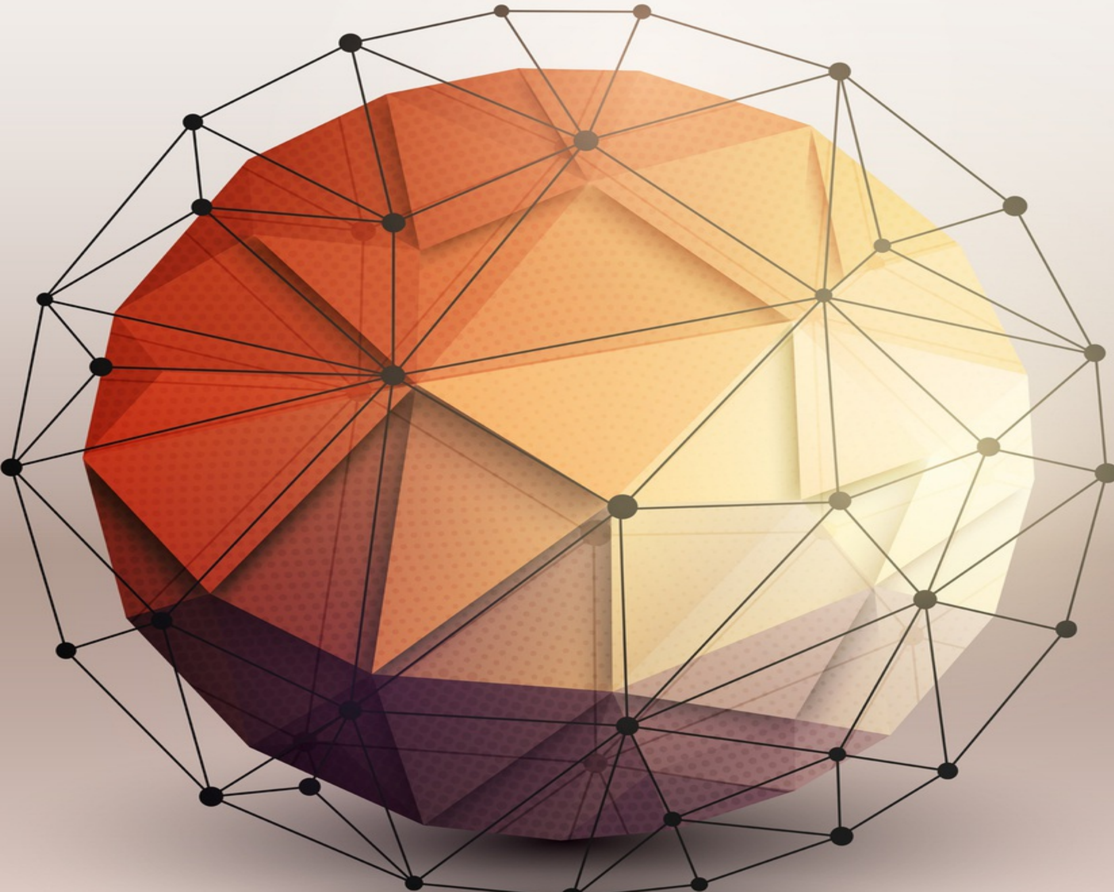
# String Theory

## SIMPLIFIED

What is Theoretical Physics?



Eliot Hawkins

# String Theory Simplified: What is Theoretical Physics?

By Eliot Hawkins

## Table of Contents

## Introduction

Theoretical physics uses the study of mathematics in order to describe specific elements of nature. Newton was known (in his era) as a "natural philosopher".

By this time, people already knew how to use the study of geometry and algebra for the construction of some simply astounding architectural works. This included some of the largest cathedrals in Europe.

However, algebra (and later, geometry) tend to only work effectively for items which are still. Newton was trying to describe items in motion and things that changed and, as a result, he invented calculus.

Today's theoretical physicists use a core technology which has developed from Gottfried Von Leibniz and Isaac Newton's calculus (created in the $18^{th}$ century).

The sun, moon, planets and stars are the most visible things in the sky and have always intrigued humans.

Newton's newly created calculus (combined with his Laws of Motion) described not only the motion of planets but also the motion of objects here on Earth.

Theoretical physicists today are often working at the frontier of known mathematics; sometimes inventing new mathematical equations much like Newton's creations in relation to calculus.

However, Newton was not just a theorist, he was also an experimenter. He spent so many hours observing nature's behavior that he sometimes neglected his health.

His Laws of Motion are not considered to be laws which nature feels forced to obey, but are actually a result of the observed behavior of nature which has then been interpreted into the language of mathematics.

The **Principle of Least Action** is the core technology used by modern physicists which was only developed on the backs of Newton and Lagrange. The following information comes from the extrapolations taken to reach this crucial new principle.

The relationship Newton discovered between acceleration and force, expressed in his arcane, personal notation of fluxions, most impacted Leibniz. Leibniz used an easier differential notation, and one in which you may be familiar:

$$F = ma = m\frac{d^2x}{dt^2}$$

When Newton alleged that Leibniz plagiarized his calculus discovery, the vastly easier integral and differential equations fell out of favor in England. The majority of calculus advances took place in Germany and France over the 100 years.

While attending the University of Basel, Leonhard Euler developed the calculations of variations – this became an important tool for theoretical physicists allowing them to discover curves which met the minimal or maximal length (given a particular set of conditions).

Joseph-Louis Lagrange applied Euler's results to the mechanics of Newton's work. The **Principle of Least Action** emerged from the work of Lagrange and Euler. It is the core technology and basis of today's theoretical physics and gave rise to the following:

"Minimizing the action of the system
derives the differential equations of motion."

Put another way, how does an object move between its start point and its finish point? What is the motion of the object between start and finish (aka its worldline)? A Langrangian! A Langrangian is simply the difference between kinetic energy and potential energy.

$$S \equiv \int L dt = \int (T - V) dt$$

Derived from the **Principle of Least Action** are the equations developed by Euler-Lagrange. There are more than one, but below is a sampling for perusal. Naturally, it is also a differential equation. These equations describe motion within a system.

$$S(q) = \int_a^b L\big(t, q(t), q'(t)\big) \mathrm{d}t$$

Newton's equation of motion (in relation to free particles with no external force) is reproduced by this particular equation which minimizes the action.

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}_i}\right) - \frac{\partial L}{\partial x_i} = 0$$

This group of methods (shown above) is called the **Lagranian Formalism of Mechanics**. William R. Hamilton, a Dublin mathematician, created **Hamiltonian Formalism** in 1834 by applying Newtonian mechanics to his work in optics. Hamilton borrowed the idea of a function's value remaining constant.

This now symbolizes the system's total energy and is named after Hamilton. Hamiltonian formalism and Lagranian formalism are related through what is known as the **Legendre Transformation**.

The Legendre transformation changes one real-valued function of a real variable into another.

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

There is no need to use classical mechanics because Lagrangian formalism (and its use in relation to differential equations) has proven to be adaptable enough to study continuous media, such as the vibrations of uninterrupted η-dimensional objects (like 1D strings and 2D membranes) and the flows of fluids.

To extend Lagrangian formalism to systems which never stop, the Lagrange function is no longer integrated over just time. Instead, a

Lagrangian density is integrated over time and all of space.

This completes the jump to classical field theory from classical mechanics, and the equation is written as shown below:

$$S = \int L(q(x^2), \frac{\partial q}{\partial x^a}) d^{D+1}x$$

When it comes to continuous media, it is important to show the Euler-Lagrange equation as follows:

$$\frac{\partial f}{\partial y} - \frac{d}{dt}\left(\frac{\partial f}{\partial \dot{y}}\right) = 0$$

The Lagrange density's functional differentiation replaces ordinary differentiation. This is known as a field and is dependent upon time and space (known as field theory).

When it comes to importance, the first equation to be used was Newton's **Law of Gravitation** (as shown below):

$$F = G\frac{m_1 m_2}{r^2}$$

Newton's Law of Gravitation is where it all started, but it took another 200 years for the modern age of telecommunications to begin. In the 1800s, mathematicians and physicists were focused on understanding magnetism and electricity.

