Einstein's General Theory of Relativity

Einstein's General Theory of Relativity

by

Asghar Qadir

Cambridge
Scholars
Publishing

Einstein's General Theory of Relativity

By Asghar Qadir

This book first published 2020

Cambridge Scholars Publishing

Lady Stephenson Library, Newcastle upon Tyne, NE6 2PA, UK

British Library Cataloguing in Publication Data A catalogue record for this book is available from the British Library

ISBN (10): 1-5275-4428-1
ISBN (13): 978-1-5275-4428-4

Dedicated To

My Mentors:
Manzur Qadir
Roger Penrose
Remo Ruffini
John Archibald Wheeler
and
My Wife:
Rabiya Asghar Qadir

Contents

List of Figures

Preface

This book was started more than 30 years ago and was ready in some rough form a couple of years after that. I was awarded a "Book Project" to write this book by the King Fahd University of Petroleum & Minerals nearly a quarter of a century ago and the book was ready in a typed form needing some serious editing over 20 years ago. You will notice my slow progress with it even then. It has "idled" since then until Ghulam Abbas, an ex-PhD student of Muhammad Sharif, an ex-PhD student of mine, complained publicly at a Conference about my unfairly withholding this book from the next generations of my students. I felt that he had a point and, fortuitously, I received an invitation by Cambridge Scholars Publishers to submit a book proposal. Knowing myself by now (being over 72 years I have had time to get to do so), I felt I had been remiss long enough and that it was only by committing myself that the book would ever see the light of day.

The book is based on my lectures on General Relativity since 1971, when I joined what was then the University of Islamabad, Pakistan, and later became the Quaid-i-Azam University, Islamabad, Pakistan. I taught the book at the local "M.Sc.", which is the equivalent of the senior years of the 4-year BS, and at the local "M.Phil.", which is the equivalent of the American MS and British M.Sc. It has been taught as a one-year, or two semester course. It is written so as to be able to teach students of the senior undergraduate of earlier postgraduate with a Physics background who have studied Special Relativity from my book *Relativity: An Introduction to the Special Theory* (World Scientific 1989) or equivalent, but do not have a sound background of Geometry. It can be used for students of Mathematics who have not studied Special Relativity but have a strong background of Geometry, by replacing the part on Geometry by chapters 2, 3, and parts of 5, 6 and 7. I would break the course off part way through Chapter 5, at section 5, and proceed for the rest of Chapter 5 and the next three chapters in the next semester. Chapter 8 of this book contains various recent developments and some other special topics (some of which could be left out from the course without any damage done to the rest of the course).

Let me also talk a bit about those to whom the book is dedicated. All my mentors said that if one cannot explain something simply, one has not understood it. My first mentor was my father, who not only started my education in Mathematics but was the person who, despite being a lawyer, first introduced me to the subject of Relativity, at the age of 9 and motivated me to try to understand the subject. I learned from him, also, that knowledge does not come by degrees but by curiosity — the desire to know. I found his understanding and knowledge of Mathematics better than many PhDs in the subject. He was very critical of anyone trying to argue by bald claims hidden behind layers of jargon.

As regards my second mentor, I cannot imagine a better PhD supervisor than Roger Penrose. When I would say something stupid, he would not say it was stupid but that he did not understand it — *and he meant it*! When the discussion led to the correct version, he never pointed out that I had been wrong. Without appearing to guide me to the solution of the problem I had been interested in addressing when I joined him, by the end of the PhD he had got me to the stage of doing what I had wanted to achieve. He was not ready to take the "accepted wisdom" as correct, but judge it for himself each time. From him I learned to do the same. From him I also learned how Mathematics could lead, not merely to correct physical consequences, but to physical insight. He always gave credit for ideas freely and never claimed it for himself. He never put his name on a paper that was not significantly his.

From my third mentor, Remo Ruffini , I learned the importance of enthusiasm for the subject, especially in talking about it and communicating it. I had been involved in the attempt to find a Quantum Theory consistent with General Relativity. Remo got me interested in Relativistic Astrophysics. I was also fortunate to see a selfless appreciation for work on the development of ideas, rather than trying to grab credit for it. He had found the mass limit at which a collapsed object must become a black hole. When he went to China, he found that Fang Li Zhi had discovered the self-same limit, but had been unable to publish it in Western journals because this was at the time of the "Cultural Revolution". Fang had published it in China. Remo publicized the discovery by Fang as a contemporaneous independent discovery. I also learned from him the importance of using humour in and human interest in communicating serious Physics.

My fourth mentor, John Archibald Wheeler, started my love affair with Physics. He did not separate off parts of Physics but saw it as a unified whole. His "poor-man's way" of seeing results was an indispensable tool for his understanding. He needed to see things simply before going for long calculations to get the correct answer. As he said "I never start a calculation unless I know the answer". And *how* did he know the answer? By the poor-man's way. He also had a knack for catchy phrases and turns of expression. He invented the terms "black hole" and "big crunch" for example. His juxtaposition of opposites would express it all, as with "magic without magic". From him I learned the importance of saying things in a way which would catch the imagination and stay with the reader (or listener).

My wife is to blame for my still being around to write the book. If it had not been for her, it is highly unlikely that I would have actually got the book written, leave alone published, as I would have died long before.

I would be remiss not to thank all my students on whom I tried out my explanations and developed them to the point where most could follow what i taught. I must particularly thank two recent students of mine: Shameen Khattak for a very thorough proof-reading of the mathematical calculations in my book, eliminating many errors in the earlier draft; and Muhammad Usman for helping with handling the LaTeX required for typing the book and with diagrams.

Chapter 1

Introduction

If you say “Relativity”, everybody thinks “Einstein” and if you say “Einstein”, everybody thinks “Relativity”. It may not be fair to Einstein to limit him to the theories of Relativity, as he was also the first person to believe in a quantum of energy and got a Nobel Prize for that work, along with his prediction for Brownian motion. Nor, for that matter, is the theory of Relativity solely developed by Einstein. The names of Poincaré and Hilbert are often mentioned as cofounders for the development of the Special and General Theories, respectively (and the name of Marcel Grossmann strangely suppressed for the latter). I will try to explain the development of the unrestricted theory, following a historical perspective, and explain why the theory should genuinely be regarded as Einstein’s creation, despite all the contributions of other researchers. However, the essential purpose of this book is to explain the unrestricted, or general, theory so that the reader can actually follow the latest developments in the theory. But first some words about the first, *restricted*, or special, theory (being restricted to constant velocity).

