Solving the Mystery

1. Introduction

There is something intellectually satisfying in using a mathematical model under carefully defined assumptions to find an answer to a problem. Learning maths and physics in school we soon come to appreciate the way in which maths can be applied to problem solving in physics.

We are also cautioned to respect the fact that mathematical models are an idealised representation of the real world which may have its limitations. For example the idea that a large mass and a small mass will fall under gravity at the same speed ignores the effect of air resistance. A feather and a rock do not fall at the same speed unless you are somewhere like the moon where there is no atmosphere.

The important point is that the terms of reference of a school maths problem are such that you are given basic assumptions, laws to follow, mathematical equations to use and, in the context of the problem solving process, a complete understanding of the problem. It is fundamentally important that when applying maths to solving a problem in physics that the descriptive context in which the mathematical model is to be used is well defined and well understood.

Where we get into difficulty in some areas of physics is in the application of mathematical models without a clear understanding of the physical interpretation of the mathematical model. We have found a mathematical equation which seems to give the right answer to some questions but when it is applied to other cases the results are not valid.

In the subsequent sections the relationship between maths and physics is studied, looking at examples of scientific development and aspects of our understanding of the real world. In some cases a mathematical relationship was detected between physical properties and this then gave an insight into the physical nature of physical entities. In other cases the new physical theory is derived from first principles by thought experiment and logical deduction. In other cases scientists will admit that they start with a guess or a hunch and then go on to complete experiments, make observations and verify the guess. Then the mathematical model can be developed and used to make further predictions.

The proposal is made that there is no mysterious connection between physics and mathematics. It is just that we do not yet have a complete understanding or description of what is going on at a fundamental level. If we had this understanding, then the relationship between experimental observation, physical description and mathematical model would become clear. So it is the lack of a correct physical description that is impeding the use of mathematics to the best advantage. In this sense and in this application, mathematics is subservient to physics in that the mathematics can only be applied once the physics is properly understood.

We could call it a search for the unification of physics. Once we have clear and comprehensive model to hand which describes the physics of the very small and the very large then the relationship between the physical description and the mathematical equations will become fully clear.
2. Making the connection

Sir Isaac Newton developed basic principles of mathematics which he used to solve physical problems. He made the connection that the force of gravity causing objects to fall to earth was the same force of gravity that kept the moon and the planets in orbit. From observations of the movement of planets he was able to develop the law of gravitation which showed that the gravitational attraction between masses was proportional to the two masses and inversely proportional to the square of the distance between them.

At the time the force of gravity remained a mystery in that there appeared to be an "action at a distance" which was unexplained. The magnitude of the force of gravity had been explained but not the underlying cause.

It was several centuries later that Albert Einstein developed the special theory of relativity and the general theory of relativity. In the general theory of relativity the force of gravity is explained as occurring as a result of the curvature of spacetime. The sun curves spacetime and this causes the earth to follow its orbit.

This illustrates the point that mysterious connections between physics and maths are solved when a deeper understanding or enlightenment is achieved.

We can consider spacetime as a prime example of the connection between physics and mathematics. When Albert Einstein developed his theories he did so methodically by considering measuring rods and clocks and using thought experiments and logical deduction. We do not question that if we put two identical rods in line we get twice the length and this is the basis of the mathematical representation of distance.

We can reasonably assume that standard identical clocks in the same frame of reference tick equal seconds sequentially and in synchronisation with each other. We have a basic mathematical representation of time. In the case of the theories of relativity, there is no longer a mystery in the fact that mathematics can be used with great accuracy to model physical processes.

In the next section we will look at examples where a mystery remains and this mystery is attributed to our lack of a complete understanding of physical reality.


The understanding of the structure of the atom has evolved over time with some difficulty. Go back just over one hundred years and the question of whether atoms existed was unresolved. Then the observation of Brownian motion and the explanation by Albert Einstein of the cause of the movement of the pollen grains on the surface of water provided a convincing argument that atoms did indeed exist. Mathematical modelling of the impact of the water molecules on the pollen grains backed up the physical conclusions.

The study of the very small scale has been limited by the difficulty of observing at this scale and also the fact which became apparent that the observations themselves disturbed the system and affected the outcome of measurement. Over time, measurement at the
macroscopic scale led to conclusions regarding the underlying nature of matter and radiation.