James Clark Maxwell discovered equations which were unified in the motion of magnetic and electric fields. These equations are known as **Maxwell's equations**:

$$\nabla \cdot E = \frac{\rho}{\epsilon_0}$$

$$\nabla \cdot B = 0$$

$$\nabla x E = -\frac{\partial B}{\partial t}$$

$$\nabla x B = \mu_0 J + \mu_0 \epsilon_0 \frac{\partial E}{\partial t}$$

Maxwell also devised that there are electromagnetic traveling wave equations as shown below:

$$\frac{\partial^2 E_x}{\partial y^2}+\frac{\partial^2 E_x}{\partial z^2}=\frac{1}{c^2}\frac{\partial^2 E_x}{\partial t^2}$$

In 1873, Maxwell hypothesized that these equations (and the electromagnetic waves which form part of these equations) provided the answer to the continuing question to light's true nature.

However, these were just mathematical and theoretical and in 1884 (possibly the most wonderful year for classical field theory).

In his lab, Heinrich Hertz studied and generated radio waves. He confirmed Maxwell's predictions and changed physics – and the world – forever. Maxwell's theoretical unification in relation to magnetism and electricity is now used to communicate (at the speed of light) across space.

This was a powerful (and stunning) achievement for both theoretical and experimental physics, and it shaped the future of telecommunications in the $20^{th}$ century.

This was the just the start. In the coming century, theoretical physics continued to challenge reality and space-time; the technological advances would become even more stunning.

Today, theoretical physicists and experimental physicists are divided into two distinct entities. Theorists are exploring parts of nature using mathematics for which there is no technology to test the validity of the theory.

Many theoretical physicists never lived to see their theories supported by experimentation. Current theorists are, as a result, forced to live with a total lack of certainty in their mission to use mathematical formulae to describe nature.

## What Types of Math?

Mathematics is the language of physics and, in order to really study physics, you need to know that the development of mathematics took extremely intelligent people 100s of years to work out.

For instance, geometry was once cutting-edge math, but now it taught as just a stepping-stone to higher mathematics.

### Algebra

Algebra is most people's first experience with using items such as constants and variables. It also gives you experience in solving and manipulating linear equations in the form of $y=ax+b$ or even more complex equations such as the quadratic formula ($ax^2+bx+c$).

### Geometry

This is no longer the geometry you learned in high school. At this level, the type of geometry you studied would have been Euclidian's two-dimensional geometry. Courses teach you how to think geometrically, use ideas like symmetry, and comprehend the properties of shapes in a flat, two-dimensional world.

### Trigonometry

This begins with the exploration of right triangles and Pythagorean's theorem (i.e. $a^2 + b^2 = c^2$). The functions cos, tan, sin are fully investigated.

### Calculus (single variable)

Calculus starts with the definition of an abstract function in relation to a single variable. It introduces the functions ordinary derived as the curve's tangent at any given point along that curve. This is known as **Inverse of Differentiation**.

### Calculus (multi variable)

This type of calculus introduces functions of more than one variable (sometimes many more than one) and gives you the chance to learn about total and partial derivatives. Within two and three-dimensional Euclidean space, the concepts of directional derivative, integration over a surface and integration along a path are developed.

## Analytic Geometry

This is the combination of algebra and geometry. Objects such as planes, spheres, and conic sections are taught by using algebraic equations. Vectors, spherical and polar coordinates are first experienced.

## Linear Algebra

This teaches you to solve linear equations and is expressed by vectors and matrices. All properties of matrices are fully studied.

## Ordinary Differential Equations

All the math up to this point has been preparation for physics. This is where physics begins. A lot of physics deals with solving and deriving differential equations, but the most important one to learn is the **Harmonics Oscillator Equation**.

## Partial Differential Equations

Since physics takes place in multiple dimensions, it is necessary to make use of partial derivatives and, as a result, partial differential equations. In early studies of physics, the first partial differential equations learned are the linear, separable ones derived in the 18$^{th}$ and 19$^{th}$ centuries.

## Methods of Approximation

Many problems arising in the area of physics can't be solved in an exact manner. Therefore, it is necessary to learn how to make clever approximations. These include power series expansion, saddle point integrations and perturbations.

## Statistics and Probability

This field of mathematics becomes much more important when quantum mechanics came on the scene. This course includes counting distinguishable objects vs. indistinguishable and man vs. variance.

## Particles and Relativity

Once electromagnetic fields were completely described through math, many physicists thought the field was dried up, with nothing left to explain. This led to two separate areas of thought – the geometric view and particle view.

The geometric view does an excellent job of describing gravity at astronomical distances. The particle view is equally good at describing three of the four forces of nature.

In 1897, J.J. Thomson's studies gave birth to electron and particle physics. Experimental physicists started developing an atomic model where the positively-charged core was orbited by negatively-charged particles.

This was a problem because Maxwell's equation stated the system shouldn't be stable. Classical field theory did not have the capability to explain the new data regarding atomic structure.

The thermal behavior of light was another mystery that arose from Maxwell's equations. Hot objects (such as a piece of charcoal) discharge light which is comprised of a distribution of different wave frequencies.

However, when physicists tried to use Maxwell's equations to describe the distribution of light frequencies, they were stymied.

At the beginning of the 20$^{th}$ century, a young, German physicist named Max Planck (in an act of desperation) made an attempt to explain the gaps in understanding regarding thermal radiation by sheer guesswork.

He called it **Quantum Hypothesis** and this theory states that the observable light spectrum is derived from a collection of individual, but identical quanta of energy. Thus began the theory of **Particle-wave Duality** and, further, **Quantum Mechanics**.

The newly learned **photoelectric effect** was explained by Einstein, using Planck's idea. Einstein's proposal theorized that light is discharged by excited electrons, in individual quanta (called photons).

The energy of an electron can be found with the following equation (where "h" is equal to Planck's constant which is known as the measurement of $6.6 \times 10^{-34}$ joule seconds):

$E = hV$

This equation gave rise to the following question:

"If a light wave can act like a particle, can particles act like waves?

Louis de Broglie came along with the idea that an electron can behave like a continuous wave and, subsequently, demonstrated it with this equation.

$$\lambda_B = \frac{h}{p}$$

As the arguments began to settle, they realized this new theory for quantum physics was describing a physical system as the **probability amplitude** for a system to be in a particular quantum state.

This amplitude is the square of a function and is referred to as a **wave function** and is also known as the answer to the **Schrödinger equation**:

$$i\hbar\frac{\partial}{\partial t}\Psi(\mathbf{r},t) = \frac{-\hbar^2}{2m}\nabla^2\Psi(\mathbf{r},t) + V(\mathbf{r},t)\Psi(\mathbf{r},t)$$

Solutions to Schrödinger's equations are interesting. The Schrödinger equation has the ability to predict what the distributions of probability are but, at its most basic level, cannot predict each measurement's exact result.

So, imagine a system containing two particles and trade those particles. The wave function will then obey this new relation:

$$|\Psi(x_1x_2)|^2 = |\Psi(x_2x_1)|^2$$

In the plus set, the two particles are bosons and, in the minus set, they are fermions. So, using the mathematics of experimental observation combined with the use of quantum mechanics, physicists learned that there are two classes of particles:

Fermions

Bosons

Bosons are the particles which transmit force and several bosons can (at any given time) occupy the same position.

The same cannot be said for fermions. Only one fermion can occupy any given position at any given time explaining why fermions are the composition particles for matter, and why solids never pass through each other.

The theory of **Pauli repulsion** defines matter's inability to share the same space and time as forces can. It also explains the structure of all the atoms in the periodic table of elements (in other words – all matter).

Particle physics was developed alongside quantum mechanics, and more observational evidence emerged that light (such as electromagnetic radiation) traveled in every direction at a single fixed speed (within a vacuum).

How do you this in physics terms? Einstein explained it via his **Theory of Relativity**. He used the geometric idea of a metric, similar to the Pythagorean Theorem.

