

Front Cover

New Ideas
that
Change Einstein's Theories

Lawrence Stephenson

'This book gives new insight into special and general relativity'
Professor C.W. Kilmister

Back Cover

Einstein's special and general theories of relativity are mathematically flawless – but the conventional physical interpretation of his mathematics has been inadequate because the “advanced potential” solution of Maxwell's equations has been misunderstood.

With the help of two of the most eminent authorities on relativity theory and electromagnetic theory, the author has incorporated the advanced potential solution into special relativity theory. The interpretation of the mathematics then agrees with common sense.

Quantum theory predicts that the *act of observation* changes the velocity of a light signal when it arrives at any material detector. The first few photons to arrive correspond to a precursor transient which establishes the new steady-state electromagnetic “near” fields surrounding the detector. The advanced potential solution then predicts that the observed arrival velocity of the light wave must be equal to c *relative to the detector*.

Clocks really do go slow, and a twin will age more slowly, when travelling at speeds close to the speed of light. But a simple deduction from Maxwell's equations indicates that the clock's speed should always be measured relative to the background provided by the distant stars. There is then no clock or twin paradox.

Einstein suggested two ways of extending general relativity using Mach's Principle. One of his approaches, which has been overlooked, may predict the particular value of the gravitational constant G we observe. A single equation establishes a common gravitational origin for the four fundamental forces of nature.

First published in Great Britain in 2009 by
Bucke Academic Publications,

Second enlarged edition 2011

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, photocopying, recording or otherwise, without prior permission of the copyright owner.

Copyright © Lawrence Stephenson 2009

The author asserts the moral right to be identified as the author of this work.

ISBN 97809522823 34

Printed and bound in Great Britain by
the MPG Books Group, Bodmin and King's Lynn

Contents

Acknowledgments		5
Preface		7
Introduction		13
Chapter 1	Background to the Problems in Physics	15
Chapter 2	Galilean and Newtonian Theories	24
Chapter 3	Mach's Principle	27
Chapter 4	The Origin of Gravity, Weak Forces, Strong Forces and Electromagnetism	32
Chapter 5	How Antennas Work	39
Chapter 6	Advanced Potentials, Quantum Theory and Special Relativity	45
Chapter 7	Time Dilation, the Clock Paradox or Twin Paradox, and Relativity	53
Chapter 8	A Summary of Special Relativity	60
Appendix	Paper 1 A Dynamical Origin for the Gravitational Constant...	63
	Paper 2 The Significance of Precursor Electromagnetic Waves....	72
	Paper 3 The Relevance of Advanced Potential Solutions....	81

In this web version of my book the page numbers differ because some blank pages at the ends of Chapters have been omitted.

Acknowledgments

I am very grateful for the advice I received from the late Professor Clive W Kilmister over a period of forty-five years. Clive, who was Head of the Mathematics Department at King's College London until his retirement, was one of the most scholarly authorities on relativity theory. Without Clive's guidance I could not have published many of my academic papers.

Secondly, I should like to thank Professor Alexander L Cullen FRS, who is now retired from being Head of the Electronic and Electrical Engineering Department at University College London. Alex is a highly respected authority on electromagnetic theory.

In 1998 I was aware that all authorities claimed that individual "advanced potential" solutions of Maxwell's equations (explained in Chapter 6) were only valid in an imaginary world where time was running backwards (*i.e.* in the opposite direction to the normal "arrow of time"). Hence, I was prepared for some criticism when I sent Alex a draft of my 1998 paper on electromagnetic "precursor transients" (explained in Chapter 5). He told me that the application of the advanced potential solution to a receiving antenna (or aerial) was invalid, and he produced an irrefutable seven-line argument to prove that it was invalid. His analysis was correct within the context of the generalized reception of electromagnetic energy, in a three-dimensional region of space, that he had naturally assumed.

There was then a three-way correspondence between Alex, Clive and myself. It was finally agreed that certain carefully specified advanced potential solutions are not only valid in the real world, but also they form an essential part of electromagnetic theory. The advanced potential solution provides the theoretical basis for two of the three assumptions which Einstein had to make when he was developing special relativity.

The correct application of the advanced potential solution requires special relativity to be limited in its application to steady-state (non-transient) electromagnetic wave signalling. Quantum theory has to be introduced to explain the precursor transient signals produced by the arrival of the first few photons. These precursor transients constitute the front edge of all pulsed

electromagnetic signals (explained in Chapter 5).

The advanced potential solution provides an explanation of why the *observed* velocity of any electromagnetic *wave*, travelling in free space, is always found to have a constant value equal to c that is independent of the velocity of the source. The advanced potential solution predicts that the presence of a *material* detector, inserted into any free space region, will change the initial arrival velocity of the front edge of any approaching signal so that, in the steady state, the observed electromagnetic wave arrives at a velocity equal to c *relative to the detector*.

However, the arrival velocities of the first few photons, associated with the initial precursor transient of a pulsed signal, are not restricted to c . The first few photons will interact with the detector and set up the new perturbation fields of the detector. It is these stored energy fields which ensure that the arrival velocity of the following steady-state electromagnetic wave must be equal to c .

Einstein's second assumption, that the velocity of light is always constant in empty space, is unnecessary and of limited validity. Furthermore, the concept of velocity retains its classical interpretation within special relativity. Any difference in the departure velocity of a signal from a source, and its observed arrival velocity at a material detector, is accounted for by the obstruction provided by the detector and a change in the radiation pressure force exerted on the detector. An intervening medium will produce similar changes.

The advice I received from Clive and Alex enabled me to produce my year 2000 paper on: "The Relevance of Advanced Potential Solutions of Maxwell's Equations for Special and General Relativity" (see Appendix, Paper 3).

The three papers in the *Appendix* have been reproduced with the kind agreement of Physics Essays. The link to their web site is: <http://physicsessays.aip.org/>

Preface

The mathematics of special relativity is logically consistent and beyond question. But there has been a failure to understand special relativity as a theory of physics. This failure has held back the development of general relativity, and it has obscured the link between special relativity and quantum theory.

In this *Preface*, and in the *Introduction*, I will indicate where there are weaknesses in relativity theory, and I will briefly outline how some new ideas may be developed. If you have a limited knowledge of the basics of special relativity and quantum theory please read Chapter 1 first.

Einstein intended special relativity to be a theory of physics. He made precise statements concerning the way in which clocks may be synchronized when they are in relative motion. However, he had to make some “time-assigning” assumptions to justify these statements. The inherent limitations imposed on special relativity by Einstein’s initial assumptions have been ignored for over a century.

There has also been a failure to understand two important implications of Maxwell’s equations. First, a crucial part of the advanced potential solution of Maxwell’s equations, which establishes the theoretical basis for two of the assumptions made in special relativity, has been wrongly declared to be invalid because of the incorrect application of a causal argument. Secondly, the major consequence for relativity theory associated with the electric field solution for a single charged body, situated in an otherwise empty Universe, has been overlooked.

New ways to develop special and general relativity may be readily achieved by examining these two neglected aspects of Maxwell’s equations in more depth, and by studying the many comments made by Einstein concerning the limitations of his theories.

It is essential to clarify the “clock paradox”, which is also referred to as the “twin paradox” (see Chapter 7). Many relativists have caused confusion by failing to admit openly that there is a very real problem with the clock paradox. There are inherent difficulties in defining, with precision, the pivotal underlying concepts of both the inertia of a moving body and an inertial frame of reference (an inertial frame is one that is not being

accelerated). These difficulties were clearly stated by Einstein. But this knowledge, on its own, is not sufficient to solve the problem of the clock paradox. It will be shown that Maxwell's equations contain more information about the origin of time dilation than appears in just the simple mathematics of the Lorentz time transformation. The full solution to the problem of time dilation involves Einstein's clarification of Mach's Principle (see Chapters 3 and 7). However, before discussing time dilation any further it is necessary to consider a still more fundamental oversight.

Although the mathematics of special relativity is flawless there is a simple mathematical reason why the physical understanding of the theory has been unsound from the outset. A monochromatic (single frequency) electromagnetic wave is unattainable in the real world because, by definition, such a wave must continue forever and is not permitted to have a beginning or an end. A truly monochromatic wave cannot carry any information, and it cannot have an observable velocity. The observed velocity of any physically realizable, quasi-monochromatic, electromagnetic wave depends on the material boundary condition imposed by the detector on the electromagnetic energy when it first arrives. On the arrival of the initial wave front the output of the observer's detector will contain additional frequency components which will be time dependent (see below for an explanation). A full solution for all of these frequency components must be made before one may predict the observed velocity of any part of the electromagnetic energy.

It is therefore essential to qualify the second assumption of special relativity: "that light is always propagated in empty space with a definite velocity c which is independent of the velocity of the source". When considering the validity of this assumption one must divide the required analysis into two distinct parts. There has been a failure to consider the separate wave and photon approaches which are essential when analysing the nature of all light signalling in special relativity.

If a pulse of light is used as a signal, and the pulse has a very rapid rise-time at its front edge, then the first thing to be detected by an observer will be a high-frequency photon having a large energy, and *not* an electromagnetic wave. Why is this so? Simply because the mathematical description of the very fast rise-time of the front edge of the pulse requires sinusoidal Fourier frequency components which will extend into the ultraviolet region and beyond, and the energy of a photon is proportional to

its frequency. What can we say about the observed arrival velocity of the first photon? The answer is, very little! Quantum theory is needed for the solution, and this theory tells us that the location of the first photon, its arrival time, and its velocity, will all be indeterminate.

Special relativity cannot be used to predict the arrival velocity of the first photon. Special relativity may be applied to all steady-state electromagnetic waves, but it cannot deal with individual photons. Both Maxwell's equations and special relativity become invalid when frequencies are in the ultraviolet region and higher because of the ultraviolet catastrophe limitation (see Chapter 6).

Nevertheless, it is important to find out what happens to the first few high-energy photons associated with the arrival of a pulsed light signal. It is the answer to this question that is crucial if we wish to fully understand the physics of special relativity.

Any observation of a light signal requires the presence of a material detector. The electromagnetic wave solution for this detector, appropriate to special relativity, must be derived from Maxwell's equations. In the limit, it is necessary to consider an infinitesimal dipole antenna (or Hertz dipole) as the detector. The insertion of the dipole into a region of free space will perturb the field pattern that had existed prior to the presence of the dipole. The first few high-energy photons to arrive will set up new perturbation (or scattered) stored energy fields that are appropriate to the dipole in the electromagnetic steady state. These perturbation fields are often referred to as the "near fields" of the dipole. Although the near fields are very weak at appreciable distances from the antenna, they will still extend *past the source and to infinity* when the final electromagnetic steady state is reached. The advanced potential solution of Maxwell's equations (see Chapter 6) then requires that the arrival velocity of the following electromagnetic wave must be equal to c *relative to the detector*. In the steady state the electromagnetic wave may be considered to be guided onto the detector, at a velocity equal to c relative to the detector, by the perturbation fields that were initially set up by the arrival of the first few high-energy photons. Quantum theory indicates the nature of the transient solution. Special relativity (based on Maxwell's equations) provides the steady-state, or near-steady-state, electromagnetic wave solution. The observed velocity of any steady-state electromagnetic wave must always be equal to c , relative to the detector, because *it is the presence of the material detector that makes it so*. The detector will produce an initial slowing down, or speeding up, of the incoming energy if there is relative motion between the source and the

detector. There will be a corresponding change in the radiation pressure force exerted on the detector.

These new ideas enable the mathematics of special relativity to be established without using two of the assumptions adopted by Einstein. A theory with the minimum number of assumptions is to be preferred. Einstein's own very clear suggestions may also be used to demonstrate this development of special relativity. Special relativity is *not* primarily based on Einstein's second postulate: "that light is always propagated in empty space with a definite velocity c which is independent of the state of motion of the emitting body". This second postulate arises directly from Einstein's third assumption of a time-assigning function, which he stated to be necessary in order to establish a definition of simultaneity when electromagnetic signals are passed between observers in two separate inertial frames of reference A and B . This third assumption states that the out-and-return transit time for a signal passing from A to B and back to A must equal $2AB/c$. After making this third assumption Einstein left the door open for a new specific interpretation of special relativity by stating: "We assume that this definition of synchronism is free from contradictions". The third assumption Einstein made is, of course, valid for the implied electromagnetic wave signalling process between observers that he adopted. But things have moved on since 1905. A signal consisting of a single photon arriving at frame B will not produce an instantaneously generated return photon. Signalling using individual photons will result in an out-and-return transit time that is greater than $2AB/c$. Quantum theory requires a modification to be made to special relativity when dealing with precursor transient signals (see Chapter 5). Quantum theory then dictates that it is the *act of observation* of an electromagnetic wave which ensures that the observed *steady-state* arrival velocity of the wave is equal to c *relative to the detector*. The predictions of special relativity are indeterminate until the electromagnetic steady state is reached.

It is the appearance of the velocity of light c in the uniform plane wave solution of Maxwell's equations that has caused major confusion in relativity theory. This confusion arises because the uniform plane wave solution only applies to an electromagnetic wave (or other function) of unchanging form. As stated earlier, such a wave cannot carry information or have an observable velocity. The boundary condition imposed by any material detector inherently generates numerous frequency components when a wave-front arrives at a detector. A full analysis at this material boundary is essential if one wishes to predict what may be observed.

Having outlined the approach that is required to clarify the physical interpretation of special relativity it is appropriate to consider general relativity. In Chapter 4 more new ideas are developed by following up proposals that were initially put forward by Einstein. Einstein indicated two independent approaches that might be used to fully incorporate his formulation of Mach's Principle into general relativity theory. But only the first of his two possible approaches has ever been seriously considered. The first approach attempts to link the origin of the local inertial reaction force, that occurs when a body is pushed, to a gravitational interaction between the body and all of the distant matter in the Universe. It will be shown in the *Introduction* that this first approach leads to unacceptable conclusions.

Einstein's second way of interpreting Mach's Principle is fully discussed in Chapter 4. His alternative proposal suggests that one possible explanation for the origin of gravity, and the particular value of the gravitational constant G we observe, might be that gravity is generated by the rotational inertial motion of our Universe against a background of distant universes. A constant free-space value of G would be predicted to occur throughout our Universe. But the theory may also predict an observable increase in the value of G when it is measured within the boundary surface of a smaller, rotating, fluid body, such as the Earth. If confirmed, such a terrestrial observation would put the study of gravitation on a completely new basis. There were significant, unexplained, variations in the observed value of G in some measurements taken at the National Bureau of Standards, Washington, D.C., that initially confirm this prediction (see Chapter 4). Further discrepancies in the observed value of G , ranging from 0.2% to 0.6%, have been noted in three recent experiments when the anticipated accuracy was of the order of 0.05%.

The specific proposal concerning the origin of gravity, which has just been outlined, is clearly speculative. Nevertheless, a fresh start needs to be made on the further development of Mach's Principle, along these lines, in view of the importance attached to this principle by Einstein.

Following this brief introduction to Mach's Principle, it is now appropriate to outline the approach that will be taken to clarify the clock paradox. In Chapter 7 a further development of Maxwell's equations is discussed which provides a rational explanation of the clock paradox. This development also requires a consideration of Mach's Principle. Relativists tend to forget some of Einstein's early comments. He made it clear that any attempt to define either the inertia that a body is given when it is pushed, or an inertial frame of reference, introduces an inherent problem which is

linked to Mach's Principle. At a really fundamental level one reaches a circular argument. Special relativity can only state that a particular frame is an inertial frame of reference if this frame is found to be the one in which a standard clock indicates the largest passage of proper time between two events. Hence one can only establish the result that one is trying to deduce by making an observation! But there is a way forward.

Special relativity is inherently subordinate to Maxwell's equations. An examination of the deeper aspects of Maxwell's equations provides a solution to the clock paradox. Mathematically, Maxwell's equations predict that a single point charge will produce an electric field extending to infinity in an otherwise empty Universe. But such a proposed situation is not tenable unless a second distributed charge, equal in magnitude and of opposite sign, is also present in the Universe to create the electric field. Hence, distant matter must be present in the Universe to validate Maxwell's concept of an electric field. This result indicates that a Lorentz time transformation correction only arises because of the motion of the clock or twin relative to a background frame of reference which is provided by distant matter in our Universe. There is then no clock or twin paradox, because background matter in the Universe is providing a primary inertial frame of reference. A standard atomic clock will run at its fastest rate when located in this primary inertial frame.

An atomic clock will also run more slowly if it is located close to a massive object as a result of a gravitational interaction. There is an overall logical consistency if gravitational time dilation depends on a specific value of G that is created by the suggested rotational motion of the matter in our Universe relative to a background frame of distant universes, and if Lorentz time dilation depends on the motion of a clock relative to the background frame provided by the matter in our Universe. At a fundamental level, with the value of G being established by the rotation of our Universe, both forms of time dilation would depend on the motion of an atomic clock in space relative to distant matter, as Mach's Principle requires. The equality of inertial mass and gravitational mass, and the equivalence principle, would then also follow as a direct consequence of Mach's Principle.

Introduction

To give a flavour of the underlying problems with physics, I will describe a very basic example of our present lack of understanding.

Consider Newton's third law of motion, which lies at the heart of physics. This law states that action and reaction are equal and opposite. For example, if I push on a car to start it moving, and give the car some inertia, I will feel an equal and opposite inertial reaction force pushing back on my hand. Although Newton's third law of motion might appear to be obvious and, with hindsight, might even be considered to be a common-sense deduction, it has hidden depths.

Physicists who deal with the fundamentals of physics like to consider what actually *causes* the inertial reaction force on my hand when I push a car. One then needs to consider Mach's Principle, a fundamental principle formulated by Einstein and one that will be discussed in some detail in Chapter 3. It is sufficient to say here that the majority of eminent physicists working in this field of study claim that the best explanation for the cause of the inertial reaction force on my hand is that it arises from a gravitational interaction force between *all of the distant matter in the Universe* and my moving hand. But relativity theory requires that any change in the gravitational interaction force, produced by distant matter, can only travel to my hand at the speed of light! How can the *instantaneous* reaction force I feel on my hand, when I push a car, be caused by the motion of my hand relative to all of the distant matter in the Universe?

The currently accepted explanation for the cause of this inertial reaction force then has to be made even more complicated by stating that the changes in the gravitational interaction forces are travelling to my hand from both *future time* as well as *past time*, as advanced gravitational waves and retarded gravitational waves. These two sets of gravitational waves then combine in *present time* to give an instantaneous inertial reaction force at my hand! Quite seriously, many leading physicists consider this to be the best explanation of the fundamental origin behind Newton's third law of motion. One has to invoke gravitational energy coming from both future time and past time to explain the instantaneous force I feel on my hand when I push

a car in present, real, time. Such an improbable concept is only tolerated because all of the alternative theories, which would relate local inertial reaction forces to distant matter, are inherently flawed.

