most bewildering aspect of quantum mechanics. Let's explore this a bit further.[4] The punchline will be that the universe is a much stranger place than you might have imagined.

2.3.1 The GHZ Experiment

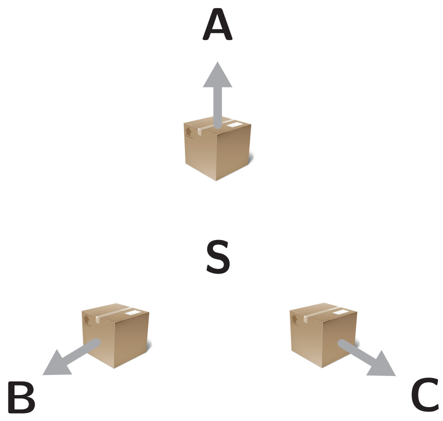
Consider three scientists:

It is conventional in the quantum mechanics literature to call them Alice (A), Bob (B) and Charlie (C).[5] They are sent to labs at three different locations. Every minute, they receive a

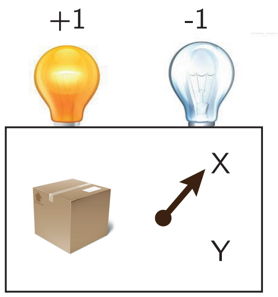
[4]This section is based on a famous lecture by Sidney Coleman (http://www.physics.harvard.edu/about/video.html), which itself was based on Greenberger, Horne, and Zeilinger (GHZ) (1990) "Bell's theorem without inequalities", Am. J. Phys. 58 (12): 1131, and Mermin (1990) "Quantum mysteries revisited" Am. J. Phys. 58 (8): 731734.

[5]Admittedly, A doesn't look much like an Alice.

package from a mysterious central station (S):



Each scientist has a machine that performs measurements of the packages. The machine has two settings, X or Y, and each measurement can give two outcomes, +1 and −1.



Alice, Bob and Charlie are told what they have to do:[6]

1. Choose the setting X or Y on the machine.
2. Insert the package into the machine.
3. Make a measurement.
4. Record whether the result is +1 or −1.
5. Go back to Step 1.

Each measurement is recorded until each scientist has a list that looks like this:

X	X	Y	X	Y	Y	X	Y	X	X	X	Y	⋯
+1	−1	+1	−1	−1	+1	+1	+1	−1	−1	+1	−1	⋯

After they each have made a bazillion measurements, Alice, Bob and Charlie get together and start looking for *correlations* in their measurements. (Since their packages came from the same source, it isn't unreasonable to expect some correlations.) They notice the following: Whenever one of them measured X, and the other two measured Y, the results *always* multiply to +1, i.e.[7]

$$\mathsf{X}_\mathrm{A}\mathsf{Y}_\mathrm{B}\mathsf{Y}_\mathrm{C} = \mathsf{Y}_\mathrm{A}\mathsf{X}_\mathrm{B}\mathsf{Y}_\mathrm{C} = \mathsf{Y}_\mathrm{A}\mathsf{Y}_\mathrm{B}\mathsf{X}_\mathrm{C} = +1\ . \tag{2.3.27}$$

[6]They are not told what's in the packages: They could be blood samples, with the machine testing for high/low glucose when the switch is on X, and high/low cholesterol when the switch is on Y. They could be elementary particles with the machine measuring the spin along x or y. Or, the whole thing could just be a hoax with the machine flashing up $+1/-1$ at random.

[7]Here, the notation $\mathsf{X}_\mathrm{A}\mathsf{Y}_\mathrm{B}\mathsf{Y}_\mathrm{C}$ means that Alice measured X, while Bob and Charlie measured Y.

Maybe this occurred because all three got the result +1; or perhaps one got +1 and the other two got −1. Since the central station doesn't know ahead of time which setting (X or Y) each scientist will choose, it has to prepare each package with definite states for both property X and property Y. The observed correlation in (2.3.27) is consistent with only the following 8 different shipments from the central station:

$$= \begin{bmatrix} \mathsf{X}_\mathrm{A} & \mathsf{Y}_\mathrm{A} \\ \mathsf{X}_\mathrm{B} & \mathsf{Y}_\mathrm{B} \\ \mathsf{X}_\mathrm{C} & \mathsf{Y}_\mathrm{C} \end{bmatrix} = \begin{bmatrix} + & + \\ + & + \\ + & + \end{bmatrix} \begin{bmatrix} - & - \\ - & - \\ + & + \end{bmatrix} \begin{bmatrix} - & + \\ - & + \\ + & - \end{bmatrix} \begin{bmatrix} - & - \\ + & + \\ - & - \end{bmatrix}$$

$$\begin{bmatrix} - & + \\ + & - \\ - & + \end{bmatrix} \begin{bmatrix} + & - \\ + & - \\ + & - \end{bmatrix} \begin{bmatrix} + & + \\ - & - \\ - & - \end{bmatrix} \begin{bmatrix} + & - \\ - & + \\ - & + \end{bmatrix} \tag{2.3.28}$$

Now, notice that (2.3.27) gives a *prediction* ... if all three scientists measure X, the results multiplied together *must* give +1. You can see this simply by multiplying the entries of the first columns in the matrices in (2.3.28). Alternative, we can prove the result using nothing but simple arithmetic:

$$\begin{aligned} (\mathsf{X}_\mathrm{A}\mathsf{Y}_\mathrm{B}\mathsf{Y}_\mathrm{C})(\mathsf{Y}_\mathrm{A}\mathsf{X}_\mathrm{B}\mathsf{Y}_\mathrm{C})(\mathsf{Y}_\mathrm{A}\mathsf{Y}_\mathrm{B}\mathsf{X}_\mathrm{C}) &= \mathsf{X}_\mathrm{A}\mathsf{X}_\mathrm{B}\mathsf{X}_\mathrm{C}(\mathsf{Y}_\mathrm{A}\mathsf{Y}_\mathrm{B}\mathsf{Y}_\mathrm{C})^2 \\ &= \mathsf{X}_\mathrm{A}\mathsf{X}_\mathrm{B}\mathsf{X}_\mathrm{C} \\ &= +1 \ . \end{aligned} \tag{2.3.29}$$

In the first equality we used that the product is associative, while the second equality follows from $(\pm 1)^2 = +1$. The final equality is a consequence of (2.3.27).

2.3.2 Completely Nuts, But True!

The GHZ experiment has been done[8]. The things measured were the spins of electrons. Here is the astonishing truth: The experimenter observed same correlations as in (2.3.27):

$$\mathsf{X}_\mathrm{A}\mathsf{Y}_\mathrm{B}\mathsf{Y}_\mathrm{C} = \mathsf{Y}_\mathrm{A}\mathsf{X}_\mathrm{B}\mathsf{Y}_\mathrm{C} = \mathsf{Y}_\mathrm{A}\mathsf{Y}_\mathrm{B}\mathsf{X}_\mathrm{C} = +1 \ . \tag{2.3.30}$$

However, instead of (2.3.29), they found

$$\mathsf{X}_\mathrm{A}\mathsf{X}_\mathrm{B}\mathsf{X}_\mathrm{C} = -1 \ . \tag{2.3.31}$$

In classical physics, this shouldn't happen, since (2.3.29) was a logical consequence of (2.3.27). Something about our basic (classical) intuition for how the universe works is *wrong*!

2.3.3 Quantum Reality

We assumed that the packages leaving the central station had definite assignments for the quantities X and Y—we listed all possibilities in (2.3.28). But in the quantum world, we cannot give

[8] Pan et al. (2000), "Experimental test of quantum non-locality in three-photon GHZ entanglement" Nature 403 (6769) 515 - 519.

definite assignments to all possible measurements. Instead we have to allow for the possibility of a superposition of states and experimental outcomes being probabilistic. It turns out that this special feature of quantum mechanics resolves the puzzle.

More concretely, Alice, Bob and Charlie were, in fact, measuring the spins of electrons along the x- and y-axes. In this case, the measurement matrices are

$$\mathsf{X} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad , \quad \mathsf{Y} = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \ . \tag{2.3.32}$$

You can check that these matrices both have eigenvalues $+1$ and -1, corresponding to the measurements. (But, X and Y do not have the same eigenvectors.) As before, we define two special state vectors corresponding to an electron spinning up or down relative to the z-axis,

$$|{\uparrow}\rangle \equiv \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad , \quad |{\downarrow}\rangle \equiv \begin{pmatrix} 0 \\ 1 \end{pmatrix} \ . \tag{2.3.33}$$

These states are not eigenstates of X and Y. It is easy to see that acting with the matrix X on the up-state turns it into a down-state and vice versa:

$$\mathsf{X}\,|{\uparrow}\rangle = |{\downarrow}\rangle \tag{2.3.34}$$

$$\mathsf{X}\,|{\downarrow}\rangle = |{\uparrow}\rangle \ . \tag{2.3.35}$$

Similarly, acting with the matrix Y also exchanges up- and down-states (up to factors of i and $-i$)

$$\mathsf{Y}\,|{\uparrow}\rangle = i\,|{\downarrow}\rangle \tag{2.3.36}$$

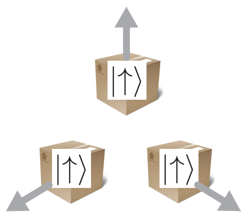
$$\mathsf{Y}\,|{\downarrow}\rangle = -i\,|{\uparrow}\rangle \ . \tag{2.3.37}$$

Now, you are told that the central station sent out the following entangled state

$$|\Psi\rangle = \tfrac{1}{\sqrt{2}}\,|{\uparrow\uparrow\uparrow}\rangle - \tfrac{1}{\sqrt{2}}\,|{\downarrow\downarrow\downarrow}\rangle \ . \tag{2.3.38}$$

This corresponds to the superposition of two states: a state with all spins up

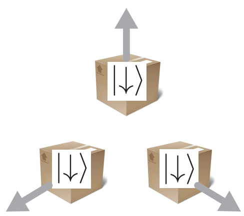
$|{\uparrow\uparrow\uparrow}\rangle \ \equiv$



,

and a state with all spins down

$|{\downarrow\downarrow\downarrow}\rangle \ \equiv$



.

As before, I am using a notation where the arrows are ordered: the first arrow corresponds to the spin of the first particle (the one sent to Alice), the second arrow corresponds to the spin of the second particle (the one sent to Bob), etc. The measurement matrix $\mathsf{X}_{\rm A}$ therefore acts on the first arrow of each state, $\mathsf{X}_{\rm B}$ on the second arrow, etc.

Just to be clear, let me give an excessively explicit example

$$\mathsf{X}_{\rm A}\mathsf{Y}_{\rm B}\mathsf{Y}_{\rm C}|\uparrow\uparrow\uparrow\rangle = \mathsf{X}_{\rm A}|\uparrow\rangle \otimes \mathsf{Y}_{\rm B}|\uparrow\rangle \otimes \mathsf{Y}_{\rm C}|\uparrow\rangle \ . \tag{2.3.39}$$

Using (2.3.34) and (2.3.36), we find

$$\begin{aligned}\mathsf{X}_{\rm A}\mathsf{Y}_{\rm B}\mathsf{Y}_{\rm C}|\uparrow\uparrow\uparrow\rangle &= (1)|\downarrow\rangle \otimes (i)|\downarrow\rangle \otimes (i)|\downarrow\rangle \\ &= (1)(i)(i)|\downarrow\downarrow\downarrow\rangle \\ &= -|\downarrow\downarrow\downarrow\rangle \ . \end{aligned} \tag{2.3.40}$$

The state in (2.3.38) is an eigenvector of $\mathsf{X}_{\rm A}\mathsf{Y}_{\rm B}\mathsf{Y}_{\rm C}$ and $\mathsf{Y}_{\rm A}\mathsf{X}_{\rm B}\mathsf{Y}_{\rm C}$ and $\mathsf{Y}_{\rm A}\mathsf{Y}_{\rm B}\mathsf{X}_{\rm C}$. And, importantly, it is also an eigenvector of $\mathsf{X}_{\rm A}\mathsf{X}_{\rm B}\mathsf{X}_{\rm C}$. Let us check that this gives rise to the observed correlations: For instance,

$$\begin{aligned}\mathsf{X}_{\rm A}\mathsf{Y}_{\rm B}\mathsf{Y}_{\rm C}|\Psi\rangle &= \mathsf{X}_{\rm A}\mathsf{Y}_{\rm B}\mathsf{Y}_{\rm C}\Big[\,|\uparrow\uparrow\uparrow\rangle - |\downarrow\downarrow\downarrow\rangle\,\Big] \\ &= (1)(i)(i)|\downarrow\downarrow\downarrow\rangle - (1)(-i)(-i)|\uparrow\uparrow\uparrow\rangle \\ &= -|\downarrow\downarrow\downarrow\rangle + |\uparrow\uparrow\uparrow\rangle \\ &= +1\,|\Psi\rangle \ . \end{aligned} \tag{2.3.41}$$

Similarly, we can show that

$$\mathsf{Y}_{\rm A}\mathsf{X}_{\rm B}\mathsf{Y}_{\rm C}|\Psi\rangle = \mathsf{Y}_{\rm A}\mathsf{Y}_{\rm B}\mathsf{X}_{\rm C}|\Psi\rangle = +1|\Psi\rangle \ . \tag{2.3.42}$$

So, whenever exactly one scientist measures X, the results indeed multiply to give $+1$. However, when all three scientists measure X, we get

$$\begin{aligned}\mathsf{X}_{\rm A}\mathsf{X}_{\rm B}\mathsf{X}_{\rm C}|\Psi\rangle &= \mathsf{X}_{\rm A}\mathsf{X}_{\rm B}\mathsf{X}_{\rm C}\Big[\,|\uparrow\uparrow\uparrow\rangle - |\downarrow\downarrow\downarrow\rangle\,\Big] \\ &= (1)(1)(1)|\downarrow\downarrow\downarrow\rangle - (1)(1)(1)|\uparrow\uparrow\uparrow\rangle \\ &= |\downarrow\downarrow\downarrow\rangle - |\uparrow\uparrow\uparrow\rangle \\ &= -1\,|\Psi\rangle \ . \end{aligned} \tag{2.3.43}$$

The classical expectation, eq. (2.3.29), is wrong. But, it makes total sense in the quantum world.

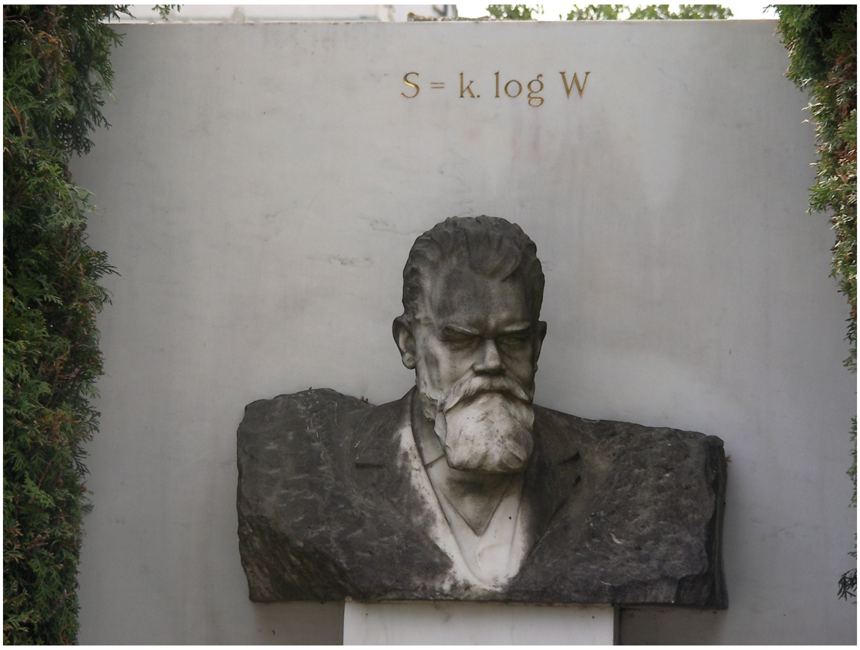
It was important that the spin states of the three particles weren't independent, but that they were "entangled" in the state (2.3.38)—a superposition of all spins up $|\uparrow\uparrow\uparrow\rangle$ and all spins down $|\downarrow\downarrow\downarrow\rangle$. No matter how far the scientists are separated in space, this entanglement of spin states is reflected in their measurements. It is how the world works.

3 Statistical Mechanics

"Ludwig Boltzmann, who spent much of his life studying statistical mechanics, died in 1906 by his own hand. Paul Ehrenfest, carrying on the work, died similarly in 1933. Now it is your turn to study statistical mechanics."

David Goodstein

Today, I want to tell you about **entropy** and the **Second Law** of thermodynamics. This will address deep questions about the world we live in, such as why we remember the past and not the future. It will also have relations to information theory, computing and even the physics of black holes. Along the way, we will encounter one of the most important formulas in all of science:



3.1 More is Different

Suppose you've got theoretical physics cracked—i.e. you know all the fundamental laws of Nature, the properties of the elementary particles and the forces at play between them. How can you turn this knowledge into an understanding of the world around us?

Consider this glass of water:

It contains about $N = 10^{24}$ atoms. In fact, any macroscopic object contains such a stupendously large number of particles. How do we describe such systems?

An approach that certainly won't work, is to write down the equations of motion for all 10^{24} particles and solve them. Even if we could handle such computations, what would we do with the result? The positions of individual particles are of little interest to anyone. We want answers to much more basic questions about macroscopic objects. Is it wet? Is it cold? What colour is it? What happens if we heat it up? How can we answer these kind of questions starting from the fundamental laws of physics?

Statistical mechanics is the art of turning the microscopic laws of physics into a description of the macroscopic everyday world. Interesting things happen when you throw 10^{24} particles together. More is different: there are key concepts that are not visible in the underlying laws of physics but *emerge* only when we consider a large collection of particles. A simple example is *temperature*. This is clearly not a fundamental concept: it doesn't make sense to talk about the temperature of a single electron. But it would be impossible to talk about the world around us without mention of temperature. Another example is *time*. What distinguishes the past from the future? We will start there.

3.2 The Distinction of Past and Future

Nature is full of irreversible phenomena: things that easily happen but could not possibly happen in reverse order. You drop a cup and it breaks. But you can wait a long time for the pieces to come back together spontaneously. Similarly, if you watch the waves breaking at the sea, you aren't likely to witness the great moment when the foam collects together, rises up out of the sea and falls back further out from the shore. Finally, if you watch a movie of an explosion in reverse, you know very well that it's fake. As a rule, things go one way and not the other. We remember the past, we don't remember the future.

Where does this irreversibility and the arrow of time come from? Is it built into the microscopic laws of physics? Do the microscopic laws distinguish the past from the future? Things are not so simple. The fundamental laws of physics are, in fact, completely reversible.[1] Let us take

[1] This isn't quite true for processes involving the weak nuclear force, but this isn't relevant for the present discussion.

the law of gravity as an example. Take a movie of the planet going around a star. Now run the movie in reverse. Does it look strange? Not at all. Any solution of Newton's equations can be run backward and it is still a solution. Whether the planet goes around the star one way or the opposite way is just a matter of its initial velocity. The law of gravitation is time reversible. Similarly, the laws of electricity and magnetism are time reversible. And so are all other fundamental laws that are relevant for creating our everyday experiences.

So what distinguishes the past from the future? How do reversible microscopic laws give rise to apparent irreversible macroscopic behaviour? To understand this, we must introduce the concept of **entropy**.

3.3 Entropy and The Second Law

> *"If someone points out to you that your pet theory of the universe is in disagreement with Maxwell's equations—then so much the worse for Maxwell's equations. If it is found to be contradicted by observation—well these experimentalists can bungle things sometimes. But if your theory is found to be against the Second Law of Thermodynamics I can give you no hope; there is nothing for it but to collapse to deepest humiliation."*
>
> Sir Arthur Eddington

3.3.1 Things Always Get Worse

We start with a vague definition of entropy as the amount of *disorder* of a system. Roughly, we mean by "order" a state of purposeful arrangement, while "disorder" is a state of randomness. For example, consider dropping ice cubes in a glass of water. This creates a highly ordered, or low entropy, configuration: ice in one corner, water in the rest. Left on its own, the ice will melt and the ice molecules will mix with the water.

The final mixed state is less ordered, or high entropy.

Similarly, the natural tendency of coffee and milk is to mix, but not to unmix

These basic facts of life are summarized in the Second Law of Thermodynamics:[2]

The entropy of an isolated system always increases.

To the physicists of the late 19$^{\text{th}}$ century the Second Law was a serious paradox. They knew that the microscopic laws of physics are time reversible. So if entropy can increase, the laws of physics say it must be able to decrease. Yet, experience says otherwise. Entropy always increases.

3.3.2 Probability

This is where Ludwig Boltzmann's genius came in. He realized is that the Second Law is not a law in the same sense as Newton's law of gravity or Faraday's law of induction. It's a *probabilistic* law that has the same status as the following obvious claim: if you flip a coin a million times you will not get a million heads. It simply won't happen. But is it possible? Yes, it is—it violates no law of physics. Is it likely? Not at all. Boltzmann's formulation of the Second Law is just that. Instead of saying entropy does not decrease, he said that

entropy probably doesn't decrease.

This is where the difference between 1 (or a few) and 10^{24} is important again. It is much more likely for a handful of particles to spontaneously do crazy things than for 10^{24} particles. The Second Law *emerges* for a large number of particles.

This also implies that if you wait around long enough, you will eventually see entropy decrease:

- By accident, particles and dust will come together and form a perfectly assembled bomb. But, how long does it take for that to happen? A very long time. A lot longer than the time to flip a million heads in a row, and even a lot longer than the age of the universe.
- Imagine I drop a bit of black ink into a glass of water. The ink spreads out and eventually makes the water grey. Will a glass of grey water ever clear up and produce a small drop of ink? Not impossible, but very unlikely.
- The air in this room is uniformly distributed. Is it possible that all air molecules spontaneously collect in one corner of the room, leaving the rest a vacuum? Not impossible, but very unlikely.

[2] The First Law is the conservation of energy.

3.3.3 Counting

Let us now get a bit more precise and less philosophical. First, we want to give definitions of three related concepts: *microstates*, *macrostates* and *statistical entropy.*

As a concrete example, consider a collection of N particles in a box. If a particle is in the left half of the box, we say that it is in state L. If it is in the right half, we call its state R. We specify a *microstate* of the system by making a list of the states of each particle, whether its left (L) or right (R). For instance, for $N = 10$ particles, a few possible microstates are

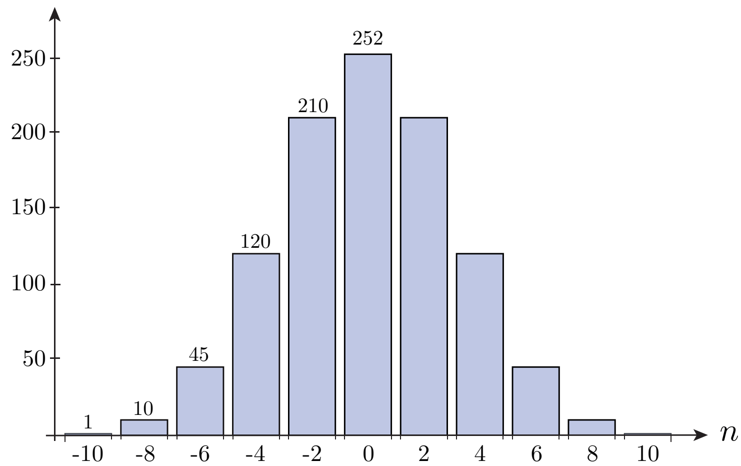
$$\begin{array}{c} L\,L\,L\,L\,L\,L\,L\,L\,L\,L \\ R\,L\,L\,L\,L\,L\,L\,L\,L\,L \\ L\,R\,L\,L\,R\,R\,L\,L\,L\,L \\ \cdots \end{array}$$

The total number of possible microstates is 2^N (two possibilities for each particle). For $N = 10^{24}$ particles, this is a ridiculously large number, $2^{10^{24}}$. Luckily, we never need a list of all possible microstates. All macroscopic properties only depend on the relative number of left and right particles and not on the detail of *which* are left and right.

We can collect all microstates with the same numbers of L and R into a single *macrostate*, labelled by one number

$$n \equiv N_L - N_R \ . \tag{3.3.1}$$

How many microstates are in a given macrostate? Look at $N = 10$. For $n = 10$ (all left) and $n = -10$ (all right), we only have 1 unique microstate each. For $n = 8$ (one right), we get 10 possible microstates since for 10 particles there are 10 ways of putting one particle on the right. For $n = 0$ (equal number on the left and the right), we get 252 microstates. The complete distribution of the number of microstates per macrostate is summarized in the following figure:



It is easy to generalize this: let $W(n)$ be the number of ways to have N particles with N_L particles on the left and N_R particles on the right. The answer is

$$W(n) = \frac{N!}{N_L! N_R!} = \frac{N!}{(\frac{N-n}{2})!(\frac{N+n}{2})!} \ . \tag{3.3.2}$$

For very large N, your calculator won't like evaluating the factorials. At this point, a normal distribution is a very good approximation

$$W(n) \approx 2^N e^{-n^2/2N} . \qquad (3.3.3)$$

So far this was just elementary combinatorics. To this we now add the fundamental assumption of statistical physics:

each (accessible) microstate is equally likely.

Macrostates consisting of more microstates are therefore more likely.

Boltzmann then defined the *entropy* of a certain macrostate as the *logarithm*[3] of the number of microstates

$$\boxed{S = k \log W} , \qquad (3.3.5)$$

where $k = 1.38 \times 10^{-23} JK^{-1}$ is Boltzmann's constant. The role of Boltzmann's constant is simply to get the units right. Eq. (3.3.5) is without a doubt one of the most important equations in all of science, on a par with Newton's $F = ma$ and Einstein's $E = mc^2$. It provides the link between the microscopic world of atoms (W) and the macroscopic world we observe (S). In other words, it is precisely what we were looking for.

Given the fundamental assumption of statistical mechanics (that each accessible microstate is equally likely), we expect systems to naturally evolve towards macrostates corresponding to a larger number of microstates and hence larger entropy

$$\frac{dS}{dt} \geq 0 . \qquad (3.3.6)$$

We are getting closer to a microscopic understanding of the Second Law.

3.3.4 Arrow of Time

With this more highbrow perspective, we now return to the question how macroscopic features of a system made of many particles evolve as a consequence of the motion of the individual particles.

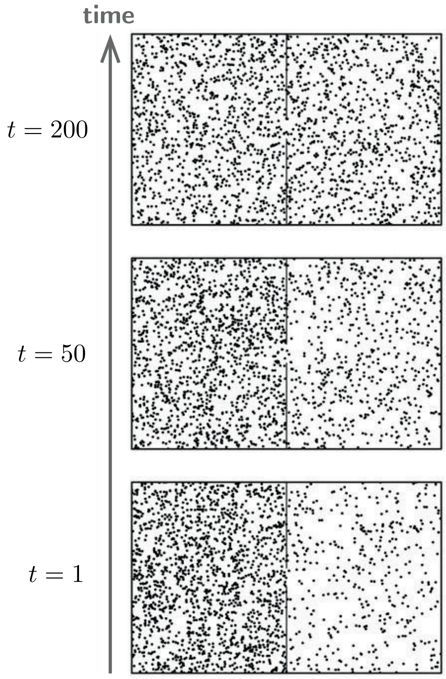
Let our box be divided in two by a wall with a hole in it. Gas molecules can bounce around on one side of the box and will usually bounce right off the central wall, but every once in a while they will sneak through to the other side. We might imagine, for example, that the molecules bounce off the central wall 995 times out of 1,000, but 5 times they find the hole and move to the other side. So, every second, each molecule on the left side of the box has a 99.5 percent chance of staying on that side, and a 0.5 percent chance of moving to the other side—likewise for the molecules on the right side of the box. This rule is perfectly time-reversal invariant—if you made a movie of the motion of just one particle obeying this rule, you couldn't tell whether

[3]Taking the logarithm of W to define entropy has the following important consequences: i) It makes the stupendously large numbers, like $W = 2^{10^{24}}$, less stupendously large; ii) More importantly, it makes entropy additive—i.e. if we combine two systems 1 and 2, the number of microstates multiply, $W_{\rm tot} = W_1 W_2$, which means that the entropies add

$$S_{\rm tot} = k \log W_{\rm tot} = k \log(W_1 W_2) = k \log W_1 + k \log W_2 = S_1 + S_2 . \qquad (3.3.4)$$

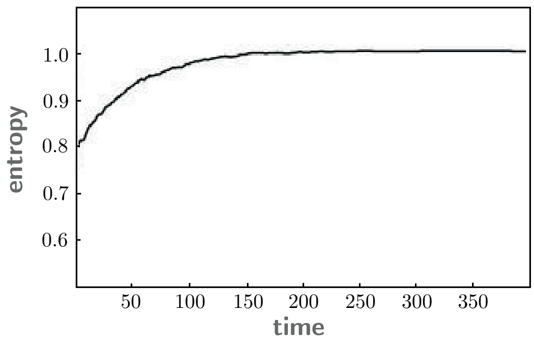
it was being run forward or backward in time. At the level of individual particles, we can't distinguish the past from the future.