Einstein's theory has a property that is **invariant under rotations** – this means that no matter how you rotate a line, the length of the line does not change. This notion of a metric is lengthened to include time with this formula:

$$dx^2 = \eta_{\mu v} dx^\mu dx^v = -(c \cdot dt)^2 + (dx)^2 + (dy)^2 + (dz)^2$$

Space-time is an object considered to be invariant under any type of rotation in space. However, it is also known to be invariant under a **Lorentz Transformation** (defined as a fresh kind of rotation for space and time).

This transformation tells us how individual observers see the world (observers who are moving with a velocity which is constant in relation to one another).

Under the Lorentz transformation, light's speed remains constant. With these observations made, quantum mechanics gave rise to something now known as RQFT (or **relativistic quantum field theory**).

For the latter half of the 20th century, physicists used the relativistic quantum field theory to describe subatomic particles and their behavior.

Einstein then moved on to gravity and tackled the Theory of Gravitation (developed by Sir Isaac Newton). Newton's constant is described as follows:

$$F = G\frac{m_1 m_2}{r^2}$$

Newton's theory successfully explained the movement of the sun, moon and other astronomical bodies, but it did not even hint at how gravitational fields change in time.

It certainly did not describe movement in a way that was in line with our brand new comprehension of Special Relativity – believing that no object has the ability to travel faster than the speed of light.

As a result, Einstein created a crazy new type of mathematics known as **non-Euclidean geometry**. This type of geometry changes the Pythagorean theorem allowing the inclusion of metrics which depend on space-time.

He extended this to include Lorentz's invariance and, as a result, created **general coordinate invariance**. By extending his theories to include Newton's theory of Gravitation, he produced his General Theory of Relativity.

This brought to physics the mathematics of **differential geometry**.

In Einstein's theory, space-time may have a curvature. Previously, it was believed that space-time was flat. In his equation:

$$R_{\mu v} - \frac{1}{2} g_{\mu v} R + g_{\mu v} \wedge = \frac{8\pi G}{c^4} T_{\mu v}$$

The equation below demonstrates that the space-time curvature ($R_{\mu v}$ and R) is found by using the momentum and total energy ($T_{\mu v}$) of all the items in space-time. Objects such as black holes, planets, radiation, etc. all make a curvature of different sizes, in space-time.

Many observational successes have proved General Relativity worthy as a description of nature; however two predictions coming from this theory have astounded both the public and scientific communities alike:

The universe is expanding

There are black holes

Both have now been observed, but the issues they encompass, at least mathematically, challenge the very nature of reality and existence.

## Differing Views

When working with quantum mechanics, many of the most typical questions answered have been concerned with quantum state types. This has allowed for transitions in a system featuring one or more particles with some type (or level) of potential energy.

Typically, the Schrödinger equation is used to find the wave function, the energies for each quantum state and any allowed transitions between the states.

Things are quite different in general relativity. Generally calculations are performed which help to find the answers to the structure and evolution of an entire universe at the same time.

The equations used turn into a system of 2$^{nd}$ order non-linear differential equations with their solutions yielding the metric of space-time.

Given the fundamental difference in questions asked and methodologies that each physicist uses between general relativity and quantum mechanics, the need to unite gravity and quantum physics with gravity would prove very difficult.

## Why Did Strings Enter the Picture?

Relativistic quantum field theory worked well in describing the behaviors of elementary particles, but the theory only works when gravity is so weak that is can be forgotten.

If you can pretend that gravity doesn't exist, it is at this point theoretically that particle theory works.

Basic relativity has produced many revelations about the universe, but it only works when we pretend that quantum mechanics is not needed to describe nature.

Obviously, there was a gap in our knowledge and **String Theory** is believed to close the gap.

It soon became apparent that Schrödinger's quantum mechanics equation wasn't Lorentz invariant. Thus, although successfully tried and tested in the 1920's, quantum mechanics wasn't an accurate enough description for nature because the system contained particles with the ability to move close to the speed of light.

The difficulty is that the Schrodinger equation is 2nd order in spatial derivatives but 1st order in time derivatives. An equation that was the same order in both was needed.

The Klein-Gordon equation is 2nd order in both space and time **and** has answers with particles with 0 spin. It is shown below:

$$\left(\hbar^2 \nabla^\mu \nabla_\mu + m^2 c^2\right)\phi = 0$$

Dirac did further work with the equation and came up the "square root" for Klein-Gordon's equation. He used gamma matrices. The solutions resulted in particles with a spin of ½:

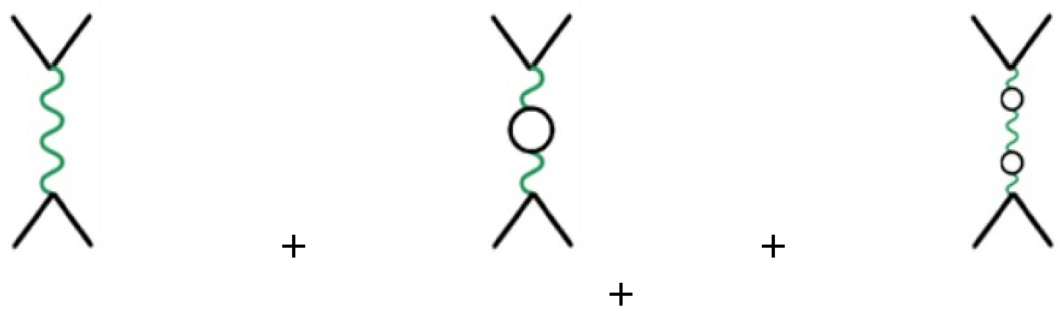
$$(\not{p} \pm mc)\psi = 0, \not{p} = y^\mu p_\mu = i\hbar y^\mu \nabla_\mu$$

$$\{\gamma^\mu, \gamma^v\} = \gamma^\mu \gamma^v + \gamma^v \gamma^\mu = 2\eta^{\mu v}$$

However, one large problem existed with this type of quantum mechanics. The answers of both the Klein-Gordon and Dirac equations have instabilities which represent the annihilation and creation of virtual particles from empty space.

Feynman, Schwinger and Tomonaga developed RQFT in the 1940's. Feynman diagrams are used to explain the behaviors of elemental particles. These diagrams properly calculate for the annihilation and creation of these virtual particles that proved to be so problematic previously.

Here is a set of Feynman diagrams showing the scattering of two electrons.




The black lines which look like a "v" represent electrons. The curved green lines represent photons, or the electromagnetic field present between electrons (in more classical terms).

Each black loop shows a photon creating positron electrons which, later, annihilate each other and create a photon. This is known as a virtual process. The amplitude shown of these scattering electrons is the aggregate of all contributions for all possible loops.

The quantum loop process also gives us a very big problem. In order to ensure that all virtual processes are accounted for, you must use every possible momentum value, from infinite momentum to zero momentum.

However, the loop integrals for a particle in D dimensions with J spin take an approximate form:

$$I_{loop} \sim \int p^{4J-8} d^D p$$

If the element of "4J+D-8" is negative, the integral will then act appropriately for levels of infinite momentum. However, the calculations lose sensibility because they give infinite answers if the quantity is "Ø" or positive.

The world we live in has "D=4", and a photon has a spin of "J=1" meaning that, in this instance of electron-electron scattering, the loop integrals are able to have infinite values.

However, they travel to infinity slowly and are able to be renormalized allowing the infinities to be reabsorbed into a small number of parameters. These parameters could be the charge or mass of the electron.

With the problem of the creation and annihilation of virtual particles solved, theoretical physicists began trying to solve other difficult problems. In the 1960's, particle physicists began to take on the strength of the nuclear force and posited a dual resonance model.

It wasn't that successful, but by the 1970's they realized that these dual models were really the quantum theories for relativistic vibrating strings. Dual models eventually became known as string theory and display some interesting mathematical behavior.

In 1971, a new category of quantum field theory was developed which analyzed the weaker levels of nuclear force. This united electromagnetism with the weak force and produced what we call the "electroweak" theory.

Taking what they just learned, they produced quantum chromodynamics (aka QCD) by applying the new knowledge to the stronger nuclear force. QCD was a renormalized theory.