What has been forgotten by most physicists is that when Einstein clarified Mach's Principle in 1916 he clearly stated that the motion of distant matter, relative to local matter, must be the cause of *either* local inertial reaction forces *or* local gravitational forces. Einstein's alternative interpretation of Mach's Principle has been almost totally ignored. Why not try to satisfy Mach's Principle by relating local gravitational forces to the motion of distant matter? In 1964 Fred Hoyle made a partial attempt at this problem, which was later shown to be unacceptable. Since 1964 this aspect of Mach's Principle has been dropped.

If Einstein's alternative interpretation of Mach's Principle is correct, and if local gravitational forces are caused by the suggested rotation of our Universe relative to a background of distant universes, then there is no problem with the instantaneous nature of inertial reaction forces. It is only necessary to meet one of Einstein's interpretations of Mach's Principle for the principle to be satisfied. Before venturing forward to discuss these new ideas in more detail it is worthwhile to briefly summarize the background to these ideas in Chapters 1, 2 and 3.

Chapter 1

Background to the Problems in Physics

The aim of physics is to produce theories and laws that enable us to understand how the Universe operates. A great deal of theoretical effort is being undertaken at the present time to achieve a better understanding of gravitation, electromagnetism and quantum field theory. It so happens that the Universe behaves in a way that enables us to produce a number of quite simply expressed mathematical laws that separately cover all of these individual topics. However, although there are some links between the individual topics, it has not been found possible to bring together all of the relevant individual theories so as to produce a single *Unified Field Theory*, or a *Grand Unified Theory*, or a *Theory of Everything*. These are all names that have been used to indicate a more generalized theory. At the outset it should be acknowledged that the use of the expression a *Theory of Everything* is a misnomer, despite its use by many physicists, as well as colloquially. The most that a *Theory of Everything* could realistically achieve would be to unite all of the presently known theories and laws associated with the various branches of physics and cosmology.

A further word of caution is necessary. The laws of physics can never be proved. All scientific laws or theories are inherently based on certain initial assumptions, which may be either stated or implied. Any given law or theory will only be valid within the limits imposed by these assumptions. The implied assumptions are sometimes hidden and are frequently only detected at a later date after analysing experimental or observational evidence.

James Clerk Maxwell produced the first successful partial *Unified Field Theory* in 1864. He combined the previously separate theories of electricity and magnetism to form the present-day theory of electromagnetism. Maxwell's four equations express mathematically four experimental observations. These observations are that an electric current produces a magnetic field, a time-varying magnetic field generates a voltage in a conductor, free electric charges may exist, but free magnetic poles do not exist. Maxwell's equations give rise to the whole of classical electromagnetic field theory and, in particular, predict the existence of

electromagnetic waves that can be radiated off into space. The frequency spectrum of electromagnetic waves includes radio and television signals, microwave communication signals, infrared radiation, all of the visible light frequencies, the ultraviolet region, and right through to X-rays and beyond.

Further deductions from Maxwell's equations led Einstein to produce the special theory of relativity in 1905. Einstein had to make three assumptions to establish special relativity. He first defined the principle of relativity. This principle states that the laws of physics should be the same in all inertial frames of reference (an inertial frame is one that is not being accelerated). Einstein then postulated that light always travels in empty space with a definite velocity c which is independent of the velocity of the source. The value of c is approximately 186,000 miles per second (or 300,000 kilometers per second). However, to make these two statements consistent with each other Einstein had to make a third assumption of some "time-assigning" functions that define how clocks might be synchronized, using light signals, by two observers who are in relative motion.

Einstein's second and third assumptions will be discussed in Chapters 5 and 6 and they will be shown to be unnecessary, provided a correct validation of the advanced potential solution of Maxwell's equations is undertaken. One only needs to assume Maxwell's equations and the principle of relativity to establish special relativity.

In order to deal with uniformly accelerated frames of reference Einstein introduced general relativity in 1916. He noted that if a body is at rest in a gravitational field, then the body feels a force exerted on it. This force will accelerate the body, provided it is free to move. However, if the body is in free fall in a gravitational field then a force is no longer felt by the body. In the frame of reference of the free-falling body gravity is no longer felt and the gravitational force has, effectively, disappeared. This fact is known as the *equivalence principle*. Hence, one can eliminate the complications associated with the acceleration of a body, produced by a gravity field, if one performs a mathematical transformation out of the original reference frame and into the new frame of the free-falling body. But the equivalence principle only holds locally. If all uniformly accelerated frames are to be equivalent then Euclidean geometry cannot hold in all of them. A switch to Riemannian geometry solves this problem. The consequence is that the presence of matter is then considered to produce a warping of space-time in the new geometry.

As a result, general relativity is able to explain the apparently

instantaneous, action-at-distance, effects of gravity that are inherent in Newtonian theory if angular momentum is to be conserved. In general relativity the warping of space-time acts as an intermediate step. Thus the presence of any massive body, such as the Sun, warps space-time over a very wide region. This initial distortion travels away from the Sun at the speed of light, producing a *steady-state* warping of space-time spread over the whole of the solar system. According to general relativity the planets do not feel the force of gravity. They just follow the equivalent of “straight-line”, free-fall, paths in the new distorted space-time geometry. All gravitational forces have been eliminated and have been replaced by a change in the geometry.

General relativity is not a theory *of* gravity. It is a theory that eliminates the complications produced by gravitational forces and the accelerations produced by gravity. Hence, one should never expect general relativity to be able to predict the origin of gravity, or to predict the value of the gravitational constant G . General relativity assumes that the degree of space-time warping, at a given distance from the Sun, is predetermined by the given mass of the Sun and the given value of G .

Einstein spent over thirty years in attempting to unite the general theory of relativity with electromagnetic theory, but he was unsuccessful. Most scientists consider that Einstein wasted his energies on his later attempts at unification. Very few physicists or mathematicians were involved with the development of any form of *Unified Field Theory* until 1980. But since 1980 there have been many fresh attempts at unification. What do these attempts involve?

To the best of our present knowledge there are only four fundamental forces in nature. These are the gravitational force, the weak nuclear force (that controls radioactivity), the electromagnetic force, and the strong nuclear force (that holds the nucleus of the atom together). Each of these fundamental forces may be considered to act through its own set of fields. A complete *Unified Field Theory* would unite the four sets of fields associated with the four fundamental forces in nature. But no such theory has yet been discovered. From the time of Maxwell’s partial unification in 1864, up to the present time, there has only been one further partial unification. This provides a link between the weak nuclear force and the electromagnetic force.

We now need to go back a long way, in the order in which theories have been developed, and examine a very important further complication. Fields are not the only way of considering how the four fundamental forces can act.

Energy may also be considered to exist in minimum size packets, or quanta. For example, the energy contained in a single quantum of light, the photon, is equal to the frequency of the light multiplied by Planck's constant. This idea, discovered by Einstein, leads onto quantum theory. It also leads to the *uncertainty principle*, which I will briefly explain.

Suppose I am trying to find out the position of an electron and, at the same time, determine its velocity. To get the position of the electron accurately I will need to examine it using electromagnetic waves that have a very short wavelength because the limit of the accuracy of the experiment will be of the order of one wavelength. The wavelength of a given monochromatic (single-frequency) wave is inversely proportional to its frequency. I will therefore be using some very high frequency photons to examine the electron. However, we see from the previous comment on photons that high frequencies mean that each individual photon will have a lot of energy. Just a single photon will disturb the electron I am trying to observe and will change its velocity. Hence, there will always be an uncertainty in the combined accuracy of obtaining the position measurement and the velocity measurement for a single electron. If I try to get the position accurately I change the velocity, and if I try to get the velocity accurately I get a poor accuracy for the position measurement. Any classical measurement of a particular physical quantity will inevitably change what is being measured very slightly. However, this type of error can usually be calculated and then allowed for. But at the quantum level the error is much more fundamental. The uncertainty principle states that there is an inherent uncertainty about the way in which electrons, other atomic particles, or photons, interact with any material boundary. This uncertainty may be an uncertainty about the combination of the position and the velocity, or an uncertainty about the combination of the time of arrival and the energy.

One may then go one stage further and produce quantum field theories, which have a very different mathematics when compared with Maxwell's field theory. Quantum field theories have been very successful in making predictions in particle physics. But they have inherent limitations because a field appears to be able to exist at a point in space and, at the same time, there is no limit to the amount of energy that may be carried by a field at this single point. This fact produces many infinities in the theory. These infinities can be ignored, or renormalized, in particle physics applications. However, in all of the attempts that have been made to unite quantum field theory and gravitational theory the infinities are a complete block to making any

progress.

One further aspect of quantum field theory is of interest. The uncertainty principle states that there is an uncertainty when attempting to observe energy and time simultaneously at a given point. There is then the theoretical possibility of virtual particles, consisting of particle-antiparticle pairs, being created in a vacuum. These virtual particles are suddenly created, and they then very rapidly recombine and disappear. But during the very short time of their existence energy may exist in the vacuum. This idea is valuable for quantum field theory, and gives further knowledge of how atoms spontaneously emit photons (*e.g.* as shown by the yellow colour produced by a sodium salt when it is heated in a flame). Vacuum energy is also involved with the study of cosmological models and black holes. It is important to realize that the concept of vacuum energy can only be confirmed at the material boundary of some form of detector (the Casimir effect).

To get over the problem of the infinities at points in quantum field theory the idea of a *String Theory* has been proposed, where a string has finite dimensions. The vibrational mode of the string determines its frequency, its energy, its size, and which elementary particle it represents. The strings are then thought to exist in a rather unusual type of space that may have eleven, or more, dimensions. Three of these dimensions would produce our normal three-dimensional space. The fourth would represent time. The remaining dimensions would be wrapped up tightly and be invisible to us. These ideas can be taken further. In a more developed form of topology the strings can become points and surfaces. So far, no proposed string theory, or superstring theory, has been developed to the stage where an experimental check could be made.

A *String Theory* might hopefully become very near to being a *Theory of Everything*. Such a theory would go even further than a full *Unified Field Theory*. As well as being able to unify quantum field theory and gravitational theory it should be able to explain the presently accepted constituents of matter – quarks and leptons. It might also be able to explain some of the basic physical constants.

Examples of the basic physical constants are: the velocity of light, the charge on the electron, the masses of the electron and the proton, Planck's constant and Newton's gravitational constant G . The gravitational constant is most unusual because it appears to be the only physical constant that is totally unrelated to any of the other constants. An interesting, and very

significant, result comes about if we combine some of the constants in a particular way. We then get some dimensionless numbers. These are numbers which are both constant numbers and they also have no units. One example is the ratio of the mass of the proton to the mass of the electron, which is equal to about 1836. Another example is the fine structure constant, which is achieved by dividing the square of the charge of an electron by the product of the velocity of light and Planck's constant. The value of the fine structure constant is approximately equal to $1/137$. It is interesting to note that, on a classical basis, the electron in a hydrogen atom travels in an orbit around the nucleus at a velocity equal to the velocity of light c divided by 137. The fine structure constant lies at the basis of calculating how atoms are built up.

We may now return back to a much simpler level. When dealing with the assumptions and limitations of special relativity, in Chapters 5 and 6, we need only consider Maxwell's equations, Maxwell's field theory, electromagnetic waves and photons. It is important to re-assess the limitations of special relativity because so much that has followed special relativity is based on Einstein's initial assumptions.

Having discussed electromagnetic waves and photons as if they really existed, it is relevant to remind ourselves of the true situation, in a way that is often overlooked or ignored:

Electromagnetic waves and photons do not exist, as such, in free space. They are simply two alternative mathematical concepts that enable one to solve for the different effects that occur at specific *material* boundaries when one attempts to observe the electromagnetic energy that had *previously* existed in the region of free space that is now being occupied by the material of a detector.

A photon can only be postulated to exist as it is emitted or absorbed at a material boundary. A photon cannot exist in free space. If photons did exist in free space one would have to have photons changing their frequencies, and hence their energy contents, as they travelled in space towards a moving frame (the Doppler effect). Similarly, a monochromatic electromagnetic wave cannot exist in a finite region of free space because, by definition, such a wave is not permitted to have a beginning or an end. The presence of any realizable, material, detector will introduce very significant additional frequency components when the front edge of a quasi-

monochromatic electromagnetic wave arrives at the detector. The continued overlooking of these simple facts lies at the heart of the problems that exist in interpreting the mathematics of special relativity. The mathematics is flawless but the interpretation has been naive.

The fact that electromagnetic waves and photons do not exist, as such, in free space has a more general consequence. When any interaction occurs between electromagnetic energy travelling in free space and a material boundary then a mathematical wave analysis is *only* appropriate if an overall steady state, or near-steady state, has been reached. A mathematical photon analysis is *only* relevant if either the electromagnetic energy arrives as a precursor transient, or if the detection process involves a transient electromagnetic transition. There is no duality. For example, in the Young's slit experiment an overall steady state has been reached and only a wave analysis is relevant. A photon analysis is not appropriate even when light energy levels are reduced to nearly zero.

Most scientists agree that all of the well-accepted theories of both physics and cosmology can only be approximations to an ultimate truth that is unlikely to be fully understood in its deepest meaning. Although the basic laws of physics may be expressed quite simply, the history of science demonstrates that these simply expressed laws may have a much deeper significance than appears at first sight, and they usually need to be extended when applied under extreme conditions.

The classic examples of such basic laws of physics are Newton's three laws of motion and Newton's law of gravitation. For over two hundred years these laws were considered to be able to produce a precise description of all possible motions of matter in the Universe.

However, since 1905 it has been found that Newton's laws have had to be extended, following Einstein's development of the special theory of relativity. But any modifications of Newton's laws are only needed if the velocities involved are significant when compared with the velocity of light, or when gravitational effects are on the scale of the Solar system. For the vast majority of problems, including sending astronauts to the moon, we continue to use Newton's laws. Nevertheless, long before any of the complications of relativity arose, Newton considered that there were deep philosophical problems with his own laws! These problems have never been solved, and will be discussed in Chapter 3.

It may well be slightly misguided to try to produce a full *Unified Field*

Theory or a *Theory of Everything* before we have solved some of the many major problems in our fundamental understanding of Newtonian theory, electromagnetic theory, relativity theory, gravitational theory and quantum theory. At a really fundamental level we do not understand many basic aspects of physics and cosmology. It would seem more logical to attack these individual major problems with much more effort before attempting the grander problem of unification. I described one of these problems in the *Introduction*.

The aim of this book is to examine three of the major individual problems. After discussing Mach's Principle in Chapter 3, an alternative interpretation of Mach's Principle will be discussed in Chapter 4 that is capable of explaining a possible origin of gravitational forces. All of the ideas I will put forward concerning the origin of gravitational forces are speculative. The reason I include them is because there are possible terrestrial observations that could be undertaken which, if positive, would totally change our understanding of the origins of gravity. This alternative explanation of Mach's Principle gets rid of the implausible explanation of the origin of the inertial reaction force that was discussed in the *Introduction*. In addition, and more importantly, it establishes a possible gravitational origin for the four fundamental forces of nature. At the atomic level all four forces appear to be *gravitational* forces arising from the specific spin and orbital rotations of atomic particles. A single equation is able to predict the wide diversity of the magnitudes of these forces.

The rest of the book is not speculative. It is shown that Maxwell's equations, which provide the basis of special relativity theory, contain much more information than has been recognized. For nearly a century physicists and electrical engineers have failed to correctly interpret the advanced potential solution of Maxwell's equations. It is the advanced potential solution that leads to the concept of advanced electromagnetic waves. Until very recently all authorities claimed that advanced electromagnetic waves could only exist in an imaginary world where time was running backwards. Having assumed this "fact" they argued that individual advanced potential solutions needed to be forcibly rejected from Maxwell's equations. However, no theoretical analysis has ever been found to justify this claim, despite a great deal of searching. Some eminent authorities even went so far as to suggest that Maxwell's equations needed to be revised.

Maxwell's equations are at the basis of electromagnetic theory and special relativity. It is important to realize that special relativity is

inherently subordinate to Maxwell's equations. Any failings of Maxwell's equations would inherently be reflected in both special relativity and general relativity. However, it may be shown that Maxwell's equations are not compromised. The incorrect rejection of the advanced potential solution relies on a causal argument, which happens to be valid in specific and commonly used circumstances. Nevertheless, a fuller analysis of the advanced potential solution shows that in other specific circumstances it is not only a valid solution, but it is also the necessary solution that forms the theoretical basis of special relativity. Special relativity only needs to assume Maxwell's equations and the principle of relativity. The other assumptions are superfluous. Special relativity is put on the same footing as quantum theory, in that *the act of observation*, using a material detector, *changes what is observed*. Uncertainty then enters into special relativity as well as quantum theory.

A deeper understanding of Maxwell's equations also throws new light on the well-known "clock paradox" or "twin paradox" of special relativity. Studying the apparently simple concept of having an electric field produced by a single charged body, located in an otherwise empty Universe, leads to an explanation of why moving clocks go slow.

Chapter 2

Galilean and Newtonian Theories

Galileo developed his theory of motion from considerations of free-fall, position and velocity. If two frames of reference are in uniform relative motion then Galilean transformations provide us with the means of calculating the effects of that motion, provided velocities are well below the velocity of light. It is interesting to note that Galilean transformations hold between *any* two frames that are in uniform relative motion. General relativity theory is developed from Galilean theory and has no direct links with either Newtonian theory or special relativity.

The Lorentz transformations of special relativity hold only between *inertial* frames that are in uniform relative motion because special relativity is developed from Maxwell's equations.

Galileo was the first scientist to make analytical statements about acceleration. He was the first to clearly state that the velocity of a body can only change if the body is acted on by a force.

Newton's theories arose from a consideration of acceleration, force and mass. Newton introduced the idea of there being a specific relationship between force, mass and acceleration. Newton's laws of motion are that, relative to an inertial frame of reference:

1. Every body continues in its state of rest, or uniform motion in a straight line, unless acted on by a force.
2. The acceleration of a body is proportional to the applied force and takes place along the line of the force.
3. Action and reaction are equal and opposite.

Newton's laws of motion were brilliant in their concept. They also contain hidden depths. In essence, Newton's laws of motion rely on the circular argument that there can be no acceleration without a force and there can be

no net force without an acceleration.