The accurate measurement of black body radiation resulted in a situation where mathematical equations could be deduced but explaining why the equations took that form was more difficult. Max Planck resorted to considering the radiation as coming from an arrangement of oscillators, each oscillating with a frequency given by the equation $E = hf$, where $E$ is the energy of oscillation and $f$ is the frequency of oscillation. Planck’s constant $h$ provided the constant of proportionality.

Max Planck did not initially believe that such quantisation actually existed. Albert Einstein studied the photo-electric effect and showed that the quantisation $E = hf$ gave a beautiful explanation of the emission of electrons from a metallic surface when exposed to light of various frequencies.

A further case where an equation was deduced from experimental results was in the absorption spectrum of hydrogen. The dark lines in the light from the sun were measured at very precise wavelengths and the relationship between the wavelengths was deduced by Johann Balmer in 1884 as described by the formula
\[
\lambda = \frac{b(m^2/n^2)}{(m^2-n^2)}
\]
where $\lambda$ is the wavelength at which the dark lines occurred and $m$ and $n$ were integers. It was not until 1913 when Niels Bohr was considering the structure of the atom that the Balmer formula was explained as caused by electrons changing energy levels.

The work of understanding the atom progressed and was led by Niels Bohr and a scientific consensus developed around the Copenhagen interpretation of quantum theory. A whole new field of quantum mechanics developed with Paul Dirac in the UK making a major contribution to organising the mathematical analysis in a rigorous way.

The relationship between physics and mathematics in quantum theory was mysterious in many ways with a lack of consensus in the interpretation of the equations of quantum theory which persist to this day.

Quantum theory makes extensive use of complex numbers in the mathematical analysis. It is true that, whenever observations are made, the results are always real numbers and this is enforced by the mathematical relationships. However, it does seem surprising that equations describing a real physical process can only be modelled by using complex numbers, complex equations and complex wave functions.

This is indeed a mystery and suggests that there must be an underlying physical description which by its nature is constrained in such a way that the equations of quantum theory operate to produce the right results. Where we detect a mystery, we should be looking for enlightenment through a new perspective on the problem.

There is another aspect of the Copenhagen interpretation of quantum theory which deserves mention and that is the probability interpretation of the wave function. This concept is based on the idea that the wave function in quantum theory can be used to determine the probability of finding a particle at any point in space.

The dual slit experiment shows the wave properties of light and results in an interference pattern on a detection screen. This experiment can be run using one light quantum (photon)
at a time and still the interference pattern appears. If a detector is placed at one of the slits to find out which slit the wave quantum used in its passage from the light source to the screen, then the interference pattern disappears. It is possible to model this experiment mathematically and use the wave function to model the probability of detecting the wave quantum at any point on the screen. However, if the wave quantum went through only one of the slits than logically it should make no difference if the other slit were closed. The logical conclusion must be that the light quantum went through both slits but this is not the conclusion of the Copenhagen interpretation.

The mystery must be solved in this case by assuming that a real physical wave passed through the apparatus and passed through both slits of the experiment. This requires a shift in our understanding of the physical reality of the nature of light.

A further mystery surrounds the connection between physics and mathematics in the context of the fundamental forces of nature. There are considered to be four fundamental forces namely: gravity, electromagnetic, strong nuclear and weak nuclear. It has been possible to create mathematical models to unify three of the four fundamental forces but gravity remains outside of the unification.

A full unification has been attempted using string theory which was developed following an insight that particles could be represented as the oscillations of vibrating strings. This had the merit that it avoided the infinities which occurred when performing calculations under the assumption of point particles. The mathematical models of string theory have been extensively developed but it is a major drawback that the theory implies the existence of 12 dimensions and no evidence or experiment can be devised to verify this hypothesis.

We have to solve the mysterious connection between the physics and the mathematics of string theory by determining the real underlying nature of the real world in the known dimensions of spacetime namely three space dimensions and one time dimension. Then we have the potential to solve the mystery of string theory by considering modes of vibration of spacetime.

The same approach to the solution to the mystery of quantum theory will be to find the underlying nature of spacetime which must be constrained in such a way that the complex number equations and functions of quantum theory are effective in giving the right results.