When Special Relativity was developed, the misnomer of the theory created a lot of confusion. What Einstein had developed was a theory that said that the simultaneity of two events that occur, was not only dependent on the positions of two observers, but also on their *relative* velocity. That an observer near one event would see that event before a more distant one, did not take an Einstein to know — being rather blatantly obvious. The German name for the theory was “the relativity of *simultaneity*”. In fact, the theory goes on to discuss those quantities that *do not* depend on relative motion. This is discussed more fully in my book on Special Relativity [1]. I will not review the Special Theory here, but do need to contrast the views in it, rather than the results, with Newton’s views.

In Newton’s view of the Universe, space and time are “absolute” entities in themselves. Space exists, whether it is occupied or not; whether anybody sees it or not; whether the person seeing it is moving or not — it just “is”. This ran counter to the usual thinking based on Aristotle’s metaphysics. To make sense of this belief, Newton invoked the existence of God as a universal observer. Despite the fact that this thinking got ingrained into us, if one thinks about it afresh, it *does* seem strange — what is meant by the existence of nothing? Newton also assumed that “time flows at a constant rate”. Again, this (now) trite observation contains in it the question of what is meant by “the flow of time", as if it were a stream flowing? When a particle in a stream is seen to move some distance in a unit of time at one stage, and a different distance at another stage, we say that the rate of flow of the stream has changed. If it is not seen to change, we say that the rate of flow is constant. How can "the rate of flow of time", then, mean *anything*? All Einstein did was to challenge these mystical beliefs, and replace them by assumptions relating to actual observation of physical quantities in some (thought) experiment. In this sense Einstein only cleared up confusion caused by unnecessary assumptions.

The unfortunate name of the theory led people to take it that Einstein had somehow argued that *everything* — even ethical values — is relative. Since Relativity was regarded as "scientifically proved", it was claimed that all certainty in life and reality was lost. He was regarded as a new Shakespearian Prospero who had made the World tempestuous saying "We are such stuff as dreams are made on" (The Tempest Act 4, Scene 1). Ironically, it was the Quantum Theory that actually destroyed Victorian certainty, by saying that all physical predictions are only probabilistic and not deterministic. Probability was already used for Statistical Mechanics, but only as a way of getting approximate results for something that could be known more precisely in principle. Quantum Theory insisted that it *could not* be known. Though Einstein had been one of its founders, he strongly disagreed with this probabilistic interpretation of the theory. Nevertheless, to an epitaph for Newton:
'Nature and her laws lay hid in night,
God said "Let Newton be!" − and all was light!',
someone added the couplet for Einstein:
'But not for long, the Devil howling "Ho!
Let Einstein be!", restored the status quo.'
That couplet would have applied better to Niels Bohr and Werner Heisenberg, who had pioneered the view of an inherent probability in the laws of Nature. But I suppose, "Let Bohr and Heisenberg be", would not go down that well, as it loses the meter.

Actually, Einstein was very clear that accelerated motion is *not relative*, in that it does not depend on the velocity of the observer. He talked of this by giving the example that if a train is moving smoothly a passenger will feel nothing but if the train speeds up or slows down, the passenger will be pushed back or to the front. It must have taken a lot of imagination for Einstein to think of a train of those days moving smoothly. But then, there were no planes in those days. (The Wright brothers had taken their maiden flight but that was pretty well all.) Galileo had a better example for the relativity of uniform linear motion with a boat moving in a calm sea, which was presumably modified by Einstein to a more "modern" example for the times. My point is that Einstein had realized that accelerated motion would be detectable from within a closed laboratory. This led him to focus on another mystical belief of Newton's view — to do with his law of gravity. Newton's view was that the force of gravity of a mass is instantaneously felt at a distance. Thus, if the Sun were to suddenly disappear, the Earth would be released from its orbit and go flying off at a tangent. Just imagine: you are seeing the Sun in the sky, and then suddenly, 8 minutes 20 seconds later, you see it go shooting off and disappear. (Bear in mind that it would take light that long to reach the Earth from the Sun.) Of course, that is an absurd example, as the Sun could not suddenly cease to exist, so one may not bother about it. However, what if the Sun were just accelerated away? How could the information reach the Earth instantaneously, as nothing can go faster than light? There must be gravitational disturbances that travel at the speed of light — gravitational waves!

Einstein noted one other point. In Newton's laws the mass appears in two ways: in the second law of motion as inertia; and in the law of gravitation as a sort of gravitational charge. Attempts had been made to try to find a difference between the two, but (as I shall be mentioning shortly) had given a null result. Einstein realized that this could not be accidental, they must really *be the same*. He, therefore, stated "the principle of equivalence" that the gravitational and inertial masses are identical. That means that at one point one cannot distinguish between a gravitational effect and the effect of acceleration. In modern terms, if one is in a closed laboratory, it will not be possible to tell, by any experiment, if the laboratory is being accelerated by a rocket or being held up against the pull of the gravity of a planet. Conversely, if the laboratory is in a lift, by cutting the cable holding the lift up and letting the lift fall freely, we will have "switched off gravity". He later described this realization as "the happiest thought of my life". Imagine him pondering this matter with soft music playing in the background. As he ponders and comes closer to the realization, the music speeds up and grows in volume. Then, when he is struck by *the thought*, the music reaches a crescendo with a clashing of cymbals. General Relativity has arrived!