This left a single force – gravity – which could not be renormalized. Physicists tried for a very long time but there was one major issue. Classical gravitational waves have a spin of "J=2", therefore a graviton's spin should also be "J=2". A graviton is defined as the quantum particle carrying the gravitational force.

However, for "J=2" and "D=4", the loop integral for the gravitational force (i.e. "4J-8+D=D") became infinite, just like the $4^{th}$ power of momentum. This was not okay!

However, particles in string theory arise from excitations in the string, and one of those excitations is a particle with zero mass and a "J=2" spin. If there did exist a suitable gravitational quantum theory, then the particle which carried the force would have a mass of zero and a "J=2" spin. The question was finally asked in 1974:

"Is string theory simply a theory of quantum gravity?"

This insight led early string theorists to shift the application of string theory from the strong nuclear force to quantum gravity. But it wasn't enough that a graviton was foretold by string theory. Anyone can add a graviton to the equations used, but that doesn't mean that it is observable in nature.

One major benefit to string theory is the analog of the Feynman diagram. In string theory, the integrals loop over an incredibly sleek surface while the loop integrals lack the infinite momentum, zero distance of the particle loop integrals. This means that infinite momentum does not necessarily mean a distance of zero.

String tension and the related parameter **α'** (pronounced alpha prime), are fundamental to string theory. String tension requires minimum length of observable time for a quantum string.

The behavior of zero-distance issues that plagued quantum field theory has now become irrelevant to string theory. This makes string theory a very attractive applicant for the theory of quantum gravity.

However, if string theory is going to explain quantum gravity, then the minimum length scale must be Planck's length. This is the length scale made by combining the speed of light, Planck's constant, and Newton's constant:

$$L_P = \sqrt{\frac{\hbar G_N}{c^3}} = 1.6\ x\ 10^{-33} cm$$

Bridging the gap between general relativity and quantum mechanics was a huge obstacle to be overcome and this was achieved in the physics of the late 20th century.

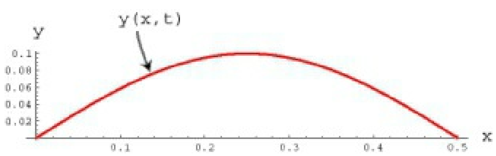
The mathematics is complex and abstract but many young physicists learn it because it is necessary in order to study interacting strings and the quantum theory behind them.

## So What is String Theory Then?

Picture in your mind a violin that has been tuned. The strings are taut, under tension across the violin. Depending on the amount of tension on the string, and how those string are plucked, a variety of musical notes can be created.

These notes could be the **excitation modes** of the violin string under tension.

Pythagoras actually described the motion of his lyre's strings and, as a result, today we have more precise mathematics to explain it:




**x** This equals the distance along the string

**y** This equals the height of the string as the string vibrates (or oscillates)

**t** This equals time

The equation for motion is known as the 1D wave equation:

$$\frac{\partial^2 y(x,t)}{\partial r^2} = \sqrt{\frac{T}{\mu}}\frac{\partial y^2(x,t)}{\partial x^2} = v_W{}^2\frac{\partial^2 y(x,t)}{\partial x^2}$$

When solving motion equations, it is important to understand the boundary conditions (i.e. we need to know how long the string is).

Suppose the string has been fixed at both ends and its length is unstretched. The equation would be written as follows:

$$y(x,t) = \sum_{n=1}^{\infty} (a_w \cos\frac{n\pi v_w t}{L} + b_w \sin\frac{n\pi v_w t}{L}) \sin\frac{n\pi x}{L}$$

A normal mode has a wavelength of some integral fraction which is twice the wavelength:

$$\lambda_N = \frac{2L}{n}$$

The frequency of this normal mode can then be displayed as:

$$f_n = \frac{n v_w}{2L}$$

Normal modes are what, when the violin is played, we hear as notes. As the tension of the string increases, the string wave velocity also increases, in turn increasing the string's normal frequency.

That is why a violin string sound gets higher as the string is tightened.

In string theory, we observe elementary particles in particle accelerators. These particles are the excitation modes of elementary strings. Similar to violin playing, in string theory, the string must be under tension to get excited.

However, in string theory, the strings are floating around in space-time. Because these strings are relativistic, the equation needs to use coordinates that have a Lorentz transformation. This presents an issue because the strings are floating in space-time. They oscillate in time and space and sweep out a 2D surface in space-time. This is known as a world sheet (as opposed to the world line) of a particle).

In a relativistic string, we have to consider the string's **world sheet** as a **2D space-time** of its own. In this manner, the division between time and space depends on the observer. The equation for motion of relativistic strings is:

$$X^i(\sigma,\tau) = x^i + x^i\tau + i\sqrt{2\alpha'} \sum_{n \neq 0}^{\pm\infty} \frac{1}{n} \alpha_n^i \left(\cos\frac{n\pi c\tau}{L} - \sin\frac{n\pi c\tau}{L}\right) \cos\frac{n\pi\sigma}{L}$$

Since strings are about the size of the Planck length (or one millionth of one billionth of one billionth of one billionth of one centimeter), physicists must come up with some very clever ways of studying them.

There is more than one type of string theory and they are categorized according to two variables – if the strings are closed or open and if the particle spectrum also includes fermions. If they are included, then the theory must have **supersymmetry**. Supersymmetry means that every boson has a fermion.

Therefore, supersymmetry finds a relation between the particles that make up matter and particles that transmit force.

Supersymmetric partners have not been seen in particle experiments. However, theorists are not concerned. They think this is caused by supersymmetric particles being too massive to detect at current accelerators.

However, in the coming decade, current accelerators could find high-energy supersymmetry. This would be good corroboration that string theory is a relevant mathematical model for nature, at least on small-scale distances

The above equation for motion of relativistic strings isn't tied down on either end, making it floppy and open. Closed string boundaries are considered to be periodic and the solution to this oscillation looks just like two open-string oscillations but moving in opposite directions around the string.

This produces right-movers and left-movers. This difference becomes very important later in the string theory for **Supersymmetric Heterotic**.

Boundary conditions are of extreme importance for string behavior. Strings can be open (having ends which travel at the speed of light) or closed (having their ends joined up in a circle).

Within a closed string, only one of the particle states has zero mass and two units of spin (in the same way as a graviton).

## How Many String Theories are There?

To build a string theory, you must start with the basic ingredient – a teeny wiggling string. Next, you need to ask:

"Should the string be open or closed?"

The next question should be:

"Will I settle for just bosons or do I also need fermions?"

If you only want bosons, then you get bosonic string theory. When your requirements demand fermions as well, then you will need supersymmetry. This theory will be called superstring theory.

The final question you should ask is:

"Can I perform quantum mechanics sensibly with this theory?"

For bosonic string theory, the answer is only **Yes** if there are 26 space-time dimensions. For superstring theory, the number of dimensions decreases to 10.

Working down far enough to get to the four dimensions of we observe in our world is much more difficult.

There are four types of bosonic string theories, all of which need 26 space-time dimensions to propagate. The string theory produced is dependent upon the boundary conditions employed for solving the equation of motion.

There are two categories of choices:

Are the strings closed or open?

Are the strings unorientable or orientable? Orientable means you can determine in which direction you are traveling on the string.

There are four possible combinations of these variables and they produce the four different categories of bosonic string theories. All of these theories are considered to be unstable as they have the lowest mode of excitation.

This excitation mode is a tachyon with a level of $M^2=-1/a'$ and it does not lend stability.

Gravity is always included because the massless particle spectrum always includes a graviton. Even though this theory is unstable, this is acceptable because it is not a sensible theory.

The physical world has matter and is therefore made up of fermions. Since bosonic string theory excludes fermions, it cannot accurately describe observable nature.

The next step is to add fermions to this string theory in order to satisfy the fact that there are both forces and matter is nature. However, adding fermions imports a fresh set of negative norms (i.e. ghosts).

String theorists determined that these undesirable ghost states uncouple themselves from the string theory spectrum when two conditions are met:

There are 10 space-time dimensions

The theory is supersymmetric

Bosons have simpler boundary conditions than fermions, so learning all the various possible superstring theories engaged them for quite a while.