After considering the simplest linear statement of Newton's second law, namely, that the applied force exerted on a body equals the product of the mass of the body times the acceleration given to the body, one develops both sides of the equation, to suit more complex problems, so as to make sure an equality is maintained. If an equality is not achieved, one does not consider that Newton's laws of motion have failed, one simply looks for another term to restore the equality. As a first step one switches into three dimensions by treating force and acceleration as vector quantities. One may then consider a rotating frame. For the resulting curved path one adds a centripetal force term (to hold the mass in its curved path) and, if the radius or angular velocity change, one adds a further Coriolis force term. The Coriolis force makes bath water rotate anti-clockwise as it goes down the plug hole, in an ideal bath, in the northern hemisphere and clockwise in the southern hemisphere. It also has a major effect in making weather patterns rotate as they move across the Earth's surface.

Finally, one may add the completely new concept of a gravitational force term by introducing Newton's law of gravitation. Hence, given certain limitations, Newton's laws of motion can never be wrong – the beauty of them is that they act as a guide to achieving an inevitable equality. But Newton's laws tell us nothing about the fundamental natures of mass, gravitation, or even force.

Newton's laws rely on four primary assumptions:

1. The mass of a body is assumed to have a constant value.
2. Gravitational mass is assumed to be equal to inertial mass.
3. The gravitational constant G is assumed to be a universal constant.
4. An unvarying, universal, time-scale is assumed to exist.

All of these assumptions are open to experimental testing. Newtonian mechanics does not distinguish between laboratories that are in uniform motion, and Galilean transformations will apply to calculations used to relate the results applying to two frames which are in uniform relative motion. Newton did not need to assume that an absolute frame of reference existed in space, although he frequently referred to the background

provided by the Universe. Newtonian mechanics is also time symmetric, and all results will be the same if the “arrow of time” is reversed.

There are two main problems with Newtonian theory. First, Newton’s law of gravitation requires changes in gravitational forces to be transmitted instantly between distant bodies (action-at-a-distance) if angular momentum is to be conserved. It was mentioned in Chapter 1 that general relativity gets over this problem by introducing the geometry of warped space-time as an intermediate step. But note that a *steady-state* warping of space-time over a given region, by any massive body, is initially set up at the speed of light. Subsequently, the effects of gravitation on any other body will appear to occur locally.

Secondly, although Newton’s laws can readily deal with rotational motion (unlike general relativity) they cannot explain the reason behind the absolute nature of rotational motion. This problem will be discussed in the next chapter.

Chapter 3

Mach's Principle

It was Newton who, having built up apparently impregnable theories of mechanics and gravitation, suggested that there were very deep philosophical problems with the concepts involved in describing inertial and gravitational accelerations. These problems are particularly significant when one considers the absolute nature of rotational acceleration.

Newton illustrated the most important of these problems by describing what has come to be known as “Newton’s bucket” experiment. Consider how the surface of the water, in a half-filled bucket of water, changes as the bucket is rotated about a vertical axis that passes through the centre of the bucket. When the bucket and the water are at rest the surface of the water is flat (except for surface tension and Earth curvature and rotation effects). An initial acceleration of the bucket produces negligible change in the surface of the water. But as the water is gradually accelerated, by the effects of drag, the water begins to rise up the sides of the bucket. Finally, at a constant angular velocity of both the bucket and the water, when the relative velocity between the bucket and water is zero, the curvature of the surface of the water is at a maximum. It is clear that the curvature of the water surface does not depend on the relative velocity of the bucket and the water. Neither does it depend on the relative velocity of the water and the Earth. The distortion of the surface of the water, due to rotation, depends on the velocity of the water relative to the background provided by the distant stars. But why is this so? How can the water know that it is the relative velocity with respect to distant stars which is the determining factor in controlling the surface shape?

Another simple demonstration of the absolute nature of accelerated motion is illustrated by a Foucault’s pendulum. Such a pendulum can be seen swinging in most science museums. If a pendulum bob is suspended by a very long wire, and set swinging over a long period of time, the line of swing is observed to remain constant relative to the distant stars. The

surface of the Earth is seen to gradually rotate under the line of swing of the pendulum.

Ernst Mach examined many problems associated with the absolute nature of acceleration. He produced some important conclusions about the way that the distant matter in the Universe may be affecting the fundamental laws of physics. One of his very simple arguments, that showed the difficulties with Newtonian theory, went as follows. Consider two fluid stars in the sky, that are well separated from both each other and also any other matter. One of these stars is stationary and has a spherical shape. The other star is spinning about an axis that passes through the centre of the first star, and it has an oblate shape due to its rotation. Now let all of the rest of the matter in the Universe be gradually removed. In the limit we are left with only one spherical star and one oblate star in the Universe. One can detect that the latter star is spinning, because of its oblate shape, about an axis that is common to both stars. But there is no external frame of reference. To have one star spinning and oblate, and the other star stationary and spherical, is not acceptable. As they can only relate to each other, why is there a difference between them? One may go one stage further and remove the spherical star. We now have a single star spinning in an otherwise empty Universe. The fact that it is spinning may be detected by local observation on the star itself, because it is oblate, but what is it spinning with respect to?

Mach's answer to this problem was to postulate that our basic concepts of the mass of a body, the acceleration a body is given, and the gravitational attraction between masses, must somehow depend on specific causes lying well outside the immediate system being considered. Mach suggested that it is the totality of matter in the Universe which, in some way yet to be defined by experiment, gives rise to our local concepts of mass, gravitation and acceleration forces. Mach is suggesting that if a distant part of the Universe were to be removed then some of our basic laws of physics might change – for example, the value of the gravitational constant G might change.

Mach also suggested an extreme deduction that if, in Newton's bucket experiment, the bucket were to be kept stationary and the whole of the rest of the Universe were to be rotated, then the water would rise up the side of the bucket, just as it does in the usual version of the experiment. These very general ideas put forward by Mach are open to many different interpretations.

If an attempt is to be made to extract a specific new theory from Mach's Principle, then it is necessary to define some of his arguments rather more precisely. It is fair to state that the main thrust of Mach's original argument was to propose that any changes in the *static* distribution of matter in the Universe may produce variations in the local *inertial* effects felt by any body. Having then decided on the essence of Mach's argument, can we be sure that we have exhausted all possible interpretations of the anomalies associated with rotational motion? The answer is no. We may invert the argument previously made and suggest that it may be the *inertial motion* of distant matter in the Universe that gives rise to local *static* effects. Such a change in the local static effects could be a change in the value of the gravitational constant G .

Einstein was greatly influenced by Mach and he attempted to fully incorporate his ideas into relativity theory. However, any Machian effects in general relativity cannot explain the two-stars problem outlined above.

But Einstein was very clear about Mach's two-stars problem. In his 1916 paper he stated what has now become to be known as Mach's Principle:

We have to take it that the general laws of motion, which in particular determine the shapes of (the two stars) S_1 and S_2 , must be such that the mechanical behaviour of S_1 and S_2 is partly conditioned, in quite essential respects, by the distant masses which we have not included in the system under consideration. These distant masses and their motions relative to S_1 and S_2 must then be regarded as the seat of the causes, which must be susceptible to observation, of the different behaviour of our two bodies S_1 and S_2 .

Put in modern terminology, Einstein is stating that the distant masses *and their motions relative to local masses* must affect either Newton's laws of motion, or Newton's law of gravitation, or both. But Einstein was unable to fully embrace this concept within general relativity. General relativity can introduce Machian effects, but these effects can always be transformed away by a suitable change of coordinate system and geometry. General relativity is too general!

Physicists have usually ignored part of Einstein's statement concerning Mach's Principle and they have assumed that the law of gravitation is fixed. As was stated in the *Introduction*, unconvincing attempts have been made to show how the local laws of motion might be affected by the relative

motion of distant matter. It is currently postulated that the origin of any local inertial reaction force is a gravitational interaction between the body being pushed and the distant matter in the Universe. But inertial reaction forces are instantaneous. It then has to be assumed that the changing gravitational forces from this distant matter must be continually flowing from both future time and past time, as advanced and retarded gravitational waves, so as to artificially combine in present time and thus produce an instantaneous inertial reaction force.

Additionally, the standard approach of trying to relate local inertial reaction forces to gravitational waves coming from distant matter does not embody Einstein's belief that atomic particles are gravitationally stabilized.

Why has Einstein's other option for developing Mach's Principle been ignored? Instead of assuming that local inertial reaction forces are produced by a gravitational interaction with distant matter, one may propose that local gravitational forces are generated by the relative motion of distant matter. This new approach, which is discussed in the next chapter, suggests that the motion of the distant masses, relative to local masses, may change the law of gravitation in some way. The most likely change would be a variation in the value of the gravitational constant G . After the distant masses have established a steady-state value of G , inertial reaction forces would be instantaneous against this background of established gravitation. To meet Einstein's interpretation of Mach's Principle one only has to relate *either* Newton's laws of motion *or* Newton's law of gravitation to the effects of distant matter. However, in using the proposed new approach, the manner in which G is generated must not lead to any conflict with general relativity, and the well established gravitational effects that have been observed in the laboratory, the solar system, and the Universe at large.

While leaving general relativity and the free-space value of G unchanged, the new approach proposes that the rotational motion of the quasi-fluid Earth, relative to the distant masses in the Universe, may give rise to an increase in the value of G equal to about 0.4% when it is measured well below the surface of the Earth. There is already some evidence for this increase that has been provided by both Frank Stacey, who took measurements in an Australian mine, and myself (see next chapter). Some recent measurements, obtained at many different laboratories across the world, show unexpected variations of G of up to 0.6%.

In the next chapter we start by examining the origin of gravity from a different point of view. Instead of invoking Mach's Principle at the outset

we consider the experimental evidence. The most accurate experiments ever undertaken in physics demonstrate that the gravitational mass of any given body is precisely equal to its inertial mass. Why should these apparently very different forms of mass be identical?

Chapter 4

The Origin of Gravity, Weak Forces and Strong Forces

Physicists have no idea why the gravitational mass of a given body is equal to its inertial mass. Until this basic problem of classical physics is sorted out, it is most unlikely that gravitational theory can be united with quantum theory. The fundamental interpretations of gravitational theory have hardly changed since 1916, whereas the interpretations of quantum theory have advanced by leaps and bounds over the years since 1916. It is therefore likely that the problems with the unification of the two theories will be solved by a better understanding of gravitational theory.

Because we are so familiar with the concept of the mass of a body, we tend to forget that there is a fundamental difference between the gravitational mass of a body and its inertial mass. Consider a simple pendulum:

When the pendulum is given a small push, to start it swinging, the pendulum bob would continue to revolve continuously forever, in a vertical circle, (assuming the right sort of pivot), if it were not for the effects of gravity and friction. The restoring force, which makes the pendulum swing to and fro, is the force of gravity acting on the bob. When one calculates the formula for the periodic time of swing of the pendulum, one finds that the periodic time depends not only on the length of the pendulum and the acceleration due to gravity, but also on the ratio of the gravitational mass of the bob to its inertial mass. However, whatever material is chosen for the bob, this ratio is always found to be unity. Hence, *experiment* shows that the gravitational mass of any given body is equal to its inertial mass.

Why should all of the different atomic particles within a body interact in just the same way with a gravitational field as they do when the surface of the body is given a push? There is no agreed explanation. We have had to

accept that the gravitational mass of a given body is identical to its inertial mass.

But we can go further that just accepting this equality as a fact. When it is found that any two items are identical, we always deduce that the two items must be fundamentally related to one another and have a common origin.

The obvious question then arises, how may the two identical forms of mass be considered to have a common origin? If one accepts Einstein's statement concerning Mach's Principle, then one is left with only two options. Either every local inertial reaction force must be caused by an instantaneous gravitational interaction with distant matter (the currently accepted viewpoint of many physicists), or the value of the gravitational constant G must result from the motion of local matter relative to distant matter.

Up until now everyone has considered it ridiculous to suggest that the large-scale inertial motion of matter could generate gravitational forces. However, this idea may not be so outrageous as it appears at first sight.

In 1964 Hoyle tried to partially introduce Mach's ideas by suggesting that the mean density of matter throughout the Universe might establish the value of G . Hoyle predicted that if the mean density of the Universe halved then the value of G would double. It seems logical to extend Hoyle's idea to include the *motion* of distant matter, as well as its density, when predicting G .

A number of motions of distant matter, that would fully incorporate Mach's Principle and establish a specific free-space value for G , are possible. Alternative motions of distant matter range from the radial motion associated with the Hubble expansion of all matter in the Universe, to the rotational motion associated with the possible rotation of the Universe as a whole. In one of two possible models I have proposed in the past (see Appendix, Paper 1) it was suggested that the rotation of a galaxy, relative to the distant galaxies in the Universe, might generate the value for G that applied just within that galaxy. Such a model might have enabled one to reject both dark matter and dark energy at a stroke! However, comments from readers of my web site have shown that recent astronomical observations may contradict such a galactic model on two grounds.

First, it is now known that irregular galaxies do not rotate significantly, and yet they have a vigorous ongoing star formation that conforms, according to current theory, with the usual value of G . Secondly, our galaxy has only rotated through about 30 revolutions since it was formed. But,

according to current theory, the spiral arms should not have had time to form within so few revolutions! Hence the galaxy is not a good starting point for us to commence any search for the origin of gravity.

Current astronomical evidence appears to be clear. When related to the currently accepted models for the formation of stars in our Universe, all astronomical observations indicate that the free-space value of G , within the whole of the observable Universe, has been virtually constant for the past 10^{10} years. This requirement puts severe restrictions on any inertial model for the origin of G . But all is not lost. We are becoming familiar with the fact that the Universe we observe may be a smaller entity than we have previously thought, and there may be many universes within a much wider horizon.

If we wish to relate the creation of the particular free-space value of G in our Universe to an inertial motion relative to distant matter, then the most likely model is one where the value of G is being directly created by the rotation of our whole Universe against the background of other distant universes. It is unlikely that we should ever be able to detect this rotation directly, and there are difficulties with this model. Nevertheless, in view of the problems with all of the alternative interpretations of Mach's Principle, it is worth pursuing.

There may be further consequences if gravity is created in this way. It is possible that small perturbations in the observed local value of G might occur within the boundary surface of a near-fluid body, such as the Earth, which is rotating relative to the distant matter in just our Universe.

First, a possible observable variation of G was proposed in Paper 1 (see Appendix). It was suggested that the rotation of an idealized fluid Earth might produce an abrupt, discontinuous, increase of about 0.4% in the value of G when crossing the boundary surface into the interior. For the real, near-fluid, Earth it is likely that most of this 0.4% increase of G will occur within a shell that straddles the surface level of a hypothetical, idealized, fluid Earth. The thickness of this shell might vary from around a few meters in land areas close to sea level, to a few kilometers in moderately hilly areas. To be certain of observing this spatial change in G two separate measurements of G would have to be made at points about 1km above and below sea level, at carefully selected sites. Many recent measurements of G have differed by up to 0.6% from the expected mean value.

But there may be further consequences if the shell is only a few meters thick at some particular land areas. Such land areas might be flat land areas close to sea level in elevation. As the Earth rotates and orbits the sun, some

of these locations, which are fixed on the Earth's land surface, may pass back-and-forth across the hypothetical boundary surface of an idealized, fluid Earth. At these points a monthly and annual variation in the value of G might be anticipated.

I analysed all twenty six pairs of the original laboratory test results for G published by the National Bureau of Standards, Washington in 1930 and 1942 (L. M. Stephenson, Proc. Phys. Soc., **90**, 601, 1967, and Found. Phys., **6**, 143, 1976). The analysis demonstrated that a 0.2% sinusoidal annual variation in G at Washington, with a maximum occurring at the vernal equinox and a minimum at the autumnal equinox, reduced the spread of the mean values of the three sets of G results published in 1930 by a factor of 3. For two more accurate sets of G results, published in 1942, the spread of the mean values of the two widely separated sets was reduced by a factor of 12. A statistical analysis shows that the probability of these reductions in the spreads of the means arising by chance, if an arbitrary variable had been applied, is less than one in ten thousand. The probability of the existence of an annual variation of G at Washington is therefore significant.

It is interesting to note that, when publishing the 1942 G results, Heyl and Chrzanowski concluded with the remark: "...what is needed to account for the observed anomaly in the results with the two (suspension) filaments is a regular variation of such nature as to be incredible." Heyl and Chrzanowski could not have seriously considered the possibility of a 0.2% annual variation of G , as they would have assumed that this variation would also require a totally unacceptable 0.2% annual variation in the acceleration due to gravity g . It is well known that pendulum clocks do not vary in their timekeeping by 0.2% over the course of a year!

But most of the proposed 0.4% increase of G is predicted to occur only when crossing a narrow shell, ranging in thickness from a few metres to a few kilometres at the Earth's surface. This specific local increase of G will produce a much smaller increase in the value of g . The value of g depends on the mean value of G taken over the total path length between the centre of the Earth and a test particle. The related increase of g will be about 1 part in 10^6 when crossing a 1km thick shell at the Earth's surface. It is important to realize that this suggested change in g is much smaller than, and additional to, any normal change of g which occurs with variations of height both above and below the Earth's mean surface level.

For the particular flat land area locations close to sea level, which might lie within a possible 10m thick shell, the related annual variation of g within the shell will be about 1 part in 10^8 . This annual variation of g might be

observable by comparing the readings of two gravimeters. The accuracy of gravimeters is now of the order of $2\mu\text{Gal}$, which is equivalent to a change in g of 2 parts in 10^9 . If one gravimeter were to be located at the site in Washington where Heyl and Chranzanowski undertook their G experiments, then this gravimeter might show an annual variation in its reading of about $10\mu\text{Gal}$ for the assumed shell thickness of 10m at this point. There would be a maximum at the vernal equinox and a minimum at the autumnal equinox. A second identical gravimeter, located in the same region, but at a height of at least 100m above or below the first gravimeter, would show negligible annual variation in its reading due to this specific effect in the 10m shell. The proposed annual variation of g would be additional to the normal annual variations of g , due to solar tides, of about $40\mu\text{Gal}$. However, the solar tide effect produces a minimum value for g at both equinoxes.

Within the past sixteen years many measurements of G have been made, in laboratories across the world, that show some unexpected variations of G of up to 0.6%, when errors of 0.05% were expected. It would be worthwhile to study these results to see if any annual variations of G were present for those laboratories that were close to sea level, and also to check whether these variations were dependent on latitude.

It was also shown in Paper 1, using equation (4), that the spin of an electron is predicted to create a very large value for G *within* the electron that is more than sufficient to stabilize the electron against the internal electrostatic repulsion force. Any spinning atomic particle would be similarly stabilized gravitationally, in a way that Einstein believed should occur. This result removes the need to either ignore the stability problem of the electron, or to assume arbitrary short-range “strong” forces to explain the stability. In obtaining the gravitational stability result for the electron, given in Paper 1, the classical value for the Bohr radius of the electron was used. I defend the use of the classical value for this case. “The complementary views provided by both classical and quantum pictures are both essential to the understanding of nature” (a quote by Freeman J Dyson concerning the analysis of the uranium 236 nucleus).