4. Fundamental physics

I can think of no other way of illustrating the key point that mysteries are solved by creating a new understanding than by describing a new approach to fundamental physics.

We start with the nature of light. We can measure the speed of light at just under 300,000 kilometres per second. We know that light is a wave and is emitted and absorbed in quanta of energy. So a wave quantum of light with energy given by $E = hf$ travels a distance of 30cm in one nanosecond. So we have to understand the means of transmission of this energy. For waves to occur something must be waving.

We can talk about light as an electromagnetic wave but it is not valid to think that it is the electromagnetic field that is the means of propagation. A field is a secondary or tertiary effect with some other underlying cause.
Understanding the means of propagation of light has a long history and at one stage it was proposed that there was a luminiferous aether which pervaded all of space and acted as the medium for wave propagation of light. The analogy with sound waves meant that it should be possible to measure differences in the speed of light under different rates of movement through the aether. The Michelson Morley experiment showed that the speed of light did not vary in different directions through the apparatus even when the experiment was changed in orientation relative to the movement of the earth through space and at different times of the year.

Subsequently the analysis provide by Albert Einstein in the special theory of relativity showed that the speed of light is a constant for all uniformly moving observers. So the idea of a luminiferous aether pervading all of space was not correct. However, it still leaves the unresolved question of the means of propagation of light.

This is where we must make an informed guess. We know that the equations of general relativity describe a relationship between mass/energy and spacetime curvature. We know that there is a solution to these equations which represents waves travelling at the speed of light and when at a low frequency these are called gravitational waves. We know that there is no specific frequency limitation on the application of the general relativity equations so it is reasonable to propose that the means of propagation of light is spacetime itself. A wave variation in the fabric of spacetime propagates to an adjacent position at the speed of light.

So we take as our starting point the concept that light is a wave disturbance of spacetime. We have to revisit the interference experiment and understand that it is a real physical wave passing through both slits of the interference experiment and it is only when the dispersed wave arrives at a detection screen that the wave energy is absorbed by a single atom so that the energy becomes located at one place namely as part of the energy of an electron of an atom. The quantum theory collapse of the wave function is really an interaction between a dispersed wave and an atom in a detector screen.

Light is emitted when an atomic electron drops from a higher energy state to a lower energy state. What do we think could be the cause of the emission of a wave disturbance of spacetime? The answer must be that the electron is a looped wave disturbance of spacetime. The electron is like a light wave caught in a closed loop and with energy $E = hf$ travelling at the speed of light $c$ in the loop.

This concept is very similar to the ideas proposed by Louis De Broglie when he showed that if we represented electrons as matter waves described by the equation $E = hf$ then the whole analysis of the energy levels of hydrogen could be completed by recognising that there must be a whole number of wavelengths in the looped matter wave. The matter waves of Louis De Broglie were considered to be standing waves or pilot waves rather than a wave disturbance of spacetime so the physical interpretation is quite different but the mathematical analysis is the same.

We need to satisfy ourselves that the idea of an electron as a looped wave disturbance in spacetime is stable. An electron thought of as a particle in orbit around a nucleus would radiate away all its energy. However, a looped wave has energy dispersed throughout the loop and experimental evidence shows that it takes energy to increase the radius of the electron ‘orbit’ meaning that the loop will tend to contract to the lowest possible radius.
If the electron is a looped wave in spacetime, then we can reasonably propose the same hypothesis for the proton and the neutron. We can adopt the idea that all the particles that make up the atom are looped waves in spacetime. This leads to a deeper understanding of the relationship between mass, energy and spacetime curvature. The general theory of relativity fully explains the property mass in accordance with our experience of gravitational mass and inertial mass. The one additional part that is added by the spacetime wave theory is that we have now explained how mass curves spacetime by itself being a moving wave of spacetime curvature.

Also, each individual particle in an atom (proton, neutron, electron) being a looped spacetime wave reacts to a generally curved spacetime environment by preferring a lower energy position in space.

We have discussed the electron as a looped spacetime wave disturbance. The wave disturbance of spacetime represents a variation in spacetime curvature propagating in the loop. If we take one wavelength in the loop as the waves pass, space itself is compressed and expands as the change in space curvature passes. If there were no variation in the time dimension then the passage of the wave could be represented by a simple sine wave and the net compression or expansion of space would be zero.