The easiest way to inspect a superstring theory is to journey to what is known as superspace. In superspace, you will find a set of anti-commuting coordinates and a set of normal commuting coordinates.

There is still the option of closed versus open and unoriented vs. oriented. However, now there is the option of fermions to help differentiate one superstring theory from another. The superspace coordinates can only spin two directions known as handedness:

The spin axis will be going in the opposite direction of momentum

The spin axis will be going in the same direction of momentum

None of the superstring theories have problems with tachyons (which can really present themselves in bosonic string theories) and all of these theories contain gravity.

The five superstring theories (with their combinations) seemed to give the only theories workable in the real world and the quantum world. Then, in 1984, a strange idea occurred – why not combine a bosonic theory and a superstring theory?

This required the use of a closed string from superstring theory, with either handedness, and a string from bosonic string theory. Combine a bosonic string, moving in a specific direction with a supersymmetric string, moving in the opposite direction.

These combinations are known as heterotic superstring theories. Bosonic strings can only exist in 26 dimensions and supersymmetric strings exist only in 10, so it seems crazy to combine them.

However, those extra dimensions aren't really considered to be space-time dimensions.

There are only two kinds of heterotic theories which prove quantum mechanics as consistent:

SO(32)

$E_8XE_8$.

The second theory give above was considered to be the **only** string theory which gave credence to the realism of physics. However, during the 1990s, more possibilities based on other studies and theories were discovered.

## How are String Theories Related?

Once upon a time, theoretical physicists thought that of the many different superstring theories, the ones listed below were the only workable theories:

Type I

Type IIA

Type IIB

SO(32)

$E_8 X E_8$

They believed that, out of this list, one of these types would be the correct **Theory of Everything**. That would be the one whose low-level energy limit, and ten dimensions would, one day, compact down to the four dimensions which matched the physics in this world.

The alternative theories would simply be theories that were rejected because they did not happen to mathematically describe nature.

However, this view has come to be seen as naïve. Physicists now realize the five major theories are, in fact, interconnected. It's almost seems like each is part of a more fundamental hypothesis, and there is really only one.

For instance, if two of these theories are connected by a duality transformation, it means that one of the theories would be transformable to look identical to the other theory. These theories are referred to as dual to each other.

Quantities are linked within these dualities which physicists thought were previously separated:

Weak and strong coupling strengths

Large and small distance scales

In both quantum particle physics and classical field theory, these quantities mark the specific limits of action within a system. However, as mentioned previously, strings display some curious behaviors.

One of these behaviors is the ability to confuse the differences between strong and weak, large and small, etc. It is in this way that the five extremely different string theories are actually related and can form into a single larger fundamental theory.

## Large and Small distances

How can strings make a large distance look small? There is a **T-duality** that confuses our ability to differentiate between small and large distances.

This duality occurs because of the compactification into ten space-time dimensions to allow for superstring theory. Imagine you are in those ten dimensions which are made up as follows:

9 space dimensions

1 time dimension

Make a circle out of one of the nine dimensions, and then imagine a particle traveling around the circle. The particle must travel in a certain direction for a distance (as shown below) in order to get back to the starting point.

$L = 2pR$

The total energy of the particle is determined by its quantized momentum as it circumnavigates the circle.

However the strange thing regarding string theory is that the winding modes and momentum modes can be interchanged so long as the radius "R" of the circle is also exchanged with the quantity "$L_{st}^2/R$".

A string behaves differently because a closed string can travel the circumference of the circle, like a particle, but it can also wrap around the circle. The number of times the string wraps around the circle is known as the winding number (which is also quantized).

So exchanging a winding mode with a momentum mode of the string exchanges a scale of a long distance with one a small distance. This is the T-duality! This relates one superstring theory to another meaning that we can:

Take two superstring theories

Compactify them on a circle

Switch the winding and momentum modes

This will switch the distance scale and will then change one theory into the other! T-duality takes Type IIA superstring theory and relates it to Type IIB superstring theory.

This also works with two heterotic theories – heterotic SO(32) superstring theory relates to heterotic $E_8XE_8$ superstring theory. The T-duality is in direct opposition to how physics looked and worked in the days of Newton and Kepler; however it is a reasonable outcome for a quantum theory of gravity, since gravity originates in the metric tensor field that lets us know the distances between events in space-time.

## Strong and Weak Coupling

A coupling constant tells us how strong an interchange is. For instance, Newton's constant tells us how strong the gravitational interaction is with every force having a coupling constant.

Quantum physicists study the electromagnetic force that has a coupling constant is direct proportion to the square of the electric charge. Physicists can't calculate the entire action of electromagnetism so they break it down into pieces, each of which has a different power of the coupling constant ahead of it.

If this coupling constant stays small, the different smaller equations make a good approximation. But as this constant gets larger, it ceases to be an approximation of real physics.

The same thing happens in string theory because string theories possess a coupling constant. String theories have two types of expansion:

An expansion in powers of the string parameter

A quantum loop expansion for string scatter amplitudes in d-dimensional space-time

The coupling constant is not a number, as in particle theories, rather it is one of the string's oscillation modes and is known as a dilaton. You can exchange a large constant for a small one, by interchanging the dilaton field minus itself.

The theory with the strong coupling constant can't be appreciated by means of expanding a series, the weak coupling theory can. So, we only need to understand the weak theory and that is the same as understanding the strong theory.

This allows an understanding of the strong coupling limit that would be impossible by other means. This is S-duality!

As with T-duality, some string theories share duality that are unexpected. For instance, the Type I theory is S-dual to the SO(32) theory. This is surprising because Type I has both open and closed strings while heterotic theories only have open strings.

Why does this work? At very strong coupling limit, SO(32) string theory has open strings. But these same strings are highly unstable in the weak-coupling limit of the theory; this is the limit in which heterotic string theory is understood.

## What Does It Mean?

The T-duality and S-duality are unique to string theory. Particles do not have the ability for these types of duality because they are not able to be wrapped around a circle. If string theory is an accurate description of Nature then, on a deeper level, the difference between large and small distance scales is not fixed but fluid.

Distance may actually depend on the types of probe used to measure and how we measure the state of those probes.

The S-duality gives the impression that it goes against the grain of traditional physics but it is actually quite a reasonable outcome for gravity's quantum theory.

Einstein's findings gave us the theory that gravity is all about how an object's size is measured and that the interaction magnitudes are measured in curved space-time.

Einstein's theory of gravity has already proven to be a reasonable theory of gravity so the S-duality is not so radical after all.

# Is There a More Fundamental Theory?

There were more surprises to come from string theory. One was that these theories were not just applicable to 1D objects, as previously thought. The additional higher dimensional objects with dimensions from 0-9 are also included within this theory.

To understand these, it is beneficial to understand Maxwell's equations in differential form because these tell us the sources of charge are 0D objects.

According to Maxwell's equations, gauge field strengths that are "**p+2**"-forms, have sources that are p-dimensional objects. These are known as **p branes**.

Mathematics informs us Maxwell's equations couple electrically to points, or **zero branes**, as we call them in string theory.

The same math is effective in any space-time dimension, so we also know Maxwell's equations couple to point charges in all space-time dimensions.

What is a p brane composed of? It is a space-time object which provides the answer to Einstein's equation for the superstring theory low energy limit. The non-gravitational fields' energy density is confined to a p-dimensional subspace of the 9 dimensions of space.

An example follows. In an electrically charged solution with the electromagnetic field energy density distributed along a line in space-time, the 1D line would be considered a p brane with a value of "**p=1**".

P brane space-time is unstable. Supersymmetry can stabilize the p branes for a little while. Two of the most important p branes in string theory are:

d=11 (a two-brane)

d=10 (a five-brane)

It took quite some time to discover d branes because they were buried deep within the mathematics of the T-duality but were discovered while studying open strings in T-duality.

Unlike closed strings, open strings are not equipped with winding modes around some of the more compact dimensions.

This may lead you to think that, in the presence of circular dimensions, they behave in the same manner as particles. However, open strings have a very interesting stringiness within a circular dimension's proximity.

