Chapter 1

The Limits of Physics

1.1 Our Scientific Legacy

From time immemorial, human beings have observed the universe around them and have tried to understand and explain the phenomena they saw. In the process they built models, that is described new events in terms of concepts they already knew. This has been an ongoing process till date.

The earliest known model builders of the universe were the composers of the hymns of the Rig Veda, some ten thousand years ago. With amazing insights, they described the Earth and sky as two bowls [1]. They went on to describe the Sun as a star of the daytime sky and even asked, how is it that though the Sun is not bound it does not fall down? And so on and so on. For the ancient Egyptians of four to five thousand years ago, the sky was supported, at the extreme ends by mountains.

Perhaps the earliest model of what we today call microphysics was proposed by the ancient Indian thinker Kanada who lived around the seventh century B.C. For him the universe was made up of ultimate sub constituents which were in perpetual vibration [2]. Later Greeks also had an Atomic Theory, which they may or may not have acquired from India. But there was a crucial difference. Their atoms were static.

Our legacy of modern science came from these Greeks who built up over a few centuries, an even more complex cosmic scheme in which the Earth was at the centre, surrounded by a series of transparent material spheres to which the various heavenly objects like the Sun, Moon, planets and stars were attached. The material spheres were necessary, for, otherwise they would have had to explain why the Moon doesn't crash down on to the Earth, for example. These were spheres because Plato had preached that the circle (or sphere) was a perfect object, due to its total symmetry.

Furthermore these spheres would be in rotation to explain the drama of heavenly motions.

As the observations became more and more precise, the above simple model, first put forward by Anaximenes around 500 B.C. needed modifications [3]. For instance the centre of a sphere would not coincide exactly with the Earth, but rather would be eccentric, that is, slightly away from it. Then the spheres themselves had to carry additional spheres called epicycles, themselves spinning and the objects were placed on top of the epicycles. Ptolemy the Librarian of Alexandria compiled all this knowledge in two astronomical treatises, only one of which, the Great Astronomer or Al Megast survived. The Ptolemaic universe was a complicated tangle of such spheres and epicycles, undergoing complex circular motions.

These basic ideas survived for nearly two thousand years, till the time of Kepler, in fact. Early in the seventeenth century, Kepler noticed that the Greek model differed from observation by just eight minutes of arc, for the orbit of Mars. Kepler had inherited the meticulous observations of Tycho Brahe, and a lesser mortal would have attributed this minor discrepancy to an error in observation. On the contrary, Kepler was convinced that the observations were correct and that the discrepancy pointed to a reformation of Astronomy. Clearly the limit of the validity of the Greek model had been reached.

Kepler proposed his first two laws of planetary motion around 1608. Some years later the third law followed. Crucially the orbits were ellipses. With a single ellipse Kepler could explain the minute discrepancy between theory and observation, for the planet Mars. What the Greeks had tried to do was, approximate a simple elliptical motion by a series of complicated circular motions. The larger implication of this minute correction was this: The ellipse destroyed the crystal spheres of the Greek model and the age old question was once again thrown open: Why dont the Moon, the planets and so on crash down?

This question was answered by Newton who needed the laws of mechanics which had been developed a little earlier by Galileo. He introduced his Theory of Gravitation. Kepler's purely observational laws could now be explained from theory.

Newtonian Mechanics dominated the scientific scene for a few centuries. There was an absolute space, while time was separate and reversible. The equations of mechanics were valid if the time t were replaced by $-t$. Another important concept implicit in Newtonian Mechanics and gravitation theory was action at a distance. Every object exerted instantaneously a gravitational force on every other object.

However, in the nineteenth century, a new discipline was born, which also had a new ethos−−rather than being an abstract study of the universe, this new discipline, Thermodynamics was a child of the industrial era. In the words of Toffler [4]:

"In the world model contributed by Newton and his followers, time was an after thought. A moment whether in the present, past, or future, was assumed to be exactly like any other moment...

"In the nineteenth century, however, as the main focus of physics shifted from dynamics to thermodynamics and the Second Law of Thermodynamics was proclaimed, time suddenly became a central concern. For, according to the Second Law, there is an inescapable loss of energy in the universe. And, if the world machine is really running down and approaching the heat death, then it follows that one moment is no longer exactly like the last. You cannot run the universe backward to make up for entropy. Events over the long term cannot replay themselves. And this means that there is a directionality or, as Eddington later called it, an "arrow" in time. The whole universe is, in fact, aging. And, in turn, if this is true, time is a one-way street. It is no longer reversible, but irreversible.

"In short, with the rise of thermodynamics, science split down the middle with respect to time. Worse yet, even those who saw time as irreversible soon also split into two camps. After all, as energy leaked out of the system, its ability to sustain organized structures weakened, and these, in turn, broke down into less organized, hence more random elements. But it is precisely organization that gives any system internal diversity. Hence, as entropy drained the system of energy, it also reduced the differences in it. Thus the second Law pointed toward an increasingly homogeneous−−and, from the human point of view, pessimistic−−future.

"... time makes its appearance with randomness: "Only when a system behaves in a sufficiently random way may the difference between past and future, and therefore irreversibility, enter its description." In classical or mechanistic science, events begin with "initial conditions," and their atoms or particles follow "world lines" or trajectories. These can be traced either backward into the past or forward into the future. This is just the opposite of certain chemical reactions, for example, in which two liquids poured into the same pot diffuse until the mixture is uniform or homogeneous. These liquids do not de-diffuse themselves. At each moment of time the mixture is different, the entire process is "time-oriented."

In a sense these "thermodynamic" ideas were anticipated in the nineteenth century itself, through the work of Poincare and others, working in the field of celestial mechanics rather than industrial machines. Were the orbits of the planets or other celestial objects really unchanging in time? Poincare realized that celestial mechanics had been worked out under the banner of what may be called the two body problem. The orbit of the earth round the Sun, for example, would be more or less unchanging, if the Earth and the Sun were the only two objects in the universe. Even with a third planet, we have to consider the three body problem, which as Poincare realized had no analytical solution. He had laid the ground for what has subsequently come to be known as the chaos theory. As Prigogine was to say much later [5]:

"Our physical world is no longer symbolized by the stable and periodic planetary motions that are at the heart of classical mechanics. It is a world of instabilities and fluctuations..."

Definitely the limits of Newtonian Mechanics had been reached.

The nineteenth century also saw the birth and development of yet another discipline, Electromagnetism. Now at this stage the earlier action at a distance concept had to be abandoned. Maxwell's work introduced the new paradigm of a field. Earlier an electric charge was conceived of as acting on another charge, via the Coulomb force, very much like Newton's gravitational force. This idea is correct, if the two charges are at relative rest, a situation which does not exist in the real world. When the charges move, more correctly accelerate, the interaction of one charge travels through the intervening medium, in the form of electromagnetic waves, which impinge upon the other charge at a later time, unlike the instantaneous action of Newtonian gravitation.

The stage had now been set for Einstein's Special Theory of Relativity. The Special Theory of Relativity introduced two ideas, one of which appeared to be self contradictory. Nevertheless these two ideas explained the puzzling and indeed otherwise inexplicable consequences of the Michelson-Morley and similar experiments. The point was that light had been thought of as electromagnetic radiation traveling with the same speed. In this case its speed would be different for different observers in relative motion. However the Michelson-Morley experiments showed that this was not so.

Einstein proposed at the turn of the twentieth century that the velocity of light would be the same in all directions——a relatively easy idea to digest. But then he also had a second postulate——the speed of light would be the same for all observers, moving with a uniform velocity with respect to one another. How could this be? It blatantly contradicted Newtonian Mechanics. Einstein could show that this was a contradiction if we retained the Newtonian concepts of space and time. If on the other hand, we realized that space and time get mixed up and that the lengths of the intervals of spacetime, which had been taken to be the same for all observers in Newtonian Mechanics, were on the contrary different for different observers, the contradiction would be removed. Clearly the limits of Newtonian Mechanics had been reached yet again.