One is left with the important deduction that if the internal stabilizing force for all spinning atomic particles is a gravitational force then an initial link has been established between quantum theory and gravitational theory.

Specific gravitational origins for the two remaining fundamental forces, calculated using equation (4) in Paper 1, have also been proposed (L. M. Stephenson, J. Phys. A, **2**, 475, 1969). By inserting the angular velocity of

the electron in the first Bohr orbit, and the classical density of the electron, one obtains a gravitational force that equates with the magnitude of the weak force. By inserting this same angular velocity, and the density of the electron when spread out over the volume of the Bohr atom (which simulates the wave nature of the electron) one achieves a gravitational force that equates with the magnitude of the electromagnetic force within an atom. Hence, by combining the only two angular velocities of the electron in the Bohr atom with the only two densities of the electron in the atom, a *single equation* predicts three specific *gravitational* forces within the atom which correspond to the known magnitudes of the weak, electromagnetic and strong forces. It is inconceivable that these three correspondences arise by chance.

Details of the calculations from equation (4) are given on the next page.

Conclusions

If one accepts Einstein's statement concerning Mach's Principle, then one is left with only two viable options. Either every local inertial reaction force must be caused by an instantaneous gravitational interaction with distant matter, or the value of G must result from the motion of local matter relative to distant matter. For the second option, the most likely relevant motion would be a rotational motion of our Universe relative to distant universes. Some observable, local variations in G might occur.

A common gravitational origin for each of the four fundamental forces in nature is predicted, based on equation (4) in Paper 1. All of these forces appear to arise from the rotational inertial motion of matter. The usual gravitational force probably arises from the rotation of the Universe as a whole. Weak, electromagnetic and strong forces all seem to be additional, localized *gravitational* forces that arise directly from the spin and orbital angular velocities of electrons, and other atomic particles, within atoms.

Derivation of Gravitational Coupling Constants

Equation (4) from Paper 1 predicts the free-space value of G as:

$$G = \frac{\omega^2}{2} \rho k \quad (4)$$

where ω is the angular velocity of the Universe, ρ is the mean density, and k is a boundary shape constant (≤ 1). It is proposed that the free-space value of G is being generated by the rotational motion of the Universe.

Other values of G may be generated within the boundary surfaces of smaller rotating bodies. For the values of ω which apply to an electron's orbital and spin angular velocity in a Bohr atom this single equation predicts three further localized values of the gravitational constant at the atomic level, denoted by $n = 2$, $n = 3$ and $n = 4$.

Coupling constant
($e^2/m^2 = 1$)

$n = 1$ The usual free-space value for the gravitational constant G 10^{-39}

For an electron in a Bohr hydrogen atom of radius a_0 and with $n = 3$ relating to an electron being spread out over the volume of the atom to simulate the wave nature of the electron:

$$n = 2 \quad \omega_{\text{orbit}} = \alpha c/a_0 \quad \rho_e = 3m/4 r_e^3 \quad G_w = 2\alpha^6 e^2/3m^2 \quad 10^{-13}$$

$$n = 3 \quad \omega_{\text{orbit}} = \alpha c/a_0 \quad \rho_e(r_e/a_0)^3 = 3\alpha^6 m/4 r_e^3 \quad G_{em} = 2e^2/3m^2 \quad 1$$

$$n = 4 \quad \omega_{\text{spin}} = h/4\pi m r_e^2 \quad \rho_e = 3m/4 r_e^3 \quad G_s = e^2/6\alpha^2 m^2 \quad 10^3$$

$$\text{Where:} \quad r_e = e^2/mc^2 \quad \alpha = 2\pi e^2/hc \quad r_e/a_0 = \alpha^2$$

The last three coupling constants correspond with the relative magnitudes of weak, electromagnetic and strong forces. They arise from a single equation when combining the three viable combinations of ω_{orbit} , ω_{spin} , ρ_e and $\rho_e(r_e/a_0)^3$. Hence, a gravitational origin for the weak, electromagnetic and strong forces within an atom is indicated.

Chapter 5

How Antennas Work

Einstein made three assumptions when developing special relativity. These assumptions were essential to enable him to predict the transit times of light signals passing between inertial frames of reference. But light signals are only one particular form of electromagnetic signalling. Hence, an appreciation of the precise way in which all electromagnetic signals are transmitted and received is crucial if we are to understand the significance of Einstein's assumptions. We therefore need to consider both transmitting and receiving antennas in some detail.

In 1963 I was asked to write an article for the "New Scientist" magazine on "How Aerials Work". However, for most technical discussions nowadays one usually refers to an aerial as an antenna. Although I had lectured on antennas, as part of an electromagnetic theory course to university students, writing this article forced me to think about discussing antennas in a simpler way. The action of a dipole antenna, of the type that may be used to receive television signals, is more complicated than it seems. However, these complications may be described quite simply.

A dipole receiving antenna consists of a metal rod split in the middle. The output is taken from the break in the middle using a coaxial cable. The dipole is made half a wavelength long, for the particular frequency of the signal being received, and it is set vertically to receive a vertically polarized signal, or horizontally to receive a horizontally polarized signal. To make the antenna more directional, and hence more sensitive if it is correctly aligned, one may add a passive reflector behind the dipole and a number of passive director rods in front of the dipole. All of these separate elements of the antenna are mounted onto a supporting tube. This then becomes the well known Yagi antenna array one sees on many chimneys. However, we need only consider the basic single dipole antenna for our discussion.

A key point to note at the outset is that the dipole could be made of vanishingly thin wire (provided the wire is a good electrical conductor) and

it would still receive electromagnetic wave signals in just the same way. The dipole then has a negligible physical cross-sectional area facing the transmitter. How can the antenna receive energy, in just the same way, when it offers a negligible area facing the transmitter? Theory tells us that the antenna still has an *effective* area which is roughly equal to the square of its length, provided it is half a wavelength long for the frequency of the signal being received.

But is the theory right? Of course it is, provided one uses the theory within the limits of the assumptions that were made initially. The theory is based on receiving steady-state electromagnetic waves having a particular frequency. However, as is well known, electromagnetic energy is not always received as a wave. Quantum theory tells us that the energy sometimes appears to come in packets, called photons. Considerable complications to the theory then start to emerge if the frequency of these photons does not correspond to the design frequency of the antenna. Note that the energy of an individual photon is equal to its frequency multiplied by Planck's constant.

A diversion is now necessary to achieve a basic understanding of how a quantum approach may be needed when considering the arrival of a pulsed signal at a dipole antenna. It is beneficial to consider initially the simpler problem of how electromagnetic energy travels in a dielectric medium. This problem is also not as straightforward as it might appear at first sight.

When an electromagnetic wave enters a block of dielectric, such as the polythene used to fill a coaxial cable, the wave slows down to a velocity equal to c divided by the square root of the relative permittivity (or dielectric constant) of the dielectric ($c/\sqrt{\epsilon_r}$), where c is the velocity of light in free space. If, instead of an electromagnetic wave signal, a pulsed signal is sent down the cable, with a pulse length of only about one microsecond, the transmission velocity of the whole pulse will still be equal to $c/\sqrt{\epsilon_r}$.

But things get much more interesting if we consider what happens to the very front edge of a nearly-perfectly-shaped rectangular pulse of energy. The very first part of the energy (known as the first precursor transient) will travel through the dielectric at a velocity equal to c , the velocity of light in free space, no matter what type of dielectric is filling the cable. The reason for this effect is simple. A dielectric can only exhibit a relative permittivity greater than unity after some electromagnetic energy has interacted with the dielectric. It is the energy extracted from the first precursor transient which causes electronic, atomic, and molecular dipoles to be formed in the

dielectric material. It is only after the formation of these dipoles that the dielectric can exhibit a relative permittivity greater than unity.

However, it is difficult to demonstrate this effect experimentally. Any detector used to observe the arrival of a precursor transient will itself contain some material in which electrons have to respond, and then a detecting instrument has to change its reading to record this response. It is fairly easy to demonstrate the effects of molecular polarization of the dielectric, much trickier to demonstrate electronic polarization, and it is *impossible to detect* the very front edge of the first pre-cursor. We will discuss this last point in a moment.

Hence we may conclude that it is the *interaction* of the front edge of a pulse of electromagnetic energy with any dielectric, over a finite period of time, which is the direct cause of the slowing down of the main energy content of the pulse to $c/\sqrt{\epsilon_r}$, relative to the dielectric material.

But where does quantum theory come in? The answer is that the first precursor transient, and the arrival of the first few photons of energy, are one and the same thing. We spoke of the arrival of a nearly-perfectly-shaped rectangular pulse of energy at the dielectric. If one examines the front edge of an energy pulse Fourier analysis tells us that it is made up from a large number of sinusoidal frequency components. This is why one needs a wide bandwidth amplifier for amplifying a radar pulse, or for a high quality audio system. The front edge of a hypothetical, truly perfect, rectangular pulse would need an infinite frequency component. But the energy contained in each photon is equal to Planck's constant multiplied by the frequency. An infinite frequency photon would have infinite energy, and so a perfect rectangular pulse is an impossibility. The arrival of the first precursor transient corresponds with the arrival of the first few high-energy photons at the dielectric. It is the energy from these first few photons that sets up the dipoles in the dielectric material. Note that quantum theory also tells us that there will be an uncertainty about the arrival time of the first few photons.

We may now return to analyse how either a transmitting antenna or a receiving antenna operates. It will be shown that a receiving antenna *cannot produce any output* until some stored energy components are built up by a precursor transient. The approach velocity of the wave is then changed.

First, let us consider the radiation of some electromagnetic energy out into free space. To achieve this radiation of energy a transmitting antenna has to be used. Theory then shows that the electric and magnetic field components associated with such an antenna are quite numerous. Not only

does the antenna produce field components which give rise to a net radiation of energy out into free space (which travels away at c), but the antenna also produces many additional “near-field” components. Provided the antenna has been transmitting for some time (the steady-state condition), then these additional field components do not give rise to any net radiation of energy, but simply correspond to energy oscillating backwards and forwards in space, either radially or circumferentially. However, when a transmitting antenna is first switched on (the transient state) these additional field components have to be set up, at any given point in space, before any energy is radiated further out into space. One can visualize these additional field components in the same way as the stored energy in a dielectric, and they act as an invisible web of stored energy which *requires* the radiated electromagnetic waves to travel away from the transmitting antenna at a velocity equal to c relative to the transmitting antenna. We appear to have an old-fashioned source theory of radiation, provided that we are talking about the steady-state propagation of electromagnetic *waves*.

Now we can take the crucial step and consider a receiving antenna. It is important to realize that a receiving antenna does much more than simply pick up the energy that would otherwise have passed by in free space. The receiving antenna interacts with the electromagnetic energy and produces major changes in the field distribution that existed prior to the antenna's presence. A golden rule in science is to appreciate that it is impossible to observe any quantity without changing the quantity that is being observed. For most macroscopic experiments the change produced by measuring any given quantity is small. But in the case of a receiving antenna, that is moving relative to the emitter of energy (or vice-versa), there is a major effect.

When a receiving antenna detects some electromagnetic energy a current is induced in the antenna. According to well-established reciprocity theory, the steady-state field solution for an antenna receiving energy from free space is directly related to *the current induced in the receiving antenna* and is of exactly the same form as the field solution for an identical transmitting antenna. Thus the directional pattern of a given antenna when receiving energy is identical to the field pattern when transmitting energy. Hence, as well as the field components associated with the net transfer of energy into the receiving antenna, there will be additional “near-field” components that do not give rise to any net transfer of energy in the steady state. However, on the initial reception of an energy pulse at the receiving antenna, these

additional field components will have to be set up before any energy can be extracted from the antenna. It is because of this *interaction* of a receiving antenna with the arriving energy that these additional field components can again be visualized as an invisible web of energy which *requires* the electromagnetic waves to enter into the receiving antenna at a velocity equal to c relative to the antenna. In the steady state the additional field components produced by the current in a receiving antenna (the perturbation fields) extend out into space past the transmitting antenna and on to infinity. Hence, it is the current induced in a receiving antenna itself which controls the approach velocity of electromagnetic waves into the antenna in the steady state. We then have a source theory of radiation which is appropriate to the steady-state radiation of electromagnetic waves, and a sink theory of reception for the steady-state reception of electromagnetic waves.

What happens if a receiving antenna is moving steadily towards a source of electromagnetic waves at a relative velocity equal to $0.5c$? There will be no change in the observed velocity of the energy entering the receiving antenna, *in the steady state*, as the electromagnetic wave must enter the receiving antenna at a velocity equal to c relative to the current *already induced* in the receiving antenna by an earlier precursor transient. The additional field components associated with the current in the receiving antenna may still be considered to require the energy to flow into the antenna at a velocity equal to c . Hence, again it is the receiving antenna itself, and the current induced in it, which controls the approach velocity of an electromagnetic wave in the steady state. Note, however, that there will be the usually expected Doppler shift of the frequency of the electromagnetic wave.

But what happens in the transient state? What happens if an observer, with a receiving antenna, is travelling towards a transmitting antenna at a velocity equal to $0.5c$ and the transmitter is suddenly switched on? The initial precursor transient, associated with the very front edge of a pulse of energy, will travel out from the transmitter at a velocity of c relative to the transmitting antenna and a velocity of $1.5c$ relative to the receiving antenna. But the velocity of the first precursor transient cannot be observed. Before any energy can be detected by the receiving antenna the additional field components must be set up by the induced current in the receiving antenna – this is analogous to the dipoles having to be formed in a dielectric before the dielectric can affect the propagation velocity of an electromagnetic wave. In technical parlance, the effective area of the antenna will be equal

to zero until the scattered field has had time to expand.

It is only after the additional field components have been set up around the receiving antenna that energy may be detected in the steady state, and this energy will arrive at a velocity of c relative to the receiving antenna. The *observed* velocity of light will always be found to be equal to c in the steady state – this is all that special relativity theory both assumes and requires. *It is the act of observation* of an electromagnetic wave, using a material detector, which causes the observed steady-state velocity of the wave to be equal to c .

But note that there will be a time delay after a precursor transient arrives and before a signal can be detected. A signal can only be detected after the arrival of the first precursor transient has set up the “near-fields”. There will also be an uncertainty associated with the time of arrival of the first few photons which make up the first precursor transient.

The argument just given, which is based on an analysis using well-established reciprocity theory, is formally valid as well as giving a good visual picture of why the *observed* velocity of light is always found to be equal to c *in the steady state*. However a more formal approach, using advanced potentials is needed to confirm the conclusions that have been reached.

Chapter 6

Advanced Potentials, Quantum Theory and Special Relativity

The advanced potential solution of Maxwell's equations (which will be explained later) has produced major intellectual difficulties for electrical engineers and physicists for over a century. Until very recently all authorities had claimed that individual advanced potential solutions were invalid if attempts were made to apply them to receiving antennas in real time in the real world. This rejection relied on a simple causal argument because no theoretical analysis to eliminate the advanced potential solution, based on Maxwell's equations, could be found.

To be fair to the critics of the advanced potential solution, there is a significant problem with using the advanced potential solution. This problem arises from the slightly more complex geometry involved when receiving energy, when compared with the case for transmitting energy.

It is interesting to note that electrical engineers have needed to solve for receiving antennas, as well as transmitting antennas, for the best part a century. They have inadvertently managed to make a valid use of the advanced potential solution by applying well-established reciprocity arguments to the retarded potential solution. This approach neatly avoids having to analyse the full three-dimensional geometrical problem. In essence, they have considered just the energy that reaches a receiving antenna, without having to bother with the wasted energy from the transmitter that goes off into free space.

Advanced potential solutions are essential to understanding the relationship between special relativity, quantum theory and the limitations imposed on Maxwell's equations and special relativity by the ultraviolet catastrophe effect (explained on p51). There has been a failure to recognize that the ultraviolet catastrophe effect applies equally to the *reception* of electromagnetic radiation as well as to the radiation of electromagnetic

energy.

Retarded and Advanced Potential Solutions

Maxwell's equations express mathematically four experimental observations. These are: that an electric current produces a magnetic field, a time-varying magnetic field generates a voltage in a conductor, free electric charges may exist, but free magnetic poles do not exist.

Both retarded and advanced potential solutions appear whenever Maxwell's equations are applied to predicting the electric and magnetic fields associated with the radiation of energy produced by a transmitting antenna (or aerial). The simplest case to analyse is when an alternating current flows in an infinitesimal, thin-wire, antenna (a Hertz dipole). One may then integrate to obtain the solution for a given current distribution in any larger antenna.

During the analysis for a transmitting antenna the advanced potential solution is ignored. The retarded potential solution is selected to predict the fields produced at a distance r from an antenna when the antenna is transmitting energy. The fields are "retarded" because of the time taken for any signal to get from the antenna to the point in question. This time delay, t , is equal to r/c , where c is the velocity of light. The retarded potential solution is logical because one expects there to be a time delay before a signal reaches a distant observer.

The retarded potential solution involves only $+t$, whereas the advanced potential solution involves only $-t$. The advanced potential solution is mathematically identical to the time-reversed retarded potential solution. Until recently all authorities had ruled out the use of any advanced potential solution on its own, because they considered that the solution applied only when time was running backwards (*i.e.* in the opposite direction to the usual direction of the arrow of time, when an event would occur before the action that caused it, and a glass would lie broken on the floor before a person had dropped it). Such an effect can be demonstrated by taking a film with a motion picture camera and then running the film backwards. For example, when the picture is run backwards any rifle would be seen to be accepting bullets coming into its barrel from some distance, rather than the real-time sequence where bullets are fired from the barrel and travel into the distance.

However, an antenna, unlike a rifle, is a symmetrical device that has the ability to either transmit or receive energy. What happens if our antenna is now used to receive energy from some distant source?

The advanced potential solution predicts that any change in the fields at a distance r from the antenna occurs before the related current change in the antenna. The fields are then advanced in time by r/c relative to the current in the antenna. For a transmitting antenna this *would* mean that time was running backwards. However, we now wish to deal with a receiving antenna and not a transmitting antenna. For a receiving antenna we expect any change in the signal fields at a distance r to be advanced in time relative to the current change which is later going to be induced in the antenna by these fields. Analytically, in the time-delay equation, $t = r/c$, the sign of r is now negative *because the direction of propagation of the signal has been reversed relative to the antenna*. But time continues to flow in its usual direction in the real world. At first sight we then appear to have an obvious application for the advanced potential solution. We use the retarded potential solution for a transmitting antenna and the advanced potential solution for a receiving antenna.