Normal open string boundary conditions occur due to the requirement that no momentum enter or exist from the ends of an open string. These boundary conditions are called the **Neumann Boundary Conditions** meaning that open strings with their ends fixed to static p branes behave differently and are then called d branes.

A d brane is like a collection of excited strings. They are important because they help our understanding of black holes; especially the counting of quantum states which led to our understanding of the black hole entropy. More on black holes later!

## How Many Dimensions?

In the early days of string theory, the small community of theoretical physicists did not pay it much attention.

The 11-dimension theory of supergravity was the most accepted string theory whereby gravity is combined with supersymmetry. This 11D space-time was then compactified onto a small 7D sphere, leaving only 4 observable dimensions.

However, this theory could not become a unified particle physics theory due to having an unreasonable limit in the theory of point particles.

The 11D theory was eventually abandoned; however it had a habit of resurrecting itself as a contender for superstring theory within the ten dimensional arena.

How does a superstring with 10 dimensions of space-time turn into a super-gravity theory with 11 dimensions instead? We have already learned that superstring theories and the duality between them can produce very different theories.

So some duality relationship must exist that explains how the 11D theory and the 10D theory are really equated. We already know that there is a link between the various string theories, and we suspect that more individual limits of a more fundamental theory exist.

It was, therefore, believed that this other theory could exist within the 11 space-time dimensions? This belief became known as the **M theory**.

## M Theory

In scientific terms, **M theory** was the unknown 11-dimensional theory which posed a low limit of energy, and would then become the super-gravity theory in the 11 dimensions discussed earlier.

But many have started working with the M theory to identify any unknown theory pre-supposed to be part of the fundamental theory of everything.

We don’t know much about the M theory basics, however we have learned quite a bit about the 11D M theory. Within this theory, there exists extended objects referred to as **M branes**, not d branes.

A specific type of M brane, in this theory, has two space dimensions and is so-called an M2-brane.

Contemplate M theory with the 10$^{th}$ dimension of space-time compactified into a circle. If we wind around a circle, a space-time dimension that consists of this M2 brane, then we get a result which is equal to a one brane of the superstring theory Type IIA.

# Cosmology

## How Old is the Universe?

The universe’s age has been shrouded in mythological, scientific and religious inquiry for centuries. Sir Isaac Newton postulated that the Universe was only a couple thousand years old. Einstein believed it was eternal and ageless.

However, in 1929, observational evidence verified that neither man was correct.

To comprehend this observational evidence, you need to understand something called the **Doppler Shift**. Imagine listening to a train as it approaches and leaves the station.

A train approaching the person would start at a low pitch and rise to a higher pitch. A train leaving the person would start at a high pitch and descend to a lower pitch. This pitch change, depending on its location to the listener is known as the Doppler shift.

A Doppler shift not only happens with sound but also with light. Light that approaches Earth will appear to have a higher frequency level than light that recedes.

So, in 1929, observation of remote galaxies taught us that the light from these galaxies was receding. If **all** of these distant galaxies observed were moving away from us, then the universe (as a whole) was also moving away from us; it was expanding.

Since the universe expands, where did it start expanding from? This expansion leads us to believe that the universe has a bounded age and is not ageless and eternal, as Einstein previously thought.

So, what **is** the universe's age?

From radioactivity studies, we know that the Sun, Earth and our entire solar system were formed approximately 4 1/2 billion years ago. The universe has to be a minimum of twice that age because our solar system did not have the ability to form until the Milky Way had taken shape.

The formation of the Milky Way took several billion years. Because we do not have the ability to perform radioactive testing on the more remote galaxies and stars, we need to perform other types of measurement.

Instead, we measure the distance and brightness of stars and their light's red shift in order to make an approximation. The most mature star clusters, so far discovered, are 12-15 billion years old.

It is safe to estimate that the universe has a minimum age of at least 15 billion years, but we do believe that it is any more than 20 billion years old.

There is still a lot of research being done into this topic by cosmologists and astrophysicists. There will likely be radical discoveries in the near future.

# Structure of the Universe

## What is Space-Time Geometry?

Imagine a very big ball. Your brain interprets the ball in a 3-dimensional world, however, the outside of the ball is considered to be 2-dimensional due to the fact that along the surface, there are only two directions of motion independent of one another.

Now imagine you're quite tiny and that you live on that ball surface. You may think you weren't on a ball at all, but on a big level 2-dimensional plane; in much the same way as humans thought that the Earth was flat. It doesn't look round unless seen from outer space.

However, if you carefully measure the outer distance on the ball, you would soon learn that you weren't living on a level surface at all; rather you're living on the curved surface of a gigantic sphere.

The same idea of the ball's curvature can apply to the entire universe at once. Einstein's Theory of Relativity provided that very breakthrough. Space and time are synthesized into one geometric entity referred to as space-time.

Space-time has geometry and it may have a curved surface, just like that ball.

When you look at the surface of a big ball as one entity, you experience the whole space of the sphere at one time.

Mathematicians don't like to describe a part of the sphere; rather they prefer to describe the whole sphere.

The tricky aspect of describing space-time is that you are describing the whole of time and the whole of space in a single mathematical entity.

## What Determines Space-Time Geometry?

Theoretical physicists try to find the motion equation which best describes what they observe in nature. The classical equation used for the motion of space-time is known as the Einstein equation.

Quantum behavior is not figured which is why this is a classical equation. Space-time geometry is treated classically, without any of the probability-based quantum behavior. It is therefore, at best, only an approximation.

The Einstein equations tell us that the space-time curvature in any given direction is directly proportional to the momentum and energy of everything in space-time that isn't actually space-time.

This means that the Einstein equation is used to tie non-gravity to gravity and non-geometry to geometry.

## What’s the Geometry of OUR Space-time?

Space-time includes the entirety of space and the entirety of time. So, space-time holds the aggregate of what happened and the aggregate of what is to happen. It holds **everything**.

This could get us into trouble as we aren’t able to track everything that has ever happened or will happen that could alter the distribution of momentum and energy within the universe.

However, we can use approximation and abstraction to work out what the real universe would be like by using abstract models. This works pretty well at great distances.

However, to solve these equations, certain assumptions need to be simplified in order to better understand the space-time curvature.

We assume that space can be separated into space and time (although this may not be the case around black holes where they get all twisted up, but it works over the rest of the universe). So we describe space-time as space ever-changing with time.

The next assumption we use is what helps us to understand the theory behind the Big Bang. We assume that space appears to be the same in every direction at all points.

Isotropic can be defined as appearing to be the same in every direction.

Homogeneous can be defined as appearing the same at every point.

We can then make the assumption that space is both isotropic and homogeneous.

Finally, cosmologists think about the three energy types that have the ability to curve space-time:

Vacuum energy

Radiation

Matter

Once all of these assumptions have been made, Einstein's equation then comes down to two basic differential equations – both of which are easy to solve with some basic calculus knowledge.

These equations tell us about geometry and how the mass of space changes over any given period.

## Open, Closed or Flat?

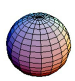
Since space appears to be the same at every point, it must have a constant curvature and that curvature must be the same at all points. This assumption simplifies the theory down to the following three options:

Zero curvature

Negative curvature

Positive curvature

When there is just radiation or matter present (i.e. no vacuum energy), we can calculate the time evolution of the space-time from the curvature of space.

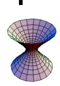


**Positive** – The cosmological scheme in which space has a constant curvature which is positive is referred to as a **closed** universe.

This universe looks like a ball. Space expands from a volume of zero and grows in the same way as a Big Bang event until it eventually reaches its maximum capacity. It then starts the Big Crunch and contracts back to zero volume.

**Zero** – Space with a curvature of zero is appropriately dubbed as a **flat** space and can also be known as an **open** universe. A flat space extends in an infinite route in all directions and is non-compact; it expands infinitely in time.

**Negative** – Infinite volume is also found is space with a negative curvature and is known as an **open** universe. Space-time also expands infinitely in time within this universe.



Another way to think of this is to imagine two cones attached to each other by their pointy ends.

So, what defines a universe as closed or open?

A flat universe has a critical density equal to 1

A closed universe has a critical density greater than 1

An open universe has a critical density less than 1

To discuss which one describes our universe, it is important to first discuss dark matter.