When Einstein proposed his Special Theory of Relativity, there were two ruling paradigms, which continue to hold sway even today, though not so universally. The first was that of point elementary particles and the second was that of space time as a differentiable manifold. Further, Einstein's work introduced the concept of causality−−no signal could travel faster than light. So, the effect−−gravitational, electromagnetic, whatever−−of one object would be felt by another object at a later time, and not instantaneously, as in the earlier theory. That is, the signals would be retarded. Little wonder therefore that as the relativistic theory of the electron developed, there were immediate inconsistencies which were finally ostensibly resolved only with the intervention of Quantum Theory. This was because, historically the original concept of the electron was that of a spherical charge distribution [6–8]. It is interesting to note that in the non-relativistic case, it was originally shown that the entire inertial mass of the electron equalled its electromagnetic mass. The question came up, was this a meeting between electromagnetism and mechanics? This motivated much work and thought in this interesting direction. To put it briefly, in non relativistic theory, we get [6],

$$\text{Kinetic} \quad \text{energy} = (\beta/2)\frac{e^2}{Rc^2}v^2,$$

where R is the radius of the electron and β is a numerical factor of the order of 1. So we could possibly speak of the entire mass of the electron in terms of its electromagnetic properties.

It might be mentioned that it was possible to think of an electron as a charge distribution over a spherical shell within the relativistic context too, as long as the electron was at rest or was moving with a uniform velocity. However it was necessary to introduce, in addition to the electromagnetic force, the Poincare stresses−−these were required to counter balance the mutual repulsive "explosion" of the different parts of the electron.

When the electron in a field is accelerated, the above picture no longer holds. We have to introduce the concept of the electron self force which is given by, in the simple case of one dimensional motion,

$$F = \frac{2}{3}\frac{e}{Re^2}\ddot{x} - \frac{2}{3}\frac{d}{dt}\ddot{x} + \gamma\frac{e^2R}{c^4}\ddot{x} + 0(R^2) \qquad (1.1)$$

where dots denote derivatives with respect to time, and R as before is the radius of the spherical electron. More generally (1.1) becomes a vector equation. In (1.1), the first term on the right side gives the electromagnetic mass of the earlier theory. As can be seen from (1.1), as R the size of the electron $\to 0$ the first term $\to \infty$ and this is a major inconsistency. It was the first of a series of infinities that has plagued twentieth century physics. In contrast the second term which contains the non Newtonian third time derivative remains unaffected while the third and following terms $\to 0$. It may be mentioned that the first term (which $\to \infty$) gives the electromagnetic mass of the electron while the second term gives the well known Schott term (Cf.ref.[6, 7, 9]). Its presence is required however because it compensates the energy loss due to radiation by the accelerated electron. In any case it is possible to develop a model of an extended electron consistent with relativity on these lines, but at the expense of introducing non electromagnetic forces.

Let us now see how it was possible to rescue the relativistic electron theory, though at the expense of introducing some unphysical concepts.

1.2 The Advanced and Retarded Fields

To proceed, from a classical point of view a charge that is accelerating radiates energy which dampens its motion. This is given by the second term on the right side of (1.1). Dirac proposed in 1938 a phenomenological equation that overcomes the infinite (electromagnetic) mass in (1.1). The Lorentz Dirac equation, which in units $c = 1$, and τ being the proper time, while $\imath = 1, 2, 3, 4$, is (Cf.[10]),

$$m\frac{d^2x^\imath}{d\tau^2} = eF^\imath_k\frac{dx^k}{d\tau} + \frac{4e}{3}g_{lk}\left(\frac{d^3x^\imath}{d\tau^3}\frac{dx^l}{d\tau} - \frac{d^3x^l}{d\tau^3}\frac{dx^\imath}{d\tau}\right)\frac{dx^k}{d\tau}, \qquad (1.2)$$

This holds for a point charge, m being a "renormalized mass" that absorbs the infinity. Here is the precursor of renormalization, that has gone hand in hand with the infinities of twentieth century physics. The first term gives the usual external field while the second term does not come from the Lagrangian (which gives the first term and the Lorentz force)−−it comes by putting in energy conservation (due to radiation loss) by hand. Equation

(1.1) can be written as

$$m\frac{d^2x^\imath}{d\tau^2} = e\{F^\imath_k + R^\imath_k\}\frac{dx^k}{d\tau} \qquad (1.3)$$

where

$$R^\imath_k \equiv \frac{1}{2}\{F^\imath_{k(ret)} - F^\imath_{k(adv)}\} \qquad (1.4)$$

In (1.4), $F_{(ret)}$ denotes the retarded or causal field allowed by relativity, as alluded to. $F_{(adv)}$ on the other hand is the advanced field that is unphysical, in the sense that it is not sanctioned by relativity. While the former is the causal field where the influence of a charge at A is felt by a charge at B at a distance r after a time $t = \frac{r}{c}$, the latter is the advanced field which acts on A from a future time. In effect what Dirac showed was that the radiation damping term in (1.2) or (1.3) is given by (1.4) in which an antisymmetric difference of the advanced and retarded fields is taken. Let us elaborate a little further.

The Maxwell wave equation has two independent solutions, one having support on the future light cone, this is the retarded solution and the other having support on the past light cone which has been called the advanced solution. The retarded solution is selected to describe the physical situation in conventional theory taking into account the usual special relativistic concept of causality. This retarded solution is physically meaningful, as it describes electromagnetic radiation which travels outward from a given charge with the speed of light and reaches another point at a later instant. It has also been called for this reason the causal solution. On the grounds of this causality, the advanced solution has been rejected, except in a few formulations like those of Dirac above, or Feynman and Wheeler (F-W) to be seen below.

It must also be mentioned that Dirac's prescription lead to the so called runaway solutions, with the electron acquiring larger and larger velocities in the absence of an external force [11]. This he related to the infinite self energy of the point electron.

To elaborate further, we use the difference of the advanced and retarded fields in (1.1), in the following manner: We use successively $F_{(ret)}$ and $F_{(adv)}$ in (1.1) and take the difference in which case the self force becomes (Cf.[9])

$$F = -\frac{2}{3}\frac{e^2}{c^3}\frac{d}{dt}(\ddot{x}) + 0(R)$$

In the above, the troublesome infinity generating term of (1.1) is absent, while the third derivative term is retained. On the other hand this term is required on grounds of conservation of energy, due to the fact that an accelerated electron radiates energy (Cf.[12]). Except for the introduction of advanced fields, we have infinity free results. However, in this formulation too, there is no electromagnetic mass term, and further, as will be seen below, we have to extend our considerations to a small neighborhood of the electron, and not just the point electron itself. To see this in detail, we observe that the well known Lorentz-Dirac equation (Cf.[6]), can be written as

$$ma^\mu(\tau) = \int_0^\infty K^\mu(\tau + \alpha\tau_0)e^{-\alpha}d\alpha \tag{1.5}$$

where a^μ is the acceleration and

$$K^\mu(\tau) = F^\mu_{in} + F^\mu_{ext} - \frac{1}{c^2}\bar{R}v^\mu,$$

$$\tau_0 \equiv \frac{2}{3}\frac{e^2}{mc^3} \sim 10^{-23} sec \tag{1.6}$$

and

$$\alpha = \frac{\tau' - \tau}{\tau_0},$$

where τ denotes the time and $\bar{R}$ is the total radiation rate.