But things are not so simple, and this is why problems have arisen. The analysis required when dealing with a receiving antenna is inherently more complicated than the analysis required for the simple case of having just a single transmitting antenna situated in an otherwise empty Universe. As we said earlier, when receiving energy there must be a minimum of both a transmitting antenna and a receiving antenna situated in an otherwise empty Universe. Most of the energy radiated by the transmitting antenna disappears off into free space and does not reach the receiving antenna. If we try to use the advanced potential solution for the reception of energy from an infinite, three-dimensional, region of free space, it is easy to show that it is an invalid solution. Consider the following simple argument.

Suppose one has a transmitting antenna and a receiving antenna, separated by a finite distance in an infinite free-space region, with both antennas being situated inside a very large imaginary sphere. The energy everywhere on the imaginary spherical surface must be travelling outwards. Hence, any fields arising from an individual advanced potential solution must be ruled out as they would result from energy travelling inwards.

However, there is a unique, and important, three-dimensional problem where an individual advanced potential solution must be applied. But, as might be expected from the above argument, one has to restrict the energy that may escape from the system through the imaginary spherical surface. This unique problem will now be described.

Let a Hertz dipole antenna radiate a single pulse of electromagnetic

energy into a free-space region which is bounded by a very large, spherical, conducting shell centred on the dipole. The relationship between the current in the dipole and the fields of the outgoing pulse is given by the retarded potential solution. On reaching the spherical, conducting shell the pulse is reflected and the direction of travel of the pulse is reversed. For the returning reflected pulse the relationship between the fields of the pulse and the current that is later going to be induced in the dipole is given by the advanced potential solution. But time continues to flow in the usual direction in the real world. The advanced potential solution is needed to account for the direction of travel of the pulse being reversed relative to the dipole. It is a valid solution because no energy escapes from the closed system. This specific three-dimensional configuration is the only case where an individual advanced potential solution may be used in three dimensions. But, it is essential that Maxwell's equations should predict an individual advanced potential solution, applying in real time, to cover this case. It was incorrect for many authorities to claim that the advanced potential solution demonstrated a weakness in Maxwell's equations. The negative sign associated with the time delay of the return signal must be interpreted as applying to the reversal of the direction of propagation of the signal in real time. Although the negative sign for the time delay could also be interpreted as a reversal in the direction of the arrow of time for a signal travelling in the positive direction of r , this solution must be discarded when one is dealing with any analysis applying in real time in the real world.

The Relationship with Special Relativity and Quantum Theory

Now comes the key point. If we use the advanced potential solution for the reception of energy in only one dimension there are no difficulties. By stating that we are using one dimension we are stipulating that a signal is being beamed directly from a transmitting antenna to a receiving antenna with only an insignificant loss of energy in between. The simple causal argument given above, that forbids advanced potential solutions when energy may be radiated off into free space, is no longer relevant. There is then no problem in using the advanced potential solution to relate the fields in the beam to the current that is later going to be induced in the receiving antenna.

Why is this very limited version of the advanced potential solution important? The answer is clear. A one-dimensional signalling system is

directly related to the assumptions made in establishing special relativity theory. Special relativity assumes that the velocity of light is always equal to c “in empty space” and independent of the source velocity. Einstein realized that such a statement was meaningless on its own, as one cannot measure the velocity of light “in empty space”. He then set up some additional time-assigning assumptions which involved signalling between inertial frames of reference. In particular, Einstein assumed that the out-and-return transit time of a signal travelling between two inertial frames A and B is equal to the total round-trip distance divided by the velocity of light c . It has not been appreciated that this signalling system is only applicable to steady-state, or near-steady-state, electromagnetic wave signals, and will not be valid for the front edge of a pulsed signal having an extremely rapid rise-time. For such signals precursor transients are produced in the material of any detector. Time delays will then arise at the reception of the signal at A and the re-radiation of the signal back to B .

A precursor transient occurs with the arrival of a nearly perfect, rectangular pulse of electromagnetic energy at any material boundary. As was mentioned in the last chapter, the front edge of a rectangular pulse may be analysed in terms of the sum of a large number of sinusoidal (Fourier) frequency components. For a perfectly shaped rectangular pulse the frequency components would have to extend to infinity and an infinite frequency photon, having infinite energy, would be required! In practice, some of the Fourier frequency components that make up a near-perfect pulse signal will be in the ultraviolet region or higher. This is where the well known "ultraviolet catastrophe" effect enters, and this effect limits the validity of both Maxwell's equations and special relativity when predicting the amounts of electromagnetic energy being radiated *and received* at very high frequencies. I will briefly explain the ultraviolet catastrophe effect.

Classical radiation theory originally demonstrated the ultraviolet catastrophe problem when analysing the radiation of energy from a black body. This classical theory predicted that the radiated power per unit frequency increased as the square of the frequency, and the total power radiated tended to infinity as the frequency is increased! It was clear that classical radiation theory could not be applied to frequencies in the ultraviolet region and beyond. The problem has to be solved using quantum theory. But there is another, unrecognized, ultraviolet catastrophe problem associated with the *reception* of precursor transient signals, as these signals also contain frequency components in the ultraviolet region and above.

Quantum theory is essential for analysing both the reception and the radiation of frequencies in the ultraviolet region and beyond. This brings us back to Einstein's signalling system that is needed to set up special relativity.

In order to observe the velocity of a pulsed electromagnetic signal, a *material* detector has to be placed in an otherwise empty region of space. As mentioned earlier, all detectors of electromagnetic waves may be analyzed in terms of an infinitesimal, or Hertz, dipole antenna. But such an antenna will only have a finite effective area, and hence can only start to detect a pulsed signal, once the scattered (or perturbation) field has had time to expand after the arrival of the pulse. After the scattered field has been generated by the arrival of the first few high-energy, high-frequency photons, the solution for the received signal is then given by the advanced potential solution of Maxwell's equations. This solution predicts that the steady-state arrival velocity of the signal is equal to c relative to the detector.

For a nearly perfect rectangular pulse signal, both Maxwell's equations and special relativity are unable to predict the initial precursor transient solution associated with the initial scattering of the field by the dipole. The precursor transient solution has to come from quantum theory, and quantum theory predicts that there will always be an uncertainty about the time of arrival of the first few photons associated with the signal.

It may be noted here that the transients observed with the pulses used for typical digital communication systems and radars are very far from being precursor transients, and special relativity will correctly apply to the near-steady-state analysis applicable to these systems. However, the latest systems are reaching the level where precursor transients are significant.

In essence, the observed velocity of a signal approaching a material detector is equal to c because *the act of observation makes it so in the electromagnetic steady state*. The advanced potential solution requires that any steady-state measurement of the velocity of light must yield a value of c . There is no need to make any assumption about the velocity of light in empty space. As well as demonstrating that special relativity and quantum theory are separate parts of a common solution, this analysis also puts special relativity on the same footing as quantum theory, in that the act of observation changes what is observed.

There is a common misconception that the appearance of the velocity c in the uniform plane wave solution of Maxwell's equations indicates that

the velocity of light must always be equal to c in free space. But this deduction is incorrect. Any hypothetical uniform plane wave experiment would require an unrealizable infinite plane source and an unrealizable infinite plane detector. Such an infinite plane detector hides the detailed nature of the interaction of a signal with any finite detector. When attempting to analyse special relativity, and its relationship with both Maxwell's equations and quantum theory, it is essential to consider how realistic observations of both the constant velocity of light, and precursor transients, may be made. To analyse how the steady-state velocity of light is observed one must always start with a Hertz dipole detector, and then integrate up for a larger receiving antenna. It is the material of the dipole that dictates that the steady-state velocity of any approaching wave is equal to c relative to the dipole.

Conclusions

The observation of any electromagnetic wave travelling in free space must be based, in the limit, on analysing the output of an infinitesimal, thin-wire (or Hertz) dipole antenna. The individual advanced potential solution of Maxwell's equations requires that the observed velocity of the electromagnetic wave must be equal to c relative to the Hertz dipole, and independent of the source velocity. No assumption is needed regarding the velocity of light in free space.

However, for the case of a nearly perfect rectangular-pulse signal a precursor transient precedes the steady-state condition associated with the arrival of the electromagnetic energy. The initial precursor transient consists of the arrival of a few individual, high-energy, photons. These photons will not produce any output from the Hertz dipole, but will be absorbed in the form of energy stored in the electric and magnetic near-fields around the dipole. Any analysis of these first few photons requires the application of quantum theory, and there will be an uncertainty about the time of arrival of the photons. It may be noted that a Hertz dipole is transparent to any signal until its *effective* area becomes finite and this only occurs after absorption of energy from the precursor transient to form the perturbation (or near) fields. In other words, until the scattered field has had time to expand.

It becomes clear why, when developing special relativity, Einstein had to assume that the velocity of light was equal to c in empty space, and why he also had to assume a time-assigning function that restricted signalling

between inertial frames to the use of steady-state electromagnetic waves. In 1905 quantum theory and precursor transients were unknown and, even in 1998, individual advanced potential solutions were misunderstood. Both of Einstein's assumptions are now unnecessary, but the mathematics of special relativity is perfectly valid provided it is restricted to the steady-state, electromagnetic wave, solution.

We may now summarize how electromagnetic waves propagate. The retarded potential solution of Maxwell's equations predicts that the departure velocity of an electromagnetic wave from a source must be equal to c . The advanced potential solution predicts that the arrival velocity of an electromagnetic wave at a detector must be equal to c . When there is relative motion between the source and the detector, the arrival velocity of the electromagnetic wave at the detector will still be equal to c , and independent of the source velocity. But how can this be so? Simply because the insertion of a material detector into a region of free space is a major change. The act of detecting the signal will bring the wave to an abrupt halt. A radiation pressure force is exerted on the detector and an equal reaction force is exerted on the wave. These forces will change if the source is in motion relative to the detector, and will ensure that the steady-state arrival velocity of the signal (*i.e.* the arrival of the wave) is equal to c relative to the detector. In the true steady state (after infinite time) the perturbation fields produced by the presence of the detector will extend past the source and to infinity. It is then not surprising that the steady-state approach velocity of a wave must be equal to c relative to both the detector and the perturbation fields, and independent of the source velocity.

The preceding analysis, using advanced potentials, was based on an initial assumption that any signal is beamed directly from a source to a detector, with only an insignificant loss of energy in between. However, the advanced potential analysis may also be applied to the general case of a wide-angle transmitting antenna that radiates energy into a three-dimensional free-space region. Provided there is linearity one may consider the component of the energy that is intercepted by the detector on its own, because this is the relevant observed part of the signal to which the advanced potential solution may be applied. The remaining component of the energy, that escapes to infinity, then requires a separate analysis.

Chapter 7

Time Dilation, the Clock Paradox, the Twin Paradox and Relativity

In the last chapter it was shown that special relativity is inherently a steady-state, or near-steady-state, theory. All precursor transient electromagnetic problems have to be analysed using quantum theory. Special relativity's major assumption – that the velocity of light is constant in empty space and independent of the source velocity – is unnecessary. The observed *steady-state* velocity of light in free space is always equal to c solely because *the presence of the observer's material detector makes it so*. This last deduction arises from the correct application of the advanced potential solution of Maxwell's equations.

Now let us consider the long-standing problem that exists in special relativity, namely time dilation. This problem is also often referred to as the “clock paradox” or the “twin paradox”. Although relativists declare that there is no paradox they still get into trouble when some aspects of Mach's Principle are raised. Einstein showed that there are difficulties in defining both inertia and an inertial frame of reference. At a really fundamental level one reaches a circular argument. Special relativity can only state that a particular frame is an inertial frame if this frame is found to be the one in which a standard clock indicates the largest passage of proper time between two events. Hence, one can only establish the result that one is trying to deduce by making an observation!

But let us look at the problem of the clock paradox and the twin paradox at a simpler level. There is no need to get involved with the epistemological difficulties within special relativity to solve this problem.

If an atomic clock is sent off on a long out-and-return journey, at a constant very high velocity relative to the Earth (except for three short acceleration periods at the start, at the turn-round point, and at the end) then this travelling clock will record the passage of less time for the journey than an identical atomic clock which stays behind on the Earth. Likewise, a

travelling identical twin, going on a similar out-and-return journey, will appear younger than the stay-at-home brother or sister when they are reunited. For brevity in the following discussion I will state that the travelling clock “goes slow”, but the more precise statement is that a travelling clock “records the passage of less time for the given journey”.

This effect was demonstrated very clearly by Hafele and Keating in 1971. They organized an experiment where some atomic clocks were carried around the world in two commercial aircraft. By making use of the fact that the Earth’s spin velocity is about 1000mph at the equator, one aircraft travelled round the world at a net speed of about 1500mph, and the other at a net speed of about 500mph. The slowing of the atomic clocks in both of the aircraft, when compared with atomic clocks on the Earth’s surface, accorded with time dilation predictions, although the recorded changes in clock readings were only slightly greater than the usual random variations of the clock readings. It is worth noting that appreciable gravitational corrections also had to be made to the clock readings because of the reduced gravitational field at the height of the aircraft. Clocks will run faster in a weaker gravitational field.

Following this early experiment, further confirmation is given by the fact that all of the clocks used in satellites to provide satellite navigation systems are adjusted so as to allow for the time dilation of the moving clock. However, this confirmation of time dilation is limited to the case of clocks that are accelerating in a free-fall orbit.

It seems difficult to explain the predicted asymmetrical slowing of a complete clock, or the reduced ageing of an identical twin, when using what appears to be a completely symmetrical theory like special relativity. Consider the case of the out-and-return journey experiment for identical twins. The Earth-bound twin sees the other twin travel out and back at a high velocity. But the travelling twin will see the Earth-bound twin appear to travel out and back at a high velocity! However, there is a major difference. The travelling twin experiences some short acceleration periods at the start, at the midway turn-around point, and at the end of his or her journey. The predicted time dilation is considered to arise as a direct result of the accelerations that have to be given to the clock, or twin, to achieve the two constant-velocity periods. But note that, because the acceleration periods are very short, the time dilation which occurs *during* the acceleration periods is neglected. All of the calculated time dilation occurs during the long constant-velocity periods when the clock is in an inertial

frame. It is because of this particular analysis that many physicists still feel slightly uneasy about the clock paradox or twin paradox. It is therefore worthwhile to develop the problem of a travelling atomic clock a little further.

We have the prediction, agreed by all relativists, that if a standard atomic clock is accelerated, and it then travels at a constant very high velocity relative to the Earth, it will run more slowly than an identical clock which has been left behind on the Earth. But the Earth is moving relative to the Sun, the Sun is moving relative to our galaxy, and our galaxy is moving relative to the background frame given by the distant galaxies. All of these bodies have had a previous acceleration history.

A clear question arises. When considering only relative motion effects, where could one locate a standard atomic clock so that we may be certain that it will run at the fastest rate compared with all other identical clocks? Based on our present knowledge, this location would appear to be the inertial frame that is stationary relative to the background given by distant matter in the Universe. The slowing of all moving clock readings should then be referred to this material preferred inertial frame. Hence, there appears to be a unique preferred inertial frame of reference in our Universe that establishes a primary time standard.

Few relativists like to admit to such an unambiguous statement. *But the majority of relativists do agree that it is unlikely that the clock paradox or twin paradox would exist if the framework given by the distant galaxies disappeared.* The reason that most relativists are prepared to go this far is because they accept Mach's Principle. Mach's Principle states that the whole set of inertial frames is determined by distant matter. If most relativists are prepared to go this far it seems strange that they will not accept a primary time scale, based on a specific preferred inertial frame which is directly related to distant matter in our Universe.

If most relativists agree that the clock paradox only arises because of the presence of distant matter in our Universe, is there any other way that they can be persuaded of the significance of this distant matter and can we also decide on which matter is relevant?

It may be possible to convince relativists by going back to the fundamentals of Maxwell's equation and special relativity. Special relativity is inherently subordinate to Maxwell's equations, and any new information coming out of Maxwell's equations may be relevant.

Time dilation comes out of Maxwell's equations. Maxwell's equations

on their own, without using any assumptions from special relativity, predict that a time transformation occurs for electromagnetic waves travelling between moving inertial frames of reference. This deduction comes from Lorentz transformations, which are a direct consequence of Maxwell's equations, and are an essential requirement if Maxwell's equations are to apply in all inertial frames.

At first sight, one could logically restrict the application of the time transformation prediction of Lorentz transformations to electromagnetic waves travelling between inertial frames, and exclude any application to atomic clocks or twins. The form of the analysis used to predict electromagnetic waves from Maxwell's equations might suggest that this is all that these equations, and Lorentz transformations, should predict. When only electromagnetic waves are being considered the acceleration histories of the frames are irrelevant because complete symmetry exists when there is an interchange of electromagnetic wave signals between inertial frames. It is the frequencies of the waves that change and the Lorentz time transformation predicts the relevant frequency shifts for both transverse and longitudinal motion of a moving source. This approach, of limiting Maxwell's equations and special relativity to just electromagnetic waves, is also logical in terms of the assumptions made in special relativity. We have shown earlier that the time-assigning-function assumptions adopted by Einstein, when establishing special relativity, are based solely on a consideration of electromagnetic wave signals passing between inertial frames. These assumptions are not valid for pulse signalling, where precursor transients are relevant, because quantum theory is then necessary. On this restricted basis special relativity would be a true relativity theory. It may be noted that this restricted approach is able to explain the increased lifetime of a mu-meson travelling at a high velocity because a mu-meson is not a complete clock. The "ticking" mechanism of a mu-meson "clock" is located in the moving frame whereas the read-out mechanism is located in the observer's stationary frame (L.M.Stephenson, J.Phys.A., **3**, 368, 1970).

But if this restricted approach to special relativity were to be adopted then it would be necessary to introduce a completely new theory to explain the time dilation applicable to complete clocks and identical twins.

However, there is a second choice. Is it possible that Maxwell's equations contain within them a built-in knowledge of the existence of distant matter in the Universe? At first sight one might imagine that a few equations, developed to deal with local electric and magnetic fields, could

not contain within them an inherent relationship between the local fields and distant matter in the Universe. But they do.

Free magnetic poles are not observed in nature and are not permitted in Maxwell's equations. Nevertheless, free electric charges are observed and their effects are included in Maxwell's equations. It is important to appreciate that any analysis involving free electric charges requires careful interpretation.

Mathematically, one may apply Maxwell's equations to a single point charge situated in an otherwise empty Universe. The predicted electric field of this point charge will extend to infinity. *But the existence of such a single point charge, in an otherwise empty Universe, is not physically possible.* Hence the mathematics needs to be interpreted carefully.

Mach's Principle also raises objections to the concept of having either a single point charge, or a single transmitting antenna, situated in an otherwise empty Universe. Neither concept has any valid physical meaning in an empty Universe. A sink for the electric field, or for the energy, must also be present.