## Dark Matter?

The matter within the universe is made up of stars or other objects that emit lights of a wavelength that is detectable by our telescopes, eyes or some other instruments.

However, in the last few decades, astronomers observed evidence of another kind of previously undetectable matter.

For instance, to gravitationally bind most galaxies would take more matter than the visible matter we have observed.

In fact, astronomers and physicists now think the majority of universal matter is invisible (aka dark matter). This information is of extreme importance to cosmology.

So, what is it composed of? If it is composed of quarks (in the same way as ordinary matter), then the early days of the universe produced more deuterium and helium than can currently exist in the UNIVERSE.

Particle physicists think dark matter may be particles of a supersymmetric nature which are extremely heavy; however, these do not couple well to the particles we now observe in accelerators.

The observable matter in our universe consists of a far lesser amount than closure density, so if nothing else existed, we would have an open universe. Is dark matter enough to bind our universe as closed?

Studies indicate that even with the inclusion of dark matter, the total amount of matter would only amount to an approximately closure density level of 30% – making it open, right?

The cosmological constant also needs to be considered.

## What is the Cosmological Constant?

Einstein wasn’t always pleased with the end product of his work. He didn’t like that the universe was expanding and so tried to fix it with a cosmological constant, which fixed his equations so that the universe remained the same size.

Pictures from Hubble showed the universe was certainly expanding, and the cosmological constant was abandoned. But it was reawakened by relativistic quantum physicists because sometimes this constant would occur naturally from virtual particle quantum oscillations.

This may be the origin of the vacuum energy of space-time. However, one difficulty in quantum theory is preventing vacuum energy overabundance which makes it one of the reason why physicists examine theories of sypersymmetry.

This cosmological constant acts to either increase or decrease the expansion of the universe and is dependent upon whether it is negative or positive.

When a space-time containing radiation and matter are added (in addition to the cosmological constant), the scenarios of open or closed become overly simplistic.

## What's the Final Answer?

Our latest measurements tell us a few things about our universe.

**It is actually amazingly flat** – As the universe cooled following the Big Bang, it didn't do so in a perfectly smooth manner meaning that there are bumps in the surface of the universe.

However, we can tell that the universe is otherwise flat and should expand forever.

**There is a cosmological constant** – This consists of an energy vacuum, or a substance exactly like it. This means that there is a cosmological constant to make the universe expand in time. We observe this in a distant supernovae red shift.

**Most of the matter in the universe is invisible and dark (i.e. dark matter)** – Studies of galactic motion indicate that the visible matter in planets, stars, interstellar gas, and galaxies only account for a small portion of the overall energy density of the Universe.

Currently, the density of vacuum energy is approximately twice as massive as the density of the vacuum energy emitted from dark matter. Visible matter is almost negligible.

## The Big Bang

The Big Bang is a term for the beginning of the universe's expansion.

The current astronomical and physics-based knowledge suggests that the universe began many billions of years ago.

This beginning arose from a state of it being very small and very hot. It then mushroomed and evolved into what we know as the universe today.

Our understanding of the Big Bang theory is a combination of the Theory of Relativity and our knowledge of particle physics. Einstein's equations tell us how quickly the universe should expand, to what size and at what time. Our estimates about the density of both radiation and matter of our early universal beginnings come to us from ancient light that reaches our night skies.

Here is a walkthrough of the Big Bang as we currently understand it.

### The Beginning: Plank Era and Inflation Era

**TIME:** Planck Time: about $10^{-43}$ seconds

If string theory is correct then, at this point, the universe doesn't have much in the way of geometry. Duality relations dictate that space-time geometry is not fundamental; rather it emerges as we zoom out to distances larger than the Planck length.

After the Plank era, cosmologists believed there was a period of inflation. Both the physics of the Plank Era and Inflation Era were difficult to understand through our, then, current knowledge of physics.

We simply were not able to perform physics at sizes smaller than a Plank length. These two ages are filled with mystery.

## Radiation Fills the Universe

**TIME:** between $10^{-12}$ and $10^{-10}$ seconds

The Big Bang theory's official beginning was at this point. At the end of the inflation Era, the universe ended up being left in a dense, hot and small quantum state. The vacuum energy changed into a smoldering photon soup, gluons and other basic particles.

According to Einstein's general relativity equations, universal expansion has the ability to occur by energy density in the form of radiation and matter.

During this initial Big Bang phase, the part of energy consisting of radiation was so much larger than the matter of energy density that we can disregard matter for a little while.

## Quarks Outnumber Anti-quarks

**TIME:** $10^{-11}$ seconds

The teeny universe (still expanding) was filled with particles and anti-particles, which would annihilate back into radiation. Quarks and anti-quarks were created and annihilated back into radiation at a high rate.

However, as the Universe expanded and cooled, a larger number of quarks were left over than anti-quarks.

We are sure of this series of events because quarks (today) are clearly more numerous than anti-quarks. However, this dominance of quarks can't be explained using physics' Standard Model of Particles and we have not been able to study it in accelerators yet.

Continued study of the early Universe might lead to greater understanding of particle physics.

## Weak Nuclear Bosons Become Massive

**TIME:** $10^{-10}$ seconds

At this point in time, the universe is expanding and cooling and the particle density average is dropping to the levels of the weak nuclear force. Now, something outstanding happens to those particles which have been transmitting that weak nuclear force.

Elementary particle physics has taught us that at a specific energy scale (the weak nuclear force's energy to be exact), bosons become very heavy.

Above that energy force, they are massless, but below that energy scale they are very massive and heavy with a weak nuclear force that can only act over small distances.

As a result, cosmologists think that when the energy levels dropped below the scale of the weak nuclear force, the bosons became heavy and no longer had a range of infinity. Now they were slow and their force was restricted to a small range.

## Quarks and Gluons are Confined

**TIME:** $10^{-4}$ seconds

The universe is now much cooler and quarks and gluons do something incredible. Up until this point, they have been zipping around without constraint.

They go through a universal-wide transformation where every quark and gluon becomes confined together inside baryons and mesons, such as neutrons and protons. Before this, protons, neutrons and mesons didn't exist.

Particle physicists have only ever measured gluons and quarks trapped inside mesons and baryons. No one has ever assessed one zipping around by itself.

But the Quantum Chromodynamics (QCD) theory postulates that at a high enough temperature, gluons and quarks become de-confined and can roam around freely.

## Proton and Neutron Ratio is Fixed

**TIME:** 1 second

By the end of this time, what we have remaining is a level of about 7 protons for every neutron. Particle physicists have known for a while that neutrons which sit around doing nothing will simply decay into the following items:

A proton

An electron

An electron antineutrino

However, a proton sitting around doing nothing will not decay at all.

We can **force** protons to change by hitting them with extra energy, but neutrons will change all by themselves.

As the universe cooled, there was no longer any additional energy which would allow protons to change to neutrons. However, neutrons could still spontaneously change into protons.

In the end, this left us with 7 times more protons than the neutrons available. Because a hydrogen nucleus has one proton and no neutrons, and the nucleus of helium has two protons and two neutrons, an excess of hydrogen in relation to helium became a direct consequence of the Big Bang.

## Protons and Neutrons Form Nuclei

**TIME:** 100 seconds

At this point the average temperature is cool enough for neutrons and protons to stick together. They can now make the nuclei of the lighter elements like hydrogen, helium and lithium.

If the temperature is too high the protons and neutrons will then move too fast and don't spend enough time together, and thus can't stick together.

Nucleosynthesis sets the stage for the formation of stars and galaxies.

## Matter Dominates Over Radiation

**TIME:** 10,000 years

The universe continues expanding and cooling. At this time, lots more matter has been created by high-energy radiation. The universe expands and matter suffers less energy loss than the radiation.

In equations from this time, there is so much more matter than radiation that we can ignore radiation and are only concerned with matter. At the end of this process, photons scatter less with matter and more with each other. The protons thermalize and start acting as thermal black body radiation. We can still measure this cosmic background radiation today.