Unfortunately, the physical theory required some mathematics that Einstein had failed to pick up in his stay at the ETH Polytechnic at Zurich − namely Geometry. Herman Minkowski had taught a course on the subject that Einstein had studiously bunked. When Minkowski had recast Einstein's Special Theory in geometric terms as kinematics in a 4−d spacetime, Einstein had rejected the development, saying that it lost the physical understanding and obfuscated it in mathematics. Now, when he needed to make a workable theory from "the happiest thought of his life", he did not know how to do so. He went to his friend Marcel Grossmann, to teach him the required mathematics. After various failed attempts using other types of Geometry, such as Affine Geometry and Teleparallelism, they hit on Differential Geometry as the language for the theory. Two papers were published by them in 1913 and 1914 [2, 3], in which the theory was almost fully formulated using the (earlier) much hated Differential Geometry.

At the time most theoretical physicists used the Euler-Lagrange (EL) equations generalized to deal with the fields of James Clerk Maxwell. What remained for the theory was to formulate it in these terms, rather than what was regarded as the unfamiliar language of Geometry used in the Einstein-Grossmann papers. Einstein had been in correspondence with David Hilbert, who was abreast with Einstein's work so far. In 1915, the two of them independently took the next step of the field-theoretic formulation, in which a correction of the earlier papers was given. There were four by Einstein and one by Hilbert published in 1915 [4, 5]. The paper in which all errors were removed by Einstein appeared in 1916 [6]. Hilbert acknowledged Einstein as the originator of the theory in his paper, claiming only to axiomatize the foundations of Physics, but people claim priority for his work because Einstein's paper appeared (a bit) later. (Perhaps, Hilbert and the claimants being Christians, and Einstein being a Jew, had something to do with it. Being a Muslim, I can be objective, as no Muslim ever came close to contributing anything to Relativity till *very* much later.) It *is* Einstein's General

Theory of Relativity.

While there are many, and varied, successes of both SR and General Relativity (GR), there are problems with this theory. The most glaring is the fact that it singles out Gravity from the other fundamental forces of Nature. Einstein's "happiest thought" used Gravity to generalize the theory of uniform linear motion to arbitrary motion. In SR, forces were dealt with ignoring the philosophical problem that the force resulted in acceleration, which did not allow velocities to remain constant. One could argue that it was only the object that was accelerated and not the observer. However, by the spirit of Relativity, the object is an equally good observer. One could say that in the SR view *all forces are equally "bad"*. By using Gravity, arbitrary motion is incorporated, *but only Gravity is "good" and the other forces remain "bad"*. (It reminds one of George Orwell's *Animal Farm* where an animal revolution was started on the slogan that "All animals are equal". Then the pigs take over the revolution and the slogan is modified to "All animals are equal – but some are more equal than others". All forces are equal, but Gravity is the most equal force.) This deficiency bothered Einstein. At the time the only other fundamental force known was Electromagnetism, and he tried to extend GR to incorporate this force in a Unified Field Theory. Soon afterwards, Hilbert tried and then others joined in the attempt. Despite various claims there has been no philosophically satisfactory attempt that is free of problems. Of course, soon afterwards it became clear that there was another fundamental force responsible for the decay of heavy atomic nuclei and then one to hold the nucleus together. These are called the "weak" and the "strong" nuclear forces. Einstein never believed that they were fundamental and hoped to demonstrate that they were "effective forces" that were approximate descriptions of the interaction of Gravity and Electromagnetism.

Another problem was the relationship of Relativity to the Quantum Theory. Paul Maurice Dirac had developed a procedure to convert a classical field theory to a quantum version, the so-called "quantization of the classical field". This method proved extremely successful for the quantization of the electromagnetic field, leading to Quantum Electrodynamics (QED). When he attempted the quantization of the gravitational field, Dirac obtained meaningless answers. He was ready to ascribe them to the same cause as the meaningless answers provided for QED, yielding infinite probabilities. However, others managed to address the issue for QED and obtain correct answers by making the infinities irrelevant. The Quantum methods were applied to the nuclear forces to provide a beautifully elegant way of dealing with them as fundamental forces, and the procedure for rendering the infinities harmless worked well for them. Abdus Salam, Sheldon Glashow and Steven Weinberg managed to provide a unified theory of the electromagnetic and weak nuclear forces to a single "electro-weak" force. This force was compatible with the strong nuclear force, so that the three can be put together as "the Standard Model" of Particle Physics. There have been attempts to provide a "Grand Unified Theory" of the three forces put together as a neat whole package, but there are problems with the attempts that I will not go into here. However, the Quantum methods failed when applied to gravity and it was shown that they were always doomed to fail. There seems to be a much deeper tension between the two theories that precludes their marriage.

Despite its deficiencies, and all its difficulties, Relativity is "the only game in town" to provide answers for questions involving gravity alone. Further, it is not as if one cannot get answers when other forces interact with gravity, but only that the methods are not *philosophically* satisfactory. Since both fundamental theories have philosophical problems, the hope is that when a correct theory arrives it will resolve all the problems. People on one end of the spectrum expect that one or more of he tenets of Relativity will need to be altered to make it compatible with the Quantum Theory; and those on the other end expect that Quantum Theory is the one that needs modification. Many have been claiming that Superstring Theory, or one of its derivatives like "brane theory" or "M-theory", will be "the Holy Grail". At the other end, many people follow Einstein's belief that Quantum Theory is flawed and when it is "corrected", or "completed", the true theory will be found. Generally, those who come to the problem from the Quantum side are of the former type and those who come from Relativity are of the latter type. Probably, both theories will need to be modified.

Almost immediately after his first complete formulation of GR, in 1916 he demonstrated that the theory necessarily required the existence of gravitational waves that travel at the speed of light [7]. The problem with detecting them is that they are about 10^{38} (a hundred trillion, trillion, trillion) times weaker than electromagnetism. With the recent discovery of gravitational waves in 2016, this prediction was verified a century after it was made! That seems to me to be something of a record.