It can be seen that equation (1.5) differs from the usual equation of Newtonian Mechanics, in that it is non local in time. That is, the acceleration $a^\mu(\tau)$ depends on the force not only at time τ, but at subsequent times also. Let us now try to characterize this non locality. We observe that τ_0 given by equation (1.6) is the Compton time $\sim 10^{-23} secs$. This is the precursor of Quantum Theory. So equation (1.5) can be approximated by

$$ma^\mu(\tau) = K^\mu(\tau + \xi\tau_0) \approx K^\mu(\tau) + \xi\tau_0\dot{K}^\mu(\tau) + \cdots \tag{1.7}$$

Thus as can be seen from (1.7), the Lorentz-Dirac equation differs from the usual local theory by a term of the order of

$$\frac{2}{3}\frac{e^2}{c^3}\dot{a}^\mu \tag{1.8}$$

the so called Schott term. It is well known that the time component of the Schott term (1.8) is given by (Cf.ref.[6])

$$-\frac{dE}{dt} \approx \bar{R} \approx \frac{2}{3}\frac{e^2c}{r^2}\left(\frac{E}{mc^2}\right)^4,$$

where E is the energy of the particle. Whence integrating over the period of non locality $\sim \tau_0$ the Compton time, we can immediately deduce that r the scale of spatial non locality is given by

$$r \sim c\tau_0,$$

which is of the order of the Compton wavelength as indeed can be expected. So far as the breakdown of causality is concerned, this takes place within a period $\sim \tau$, the Compton time as we briefly saw [6, 11].
In the F-W formulation on the other hand, the rest of the charges in the universe react back on the original electron through their advanced waves, which arrive (from the future) at the given charge at the same time as the given charge radiates its electromagnetic waves. More specifically, when an electron is accelerated at the instant t, it interacts with the other charges at a later time $t' = t + r/c$ where r is the distance of the other charge$--$these are the retarded interactions. However the other charges react back on the original electron through their advanced waves, which will arrive at the time $t' - r/c = t$. Effectively, there is instantaneous action at a distance. It must be mentioned that in the F-W formulation there is no self force (and therefore the electromagnetic mass and the infinite term$--$the first term on the right side of (1.1)) or radiation damping. This is provided instead by the action of all other charges in the universe on the original charge.
Let us throw further light on all this. There are two important inputs which we can see in the above formulation. The first is the action of the rest of the universe at a given charge and the other is spacetime intervals which are of the order of the Compton scale. In fact we can push the above calculations further. The work done on a charge e at O by the charge P at a distance r in causing a displacement $x \sim l$ is given by

$$\frac{e^2 l}{r^2}$$

Now the number of particles at distance r from O is given by

$$n(r) = \rho(r) \cdot 4\pi r^2 dr$$

where $\rho(r)$ is the density of particles. So the total work is given by

$$E = \int \int \frac{e^2}{r^2} l 4\pi r^2$$

which can be shown to be $\sim mc^2$. This is because,

$$\rho(r) = N/R^3,$$

where N is the total number of particles in the universe, R now is its radius and anticipating a result from Chapters 2 and 3,

$$R \sim \sqrt{N} l,$$

where l is given by (1.6).

Wheeler and Feynman thus reformulated the above action at a distance formalism in terms of what has been called their Absorber Theory. In their formulation, the field that a charge would experience because of its action at a distance on the other charges of the universe, which in turn would act back on the original charge is given by

$$\bar{R}e = \frac{2e^2 d}{3dt}(\ddot{x}) \tag{1.9}$$

The interesting point is that instead of considering the above force in (1.9) at the charge e, if we consider the response at an arbitrary point in its neighborhood as was shown by Feynman and Wheeler (Cf.ref.[13]) and, in fact a neighborhood at the Compton scale, as we saw above and was argued by the author [14], the field would be precisely the Dirac field given in (1.3) and (1.4).

The net force emanating from the charge is thus given by

$$F^{ret} = \frac{1}{2}\left\{F^{ret} + F^{adv}\right\} + \frac{1}{2}\left\{F^{ret} - F^{adv}\right\} \tag{1.10}$$

which is the acceptable causal retarded field. The causal field now consists of the time symmetric field which implies no radiation of the charge together with the Dirac field, that is the second term in (1.10), which now represents the response of the rest of the charges. Interestingly in this formulation we have used a time symmetric field, viz., the first term of (1.10) to recover the retarded field with the correct arrow of time.

Feynman and Wheeler stressed that the universe has to be a perfect absorber or to put it simply, every charged particle in the universe should respond back to the action on it by the given charge in our instantaneous action at a distance scenario. In the Feynman-Wheeler formulation to reiterate there is no electromagnetic mass and also no radiation damping−−we have finally the retarded field; but within the context of the Instantaneous Action at a Distance. In any case, it was realized that the limits of classical physics are reached in the above considerations, at the Compton scale. However, we will now argue that there is actually a convergence between Classical and Quantum Physics.

There are two important inputs which we would like to re-emphasize in the above more recent formulation. The first is the action of the rest of the universe at a given charge and the other is minimum spacetime intervals which are of the order of the Compton scale. The minimum spacetime interval removes, firstly the advanced field effects which take place within the Compton time and secondly the infinite self energy of the point electron disappears due to the Compton scale. We thus bypass renormalization. This would be an important idea in the rest of the book.

1.3 Quantum Mechanical Considerations

The Compton scale comes as a Quantum Mechanical effect, within which we have zitterbewegung effects and a breakdown of causal physics [15]. Indeed Dirac had noted this aspect in connection with two difficulties with his electron equation. Firstly the speed of the relativistic Quantum Mechanical electron turns out to be the velocity of light. Strictly speaking, this would imply an infinite mass for the electron. Secondly the position coordinates become complex or non Hermitian. This is physically meaningless. His explanation was that in Quantum Theory we cannot go down to arbitrarily small spacetime intervals, for the Heisenberg Uncertainty Principle would then imply arbitrarily large momenta and energies. So Quantum Mechanical measurements are an average over intervals of the order of the Compton scale. Once this is done, we recover meaningful physics. All this has been studied afresh by the author more recently, in the context of a fuzzy non differentiable spacetime and noncommutative geometry [16]. This indeed, will be the theme of this book. We will first argue that there is a convergence between preceding considerations and Quantum Mechanical Theory. The Compton scale that surfaces in both these considerations, already gives a hint of this.

Weinberg too notices the non physical aspect of the Compton scale [17]. Starting with the usual light cone of Special Relativity and the inversion of the time order of events, he goes on to add, “Although the relativity of temporal order raises no problems for classical physics, it plays a profound role in quantum theories. The uncertainty principle tells us that when we specify that a particle is at position x_1 at time t_1, we cannot also define its velocity precisely. In consequence there is a certain chance of a particle getting from x_1 to x_2 even if $x_1 - x_2$ is space-like, that is, $|x_1 - x_2| > |x_1^0 - x_2^0|$. To be more precise, the probability of a particle reaching x_2 if it starts at

x_1 is nonnegligible as long as

$$0 \leq (x_1 - x_2)^2 - (x_1^0 - x_2^0)^2 \leq \frac{\hbar^2}{m^2} \cdots \tag{1.11}$$

where $\hbar$ is Planck's constant (divided by 2π) and m is the particle mass. (Such space-time intervals are very small even for elementary particle masses; for instance, if m is the mass of a proton then $\hbar/m = 2 \times 10^{-14} cm$ or in time units $6 \times 10^{-25} sec$. Recall that in our units $1 sec = 3 \times 10^{10} cm$.) We are thus faced again with our paradox; if one observer sees a particle emitted at x_1, and absorbed at x_2, and if $(x_1 - x_2)^2 - (x_1^0 - x_2^0)^2$ is positive (but less than or $= \hbar^2/m^2$), then a second observer may see the particle absorbed at x_2 at a time t_2 before the time t_1 it is emitted at x_1.

"There is only one known way out of this paradox. The second observer must see a particle emitted at x_2 and absorbed at x_1. But in general the particle seen by the second observer will then necessarily be different from that seen by the first."