For the case of the "single" point charge there must be distant *matter* in the Universe in order to create the local single point charge. A second distributed charge, equal in magnitude and of opposite sign, will inherently be created on this distant background matter. If the single point charge is accelerated to a high velocity v , relative to the background matter frame, and is then oscillated, the resulting electromagnetic waves will be time transformed when compared with the radiation predicted for an identical experiment where $v = 0$.

Maxwell's equations contain more information than is immediately obvious provided they are always applied after the necessary *material* boundary conditions have been established. The prior application of these material boundary conditions is particularly important whenever one is considering either the reception of electromagnetic energy, or the creation of a static electric field in space.

Mach's Principle suggests that the concept of radiating energy from a single transmitting antenna into an otherwise empty Universe, where no absorber is present, is a dubious concept. However, it is even clearer that the concept of having a single point charge, producing an electric field in an otherwise empty Universe, is a totally flawed concept.

The need for a material background frame of reference is a direct consequence of examining the nature of the required boundary conditions

for Maxwell's equations in more depth. The required material background frame may seem to be more tenuous when related to an overall neutral body, such as an atomic clock, but a material background frame of reference must be there in principle. One might then expect an atomic clock to go slow when travelling relative to a substantial material background frame. Time dilation is then readily explained in terms of motion relative to this specific background frame. There is no "clock paradox" or "twin paradox".

Maxwell's equations contain all of the information that is necessary to solve all steady-state, and near-steady-state, relativistic problems involving inertial frames and material boundaries. No further assumptions from special relativity, concerning the velocity of light in empty space and time-assigning functions, are necessary. In addition, Maxwell's equations predict the need for a material background frame when applied to the real world.

The mathematics of special relativity may still be used for all problems involving inertial frames (the mechanical steady state) provided that any electromagnetic wave propagation problem is limited to the electromagnetic steady state or the near steady state. However, this mathematics arises *directly* from Maxwell's equations. One only needs to assume the principle of relativity, and Maxwell's equations, to fully establish the effects of relative motion at velocities approaching c .

It is worth noting that all time transformations in general relativity must also have their origins in Maxwell's equations. How does this deduction affect general relativity? As an example, consider the problem of predicting the orbits of the planets in the solar system. General relativity is unable to give a complete solution for any rotational problem, but can only give relativistic corrections to the Newtonian solution. The Newtonian solution for planetary orbits appears to be inadequate because, using Galilean geometry, it predicts that gravitational forces must act instantaneously over very large distances (the action-at-a-distance requirement) if angular momentum is to be conserved. But Maxwell's equations suggest that it should be impossible to transmit any information faster than the velocity of light c . As was mentioned earlier, general relativity solves this problem by introducing an intermediate step in the analysis. Initially, the presence of all of the matter in the solar system warps local space-time, and this distortion of space-time travels over the region at a velocity equal to c . In the *steady state* the warped space-time then appears to act on the individual planets instantaneously and predicts the relativistic corrections to their Newtonian

orbits. General relativity is combining two distinct effects. Relativistic space distortion is primarily a local effect arising from the presence of local matter in the solar system. It is related to the coordinate system of the solar system. Relativistic space distortion arises from the assumption of the equivalence principle in general relativity. Relativistic time distortion arises from the motions of the planets relative to the coordinate system associated with the background of distant matter in the Universe, and arises from the assumptions of Maxwell's equations.

Gravitational effects can always be transformed away (or materialized from nowhere) in general relativity by a suitable change of the coordinate system and geometry. Although Machian effects may appear in general relativity, they can always be transformed away. This is why general relativity is unable to incorporate the stronger form of Mach's Principle, as is illustrated by its inability to explain the classic two fluid stars problem. However, Maxwell's equations require the presence of substantial distant matter to provide time dilation, and it is Maxwell's equations that enable general relativity to be grounded in this background material in the Universe.

Chapter 8

A Summary of Special Relativity

Electromagnetic waves and photons do not exist, as such, in free space. They are simply two alternative mathematical concepts that enable one to solve for the different effects which occur at a specific material boundary when one attempts to observe the electromagnetic energy that had previously existed in a free-space region which is now occupied by a detector. In particular, a photon can only be observed as it is absorbed at a material boundary.

It has been shown that quantum theory provides the transient solution, and the advanced potential solution of electromagnetic theory (or special relativity) provides the near-steady-state and steady-state solutions, for any problem involving the detection of electromagnetic energy.

The process of signalling using electromagnetic energy has inherent complications which are not always widely recognized. Long before one considers deliberately introducing pulsed electromagnetic signals there is a much simpler problem.

A truly monochromatic electromagnetic wave is an impossibility because, by definition, such a wave is not permitted to have a beginning or an end. A monochromatic electromagnetic wave cannot carry any information and hence cannot have an observable velocity. This is why the appearance of the velocity c in the uniform plane wave solution of Maxwell's equations for free space is not directly relevant when considering the observed velocity of light using a material detector. It is unfortunate that it was the appearance of the velocity c in the uniform plane wave solution, together with the observed constancy of the velocity of light in steady-state diffraction experiments, which initiated all of the developments that culminated in Lorentz transformations and special relativity.

It is interesting to look at the history of special relativity. Back in 1905 Einstein was presented with a problem. He required a fully encompassing

theory that explained why the observed velocity of light appeared to be independent of either an aether or the velocity of the source. Lorentz transformations, on their own, offered an arbitrary way of maintaining Maxwell's equations in all inertial frames, but they offered no intuitive explanation of why this particular set of transformations is correct.

Einstein first defined the principle of relativity S that all physical laws should be the same in all inertial frames of reference. Einstein then proceeded to make further assumptions. His second postulate, which is often incorrectly seized upon as being the basis of special relativity, was:

“that light is always propagated in empty space with a definite velocity c which is independent of the state of motion of the emitting body”.

But this postulate could equally well have been worded:

that light is always observed to be propagated in empty space, in all electromagnetic wave experiments, with a definite velocity c which is independent of the state of motion of the emitting body.

In other words, the second postulate of special relativity is a statement of the observational facts that were available. It has been shown that this statement is a direct consequence of the advanced potential solution of Maxwell's equations.

The firm basis of special relativity followed when Einstein discussed how to define the simultaneity of the observation of two events. In 1905 Einstein would have had no knowledge of precursor transients, advanced potentials, or quantum theory. He was therefore forced to make some time-assigning assumptions relating to electromagnetic signalling between inertial frames using light signals. In particular, he assumed that the out-and-return transit time of a light signal going from frame A to frame B and back is equal to $2AB/c$. This assumption is not valid for the case of a signal involving precursor transients and individual photons. Einstein was fully aware that his time-assigning-function assumptions might limit special relativity because he stated: “We assume this definition of synchronism is free from contradictions...”.

The mathematics of special relativity is valid provided the analysis is restricted to steady-state or near-steady-state electromagnetic wave signalling between inertial frames. But the ultraviolet catastrophe limitation

forbids the application of Maxwell's equations, or special relativity, to transient signals that contain significant components of electromagnetic energy at ultraviolet frequencies or above. Quantum theory is then essential.

The key fact is that no observation of radiated electromagnetic energy, present in a free-space region, is possible until a material detector (or energy sink) is inserted into the free-space region. The energy that is going to be observed is then totally stopped in its tracks. Maxwell's equations predict that such a change in the material boundary conditions has major consequences, and will significantly perturb the energy distribution that existed in free space prior to its observation. These perturbations in the fields will extend *past the source* and to infinity in the true steady state. It is the act of observation which controls many of the detailed aspects associated with the energy that is observed. In particular, it is the presence of a material detector which requires that the observed steady-state velocity of any electromagnetic wave must always be equal to c relative to the detector, and independent of the velocity of the source.

It is also essential that the Lorentz time transformation is correctly interpreted in view of a further restriction imposed on Maxwell's equations. A material background in the Universe must be present if Maxwell's equations are to be valid. The concept of having an electric field produced by a single point charge, situated in an otherwise empty Universe, is meaningless. Other matter must be present in the Universe for the divergence of an electrostatic field to have any meaning based on causality.

A clock will go slow, and a twin will age more slowly, only when travelling at high velocities relative to a substantial material background.

Appendix

The origin of gravity, and of the gravitational constant G , was discussed in Chapter 4. It was concluded that the most likely origin of gravity is for it to be related to the rotation of the Universe as a whole against a background of other universes. In *Paper 1*, that follows below, one possible origin of G was based on the rotation of a galaxy, which has now been discounted because of observational evidence. Nevertheless, the substance of the analysis remains the same for all rotational motions. However, the new interpretation does not reject the need for dark matter and dark energy to explain observational evidence.

Paper 1

Physics Essays, volume 16, number 2, 2003

A Dynamical Origin for the Gravitational Constant that Explains Gravitational and Inertial Mass Equality and Rejects Dark Matter and Dark Energy

Lawrence M. Stephenson

Abstract

There is no accepted theory that explains why the gravitational mass and the inertial mass of a given body are equal. Any new theory that predicted a dynamical origin for the gravitational constant should also predict that the

gravitational mass of a body is equal to its inertial mass. Two dynamical theories are outlined. One of these theories does not require the assumptions of dark matter and dark energy to explain cosmological observations.

Key words: gravitational mass, gravitational constant, dark matter, dark energy.

1. INTRODUCTION

There is a fundamental difference, in principle, between the gravitational mass and the inertial mass of a given body. However, the most accurate experiments¹ ever performed in physics have demonstrated that, for all of the materials tested, the gravitational mass of a body is found to be identical to its inertial mass to an accuracy of one part in 10^{11} .

Because neither Newtonian theory nor general relativity (see below) is able to explain the physical reason for the observed equality of gravitational and inertial mass, it was suggested in 1986² that a fifth force could exist (the four basic forces being the electromagnetic force, the strong and weak nuclear forces, and the gravitational force). To check this idea a comparison was made of the gravitational and inertial masses of materials having very different nuclear constitutions. If a difference had been found, then it would indicate both the existence of a fifth force and an observable difference between the inertial mass and the gravitational mass for a few specific materials. But no difference was found. As a result of these latest experiments, it now seems certain that there is an undiscovered theoretical link between inertial and gravitational mass.

The lack of a theory to link gravitational and inertial mass is not the only problem. A secondary mystery concerning mass surrounds the value of the gravitational constant G , which controls the magnitude of the gravitational force of attraction between any two given bodies. All of the other constants of physics, such as the charge and mass of the electron, the radius of an atom, the velocity of light, and Planck's constant are interrelated. But G stands out as a totally isolated constant. Again, there is no accepted theory that explains either why G should stand alone or why it has the particular value it has.

The basis of general relativity lies in the *experimental* observation that a gravitational field imparts the same acceleration to all bodies. Hence a gravitational acceleration of a test particle cannot be distinguished from an

inertial acceleration of the particle. It is then sometimes incorrectly argued that the introduction of “mass” is unnecessary, and that general relativity itself proves that the inertial mass of a body is equal to its gravitational mass. But what would be the case if a fifth force had been found to exist and, for some materials, the inertial mass of a body had been observed to be different from its gravitational mass? It is clear that general relativity cannot offer a theoretical explanation of why inertial mass and gravitational masses are identical. Also, general relativity adopts the Newtonian assumption that the value of G must be a universal constant. But there is nothing in general relativity that either establishes a value for G or proves that G is a universal constant³.

Mass is the fundamental quantity upon which the whole of physics is built, and yet we have no understanding of its true nature. What can be done to achieve a theoretical understanding of why the gravitational mass of a body is found to be identical to its inertial mass? When any two items are observed to be identical, we deduce that they must be linked to some form of common origin. One might then expect inertial mass and gravitational mass to have a common origin. Two separate theories, which are based on assuming such a common origin, will now be considered. Both of these theories also yield a particular value for G .

2. MACH’S IDEAS ON THE ORIGIN OF INERTIAL FORCES

Mach was troubled by the deeper problems associated with rotational motion. One of the many early arguments that showed the difficulties with Newtonian theory went as follows. Consider two fluid stars that are well separated in the sky. One of these stars is stationary and has a spherical shape. The second star is spinning about an axis that passes through the center of the first star, and has an oblate shape due to its rotation. Now let all of the rest of the matter in the Universe be gradually removed. At the limit one is apparently left with one spherical star and one oblate star. One can deduce that the latter are spinning because of its oblate shape. But with all other matter removed there is no external frame of reference. To have one star spinning and oblate, and the other star stationary and spherical, is not acceptable. As they can only relate to each other, why is there a difference between them? One can go further and remove the stationary star. One now has a single star spinning in an otherwise empty Universe, but what is it spinning with respect to?

On a Newtonian basis the above argument is plausible, but Mach

considered that removing parts of the Universe might have dramatic effects. Mach's answer to this problem was to postulate that our basic concepts of the mass of a body, the inertial acceleration of the body, and the gravitational attraction within the body must all depend on causes lying well outside the immediate system being considered. Mach suggested that the totality of matter making up the whole of the Universe, in some way yet to be defined by experiment, gives rise to the concepts of mass and acceleration forces. Mach is suggesting that, if distant parts of the Universe could be removed, then some of the basic laws of physics might change. For example, the value of the gravitational constant G might change. Such a hypothesis is close to the gravitational theory of Hoyle and Narlikar⁴, who suggested that G might be inversely proportional to the mean density of the Universe. Hence, if half of the Universe disappeared, then the value of G would double. However, it is unlikely that any experimental evidence will be forthcoming if the only suggested test is to remove a substantial part of the Universe. Hoyle and Narlikar's theory is limited to defining just the value of G , and it does not produce any link between inertial mass and gravitational mass. It is necessary to develop Mach's ideas still further to find the origin of the equality of gravitational and inertial mass.

3. A DYNAMICAL ORIGIN FOR GRAVITATIONAL MASS

Hoyle and Narlikar's theory involves relating G to the total *static* distribution of matter in the Universe. But possible alternative theories exist that relate G to the *dynamic* distribution of matter in the Universe. As it was the rotational motion of matter that caused many of the difficulties that Mach highlighted, there is a good chance that a dynamical theory for G might be more appropriate, especially one that involves rotational motion.

To achieve a dynamical theory for G one needs to postulate that it is the motion of distant matter, on a very large scale, relative to the cosmic background frame, that *produces* all of the terrestrial gravitational forces that we observe. If there is no motion of distant matter, then the value of G will go to zero and all gravitational effects will disappear. Hence the gravitational mass of any given body will be directly related to the motion of distant matter. As well as predicting a value for G , this type of dynamical theory indicates that the gravitational mass of a body should inherently be equal to its inertial mass. If the origin of the local gravitational mass of a body arises from the (inertial) motion of distant matter, then one would expect the gravitational mass of the body to be equal to its inertial mass

because the apparently different types of mass have a common origin.

Any new dynamical theory for G must be compatible with our knowledge that locally moving matter does not produce a significant increase in locally observed gravitational forces. Fortunately, the detailed work of Hoyle and Narlikar supports a dynamical theory based on the motion of distant matter, because they showed that local matter has a negligible effect on G compared with distant matter.

The obvious way of transforming this idea of a dynamical origin for G into an analytical statement is to look initially at the dimensions of the gravitational constant. The dimensions of G may be expressed as:

$$G \sim 1/\rho T^2, \quad (1)$$

where ρ represents the dimensions of density and T the dimensions of time. It can immediately be seen that Hoyle and Narlikar's gravitational theory, in which G is inversely proportional to the mean density of the Universe, is based on carrying over just a part of this dimensional relationship into a full theory. In a similar way, Dirac made the well-known suggestion that G may be slowly decreasing with time, with a time constant of the same order of magnitude as the age of the Universe. Hence Dirac's suggestion is based on carrying over the other part of this dimensional relationship into a formal relationship.

If we are searching for a dynamical theory for G , based on the inertial motion of distant matter, then the above dimensional relationship may be changed slightly to become:

$$G \propto v^2/r^2\rho \quad (2)$$

where v is now the velocity of distant matter at a distance r . Note that, as we require a theory relating G to both the density and the inertial motion of distant matter, we have changed from making a simple dimensional statement about G to making G proportional to all of the quantities involved.

Alternatively, for a possible rotational motion of distant matter, (1) may be restated as:

$$G \propto \omega^2/\rho \quad (3)$$

where ω is the angular velocity of the matter relative to the cosmic background frame.

Let us consider (2) first. Suppose that G is related to the recessional velocity of distant matter in the Universe associated with the expansion of the Universe. For the usually accepted linear expansion of the Universe, proposed by Hubble, v/r is a constant and hence G will be inversely proportional to the density of matter. One interpretation of (2) then reduces to a new dynamical theory of gravitation that is similar in its outcome to the static theory of Hoyle and Narlikar. However, this dynamical theory explains both the value of G and the equality of inertial and gravitational mass. If this particular dynamical theory happens to be correct, then, again, as with Hoyle and Narlikar's theory, it is unlikely that any direct experimental test could be performed to establish the theory.

But (3) is much more interesting. This equation allows G to be related to the *rotational* inertial motion of distant matter. If we ignore the concept of the whole Universe rotating, then the next level down where rotation occurs is with galaxies or groups of galaxies. Hence one may propose a second dynamical theory of gravitation, based on the rotational motion of a galaxy, by making use of the known Newtonian stability condition for an idealized, gravitationally stable galaxy, which is:

$$G = \frac{\omega^2}{2} \rho k \quad (4)$$

where ω is the angular velocity of the galaxy, ρ is the mean density, and k is a boundary shape constant (≤ 1). In this second dynamical theory for G the value of G is no longer assumed to be a universal constant. It is proposed that the value of G is being specifically *created* by the rotational inertial motion of the matter within the galaxy. Even within our own galaxy, the Milky Way, the value of G need not be constant – the value will depend on the values of ω and ρ for the particular region. If G is created in this way, the predicted value of G will only apply within the boundary surface of the galaxy, and the value of G will be zero in any region of intergalactic space where no rotation of matter is taking place. This theory was originally proposed in 1965, and was elaborated on in 1969⁵ and 1976⁶. But at that time there were no observations that were considered to be unexplained by current theories. The situation has now changed as a result of two cosmological anomalies. Cosmologists have found that the gravitational attraction of visible matter is insufficient to account for the orbital

velocities of stars in individual galaxies, and of galaxies within clusters of galaxies. To explain this discrepancy vast amounts of dark matter are assumed to be present. More recently, Krauss⁷ has provided evidence that the gravity of visible matter is not consistent with an observed accelerating expansion of the Universe; he proposes that an assumption of large amounts of dark energy, associated with a cosmological constant, is required to account for this second discrepancy. No direct observational evidence has been found to account for either dark matter or dark energy, and alternative theories cannot be discounted. In particular, many cosmologists are reluctant to accept a cosmological constant in general relativity. The new interpretation of (4) eliminates the need for both dark matter and dark energy. The need for dark matter is removed because the new interpretation of (4) predicts that a galaxy is self-stabilizing in all regions. For a given value of angular velocity, (4) predicts that G will vary inversely with the value of the density of the region of the galaxy being considered. Hence in the low-density, outer regions of the galaxy there will be an increased gravitational attraction to correspond with the high angular velocities of stars in these regions. Artificial dark matter is not needed to explain galactic stability.