However, let's look at the evolution from a more macroscopic perspective:



The box has $N = 2{,}000$ molecules in it, and starts at time $t = 1$ with 1,600 molecules on the left-hand side and only 400 on the right. It's not very surprising what happens: because there are more molecules on the left, the total number of molecules that shift from left to right will usually be larger than the number that shift from right to left. So after 50 seconds we see that the numbers are beginning to equal out, and after 200 seconds the distribution is essentially equal. This box clearly displays an arrow of time. Even if we hadn't labelled the different distributions in the figure with the specific times to which they correspond, you wouldn't have any trouble guessing that the bottom box came first and the top box came last. We're not surprised when the air molecules even themselves out, but we would be very surprised if they spontaneously congregated all on one side of the box. The past is the direction of time in which things were more segregated, while the future is the direction in which they have smoothed themselves out. It's exactly the same thing that happens when a ice cube melts or milk spreads out into a cup of coffee.

It is easy to see that this is consistent with the Second Law. Using eqs. (3.3.2) and (3.3.5), we can associate an entropy with the system at every moment in time. A plot of the evolution looks as follows



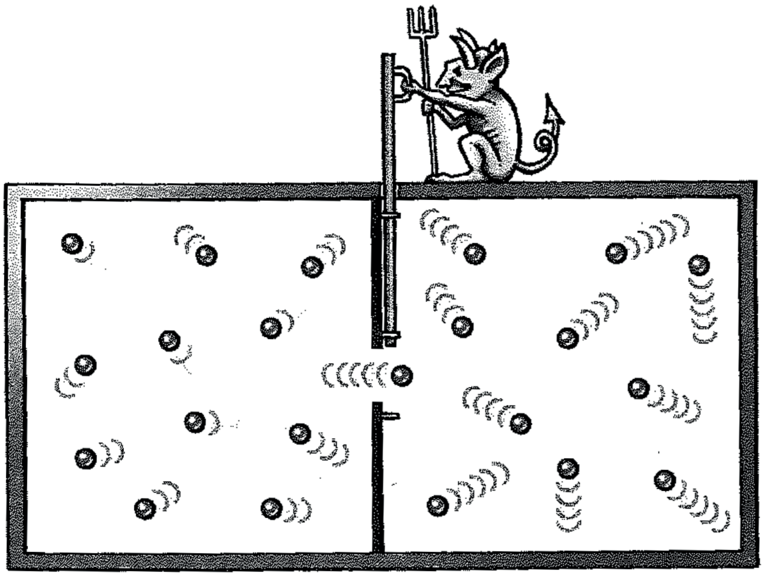
3.4 Entropy and Information

"You should call it entropy, for two reasons. In the first place, your uncertainty function has been used in statistical mechanics under that name, so it already has a name. In the second place, and more important, no one knows what entropy really is, so in a debate you will always have the advantage."

John von Neumann to Claude Shannon.

3.4.1 Maxwell's Demon

In 1871, Maxwell introduced a famous thought experiment that challenged the Second Law.



The setup is the same as before: a box of gas divided in two by a wall with a hole. However, this time the hole comes with a tiny door that can be opened and closed without exerting a noticeable amount of energy. Each side of the box contains an equal number of molecules with the same average speed (i.e. same temperature). We can divide the molecules into two classes: those that move faster than the average speed—let's call them *red* molecules—and those that move slower than average—called *blue*. At the beginning the gas is perfectly mixed (equal numbers of red and blue on both sides, i.e. maximum entropy). At the door sits a demon, who watches the molecules coming from the left. Whenever he sees a *red* (fast-moving) molecule approaching the hole, he opens the door. When the molecule is *blue*, he keeps the door shut. In this way the demon 'unmixes' the red and blue molecules, the left side of the box gets colder and the right side hotter. We could use this temperature difference to drive an engine without putting any energy in: a perpetual motion machine. Clearly this looks like it violates the Second Law. What's going on?

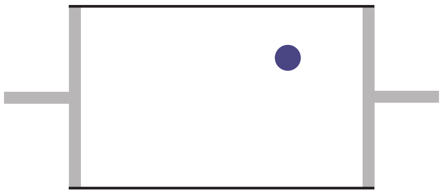
Maxwell's demon and its threat to the Second Law have been debated for more than a century. To save the Second Law there has to be a compensating increase in entropy somewhere. There is only one place the entropy could go: into the demon. So does the demon generate entropy in carrying out his demonic task? The answer is yes, but the way that this work is quite subtle and was understood only recently (by Szilard, Landauer and Bennett).[4] The resolution relies on a fascinating connection between statistical mechanics and information theory.[5]

3.4.2 Szilard's Engine

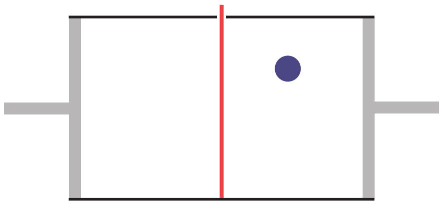
In 1929, Leo Szilard launched the demon into the information age.[6] In particular, he showed that

information is physical.

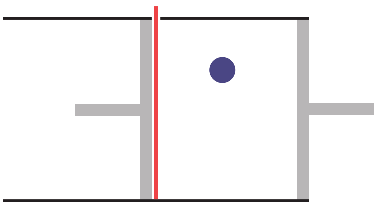
Possessing information allows us to extract useful work from a system in ways that would have otherwise been impossible. Szilard arrived at these insights through a clever new version of Maxwell's demon: this time there is only a single molecule in the box. Two walls of the box are replaced by movable pistons



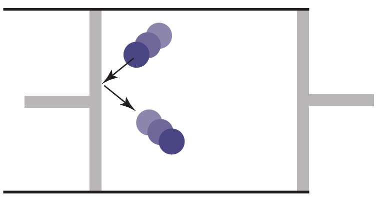
A partition (now without a hole) is placed in the middle. The molecule is on one side and the other side is empty



The demon measures and records on what side of the partition the gas molecule is, gaining one *bit* of information. He then pushes in the piston that closed off the empty half of the box



In the absence of friction, this process doesn't require any energy. Note the crucial role played by information in this setup. If the demon didn't know which half of the box the molecule was in, he wouldn't know which piston to push in. After removing the partition, the molecule will push against the piston and the one-molecule gas "expands"

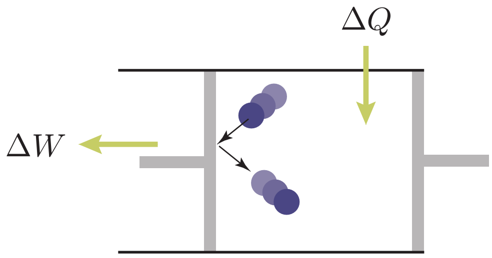


[4]One of the giants in this story wrote a nice article about it: Bennett, "Demons, Engines and the Second Law", Scientific American 257 (5): 108-116 (1987).

[5]William Bialek recently gave a nice lecture of the relationship between entropy and information. The video can be found here: http://media.scgp.stonybrook.edu/video/video.php?f=20120419_1_qtp.mp4

[6]Szilard, "On the Decrease of Entropy in a Thermodynamic System by the Intervention of Intelligent Beings."

In this way we can use the system to do useful work (e.g. by driving an engine). Where did the energy come from? From the heat Q of the surroundings (with temperature T),



The work done when the gas expands from $V_i = V$ to $V_f = 2V$ is given by a standard formula in thermodynamics:

$$\Delta W = kT \log\left(\frac{V_f}{V_i}\right) = kT \log 2 \ . \tag{3.4.7}$$

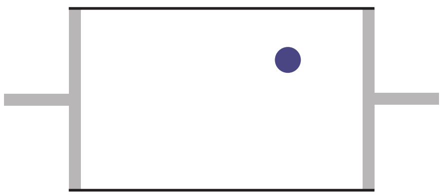
Recall that $\mathrm{d}W = F\mathrm{d}x = p\mathrm{d}V$, where p is the pressure of the gas. The integrated work done is therefore

$$\Delta W = \int_{V_i}^{V_f} p\,\mathrm{d}V \ .$$

Using the ideal gas law for the one-molecule gas, $pV = kT$, we can write this as

$$\Delta W = \int_{V_i}^{V_f} \frac{kT}{V}\,\mathrm{d}V = kT \log\left(\frac{V_f}{V_i}\right) \ .$$

The system returns back to its initial state



This completes one cycle of operation. The whole process is repeatable. Each cycle would allow extraction and conversion of heat from the surroundings into useful work in a cyclic process. The demon seems to have created a perpetual motion machine of the second kind.[7] In particular, in each stage of the cycle the entropy decreases by $\Delta S = \Delta Q/T$ (another classic formula of thermodynamics). Using $\Delta Q = -\Delta W$, we find

$$\Delta S = -k \log 2 \ . \tag{3.4.8}$$

Szilard's demon again seems to have violated the Second Law.

3.4.3 Saving the Second Law

In 1982, Charles Bennett observed that Szilard's engine is not quite a closed cycle.[8] While after each cycle the box has returned to its initial state, the mind of the demon has not! He has

[7]It is of the 'second kind' because it violates the 'second' law of thermodynamics. A perpetual motion machine of the first kind violates the 'first' law—the conservation of energy.

[8]Bennett, "The Thermodynamics of Computation".

gained one bit of recorded information. The demon needs to erase the information stored in his mind in order for the process to be truly cyclic. However, Rolf Landauer had shown[9] in 1961 that the erasure of information is necessarily an irreversible process.[10] In particular, destroying one bit of information increases the entropy of the world by at least

$$\Delta S \geq k \log 2 \ . \tag{3.4.9}$$

So here is the modern resolution of Maxwell's demon: the demon must collect and store information about the molecule. If the demon has a finite memory capacity, he cannot continue to cool the gas indefinitely; eventually, information must be erased. At that point, he finally pays the entropy bill for the cooling he achieved. (If the demon does not erase his record, or if we want to do the thermodynamic accounting before the erasure, then we should associate some entropy with the recorded information.)

3.5 Entropy and Black Holes*

I want to end this lecture with a few comments about entropy, black holes, and quantum gravity.

3.5.1 Information Loss?

Every black hole is characterized by just three numbers: its *mass*, its *spin* and its electric *charge*. It doesn't matter what created the black hole; in the end all information is reduced to just these three numbers. This is summarized in the statement that

Black holes have no hair.

This means that if we throw a book into a black hole, it changes the mass (and maybe the spin and charge) of the black hole, but all information about the content of the book seems lost forever. Do black holes really destroy information? Do they destroy entropy? Do they violate the Second Law?

3.5.2 Black Holes Thermodynamics

The Second Law could be saved if black holes themselves carried entropy and if this entropy increased as an object falls into a black hole. In 1973, Jacob Bekenstein, then a graduate student at Princeton, thought that this was indeed the solution. In fact, there were tantalizing analogies between the evolution of black holes and the laws of thermodynamics. The Second Law of thermodynamics states that entropy never decreases. Similarly, the masses of black holes (or equivalently the area of the event horizon, $A = 4\pi R^2 \propto M^2$) never decreases. Throw an object into a black hole and the black hole gets bigger. Bekenstein thought that this was more than just a cheap analogy. He conjectured that black holes, in fact, carry entropy proportional to their size,[11]

$$S_{\rm BH} \propto A \ . \tag{3.5.10}$$

[9] Landauer, "Irreversibility and Heat Generation in the Computing Process".

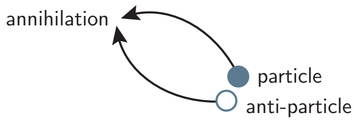
[10] In other words, you can't erase information if you are part of a closed system operating under reversible laws. If you were able to erase information entirely, how would you ever be able to reverse the evolution of the system? If erasure is possible, either the fundamental laws are irreversible—in which case it is not surprising that you can lower the entropy—or you're not really in a closed system. The act of erasing information necessarily transfers entropy to the outside world.

[11] Bekenstein, "Black Holes and the Second Law".

3.5.3 Hawking Radiation

Stephen Hawking thought that this was crazy! If black holes had entropy, they also had a temperature, and you could then show that they had to give off radiation. But everyone knows that black holes are black! Hawking therefore set out to prove Bekenstein wrong. But he failed! What he found[12] instead is quite remarkable: Black holes aren't black! They do give off radiation and do carry huge amounts of entropy.

The key to understanding this is quantum mechanics: In quantum mechanics, the vacuum is an interesting place. According to Heisenberg's uncertainty relation, nothing can be completely empty. Instead particle-antiparticle pairs can spontaneously appear in the vacuum. However, they are only *virtual particles*, living only for a short time, before annihilating each other



Most pop science explanations of this effect are completely wrong (try googling it!), so it is worth giving you are rough outline of the correct argument. We start with the following version of Heisenberg's uncertainty principle

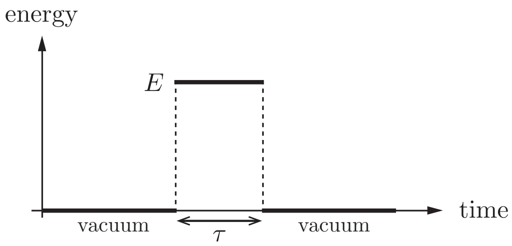
$$\Delta E \Delta t \geq \frac{\hbar}{2} \ .$$

This means the following:

To measure the energy of a system with accuracy ΔE*, one needs a time* $\Delta t \geq \frac{\hbar}{2\Delta E}$.

In other words, to decrease the error in the measurement, we perform an average over a longer time period. But this increases the uncertainty in the time to which this energy applies.

Now consider a particle-antiparticle pair with total energy E spontaneously appearing out nothing:



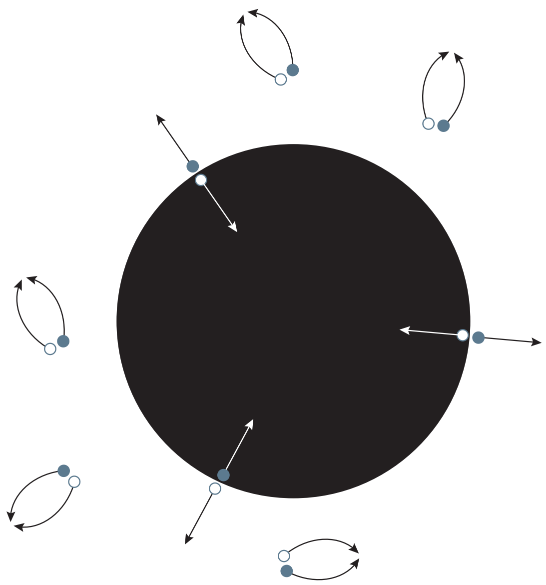
If the lifetime τ of the excited state is *less than* $\frac{\hbar}{2E}$, then we don't have enough time to measure the energy with an accuracy smaller than E. Hence, we can't distinguish the excited state from the zero-energy vacuum. The Heisenberg uncertainty principle allows non-conservation of energy by an amount ΔE for a time $\Delta t \leq \frac{\hbar}{2\Delta E}$.

The story changes if the particle-antiparticle pair happens to be created close to the event horizon[13] of a black hole. In that case, one member of the pair may fall into the black hole and

[12] Hawking, "Black Hole Explosions?"

[13] The event horizon is the point of no return. Nothing, not even light, can escape from inside the event horizon: see Lecture 6.

disappear forever. Missing its annihilation partner, the second particle becomes *real* :



An observer outside the black hole will detect these particles as *Hawking radiation*.

Analyzing this process, Hawking was able to confirm Bekenstein's guess (3.5.10). In fact, he did much more than that. He derived an exact expression for the black hole entropy

$$S_{\rm BH} = \frac{1}{4}\frac{A}{\ell_{\rm p}^2} \ , \tag{3.5.11}$$

where $\ell_{\rm p}$ is the Planck length, the scale at with the effects of quantum mechanics and gravity become equally important (see Lecture 6). In terms of the fundamental constants of quantum mechanics ($\hbar$), relativity (c) and gravity (G), the Planck length is

$$\ell_{\rm p} = \sqrt{\frac{\hbar G}{c^3}} \approx 1.6 \times 10^{-35}\,{\rm m} \ . \tag{3.5.12}$$

Eq. (3.5.11) is a remarkable formula: it links entropy and thermodynamics (l.h.s.) to quantum gravity (r.h.s.). It is therefore the single most important clue we have about the reconciliation of gravity with quantum mechanics.

3.5.4 Black Holes in String Theory

The great triumph of Boltzmann's theory of entropy was that he was able to explain an observable macroscopic quantity—the entropy—in terms of a counting of microscopic components. Hawking's formula for the entropy of a black hole seems to be telling us that there are a very large number of microstates corresponding to any particular macroscopic black hole

$$S_{\rm BH} = k \log W_{\rm BH} \ . \tag{3.5.13}$$

What are those microstates? They are not apparent in classical gravity (where a black hole has no hair). Ultimately, they must be states of quantum gravity. Some progress has been made on this in string theory, our best candidate theory of quantum gravity. In 1996, Andy Strominger and Cumrun Vafa derived the black hole entropy from a microscopic counting of the degrees of freedom of string theory (which are strings and higher-dimensional membranes). They got eq. (3.5.11) on the nose, including the all important factor of 1/4.

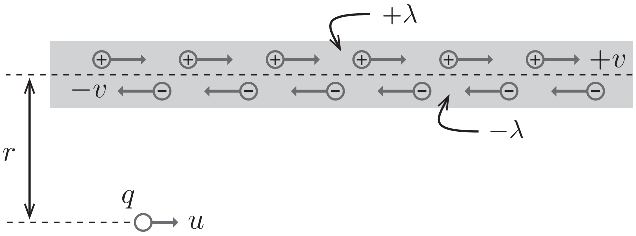
4 Electrodynamics and Relativity

The first time I experienced beauty in physics was when I learned how Einstein's **special relativity** is hidden in the equations of Maxwell's theory of **electricity** and **magnetism**. Today, it is my privilege to show this to you. In the end, we will speculate how **gravity** might fit into this.

4.1 Relativity requires Magnetism

Imagine you had never heard of magnetism. But you happen to know electrostatics and relativity. It is then possible to show that there *has* to be such a thing as magnetism.

Consider a string of positive charges moving to the right with velocity v and negative charges moving to the left with velocity $-v$:

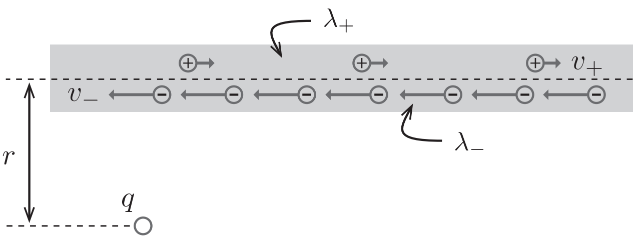


At a distance r, there is a point charge q travelling to the right at speed $u < v$. The charges are close enough together so that we may regard them as continuous line charges with densities $\pm\lambda$ (= charge/length). The net current to the right is

$$I = 2\lambda v \ . \tag{4.1.1}$$

Because the positive and negative line charges cancel, there is *no electrical force on* q (in the rest frame of the wire).

Now consider the same situation from the point of view of an observer comoving with the charge q (i.e. an observer travelling to the right with speed u):



The velocity of the positive charges is now smaller than the velocity of the negative charges. The Lorentz contraction of the spacing between negative charges is more severe than that between positive charges; therefore, *the wire, in this frame, carries a net negative charge!* There is now an electric force on the charge. In the box below I show you how to compute this force in terms of the charge density in the original frame. Transforming this force back to the original frame leads to

$$F = -quB \ , \qquad \text{where} \quad B \equiv \frac{\mu_0 I}{2\pi r} \ . \tag{4.1.2}$$

We have derived the Lorentz force using only electrostatics and a sequence of relativistic transformations. I encourage you to look at the details of the calculation. It is neat.

Let us call the original frame $\mathcal{S}$ and the new frame $\mathcal{S}'$. By the Einstein velocity addition rule, the velocities of the positive and negative charges in $\mathcal{S}'$ are

$$v_\pm = \frac{v \mp u}{1 \mp vu/c^2} \ . \tag{4.1.3}$$

If λ_0 is the charge density of the positive line charge in its own rest frame, then

$$\lambda_\pm = \pm(\gamma_\pm)\lambda_0 \ , \qquad \text{where} \quad \gamma_\pm = \frac{1}{\sqrt{1 - v_\pm^2/c^2}} \ . \tag{4.1.4}$$

Of course, λ_0 is not the same as λ, but

$$\lambda = \gamma \lambda_0 \ , \qquad \text{where} \quad \gamma = \frac{1}{\sqrt{1 - v^2/c^2}} \ . \tag{4.1.5}$$

With a bit of algebra, you can show that

$$\gamma_\pm = \frac{1 \mp uv/c^2}{\sqrt{1 - u^2/c^2}} \times \gamma \ . \tag{4.1.6}$$

The net line charge in $\mathcal{S}'$ is

$$\lambda_{\rm tot} = \lambda_+ + \lambda_- = \lambda_0(\gamma_+ - \gamma_-) = -\frac{2uv}{c^2} \times \frac{\lambda}{\sqrt{1 - u^2/c^2}} \ . \tag{4.1.7}$$

This creates an *electric* field

$$E' = \frac{1}{\epsilon_0} \frac{\lambda_{\rm tot}}{2\pi r} \ . \tag{4.1.8}$$

Hence, in the frame $\mathcal{S}'$ there is an *electrical force on q*:

$$F' = qE' = -\frac{qu}{\sqrt{1 - u^2/c^2}} \times \frac{2\lambda v}{2\pi r} \times \frac{1}{\epsilon_0 c^2} \ . \tag{4.1.9}$$

But if there is a force on q in $\mathcal{S}'$, there must be one in $\mathcal{S}$. In the relativity course last term, you learned how to transform forces between the frames. Since q is at rest in $\mathcal{S}'$, and F' is perpendicular to u, the force in $\mathcal{S}$ is given by

$$F = \sqrt{1 - u^2/c^2}\, F' = -qu \times \frac{2\lambda v}{2\pi r} \times \frac{1}{\epsilon_0 c^2} \ . \tag{4.1.10}$$

The charge is attracted toward the wire by a force that is purely electrical in $\mathcal{S}'$ (where the wire is charged, and q is at rest), but distinctly *non*-electrical in $\mathcal{S}$ (where the wire is neutral). Taken together, then, electrostatics and relativity imply the existence of another force. This other force is, of course, the *magnetic* force. Expressing λv in terms of the current (4.1.1), we get

$$F = -qu \left(\frac{\mu_0 I}{2\pi r}\right) \ , \tag{4.1.11}$$

where we have defined $\mu_0 \equiv (\epsilon_0 c^2)^{-1}$. The term in parentheses is the magnetic field of a long, straight wire, and the force is precisely what we would have obtained by using the Lorentz force law in $\mathcal{S}$,

$$F = -quB \ , \qquad \text{where} \quad B \equiv \frac{\mu_0 I}{2\pi r} \ . \tag{4.1.12}$$

We have derived the magnetic force between a current-carrying wire and a moving charge without ever invoking the laws of magnetism.

4.2 Magnetism requires Relativity

4.2.1 Maxwell's Equations

All of electricity and magnetism is contained in the four Maxwell equations:

$$\vec{\nabla} \cdot \vec{E} = \frac{\rho}{\epsilon_0} \ , \tag{M1}$$

$$\vec{\nabla} \cdot \vec{B} = 0 \ , \tag{M2}$$

$$\vec{\nabla} \times \vec{E} = -\frac{\partial \vec{B}}{\partial t} \ , \tag{M3}$$

$$\vec{\nabla} \times \vec{B} = \mu_0 \vec{J} + \mu_0 \epsilon_0 \frac{\partial \vec{E}}{\partial t} \ . \tag{M4}$$

where ρ and $\vec{J}$ are the total charge and current densities, which satisfy the continuity equation

$$\frac{\partial \rho}{\partial t} = -\vec{\nabla} \cdot \vec{J} \ . \tag{C}$$

The parameters ϵ_0 and μ_0 are simply constants that convert units.

I am assuming that you have seen these before[1], but let me remind you briefly of the meaning of each of these equations:

- *Gauss' Law*

  (M1) simply states that "electric charges produce electric fields". Specifically, the divergence of the electric field $\vec{E}$ is determined by the charge density ρ (= charge/volume).

- *No Monopoles*

  Unlike electric charges, we have never observed isolated magnetic charges ("monopoles"). Magnetic charges always come in pairs of plus and minus (or North and South). (If you cut a magnet into two pieces, you get two new magnets, each with their own North and South poles.) The net magnetic charge density is therefore zero. (M2) embodies this fact, by having zero in place of ρ in the equation for the divergence of the magnetic field $\vec{B}$.

- *Faraday's Law of Induction*

  (M3) describes how a time-varying magnetic field creates ("induces") an electric field. This is easily demonstrated by moving a magnet through a loop of wire. This induces an electric field, which forces a current through the wire, which might power an attached light bulb.

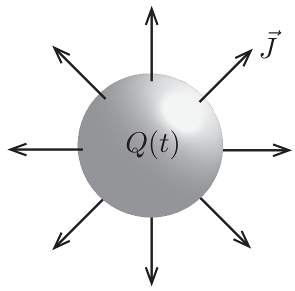
- *Ampère's Law*

  It was known well before Maxwell that an electric current produces a magnetic field—e.g. Oersted discovered that a magnet feels a force when placed near a wire with a current flowing through it. However, the equation that people used to describe this, $\vec{\nabla} \times \vec{B} = \mu_0 \vec{J}$, was wrong. Maxwell pointed out that the original form of Ampère's law was inconsistent with the conservation of electric charge, eq. (C), (do you see why?). To fix this he added an extra term to (M4). This extra term implies that a time-varying electric field also produces a magnetic field. This led to one of the most important discoveries in the history of physics.

[1]See Example Sheet 3 of your course on *'Vector Calculus'*.

- *Conservation of Charge*

  Charges and currents source electric and magnetic fields. But, a current is nothing but moving charges, so it is natural to expect that ρ and $\vec{J}$ are related to each other. This relation is the continuity equation (C). Consider a small volume V (bounded by a surface S) containing a total charge Q.



  The amount of charge moving out of this volume per unit time equals the current flowing through the surface:

$$\frac{\partial}{\partial t}Q = -\int \mathrm{d}\vec{S}\cdot\vec{J}\ . \tag{4.2.13}$$

  Eq. (C) embodies this locally (i.e. for every point in space),

$$\frac{\partial}{\partial t}Q = \frac{\partial}{\partial t}\int \mathrm{d}V\,\rho = -\int \mathrm{d}V\,\vec{\nabla}\cdot\vec{J} = -\int \mathrm{d}\vec{S}\cdot\vec{J}\ , \tag{4.2.14}$$

  where we used Stokes' theorem in the last equality.

Maxwell's equations are ugly! As we will see, this is because space and time are treated separately, while we now know that Nature is symmetric in time and space. By uncovering this hidden symmetry we will be able to write the four Maxwell equations in terms of a single elegant equation. In the process, we will develop a unified relativistic description of electricity and magnetism.