There is another way to view (1.11). The light cone of special relativity viz., $(x_1 - x_2)^2 - (x_1^0 - x_2^0)^2 = 0$ now gets somewhat distorted because of Quantum Mechanical effects.

Let us now consider the above in the context of a non zero photon mass. Such a mass $\sim 10^{-65} gms$ was rather recently deduced by the author, and it is not only consistent with experimental restrictions, but also predicts a new effect viz., a residual cosmic radiation $\sim 10^{-33} eV$, which in fact has been observed [18–22]. We will come back to this in detail in later Chapters, particularly Chapter 4. Such a photon would have a Compton length $\sim 10^{28} cms$, that is the radius of the universe itself.

This would then lead to the following scenario: An observer would see a photon leaving a particle A and then reaching another particle B, while a different observer would see exactly the opposite for the same event−−that is a photon leaves B and travels "backward" in time to A, as in the Weinberg interpretation. This latter gives the advanced potential. We are back with the Feynman-Wheeler instantaneous action scenario. The distinction between the advanced and retarded potentials of the old electromagnetic theory thus gets mixed up and we have to consider both the advanced and retarded potentials [13]. Thus, two charged particles interacting via the exchange of photons will be described as above, using (1.11). Indeed in Quantum Field Theory this is described as the exchange of virtual photons.

We consider this in a little more detail: The advanced and retarded solutions of the wave equation are given by the well known advanced and retarded potentials given by, in the usual notation, the well known expression

$$A^{\mu}_{ret(adv)}(x) = \frac{1}{c}\int \frac{j^{\mu}(x')}{|r-r'|}\delta\left(|r-r'| \mp c(t-t')\right) d^4x'$$

(The retarded part of which leads to the Lienard Wiechart potential of earlier theory).

It can be seen in the above that we have the situation described within the Compton wavelength, wherein there are two equivalent descriptions of the same event−−a photon leaving the charge A and reaching the charge B or the photon leaving the charge B and reaching the charge A. The above expression for the advanced and retarded potentials immediately leads to the advanced and retarded fields (1.4) and (1.10) of the F-W description except that we now have a rationale for this formulation in terms of the photon mass and the photon compton wavelength rather than the perfect absorber ad hoc prescription. In fact there is now an immediate Quantum Mechanical explanation in this of the Instantaneous Action At a Distance Theory alluded to. Thus these considerations reconcile the Quantum Mechanical and Classical pictures. We note however that as the photon mass is so small, the usual theory is still a good approximation.

To sum up [13], the Feynman Wheeler Perfect Absorber Theory required that every charge should interact instantaneously with every other charge in the universe, that is that the universe must be a perfect absorber of all electromagnetic fields emanating from within. If this condition were satisfied, then the nett response of all charged particles along the future light cone of the given charge is expressed by an integral that converges. We have argued that this ad hoc prescription of Feynman and Wheeler as embodied by the inclusion of the advanced potential is automatically satisfied if we consider the photon to have a small mass $10^{-65} gms$ which is consistent with the latest experimental limits−−this leading to the effect mentioned by Weinberg within the Compton wavelength, which is really the inclusion of the advanced field as well.

To put the above in different words, when we talk of two (charged) particles A and B and the instant t, we are attributing the same t to A and the distant B. This is consistent with Special Relativity. This enables us to talk of an advanced wave leaving B at $t + \Delta t$ and travelling "backward" in time to reach A at t. This simultaneity however, breaks down within the Quantum Mechanical Compton time. We could very well describe the event as an ordinary retarded wave leaving B at $t - \Delta t$ and reaching A at t.

1.4 The Limits of Special Relativity

What we have witnessed above is that it is still possible to rescue the classical relativistic theory of the electron, but at the expense of introducing the advanced fields into the physics, fields which have been considered to be unphysical.

Another perspective is, as seen above, that there is instantaneous action at a distance, which apparently goes against relativistic causality. But let us now note that in both the Dirac and the Feynman-Wheeler approaches, we are no longer dealing with point particles alone, but rather with a small neighborhood of such a point particle, a neighborhood of a Compton length dimension. Furthermore within the Compton scale, relativistic causality breaks down as embodied in (1.11).

We can then reformulate the above considerations in the following manner: The limit of applicability or the limit of validity of the relativistic electron theory as also the Special Theory of Relativity is the Compton scale of a particle. The points within the Compton scale no longer obey Special Relativity and see a non relativistic, instantaneous action at a distance universe. Indeed Rohrlich notes [23], "... the notion of a "classical point charge" is an oxymoron because "classical" and "point" contradict one another: Classical physics ceases to be valid at sizes at or below the Compton wavelength and thus cannot possibly be valid for a point object..."

1.5 Discussion

Let us sum up the foregoing considerations. In Classical Physics the point electron leads to infinite self energy via the electromagnetic mass term e^2/R, where R is the radius which is made to tend to zero. If on the other hand R does not vanish, in other words we have an extended electron, then we have to introduce non electromagnetic forces like the Poincare stresses for the stability of this extended object, though on the positive side this allows the radiation damping or self force that is required by conservation laws.

Dirac could get rid of these problems by introducing the difference between the advanced and retarded potentials in his phenemenological equation in which the infinity was absorbed into a renormalized point particle mass: This was the forerunner of the renormalization theory and was the content of the Lorentz-Dirac equation. The new term represents the radiation damping effect, but we then have to contend with the advanced potential or equivalently a non locality in time. However this non locality takes place within the Compton time, within which the electron attains a luminal velocity.

The Lorentz-Dirac equation also had unsatisfactory features like the non-Newtonian derivative of the acceleration, the non locality in time and the run away solutions, features confined to the Compton scale.

The Feynman-Wheeler approach bypasses the infinity and the extended electron self force−−but the mass is no longer electromagnetic. Moreover the nett result is that there is only the desired retarded potential, but an instantaneous interaction with the rest of the charges of the universe has to be invoked. It is this interaction with the remaining charges which leads to the point electron's self energy. Surprisingly however the interaction with the rest of the charges in the immediate vicinity of the given charge in the Feynman-Wheeler formula gives us back the Dirac antisymmetric difference with its non locality within the Compton scale. There is thus a reconciliation of the Dirac and the Feynman Wheeler approaches, once we bring into the picture, the Compton scale.

Outside this scale, however, the theory is causal that is uses only the retarded potential because effectively the advanced potential gets canceled out as it appears as the sum of the symmetric and antisymmetric differences.

The final conclusion was that in a classical context a totally electromagnetic electron is impossible as also the concept of a point electron without introducing additional "unphysical" concepts including action at a distance. It was believed therefore that the electron was strictly speaking the subject of Quantum Theory.