The need for dark energy is removed because this dynamical theory for G predicts that G is zero outside the boundary of any individual galaxy or any rotating galaxy cluster. Without gravitational attractions between individual galaxies, or galaxy clusters, one would expect a faster expansion of the Universe than current theory predicts, where G is assumed to be a universal constant and gravitational attraction between galaxies will slow expansion.

This theory may also give rise to terrestrially observable effects. It is proposed that it is only for gravitationally-stable, rotating, *fluid* bodies, like the galaxy, that G will be created by the rotational motion of matter as given by (4). In a solid rotating body, such as a flywheel, additional centrifugal forces can be developed because G is not limited by the fluid boundary-shape condition. Hence (4) will not apply to a flywheel. However, although the Earth is semi-rigid, its boundary shape is very near to that which would be adopted by a truly fluid Earth. If (4) applies to the Earth, then for a truly fluid Earth an additional increase in G of approximately 0.4% will occur when passing across the boundary surface into the interior. To observe the possible predicted increase in G below the Earth's surface it is necessary to perform two full G experiments. The first would be at a

height of 1-2km above the Earth's mean surface level (at a location not in an anomalous g gravity high). The second experiment would be performed down a 1-2km deep mineshaft (at a location not in an anomalous g gravity low). Such experiments would establish whether there is an increase in G when crossing the mean boundary surface into the interior.

At the other extreme to predicting galactic stability, (4) predicts that the electron, or any other spinning atomic particle, is stabilized by gravitational forces⁵ and not by arbitrary short-range forces. If one substitutes a value for ρ deduced from the classical radius e^2/mc^2 and a value for ω deduced from the angular momentum $h/4$, then the internal value of G for an electron comes out to 10^{45} times the terrestrial value for G . Since the ratio of the electromagnetic coupling constant to the usual gravitational coupling constant is about 10^{43} , the internal values of the gravitational forces within an electron are then predicted to be more than sufficient to stabilize the electron against the electrostatic repulsion forces. Einstein was convinced that atomic particles were stabilized by gravitational forces and he attempted to account for this stability by renormalizing the gravitational field at the boundary of the particle. As the nature of this renormalizing process was ill defined, arbitrary short-range forces are now assumed in order to explain electron stability. It seems likely that Einstein's instincts concerning atomic particle stability were correct.

4. CONCLUSION

The most accurate experiments ever performed in physics indicate that the gravitational mass of a body is identical to its inertial mass. A theoretical link would be anticipated, but neither Newtonian theory nor general relativity is able to provide this link. If the gravitational mass of a body is to be directly related to its inertial mass, then the gravitational constant, which controls the gravitational mass of the body, must have a dynamical origin. One particular dynamical origin for the gravitational constant, based on the rotation of galaxies, explains two cosmological anomalies, in addition to linking gravitational and inertial mass.

Acknowledgement

I am very grateful to Professor C.W.Kilmister for some valuable comments, and to both referees for suggesting some clarifications.

Received 6 June 2003

References

1. P.G.Roll, R.Krotov and R.H.Dicke, Ann. Phys. (NY) **26**, 442 (1964).
2. E.Fischbach, *et al.*, Phys. Rev. Lett. **56**, 3 (1986).
3. L.M.Stephenson, Phys. Essays **13**, 138 (2000).
4. F.Hoyle and G.Narlikar, Proc. Roy. Soc. London , Series A **282**, 191 (1964).
5. L.M.Stephenson, J. Phys. A., Gen. Phys. **2**, 475 (1969).
6. *Idem*, Found. Phys. **6**, 143 (1976).
7. L.M.Krauss, Astrophys. J. **501**, 461 (1998).

Lawrence M. Stephenson

Department of Electronic and Electrical Engineering
University College London, WC1E 7JE, England

Paper 2

Physics Essays, volume 11, number 4, 1998

The Significance of Precursor Electromagnetic Waves in Special and General Relativity

Lawrence M. Stephenson

Abstract

Although the mathematics of special relativity is not open to question, the interpretation of the mathematics, especially regarding its relationship with quantum theory, is far from complete. A study of precursor electromagnetic waves, which are closely associated with quantum effects, offers a new insight into special and general relativity.

Key Words: special relativity, general relativity.

1. INTRODUCTION

Special relativity is based on two primary assumptions. The first assumption is the *principle of relativity*, which states that the laws of physics should be the same in all inertial frames of reference. The second assumption is a postulate concerning the velocity of light in empty space, namely: light is always propagated in empty space at a definite velocity c relative to any mathematical, matter-free, inertial frame of reference. But it is impossible to test this postulate experimentally because any observation of the velocity of light requires the presence of a material detector. Einstein was clearly aware of the imprecision involved in defining the velocity of light relative to empty space. He introduced additional assumptions by establishing a viable system of signalling and time assigning, so as to achieve a method of synchronizing a number of clocks located in inertial frames which are in relative motion. Einstein recognized

that these additional time-assigning assumptions might create limitations for the theory because he concluded: "We assume that this definition of synchronism is free from contradictions...".

Einstein's additional time-assigning assumptions are essential to qualify the primary assumption concerning the velocity of light in empty space. It is incorrect to assume that special relativity is inherently a mathematical theory of matter-free, four-dimensional, space-time, and then to introduce matter back into the theory at a later stage. Such an approach can offer no explanation of the one-dimensional nature of the Lorentz transformation, and it requires a further assumption of a mathematical rotation of axes, to form the Lorentz group, in order to maintain invariance of transformation in three spatial dimensions. If special relativity is a theory of physics, then it is reasonable to require that only material inertial frames should be considered.

It is possible to avoid all assumptions in special relativity, other than the principle of relativity, if one returns to Einstein's approach and reassesses Maxwell's equations in terms of the effects that precursor transients have on any realistic experimental measurement of the velocity of light.

2. PRECURSOR TRANSIENTS

Precursor transients arise whenever a step-function of electromagnetic energy, initially traveling in free space, interacts with any material that is able to store energy. The best known example⁽¹⁾ of a precursor transient occurs when a step function of electromagnetic radiation enters a slab of dielectric medium. The front edge of the first precursor will continue to travel in the medium at c , the velocity of light in free space, because a dielectric cannot exhibit a value of relative permittivity greater than unity until dipoles have been formed in the medium. The formation of these dipoles takes a significant time and depends on energy being extracted from the precursor. Although classical electromagnetic theory may be used to find the velocity in the dielectric of both the front edge of the first precursor and the observed signal, there are difficulties in analyzing how the precursor transients develop before the arrival of the true signal. A definitive analysis is limited by the usual restrictions placed on Maxwell's equations at the atomic level and approximations are necessary concerning the formation of the dipoles.

A precursor transient will also arise whenever a step-function of radiated electromagnetic energy, traveling in free space, is observed using a

practical, material detector. In the limit, the analysis of any electromagnetic wave detector may be reduced to considering a Hertz dipole antenna. Consider the effect of such a simple, thin-wire, receiving antenna. A thin-wire antenna distorts the free-space field pattern that would have existed in its absence and, in the steady state, the antenna will be surrounded by new "near" fields which contain stored energy. On the arrival of a step-function of energy at the antenna the precursor transient introduces a time-delay before the signal can be detected; during this time-delay energy from the precursor transient is building up the near-field stored energy around the antenna. Looked at another way, a thin-wire antenna presents a negligible physical area. Even on a classical field theory approach, which assumes an instantaneous response of the antenna to the electric field stimulus, it is still predicted that the antenna will have a negligible *effective* area until the scattered field has had time to expand and the near fields have begun to be established. Until the effective area of the antenna is finite no signal can be detected.

The time-delay introduced by the precursor transient, before a signal can be detected, may also be readily examined on a quantum basis. A single photon, appropriate to an infinite-frequency Fourier component, associated with a perfect step-function of energy will have infinite energy. Even for a near-perfect step-function of energy a single photon will have sufficient energy to establish the stored energy in the near fields produced by the steady-state current induced in the antenna. But the precise time of arrival of the first photon is an indeterminate quantity. Hence, both quantum theory and classical electromagnetic theory predict a finite time-delay before a step-function signal can be detected or reradiated by any material object situated in free space.

The prediction of such a finite time-delay, before a step-function of energy can be detected or reradiated, has an impact on special and general relativity. In order establish special relativity, and a method of signaling between two observers *A* and *B*, Einstein made two assumptions. The time-assigning assumption *defines* the one-way transit time of a signal traveling between *A* and *B* as being equal to one half of the out-and-return transit time. In addition, special relativity assumes that the out-and-return transit time is equal to $2AB/c$. But for a practical, material reflector at *B* the observed out-and-return transit time of a true square-edge signal pulse will be greater than $2AB/c$ because of the extra precursor transient time-delay introduced during the process of detection and reradiation. It is not until the

new near fields have begun to be established at the material surface of the reflector at B that a return signal can be generated. Hence, if special relativity is considered to be based on the signaling system initially envisaged by Einstein to establish clock times in inertial frames, then special relativity is invalid for any case where a step function of radiant energy is used to signal between material inertial frames.

3. DISCUSSION

The conclusion just reached may be viewed in two distinct ways. At a trivial level it indicates that, for non-ideal circumstances, a *measurement* problem may exist when applying special relativity. It is necessary to analyze any measurements appropriately. If an inertial frame containing a material detector exists in a free-space region, the steady-state, or near-steady-state, solution for a signal received by the detector will be given by Maxwell's equations and special relativity. But the extreme precursor transient solution associated with a near-perfect step-function signal, and the arrival of the first few photons, can only be achieved using quantum theory.

Hence, when analyzing an inertial frame containing a material detector of energy, it is clear that, from a measurement point of view, special relativity must be restricted to steady-state, or near-steady-state, solutions.

But such a trivial explanation of the significance of precursor transients is at variance with the essential part played by the presence of a material sink in any observation. Special relativity should be based on studying observable quantities. However, no observation is possible in a non-material inertial frame. To make an observation a material detector has to be introduced into a free-space region. The resulting radically changed steady-state field pattern then has a further specific consequence. The steady-state approach velocity of an electromagnetic wave to an infinitesimal (Hertz) dipole receiving antenna must be equal to c , and independent of the source velocity, as a result of Maxwell's equations. The solution is derived from the retarded potential⁽²⁾ for a transmitting Hertz dipole followed by the application of reciprocity for a linear, isotropic, homogeneous, free-space region. This result may also be deduced from the advanced potential solution⁽³⁾, provided one avoids causal restrictions by limiting the analysis to signal energy that is beamed directly from the transmitter to the detector. Hence, Maxwell's equations dictate that the observed steady-state velocity of light must be equal to c , relative to an

observer's material detector, *because the act of observation makes it so*. In quantum theory the effect of the observer on the outcome of any observation is inherently significant. But in special relativity the effect of the presence of a material observer on the outcome of an observation has not been previously considered. It is now clear that it is the material of the observer's detector that ensures that the approach velocity of any electromagnetic wave is equal to c , relative to the observer, in the steady state.

It is often implied that the uniform plane wave solution of the wave equations demonstrates that the velocity of light must be equal to c "in empty space". Such a deduction is misleading as no observation can be made in empty space. The uniform plane wave solution is a useful idealization, but it is only valid for the unrealizable case of an infinite plane source and an infinite plane detector. The solution for an infinite plane detector fails to illustrate how a practical detector interacts with the wave in any finite region of space. However, the solution for a realizable detector, outlined above, shows that the steady-state velocity of any observed electromagnetic wave traveling in free space must be equal to c relative to an infinitesimal dipole.

As this last point is so important it is worth elaborating further. If a finite-sized transmitting antenna is used, instead of a Hertz dipole, then the steady-state, free-space signal velocity of any radiated electromagnetic wave will be marginally less than c close to the antenna and will only asymptote to c at infinity. At first sight this deduction might be thought to imply a source theory of propagation, but exactly the same result applies to the observer's viewpoint of the steady-state approach velocity of an electromagnetic wave to a finite-sized receiving antenna. To deduce what will be observed it is the solution relating to the sink that is required, not the solution relating to either empty space or the source. Special relativity is able to ignore the importance of considering the observer's material sink of the energy, and the effect of precursor transients, by making two assumptions. But the consequence of these assumptions is to restrict special relativity to near-steady-state solutions. Furthermore, there is then no need for Maxwell's equations to be Lorentz invariant under all conditions. Lorentz transformations are simply the transformations required to allow for the steady-state boundary conditions imposed by a material detector when in uniform translatory motion.

Care is needed if special relativity is applied to transient problems. In

addition to steady-state analysis, special relativity may be applied to most established transient signal analysis because, in nearly all cases, the significant part of the signal is observed after the arrival of the first few photons and these photons will have established the near fields of the detector. However, special relativity and Maxwell's equations will not apply to the precursor transient. As the main relativistic aspects of time appearing in general relativity are wholly dependent on the relativistic aspects of time developed in special relativity, care is also needed when applying general relativity to some transient problems. General relativity may have limitations when applied to a theoretical examination of the origins of a "Big-Bang" type of Universe, as the initial stage of this problem may involve precursor transients. But general relativity will apply to the three main tests of the theory because the perihelion precession of the planet Mercury, the deflection of a light beam by the Sun, and the gravitational red-shift of the frequency of emitted radiation, are all steady-state phenomena.

The proposed alternative way of interpreting special relativity, based on considering only material detectors, goes further and explains why the relationship between Lorentz transformations and the rotation group is different from that holding for Galilean transformations. In the limit, Maxwell's equations require that any observation of an electromagnetic wave may be based on analyzing a Hertz dipole receiving antenna, followed by integration for any larger antenna. A one-dimensional current element is thus always associated with the practical radiation or detection of energy. Three-dimensional, spherically-symmetrical, phase-coherent radiation of electromagnetic energy is geometrically unobtainable because it is impossible to arrange a current distribution within a spherical shell of material that appears identical when viewed from all directions and which radiates.

The Lorentz transformation is a convenient way of abbreviating the steady-state solution for a moving, one-dimensional, current element. To maintain invariance of transformation in three spatial dimensions an additional mathematical rotation of the axes has to be introduced to form the Lorentz group, whose structure differs from the Galilean group. Physical derivations of the Lorentz transformation (as distinct from abstract approaches based on the assumption of the group structure) always deduce it in one dimension. The full group then arises from more or less clumsy arguments involving rotations. This is not a sign of inadequacy in the

mathematics; it is because there is no physical rotation of the axes of a material body. Terrell's analysis⁽⁴⁾ deserves much more credit. He showed that the mathematical rotation is precisely allowing for the fact that a moving, incandescent, *material* body intercepts some of the radiation produced at the leading edge. The perceived diameter of such a body, if spherical, is not Lorentz contracted in the direction of motion. The diameter in the direction of motion has to be mathematically rotated so that it coincides with the point sources of radiation whose outputs, after different transit times, define the leading and trailing edges of the circular image that appears on an observer's flat screen. Hence, all mathematical aspects of the Lorentz transformation summarize the full steady-state analysis arising from the application of Maxwell's equations to a particular class of realizable experiments employing material bodies in uniform translation.

By restricting special relativity to material frames it is possible to anchor both special and general relativity in the real Universe. Any material frame will have a finite mass. A uniformly moving frame of finite mass cannot have achieved its motion unless there was, at some time in the past, an exchange of momentum between the material of the frame and the rest of the Universe. It is then clear that a clock will only show a time difference, as a result of its motion, following an exchange of momentum between the clock's material inertial frame and the rest of the Universe, which accords with Mach's principle. If special and general relativity are anchored in the real Universe, a direct relationship between general relativity and other gravitational models is available^(5,6). One model indicates that the value of the gravitational constant may be related to the density and angular velocity of the galaxy.

4. CONCLUSION

Conventional special relativity assumes that light travels in empty space at a velocity c relative to any hypothetical, matter-free, inertial frame. Such a proposition can never be checked experimentally because any observation of the velocity of light requires the presence of a material detector. Einstein partially recognized this fact by additionally assuming the need for a viable system of signaling and time assigning. But he did not consider the limitations imposed on special relativity by the presence of material detectors. It is incorrect to assume initially that special relativity is inherently a mathematical theory of matter-free, 4-dimensional, space-time, and then to reintroduce matter back into the theory at a later stage.

Two alternative approaches to special relativity are possible, both based on considering the use of practical, material detectors to make observations of the velocity of light. First, one may return to developing special relativity so as to be consistent with Einstein's suggested methods of electromagnetic signaling and time assigning. If special relativity is based on physically realizable observations, made between material inertial frames, then any application of the theory must be restricted to steady-state or near-steady-state electromagnetic analysis. This restriction arises because the initial assumptions of special relativity forbid its application to the precursor transient which arises when a true step function of electromagnetic radiation is incident on any material medium. Special relativity then becomes the steady-state part of the solution to any given material boundary value problem. Quantum theory provides the extreme transient solution associated with the arrival of the precursor.

A second alternative approach to special relativity is also possible that is independent of any argument concerning precursor transients. One may examine the implications of Maxwell's equations for a material receiving antenna. For a material inertial frame the assumption of a constant value of c in empty space becomes redundant because Maxwell's equations dictate that the free-space, steady-state, approach velocity of an electromagnetic wave to an observer's receiving Hertz dipole must be equal to c and independent of the source velocity. In this context, Maxwell's equations only need to be Lorentz invariant to meet the requirement of the steady-state solution applicable to the boundary conditions imposed by an observer's material detector. Not only does this approach eliminate the need for any assumption concerning the velocity of light in empty space, but it also explains the one-dimensional characteristic of Lorentz transformations in terms of the one-dimensional nature of the current element always associated with the radiation or reception of electromagnetic waves.

Special relativity may be considered to be a theory of physics arising directly from the application of Maxwell's equations to the material boundary condition imposed by an observer's receiving antenna. There is no need for special relativity to be based on an experimentally unverifiable initial proposition concerning the constancy of the velocity of light in empty space. Special relativity then becomes directly related to quantum theory in the sense that it provides the final steady-state solution of any given material boundary problem for which quantum theory has provided the transient solution.

It is essential to appreciate that physics has moved on from basic Maxwell's equations to quantum electrodynamics, and from Einstein's original form of general relativity to the beginnings of quantum gravity. Nevertheless, the limitations and assumptions introduced at the formulation of any theory must always be considered when interpreting the final predictions of the theory.