4.2.2 Let There Be Light!

From (M3) we see that a time-dependent magnetic field, $\vec{B}(t)$, produces an electric field. Similarly, (M4) implies that a time-dependent electric field, $\vec{E}(t)$, creates a magnetic field. So we see that "a changing magnetic field produces an electric field, which produces a magnetic field, which produces an electric field, which ...". Once we set up a time-dependent electric or magnetic field, it seems to allow for self-sustained solutions that oscillate between electric and magnetic. Since the oscillating fields also propagate through space, we call them *electromagnetic waves.*

Let us describe these electromagnetic waves more mathematically. We restrict to empty space (i.e. a perfect vacuum with no charges and currents, $\rho = \vec{J} = 0$). The Maxwell's equations then are

$$\vec{\nabla}\cdot\vec{E} = 0\ , \tag{M1$'$}$$

$$\vec{\nabla}\cdot\vec{B} = 0\ , \tag{M2$'$}$$

$$\vec{\nabla}\times\vec{E} = -\frac{\partial\vec{B}}{\partial t}\ , \tag{M3$'$}$$

$$\vec{\nabla}\times\vec{B} = \mu_0\epsilon_0\frac{\partial\vec{E}}{\partial t}\ . \tag{M4$'$}$$

Now take the curl of the curl equation for the $\vec{E}$-field, i.e. $\vec{\nabla}\times$(M3′),

$$\vec{\nabla} \times (\vec{\nabla} \times \vec{E}) = -\frac{\partial}{\partial t}\vec{\nabla} \times \vec{B} = -\mu_0\epsilon_0 \frac{\partial^2 \vec{E}}{\partial t^2} \ , \tag{4.2.15}$$

where we have used (M4′) in the final equality. We use a vector identity to manipulate the l.h.s. of this expression,

$$\vec{\nabla} \times (\vec{\nabla} \times \vec{E}) = \vec{\nabla}(\vec{\nabla} \cdot \vec{E}) - \nabla^2 \vec{E} \tag{4.2.16}$$

$$= -\nabla^2 \vec{E} \ , \tag{4.2.17}$$

where we have used (M1′) in the final equality. We get

$$\mu_0\epsilon_0 \frac{\partial^2 \vec{E}}{\partial t^2} - \nabla^2 \vec{E} = 0 \ . \tag{4.2.18}$$

This partial differential equation is the *wave equation* for the $\vec{E}$-field.[2] Try substituting

$$\vec{E}(t,x) = \vec{E}_0 \sin\left[2\pi\left(\frac{t}{T} - \frac{x}{\lambda}\right)\right] \ , \tag{4.2.20}$$

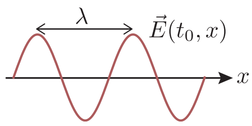
where T and λ are constants. You will find that this is a solution of (4.2.18), if

$$v \equiv \frac{\lambda}{T} = \frac{1}{\sqrt{\mu_0\epsilon_0}} \ , \tag{4.2.21}$$

where v is the speed of the wave.

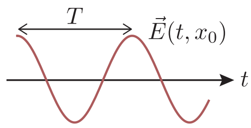
It is easy to see that (4.2.20) indeed describes a wave:

First, consider a snapshot of the solution at a fixed time $t = t_0$. The electric field $\vec{E}(t_0, x)$ varies periodically through space:



The solution repeats every distance λ along the x-axis.

Next, imagine sitting at a fixed point $x = x_0$. The electric field at that point, $\vec{E}(t, x_0)$, oscillates in time:



The solution repeats with a time period T. Hence, (4.2.20) indeed describes a wave propagating along the x-axis with speed $v = \lambda/T$ given by (4.2.21).

As boring as eq. (4.2.21) looks, it is an absolutely remarkable result! Through the genius of Einstein it led to the special theory of relativity.

[2] Similar manipulations for the curl equation for the $\vec{B}$-field, eq. (M4′), give the wave equation for the $\vec{B}$-field

$$\mu_0\epsilon_0 \frac{\partial^2 \vec{B}}{\partial t^2} - \nabla^2 \vec{B} = 0 \ . \tag{4.2.19}$$

4.2.3 Racing A Light Beam

Playing with coils and metal plates, you would find that $\epsilon_0 = 8.85 \times 10^{-12}\,\mathrm{C^2/Nm^2}$ and $\mu_0 = 4\pi \times 10^{-7}\,\mathrm{N/A^2}$. Eq. (4.2.21) then predicts that all electromagnetic waves propagate with

$$c \equiv \frac{1}{\sqrt{\mu_0 \epsilon_0}} = 3 \times 10^8\,\mathrm{m/s}\ . \tag{4.2.22}$$

But, this is just the speed of *light*. Electromagnetic waves *are* light!

Eq. (4.2.22) looks innocent, but notice that I never mentioned the speed of the observer. According to Maxwell's equation,

the speed of light is independent of the motion of the observer.

This flies in the face of Newton's law of velocity addition. The light emitted from a moving spaceship travels at exactly the same speed as the light from a stationary source. This means that you can never catch up with a light ray. No matter how fast you run after a light ray, it will always recede from you at speed c. It would have been easy to dismiss this craziness as a flaw of Maxwell's theory. However, Einstein dared to accept this strange feature of light as a fundamental principle of Nature and built his theory of relativity around it.

4.2.4 My Time Is Your Space

To give up Newton's simple law of velocity addition, means to give up on absolute measurements of time and space. In other words, in order for two observers $\mathcal{O}$ and $\mathcal{O}'$ in motion relative to each other to agree on the speed of light, they have to *disagree* on measurements of time (Δt vs. $\Delta t'$) and space (Δx vs. $\Delta x'$). However, that disagreement is of a very specific kind, such that everybody agrees on the value of the speed of light

$$c = \frac{\Delta x}{\Delta t} = \frac{\Delta x'}{\Delta t'}\ . \tag{4.2.23}$$

Eq. (4.2.23) can be rewritten as follows

$$c^2 \Delta t^2 - \Delta x^2 = c^2 (\Delta t')^2 - (\Delta x')^2 = 0\ . \tag{4.2.24}$$

In fact, even for objects not moving at the speed of light, observers will disagree on Δx and Δt, but will always agree on the combination

$$\Delta s^2 \equiv c^2 \Delta t^2 - \Delta x^2 = c^2 (\Delta t')^2 - (\Delta x')^2\ . \tag{4.2.25}$$

More important than the fact that space and time are relative, is the fact that space and time are related by a special *symmetry* that guarantees the invariance of Δs^2. This symmetry is called *Lorentz symmetry*.

As you know, an elegant way to make this symmetry manifest is to combine space and time coordinates into a single four-dimensional *spacetime* coordinate

$$t\,,\ \vec{x} \qquad \Rightarrow \qquad X^\mu = \begin{pmatrix} \overset{\text{time}}{ct} \\ \underset{\text{space}}{\vec{x}} \end{pmatrix}\ . \tag{4.2.26}$$

The "square" of the spacetime four-vector is

$$s^2 \equiv \eta_{\mu\nu} X^\mu X^\nu \ , \tag{4.2.27}$$

where I have introduced the *Minkowski metric*,

$$\eta_{\mu\nu} \equiv \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} . \tag{4.2.28}$$

Eq. (4.2.27) uses the Einstein summation convention, so that repeated indices are summed over. You should remember that the coordinates of different observers are related by a *Lorentz transformation*

$$X'_\mu = \Lambda^\nu{}_\mu(v)\, X_\nu \ . \tag{4.2.29}$$

You can think of these as "rotations" in the 4d spacetime, that leave $s^2 = X^\mu X_\mu$ the same. This is analogous to 3d spatial rotations

$$x'_i = R^j{}_i(\theta)\, x_j \ , \tag{4.2.30}$$

that leave the magnitude of the position 3-vector, $\ell^2 = x^i x_i$, the same.

4.3 Relativity unifies Electricity and Magnetism

I promised you that the Maxwell equations become beautiful when written in a relativistic way. Let's see how that works.

4.3.1 Relativistic Electrodynamics

First, consider Gauss' law for the magnetic field, eq. (M2),

$$\vec{\nabla} \cdot \vec{B} = 0 \ . \tag{4.3.31}$$

Since "div curl = 0", this is solved automatically if $\vec{B}$ is written as the curl of a vector field $\vec{A}$,

$$\boxed{\vec{B} = \vec{\nabla} \times \vec{A}} \ . \tag{4.3.32}$$

Next, we use eq. (4.3.32), to rewrite Faraday's law for the electric field, eq. (M3), $\vec{\nabla} \times \vec{E} = -\dot{\vec{B}}$, as

$$\vec{\nabla} \times \left[\vec{E} + \dot{\vec{A}}\right] = 0 \ . \tag{4.3.33}$$

Since "curl grad = 0", we can write $\vec{E} + \dot{\vec{A}}$ as the gradient of a scalar function. The Maxwell equation (M3) is hence solved if $\vec{E}$ is written as

$$\boxed{\vec{E} = -\vec{\nabla}\phi - \dot{\vec{A}}} \ . \tag{4.3.34}$$

The choice of vector field $\vec{A}$ is not unique. We can add the gradient of any scalar function $\alpha(t,\vec{x})$ to $\vec{A}$ without changing the magnetic field $\vec{B} = \vec{\nabla} \times \vec{A}$ (since "curl grad = 0"),

$$\vec{A} \mapsto \vec{A}' = \vec{A} + \vec{\nabla}\alpha \ . \tag{4.3.35}$$

The electric field $\vec{E} = -\vec{\nabla}\phi - \dot{\vec{A}}$ also remains unchanged under the transformation (4.3.35), if ϕ simultaneously transforms as

$$\phi \mapsto \phi' = \phi - \dot{\alpha} \ . \tag{4.3.36}$$

We hence have the freedom to pick any scalar α without changing $\vec{E}$ and $\vec{B}$. This is equivalent to the freedom of fixing $\vec{\nabla} \cdot \vec{A}$, since

$$\vec{\nabla} \cdot \vec{A}' = \vec{\nabla} \cdot \vec{A} + \nabla^2 \alpha \ . \tag{4.3.37}$$

A particularly useful choice is the so-called *Lorenz gauge*

$$\vec{\nabla} \cdot \vec{A} = -\frac{\dot{\phi}}{c^2} \ . \tag{4.3.38}$$

We have shown that the two Maxwell equations (M2) and (M3) are automatically satisfied if the two vector fields $\vec{E}$ and $\vec{B}$ are expressed in terms of a scalar field ϕ and a new vector field $\vec{A}$. But, a scalar and a vector can be combined into a four-vector. Just like t and $\vec{x}$ were combined into $X^\mu = (ct, \vec{x})$. Let's try and define the *vector potential*

$$\phi\,,\,\vec{A} \qquad \Rightarrow \qquad A^\mu = \begin{pmatrix} \overset{\text{electricity}}{\phi/c} \\ \underset{\text{magnetism}}{\vec{A}} \end{pmatrix} \ . \tag{4.3.39}$$

A single four-vector describes both electric and magnetic fields. As we will see in Lecture 5, in quantum mechanics, A_μ is actually more fundamental than $\vec{E}$ and $\vec{B}$.

The transformations (4.3.35) and (4.3.36) combine into a single equation

$$A_\mu \to A'_\mu = A_\mu - \partial_\mu \alpha \ , \tag{4.3.40}$$

where we have defined the spacetime derivative

$$\partial_\mu \equiv \frac{\partial}{\partial X^\mu} = \left(\frac{1}{c}\frac{\partial}{\partial t}, \frac{\partial}{\partial \vec{x}}\right) \ . \tag{4.3.41}$$

The Lorenz gauge condition (4.3.38) becomes

$$\partial^\mu A_\mu = 0 \ . \tag{4.3.42}$$

Two Maxwell equations are automatically satisfied, but what about the other two? (M1) and (M4) have source terms: the charge density ρ (a scalar) and the current density $\vec{J}$ (a vector). Again, a scalar and a vector make a four-vector

$$\rho\,,\,\vec{J} \qquad \Rightarrow \qquad J^\mu = \begin{pmatrix} \overset{\text{electricity}}{c\rho} \\ \underset{\text{magnetism}}{\vec{J}} \end{pmatrix} \ . \tag{4.3.43}$$

As I show in the insert below, expressed in terms of the four-vectors A^μ and J^μ, the Maxwell equations (M1) and (M4), unify into a single pretty equation[3]

$$\boxed{\Box A^\mu = \mu_0 J^\mu} \ , \tag{M5}$$

where $\Box$ is shorthand for $\eta^{\mu\nu}\partial_\mu\partial_\nu = \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \nabla^2$. Notice that $\Box$ is the same combination that appears in the wave equation (4.2.18). Hence, there will be similar wave solutions for A_μ.

Substitute eq. (4.3.34) into the Maxwell equation (M1),

$$\vec{\nabla}\cdot\vec{E} = -\frac{\partial}{\partial t}\vec{\nabla}\cdot\vec{A} - \nabla^2\phi = \frac{\rho}{\epsilon_0} \ , \tag{4.3.44}$$

or

$$\left(\frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \nabla^2\right)\phi = \frac{\rho}{\epsilon_0} + \frac{\partial}{\partial t}\left(\vec{\nabla}\cdot\vec{A} + \frac{\dot{\phi}}{c^2}\right) \ . \tag{4.3.45}$$

In the Lorenz gauge (4.3.38), this becomes

$$\Box\phi = \frac{\rho}{\epsilon_0} \ . \tag{4.3.46}$$

Substituting eqs. (4.3.32) and (4.3.34) into (M4) gives

$$\vec{\nabla}\times\vec{B} = \vec{\nabla}\times(\vec{\nabla}\times\vec{A}) = \mu_0\vec{J} + \epsilon_0\mu_0(-\ddot{\vec{A}} - \vec{\nabla}\dot{\phi}) \ . \tag{4.3.47}$$

Using the identity $\vec{\nabla}\times(\vec{\nabla}\times\vec{A}) = \vec{\nabla}(\vec{\nabla}\cdot\vec{A}) - \nabla^2\vec{A}$, we get

$$\left(\frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \nabla^2\right)\vec{A} = \mu_0\vec{J} - \vec{\nabla}\left(\vec{\nabla}\cdot\vec{A} + \frac{\dot{\phi}}{c^2}\right) \ . \tag{4.3.48}$$

In the Lorenz gauge (4.3.38), this becomes

$$\Box\vec{A} = \mu_0\vec{J} \ . \tag{4.3.49}$$

Eqs. (4.3.46) and (4.3.49) combine into eq. (M5).

Let us summarise what we have achieved:

1. We unified electric ($\vec{E}$) and magnetic ($\vec{B}$) fields into the 4-vector potential A_μ.
2. We reduced the 4 Maxwell equations to just 1.

4.3.2 A Hidden Symmetry*

Eq. (M5) is more than just a pretty way of representing all the information of the Maxwell equations. It has *Lorentz symmetry*! Everything is packaged into 4-vectors without any lose ends. Space and time appear precisely in the symmetric way required by relativity.

The 4-vectors A^μ and J^μ, transform in the same way as X^μ under Lorentz transformation, cf. eq. (4.2.29),

$$A^\mu \mapsto A^{\mu\prime} = \Lambda^\mu{}_\nu A^\nu \ , \tag{4.3.50}$$

$$J^\mu \mapsto J^{\mu\prime} = \Lambda^\mu{}_\nu J^\nu \ . \tag{4.3.51}$$

[3] Eq. (M5) uses the Lorenz gauge (4.3.38). It looks a little less pretty in a general gauge: $\Box A^\mu = \mu_0 J^\mu + \partial^\mu(\partial^\nu A_\nu)$. This can be rearranged into $\partial_\nu F^{\nu\mu} = \mu_0 J^\mu$, where $F^{\nu\mu} \equiv \partial^\nu A^\mu - \partial^\mu A^\nu$.

The operator $\Box$ doesn't have a free index, so it doesn't transform. Explicitly,

$$\Box \mapsto \Box' = \eta_{\mu\nu}\partial^{\mu\prime}\partial^{\nu\prime} = \eta_{\mu\nu}\Lambda^{\mu}{}_{\alpha}\Lambda^{\nu}{}_{\beta}\partial^{\alpha}\partial^{\beta} = \eta_{\alpha\beta}\partial^{\alpha}\partial^{\beta} = \Box \ . \tag{4.3.52}$$

Hence,

$$\Box A^{\mu} \mapsto \Box' A^{\mu\prime} = \Lambda^{\mu}{}_{\nu}\Box A^{\nu} \ . \tag{4.3.53}$$

Combining (4.3.53) and (4.3.51), we see that if (M5) holds in the frame $\mathcal{S}'$, then it also holds in the frame S,

$$\Box' A^{\mu\prime} = \mu_0 J^{\mu\prime} \quad \Leftrightarrow \quad \Box A^{\mu} = \mu_0 J^{\mu} \ . \tag{4.3.54}$$

This is the defining feature of a theory that is consistent with Einstein's relativity. Its equations take the *same form* in all inertial frames.[4]

It is remarkable that Maxwell's theory automatically was of that form and didn't need any fixing. This isn't true for many other theories. Consider, for example, Coulomb's law

$$m\frac{d^2x}{dt^2} = \frac{1}{4\pi\epsilon_0}\frac{qQ}{x^2} \ . \tag{4.3.55}$$

No matter how much massaging you do, you will never be able to write this as an equation involving only four-vectors. It also easy to see that (4.3.55) violates the basic principle of relativity that nothing travels faster than the speed of light. Changing the source at $x = 0$ *instantly* affects the solution *everywhere* in space. The famous Coulomb's law is therefore only an approximation. The exact result follows from Maxwell's equations.

4.4 More Unification?*

We are done with the basic story of the unification of electricity and magnetism through relativity. However, it is tempting to speculate a little further and ask how gravity, the second fundamental force of Nature, could fit into this. Let me emphasize that I am describing this just for fun and it is not clear whether this has anything to do with reality.

4.4.1 Kaluza-Klein Theory

In 1919, Theodor Kaluza send a letter to Einstein. He described to him a proposal for unifying gravity and electromagnetism. Einstein liked the idea.

In this lecture, we have seen that the fundamental field in electromagnetism is the four-vector potential $A_\mu(X)$. Here, I have indicated that A_μ depends on the spacetime coordinate X^μ. In Lecture 6, we will see that gravity is described by a more complicated object called the *metric* $g_{\mu\nu}(X)$. This is like the Minkowski metric (4.2.28), except that now all entries of the "matrix" can depend on the spacetime location X^μ. In Lecture 6, we will have much more to say about $g_{\mu\nu}$. For now, just think of $g_{\mu\nu}$ as a 4×4 matrix whose entries vary through spacetime. Since $g_{\mu\nu}$ is symmetric it has 10 independent components.

It is a fundamental theorem of algebra that

$$15 = 10 + 4 + 1 \ . \tag{4.4.56}$$

[4] These days we usually reverse the logic. When we come up with new theories, we don't dare to write down anything that doesn't have Lorentz symmetry.

Kaluza noticed that 15 is the number of independent components of a 5×5 symmetric matrix. A 5×5 symmetric matrix G_{MN} has enough room to fit both the vector potential and the spacetime metric[5]

$$\begin{pmatrix} G_{00} & G_{01} & G_{02} & G_{03} & G_{04} \\ G_{10} & G_{11} & G_{12} & G_{13} & G_{14} \\ G_{20} & G_{21} & G_{22} & G_{23} & G_{24} \\ G_{30} & G_{31} & G_{32} & G_{33} & G_{34} \\ G_{40} & G_{41} & G_{42} & G_{43} & G_{44} \end{pmatrix} = \begin{pmatrix} \overbrace{\begin{matrix} g_{00} & g_{01} & g_{02} & g_{03} \\ g_{10} & g_{11} & g_{12} & g_{13} \\ g_{20} & g_{21} & g_{22} & g_{23} \\ g_{30} & g_{31} & g_{32} & g_{33} \end{matrix}}^{\text{gravity}} & \begin{matrix} A_0 \\ A_1 \\ A_2 \\ A_3 \end{matrix} \\ \underbrace{\begin{matrix} A_0 & A_1 & A_2 & A_3 \end{matrix}}_{\text{electromagnetism}} & S \end{pmatrix} \tag{4.4.57}$$

In more compact form,

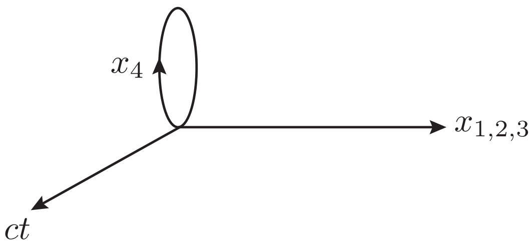
$$G_{MN} = \begin{pmatrix} g_{\mu\nu} & A_\mu \\ A_\mu & S \end{pmatrix} \tag{4.4.58}$$

Kaluza went on to speculate about the physical meaning of the object G_{MN}. He suggested it may be the metric of a five-dimensional spacetime. More than that: he showed that the equation of gravity in 5D could be split into the equations of Einstein's theory gravity in 4D and Maxwell's theory of electromagnetism:

$$\text{5D gravity} = \text{4D (gravity + electromagnetism)} \,. \tag{4.4.59}$$

Einstein was intrigued that gravity in five dimensions unifies gravity and electromagnetism in four dimensions.

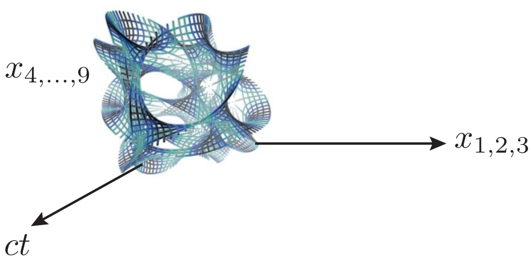
Kaluza's theory had an obvious flaw. The world isn't five-dimensional! Or, is it? In 1926, Oscar Klein pointed out that we wouldn't notice that the world is five-dimensional if one of the space dimensions, say x_4, is curled up into a small circle:



This became known as Kaluza-Klein theory.

4.4.2 String Theory

String theory requires extra dimensions of space to be self-consistent. Most versions of the theory have six extra dimensions. They are curled up into a small ball, maybe as small as 10^{-34} cm:

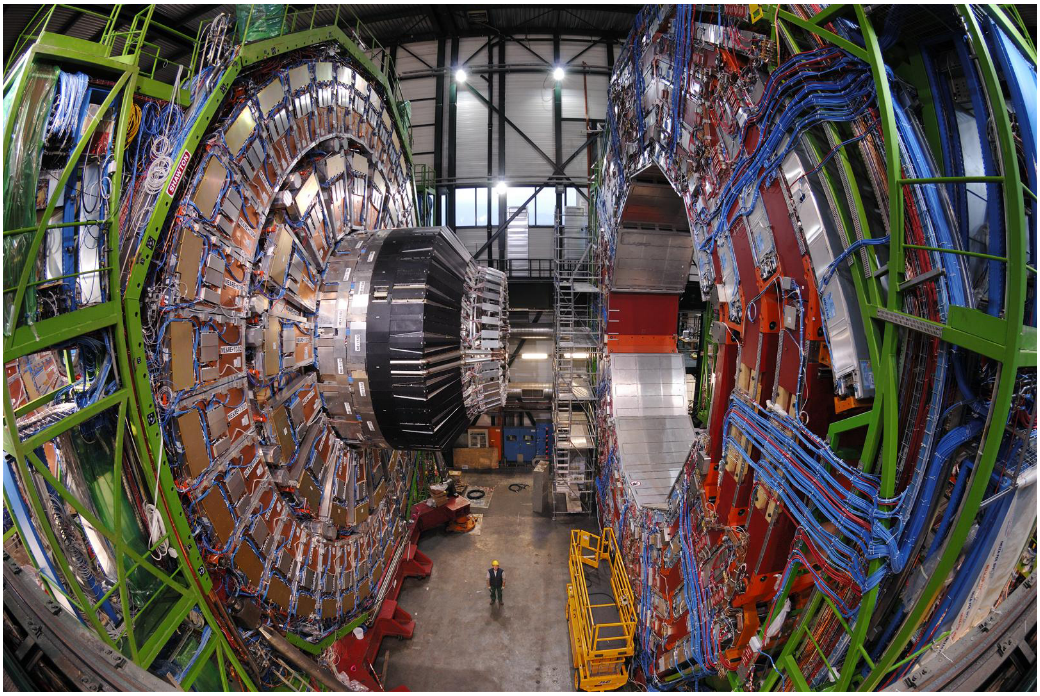


[5]In fact, G_{MN} has room to spare for an extra scalar field $S(X)$. We will ignore this detail.

Now there is enough room to unify all forces of Nature: gravity, electromagnetism, weak and strong nuclear force all become one. Unfortunately, we haven't yet found ways to test the theory experimentally. It therefore remains an intriguing, but highly speculative, endeavor.

5 Particle Physics

This lecture is about **particle physics**, the study of the fundamental building blocks of Nature and the forces between them. We call our best theory of particle physics the **Standard Model** (in capital letters). This pompous name reflects our confidence in the theory. Recently, the Large Hadron Collider (LHC) at CERN discovered the last puzzle piece of the Standard Model: the Higgs particle. In this lecture, you will learn why the Higgs is so important.



5.1 The Standard Model

5.1.1 A New Periodic Table

We have seen how quantum mechanics explains the periodic table of elements. These elements used to be considered the building blocks of Nature. However, these days we have simplified things quite a bit. We now know that atoms aren't indivisible. Each atom consists of a nucleus, surrounded by a swarm of *electrons* (e). The nucleus carries by far the majority of the mass but takes up a tiny space. The nucleus is composed of *protons* (p) and *neutrons* (n). And each proton and neutron is itself made out of three *quarks* (q). In fact, there are two different types of quarks: due to a lack of imagination they're called "up" (u) and "down" (d).

The proton contains two up quarks and a down quark

$$p = (uud) \ ,$$

and the neutron contains two down quarks and an up quark

$$n = (ddu) \ .$$

So we've reduced the 110 elements in the periodic table to just 3: an electron and two quarks. In fact, there's one further particle that we should add called the (electron) *neutrino* (ν_e). It doesn't live inside atoms, but it is created in certain radioactive processes, most notably in beta decay and inside the Sun. Right now, about 60 billion of them are streaming through every square centimetre of your body every second. Virtually all of them sail straight through you, which is why you've never seen or felt one.

These four particles give us a new periodic table:

charge		mass
$+\frac{2}{3}$	u	~ 4
$-\frac{1}{3}$	d	~ 8
-1	e	1
0	ν_e	~ 0

That's pretty nice. Four is a good number. Except that Nature did something very odd, and no one really knows why. It repeated this pattern twice more. These copies (called "generations") have exactly the same properties as the first lot—which means the same charge, and the same properties under other forces—but their masses are different.

charge		mass		mass		mass
$+\frac{2}{3}$	u	~ 4	c	$\sim 2,000$	t	$\sim 340,000$
$-\frac{1}{3}$	d	~ 8	s	~ 200	b	$\sim 8,000$
-1	e	1	μ	~ 200	τ	$\sim 3,000$
0	ν_e	~ 0	ν_μ	~ 0	ν_τ	~ 0

Why does this happen? We have absolutely no idea. Not a whole lot would change if this wasn't the case.

The masses of these particles look rather random. Why these numbers? Notice that the mass of the up quark is smaller than that of the down quark, while it is the opposite in the other two generations. This is one of the most important inversions in science! It means that neutrons are heavier than protons. As a consequence, if neutrons are left on their own, they decay into protons in about 14 minutes. If this down/up mass ratio was inverted then protons would decay into neutrons instead; the nuclei of atoms would be unstable and there wouldn't be any interesting chemistry to speak of. There would then also be no biology and no life. So it's indeed an important property. But is it an important question?!! Should we try to answer it, or is it just one of those fortunate accidents of life that we have to accept?

Finally, for each type of particle there is a corresponding *anti-particle*—i.e. particles with the same mass, but opposite charge. This doubles the number of particle species. When a particle and an anti-particle come in contact, they annihilate. This process was particularly efficient in the early universe when the density of particles was very high. In fact, it is a big mystery why there is any matter left at all. Something in the early universe must have created an imbalance between matter and anti-matter, so that some matter survived the annihilation process. We don't know how that happened.

5.1.2 Four Forces Bind Them All

The particles of the Standard Model interact with each other through four fundamental forces:

- *Electromagnetism*

  Electric forces are the reason why a table feels solid although it is mostly empty space. When you push against a table, you try to push the outer electrons in the atoms of your hand into the outer electrons in the atoms that make up the table. But like charges repel, so the electrons resist being pushed into each other.

- *Strong Nuclear Force*

  The nuclei of heavy elements consist of a large number of protons and neutrons. However, protons carry positive electric charge, so they want to repel each other. What holds the nuclei together? It's the strong nuclear force. It acts between the protons and neutrons to counterbalance the electric repulsion.

- *Weak Nuclear Force*

  As mentioned before, free neutrons are unstable, decaying into protons within about 14 min. What force triggers this decay? The lifetimes of particles that decay via the strong force is typically fractions of seconds, not minutes. We need a weaker force. We call it the weak nuclear force.