Nevertheless in Dirac's relativistic Quantum Electron, we again encounter the electron with the luminal velocity within the Compton scale, precisely what was encountered in Classical Theory as well, as noted above. This again is the feature of a point space time approach. At this stage a new input was given by Dirac−−meaningful physics required averages over the Compton scale, in which process, the unphysical zitterbewegung effects were eliminated. Nor has Quantum Field Theory solved the problem−−one has to take recourse to renormalization, and as pointed out by Rohrlich, one still has a non electromagnetic electron. In any case, it appears that further progress would come either from giving up point spacetime or from an electron that is extended (or has a sub structure) in some sense [8, 7, 11, 6]. From this point of view the relativistic theory of the electron is inconclusive to date. As noted by Feynman himself in his famous Lectures on Physics

(Vol II), "We do not know how to make a consistent theory−−including the Quantum Mechanics−−which does not produce an infinity for the self energy of the electron, or any point charge. And at the same time there is no satisfactory theory that describes a non-point charge..."
In the words of Hoyle and Narlikar [10], "...it was believed that the problem of the self force of the charge would not be solved except by recourse to Quantum Theory... This hope has not been fully realized. Quantum Field Theory does alleviate the self energy problem but cannot surmount it without introducing the renormalization programme..."
Indeed Dirac himself was unhappy with renormalization, which he termed an accident. He expressed his confidence that it would be disproved eventually.
In his words, "I am inclined to suspect that the renormalization theory is something that will not survive in the future, and that the remarkable agreement between its results and experiments should be looked on as a fluke..."
We have pointed out that the important point however is that all this can be explained consistently in Quantum Mechanical terms in the context of the photon having a non zero mass, consistent with experiment $\sim 10^{-65} gms.$
So there is convergence between the Dirac and the Feynman-Wheeler approaches if we consider the fact that special relativity, as seen above, does not hold within the Compton wavelength. This explains the non locality in time. This justifies the use of the advanced potential or non locality in time of the Lorentz-Dirac approach or also the fact that a point inside the Compton wavelength sees a non relativistic instantaneous action at a distance universe around it−−this is the instantaneous action at a distance of the Feynman-Wheeler approach. Furthermore, the radiation of photons emitted by the accelerated electrons (in the Dirac self force) are meaningful only if they impinge on other charges as in the Field Theory.
We now briefly re-emphasize the following.
1. In classical relativistic theory, there appeared an impasse. We could get a special relativistic electron with cohesive forces in an extended model but at the expense of the purely electromagnetic electron. On the other hand point electrons were not meaningful as their self energy diverged. Consequently the structure dependent terms for example in (1.1) had to be taken seriously.
2. We have arrived at the Compton scale from two different approaches. Classically, there was the electron radius and Quantum Mechanically the Compton length, both of the same order except for a factor of the order of

the fine structure constant:

$$\hbar/mc \sim \beta \cdot e^2/mc^2$$

The left side has the Quantum Mechanical Planck constant while the right side has merely classical quantities. We could consider this to be a derivation of the rough value of the Planck constant of Quantum Mechanics, in an order of magnitude sense. We will return to this point in a later Chapter.
3. In any case the above considerations at the Compton scale lead in recent studies to a noncommutative geometry and the limit to a point particle no longer becomes legitimate. This will be extensively discussed in this book.

1.6 The Quantum Universe

The advent of Quantum Mechanics threw up several, what may be called counter intuitive ideas and even Einstein could not reconcile to them. One of these ideas was the wave particle duality. Another was Heisenberg's Uncertainty Principle: surprisingly it would not be possible to measure simultaneously and accurately the position and momentum of a particle. This was related to wave particle duality itself. Yet another was that of the collapse of the wave function in which process causality becomes a casuality. To put it simply, if the wave function is a super position of the eigen states of an observable, then a measurement of the observable yields one of the eigen values no doubt, but it is not possible to predict which one. Due to the act of observation, the wave function instantly collapses to any one of its eigen states in an acausal manner. To put it another way, the wave function obeys the causal Schrodinger equation, for example, till the instant of observation at which point, causality ceases. Indeed, we saw that this was true within the Compton scale itself.

Another important counter intuitive feature of Quantum Mechanics is that of non locality. In fact Einstein with Podolsky and Rosen put forward in 1935 his arguments for the incompleteness of Quantum Mechanics on this score [24, 25]. This has later come to be known as the EPR paradox. To put it in a simple way, without sacrificing the essential concepts, let us consider two elementary particles, for example two protons kept together somehow. They are then released and move in opposite directions. When the first proton reaches the point A its momentum is measured and turns out to be say, $\vec{p}$. At that instant we can immediately conclude, without any further measurement that the momentum of the second proton which is at the point B is $-\vec{p}$. This follows from the Conservation of Linear Momentum, and is perfectly acceptable in Classical Physics, in which the particles possess a definite momentum at each instant.

In Quantum Physics, the difficulty is that we cannot know the momentum at B until and after a measurement is actually performed, and then that value of the momentum is unpredictable. What the above experiment demonstrates is that the proton at B instantly came to have the value $-\vec{p}$ for its momentum without any further measurement, when the momentum of the proton at A was measured. This "instant" or "spooky action at a distance" feature was unacceptable to Einstein.

In Quantum Theory however this is legitimate because of another counter intuitive feature which is called Quantum Non-separability. That is, if two systems interact and then separate to a distance, they still have a common state vector. This goes against the concept of locality and causality, because it implies instantaneous interaction between distant systems. So in the above example, even though the protons at A and B may be separated, they still have a common wave function which collapses to some value with the measurement of the momentum of any one of them and self-consistently provides an explanation of the fact that the momentum of the other particle is automatically known without requiring another measurement. This non-separability has been characterized by Schrodinger in the following way: "I would not call that *one*, but rather *the* characteristic of Quantum Mechanics." For Einstein however this was like spooky action at a distance. All this has been experimentally verified since 1980 which sets at rest Einstein's objections.

However this "entanglement" as it is called these days, between distant objects in the universe, does not really manifest itself though it is perfectly legitimate and observable in a universe that consists of let us say just two particles. But a measurement destroys the entanglement. Now in the universe at large as there are so many particles and correspondingly a huge amount of interference, the entanglement is considerably weakened. This was the crux of Schrodinger's arguments. What is these days called decoherence works along these lines. This is in fact the explanation for the famous "Schrodinger's Cat" paradox.

This paradox can be explained in the following simple terms: A cat is in an enclosure along with, let us say a microscopic amount of radioactive material. If this material decays, emitting let us say an electron, the electron would fall on a vial of cyanide, releasing it and killing the cat in the process. Let us say that there is a certain probability of such an electron being emitted. So there is the same probability for the cat to be killed. There is also a probability that the electron is not emitted, so that there is the same probability for the cat to remain alive. The cat is therefore in a state which is a superposition of the alive and dead states. It is only when an observer makes an observation that this superposed wave function collapses into either the dead cat state or the alive and kicking cat state, and this happening is acausal. So it is only on an observation being made that the cat is killed or saved, and that too in an unpredictable manner. Till the observation is made the cat is described by the superposed wave function and is thus neither alive nor dead.

The resolution of this paradox−−it is a paradox−−is of course quite simple. The paradox is valid if the system consists of such few particles and at such distances that they do not interact with each other. Clearly in the real world this idealization is not possible. There are far too many particles and interferences taking place all the time and the superposed wave function would have collapsed almost instantly. This role of the environment has come to be called de-coherence. We will return to this point shortly.

The important point is that all of Classical and Quantum Physics is based on such idealized laws as if there were no interferences present, that is what we have called a two body scenario, is implicit. Clearly this is not a real life scenario.

1.7 The Strong and Weak Interactions

A major achievement of the twentieth century has been the incorporation of three of the four fundamental interactions, viz., electromagnetism, weak interactions and the strong interaction within a unified mathematical framework. This framework is the non Abelian gauge field theory [26–31]. Though the three forces remain different, the underlying mechanism is the same. From this point of view they could be thought to be different aspects of a single underlying process.

Thus there are leptons and there are quarks. The difference between these sets of particles which are perceived today arise because the Universe has become cold. At sufficiently high energies $\sim 10^{15} GeV$, leptons and quarks would be interchangeable and so also all the three forces would have the same strength. It must be mentioned that the above energy is still beyond the reach of foreseeable accelerators.

Apart from leptons and quarks, which are Fermions, or "material" particles, the fields are mediated by Bosons. These are the photons for electromagnetism, the W and Z Bosons for weak interactions and the gluons for the strong interactions.

Quarks were conceived following the work of Gellmann, Ne’eman and Zweig in the sixties. The motivation had been the overabundance of resonances observed in hadron or strong interaction collisions. These resonances could be classified on the one hand according to the Regge trajectories that plot the angular momentum J versus the mass squared, M^2 [32]. We will touch upon this briefly again. On the other hand, there was the SU(3) classification scheme which related particles of the same spin but different quantum numbers by introducing elemental entities−−the quarks−−whose combinations could account for all observed hadrons.