Acknowledgment

I am very grateful to Professor C.W.Kilmister and Professor A.L.Cullen for many invaluable suggestions and comments. I am also grateful to a superb referee whose detailed comments have resulted in many improvements to the text.

Received 15 May 1998

References

1. J.A.Stratton, Electromagnetic Theory, (McGraw Hill, New York, (1941).
2. J.D.Jackson, Classical Electrodynamics, (John Wiley & Sons, 1962).
3. L.M.Stephenson, Found. Phys., **8**, 921 (1978).
4. J.Terrell, Phys. Rev., **116**, 1041 (1959).
5. L.M.Stephenson, Found. Phys., **6**, 143 (1976).
6. L.M.Stephenson, J.Phys.A., Gen. Phys, **2**, 475 (1969).

Lawrence M. Stephenson

Department of Electronic and Electrical Engineering
University College London, WC1E 7JE, England

Paper 3

Physics Essays, volume 13, number 1, 2000

The Relevance of Advanced Potential Solutions of Maxwell's Equations for Special and General Relativity

Lawrence M. Stephenson

Abstract

Although the mathematics of special relativity is not open to question, the interpretation of the mathematics, especially regarding its relationship with quantum theory, is far from complete. A study of advanced potentials offers a new insight into special and general relativity.

Key words: special relativity, general relativity

1. INTRODUCTION

Maxwell's equations predict symmetrical retarded and advanced potential solutions⁽¹⁾ which respectively relate to the radiation or absorption of energy by a Hertz dipole antenna. The retarded potential solution is associated with the radiation of energy from a single source into an infinite region of free space, and is generally accepted as being valid in classical electromagnetic theory. But the advanced potential solution produces difficulties if attempts are made to apply it to the generalized absorption of energy by a Hertz dipole antenna. These difficulties will be discussed later. As a result, all authorities rule out any valid application of individual advanced potential solutions^(2,3), but they are only able to deduce this result by appealing to causality arguments. Hoyle and Narlikar⁽²⁾ even go so far as to suggest that the existence of individual advanced potential solutions, which are never specifically excluded by the theory at a later stage in the analysis, indicate

that there is a fundamental weakness in Maxwell's equations. However, it will be shown that this conclusion is not wholly correct.

It is important to appreciate that the advanced potential solution is a physically significant solution which applies, in certain circumstances, to positive energy, travelling in positive real time (*i.e.* with the usual direction of the arrow of time) towards a Hertz dipole antenna. First, it may be noted that for any receiving antenna the fields associated with the energy approaching the antenna will be advanced in time relative to the corresponding current which is later induced in the antenna by these fields. All that is signified by the time-reversal of the advanced potential solution, relative to the retarded potential solution, is that one is now treating the fields in free space as the cause and the current in the antenna as the effect. Thus, the advanced potential solution would be expected to apply to a sink of electromagnetic energy⁽⁴⁾ in limited circumstances. It is true to state that the passing of an alternating current in an antenna cannot produce advanced waves converging on the antenna from infinity⁽³⁾. Nevertheless, suitably positioned advanced waves, converging on the antenna from a distance, can produce an induced current in the antenna.

One trivial example of a valid advanced potential solution exists, where all of the energy is contained in a closed system. Let a Hertz dipole antenna radiate a single pulse of electromagnetic energy into a free-space region which is bounded by a very large, spherical, conducting shell centred on the dipole. The need for both the retarded and advanced potential solutions is immediately clear. The relationship between the current in the dipole and the fields of the outgoing pulse is given by a retarded potential solution. For the returning reflected pulse the relationship between the fields of the pulse and the current which is later induced in the dipole is given by an advanced potential solution. The advanced potential solution is mathematically identical to the time-reversed retarded potential solution. But such a time reversal does not necessarily imply that real time is running backwards. For this example, the advanced potential solution applies to the dipole when acting as a sink of energy and to the reflected pulse as it travels with time, in the normally accepted direction of the arrow of time, towards the dipole. There is no question of negative energy being involved, or real-time reversal⁽³⁾, if the advanced potential solution is applied to the current induced in the antenna by the positive energy received by the dipole from the proposed reflected pulse. It is essential that Maxwell's equations should predict an advanced potential solution to cover this case. However, such a

valid advanced potential solution is based on an impractical source of energy and the advanced potential solution has no practical applications in three dimensions. All other three-dimensional solutions have to be eliminated by a direct appeal to the logic of cause and effect, as Maxwell's equations, on their own, are unable to eliminate these solutions. The argument used is simple. Suppose one has a transmitting antenna and a receiving antenna, separated by a finite distance in an infinite free-space region, both being situated inside a very large imaginary sphere. The energy everywhere on the spherical surface must be travelling outwards. Hence, any fields associated with an individual advanced potential solution are ruled out as they would result from energy travelling inwards. However, it is clear that such a deduction, based on causality arguments, only applies to problems where energy may escape to infinity.

It is usually considered that there are no problems associated with individual retarded potential solutions, because it is assumed that a physical meaning may be attached to the concept of an isolated source radiating energy into an infinite region of free space. But there are deeper, Machian-like, arguments that question the significance of having just a source of energy in an otherwise empty Universe, where no absorber is present. It seems likely that both retarded and advanced potential solutions require a further physical interpretation.

At this stage it is interesting to note that Wheeler and Feynman⁽⁵⁾ have demonstrated that valid solutions exist at the microscopic photon level for a propagator which produces a combination of half-retarded and half-advanced potentials, provided that all electromagnetic disturbances are ultimately absorbed. Their approach throws no light on the macroscopic problem of interpreting the validity of individual advanced potential solutions associated with the absorption of energy by a receiving antenna. However, there is an indication from their approach that valid advanced potential solutions will only exist in a system where no energy is permitted to escape to infinity.

The point being made in this paper is that, if one limits any analysis to one-dimensional solutions, there are valid individual advanced potential solutions when both a transmitting dipole and a receiving dipole are present in a free-space region. There are then cases where no energy can escape to infinity. One solution is for the case of an open resonator⁽⁶⁾. Another solution applies to a detector where signal energy is beamed directly from a source to the detector⁽⁴⁾, and no energy is lost in the transfer. It is

permissible to use the advanced potential solution if all of the signal energy is transferred from the source to a chosen sink of the energy. Furthermore, it will be shown that the advanced potential solution is critically significant when analysing the reception of an electromagnetic signal if there is relativistic motion of the source, as it then explains the limitations of the assumption made in special relativity concerning the constancy of the velocity of light relative to empty space.

To observe a signal a material detector has to be introduced into a free-space region. But the presence of the detector will change the steady-state field pattern out to infinity. Hence, it is essential to consider how the act of observation may change the observed velocity of the signal. Consider the proposition: "no physical meaning may be attached to the statement that an electromagnetic signal travels at a given velocity relative to its source until the signal is observed". This proposition is contrary to the usual method of classical physics. Nevertheless, this proposition bears a very close relationship to the assumption made in special relativity, namely: "that light is always propagated in empty space at a definite velocity c which is independent of the source velocity". Special relativity's indeterminate assumption, that light travels at c relative to empty space, is unnecessary. In place of this assumption one may introduce a valid advanced potential solution by directly applying Maxwell's equations to the boundary condition imposed by the material detector that must be present when a signal is observed.

Let a source of electromagnetic waves produce a signal consisting of an ideal cylindrical beam of energy in free space, all of which is intercepted by a Hertz dipole receiving antenna. Such an arrangement is possible if a parabolic reflector is placed behind the receiving antenna. Provided no energy is lost in the transfer of the signal, one may apply the advanced potential solution to the receiving antenna and deduce that the steady-state velocity of the approaching wave will be equal to c . The advanced potential solution requires that this velocity must be relative to the receiving antenna. The limitation of the advanced potential solution to the steady state arises from the inherent finite response time of a receiving antenna. If the signal to be observed consists of a pulse of energy then, on a classical field theory approach, one assumes an instantaneous response of a Hertz dipole to the electric field stimulus. But the dipole will have a negligible effective area until the scattered field has had time to expand⁽⁷⁾. New "near" fields will become established around the dipole which correspond to the near fields

of a transmitting dipole. Hence, Maxwell's equations are restricted to predicting that the observed steady-state, or near-steady-state, velocity of a signal will be c in free space. Quantum theory must be used to predict the velocity of the initial transient associated with the arrival of the first few photons. In passing, it may be noted that quantum theory indicates that an indeterminate time delay will occur before the first photon may be detected. To conclude, Maxwell's equations dictate that a Hertz dipole is initially transparent to a signal pulse. The act of observation with a material dipole generates a changed steady-state field pattern, and it is these new fields which determine the observed steady-state approach velocity of the signal. The advanced potential solution requires that this velocity must be equal to c relative to the detector.

If one extends this argument to the case where there is uniform translatory motion of the receiving antenna, then the advanced potential solution of Maxwell's equations still dictates that the creation of a sink of energy ensures that the steady-state approach velocity of the signal must be equal to c relative to the current induced in the material of the receiving antenna and independent of the source velocity. No assumption concerning the velocity of light in empty space is necessary when considering steady-state observations.

2. THE ONE-DIMENSIONAL LIMITATION PLACED ON LORENTZ TRANSFORMATIONS

A direct application of Maxwell's equations to the analysis of uniform translatory motion in special relativity only yields a one-dimensional Lorentz transformation. All physical derivations of the Lorentz transformation deduce it in one dimension. To maintain invariance of transformation in three spatial dimensions an additional rotation of the axes has to be assumed to form the Lorentz group, whose structure differs from the Galilean group. The assumption of this rotation of the axes has generally been considered to imply that a physical rotation of an inertial frame occurs when it is travelling at a relativistic velocity. For a material body located in the inertial frame, such a physical rotation would require a torque and an energy input, but there is no apparent source for this energy. These deductions arise solely because it has been assumed that any analytical deductions from Maxwell's equations do not require to be further interpreted in terms of their physical validity.

However, Terrell⁽⁸⁾ has produced an analysis which eliminates all

difficulties associated with the rotation of axes in the Lorentz group. Terrell's analysis has not received sufficient credit because it has not been appreciated that it provides the physical interpretation of the mathematics of the Lorentz group. It is well known that Terrell has shown that there is no visible Lorentz contraction in the direction of motion. But his analysis also requires that there is no physical rotation of the axes of a body in uniform translatory motion. If one considers observing a spherical, incandescent, material, moving body then the leading edge of the body will *appear* to trap some of the visible radiation. From an observer's point of view, the photons that define the leading and trailing edges of the body, on the image he or she observes on a flat screen, come from points on the body corresponding to a rotated diameter. Hence, the rotation of the axes of an incandescent moving body is a mathematical requirement which allows for the *apparent* distortion of the paths of the photons reaching the observer. Nevertheless, no actual interception of photons occurs at the leading edge of the body in the frame of reference of the body.

3. CONCLUSIONS

No observation of radiated electromagnetic energy is possible until a material sink is inserted into a free-space region. Fields and photons do not exist, as such, in free space. They are simply two alternative mathematical concepts which enable one to solve for different parts of any problem relating to the observation of electromagnetic energy in free space using a material detector. Quantum theory provides the extreme transient solution and electromagnetic wave theory provides for the near-steady-state and steady-state solution.

One may extend these concepts to cover the case where a source and a detector are in relative motion. Maxwell's equations cannot be used to analyse the problem associated with the detection of an electromagnetic wave unless a material sink is first introduced into a free-space region. To obtain the solution for what is observed it is then necessary to produce an analysis applicable to a material detector. The advanced potential solution dictates that the steady-state approach velocity of a signal beamed onto a Hertz dipole must be equal to c , relative to the detector, and independent of the source velocity. There is no need to make any assumption concerning the velocity of light "in empty space". It is the material that must be present to form a sink of energy which ensures that the observed steady-state velocity of light is always c . It is, therefore, the act of observation with a

material detector which governs the observed steady-state velocity of light. The advanced potential solution is vital because it provides the essential solution for the observed steady-state velocity of the energy received by a material detector. However, for an extreme transient, where only the quantum nature of the energy is relevant⁽⁷⁾, the observed velocities are not limited to c . Such superluminal velocities have been observed, and tunnelling velocities of individual photons as large as $4.7c$ have been noted^(9, 10). There is no reason why Maxwell's equations (and hence special relativity) should restrict the velocity of such a transient to c .

Terrell's analysis demonstrates that the complication of the rotation of axes, in the formation of the Lorentz group, arises because an observer viewing a moving, incandescent, material body considers that the energy emitted at the leading edge is intercepted by the body. A mathematical rotation of the axes of the body has to be made to allow for this apparent interception of the energy. But there is no actual interception of the energy. Hence, there is no physical rotation or contraction of a moving body. This deduction has clear consequences. If there is no physical rotation of the body then there can be no anomalous energy required to produce the rotation.

The derivation of special relativity from the advanced potential solution restricts the theory to material inertial frames of reference. All anomalies associated with the theory are then eliminated. Thus, if special relativity is restricted to material inertial frames, both special and general relativity may be anchored in the real Universe. Any material frame must have a finite mass. But a uniformly moving frame of finite mass could not have achieved its motion unless there was, at some time in the past, an exchange of momentum between the material of the frame and the rest of the Universe. A clock will then only show a time difference, as a result of its motion, following an exchange of momentum between the clock's material inertial frame and the rest of the Universe, which accords with Mach's principle. Special relativity is thereby anchored in a unique preferred inertial frame provided by the distant stars. As all aspects of time transformation arising from motion in general relativity have their origins in special relativity, general relativity will also be anchored in the real Universe.

If special and general relativity are anchored in the real Universe then it is likely that a classical theory must exist to explain both the origins of the value of the gravitational constant, G , and also the observed identical nature of the gravitational mass and inertial mass of any given body. The

gravitational theory of Hoyle and Narlikar⁽¹¹⁾ proposes that the value of G may be inversely proportional to the mean density of all matter in the Universe. Their theory is viable because they show that local matter would have an insignificant effect on the value of G . However, to explain both the origin of G , and the equality of inertial mass and gravitational mass, it is necessary to extend their static theory so as to produce a dynamical theory relating G to the density and the motion of distant matter. Only two large-scale, inertial motions of distant matter exist. One is the expansion velocity of a given region, and the other is the rotational inertial motion of a given galaxy. We will consider the consequences if the terrestrial value of G were to be generated, in a classical cause-and-effect manner, by the rotational inertial motion of distant matter in our galaxy. If the value of G had its origin in the rotational inertial motion of distant matter in our galaxy, then the equality of gravitational mass and inertial mass would follow as a necessary consequence. A specific dynamical theory for the value of G is possible^(12, 13). This theory relates the free-space value of G within any gravitationally-stable, rotating galaxy to the mean angular velocity (ω) and density (ρ) of the galaxy, such that $G = \omega^2 / 2\rho$. The theory will also apply to gravitationally-stable, rotating galactic clusters. G will then have different values in other galaxies and galactic clusters. It is interesting to note that an idealized, gravitationally-stable, rotating fluid body has two unique characteristics. First, all particles within the body are in continuous free fall. Secondly, the equivalence principle applies throughout the volume of the body as a whole and not just at individual points.

General relativity assumes that G is a universal constant, but this assumption may be unnecessarily restrictive as it requires the assumption of dark matter to explain the stability of many galactic clusters. There is no inherent requirement in general relativity that insists that G should be a universal constant, and this deduction may be easily demonstrated. Suppose that, on another planet in some distant galaxy, G had a value that was, say, 10 percent larger than our terrestrial value. Provided that the law predicting the value of G is the same in all inertial frames then the principle of relativity is satisfied. On this distant planet scientists would gradually develop the laws of physics, up to and including general relativity, and all of the tests of general relativity would be satisfied with the larger value of G . There is, therefore, no inherent physical reason for continuing with the assumption, originally made by Newton, that G is a universal constant. The

larger value of G , predicted by the theory for all galactic clusters which have a low density, would then explain their stability without the need to assume the presence of dark matter.

The proposed dynamical gravitational theory is fragile to experimental testing, as it predicts that a perturbation of G will arise from the rotation of the near-fluid Earth. A terrestrial measurement of G will indicate a value that is about 0.4 percent higher when it is measured below the geoid compared to when it is measured in free space above the geoid. The Bouguer anomalies present in observations of the acceleration due to gravity (about a 0.4 percent increase below the geoid) might then arise partly from actual variations in G . Such a specific check on G has never been made, but early measurements are consistent with this prediction⁽¹⁴⁾. Two very recent measurements^(15,16), made in Colorado, USA, and Wuppertal, Germany, show unexpected differences in the measured values of G well in excess of the expected experimental error of about 0.05 percent. While these experiments are still being undertaken it would be valuable to conduct two additional experiments, using the same apparatus, one located at a height of about 2km above the geoid and the other located at about 2km below the geoid.

Acknowledgement

I am very grateful to Professor C. W. Kilmister and Professor A. L. Cullen for many invaluable comments and suggestions. I also wish to thank a referee for proposing a number of clarifications.

Received 1 February 2000

References

1. J. D. Jackson, *Classical Electrodynamics*, (John Wiley & Sons, New York, 1962).
2. F. Hoyle and J. V. Narlikar, *Action at a Distance in Physics and Cosmology*, (Freeman, San Francisco, 1974).
3. J. G. Cramer, *Phys. Rev. D* **22**, 362 (1980).
4. L. M. Stephenson, *Found. Phys.* **8**, 921 (1978).
5. J. A. Wheeler and R. P. Feynman, *Rev. Mod. Phys.* **21**, 425 (1949).
6. A. L. Cullen, *Proc. R. Soc.* **366**, 155 (1979).
7. L. M. Stephenson, *Phys. Essays* **11**, 585 (1998).

8. J. Terrell, Phys. Rev. **116**, 1041 (1959).
9. A. Enders and G. Nimtz, Phys. Rev. E **48**, 632 (1994).
10. G. Nimtz, Gen. Rel. & Grav. **31**, 737 (1999).
11. F. Hoyle and G. Narlikar, Proc. Roy. Soc. **A282**, 191 (1964).
12. L. M. Stephenson, Found. Phys. **6**, 143 (1976).
13. L. M. Stephenson, J. Phys. A., Gen. Phys. **2**, 475 (1969).
14. L. M. Stephenson, Proc. Phys. Soc. **90**, 601 (1967).
15. J. E. Faller and J. Schwarz, Measurement Sci. & Tech. **6**, 478 (1999).
16. U. Kleinevoff, H.Meyer, A.Schumacher, S.Hartmann, Measurement Sci. & Tech. **6**, 492 (1999).

Lawrence M. Stephenson

Department of Electronic and Electrical Engineering
University College London, WC1E 7JE, England