However, we have to take into account the wave variation in the time dimension which is synchronised with the wave variation in the space dimension. A wave in the time dimension means that the rate of passage of time varies during the wave cycle so that the time spent in an expanded state and the time spent in a compressed state is not equal and there is a net result of expansion or compression of space. It can further be shown that the net effect of all the wavelengths in the loop is independent of the mass of the particle.

The effect of the time variation is equivalent to the property of electric charge. Two electrons in proximity will have a lower energy state if they are moved further apart because of the effect on compression/expansion of space. The electric charge of an electron and a proton is distributed throughout the circumference of the looped wave. The underlying mathematical calculation for this is to be found by searching in Google for Richard Lewis The Unification of Physics.

As part of the process of unification we have to find an approach to the unification of the four fundamental forces. The approach is to recognise that all four forces are caused by energy differences in the configuration of objects in spacetime. The force of gravity can be viewed in this way by considering the energy associated with any element of mass in two adjacent positions. Objects will tend to move to a position of lower energy and this appears to us as a force, in this case the force of gravity.

We have discussed the electrostatic force as due to variations in the compression of space due to wave variations in the time dimension. This leads to a lower energy state when electrons are moved further apart.

The magnetic force arises in every case due to a moving electric charge and always occurs as a magnetic dipole. A magnetic monopole cannot exist. The strong nuclear force arises due to the energy difference (mass deficit) associated with a neutron and a proton in close proximity. The weak nuclear force occurs due to the difference in energy state of a neutron before and after decay. Thus all four fundamental forces are caused by the energy
difference between objects in different positions in spacetime with a universal tendency to move to a lower energy position giving rise to an apparent force.

How does this new interpretation of the physical world help us to solve the mysteries which we have identified? It answers the important question of how light waves are propagated. It explains the properties of mass and charge. It provides an understanding that the experiments that we conduct are with real physical waves. It identifies the important active role of the detection mechanism in an experiment. It provides unification of the four fundamental forces. It provides an interpretation of quantum theory albeit one which requires further work to show how the spacetime wave theory describes the underlying cause for the results obtained in the mathematical models of quantum theory.

These loops in spacetime should be modelled using the equations of general relativity which relate spacetime curvature to a mass energy distribution. The direction of search for a single equation to describe all fundamental physics is resolved by recognising that we already have this equation within the general theory of relativity. We just have to learn how to apply the equations of general relativity to the physical reality which described above.

The process of understanding has been built up by the development of ideas and experiments over many years and the process of scientific investigation has taken many forms with many different methods.

Once we have obtained a clear picture of the fundamental properties of nature, we are then in a position to construct a comprehensive model where we start with a complete description of the physical world. Then we show how this description relates to the physical properties that we can measure and the experiments that we can perform. Then the underlying mathematical equations can be used to provide quantitative results which can be verified by experiment.

This top down approach could not have been adopted to achieve the understanding in the first place. However, once in place it eliminates the mystery of the connection between physics and mathematics. The whole becomes clear and coherent and the reason why the mathematical models work is apparent.

**5. The evolution of the universe**

The underlying assumption that we adopt when using mathematics in the application to cosmology is that the laws of physics hold throughout the universe. This is a realistic assumption and has been of immense benefit. The general theory of relativity was constructed through the careful use of thought experiment and mathematical analysis and verified experimentally through observations within the solar system. However, it has been applied to model black holes, distant stars, and the entire universe.

The application of mathematics to problems in physics depends fundamentally on any assumptions that are implicitly or explicitly built into the model. We observe the universe and observe that it is expanding and so we naturally think back in time to when the universe was smaller and hotter. We can start to make use of mathematical models to model the effect of gravitational attraction on this system as it then evolves forward in time.
To make the mathematical models useable, cosmologists made simplifying assumptions. The cosmological principle is that 'Viewed on a sufficiently large scale, the properties of the Universe are the same for all observers.' Related to this assumption is the hypothesis that the universe is finite but there is no space boundary. This is justified on the assumption that the universe is a three dimensional surface in a universe which has four space dimensions and one time dimension. It is important always to remember that these are assumptions which are not proven.