It is now believed that there are six kinds of quarks: The down (d), the up (u), the strange (s), the charmed (c), the bottom (b) and the top (t). We attribute to the quarks three colours, red, green and blue which are generalizations of the positive and negative charges. It is these colours which characterize strong interaction and hence this field has come to be known as Quantum Chromo Dynamics (QCD). It may be observed that the leptons do not have any colour and so they do not participate in the strong interactions.

A peculiarity of quarks is their fractional charge−−they have either the charge $\frac{1}{3}$ or the charge $\frac{2}{3}$ with their corresponding anti particles having opposite charges. So quarks can combine in two different ways to form hadrons, that is particles like protons and neutrons: Either as quark antiquark pairs or as a triplet of quarks, such that the total charge is either one or zero.

In electromagnetism, or Quantum Electro Dynamics (QED), two charged particles interact by the exchange of a photon, more correctly a virtual photon [33] as noted earlier. This exchange takes place within the Heisenberg Uncertainty time. There is a conservation of electric charge in the process. This combined with the masslessness of the photon is characteristic of the $U(1)$ Group which characterizes QED.

QCD is modelled on QED. However QCD which is described by the SU(3) group is more complicated because it describes interactions of three different colours, unlike QED which deals with just one charge. In QCD the interaction between different colours is expressed in terms of eight massless particles, the gluons, unlike the single photon of QED. Another profound difference is that the gluons do carry colour unlike the photon which is chargeless. The nett result of all this is that there is an effect opposite to that encountered in the charge screening of QED. In this latter case, an electron is surrounded by virtual electron-positron pairs. The electron attracts the positrons and repels the electrons of these pairs with the result that at large enough distances, the electron charge is shielded by the positrons and so appears reduced. In QCD on the other hand, virtual gluon pairs, themselves carrying colour are formed around a quark, no doubt. But there is now an anti screening effect as if the red component of a gluon is attracted to the red of a quark, for example. So at relatively larger distances, the colour charge of a quark increases and again contrary to the QED scenario, decreases as we approach the quark. The QCD force can therefore be compared to rubber bands−−as we stretch, the elastic force manifests itself, but if the bands slacken at close range, the force decreases and even disappears. It is as if there is confinement at large distances and freedom at shorter, asymptotic distances.

The QCD potential can be written as [30, 34]

$$V(r) = -\frac{\alpha(r)}{r} + \frac{r}{\beta^2}$$

This consists of the Coulombic part $\propto -\frac{1}{r}$ and a confining part $\propto r$. Because of this latter, which dominates for large r, free quarks cannot be observed in nature. After all, the model should explain this fact! On the other hand, the Coulombic part ensures that for small r, the inter quark force vanishes, a circumstance which is called asymptotic freedom. Professors Wilczek, Politzer and Gross were awarded the 2004 Nobel Prize in Physics for this work, done thirty years earlier.

The neutrinos are closely associated with the weak interactions. Though the neutrinos are leptons, they differ from their counterparts in that they are massless (or more precisely, as later discovered, they have a very tiny mass). A massless Fermion exhibits handedness, that is, its spin is either aligned in the direction of its motion (righthanded) or it is aligned anti parallel to its motion (lefthanded). This extra property of handedness characterizes the weak force which violates parity, unlike the other forces (though even the quarks exhibit handedness!). Only lefthanded particles and righthanded anti particles bear a weak charge while the righthanded particles and the lefthanded anti particles are neutral from the point of view of the weak interaction. This interaction acts on doublets of particles, which latter are described by the SU(2) Group, in which particles of a doublet pair can be transformed into one another. The weak interactions are mediated by the W Bosons. However a suitable mixture yields both the photon of electromagnetism and the Z° characterizing weak interactions.

This theory therefore combines electromagnetic and weak interactions and is incorporated in the SU(2) XU(1) group [35].

An important difference between the weak forces on the one hand and QED and QCD on the other is that the intermediate particles of the weak interactions, the W and Z Bosons are not massless, but rather have large masses $\sim 100 GeV$. This is characteristic of the fact that the weak charge is not invariably conserved and moreover has an extremely short range $\sim 10^{-15} cms$. We will return to this point later.

One of the problems that has plagued modern field theories is that of infinities. Indeed as we saw, this problem was encountered early in the twentieth century itself when an attempt was made to model the electron as a tiny sphere. If the radius of the sphere was then made to shrink indefinitely, the energy of the electron increased without limit as noted [6]. In QED for instance, if we approach the electron through the shield of screening positrons, the bare charge of the electron would be infinite. It is only the physically observable charge, at a distance, screened by the positron charges, which is finite. It is as if the infinite bare negative charge has been cancelled or neutralized by the infinite screening positive charge, the nett result being the observed finite physical charge. Loosely speaking this procedure is called "renormalization".

Mathematically, we encounter divergent integrals [36]. The infinities are eliminated in two steps. In the first step, called regularization, we introduce constraints, for example a cut off (or a lattice structure), to get a finite result dependent on the regularization parameter like the cut off. Counter terms (dependent on these parameters) are then added to the Lagrangian, such that they cancel the parameter dependent integrals. This generally leads to a rescaling of the mass, charge etc. This is the process that is called Renormalization.

The concept of Renormalization is unsatisfactory from the logical point of view as well as from the point of view of internal consistency. It has provoked unease among Physicists such as Dirac quoted earlier [37] or as we will see 't Hooft and several others. Its merit however, has been that phenomenologically speaking, it works.

1.8 Gauge Fields

It has now come to be recognized that the physical principle governing the fundamental interactions between the elementary particles is gauge invariance. This principle, as we shall see in greater detail later, was originally introduced by Hermann Weyl, though in a different form and with a different motivation viz., the attempt to give a unified General Relativistic description of Electromagnetism and Gravitation [38]. At that time these were the only two known interactions and electrons and protons were the only known elementary particles. Weyl's original theory was soon dismissed as adhoc. But nevertheless it was recognized that gauge invariance was a symmetry of Maxwell's equations with useful implications.

Then in the 1950s Yang and Mills (and Shaw) tried to extend gauge symmetry to other interactions. It must be emphasized that both in Special Relativity and General Relativity there are no absolute frames of reference in the Universe. The physics within a system is independent of the choice of the reference frame. However in Special Relativity this freedom of choice of reference frame is a global symmetry- the Lorentz symmetry. In General Relativity on the other hand, the reference frame is to be defined locally, that is at each and every point in the gravitational field. There are the connections−−the affine connections or Christofell symbols which relate nearby frames in General Relativity, something which is not required in Special Relativity [39].