Very soon after his final formulation of GR, in 1917 Einstein [8] tried to apply it to the Universe as a whole. At that time such discussion fell into the realms of Theology. Some monks had given physical arguments in favour of their favourite theological cosmology, but there was no interest in them among the physicists. Of course, physicists had always held their own religious beliefs but, since the time that Simon Laplace, had said "[Sire,] je n'ai pas eu besoin de cette hypothèse" ("[No, Sire,] I had no need of that hypothesis"), when Napoleon pointed out that there was no mention of God in his book on the Heavens, they did not bring it into their Physics. Einstein had obtained tests of GR that only gave very fine corrections. He probably wanted a situation where "GR would rule". As it transpired, this was a very fruitful line of enquiry and much work followed from it. The need for precise testing of Einstein's theories has not only contributed to our understanding of the physical world around us, it has driven technological development, leading to developments of telescopes and of laser interferometers in space. It used to be said that QED is the most precisely tested theory in Physics. GR has, since joined it.

Considering the wide variety of areas of Physics that Einstein made seminal contributions in, it is difficult to keep track of all that he did relevant to the development of SR and GR. I will not try to cite all his relevant papers, preferring to refer to two excellent biographies of Einstein and his work [11, 12] and the list of Scientific Publications of Albert Einstein on Wikipedia.

1.1 The Equivalence of Gravitational and Inertial Masses

It had already been noticed, by the nineteenth century, that the term "mass" appears in two separate contexts in Mechanics, with no reason to regard the term as identical in both contexts. One is the resistance offered to a force that tends to accelerate an object, called the "inertial mass". The other is as a gravitational analogue of the electromagnetic charge, a sort of "gravitational charge", called the "gravitational mass". Despite the formal similarity between the gravitational and electromagnetic force laws there are three fundamental, physical points of difference between them. First, the gravitational charge is always positive (giving an attractive force only) while there are three possible types of charge, namely positive, negative and neutral (giving attractive, repulsive and zero forces). Second, in gravity like charges attract instead of repelling as in electromagnetism. Correspondingly the field intensity and hence the potential of gravitation has the opposite sign to electromagnetism. This point had been noted by Maxwell as a serious hurdle to extending his ideas for electromagnetism to gravitation. This means that the energy associated with gravity should be negative, if the energy associated with electromagnetism is taken to be positive. Third, the electric charge is quantized, in that it occurs in multiples of a third of the electron charge, while there does not seem to be a corresponding quantization of the gravitational charge. Any attempt to unify the forces of nature and Relativity with Quantum Field Theory must take these points into account. However, for our present purposes these differences are not so important. What is important is that there is no *a priori* reason why the inertial and gravitational masses *should* be identical.

An experiment to test the identity between the inertial and gravitational masses was performed by Baron Eötvös starting in 1886 (and going on to 1909). As with the Michelson-Morley experiment it was a crucial null experiment. Unlike that experiment the null result was expected and so there was no resistance to accepting its result. Only Einstein seems to have realized its significance. Since this was the only *fact* − the only piece of experimental evidence − on which GR is based, it is worth a more detailed discussion. It holds the same position for GR that the consistency of the speed of light does for SR. In fact, it is the basis of Einstein's "happiest thought of [his] life". In basing his new, unrestricted, theory on this fact, he elevated it to the status of a principle, calling it the "principle of equivalence". This is the crucial step in separating gravity from the other forces. Henceforth unifying gravity with the other forces would be something of a non-sequitur − at least if we require GR to hold. Einstein never seems to have realized that his "happiest thought" had turned into his "saddest thought" in his subsequent attempts to unify gravity and electromagnetism. It may be that this problem is at the base of those attempts at "quantizing gravity" that also unify unify the fundamental forces by modifying GR without touching Quantum theory. Because of its importance, this principle has gone on being tested to this day with the equivalence being maintained.

To try to distinguish between the gravitational and inertial masses it is necessary to allow the same body to experience a gravitational and an inertial force simultaneously. Let the gravitational and inertial masses be denoted by M and m respectively and the corresponding accelerations by $\mathbf{g}$ and $\mathbf{a}$ respectively. The net force on the body will then be $M\mathbf{g} + m\mathbf{a}$. If the forces are parallel or anti-parallel $\mathbf{a} \propto \mathbf{g}$ or $\mathbf{a} = k\mathbf{g}$, where k is a constant which may be positive or negative. The total force will then be $(M + km)\,\mathbf{g}$. Thus we will only be able to measure the combined, effective, mass $(M + km)$ and not be able to resolve it into the two separate masses. However, if the forces are at some oblique angle we will have two components with different combinations of M and m. We could use them to solve a pair of simultaneous equations in M and m so that we could distinguish between them. For this purpose it is necessary, therefore, that the inertial force should not act straight up or down. The most convenient inertial force is the centrifugal force due to the Earth's rotation. Ofcourse, as we increase the magnitude of this force (by moving towards the equator) we decrease the angle and as we increase the angle (by moving towards a pole) we decrease the magnitude.

It is never as easy to measure the absolute values of two quantities as to measure small differences between them. The accuracy is much greater in the latter case. An example of this basic fact of experimentation is the relative ease with which the Michelson interferometer can measure path differences $\sim 10^3$ (or 10^{-7}m) while the path length can be measured with an accuracy of only $\sim$ 1mm (or 10^{-3}m). This fact is also repeatedly encountered in making Astronomical and Cosmological measurements, which I will not go into here. Eötvös also used this principle. He looked for a difference between the ratio $\alpha = M/m$ for different bodies. If the ratio is the same for all bodies it can be chosen, by appropriate choice of units, to be unity and hence gravitational and inertial masses will be identical. Since it is known that the two types of mass are more or less the same we know that α is close to unity for all bodies. Eötvös tried to measure the small difference $|1 - \alpha|$.

Notice that for the rotating Earth the magnitude, a of $\mathbf{a}$ (we will denote the magnitude of a vector by the same letter but not in boldface) is given by

$$a = R\omega^2 \cos\theta \ , \tag{1.1}$$

where R is the Earth's radius ($\approx 6.4 \times 10^8 m$), ω the angular frequency of the Earth's rotation ($\approx 7.3 \times 10^-5 rad/sec$) and θ is the latitude where the experiment is performed (being the complement of the usual polar angle in spherical polar coordinates in the Northern hemisphere). For a to be large θ must be small. However, as pointed out above, the quantity to be measured will be large when there is a large difference between the direction of $\mathbf{a}$ and $\mathbf{g}$. For concreteness take θ to be 30^o, $a \approx 0.3$ m/sec^2 while $g \approx 9.8$ m/sec^2. Thus, typically, $a \approx 3\%$ of g, which is a sizable value.