- *Gravity*

  We experience gravity every day, so it may come as a surprise that we understand it the least. For example, we don't understand why gravity is so much weaker than all other forces in Nature. A small magnet can lift a paper clip off the table, thereby overcoming the gravitational pull of the entire Earth! Gravity is an extremely weak force. The electric force between two electrons is a factor of 10^{40} times larger than their gravitational attraction. This is why gravity can usually be completely ignored when discussing elementary particles and atoms. Only when an enormous number of particles is combined, e.g. to form a planet like the Earth, does gravity become important. But, fundamentally gravity is an extremely weak force. And we have no idea why!

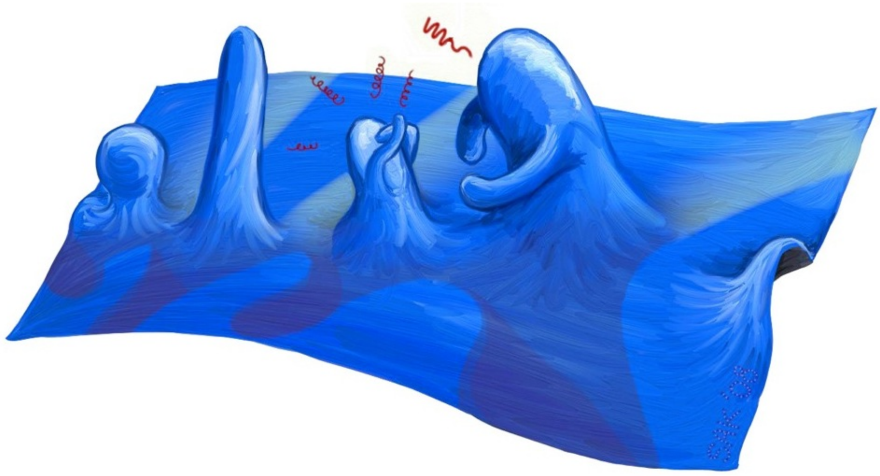
5.2 From Fields to Particles

Feynman once posed the following, slightly absurd, question: Suppose that all of human knowledge was going to be wiped out, and you could communicate only one piece of information about Nature to the fledgling civilization that starts again. What would this be? Obviously it can't be complicated mathematics—it has to be something simple that can be explained in plain language. Feynman's answer was that he would tell them about atoms. Or, more precisely, he would tell them that matter was discrete, made up of fundamental building blocks. (He would then hope that they would be smart enough to come up with quantum mechanics.)

However, in the last few decades, we have come to appreciate that particles and atoms aren't the most fundamental objects in the universe. It is *fields*. One of the key breakthroughs of the 20th century was the realization that

there is a different kind of field for each particle in Nature.

An electron field, several quark fields, and so on. What happens is that ripples in this field get tied up in knots by virtue of quantum mechanics. It is those ripples that we identify as particles.



Thus the study of particle physics is called **quantum field theory**.

5.3 From Symmetries to Forces

In addition to matter fields, there are force fields. For instance, the electromagnetic force is described in terms of electric and magnetic fields, $\vec{E}$ and $\vec{B}$. As we have seen in Lecture 4, these can be unified in a single four-vector, the vector potential,

$$A_\mu = \begin{pmatrix} A_0 \\ A_1 \\ A_2 \\ A_3 \end{pmatrix} .$$

Ripples of this field are light waves. Quantum mechanics tells us that these light waves are actually made out of particles,

$$\textit{photons}: \quad \gamma \,.$$

The two nuclear forces work in the same way. There are analogs of electric and magnetic fields and of the vector potential. Except now, each component of the vector is itself a *matrix* $(\,\cdot\,)$:

$$\mathsf{A}_\mu = \begin{pmatrix} \mathsf{A}_0 \\ \mathsf{A}_1 \\ \mathsf{A}_2 \\ \mathsf{A}_3 \end{pmatrix} = \begin{pmatrix} \mathsf{a}_0\,(\,\cdot\,) \\ \mathsf{a}_1\,(\,\cdot\,) \\ \mathsf{a}_2\,(\,\cdot\,) \\ \mathsf{a}_3\,(\,\cdot\,) \end{pmatrix} ,$$

where the amplitudes a_μ are real numbers.

To reproduce the properties of the weak force the components have to be matrices of type[1] $SU(2)$ (i.e. 2×2 unitary matrices with determinant one). How many such matrices are there? Three! (The Pauli matrices.) We therefore have *three* different vector potentials associated with the weak force, and hence three different force particles,

$$\textit{weak bosons}: \quad W^+ \ , \ W^- \ , \ Z \ .$$

What do the $SU(2)$ matrices of the weak force field act on? It has to be 2-component vectors made out of the matter particles of the Standard Model. Looking at our periodic table, an obvious guess is to pair the two quarks and the electron and the neutrino

$$Q = \begin{pmatrix} u \\ d \end{pmatrix} \qquad L = \begin{pmatrix} e \\ \nu_e \end{pmatrix} .$$

The $SU(2)$ matrices of the weak force indeed act on these vectors. (The decay of the down quark into the up quark—which happens when the neutron decays into the proton—can therefore be thought of as a rotation!)

Similarly, the strong force is associated with matrices of type $SU(3)$ (i.e. 3×3 unitary matrices with determinant one). How many such matrices are there? Eight! (The Gell-Mann matrices.) There are therefore *eight* force particles for the strong force,

$$\textit{gluons}: \quad g_1 \ , \ g_2 \ , \ \cdots \ , \ g_8 \ .$$

What do the $SU(3)$ matrices of the strong force field act on? The strong force only acts on the quarks, so we are looking for 3-component vectors with quarks as its entries. This time looking at the periodic table doesn't reveal an obvious pairing. Instead we have to postulate that each quark comes in three *flavors*:

red, *blue*, and *green* .

(This property of quarks is called the colour charge and the theory that describes it is called Quantum Chromo-Dynamics or QCD.) This triplet of colours can be arranged into a vector, e.g. for the up and down quarks

$$u = \begin{pmatrix} u_R \\ u_B \\ u_G \end{pmatrix} \qquad d = \begin{pmatrix} d_R \\ d_B \\ d_G \end{pmatrix} .$$

The $SU(3)$ matrices of the strong force act on these colour-vectors of quarks. They don't act on the leptons.

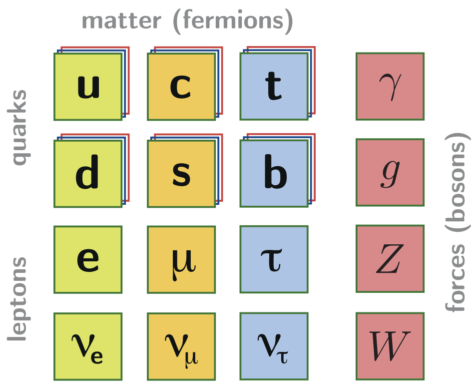
In this language, the electromagnetic force is associated with "matrices" of type $U(1)$ (complex numbers of unit norm). There is only one such matrix, and hence only one type of *photon*.

In summary, the forces of the Standard Model are determined by the matrices (or symmetry groups)

$$U(1) \times SU(2) \times SU(3) \ .$$

[1] For more on $SU(n)$ matrices see Example Sheet 4, Problem 12 of your *'Groups'* course.

For each "matrix field" we have a "force particle": a photon for the electric force, eight gluons for the strong force, and three vector bosons for the weak force. We add these to our periodic table:

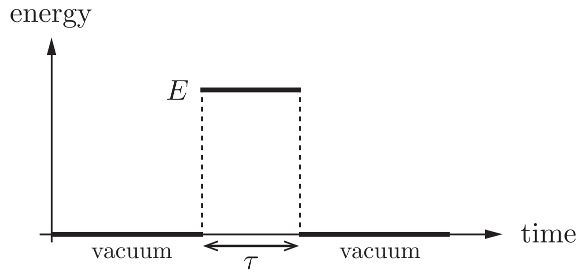


5.4 From Virtual Particles to Real Forces

We have learned enough quantum mechanics to understand the mechanism by which forces are communicated between the elementary particles. Let me describe this in a bit of detail.

5.4.1 Heisenberg's Uncertainty Principle

Recall our discussion of virtual particles and Hawking radiation in Lecture 3. There we have seen that an excited state with energy E is indistinguishable from the vacuum if its lifetime is less than $\hbar/(2E)$:

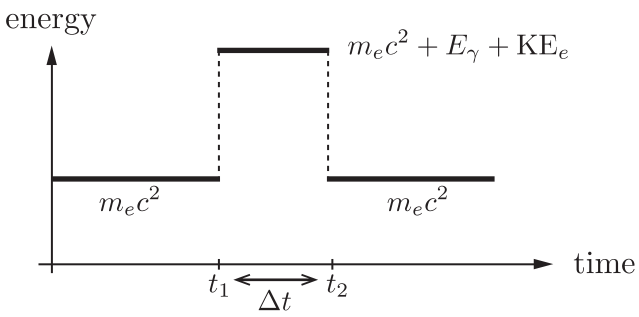


I will now show you that the same mechanism explains the forces of the Standard Model.

To warm up, consider a stationary electron with rest mass energy $m_e c^2$ (even if the electron is moving, you can always go to its rest frame). Now imagine that the electron spontaneously emits a photon of energy E_γ. To conserve momentum, the electron has to recoil with momentum equal and opposite to that of the photon. The combined electron+photon state has energy

$$E = m_e c^2 + \underbrace{E_\gamma + \text{KE}_e}_{\Delta E} ,$$

where KE_e is the kinetic energy of the recoiling electron. In classical physics, this process is clearly impossible since it violates the conservation of energy. However, in quantum mechanics, Heisenberg's uncertainty principle allows a temporary violation of energy conservation:

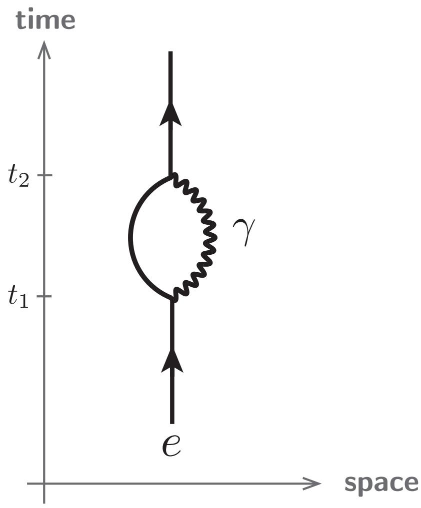


In other words, a free electron can emit a virtual photon as long as it reabsorbs the photon and returns to its original state within a time

$$\Delta t < \frac{\hbar}{2\Delta E} \ .$$

5.4.2 Feynman Diagrams

Feynman has introduced a neat way to picture such processes as spacetime graphs. We now call them *Feynman diagrams*. Our example of a stationary electron emitting a photon and recombining with it, would look as follows:

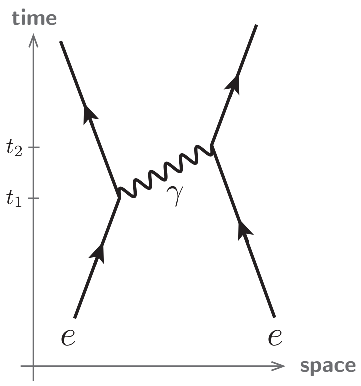


The vertical parts of the diagram correspond to the electron remaining at the same point. (If the electron was moving with a steady speed, these lines would be tilted at an angle to the vertical.) The loop in the middle corresponds to the virtual state.

Now let us consider two electrons approaching each other with constant speed.[2] At some moment t_1, one of the electrons emits a virtual photon. To conserve momentum, the electron has to recoil. Moreover, according to the Heisenberg uncertainty principle, the virtual photon must be reabsorbed within $\Delta t = t_2 - t_1 < \frac{\hbar}{2\Delta E}$. However, this time, instead of the first electron

[2]They may or may not be heading straight towards each other. On our schematic Feynman diagrams we only have one space dimension and so both possibilities would look similar.

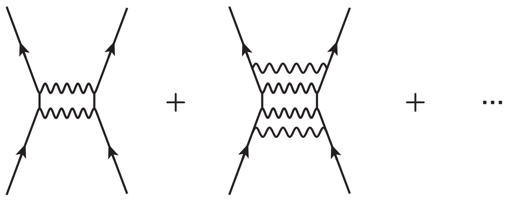
reabsorbing the photon, the second electron can absorb the photon:



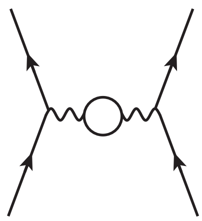
Provided this happens in such a way that the total energy of the two electrons before and after the intermediate state is the same, no quantum rule of physics is violated. The second electron recoils to the right because of the momentum it picks up from the photon. The Feynman diagram of the process looks as if the particles have repelled each other. So in quantum mechanics we can think of the electromagnetic force between two electrons as arising from the exchange of virtual photons. To understand the origin of attractive forces in this way requires more quantum mechanics than we want to get into right now. But that is details, the idea is the same.

5.4.3 QED

The photon exchange can happen more than once



Also, while in-flight, the photon may spontaneously turn into a electron-positron pair and back



Quantum mechanics associates a probability amplitude with each of these little diagrams.[3] However, it turns out that in Quantum Electro-Dynamics (QED)

the more complicated the diagram, the "less likely" it is.

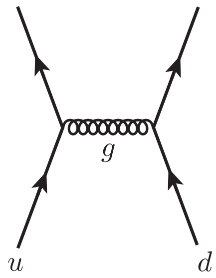
So to a good approximation we get away with only the single photon exchange diagram. In a few years, you will be able to calculate the probability amplitude for that diagram and show that it is equivalent to the Coulomb force.

[3] Compare this to Lecture 1, where we saw that a particle has a probability amplitude for every possible path. Here, there is a probability amplitude for every possible diagram that starts and ends with two electrons. Add up all the amplitudes and square the result to get the probability for the process to happen.

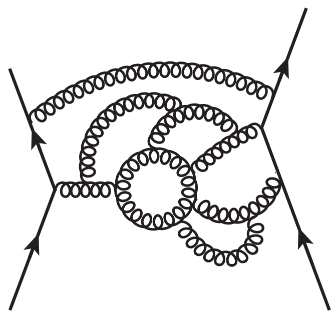
5.4.4 QCD and Confinement

The strong force acts only on the quarks. It's like electromagnetism, but with matrices for the fields. Going from numbers to matrices shouldn't make too much difference. Right?! In fact, it makes the problem completely intractable!!

We can again draw Feynman diagrams to represent the force between two quarks as the exchange of gluons



But there are crucial differences between numbers (photons) and matrices (gluons): first, there is only one photon, but there are eight gluons. Second, being matrices these gluons can interact with each other (while photons don't). So we can have very complicated diagrams, like this one



And, most importantly, in Quantum Chromo-Dynamics (QCD)

the more complicated the diagram, the "more likely" it is!

This means that the Feynman diagram way of calculating things in QCD becomes completely useless. Unlike QED, in QCD we can't just draw a few simple diagrams and get a good approximation to the exact answer. The difference between QED and QCD is like the difference between[4]

$$\tfrac{1}{2} + (\tfrac{1}{2})^2 + (\tfrac{1}{2})^3 + \cdots \ = \ 1$$

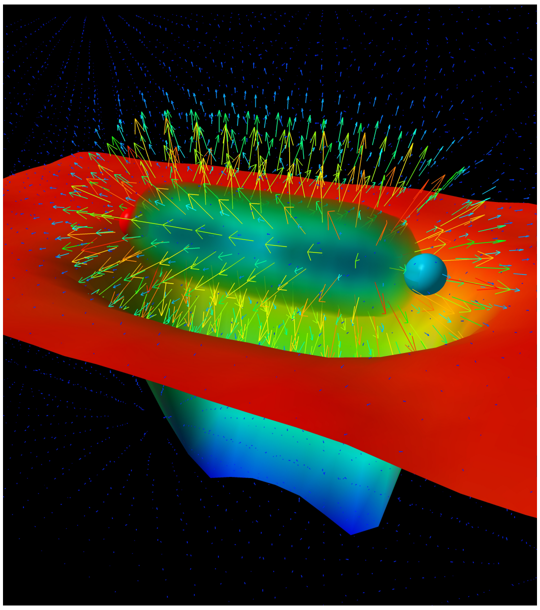
and

$$2 + 2^2 + 2^3 + \cdots \ = \ \infty \ .$$

Nobody has ever seen a free quark. Instead, quarks always come in pairs or triplets and resist being separated. In contrast, when two electrically-charged particles separate, the electric fields between them diminish quickly, allowing (for example) electrons to become unbound from atomic nuclei. However, as two quarks separate, the interacting gluons form narrow tubes (or strings), which tend to bring the quarks together as though they were some kind of rubber band.

[4]In the first case (QED), keeping only the first few terms in the sum (i.e. the simplest diagrams) gives an accurate approximation to the true answer. In QCD, you need to understand how to deal will all diagrams at once.

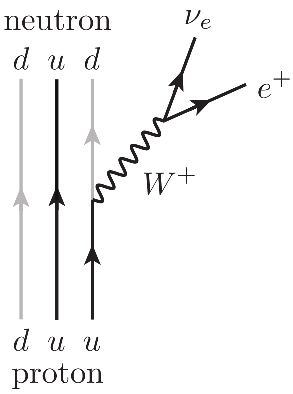
Computer simulations indeed confirm this picture:[5]



This is quite different in behaviour from electrical charge. The force between quarks does not diminish as they are separated. In fact, it increases with distance. Because of this, it would take an infinite amount of energy to separate two quarks; they are forever bound into protons and neutrons.[6] This phenomenon is called *confinement*. To prove confinement from the equations of QCD is one of the most important open problems in theoretical physics and mathematics. In fact, the Clay Institute will pay you $1 million[7] if you solve it!

5.4.5 Why the Sun Shines*

I mentioned that the weak force is responsible for the decay of free neutrons into protons. The reverse can also happen (as long as we supply some energy): the weak force can trigger the conversion of a proton into a neutron. Here is the Feynman diagram associated with this process:



[5] http://www.physics.adelaide.edu.au/theory/staff/leinweber/VisualQCD/Nobel/

[6] This also implies that when two quarks fly apart, as happens in particle accelerators, at some point it is energetically favourable for a new quark-antiquark pair to spontaneously appear, rather than to allow the tube to extend further. As a result, when quarks are produced in accelerators, instead of seeing the individual quarks in detectors, scientists see "jets" of quark pairs (mesons) and triplets (baryons).

[7] http://www.claymath.org/millennium/Yang-Mills_Theory/

An up quark turns into a down quark, while emitting a W^+ particle. The W^+ then decays into a positron (the anti-particle of the electron) and a neutrino. This process is of fundamental importance for life on Earth. Without it, the Sun wouldn't be shining.

As you know, the Sun shines because of nuclear fusion—light nuclei combine into heavier nuclei, while releasing energy in the process. The simplest fusion process is two protons combining into a helium nucleus. However, there is a puzzle. Protons are positively charged, so they don't like being pushed together. In fact, one can show that the temperature inside the Sun isn't high enough, and hence the protons aren't moving fast enough, to overcome the electric repulsion. The weak force comes to the rescue. Because of the W particle, one of the protons in the collision can convert into a neutron. The newly formed neutron and the remaining proton can get very close, because the neutrons doesn't carry electric charge. Freed from the electromagnetic repulsion they can fuse together (as a result of the strong force) to form deuteron. This quickly leads to helium formation, releasing life-giving energy in the process.

5.5 The Origin of Mass

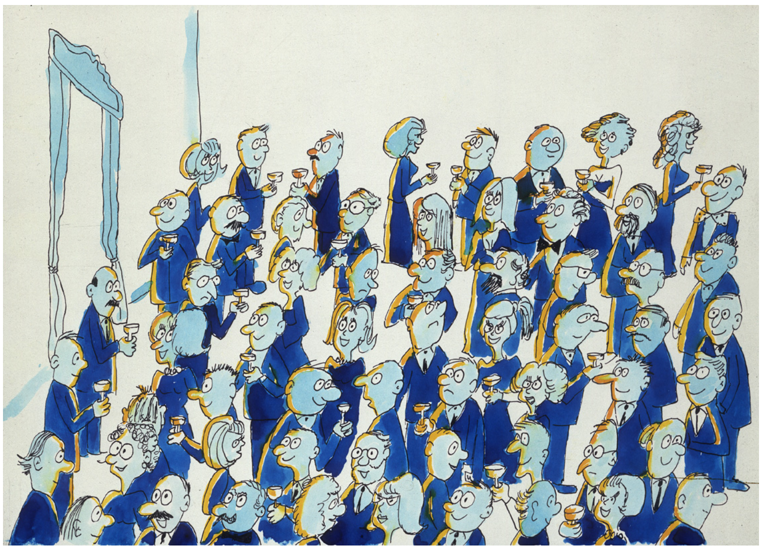
In many aspects, the carriers of the weak force, W and Z, are like the photon of the electromagnetic force. However, there is one important difference: the W and Z particles are very massive, while the photon is massless. The reason the weak bosons are massive is the same reason all the other particles in the Standard Model are massive: the Higgs field (h).

5.5.1 An Analogy

The Higgs mechanism isn't easy to explain without some additional training in physics and mathematics. I will start with a popular analogy. In §5.5.2, I will give a few more details.

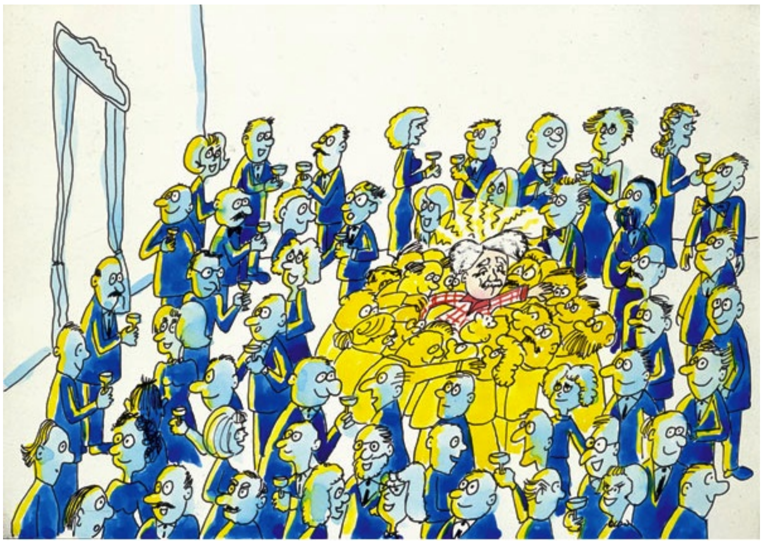
The Higgs Mechanism

Imagine that a room full of physicists, all talking to their nearest neighbours



This corresponds to space filled with the Higgs field.

When a famous scientist enters the room, the physicists cluster around him, slowing the scientist's progress:



In much the same way, the Higgs field becomes locally distorted whenever a particle moves through it. The distortion—the clustering of the field around the particle—generates the particle's mass.[8]

Higgs Boson

Now consider a rumour passing through our room full of uniformly spread physicists

Those near the door hear of it first and cluster together to get the details, then they turn and move closer to their next neighbours who want to know about it too. A wave of clustering passes through the room. Since the information is carried by clusters of people, and since it was

[8]The idea comes directly from the physics of solids. Instead of a field spread throughout all space, a solid contains a lattice of positively charged crystal atoms. When an electron moves through the lattice the atoms are attracted to it, causing the electron's effective mass to be as much as 40 times bigger than the mass of a free electron. The postulated Higgs field in the vacuum is a sort of hypothetical lattice which fills our universe.

clustering which gave extra mass to the famous scientist, then the rumour-carrying clusters also have mass. The Higgs particle is just such a clustering in the Higgs field.[9]

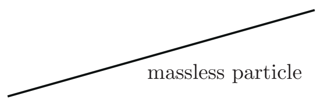
5.5.2 Beyond Cartoons*

For those who want more than just cartoons, this section has a few more details.

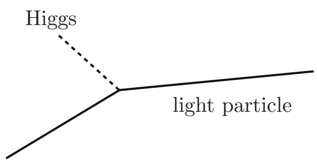
Higgs Mechanism

In the early universe, the value of the Higgs field was zero everywhere. All the particles of the Standard Model were therefore *massless* and travelled at the speed of light. As the universe expanded it cooled. At some critical point it became cold enough for the Higgs field to condense. Since then, the Higgs field has a constant value throughout space. The particles that interact with the Higgs field became massive.

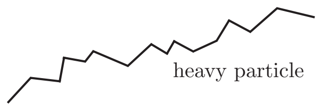
Some particles do not interact with the Higgs. This is the case for photons, gluons and gravitons. The most likely path of these particles between two points is therefore a straight line:



A light particle (e.g. the electron) has only a small probability to interact with the Higgs field. This interaction deflects the particle:



A heavy particle (e.g. the top quark or the W and Z bosons) is more likely to interact with the Higgs field and therefore experiences more deflections:



Particles get their masses because interactions with the Higgs field force them to zig-zag their way through space. All particles of the Standard Model get their masses this way. Different particles get stuck in the Higgs condensate by different amounts. But why? What gives rise to the vast difference in masses of the particles? We don't know.

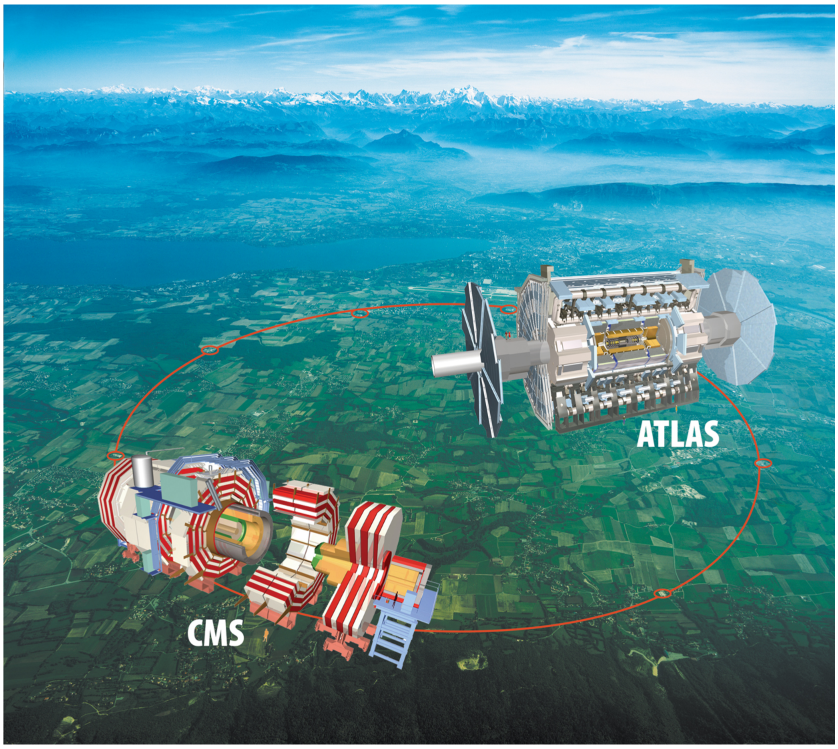
Higgs Boson

There is a particle for every field. The Higgs field is no different, so we expect a particle to correspond to an excitation of the Higgs field. Unsurprisingly, this is called the Higgs particle, or *Higgs boson*. It will be much easier to believe that the Higgs field exists, and that the mechanism for giving other particles mass is true, if we actually see the Higgs particle itself. The search for the Higgs boson has occupied experimental particle physicists for the past few decades. Last year, they found it!

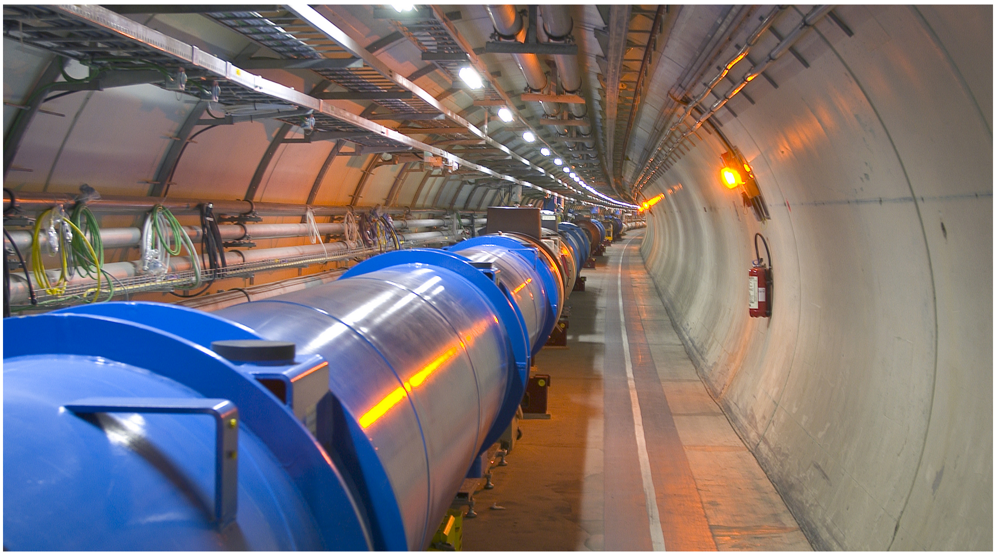
[9] Again, there are analogies in the physics of solids. A crystal lattice can carry waves of clustering without needing an electron to move and attract the atoms. These waves can behave as if they are particles. They are called phonons.