Weyl attempted to investigate if there were similar connections associated with Electromagnetism [38]. Just as in General Relativity, all physical measurements are relative, so also could the norm of a physical vector depend on its location? If so, a new connection would be required to relate the lengths of the vectors at different positions. This clearly would be a local property. It was called Gauge Invariance. Let us see how this can be expressed mathematically [28, 40]. In essence we have to multiply the norm of a vector $f^\mu(x^\mu) \equiv f(x)$ at $x \equiv x^\mu$ by a scale factor $S(x^\mu) \equiv S(x)$, which latter would represent the change in scale from point to point. So we have for a small displacement to the point $x + dx$, the equations

$$S(x + dx) = 1 + \partial_\mu S dx^\mu$$

$$Sf = f + (\partial_\mu S) f dx^\mu + \partial_\mu f dx^\mu$$

If f is a constant vector, then we have on the right

$$(1 + \partial_\mu S) f dx^\mu$$

As can be seen from the above, the derivative $\partial_\mu S$ is the new mathematical connection associated with the gauge transformation. Weyl identified this connection with the electromagnetic potential A_μ. This is motivated by the fact that a second gauge change with a scale factor Λ leads to

$$\partial_\mu S \to \partial_\mu S + \partial_\mu \Lambda$$

which mimics the behavior under a gauge transformation of the electromagnetic potential in classical theory,

$$A_\mu \to A_\mu + \partial_\mu \Lambda$$

With the advent of Quantum Theory, Weyl himself realized that his old idea could be given a new interpretation. Rather than being a change of scale, a gauge transformation could be interpreted as a phase transformation. This is because if

$$\psi \to \psi e^{-\imath\lambda} \tag{1.12}$$

then for the electromagnetic potential we would have

$$A_\mu \to A_\mu - \partial_\mu \lambda \tag{1.13}$$

Equation (1.12) together with equation (1.13) is a symmetry transformation of the Schrödinger equation. All this is nothing but the well known minimum coupling algorithm,

$$p_\mu \to p_\mu - eA_\mu$$

The reason that this reinterpretation of gauge transformations is acceptable is that the Quantum Mechanical phase is not a directly measurable quantity. It is now clear that Electromagnetism can be interpreted as a Quantum Mechanical local gauge theory. This time it is the local phase of the wave function which is the physical degree of freedom that depends on its spacetime position.

The modern rebirth of gauge theory stemmed from a study of the strong forces mediated by the Yukawa Meson, and Heisenberg's iso spin interpretation of the identity of neutrons and protons when electromagnetic interactions are switched off. That is the strong force was invariant in the SU(2) isotopic spin group.

The difficulty was that iso spin is not a local gauge symmetry, because it is an internal Quantum number independent of spacetime location. So there was no question of an iso spin potential connection whose Quantum would be the Yukawa Meson.

Nevertheless in 1954 Yang and Mills went ahead to treat strong interactions as a gauge invariant field theory by postulating that the local gauge group was the SU(2) iso spin group, in analogy with the electromagnetic case. This time the proposed connection was a linear combination of the angular momentum operators,

$$A_\mu = \sum_\imath A^\imath_\mu(x) L_\imath \tag{1.14}$$

This is a generalization of the electromagnetic case. In the latter, the operators $L_\imath$ are replaced by the unit matrix and the coefficients $A_\mu(x)$ are proportional to the phase change $\delta_\mu \lambda$. As can be seen from (1.14) the Yang-Mills potential is both a field in spacetime and an operator in iso spin space. It must be observed that like the electromagnetic field the Yang-Mills field is mediated by zero mass Bosons. This is because a massive intermediary would imply a term of the form $m^2 A_\mu A^\mu$, which is clearly not gauge invariant.

Let us now see how a symmetry group transformation leads us to a connection which can be identified with the gauge potential field. Indeed, for an arbitrary non-Abelian group, the symmetry transformation is given by

$$U\Psi = \exp\left(\imath q \sum_k \Theta^k(x) F_k\right)\Psi \tag{1.15}$$

In (1.15), the fact that $\Theta^k(x)$ are continuous functions of x defines the local transformation. q is the coupling constant for the gauge group in question. F_k are the generators of the internal symmetry group, satisfying the commutation relations

$$[F_\imath, F_j] = \imath \epsilon_{\imath j k} F_k \ ,$$

In (1.15) if an infinitesimal transformation of the spacetime coordinate is carried out, we get instead of the usual derivative, the gauge covariant derivative describing the changes in both the external and internal components of $\Psi(x)$ viz.,

$$D_\mu \Psi_\beta = \sum_\alpha \left[\delta_{\beta\alpha}\partial_\mu - \imath q (A_\mu)_{\beta\alpha}\right]\Psi_\alpha \tag{1.16}$$

where A_μ are given by

$$(A_\mu)_{\alpha\beta} = \sum_k (\partial_\mu \Theta^k)(F_k)_{\alpha\beta}$$

A special case of (1.16) is the U(1) electromagnetic gauge group, for which this reduces to the usual form with the minimal coupling

$$D_\mu \Psi = (\partial_\mu - \imath q A_\mu)\Psi$$

Thus for the electromagnetic gauge group the gauge covariant derivative is the familiar canonical momentum. It must be noted that the potential A_μ is both an external field and as well, an internal space operator. Furthermore in the non-Abelian gauge group, an internal operator part of the potential would contain a linear combination of the group generators, F_k

which do not in general commute. However as we saw above, the problem has been that we cannot incorporate a mass for the gauge field in an invariant manner. This is achieved by considering an additional field. In the case of weak interaction this is the Higgs field, which breaks the symmetry and leads to a mass generating mechanism.

The Theory of Relativity (Special and General) and Quantum Theory have been often described as the two pillars of twentieth century physics. Each in its own right explained aspects of the universe to a certain extent. But there are still many unanswered questions. For example spacetime singularities (like the Big Bang), termed by John Wheeler as the Greatest Crisis of Physics, the many divergences encountered in particle physics, some eighteen arbitrary parameters in the standard model, elusive monopoles (and Higgs bosons), gravitational waves and Dark Matter and and Supersymmetric particles and so on.

To quote 't Hooft (drawing a comparison with planetary orbits) [41], "What we do know is that the standard model, as it stands today, cannot be entirely correct, in spite of the fact that the interactions stay weak at ultra-short distance scales. Weakness of the interactions at short distances is not enough; we also insist that there be a certain amount of stability. Let us use the metaphor of the planets in their orbits once again. We insisted that, during extremely short time intervals, the effects of the forces acting on the planets have hardly any effect on their velocities, so that they move approximately in straight lines. In our present theories, it is as if at short time intervals several extremely strong forces act on the planets, but, for some reason, they all but balance out. The net force is so weak that only after long time intervals, days, weeks, months, the velocity change of the planets become apparent. In such a situation, however, a reason must be found as to why the forces at short time scales balance out. The way things are for the elementary particles, at present, is that the forces balance out just by accident. It would be an inexplicable accident, and as no other examples of such accidents are known in Nature, at least not of this magnitude, it is reasonable to suspect that the true short distance structure is not exactly as described in the standard model, but that there are more particles and forces involved, whose nature is as yet unclear."

Further, there has been much talk about going beyond the Standard Model, ever since the mass of the neutrino, predicted independently by the author [42] was confirmed by the Super Kamiokande experiment. For according to the Standard Model, the neutrino should be massless. Clearly, we have reached the limits of the Standard Model.

Returning to the issue of Quantum Mechanics and General Relativity it was almost as if Rudyard Kipling's "The twain shall never meet" was true for these two intellectual achievements, a view endorsed by Pauli, who went as far as to say that we should not try to put together what God had intended to be separate. For decades there have been fruitless attempts to unify electromagnetism and gravitation, or Quantum Theory and General Relativity. For, we cannot leave the Universe with a split personality−−one for the micro world and one for the macro universe. Such a dichotomic description of nature is totally unsatisfactory. For sometime it looked like String Theory would answer all questions, as we will see shortly.

1.9 Standard Cosmology

In the sixties, it was not suspected that Elementary Particle Physics would be intimately connected with Cosmology, which was at the other end of the spectrum in terms of sizes! But it was subsequently realized that further experimentation on theoretical particle models would require energies that could not be available in foreseeable particle accelerators. Fortunately the Big Bang model of cosmology provides a scenario in the early Universe where such high energies were accessible and consequently particle physics predictions become testable. The very interesting development that has emerged is that Particle Physics and Cosmology have got linked by this high energy bridge.

The so called Big Bang model arose from three main observations. The first was the discovery in the 1920s that the Universe is expanding, in the sense that the basic constituents, the galaxies (as then believed) showed red shifts. Furthermore as Hubble discovered, the farther the galaxy, the greater its speed of recession. This is Hubble's Law: $v = Hr$, where H is the Hubble constant.