The experimental apparatus consisted of a rod, which can be represented by a vector $\mathbf{b}$ with two dense bodies attached, which we denote by "1" and "2". Rotating the rod through π radians interchanges "1" and "2" and reverses the vector $\mathbf{b}$, i.e. changes $\mathbf{b}$ to -$\mathbf{b}$. Denoting the gravitational and inertial masses of the two bodies by the previous symbols with the corresponding subscripts, the net torque due to the two forces (gravitational and centrifugal) is

$$\mathbf{T} = \mathbf{b} \times [(\mathbf{M_1} - \mathbf{M_2})\mathbf{g} + (\mathbf{m_1} - \mathbf{m_2})\mathbf{a}] \ , \tag{1.2}$$

(where $\times$ represents the usual "cross product" of vectors) while the resultant of the two forces on both bodies is

$$\mathbf{F} = (M_1 + M_2)\mathbf{g} + (m_1 + m_2)\mathbf{a} \ . \tag{1.3}$$

Thus the effective torque, which is the quantity that will actually be measurable, will be the component parallel to the resultant force (see Fig. 1.1)

$$\begin{aligned} T_\parallel &= \frac{\mathbf{T}\cdot\mathbf{F}}{|\mathbf{F}|} \\ &= \frac{\mathbf{b}\wedge[(M_1-M_2)\mathbf{g}+(m_1-m_2)\mathbf{a}]\cdot[(M_1+M_2)\mathbf{g}+(m_1+m_2)\mathbf{a}]}{2[\{(M_1+M_2)\mathbf{g}+(m_1+m_2)\mathbf{a}\}\cdot\{(M_1+M_2)\mathbf{g}+(m_1+m_2)\mathbf{a}\}]^{1/2}} \ . \end{aligned} \tag{1.4}$$

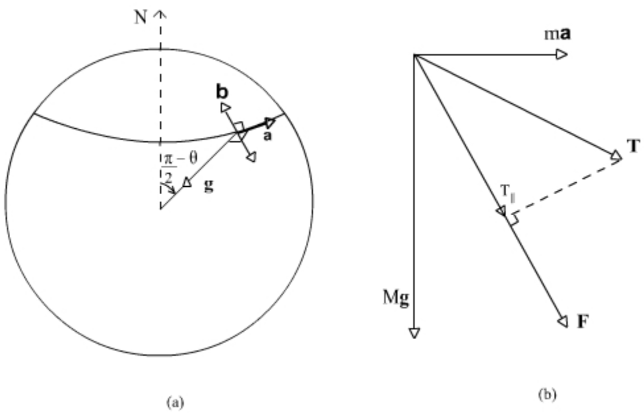

Figure 1.1: The Eötvos experiment. (a) At a point on the Earth's surface, at latitude θ a rod of length b is placed horizontally in the North-South direction. The centrifugal acceleration, a, is clearly orthogonal to the rod, b. (b) The resultant force, $F = Mg + ma$ does not act straight down. The component of the torque $\mathbf{T}$ along this force, $T_{||}$, is measured.

Now $\mathbf{g}.\mathbf{g} \sim 100, 2\mathbf{g}.\mathbf{a} \sim 5, \mathbf{a}, \mathbf{a} \sim 0.1$, in units of $(\mathrm{m/sec}^2)^2$. Thus the denominator may be approximated by $2(M_1 + M_2)g$, as the total gravitational mass is more or less the same as the total inertial mass, even if there is some slight difference between them. In the numerator, we can change the order of the scalar triple product so that the cross appears between the second and third terms. Since $\mathbf{g} \times \mathbf{g} = 0 = \mathbf{a} \times \mathbf{a}$, we are left only with a $\mathbf{g} \times \mathbf{a}$ term with coefficient $(M_1 - M_2)(m_1 + m_2) - (M_1 + M_2)(m_1 - m_2) = M_1 m_2 - M_2 m_1$. The scalar triple product $\mathbf{b}.\mathbf{g} \times \mathbf{a}$ will be maximum if $\mathbf{b} \perp \mathbf{g}$, $\mathbf{a}$. Setting this as the orientation of $\mathbf{b}$ and writing Eq.(4) in terms of α,

$$T_{||} \approx \frac{m_1 m_2(\alpha_1 - \alpha_2)}{m_1\alpha_1 + m_2\alpha_2} bR\omega^2 \sin\theta\cos\theta \; . \tag{1.5}$$

Clearly, the maximum value of $T_{||}$ will be at $\theta \approx \pi/4$ rad (or 45^o latitude). Taking $b = 1$m, $m_1 = m_2 = 1$kg, the effective torque (in mks units) is given by

$$T_{||} \approx 0.017(\alpha_1 - \alpha_2)/(\alpha_1 + \alpha_2) \approx 0.0085(\alpha_1 - \alpha_2) \; . \tag{1.6}$$

The torsion is actually measured by rotating the rod through π radians and observing the deflection of a spot of light reflected from a mirror attached to the wire whose torsion is being measured. Thus $2T_{||}$ is measured. Eötvös found that $|\alpha_1 - \alpha_2| < 10^{-8}$. Later experiments have raised this accuracy to $\sim 10^{-14}$. Even now, there remain attempts to introduce scalar fields (or even vector fields), which will lead to a fifth force, and argue that the effect is so small that it has not yet shown up.

There had been claims of an observed difference attributed to a "fifth force". Despite the great excitement generated at the time there was no satisfactory evidence for it, as the various claims were mutually contradictory [79] and the claims were later withdrawn.