5.5.3 Discovery of the Higgs

The LHC is a magnificent experiment.

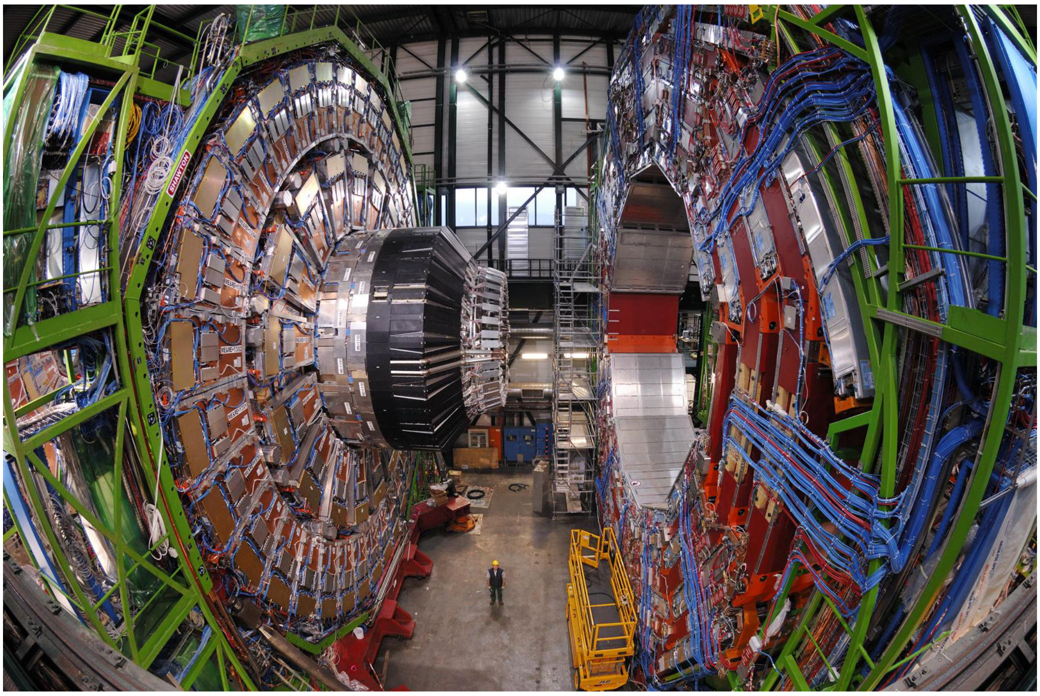


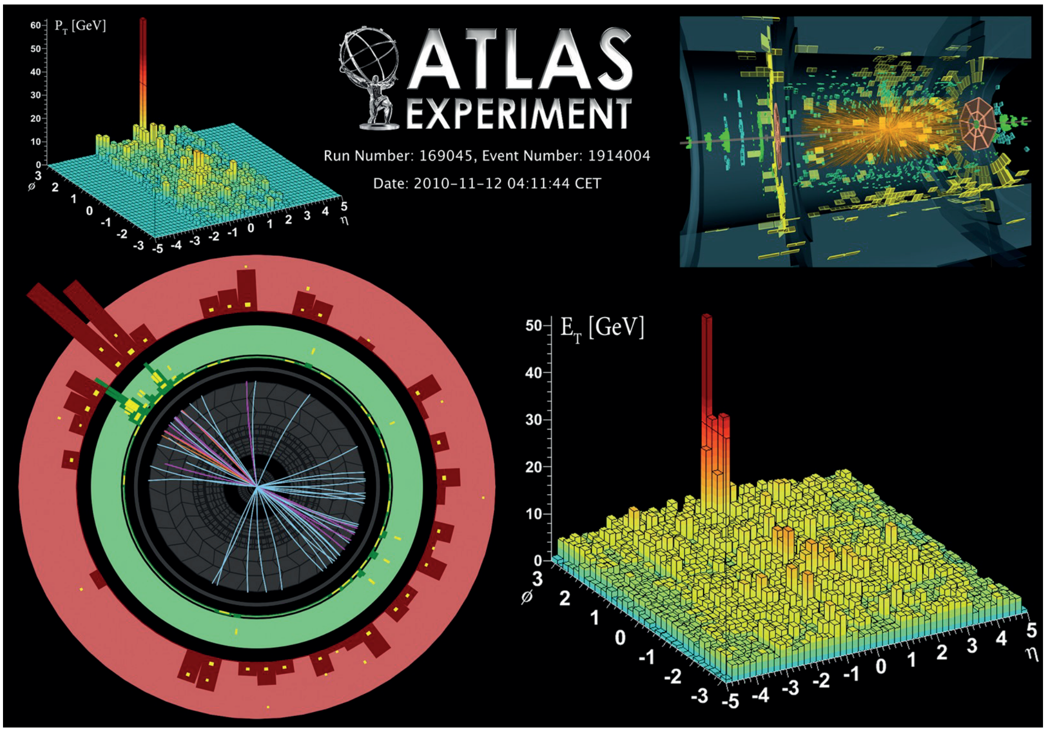
It accelerates protons and anti-protons around an underground ring seventeen miles in circumference.



A number of accelerating structures boost the energy of the particles along the way until they reach speeds close to the speed of light and energies close to 14 TeV (centre of mass)—15,000 times the energy in the mass of a proton. The beams travel in opposite directions in separate beam pipes—two tubes kept at ultra-high vacuum. They are guided around the accelerator ring by a strong magnetic field, achieved using superconducting electromagnets. These are built from coils of special electric cable that operates in a superconducting state, efficiently conducting electricity without resistance or loss of energy. This requires cooling the magnets to about $-271°C$ or two degrees above absolute zero.

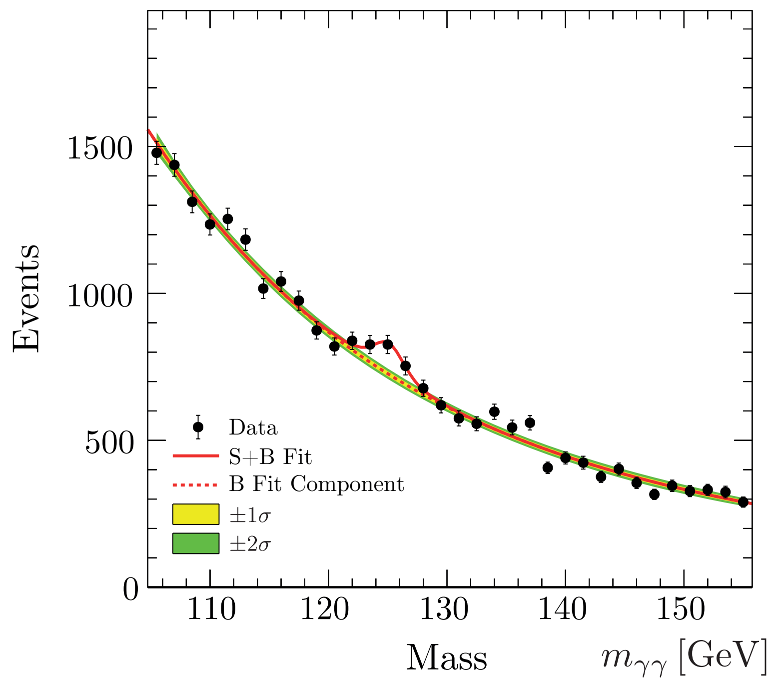
The particles collide at two locations in the ring, where two huge detectors—called ATLAS and CMS—record the debris



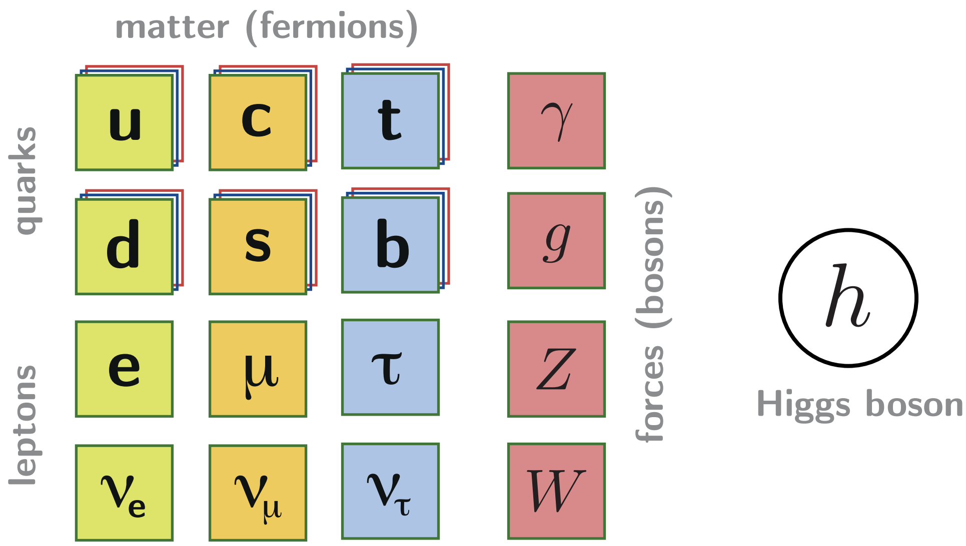


About 600 million collisions occur per second! This results in 10 Petabytes of data per year (= 20 km high stack of CDs!).

On July 4th, ATLAS and CMS announced the discovery of a “Higgs-like” particle with a mass 133 times the mass of the proton (126 GeV):



They are still doing more checks, but we all know it is the Higgs. This completes the periodic table of the Standard Model:



5.6 Beyond the Standard Model*

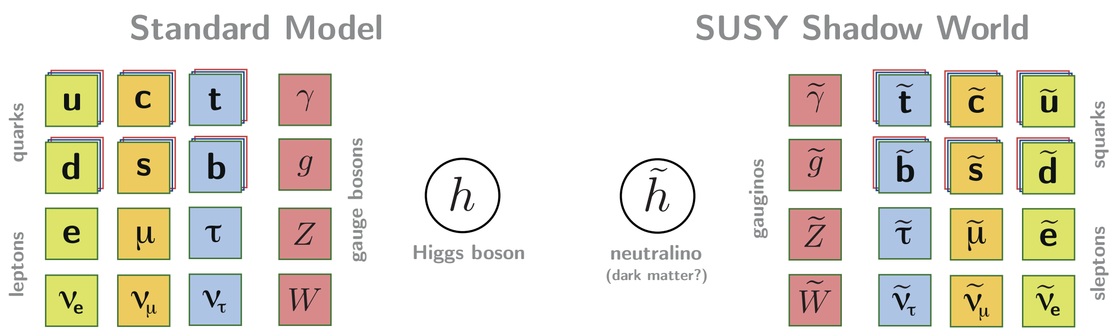
5.6.1 Dark Matter

The LHC has detected the Higgs. Is particle physics done?

There are reasons to believe that the answer is no. For example, cosmological observations suggest that the particles of the Standard Model make up less than 15% of the total matter in the universe. The majority is in some form of *dark matter* (see Lecture 7). But we have no idea what the dark matter is. A satisfactory theory of particle physics should explain this.

5.6.2 Supersymmetry

A popular extension of the Standard Model is Supersymmetry (SUSY). For reasons that I don't have time to explain, SUSY overcomes a number of theoretical shortcomings of the Standard Model.[10] According to SUSY, every particle of the Standard Model has a hidden partner. Essentially, SUSY proposes a doubling of our periodic table:



With some luck the LHC could discover a SUSY shadow world. Moreover, the lightest supersymmetric partner of the Standard Model would be stable and could be the dark matter! So things could fall into place. Or maybe they won't—the next few years will tell. Stay tuned ...

[10] Essentially, without SUSY it is hard to understand why the Higgs particle wouldn't be much heavier than it is found to be.

6 General Relativity

Today, we are going to talk about **gravity** as described by Einstein's **general theory of relativity**.

We start with a simple question:

- *Why do objects with different masses fall at the same rate?*

  We think we know the answer: an object of mass m is attracted to a second object of mass M by the gravitational force $F = GmM/r^2$. To find the acceleration of m, we apply Newton's law, $F = ma$,

$$\cancel{m} a = G \frac{\cancel{m} M}{r^2} \ . \tag{6.0.1}$$

  We notice that the mass m cancels, so the acceleration doesn't depend on it. Heavier objects feel a larger gravitational force, but they also experience more inertia. The two effects exactly cancel. Einstein wouldn't be Einstein if he didn't question this seemingly obvious fact. He noticed that the meaning of 'mass' on the left-hand side and the right-hand side of (6.0.1) is very different. We should really distinguish between the two masses by calling them something different:

$$m_\mathrm{i} a = G \frac{m_\mathrm{g} M}{r^2} \ . \tag{6.0.2}$$

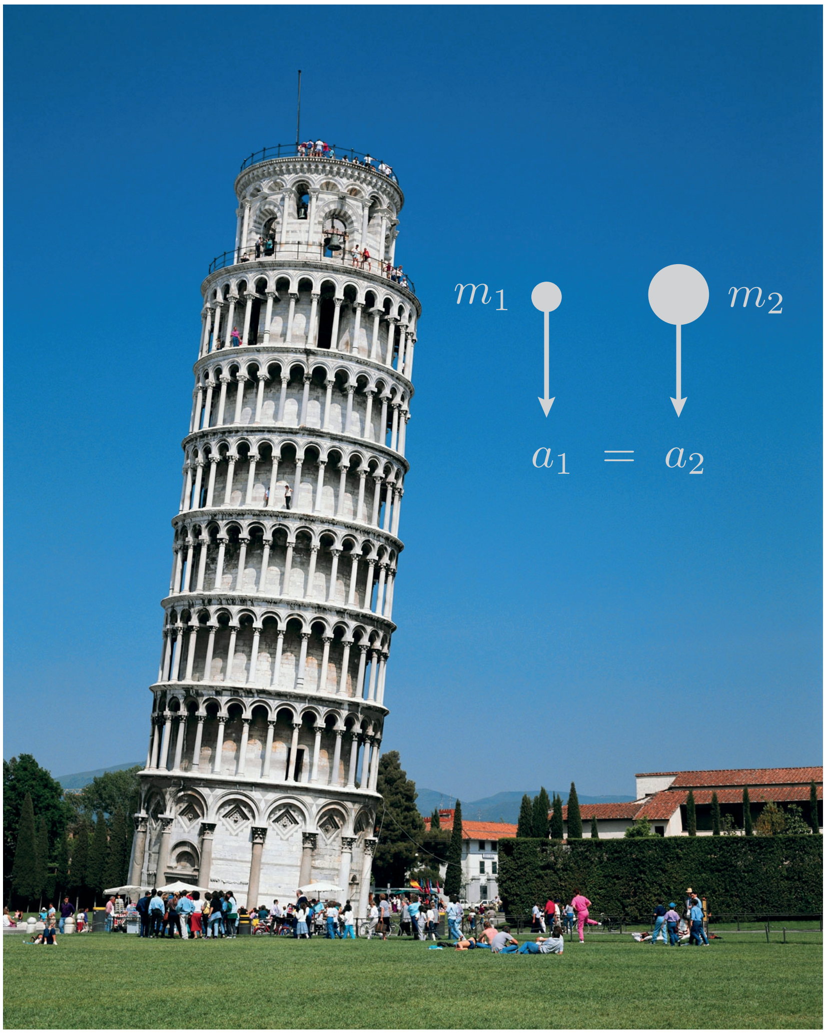
  The *inertial mass* is labelled m_i, while the *gravitational mass* is denoted m_g. You should think of m_g more like *charge* in Coulomb's law, $F \propto qQ/r^2$, and you wouldn't be tempted to cancel m_i and q in

$$m_\mathrm{i} a = \frac{1}{4\pi\epsilon_0} \frac{qQ}{r^2} \ . \tag{6.0.3}$$

This makes it clear that we should really think of $m_{\rm i}$ and $m_{\rm g}$ as distinct entities. The gravitational mass $m_{\rm g}$ is a source for the gravitational field (just like the charge q is a source for an electric field), while the inertial mass $m_{\rm i}$ characterizes the dynamical response to any forces. It is a non-trivial result, that experimentally one finds

$$\frac{m_{\rm i}}{m_{\rm g}} = 1 \pm 10^{-13} \ . \tag{6.0.4}$$

This equality of inertial and gravitational mass is responsible for the well-known fact that objects with different masses fall at the same rate under gravity.



But, is there a deeper reason *why* the gravitational force is proportional to the inertial mass?

6.1 The Happiest Thought

> *"I was sitting in a chair in the patent office in Bern when all of a sudden a thought occurred to me: "If a person falls freely he will not feel his own weight." I was startled. This simple thought made a deep impression on me. It impelled me toward a theory of gravitation."*
>
> Albert Einstein

6.1.1 Fictitious Forces

There are two other forces which are also proportional to the inertial mass. Consider a particle moving with velocity $\vec{v} \equiv \dot{\vec{r}}$ on a disc that is rotating with angular velocity $\vec{\omega} \equiv \dot{\vec{\theta}}$. From the point of view of an observer on the rotating disc, the particle experiences the following forces:

$$\text{Centrifugal force}: \qquad \vec{F} = -m_{\rm i}\vec{\omega} \times (\vec{\omega} \times \vec{r}) \ . \tag{6.1.5}$$

$$\text{Coriolis force}: \qquad \vec{F} = -2m_{\rm i}\vec{\omega} \times \vec{v} \ . \tag{6.1.6}$$

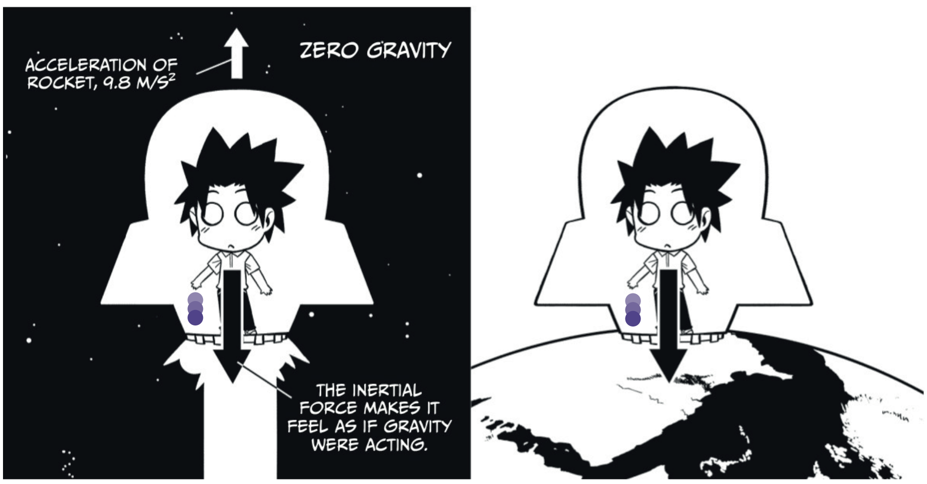
The path of the particle appears to be curved in response to those forces. However, for both of these forces we understand very well why they are proportional to the inertial mass $m_{\rm i}$. It's because these are "fictitious forces", arising in a non-inertial frame (rotating with frequency $\vec{\omega}$). To an outside observer in an inertial frame there is no force on the particle and it travels in a straight line. Could gravity also be a fictitious force, arising only because we are in a non-inertial frame?

6.1.2 Equivalence Principle

The beauty of Einstein's theories is that they start from one or two simple principles, and derive by simple logical reasoning dramatic consequences. The basic principle underlying general relativity is the *equivalence principle.* This principle states that

> *locally, it is impossible to distinguish between a gravitational field and acceleration.*

In other words, there is no experiment you could perform that would distinguish a rocket accelerating at $g = 9.8$ m/s^2 in empty space, from a stationary rocket in the gravitational field of the Earth:



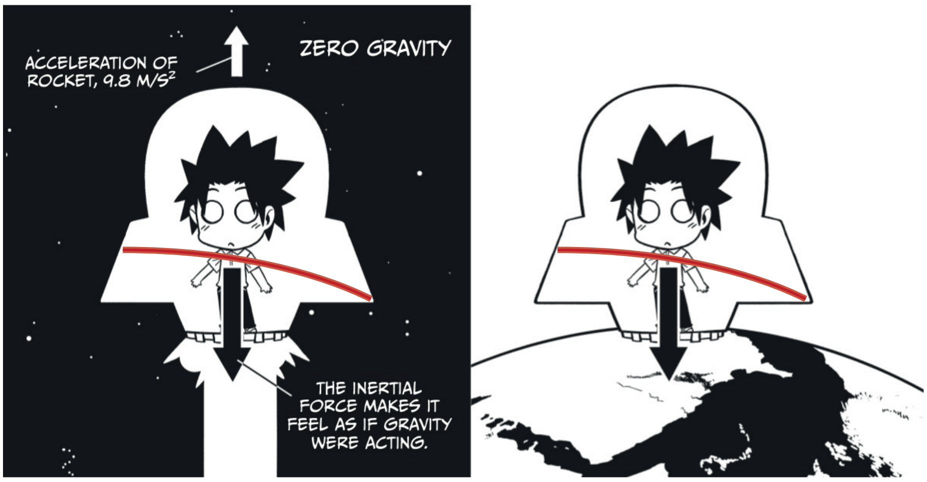
There are some obvious difficulties with trying to identify gravitation with acceleration. In Cambridge, it feels as if we're accelerating upwards. The people in New Zealand feel as if they're accelerating in the opposite direction. Why aren't we getting further apart?! Moreover, if two balls are dropped from different heights, the lower ball will accelerate faster in gravity due to the $1/r^2$ law. This seems to give a way to experimentally determine the difference between acceleration and gravitation.

However, this is not in conflict with the equivalence principle because of the careful choice of the word *local*. Here, "local" means in a small enough region of spacetime, where "small" is relative to the scale of variation of the gravitational field. The equivalence principle simply states that at every point in spacetime, we can remove or create the gravitational field by going to an accelerating frame.[1] Patching together different frames at each point in spacetime gives rise to the curvature of spacetime. But this is getting a bit ahead of ourselves ...

... instead, let's follow the equivalence principle to derive a few interesting experimental predictions:

- *Bending of light*

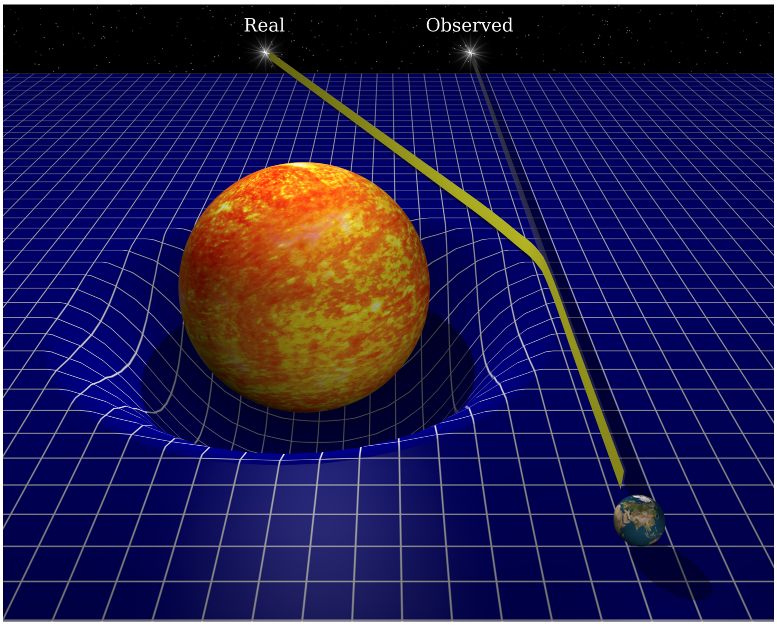
  An immediate consequence of the equivalence principle is that light bends in a gravitational field:



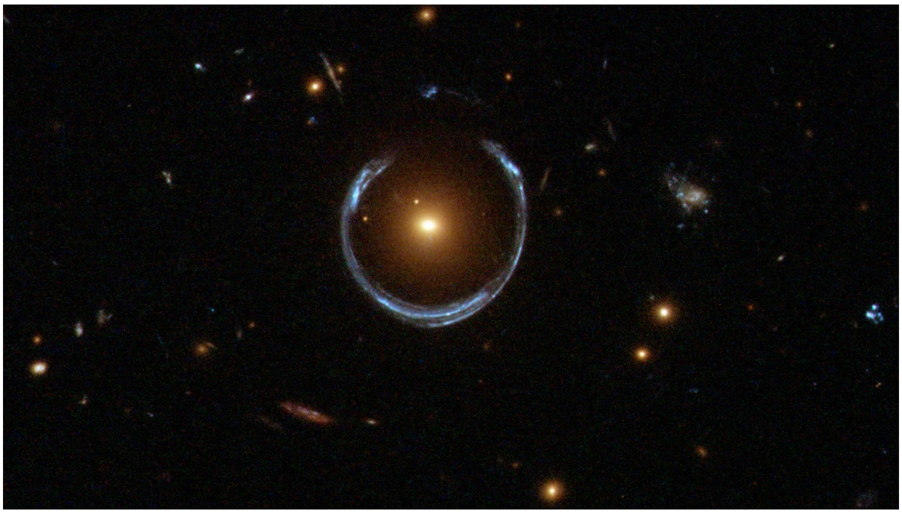
  Imagine shinning a beam of light from one side of the rocket to the other. During this time, the accelerating rocket moves forward, so the light reaches the second wall slightly below the height that it left the first wall (see the figure, if the words aren't clear enough). By the equivalence principle, the same has to be true for the stationary rocket on the surface of the Earth. Light has to bend in a gravitational field. And it does!

[1] Recall from Lecture 4, that in special relativity we can remove or create a magnetic field by boosting to another inertial frame. Gravity is similar and arises from going to a non-inertial frame.

General relativity's prediction for the bending of light was first confirmed by Eddington, who observed the deflection of starlight passing close to the Sun during a solar eclipse:[2]



Since then we have observed many examples of the gravitational bending of light, such as gravitational lensing of entire galaxies by other galaxies along the line-of-sight:



- *Gravitational redshift*

Consider what happens when we shine the light along the rocket, rather than across. Let the rocket have height h. It starts from rest, and moves with constant acceleration g. Light emitted from the bottom of the rocket at $t = 0$ is received at the top at $t = h/c$. By this time, the rocket is travelling at speed

$$v = gt = gh/c \, . \tag{6.1.7}$$

[2]To measure this effect one has to compare the relative positions of stars on the night sky (i.e. in the absence of the Sun) to the shifted positions during the eclipse (i.e. in the presence of the Sun).

Due to the Doppler effect, the received frequency f' is smaller than the emitted frequency f, by an amount

$$f' = f\left(1 - \frac{v}{c}\right) = f\left(1 - \frac{gh}{c^2}\right) . \tag{6.1.8}$$

Since $f = c/\lambda$, the observed light is more red than the emitted light, $\lambda' > \lambda$. We say that the light is *redshifted*.

By the equivalence principle, the same effect must be observed for light "climbing out" of a gravitational potential well.[3] The gravitational field Φ gives rise to the acceleration $\vec{a} = -\vec{\nabla}\Phi$. We fix the gravitational potential at the bottom of the rocket to be $\Phi \equiv 0$. For constant acceleration $a = -g$, we then find that the gravitational potential at the top of the rocket is

$$\Phi = gh . \tag{6.1.9}$$

Comparing this to (6.1.8) gives the formula for gravitational redshift

$$f' = f\left(1 - \frac{\Phi}{c^2}\right) . \tag{6.1.10}$$

From this we can derive our third, and most dramatic, prediction.