## Protons and Electrons Form Hydrogen

**TIME:** 500,000 years

By now there is a multitude of protons, electrons together with other types of light nuclei in the universe. However, until this point, the universe has been too dense and hot for the capture of electrons by a nucleus to last for very long. Now, the universe is cooler and electrons are caught, forming the first atoms:

Helium

Hydrogen

Lithium

## Hydrogen Gas Makes First Stars

**TIME:** One billion years

Since radiation has decoupled from matter, and electrons are bound to atoms, gravity becomes an important force. Small vacillations in matter density and gravitational field begin to increase and coalesce. Hydrogen gas is gathered by gravity and it fuses into the first stars.

## Stars Produce Heavier Elements

**TIME:** 2-13 billion years

When stars and galaxies took shape the only elements were hydrogen, helium and lithium. Heavier elements come from within stars. Nuclear fusion creates the heavier elements.

The heavier elements are ejected from stars as they supernova, but before they settle into old age as a white dwarf, neutron star or black hole.

The process of making and ejecting the heavier elements runs over the lifespan of a star…2-13 billion years.

## Life Evolves

That's it. Life evolved on Earth. Maybe elsewhere, maybe not.

# Black Holes

## Gravitational Collapse

As a child, most of have tried to jump as high as possible thinking that if we jumped high enough, we might just fly straight off the Earth. We never got far, though, did we?

The reason for this is because gravity pulls you back to Earth. If you were on Mars, you would be able to jump higher and even higher still if you were on the Moon.

The reason for this is because the mass of Mars and the Moon is smaller than Earth's. If you use a rocket, you can reach escape velocity and actually get off the Earth – making the space program work the way it does.

However, everything within our universe has an upper speed limit known as the speed of light. What if the planet's escape velocity was greater than this (i.e. fast enough to trap light)? The result would be a black hole.

The event horizon of a black hole is the point where light no longer has the ability to escape. Anything which slips over the event horizon will not escape – not even light.

Black holes are created by a group stars which have a mass of at least twice that of the sun.

When a star gets too old and all of its hydrogen becomes helium, and that helium is then turned into even heavier elements, the star can become one of three things:

**A white dwarf** – Created by the repulsion pressure of fermions in the star's core

**A neutron star** – Created by the neutron's repulsion pressure of the fermions which occur inside the nuclei of the core's heavy atoms

**Black hole** – The repulsion pressure of the fermions is not great enough due to the sheer size of the star's mass

Another method for the creation of black holes is by the gravitational collapse located at the center of a large group of stars.

This type of black hole can end up massively larger than our sun. There may even be one at the center of our Milky Way.

## Properties of Black Holes

Since the launch in 1990 of the Hubble Space Telescope, there have been several (if not many) observations of things which were perceived to be black holes.

However, before that time, black holes were strictly theoretical and mathematical. The Theory of Relativity provided us with the mathematical language to discuss and interpret black holes which allowed us to learn about them while they were still on paper.

## Abstract Theoretical Black Holes

We have learned a great deal about black holes through the mathematical models.

First, a black hole's surface area will never decrease; it can only increase meaning that two black holes can join together to make a single, larger one, but one black hole cannot be split in two.

Secondly, gravitational force is constant at the event horizon.

The first element makes it impossible for black holes to decay (or die) away, however, challenges to this will arise in the next section when we add quantum mechanics.

## Observable Astrophysical Black Holes

If light is not able to escape the event horizon, how will we even find and see one? It turns out that when gas and dust get sucked into a black hole, ionization occurs and huge amounts of light are released at the event horizon.

To this end, physicists and astronomers look for ionized gas and dust which are releasing light and being pulled back into something so quickly that this phenomenon could only be caused by a black hole.

However, this can be difficult to see, because black holes also pull in massive amounts of interstellar dust and are often shrouded behind huge dust clouds.

We have to get lucky and find a black hole that is either not shrouded or has a warped cloud with a little light peaking through.

## Hawking Radiation Poses New Questions

The previous properties of black holes were classic because they used Einstein's equations and did not use quantum mechanics to take particles into account.

Quantum mechanics is most easily added to classical approaches by looking at the scattering of particle in space-time curvature, where this curvature can't react to the particle scattering. This is a bit contrived but it still makes black holes do some pretty amazing stuff.

## Black holes decay!

Theoretical physicists (i.e. those who studied the scattering of quantum particles in a space-time curvature) uncovered the fact that the definition of particles and anti-particles solely depends on who is observing.

This goes against all the usual rules of the general relativity theory and implies that the particles to be counted also depend on the person doing the counting.

This led to the development of an interesting perception in relation to black holes. If, by some unlucky chance, a physicist fell in a black hole, his visibility would not allow him to see anything while passing the event horizon, and he would be crushed by the black hole's center of gravity.

But if a rope were attached to his thumbs and he was held just on the outside of that horizon, he would view quite a different, if not startling, picture. His toes would get quite burnt from the seething particle soup emitted from the black hole.

How is it possible for particles to get out of a black hole when not even light can? A particle/anti-particle pairing is created (just for a brief moment) directly outside of the black hole's event horizon.

Before this particle/anti-particle pairing can be destroyed, the anti-particle portion is pulled into and behind this event horizon, and the particle portion is expelled into the opposite direction.

The physicist (the one who is being held just at the edge of the black hole) can see these particles emitted at the event horizon. A more distant observer sees the black hole's mass decrease by one particle and will notice that the event horizon area has just gotten smaller - meaning that that black holes **can** get smaller and will decay over time.

## Where are the Quantum Microstates

Not only does a black hole die off but its expelled particles have a thermal distribution. These dead (or soon to be dead) black holes begin to look a lot like the objects which would have been studied in thermodynamics during and since the $19^{th}$ century.

The quantum revolution gave us the information that all thermodynamics could be defined in terms of bulk limits of the quantum microstate sums. For instance, steam power is a simple matter of understanding the microstates of air and water. This leads us to wonder what the quantum microstates are which aid in the creation of black holes? String theory thinks it has the answer.

## String Theory and Black Holes

Black holes provide the answers needed to Einstein's equations giving rise to the thought that any string theories which contain an element of gravity can also allow for the prediction of the existence of a black hole.

However, these string theories then provide us with more fascinating symmetries and matter types meaning that black holes also give rise to some more unusual things in string theory.

## String Theories Answer

To arrive at the answers, first we need to understand entropy and temperature. Temperature is a concept most of us use daily. Entropy is a little less common. Entropy is a logarithm. Every molecule has a quantized energy state.

We know the quantum theory of molecules and are able to count their quantum microstates. Entropy can be defined as the logarithm of the number of quantum microstates we count.

Black holes get hotter as they decay. A black hole's temperature is inversely proportional to its mass. The entropy of a black hole is 1/4 of the event horizon's area, so its entropy decreases as the black hole decreases in size.

But until the existence of string theory came about, there was no clear understanding of the relation between black hole entropy and quantum theory's microstates.

## Black Holes and Branes in String Theory

Through the relationship of string duality, it has since been realized that space-time geometry should not be considered as a fundamental concept. When forces are really strong or set at a small distance scale, there is an alternative way of describing a similar (if not same) physical system which appears to look different.

A BPS black hole is a very specific kind of black hole and it is important to string theory. A BPS black hole contains both elements of mass and charge with these elements satisfying a level of equality.

This, in turn, results in a supersymmetry which remains in the space-time near to the black hole. Supersymmetry is important as it allows chaotic quantum corrections to disappear, meaning that precise physics can be easily found through simple calculations.

We've already learned about d branes and p branes and since a single point could be considered a zero brane, this natural abstraction of a particular black hole can be defined as a black p brane.

A relationship also exists between both black d branes and p branes. In systems with a large charge, space-time geometry is an excellent way of describing a system of black p-branes.

With the required conditions fulfilled, we are able to work out the number of possible quantum states. When we return to the p brane system, we find that this system of entropy equals the black hole or p brane entropy levels.

This is an astounding outcome, for string theory in particular! However, can we finally argue that D branes offer us with the quantum microstates (at a fundamental level) of a black hole which, in turn, underlies the thermodynamics of black holes?

No. After all, there are probably very few black holes with a charge, and they are not even close to being BPS objects.