1.2 Field Theory

A crucial aspect of the development of General Relativity is that it is expressed in the language of field theory. Many of the failed attempts and competing theories were also field theoretic. Though I prefer to develop it as a theory of motion, to most people it is merely a field theory of gravity. I will, therefore, very briefly present the essentials of the subject of Field Theory here and leave the main part of the discussion to Chapter 4. For a more detailed discussion the reader is referred to Landau and Lifshitz (LL), *The Classical Theory of Fields* [14], and to the relativists' Bible, "Gravitation" [15] by Misner, Thorne and Wheeler (MTW). The field theory*par excellence* is Maxwell's theory of the electromagnetic field. Its remarkable power, elegance, success and simplicity have led to its current status of "role model". Maxwell had considered the possibility of extending field theory to incorporate gravity but concluded that it was impossible. As mentioned earlier, Einstein and others made various false starts at a field theory of gravity till the final formulation of Einstein and Hilbert. One of the most successful attempts to unify the two forces was made in two parts by Kalutza and Klein (which will be briefly discussed in Chapter 5), but it had problems. While Einstein was very enthusiastic about it at first, he later rejected it.

More recently the development of field theory has received fresh impetus from considerations of symmetry of the field in some context. By "symmetry" is meant that the field is invariant under some transformations. If these transformations form a group (as they generally do) the theory can be expressed in the form of the symmetry group. In 1918 Emmy Noether stated a theorem according to which, for each generator of the group there will be a conserved quantity (a "charge"). Of particular interest are transformations which modify the potential functions without altering the physical quantities. These are called *gauge* transformations. If the invariance is only under global transformations we have a not-so-interesting "global gauge symmetry". To illustrate the difference between the two, consider the rotation of an irregular object. A rotation through 2π radians will leave it invariant but through any other angle will change it. This is a global symmetry. There can be global symmetry of rotation through π radians. For example an ellipse when rotated about its centre through π radians is left invariant. A very much stronger symmetry is provided by a circle. Rotation through any angle, about its centre, leaves the circle invariant. This is a local symmetry. Local symmetry implies global symmetry but the converse is not true.

As a historical aside, it is of interest to note that despite her stupendous contributions in Mathematics and Physics, Emmy Noether could not be employed in a University in Germany on account of being a female. She had to teach totally uninterested (and uninteresting) students in a finishing School for "young ladies". It took Hilbert's efforts for her to be allowed to teach at the University of Göttingen.

The reason why these considerations became interesting is that they lead to non-trivial generalizations. For the above example of a local symmetry any two transformations will always commute. The group of rotations in two real dimensions, $SO(2)$, is Abelian. Similarly the group of unitary transformations in one complex dimension, $U(1)$, is Abelian. Now consider the symmetries of a sphere. It is invariant under rotations about any axis passing through its centre.

There are three independent generators. In general two arbitrary rotations will not commute, as anyone who has played with a Rubik's Cube, or Rubik's Revenge, will bear witness to. The group of rotations in three real dimensions, $SO(3)$, is non-Abelian. Similarly, the group of unimodular (with determinant 1), unitary transformations in two complex dimensions, $SU(2)$, is non-Abelian. (You may wonder why there is no $SU(1)$. The reason is that the generator of $U(1)$ is simply a phase, or complex number with magnitude 1. Hence its determinant is a phase. Making the phase angle 0 reduces to just the number 1.) Further, the symmetry group of transformations of the Euclidean plane into itself, E_2, is non-Abelian. Non-abelian gauge theories were first considered by Yang and Mills and the first example considered was $SU(2)$. Yang-Mills fields were used by Glashow, Salam and Weinberg to construct the unified electro-weak theory $SU(2)_W \otimes U(1)_Y$. Similarly, strong force has the symmetry group $SU_C(3)$. This gives the standard model symmetry group $SU_C(3) \otimes SU(2)_W \otimes U(1)_Y$. These types of considerations led to Supergravity theory, conformal field theory and then to Superstring theory. We will not go further into any of these developments.

1.3 The Lagrange Equations

In the early days Mechanics was developed to be able to predict the motion of all the bodies of the solar system known at the time. Solving for a planet in the field of an infinite mass Sun, is trivial and gives a wrong result, as the Sun does not have infinite mass. The method to correct for the finite mass was already developed by Newton in his *Principia*. In fact, Robert Hooke had proposed to him, the inverse square law for the force pulling the planets towards the Sun and Newton had generalized the idea to his law of universal gravitation, so that the planet would pull the Sun as well. For the purpose, one breaks the motion into two parts: one for the centre of mass and the other for each body orbiting about the centre of mass. However, for three bodies the 3 coupled differential equations could not be solved so simply. Lagrange developed the method of minimizing the "free energy", the difference between the kinetic and potential energies (called the Lagrangian), $L[q^i(t), \dot{q}^i(t)] = T[\dot{q}^i(t)] - V[q^i(t)]$, where $q^i(t)$ and $\dot{q}^i(t) \quad (i = 1...n)$ are the generalized positions and velocities for N particles subject to m constraints, $n = 3N - m$, as functions of time, so that L is a *functional*. A functional may have a constant value, but depend non-trivially on a function, or functions. Thus we can look for the form of the functions which give a minimum or maximum value of the functional. This is what we will be doing with the Lagrangian.

Hamilton, later, re-formulated Lagrange's mechanics in terms of the generalized positions and momenta, $(q^i(t), p_i(t))$ and demonstrated that the integral of the Lagrangian over a finite time interval, called the action S, must be minimized to obtain the path of a particle. This is called *Hamilton's principle of least action*. He further showed that there was a conserved quantity, H, which is the *sum* of the kinetic and potential energies, corresponding to the Lagrangian. This total energy is a constant of the motion as a function but is non-trivial as a functional. With the former way of looking at it we get Lagrangian mechanics, yielding the Lagrange equations, and with the latter Hamiltonian mechanic, yielding the Hamilton equations.