- *Gravitational time dilation*

  The frequency of light is the inverse of the period of oscillations of the electromagnetic wave, $f = 1/\Delta t$. Inverting (6.1.10), we therefore get a relationship between the period at emission (Δt) and the period at reception ($\Delta t'$)

$$(\Delta t')^2 = (\Delta t)^2\left(1 - \frac{\Phi}{c^2}\right)^{-2} \approx (\Delta t)^2\left(1 + \frac{2\Phi}{c^2}\right) . \tag{6.1.11}$$

  In the second equality we have assumed[4] $\Phi \ll c^2$, which holds for a weak gravitational field. This result (6.1.11) holds not just for light. In general, time goes slower at the bottom of the rocket or, by the equivalence principle, closer to the Earth, $\Delta t < \Delta t'$. If you stay at the bottom of the rocket, or closer to the Earth, you will live longer than your friend at the top of the rocket, or further from the Earth.[5] If you choose to spend your life on the ground floor, and never climb the stairs, you will live longer by a couple of microseconds.[6]

6.1.3 Black Holes

We have shown that time slows down in a gravitational field. A consequence of this is the existence of black holes, regions of spacetime in which gravity is so strong that time stands still and nothing, not even light, can escape.

Although (6.1.11) was derived under the assumption of constant acceleration, it actually holds for an arbitrary gravitational potential $\Phi(t, \vec{x})$. For a spherical object of mass M, we have $\Phi = -GM/r$ and hence

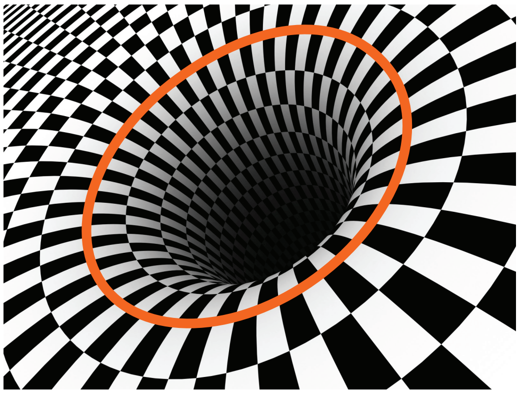
$$(\Delta t')^2 = (\Delta t)^2\left(1 - \frac{2GM}{rc^2}\right) . \tag{6.1.12}$$

[3] And it is! The effect was first measured in the Harvard physics department in 1959.
[4] This corresponds to the potential energy of a test particle, $m\Phi$, being smaller than its rest mass energy, mc^2.
[5] This "gravitational twin paradox" has been confirmed with atomic clocks on planes.
[6] Of course, this ageing effect is easily offset by the fact that you get less exercise.

We have set $\Phi \equiv 0$ at infinity, so Δt is the time interval measured by an observer far from any gravitational sources. Observers close to gravitational sources will measure $\Delta t' < \Delta t$. Something special happens at the so-called *Schwarzschild radius*

$$r_S = \frac{2GM}{c^2} \,, \tag{6.1.13}$$

where the factor in (6.1.12) vanishes. Objects whose size is smaller than r_S are *black holes*:



Notice that the Schwarzschild radius of an object depends on its mass. For the Earth, this distance is 1 cm. For the Sun, it is about 1 km. This is how much the mass of the Earth or the Sun would have to be compressed before they would form black holes.

Imagine you and your friend meet at infinity (or any other place sufficiently far from gravitational sources). Your friend then decides to jump into a black hole. She sends back signals at regular intervals $\Delta t'$ (on her clock), which you measure at intervals Δt (on your clock). As she gets closer to r_S her signals arrive less and less frequently, $\Delta t \gg \Delta t'$. In fact, at $r = r_S$, we get $\Delta t \to \infty$ for fixed $\Delta t'$. You will never see her crossing the *event horizon* at r_S. Her image will be infinitely redshifted and she will appear frozen in time (and space). What happens to her at r_S? Nothing special. For a sufficiently large black hole[7] she won't feel anything when she crosses r_S. She might not even know that it happened. But she will feel it soon enough. As she approaches the centre of the black hole the gravitational force on her feet will become much larger than the force on her head. (I forgot to tell you that she decided to jump in feet first.) She will get stretched more and more. The technical term for this is *spaghettification*. It is not a happy ending.

6.2 Gravity as Geometry

We have seen how gravity can affect how we measure time. But in relativity time and space are treated democratically, so we should expect gravity to also affect how we measure distances.

6.2.1 Spacetime

We can write down a metric, akin to the Minkowski metric that you've met in special relativity. But this time, the metric will depend on where in space we are and what matter is around.

[7]There is a black hole at the centre of our galaxy. Its mass is 4.1 million solar masses (or 8.2×10^{36} kg) and its radius is 6.25 light-hours. If the mass was uniformly distributed inside the event horizon, the density would be that of water.

We've already seen what the time component should look like:

$$\mathrm{d}s^2 = \left(1 + \frac{2\Phi(t,\vec{x})}{c^2}\right) c^2 \mathrm{d}t^2 + \cdots \tag{6.2.14}$$

Since time and space mix under Lorentz transformations, the space part should also vary with space. This means that measuring sticks contract and expand at different points in space. In general, the metric of spacetime is a 4×4 matrix, whose components can be any function of t and $\vec{x}$

$$\mathrm{d}s^2 = g_{\mu\nu}(t,\vec{x})\,\mathrm{d}X^\mu \mathrm{d}X^\nu \ . \tag{6.2.15}$$

When the metric varies in space and time, the spacetime is *curved*. In general relativity, the dynamics of gravity is encoded in the curvature of spacetime.

6.2.2 Matter Tells Space How To Curve ...

Einstein derived a famous equation that tells you how to calculate the metric for a given distribution of matter.[8] Schematically, Einstein's equation looks like this:

$$\textsf{Curvature} = \textsf{Matter} \tag{6.2.16}$$

where Curvature is a complicated function of the metric $g_{\mu\nu}$. The left-hand side takes the metric and computes a local measure of the curvature of spacetime. The right-hand side expresses how all the matter around sources the bending of spacetime. In short:

"matter tells spacetime how to curve"

The more technical form of Einstein's field equation is

$$G_{\mu\nu} = 8\pi G T_{\mu\nu} \ , \tag{6.2.17}$$

where $G_{\mu\nu}$ is the Einstein tensor and $T_{\mu\nu}$ is the stress-energy tensor. This doesn't look so bad, until I tell you what $G_{\mu\nu}$ is in terms of the metric. First, we write

$$G_{\mu\nu} \equiv R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} \ , \tag{6.2.18}$$

where $R_{\mu\nu}$ is the Ricci tensor and $R = g^{\mu\nu} R_{\mu\nu}$ is the Ricci scalar. The Ricci tensor is defined as

$$R_{\mu\nu} \equiv \partial_\lambda \Gamma^\lambda_{\mu\nu} - \partial_\nu \Gamma^\lambda_{\mu\lambda} + \Gamma^\lambda_{\lambda\sigma}\Gamma^\sigma_{\mu\nu} - \Gamma^\lambda_{\nu\sigma}\Gamma^\sigma_{\mu\lambda} \ , \tag{6.2.19}$$

where the Christoffel symbols $\Gamma^\lambda_{\mu\nu}$ are a function of the metric

$$\Gamma^\lambda_{\mu\nu} \equiv \frac{1}{2} g^{\lambda\sigma} \left(\partial_\mu g_{\nu\sigma} + \partial_\nu g_{\sigma\mu} - \partial_\sigma g_{\mu\nu}\right) \ . \tag{6.2.20}$$

You can imagine what kind of mess you get when you plug (6.2.20) into (6.2.19) into (6.2.18), and finally all into (6.2.17).

It shouldn't come as a surprise that the Einstein equation (6.2.17)—coupled non-linear PDEs for the 10 independent components of the metric—is very hard to solve. Only in a few special cases are explicit solutions known:

[8] This is analogous to the Maxwell equations which tell us how to calculate the electric and magnetic fields sourced by a given distribution of charges and currents.

- The *Schwarzschild solution* describes the static and spherically symmetry spacetime in the vacuum around an object of mass M

$$\mathrm{d}s^2 = \left(1 - \frac{2GM}{rc^2}\right) c^2\mathrm{d}t^2 - \left(1 - \frac{2GM}{rc^2}\right)^{-1} \mathrm{d}r^2 - r^2(\mathrm{d}\theta^2 + \sin^2\theta\mathrm{d}\phi^2) \ . \tag{6.2.21}$$

  This is the spacetime around a black hole, or outside a spherical star.

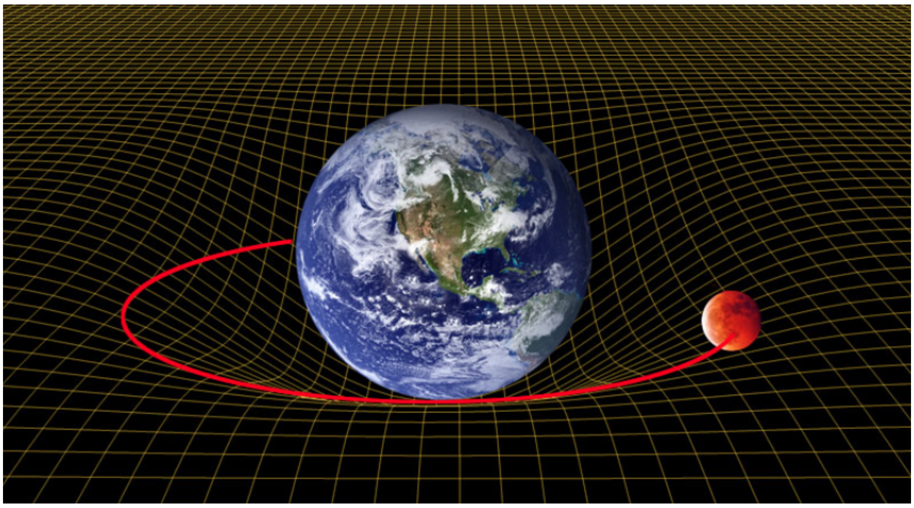
- Allowing space to expand, but assuming it is curved the same way at every point and in all directions, leads to the *Friedmann-Robertson-Walker solution*

$$\mathrm{d}s^2 = c^2\mathrm{d}t^2 - a^2(t)\left[\mathrm{d}r^2 + r^2(\mathrm{d}\theta^2 + \sin^2\theta\mathrm{d}\phi^2)\right] \ , \tag{6.2.22}$$

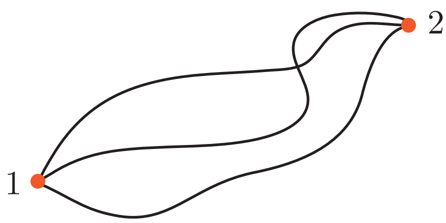
  where $a(t)$ is a monotonically increasing function of time associated with the expansion of space. This metric describes the *Big Bang*. More about that in Lecture 7.

6.2.3 ... Space Tells Matter How To Move.

Suppose we have a curved spacetime, specified by a metric. How do particles move? For example, how does the Moon move in the curved space created by the Earth?



It's simplest to describe this in terms of the principle of least action (see Lecture 1). We consider all possible paths between two points:



The (relativistic) action for the particle is simply given by

$$S = -mc \int_1^2 \mathrm{d}s \ . \tag{6.2.23}$$

This doesn't quite yet look like the usual integral over kinetic minus potential energy, but we will get there. Paths which minimise this action are called *geodesics*.

Consider a particle moving in a metric with gravitational time dilation

$$\mathrm{d}s^2 = \left(1 + \frac{2\Phi(\vec{x})}{c^2}\right) c^2\mathrm{d}t^2 - \mathrm{d}\vec{x}^2 \,. \tag{6.2.24}$$

Plugging this into the action (6.2.23), we get

$$S = -mc^2 \int_{t_1}^{t_2} \mathrm{d}t \sqrt{(1 + 2\Phi/c^2) - \dot{\vec{x}}^2/c^2} \,, \tag{6.2.25}$$

where we have pulled out $\mathrm{d}t$ from under the square root and written $\dot{\vec{x}}$ for $d\vec{x}/dt$. Expanding the square root for small velocities, $\dot{\vec{x}}^2 \ll c^2$, and weak gravity, $\Phi \ll c^2$, we find[9]

$$S \approx \int_{t_1}^{t_2} \mathrm{d}t \left[\frac{m}{2}\dot{\vec{x}}^2 - m\Phi(\vec{x}) + \cdots\right] \,. \tag{6.2.26}$$

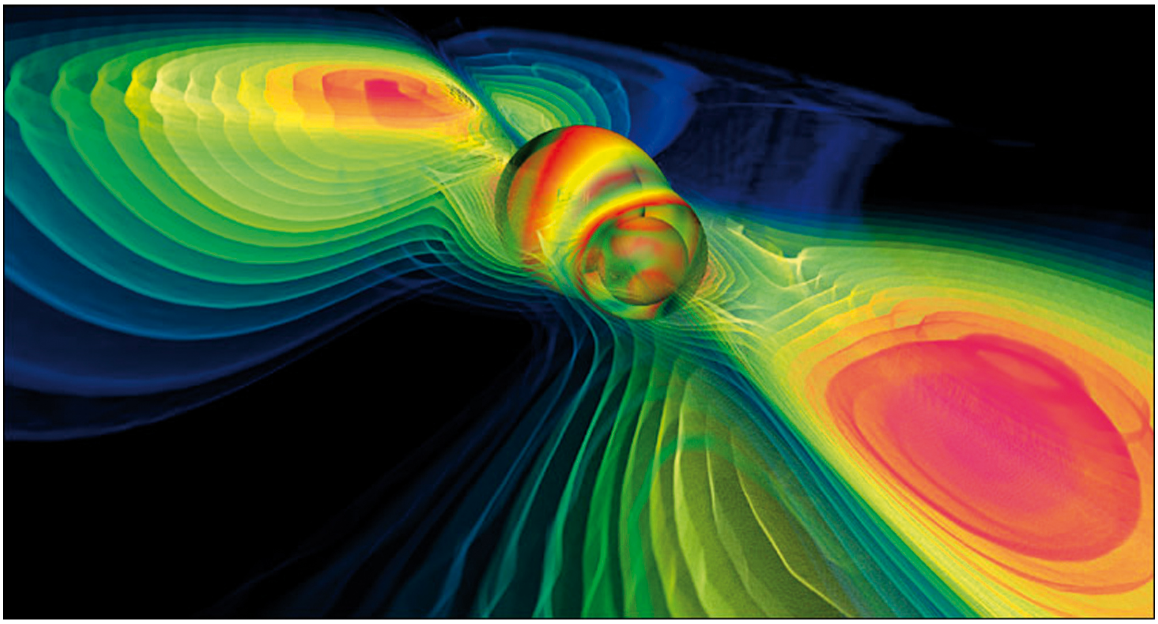
Comparing this to the familiar "kinetic energy minus potential energy" for the non-relativistic Lagrangian, we see that the time dilation in the metric induces exactly the gravitational potential energy

$$V(\vec{x}) = m\Phi(\vec{x}) \,. \tag{6.2.27}$$

At lowest order in an expansion in $\dot{\vec{x}}^2/c^2$ and Φ/c^2, general relativity therefore exactly reproduces Newtonian gravity. However, for large velocities, $\dot{\vec{x}}^2 \sim c^2$, or strong gravity, $\Phi \sim c^2$, general relativity predicts corrections to the old results: these are the dots in (6.2.26).

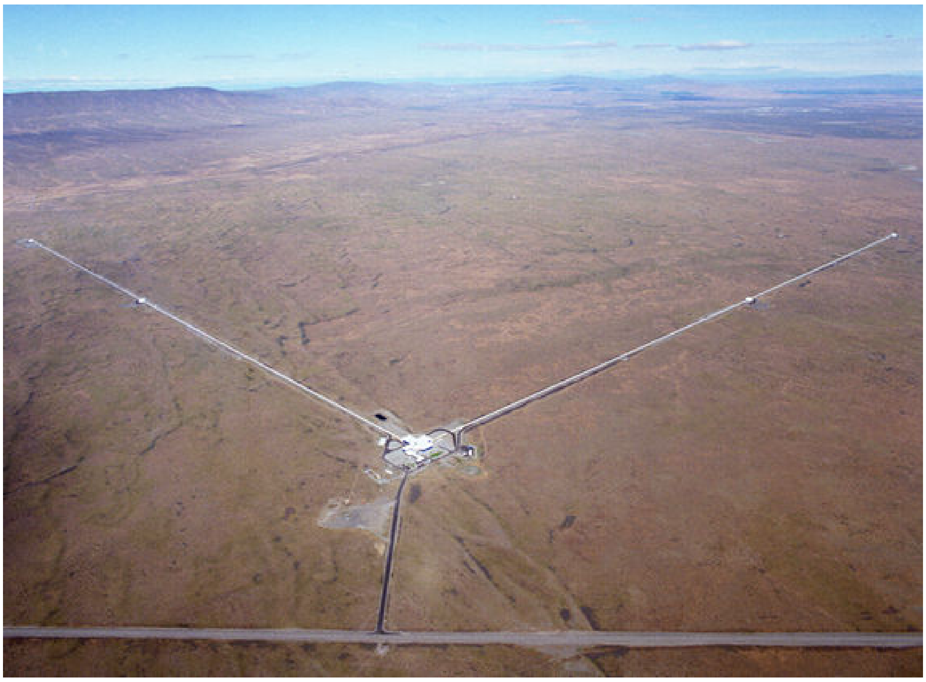
6.3 Gravitational Waves

We have seen that shaking an electric charge creates changing electric and magnetic fields that propagate away from the source as electromagnetic waves. Similarly, shaking a massive object creates ripples in spacetime that propagate away from the source as *gravitational waves*. Such gravitational waves are produced, for instance, when two black holes collide, as in this numerical simulation:



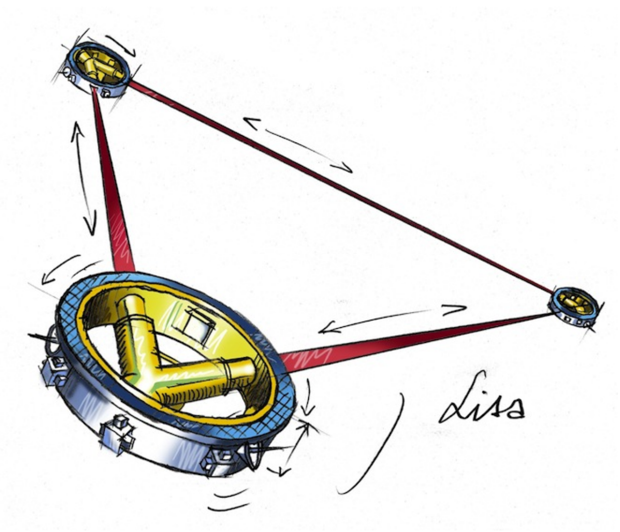
[9]We have dropped the rest mass energy mc^2 since it doesn't affect the dynamics.

The experiments that try to detect these gravitational waves are quite fascinating.



In the LIGO experiment, test masses are separated by a distance of about 1 km. When a gravitational wave passes by, the space in between the test masses stretches and contracts. Powerful lasers are used to measure this change in the separation between the test masses. How much does a typical gravitational wave change the length between the masses? 10^{-20} m !!! This is a factor of 10^6 times smaller than the size of an atomic nucleus! That LIGO is capable to measure such incredibly small changes in distance is astounding.

In the future, these experiments might go into space:

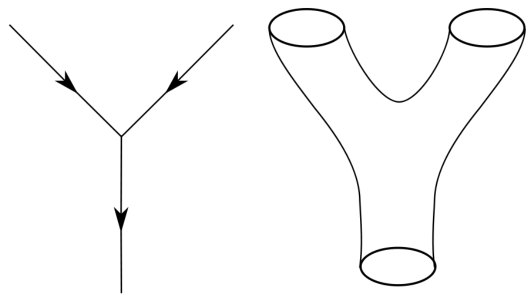


There the test masses can be separated by enormous distances (up to about 1 million km), giving the experiments access to lower frequency gravitational waves.

6.4 Quantum Gravity*

Quantum mechanics (QM) and general relativity (GR) are the two pillars of modern physics. QM applies to very small objects, like atoms and elementary particles, where the effects of gravity are miniscule. GR, on the other hand, deals with very large objects, such as stars, galaxies and the universe. On those scales, quantum mechanics plays no significant role. In most situations, we therefore use either QM or GR, but not both together. The interior of black holes and the Big Bang are two important exceptions. In these situations large masses occupy small regions of space. GR and QM are then equally important. A naive treatment of quantum field theory (QFT) for gravitons (the particles associated with the gravitational field, cf. Lecture 5) blows up in our faces. If we apply the standard QFT methods to gravity, we find negative probabilities and other nonsense.

These problems can be traced back to the fact that, in the standard approach, interactions between particles occur at points in spacetime. If the fundamental objects of the theory instead are extended objects—such as strings and membranes—then the interactions are smeared out over a finite region of spacetime:



This cures the pathologies of the standard approach. String theory is the only example of a sensible theory of quantum gravity. Unfortunately, at the time of writing, string theory hasn't provided testable predictions. It therefore remains a speculative proposal.

7 Cosmology

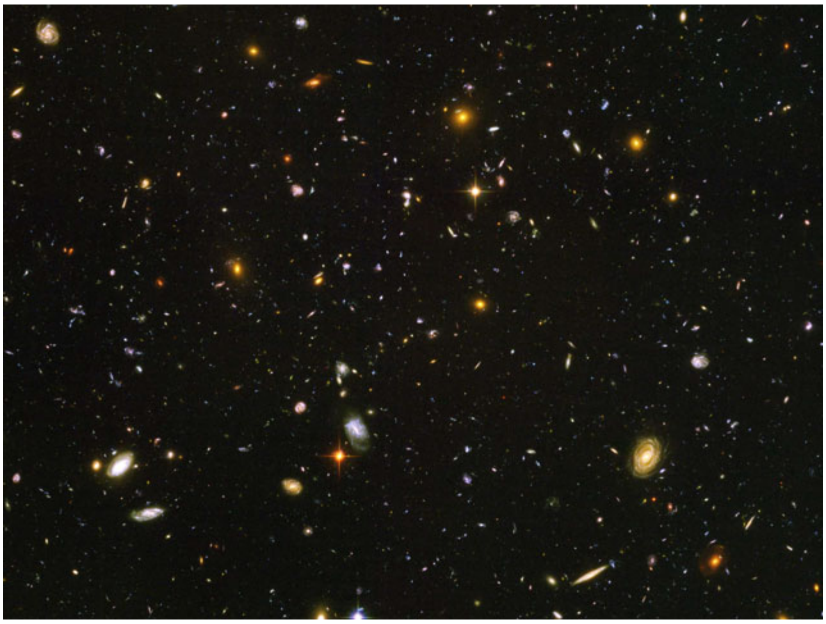
"I'm astounded by people who want to 'know' the universe when it's hard enough to find your way around Chinatown." Woody Allen

We are going to finish this course with a big topic: the entire universe. The study of the structure and evolution of the universe is called **cosmology**. Interestingly, all of the topics that we have studied in this course come together in cosmology. To describe the universe we need relativity, quantum mechanics, statistical mechanics, particle physics, and more.

Imagine you look at the night sky. You randomly select a patch of the sky only a fraction of the size of the full moon. To the naked eye it will look pitch black:

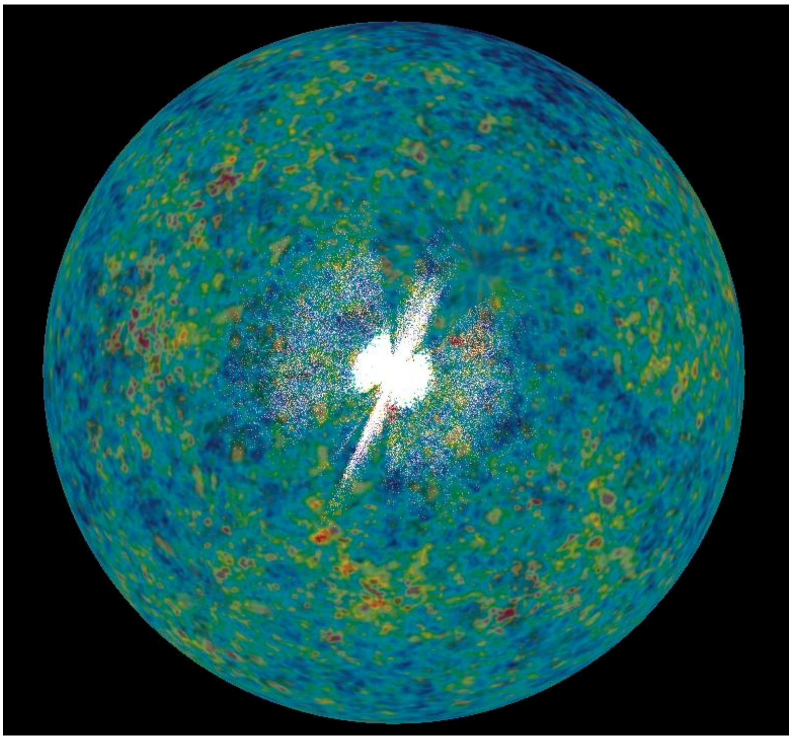
You then decide to look at the same region of the sky with the Hubble Space Telescope. But, instead of just looking at it for an instant, the telescope collects light for about 1 million seconds (= 12 days).

The result is one of the most stunning astronomical images ever produced:



Every object in the picture is an entire galaxy! A few thousand of them, each containing billions of stars similar to our Sun. Many of the stars have planets around them like our Earth. But, now remember that this tiny patch of the sky was selected at random. Any other randomly selected region in the sky would look essentially the same. From this we can estimate that the observable universe contains about 350 billion large galaxies and about 7 trillion dwarf galaxies. The total number of stars is about 30 billion trillion.

Here is the latest image of the entire observable universe:



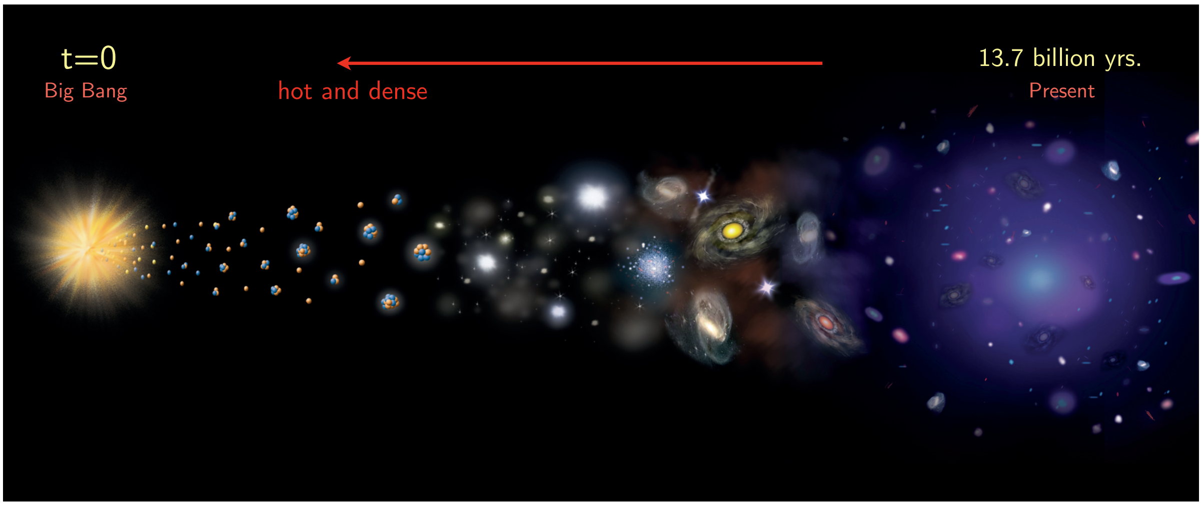
The white dots are galaxies, the surrounding sphere is the last-scattering surface of the cosmic microwave background. The goal of this lecture is to tell you where these structures came from.

7.1 The Big Bang

"What has the universe got to do with it?
You're here in Brooklyn! Brooklyn is not expanding!"

Alvy Singer's mom, Annie Hall.[1]

Modern cosmology started with the striking observation that the universe is expanding. This led to the Big Bang theory.[2] The idea is simple: Everything is getting further apart. In the past everything was therefore closer together. That's it. Well, there is a little bit more: when things get compressed they tend to heat up. Similarly, the early universe was much hotter and denser than it is today:



In the earliest moments, the average particle energies exceeded those produced in the LHC by many orders of magnitude. Going even further back in time, at some point the equations of general relativity become singular (the technical term for "blow up in your face"). We label this event the time $t = 0$. Space and time lose their familiar meanings at and before $t = 0$. We sometimes call $t = 0$ the Big Bang (singularity).