The assumptions are in place and the mathematical models are applied to model the evolution of the universe. In the process we encounter some substantial problems in that the mathematical model leads to a situation where we have to constantly revise the descriptive narrative of the evolution of the universe.

Right at the very start we have a problem in that we go back in time to a point where all the mass and energy of the universe lies within a small region and the mathematical model tells us that we have a singularity. A singularity in mathematics is something that can occur in an equation but it is most unwelcome in physics. In mathematics if you take the equation $1/r$ and try to evaluate it at the point $r = 0$ you have an infinite result.

In physics a singularity has been described as a point where the laws of physics break down but in reality it is an indication that the mathematical model describing the physics is not quite right. We would like to have laws of physics which apply everywhere and for all time and if the mathematical model has a singularity then we need a different model.

What is known as the Big Bang singularity is a huge violation of the law of conservation of energy. What is worse, the assumption regarding the formation of matter and antimatter means that we have to assume the formation (in the Big Bang singularity) of mass and energy equivalent to around one million times the total observed mass and energy of the universe today.

As we proceed, we find that certain observations are made which require us to propose modifications to the theory. We introduce the idea of inflation, of dark matter and dark energy. Even then the mathematical model is found to be so finely balanced that the probability of arriving at the universe we observe is one part in trillions of trillions. This is so unlikely that it would normally be classified as impossible.

We have our doubts about the Big Bang model of the evolution of the universe and the best way to proceed is to go back and challenge the assumptions. One very important consideration is the formation of matter in the universe. The equations of general relativity tell us that the conservation of mass and energy is only approximate where there is space curvature. Indeed we have to modify our conservation law to state that the Total Energy must be considered as mass plus energy plus spacetime curvature. This is a helpful conclusion because it means that the total energy of the universe can equate to zero and that changes in spacetime curvature can contribute to matter formation.

So we can go back and challenge the initial assumptions. We have previously adopted the idea of a time boundary and we know the intimate relationship between space and time in spacetime so we should assume the existence of a spacetime boundary. This contradicts the no boundary hypothesis and means that we can operate with a simple model of three space dimensions and one time dimension.
Starting with an empty universe with a spacetime boundary but no matter or radiation, it is the space boundary which is the cause of the expansion of space. The expansion occurs because there is nothing to stop the expansion at the boundary. The expansion is no longer fine tuned but paced and controlled by expansion at the boundary. The assumption that the boundary is expanding by one light year per year seems to fit with observation. The expansion of space is then at the rate of $1/R$ light years per year where $R$ is the current radius.

The expansion of space must lead to matter formation to balance the Total Energy equation. The way in which this occurs is a matter of conjecture but the observation of distant gamma ray bursts suggests that galaxy formation occurred as a sudden release of tension in the fabric of spacetime leading to a huge release of gamma rays with some of these gamma rays becoming looped waves in spacetime form neutrons. Neutrons decay to protons and electrons and form hydrogen atoms and molecules. So we have a galaxy formation event which results in the formation of a black hole and an expanding spherical region of hydrogen gas.

The hydrogen gas formed in the sphere of the galaxy starts to collapse under gravity and stars are formed, evolve and explode but still the stars within the galaxy halo represent around a fifth of the mass of the galaxy. The rest is hydrogen gas which produces the effect which we attribute to dark matter. The idea that star formation only occurs within the halo of the galaxy explains why we do not see star formation events in inter-galactic space.

Dark energy was introduced to explain the unexpected expansion rate of the universe and this is now explained by the expansion at the space boundary. The idea of inflation was introduced to explain the uniformity of the microwave background radiation but now we understand the microwave background radiation must have originated from the collective effect of the galaxy formation events. The cooling effect over billions of years has brought the temperature of the radiation to very similar levels throughout the sphere of our observation.

One last result from this new perspective on the universe. If the universe has a boundary then it has a centre. We can estimate our position in the universe by using the observation of the CMBR rest frame and applying Hubble’s law to find that the milky way galaxy is located just under 13 million light years from the centre of the universe. From this assumption we can go on to explain why the more distant galaxies seem to be fainter and moving more quickly than we would expect from a linear relationship.

There is a reference paper for this topic which is to be found by using Google to search for Richard Lewis The Evolution of the Universe.