Hamilton's principle that the action, S, be minimal requires that its variation be zero, i.e.

$$\delta S = \delta \int_a^b L(q^i, \dot{q}^i) dt \ , \tag{1.7}$$

where a and b are initial and final times and $q^i(a), q^i(b)$ are given constant values, so that $\delta q^i(a) = 0 = \delta q^i(b)$. The variation can be evaluated inside the integral sign to give

$$0 = \int_a^b [(\partial L/\partial q^i)\delta q^i + (\partial L/\partial \dot{q}^i)\delta \dot{q}^i] dt \ , \tag{1.8}$$

where the Einstein summation convention, that repeated indices are summed over, is used here (and throughout the book). It can be demonstrated that the operators δ and d/dt commute. Thus we can integrate the second term in the integral by parts, to obtain

$$0 = \int_a^b [\frac{\partial L}{\partial q^i} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}^i}]\delta q^i dt + \frac{\partial L}{\partial \dot{q}^i}\delta q^i \Big|_a^b \ . \tag{1.9}$$

The last term here is zero as $\delta q^i(a) = 0 = \delta q^i(b)$. Thus the expression in the integral must be zero. This will generally be true (for all a, b) only if the integrand is zero. This requirement gives the *Lagrange equations*

$$\frac{\partial L}{\partial q^i} = \frac{d}{dt}\frac{\partial L}{\partial \dot{q}^i} \ . \tag{1.10}$$

It should be borne in mind that these are *necessary*, but not *sufficient*, conditions for extrema to occur. The sufficient conditions would come from a second variation, that is seldom undertaken. There are various directions for generalizing this analysis. We have assumed that the Lagrangian has no explicit time dependence. If there were explicit time dependence there would be an extra term in the Lagrange equations. In this case energy would not be conserved. In such non-conservative systems the extra term corresponds to energy dissipation or creation. Another generalization allows the Lagrangian to depend on higher derivatives of the generalized coordinates. The above equations are second order differential equations. They would then become higher order equations of motion. There is no evidence that such a generalization is required in Physics. Yet another generalization is to deal with less rigid constraints on the generalized coordinates and to leave one, or both, of the ends free. I will not discuss any of these extensions here as they are not at all relevant for a discussion of Relativity.

1.4 Extension of the Lagrange Equations to Fields

For a very large number of particles with few constraints, i.e. very large n, it becomes convenient to take the limit $n \to \infty$. Correctly speaking, we should take the infinity to be countable (like the natural numbers), but so as to be able to use calculus, we take the continuum limit. Thus every point has a different value of $q(t)$. We thus replace the generalized coordinates by a *field*, $\phi(t, \mathbf{x})$, where $\mathbf{x}$ has replaced the index label i. Thus the field is a function of time, at each position vector $\mathbf{x}$, and hence is a function of position as well. Of course, one can generalize so that $\mathbf{x}$ may be a lower or higher dimensional vector than 3. We would, correspondingly have a lower or higher dimensional field theory. There is a complication that arises, here, in regard to the Lagrangian. To explain that I will need to digress a little bit on the theory of cardinal numbers.

The cardinal number is the transfinite extension of the "number of elements" of a finite set. By setting up one-to-one correspondences we can compare infinite sets. The cardinality of the natural numbers (the *counting* numbers) is denoted by $\aleph_0$, (read *aleph null*). A set with a one-to-one correspondence with this set is said to be *countable*. Now, by the theory of ordered sets it is known that he cardinality of the power set of a given set (the set of all its subsets) is strictly greater than the cardinality of the set, $|\exp(A)| > |A|$. It can be easily demonstrated that $|\exp(A)| = 2^{|A|}$. These statements are as true for transfinite as for finite sets. Thus $2^{\aleph_0} \equiv \aleph_1 > \aleph_0$. It can also be proved that the set of real numbers is uncountable and hence it has a cardinality greater than $\aleph_0$. From Gödel's theorem it can be shown that it is possible to choose $\aleph_1$ to be the cardinality of the set of real numbers. This choice is known as the continuum hypothesis and will be adopted henceforth. (It is possible to choose otherwise and develop a different transfinite Mathematics but we will not go into that here.)

Taking the continuum hypothesis the cardinality of the space of all functions is $\aleph_2 \equiv 2^{\aleph_1}$, as that is the set of all subsets of $\mathbb{R}$, the set of real numbers. The number of degrees of freedom of a system of N particles subject to m constraints is $n = 3N - m$. For a continuum the number of degrees of freedom is the continuous infinity, $\aleph_1$. The Lagrangian for a system of N function, for which the usual differential calculus can be used. When we deal with fields, the Lagrangian becomes a *functional* of the system, with infinitely many degrees of freedom. Replacing $q^i(t)$ by $\phi(t, \mathbf{x})$ we must replace $\dot{q}^i(t)$ by $\dot{\phi}(t, \mathbf{x})$. Notice that the dot refers to a total derivative and it not the partial derivative alone, $\dot{f} = df/dt = \partial f/\partial t + \dot{\mathbf{x}}.\nabla f$. The Lagrangian is then written as the functional $L[\phi(t, \mathbf{x}), \dot{\phi}(t, \mathbf{x})]$. The cardinality of the space of variables for the Lagrangian being higher, we can no longer use ordinary differential calculus. The differences between what is required and usual calculus are studied in *Functional Analysis*. I will not go into that here. However, to distinguish the *functional derivatives* from the ordinary derivative, I will follow the usual notation of replacing "∂" by "δ". For details the reader is referred to [16].

As before, we have taken it for granted that the Lagrangian has no explicit dependence on position or time. Its dependence only comes through the field, ϕ, and its time derivative, $\dot{\phi}$. Following the same procedure as before we arrive at the *EL equations*

$$\frac{\delta L}{\delta \phi} = \frac{d}{dt}(\frac{\delta L}{\delta \dot{\phi}}) \ . \tag{1.11}$$

It is worth mentioning, here, that the assumption that the Lagrangian has no explicit space or time dependence means that it is invariant under space and time translations. The conserved quantities given by Noether's theorem, in this case, are momentum and energy. Thus the above assumption is equivalent to the momentum and energy conservation laws!