On the other hand, from 10^{-10} seconds to today, the history of the universe is based on well understood and experimentally tested laws of particle physics, nuclear and atomic physics and gravity. We are therefore justified to have some confidence about the events shaping the universe during that time.[3] Table 7.1 shows key events in the thermal history of the universe. In the following, I want to highlight two important events:

- *Big Bang Nucleosynthesis*
- *Recombination*

[1] http://www.youtube.com/watch?v=5U1-OmAICpU

[2] Notice that what we usually call the Big Bang theory has nothing to say about how the universe started. Also, do not think of the Big Bang as an explosion. The Big Bang occurred everywhere at the same time. Every point is space is equivalent to every other point. There is no centre of the universe, although it looks as if everything is moving away from us. But the same is true for any other observer in a galaxy far, far away.

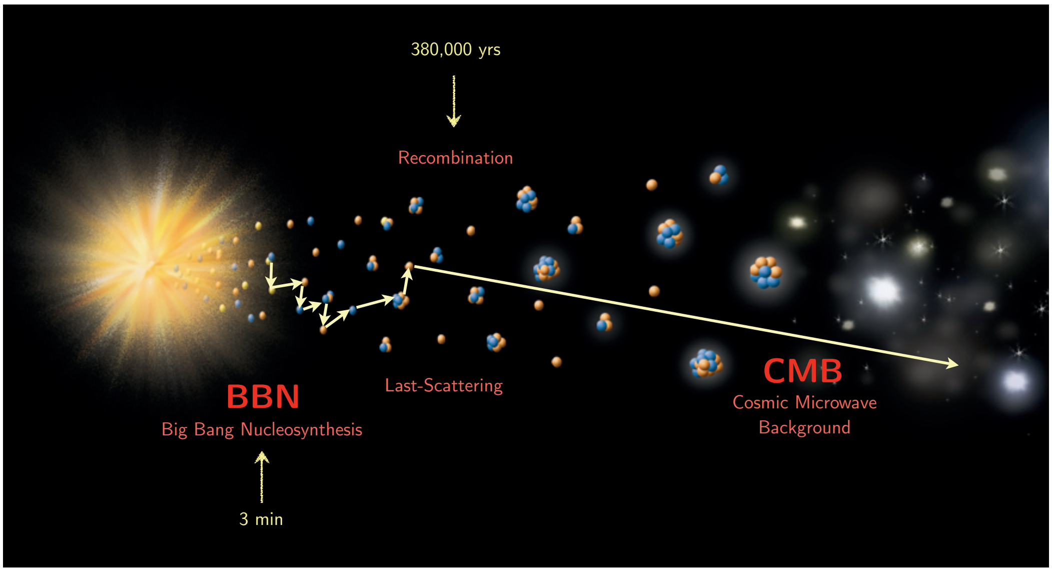
[3] I will indulge in some speculations about what happened at even earlier times in §7.3.

Table 7.1: Major Events in the History of the Universe.

	Time	Energy	
Planck Epoch?	$< 10^{-43}$ s	10^{18} GeV	
Inflation?	$\gtrsim 10^{-34}$ s	$\lesssim 10^{15}$ GeV	
Matter $\neq$ anti-matter?	$< 10^{-10}$ s	> 1 TeV	
Electroweak phase transition	10^{-10} s	1 TeV	
Protons/neutrons form	10^{-4} s	10^{2} MeV	
BBN: H, He, Li form	3 min	0.1 MeV	
			Redshift
Matter = radiation	10^{4} yrs	1 eV	10^{4}
CMB: recombination	10^{5} yrs	0.1 eV	1,100
Dark ages	$10^{5} - 10^{8}$ yrs		> 25
Reionization	10^{8} yrs		$25 - 6$
First galaxies form	$\sim 6 \times 10^{8}$ yrs		~ 10
Dark energy takes over	$\sim 10^{9}$ yrs		~ 2
Solar system formed	8×10^{9} yrs		0.5
Albert Einstein born	14×10^{9} yrs	1 meV	0

7.1.1 Nucleosynthesis

The early universe was too hot for stable atoms to exist. The matter was in the form of free electrons and nuclei (protons and neutrons). Once in a while two protons would combine into a helium nucleus. This process of Big Bang Nucleosynthesis (BBN) predicts the proportions of the light elements[4] with glorious precision: 75% H, 25% He, 10^{-5}% D and 10^{-10}% Li. We see precisely those proportions in old gas clouds. This is one of the great successes of the Big Bang theory.



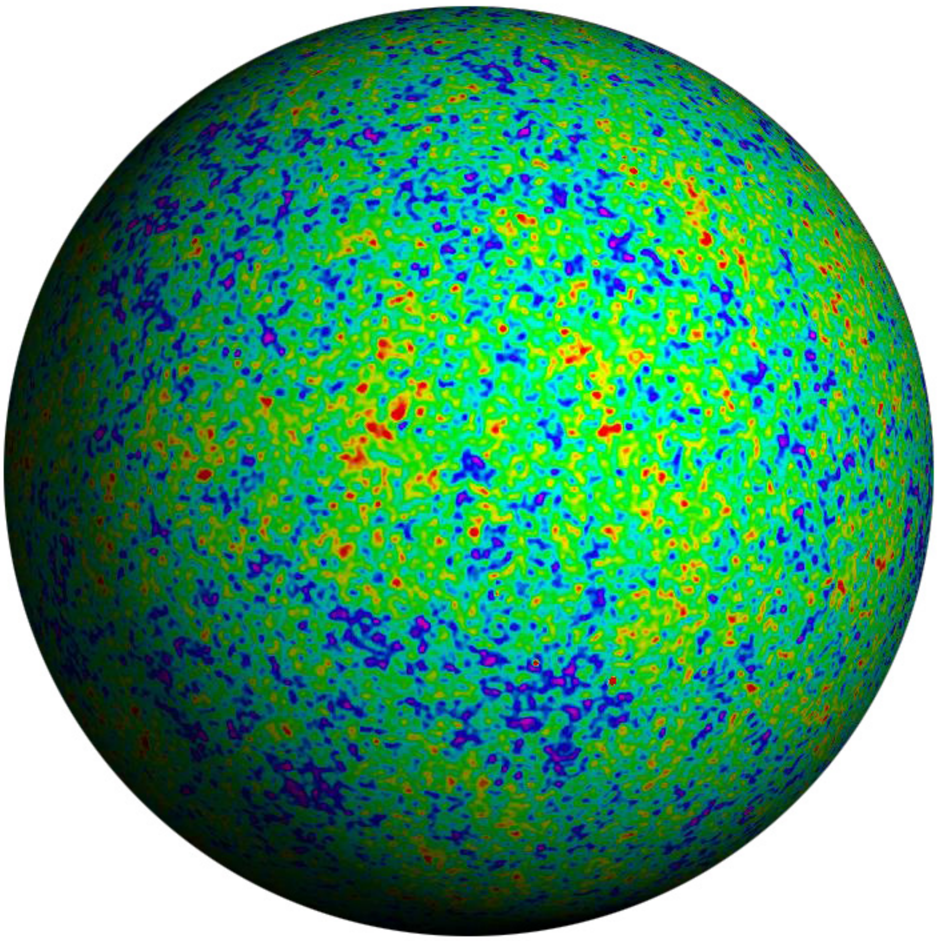
[4]Heavy elements are later produced inside stars.

7.1.2 Recombination

The earliest we can see with light is 380,000 years after the Big Bang. Before this time, the universe was still filled with the plasma of free electrons and nuclei. Light therefore couldn't propagate very far before bouncing off the charged electrons. *Recombination* marks the time when the universe cooled enough to allow the formation of the first stable atoms. At that moment, light started to stream freely. Today, some 13.7 billion years later, we receive the light from that era as the so-called *cosmic microwave background* (CMB).[5]

7.1.3 Cosmic Microwave Background

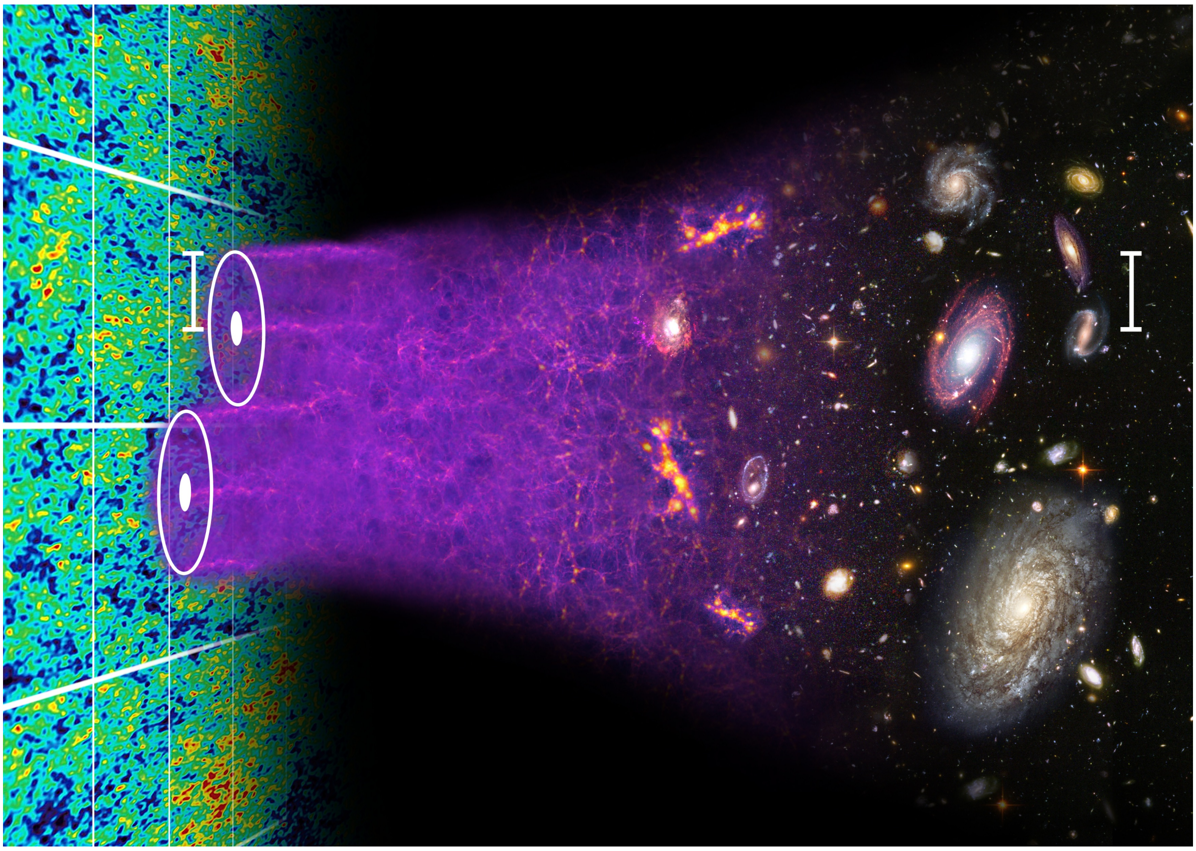
When people first observed the CMB in the 60s they found it to be completely uniform, the same temperature, about 2.7 degree Kelvin, in all directions. However, when people began to measure this radiation more and more accurately, they discovered small variations at the level of 1 part in 10 000. The CMB has spots. Parts of the sky are slightly hotter, parts slightly colder:



[5] Although the CMB photons started off very energetic at recombination, billions of years of cosmic expansion has stretched their wavelength into the microwave range.

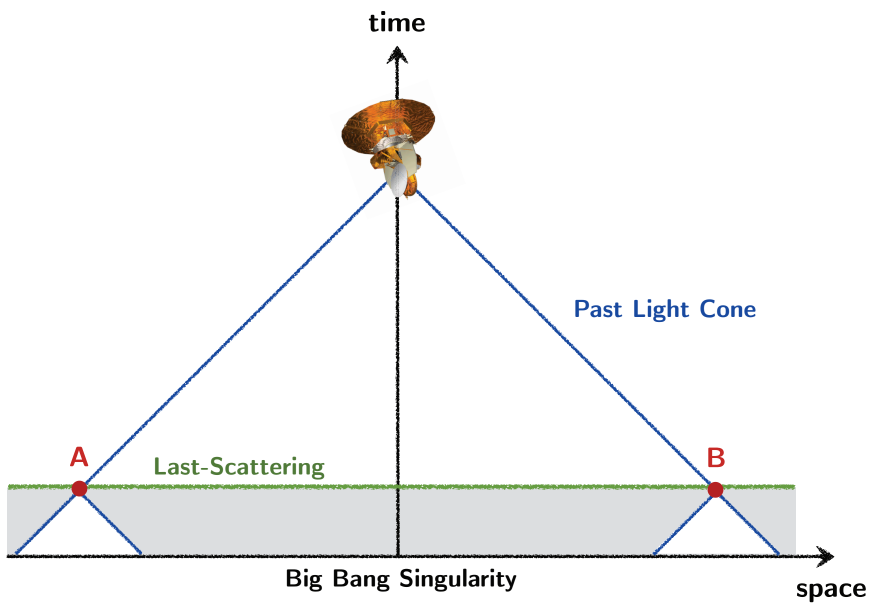
7.1.4 Gravitational Instability

The variations in the CMB temperature reflect tiny variations in the primordial density of matter. Over time and under the influence of gravity these matter fluctuations grow. The rich are getting richer. Dense regions are getting denser. Galaxies, stars and planets form.

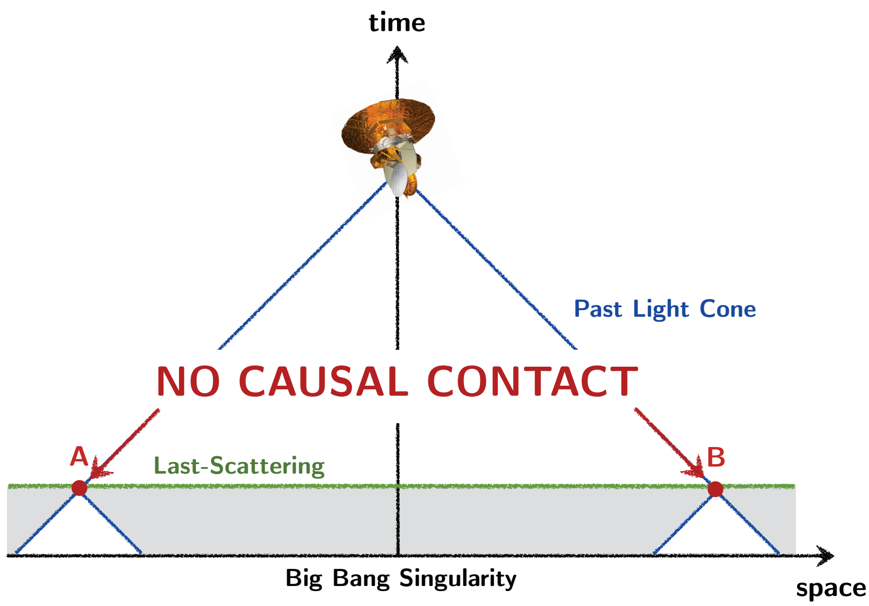


7.2 The Horizon Problem

There is a problem with this simple picture of structure formation. Consider the past light cone of an observer today. The following figure shows how this light cone intersects spatial surfaces of constant time at the moment of the initial Big Bang singularity and at last-scattering of the CMB (i.e. recombination):



Now consider two points A and B on the last-scattering surface separated by 180 degrees (i.e. opposite points on the sky). These points have their own past light cones which intersect with the Big Bang singularity in finite time. Notice that the time between the singularity and last-scattering is much smaller than the time between last-scattering and the present time. We see that the past light cones of the points A and B do not overlap. But regions of spacetime can only exchange information if they have overlapping past light cones. This means that the points A and B have never been in causal contact:

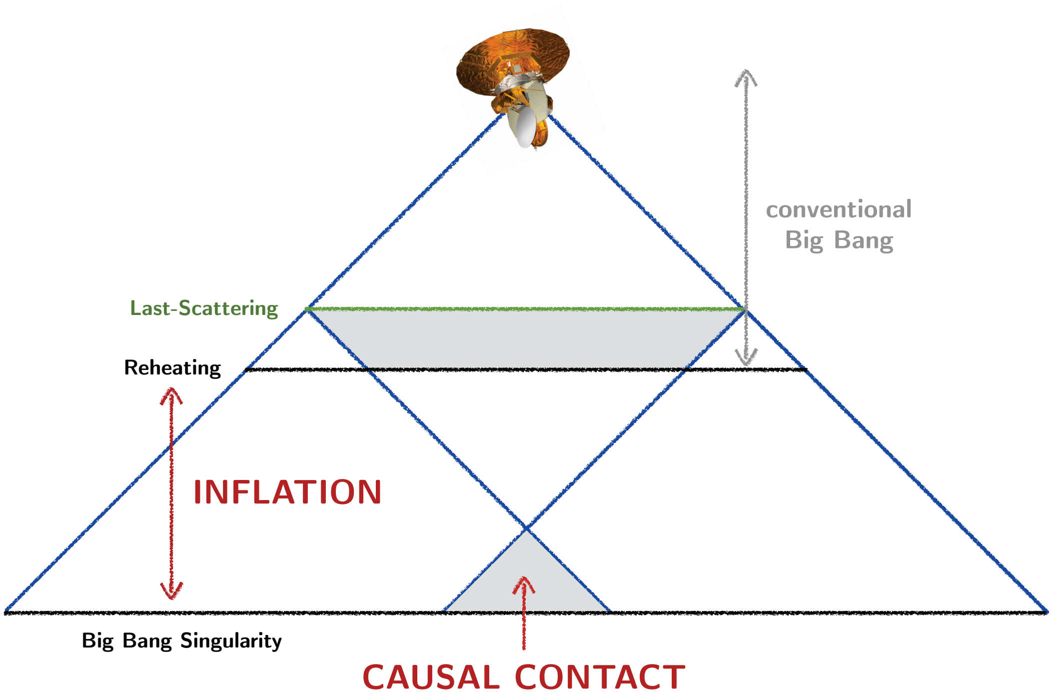


The same applies to any two points in the CMB that are separated my more than 2 degrees. It seems that the CMB is made out of many causally-disconnected patches. (In fact, 10^4 of them.) Yet, we still observe an almost perfectly uniform temperature of the CMB, even for widely separated points. How can that be? Who told these independent regions to share the same temperature? This puzzle is called the *horizon problem*.

7.3 Inflation

7.3.1 Solution of the Horizon Problem

The horizon problem would be solved if we could somehow achieve that the effective time[6] between the singularity and last-scattering is *larger* than the time between last-scattering and the present—in other words, if there was extra time *before* what we usually associate with the conventional Big Bang:



At sufficiently early times, the past light cones of any points in the CMB would then have overlapped and the whole CMB would have originated from a causally connected region of space. The uniformity of the CMB would have been given a causal explanation.

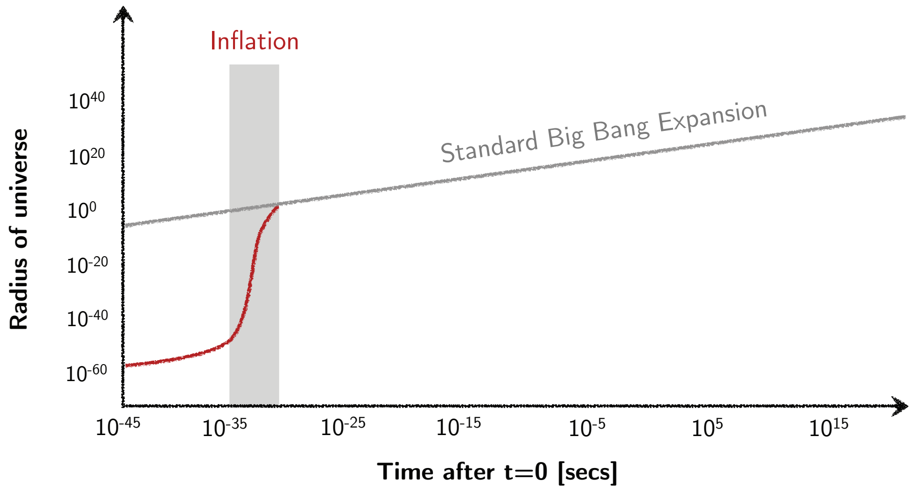
A process that achieves this solution to the horizon problem is cosmological *inflation*. During inflation the size of the universe grows exponentially:

$$ds^2 = c^2 dt^2 - e^{2Ht} d\vec{x}^{\,2} \, . \tag{7.3.1}$$

If inflation lasts long enough, CMB patches on opposite sides of the sky would have been close enough to communicate at primordial times. Because of the inflationary expansion the

[6] Technical remark for full disclosure: The effective time is called *conformal time*. The physical time between the singularity and recombination is always 380,000 years. However, it is the conformal time that is relevant for the horizon problem.

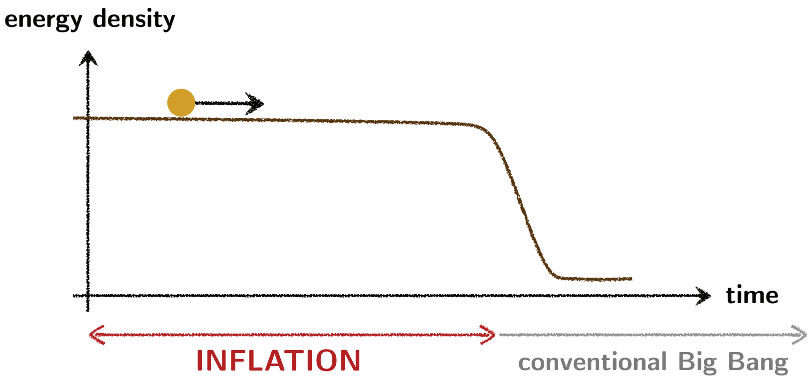
observable universe originated from a much smaller region of space than a naive extrapolation of the conventional Big Bang evolution suggests:



Measured in physical time, inflation lasted for only about 10^{-34} seconds.

7.3.2 The Physics of Inflation

From the Einstein equations, it is possible to show (see below) that the metric (7.3.1) requires a nearly constant energy density, $\rho_I = const.$

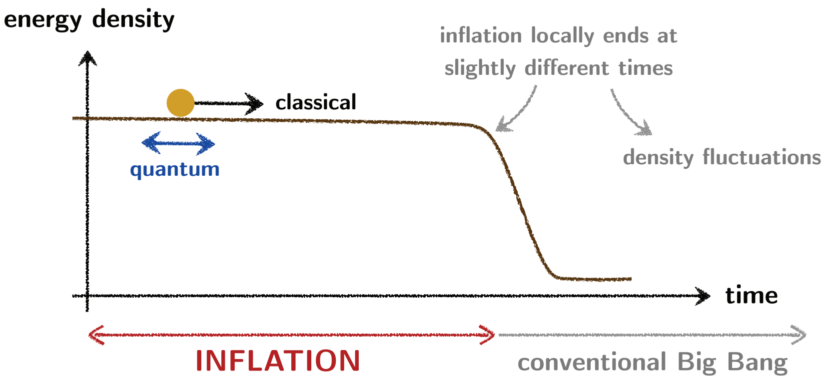


Notice that this implies somewhat exotic physics. In particular, maintaining a constant energy *density* requires that energy is created in order to compensate for the fact that the volume grows. This is challenging to arrange, but it can be done if the early universe is filled by something like the Higgs field (but it can't be the Higgs field itself).

7.4 From So Simple A Beginning

7.4.1 Quantum Fluctuations

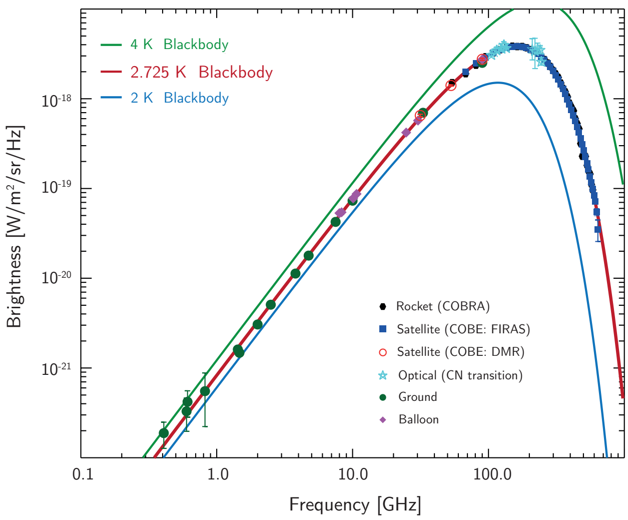
The inflationary phase is unstable and will decay within a certain time. However, in quantum mechanics this decay can't be perfectly synchronous over all of space. The uncertainty principle requires that there will be small fluctuations.



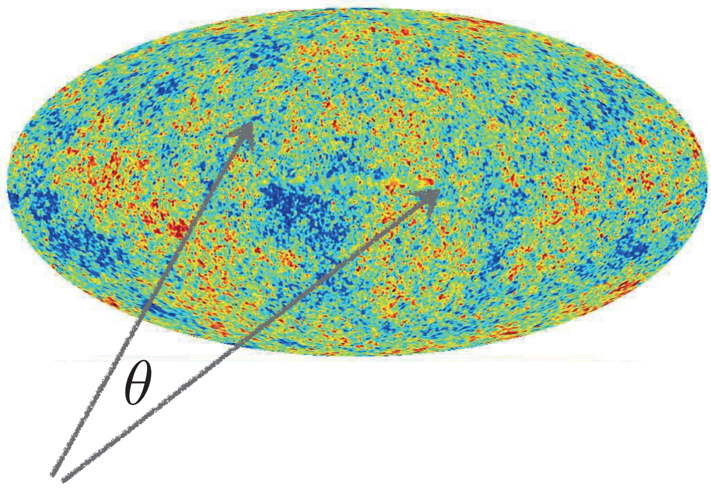
It is these fluctuations that are the source of the CMB fluctuations. To see this, imagine that the universe is filled with small clocks that measure the amount of time left before a given region of space decays. But, I told you in Lecture 2 that the uncertainty principle limits how well one can keep track of time with a very small clock. As a result, inflation will end at slightly different times in different regions. Hence, different regions of the universe grow by slightly different factors and end up at slightly different densities. The CMB photons in high density regions are more energetic than in low density regions. Voilá, the source of the CMB fluctuations is explained.

7.4.2 CMB Anisotropies

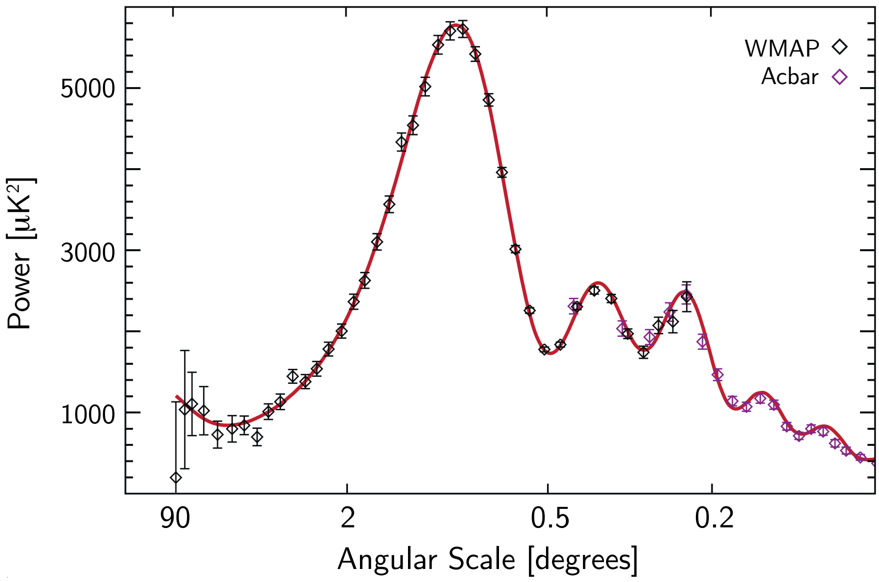
The CMB is one of the most perfect blackbodies ever observed:



The mean temperature associated with the radiation is 2.725 K. In addition, different directions on the sky show small variations in the temperature, at the level of 10^{-5} K. A key test of the inflationary hypothesis comes from studying the statistics of these CMB fluctuations. Consider two points in the CMB separated by an angle θ:



We want to know how correlated these two points are. The following plot shows this correlation as a function of the separation angle θ:



The points with error bars are data, the line is the quantum mechanical calculation (plus basic fluid dynamics and gravity to evolve the fluctuations forward in time). Something is clearly working. It is remarkable that we can now trace the origin of galaxies, some of the largest objects we know, to quantum mechanics, the physics of the very small.