1.5 Relativistic Fields

Field theory, as presented here up to now, does not accommodate Special Relativity. Time is given a special place. To be relativistic, a field theory must not refer to $\phi(t, \mathbf{x})$ and the equations must not end up with a d/dt operating on any quantity as Eq11 does. The field ϕ must be a function of the spacetime position vector, x^μ ($\mu = 0, 1, 2, 3$). Further, the only derivative available is the gradient of the field, $\phi_{,\mu} \equiv \partial\phi/\partial x^\mu$. Our Lagrangian will then have to be replaced by a *Lagrangian density*, $\mathcal{L}[\phi, \phi_{,\mu}]$. The formulation is completed by replacing Eq.(11) by

$$\delta S = 0 = \delta \int_a^b L dt = \delta \int_a^b \left(\int_V \mathcal{L} dV \right) dt = \delta \int_\Omega \mathcal{L} d\Omega \ . \tag{1.12}$$

Here $d\Omega$ is the "volume element" in spacetime and Ω is the total spacetime "volume" under consideration (see Figure 1.2). In Minkowski space, using Cartesian coordinates,

$$d\Omega = dx^0 dx^1 dx^2 dx^3 \ . \tag{1.13}$$

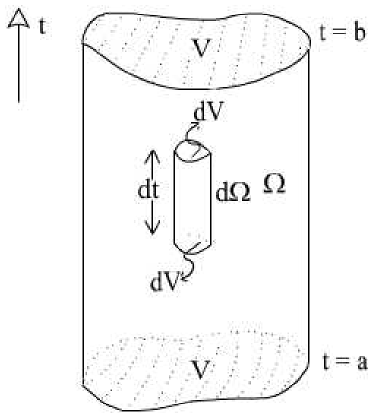

Figure 1.2: The "wotld-tube" represented by a 3-dimensional cylinder in 4-dimensional spacetime. The top and bottom "faces" are regions of 3-volume V at two times, $t = a$ and $t = b$. Inside this cylinder is a small 4-volume element.

A more general formulation will be presented later. Since that requires tensors and refers to concepts in curved spacetimes we will not go into it here. The three dimensional volume, V, traces, out a "world tube" in four dimensional spacetime.

We are now in a position to extend the EL equations to relativistic fields. We have

$$\delta\mathcal{L}[\phi, \phi_{,\mu}] = \frac{\delta\mathcal{L}}{\delta\phi}\delta\phi + \frac{\delta\mathcal{L}}{\delta\phi_{,\mu}}\delta\phi_{,\mu} \ . \tag{1.14}$$

We will need to integrate the second term by parts. Since this is not such a familiar procedure as before it needs to be elaborated a bit. For this purpose we write the second term as

$$\left.\begin{aligned}\frac{\delta\mathcal{L}}{\delta\phi_{,\mu}}\delta\phi_{,\mu} =& \frac{\delta\mathcal{L}}{}\delta\phi_{,0}\ \delta\left(\frac{\partial\phi}{\partial x^0}\right) + \frac{\partial\mathcal{L}}{\delta\phi_{,1}}\ \delta\left(\frac{\partial\phi}{\partial x^1}\right) + \frac{\partial\mathcal{L}}{\delta\phi_{,2}}\ \delta\left(\frac{\delta\mathcal{L}}{\partial x^2}\right) \\ &+ \frac{\delta\mathcal{L}}{\partial\phi_{,3}}\ \delta\left(\frac{\partial\phi}{\partial x^3}\right)\ .\end{aligned}\right\} \tag{1.15}$$

To evaluate the integral in Eq. (12), we need to integrate all four of the terms by parts, using the volume element given in Eq.(13). Let us just consider the first of the four terms. Integrating the term with respect to x^0 and using the fact that the partial derivative and the δ commute, the second expression can be treated as the function to be integrated and the first to be differentiated. Thus

$$\left.\begin{aligned}\int_\Omega \frac{\delta\mathcal{L}}{\delta\phi_{,0}}\delta(\frac{\partial\phi}{\partial x^0})(dx^3dx^2dx^1)dx^0 =& \int_V \frac{\delta\mathcal{L}}{\delta\phi_{,0}}(dx^3dx^2dx^1)\delta\phi|_a^b \\ &- (\frac{\delta\mathcal{L}}{\delta\phi_{,0}})_{,0}\delta\phi d\Omega\ ,\end{aligned}\right\} \tag{1.16}$$

where a is the initial time and b the final time. Since ϕ is fixed on the boundary, so $\delta\phi$ is zero at either end, and hence the first term vanishes. Doing the same for each of the other terms, it is obvious that Eq. (112) becomes

$$0 = \int_\Omega \left[\frac{\delta\mathcal{L}}{\delta\phi} - \left(\frac{\delta\mathcal{L}}{\delta\phi_{,\mu}}\right)_{,\mu}\right]\delta\phi d\Omega\ . \tag{1.17}$$

Since this integral is zero for arbitrary $\delta\phi$, the integrand must be zero. Hence we get the *EL equations*

$$\frac{\delta\mathcal{L}}{\delta\phi} = \left(\frac{\delta\mathcal{L}}{\delta\phi_{,\mu}}\right)_{,\mu}\ . \tag{1.18}$$

This formulation can be extended to a vector valued field, ϕ_r $(r = 1, ..., k)$. The r need not be a spacetime index. However special interest attaches to the case when the field is, itself, Lorentz covariant. In this case r will be a spacetime vector index, μ, or a tensor index like $\mu\nu$. Also, the field will be invariant under the full Poincarè group.

Another possibility is that r may be compounded of a spacetime index (or indices) and be invariant under some other symmetry. This is the case for the Yang-Mills field. To explain this I will first remind the reader of the Maxwell field. It is given by the four-vector potential, A_μ, and the Maxwell field tensor is the generalized curl of this field (as given in SR)

$$F_{\mu\nu} = A_{\nu,\mu} - A_{\mu,\nu}\ . \tag{1.19}$$

Under the gauge transformation

$$A_\mu \to \tilde{A}_\mu = A_\mu + f(x^\nu)_{,\mu}\ , \tag{1.20}$$

the field tensor remains invariant. Thus $F_{\mu\nu}$ is invariant under a further symmetry, which happens to be $U(1)$. This is called an *internal symmetry.* The corresponding Lagrangian density is

$$\mathcal{L} = \frac{1}{16\pi}(F_{\mu\nu}F^{\mu\nu} + j_\mu A^\mu)\ , \tag{1.21}$$