7.5 Breaking News: BICEP2

On March 17, 2014, an experiment at the South Pole—BICEP2—announced the discovery of gravitational waves from inflation. The result is actively debated in the physics community, but if confirmed it will be one of the greatest discoveries in the history of cosmology. I will briefly explain why the result is such a big deal.

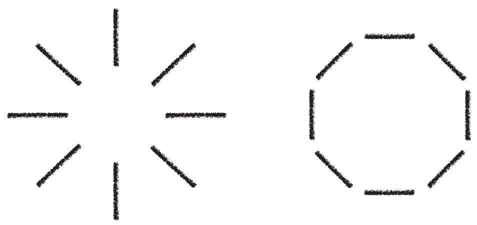
7.5.1 B-modes from Gravitational Waves

How can we be sure that something as dramatic as inflation really happened in the early universe? Remember that we are talking about the earliest moments after the Big Bang, when the energy was many, many orders of magnitude higher than what has ever been probed in experiments. The physical laws in operation at those high energies are very uncertain. Luckily, inflation makes a very clean prediction that can be tested against observations.

We have seen how quantum fluctuations affect the time when inflation end and lead to density fluctuations after inflation. Similarly, quantum fluctuations also create an anisotropic stretching of spacetime itself. These ripples in spacetime are the *gravitational waves* I mentioned at the end of Lecture 6. The strength of the gravitational wave signal depends on the precise energy at which inflation occurred. These gravitational waves leave a unique signature in the *polarization* of the CMB.

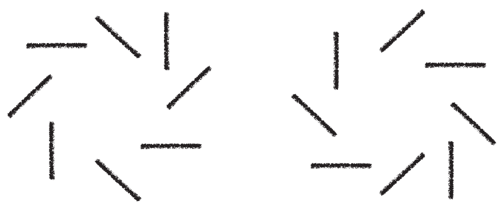
Polarization gets created when the inhomogeneous distribution of CMB photons scatters off the free electrons at recombination. Density fluctuations and gravitational waves each produce polarization, but the polarization patterns that they create are distinct. Density fluctuations only create a so-called *E-mode* pattern:

E-mode
(grad)



The polarization vectors are aligned radially or tangential around the hot and cold spots of the CMB. Gravitational waves, on the other hand, produce a *B-mode* pattern:

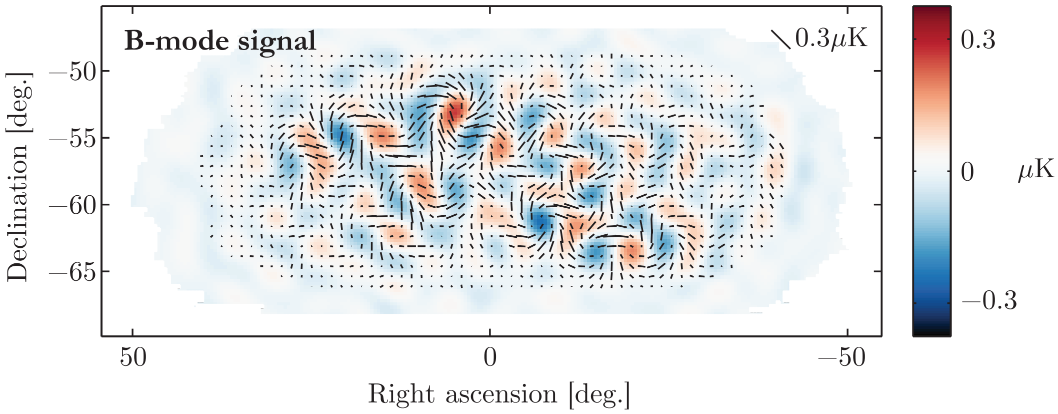
B-mode
(curl)



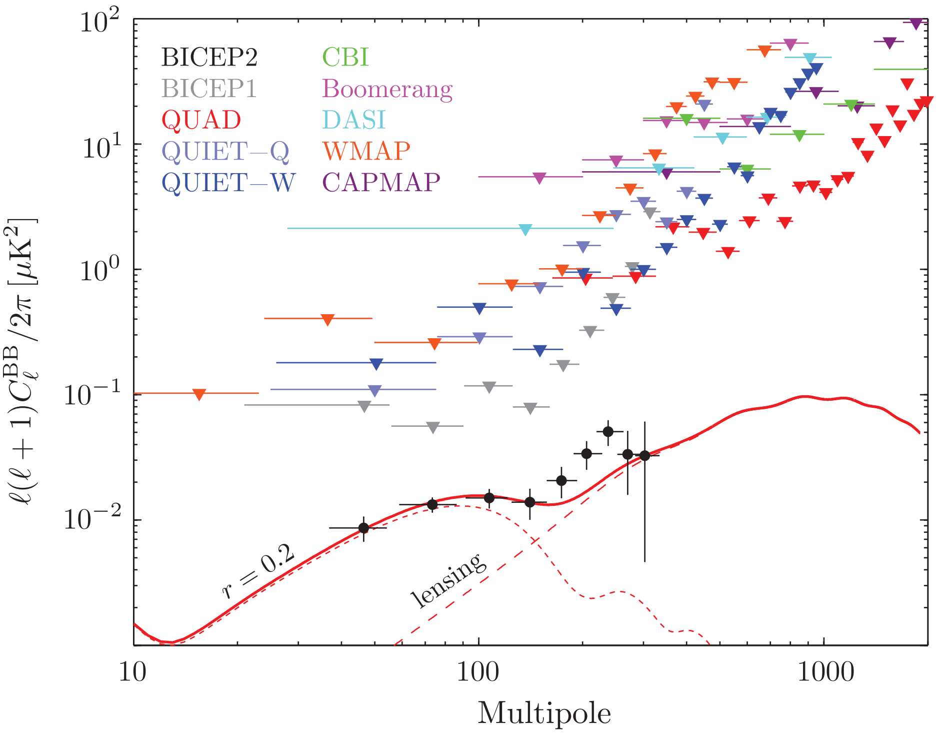
This time the polarization vectors are arranged in a *swirly* pattern around hot and cold spots in the CMB. Before March, only the E-mode pattern had been seen. Many consider the B-mode pattern the ultimate test for inflation. Some have called it the "smoking gun".

7.5.2 Have They Been Seen?

It came as an absolute shock to the physics community when the BICEP team announced in March that they had detected the elusive B-mode signal. This is the map they presented:



Can you see the swirly pattern? They also produced a beautiful measurement of the two-point correlations of the polarization signal:



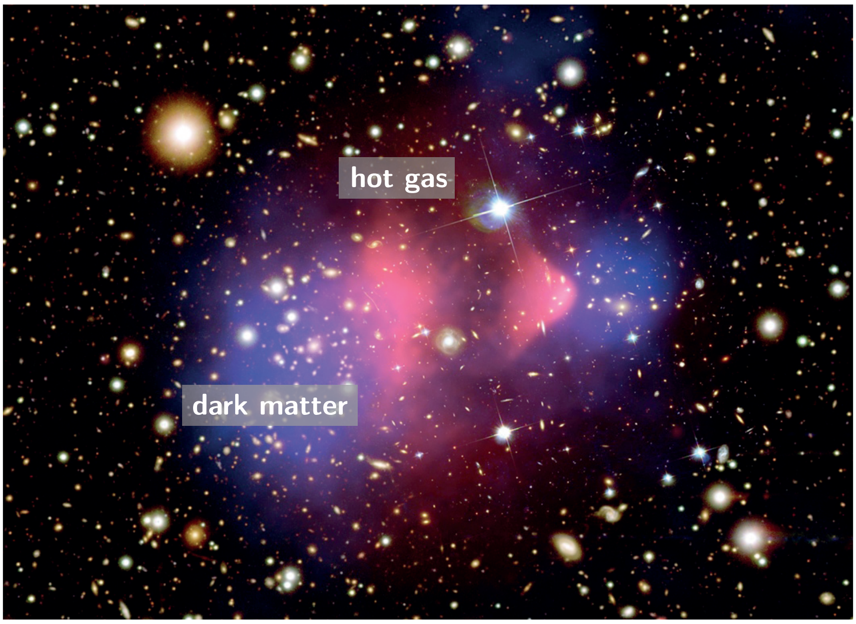
But, not everybody is convinced yet. Doubts have been raised whether BICEP2 is really seeing the primordial signal. Some believe that the team has underestimated the contamination from dust in our galaxy. This dust also produces polarised microwave radiation which could mimic the inflationary signal. We will know the answer within a year. The scientific life of a cosmologist has rarely been as exciting.

7.6 A Puzzle and A Mystery

I want to end this lecture by telling you about a puzzle and a mystery.

7.6.1 Dark Matter

The puzzle is the following: wherever we look in the universe we seem to find more invisible *dark matter* than visible atomic matter. A dramatic illustration of this is the bullet cluster:



The picture shows two clusters of galaxies that have (relatively) recently passed right through each other. It turns out that the large majority (about 90%) of ordinary matter in a cluster is not in the galaxies themselves, but in hot X-ray emitting intergalactic gas. As the two clusters passed through each other, the hot gas in each smacked into the gas in the other, while the individual galaxies and the dark matter (presumed to be collisionless) passed right through. So, the collision has swept out the ordinary matter from the clusters, displacing it with respect to the dark matter.

We now believe that 85% of the matter in the universe is dark matter. But, we don't know what it is. However, there are many theoretical ideas with fancy names, such as axions, WIMPs, neutralinos, etc. In addition, there are many future experimental tests that will help us to narrow down the range of possibilities and hopefully detect dark matter particles directly. So, dark matter is a puzzle, but one that we are quite optimistic we will solve.

7.6.2 Dark Energy

> *"Physics thrives in crisis"*
> Steven Weinberg.

You have probably heard that recently the expansion of the universe started accelerating again. This is like inflation, but at much, much lower energies. What form of *dark energy* is responsible for this?

The spacetime of the uniformly expanding universe is described by the *Friedmann-Robertson-Walker metric*,

$$\mathrm{d}s^2 = c^2\mathrm{d}t^2 - a^2(t)\mathrm{d}\vec{x}^2 \ , \tag{7.6.2}$$

where the *scale factor* $a(t)$ measures the stretching of space. General relativity tell us how $a(t)$ changes with time. This depends on the energy density ρ in the universe,

$$\left(\frac{1}{a}\frac{da}{dt}\right)^2 = \frac{8\pi G}{3}\rho \ . \tag{7.6.3}$$

Two important components are matter (dark or not) and radiation (photons or neutrinos). Matter dilutes as the universe expands: $\rho_m \propto 1/a^3(t)$ (energy is constant, so this simply describes the growth of the volume—three factors of $a(t)$). Whenever the universe is dominated by matter, we therefore find

$$\rho_m \propto a^{-3}(t) \xrightarrow{\text{eq. (7.6.3)}} a(t) \sim t^{2/3} \ . \tag{7.6.4}$$

Radiation also dilutes and in addition it redshifts. The energy density therefore decreases with an extra factor of $a(t)$, i.e. $\rho_r \propto 1/a^4(t)$. A universe dominated by radiation therefore evolves as

$$\rho_r \propto a^{-4}(t) \xrightarrow{\text{eq. (7.6.3)}} a(t) \sim t^{1/2} \ . \tag{7.6.5}$$

A common feature of the solutions (7.6.4) and (7.6.5) is that the expansion slows down:

$$\frac{d^2a}{dt^2} < 0 \ . \tag{7.6.6}$$

Another way of seeing this is to consult a second Einstein equation

$$\frac{1}{a}\frac{d^2a}{dt^2} = -\frac{4\pi G}{3}(\rho + 3p) \ . \tag{7.6.7}$$

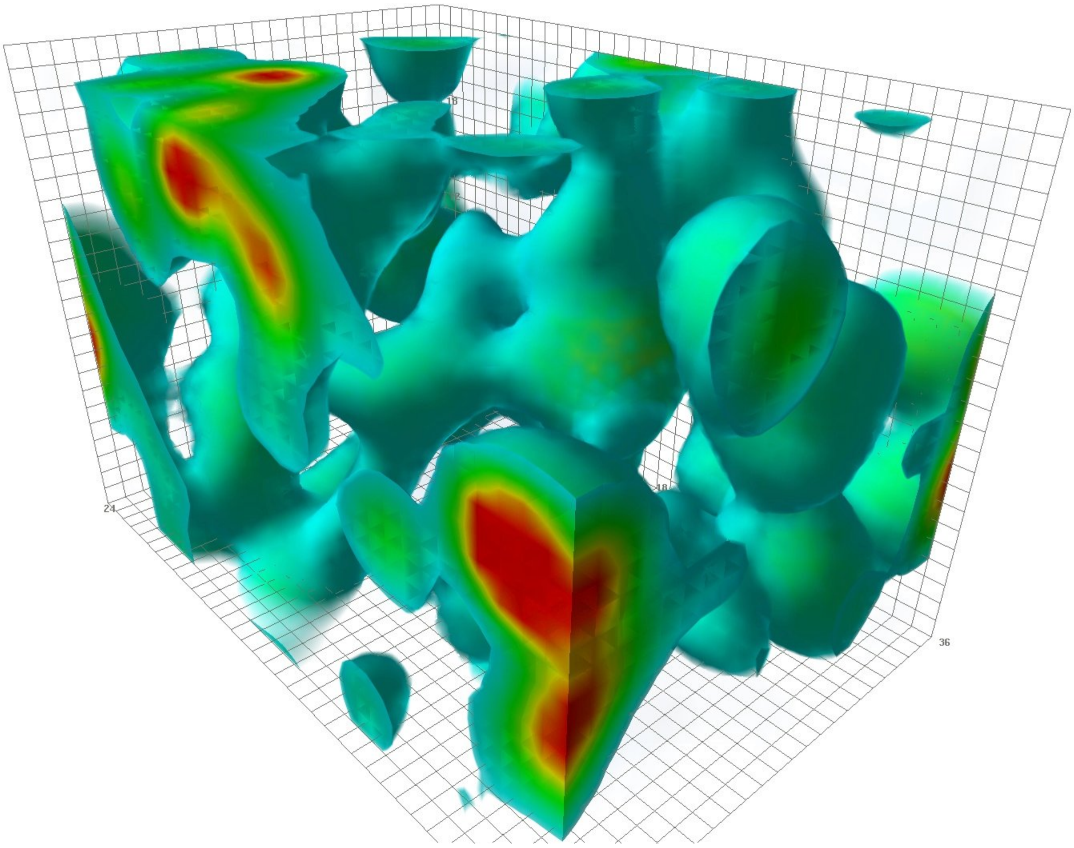
Here, p is pressure. For matter we have $p = 0$, while radiation has $p = \rho/3$. Since $\rho > 0$, it follows that both matter and radiation slow down the expansion.

In the 90s, astronomers tried to measure the rate of deceleration in order to determine exactly how much matter and radiation there was in the universe. To their surprise they found that the universe isn't slowing down at all. Instead the expansion is speeding up:

$$\frac{d^2a}{dt^2} > 0 \ . \tag{7.6.8}$$

How can that be? According to eq. (7.6.7), it requires the universe to be filled with a negative pressure fluid: $p < -\rho/3$. We call this fluid *dark energy*, but we have no idea what it is.

In fact, there is a well-known source of dark energy—quantum fluctuations of the vacuum:



However, when we use quantum mechanics to compute the size of the vacuum energy we find this

$$\rho_{\text{quantum}} \approx 10^{120} \rho_{\text{obs}} \, . \tag{7.6.9}$$

This is the worst disagreement between theory and experiment in the history of science. It is called the *cosmological constant problem*. In physics, we are not used to screwing up so badly. Especially, in quantum mechanics and relativity we have been spoiled with incredible precise predictions. So, we take this problem personally.

In fact, we do have an answer to the problem. Here it is: we postulate that in addition to the large quantum contribution ρ_{obs}, there is an equally large, but negative, classical contribution $\rho_{\text{classical}}$. We add these two contributions to get the observed dark energy density

$$\rho_{\text{obs}} = \rho_{\text{classical}} + \rho_{\text{quantum}} \, . \tag{7.6.10}$$

Numerically, this is what we do

$$10^{-120} \;=\; -\, 2.\underbrace{1568\ldots352\mathbf{1}}_{\text{120 digits}}\ldots \;+\; 2.1568\ldots352\mathbf{3}\ldots \tag{7.6.11}$$

I know, it looks ludicrous. The amount of fine-tuning seems ridiculous. Unfortunately, this is the best we have come up with. Can you help us to find a better solution?

A Symmetries and Conservation Laws

In Lecture 1, we got so close to one of the most important results in all of physics that I couldn't resist the temptation to write it down for you. It is called *Noether's theorem* and relates symmetries to conservation laws.

A.1 Conservation of Momentum

Let us start with two examples:

- **Example 1:** First, consider two identical particles at positions q_1 and q_2:

If the potential only depends on the distance between the particles, then the Lagrangian is

$$L = \frac{m}{2}\left(\dot{q}_1^2 + \dot{q}_2^2\right) - V(q_1 - q_2) \ . \tag{A.1.1}$$

It is instructive to rewrite this in terms of two new coordinates: $q_+ \equiv q_1 + q_2$ and $q_- \equiv q_1 - q_2$. The Lagrangian does not depend on q_+,

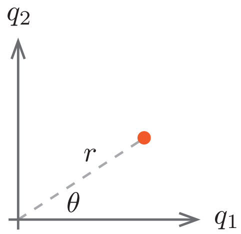
$$L = \frac{m}{2}\left(\dot{q}_+^2 + \dot{q}_-^2\right) - V(q_-) \ . \tag{A.1.2}$$

The Euler-Lagrange equation then implies

$$\frac{d\mathsf{P}_+}{dt} = -\frac{\partial V}{\partial q_+} = 0 \qquad \Rightarrow \qquad \mathsf{P}_+ = m\dot{q}_+ = m\dot{q}_1 + m\dot{q}_2 = const. \tag{A.1.3}$$

This is the conservation of momentum.

- **Example 2:** As a second example, consider a particle under the influence of a potential that depends only on the distance from the origin:



In Cartesian coordinates, the Lagrangian is

$$L = \frac{m}{2}\left(\dot{q}_1^2 + \dot{q}_2^2\right) - V(q_1^2 + q_2^2) \ , \tag{A.1.4}$$

while in polar coordinates it is

$$L = \frac{m}{2}\left(\dot{r}^2 + r^2\dot{\theta}^2\right) - V(r) \ . \tag{A.1.5}$$

The Lagrangian does not depend on θ. The Euler-Lagrange equation then implies

$$\frac{d\mathsf{P}_\theta}{dt} = -\frac{\partial V}{\partial \theta} = 0 \qquad \Rightarrow \qquad \mathsf{P}_\theta = mr^2\dot{\theta} = const. \tag{A.1.6}$$

This is the conservation of angular momentum.

The conservation laws in these examples might look like accidental properties of the potential. In reality, they are a consequence of a deep principle.

- **Example 1:** Imagine taking the two particles and shifting them both by the same distance ϵ:

$$\begin{array}{l} q_1 \mapsto q_1 + \epsilon \\ q_2 \mapsto q_2 + \epsilon \end{array} \qquad \Rightarrow \qquad \begin{array}{l} \delta q_1 = \epsilon \\ \delta q_2 = \epsilon \end{array} , \tag{A.1.7}$$

  where δq_i is the difference between the new and the old coordinates. Since the potential only depends on the difference of the coordinates, the Lagrangian doesn't change. We call the transformation (A.1.7) a *translation symmetry* of the system.

- **Example 2:** In the second example, the Lagrangian doesn't change if we rotate the particle position around the origin

$$\begin{array}{l} q_1 \mapsto +q_1 \cos\epsilon + q_2 \sin\epsilon \\ q_2 \mapsto -q_1 \sin\epsilon + q_2 \cos\epsilon \end{array} \qquad \Rightarrow \qquad \begin{array}{l} \delta q_1 = +q_2 \epsilon \\ \delta q_2 = -q_1 \epsilon \end{array} , \tag{A.1.8}$$

  where in the last expression we have assumed that ϵ is infinitesimal, so that $\sin\epsilon \approx \epsilon$ and $\cos\epsilon \approx 1$. (Any continuous transformation can be built up from a sequence of infinitesimal transformations.) We call the transformation (A.1.8) a *rotation symmetry* of the system.

A.2 Noether's Theorem

Emmy Noether proved a striking result:

for every continuous symmetry there is a conservation law.

Let's prove it.

Consider a small shift of coordinates, that may itself depend on the value of the coordinates

$$\delta q_i = f_i(q)\, \epsilon \, , \tag{A.2.9}$$

i.e. each coordinate shifts by an amount proportional to ϵ, but the proportionality factor can depend on where you are. This includes **Example 1** ($f_1 = f_2 = 1$) and **Example 2** ($f_1 = q_2$, $f_2 = -q_1$). We can calculate how much $L(q, \dot{q})$ changes under this transformation

$$\delta L = \sum_i \left(\frac{\partial L}{\partial \dot{q}_i} \delta \dot{q}_i + \frac{\partial L}{\partial q_i} \delta q_i \right) . \tag{A.2.10}$$

Now we do a bit of magic. Watch it carefully. First, we remember that $\mathsf{P}_i = \partial L / \partial \dot{q}_i$. Thus the first term is $\sum_i \mathsf{P}_i \delta \dot{q}_i$. Hold on to that while we look at the second term $\partial L / \partial q_i \, \delta q_i$. Using the Euler-Lagrange equation, $\partial L / \partial q_i = d\mathsf{P}_i / dt$, we get $\dot{\mathsf{P}}_i \delta q_i$. Combining the terms, here is what we get for the change of the Lagrangian

$$\delta L = \sum_i \left[\mathsf{P}_i \delta \dot{q}_i + \dot{\mathsf{P}}_i \delta q_i \right] . \tag{A.2.11}$$

You should convince yourself that this is the same as

$$\delta L = \frac{d}{dt} \sum_i \mathsf{P}_i \delta q_i \ . \tag{A.2.12}$$

What does all of this have to do with symmetry and conservation? First of all, by definition, symmetry means that the Lagrangian is unchanged, $\delta L = 0$. So if (A.1.7) is a symmetry, then

$$\frac{d}{dt} \sum_i \mathsf{P}_i \delta q_i = 0 \ . \tag{A.2.13}$$

Using eq. (A.1.7) for δq_i, we get

$$\frac{d}{dt} \sum_i \mathsf{P}_i f_i(q) = 0 \ . \tag{A.2.14}$$

This means that the quantity

$$Q \equiv \sum_i \mathsf{P}_i f_i(q) \ , \tag{A.2.15}$$

does not change with time. It is conserved!

To see that this makes sense, let us go back to our examples:

- **Example 1:** For $f_1 = f_2 = 1$, eq. (A.2.15) becomes

$$Q = \mathsf{P}_1 + \mathsf{P}_2 \ . \tag{A.2.16}$$

  That is just the conservation of momentum that we found before. But now we can say a far more general thing:

  For any system of particles, if the Lagrangian is invariant under simultaneous translation of the positions of all particles, then momentum is conserved.

  Sweet!

- **Example 2:** For $f_1 = q_2$ and $f_2 = -q_1$, eq. (A.2.15) becomes

$$Q = q_2 \mathsf{P}_1 - q_1 \mathsf{P}_2 = \vec{q} \times \vec{\mathsf{P}} \ . \tag{A.2.17}$$

  That is just the conservation of momentum that we found before. Again, there is a deeper thing involved than just the angular momentum of a single particle:

  For any system of particles, if the Lagrangian is invariant under simultaneous rotation of the positions of all particles, then angular momentum is conserved.

A.3 Conservation of Energy

What about the conservation of energy? To relate this to a symmetry, we have to go beyond just shifting space coordinates. The symmetry connected with energy conservation involves a shift of *time*. (Shifting time is also called ageing.)

Imagine an experiment involving a closed system far from any perturbing influences. The system has *time translation symmetry* if the outcome of the experiment doesn't depend on whether we perform it today, tomorrow or in ten years. In the language of the Lagrangian method, such a symmetry means that the Lagrangian has no explicit dependence on time. This is a subtle point: The value of the Lagrangian may vary with time, but *only* because the coordinates and velocities vary. Explicit time dependence means that the *form* of the Lagrangian depends on time.

Example.—Consider a mass m attached to a spring with spring constant k. The Lagrangian is

$$L = \frac{1}{2}\left(m\dot{q}^2 - kq^2\right) \ . \tag{A.3.18}$$

If m and k are time-independent then this Lagrangian has time-translation symmetry. It doesn't change if $t \mapsto t + \epsilon$. Now imagine that the spring constant depends on time. For example, we might heat up the spring and let it cool down. The Lagrangian then is

$$L = \frac{1}{2}\left(m\dot{q}^2 - k(t)q^2\right) \ . \tag{A.3.19}$$

This is what we mean by an explicit time dependence.

Abstractly, we allow for an explicit time dependence in the Lagrangian by adding time as a coordinate,

$$L = L(q_i, \dot{q}_i, t) \ . \tag{A.3.20}$$

The total time derivative of this Lagrangian is

$$\frac{dL}{dt} = \sum_i \left[\frac{\partial L}{\partial q_i}\dot{q}_i + \frac{\partial L}{\partial \dot{q}_i}\ddot{q}_i\right] + \frac{\partial L}{\partial t} \ , \tag{A.3.21}$$

where the final term is only non-zero if the Lagrangian has an explicit time dependence. Let's examine the various terms in (A.3.21) using the Euler-Lagrange equations. The terms in the square brackets can be written as

$$\frac{\partial L}{\partial q_i}\dot{q}_i + \frac{\partial L}{\partial \dot{q}_i}\ddot{q}_i = \dot{\mathsf{P}}_i\dot{q}_i + \mathsf{P}_i\ddot{q}_i = \frac{d}{dt}(\mathsf{P}_i\dot{q}_i) \ . \tag{A.3.22}$$

We get

$$\frac{dL}{dt} = \frac{d}{dt}\sum_i(\mathsf{P}_i\dot{q}_i) + \frac{\partial L}{\partial t} \ . \tag{A.3.23}$$

Notice that even if there is no explicit time dependence in L, the Lagrangian will nevertheless depend on time through the first term $\sum_i(\mathsf{P}_i\dot{q}_i)$. There is no such thing as conservation of the Lagrangian. However, inspection of eq. (A.3.23) reveals something interesting. If we define a new quantity H—called the *Hamiltonian*—by

$$H \equiv \sum_i(\mathsf{P}_i\dot{q}_i) - L \tag{A.3.24}$$

then eq. (A.3.23) becomes

$$\frac{dH}{dt} = -\frac{\partial L}{\partial t} \ . \tag{A.3.25}$$

We see that H varies with time *only* if the Lagrangian has an explicit time dependence. In other words:

If the Lagrangian is invariant under time translations, then the Hamiltonian is conserved.

Example.—Let us evaluate the Hamiltonian for a simple example. Consider a single particle with Lagrangian,

$$L = \frac{m}{2}\dot{q}^2 - V(q) \ . \tag{A.3.26}$$

The momentum is

$$\mathsf{P} = \frac{\partial L}{\partial \dot{q}} = m\dot{q} \ . \tag{A.3.27}$$

Using eq. (A.3.24), we get

$$H = (m\dot{q})\dot{q} - \left(\frac{m}{2}\dot{q}^2 - V(q)\right) \tag{A.3.28}$$

$$= \frac{m}{2}\dot{q}^2 + V(q) \ . \tag{A.3.29}$$

This is just the energy of the particle.

It turns out that *all* conservation laws in nature are related to symmetries through Noether's theorem. This includes the conservation of electric charge and the conservation of particles such as protons and neutrons.

References

Principle of Least Action
- Feynman, *Feynman Lectures on Physics*, Vol. I, Ch. 19
- Susskind and Hrabovsky, *Classical Mechanics*

Electrodynamics and Relativity
- Einstein, *Relativity: The Special and the General Theory*

Quantum Mechanics
- Susskind and Friedman, *Quantum Mechanics*
- Zeilinger, *Dance of the Photons*
- Feynman, *QED*

Statistical Mechanics
- Carroll, *From Eternity to Here*
- Feynman, *The Character of Physical Law*
- Gleick, *The Information*

Particle Physics
- Sample, *Massive: The Hunt for the God Particle*
- Carroll, *The Particle at the End of the Universe*

General Relativity
- Thorne, *Black Holes and Time Warps*
- Ellis and Williams, *Flat and Curved Space-Times*

Cosmology
- Weinberg, *The First Three Minutes*
- Guth, *The Inflationary Universe*
- Overbye, *Lonely Hearts of the Cosmos*

Symmetry
- Zee, *Fearful Symmetry*
- Huang, *Fundamental Forces of Nature*