Another important observation was about light element abundance−−or overabundance−−in the Universe. In the 1940s Gamow and coworkers provided an explanation for this. The early Universe must have been very hot and dense. The synthesis of light elements took place when the Universe was at a temperature of $10^9 K$. However heavier elements were formed later, inside the stars, and were strewn about by supernova explosions.

Finally there was a cosmic footprint of an explosion from a very early hot and dense state. This was the residual background radiation from that early event. In the present epoch however the earlier intense radiation would have cooled, and it was calculated that it would be in the form of microwaves. Exactly such a cosmic background microwave radiation footprint was accidentally discovered in 1965 by Penzias and Wilson. This effectively overthrew a competing model of that time−−the Steady State Model, which has now become history [43].

So the picture to emerge [43–45] was that the Universe was born in a titanic explosion or Big Bang, a name made popular by Gamow. Exactly at the time of the Big Bang some fourteen billion years ago, it is reckoned, all the matter and energy of the Universe was concentrated at a single point, where the density and curvature would be infinite. This is the Big Bang singularity. Following the Big Bang, matter and energy has been flung all round and even today the galaxies (or clusters of galaxies) are rushing outward due to that initial impact.

The question that arises is, will the expansion of the Universe continue for ever, or would it slow down to a halt and then collapse? The answer to this would depend on the mass/energy density of the Universe. If this value is greater than a critical value, then the gravitational attraction will ultimately prevail over the expansion and the Universe would collapse. But if the density is less than the critical value, the Universe would go on expanding for ever. This critical density is given by,

$$\rho_{crit} = \frac{3H^2}{8\pi G} = 2 \times 10^{-29} h^2 g/cm^3$$

Observations seem to indicate that the density of the Universe was close to the critical value. Further an observation of the speeds of rotation of the galaxies indicated that the galaxies themselves contained more matter than met the eye. This lead to "Dark Matter" being invoked. Dark matter has not been directly detected, nor can it be precisely characterized, even though there have been a number of possible candidates. For example invisible Black Holes or even difficult to detect brown dwarf stars. Exotic massive particles have also been proposed as also massive neutrinos or monopoles. With dark matter thrown in, it was believed that the Universe had the critical density to reverse the expansion.

Though the Big Bang model could explain several observations, there were subtler questions which came to haunt. These were: How come the density of the Universe, which could have been anything, is in fact so close to the critical density in a process spread over billions of years? More precisely such a close critical density today would imply that even after about a billionth of a second after the Big Bang the density was equal to the criti-

THE "THERMODYNAMIC" UNIVERSE - Exploring the Limits of Physics

cal density accurate to some twenty five decimal places. Alternatively this means that the Universe or space is very flat. This need not have been so. And then the Universe appears uniform on large scales. For instance the cosmic microwave background radiation is uniform in temperature to a high degree of accuracy. How can this be so for regions separated by such vast distances. For since the Big Bang for light itself there has not been enough time to connect them. This is called the horizon problem.

Finally how do we account for the small scale inequalities or lumps in the Universe which we see as galaxies?

In 1981 Alan Guth proposed his inflation Theory ([46, 47]). According to this there was a super fast or super rapid expansion in the early stages of the Universe, so that the size of the Universe exploded to several times its original size within a small fraction of a second.

To put it simply this super fast or exponential expansion flattens out the Universe, thus explaining the first problem. The horizon problem is also accounted for: Due to the super fast expansion or inflation, distant regions were much closer together than with an usual expansion. So they would be at the same temperature. Furthermore Quantum fluctuations in the inflation field would cause fluctuations in density, that is they would seed the formation of galaxies. Finally it may be added that given the inflationary scenario, the fact that exotic particles like magnetic monopoles are not detected is also explained. The rapid inflation would have diluted such particles and made them unobservable.

A time line of the Universe would be [48]

$$1 \qquad t = 10^{-43} secs, \quad T = 10^{32} K$$

The Planck era of Quantum Gravity would have just ended and the Universe would be described by a Grand Unified Theory

$$2 \qquad t = 10^{-35} secs, \quad T = 10^{28} K$$

The Grand Unified symmetry is broken. The size of the Universe would still be only a millimeter across

$$3 \qquad t = 10^{-10} secs, \quad T = 10^{15} K$$

At this stage electroweak symmetry is broken. Already the Universe has swelled to a size of $10^{14} cms$.

$$4 \qquad t = 10^{-5} secs, \quad T \sim 10^{12} K$$

QCD is switched off and quarks combine to form hadrons

$$5 \qquad t \sim 3 min, \quad T \sim 10^{9} K$$

Nucleosynthesis begins and nuclei of lighter elements like Helium and Lithium begin to form

$$6 \qquad t = 10^{-5} yrs, \quad T \sim 4000K$$

Electrons and nuclei combine to form neutral atoms as charged particles are no longer present. So there is no scattering of photons and radiation in general including the Cosmic Microwave Background Radiation.
Interestingly Optical and Radio Astronomy cannot probe beyond this time

$$7 \qquad t \sim 10^{9} yrs, \quad T \sim 10K$$

Galaxy formation begins

$$8 \qquad t \sim 10^{10} yrs, \quad T \sim 2.7K$$

This is the Universe of today.
The above was the model till 1997. That year, the author put forward an alternative model which in fact went against the then existing belief. On the contrary, this model predicted a dark energy driven accelerating ever expanding Universe. In 1998 dramatic confirmation for the new model came from the observations of Perlmutter, Schmidt, Kirshner and others. We will come back to all this in Chapter 3. Clearly the limits of Standard Big Bang Cosmology had been reached in 1997.

1.10 Bosonic Strings

We have already noted that String Theory (and its derivatives) held the promise of unifying gravitatioin with other fundamental forces. Let us begin with T. Regge's work of the 1950s referred to earlier [32, 49, 50] in which he carried out a complexification of the angular momentum and analyzed particle resonances. As is well known, the resonances could be fitted by a straight line plot in the (J, M^2) plane, where J denotes the angular momentum and M the mass of the resonances. That is we have

$$J \propto M^2, \tag{1.17}$$

Equation (1.17) suggested that not only did resonances have angular momentum, but they also resembled extended objects. This was contrary to the belief that elementary particles were point like. In fact as we saw, at the turn of the twentieth century, Poincare, Lorentz, Abraham and others had toyed with the idea that the electron had a finite extension, but they had to abandon this approach, because of a conflict with Special Relativity.

The problem is that if there is a finite extension for the electron then forces on different parts of the electron would exhibit a time lag, requiring the so called Poincare stresses for stability [6, 7, 33].

In this context, it may be mentioned that in the early 1960s, Dirac came up with an imaginative picture of the electron, not so much as a point particle, but rather a tiny closed membrane or bubble. Further, the higher energy level oscillations of this membrane would represent the "heavier electrons" like muons [51].

Then, in 1968, G. Veneziano came up with a unified description of the Regge resonances (1.17) and other scattering processes. Veneziano considered the collision and scattering process as a black box and pointed out that there were in essence, two scattering channels, s and t channels. These, he argued gave a dual description of the same process [52, 53].

In an s channel, particles A and B collide, form a resonance which quickly disintegrates into particles C and D. On the other hand in a t channel scattering, particles A and B approach each other, and interact via the exchange of a particle q. The result of the interaction is that particles C and D emerge. If we now enclose the resonance and the exchange particle q in an imaginary black box, it will be seen that the s and t channels describe the same input and the same output: They are essentially the same.

There is another interesting hint which we get from Quantum Chromo Dynamics that we encountered. Let us come to the inter-quark potential [30, 34]. There are two interesting features of this potential as noted. The first is that of confinement, which is given by a potential term like

$$V(r) \approx \sigma r, \quad r \to \infty,$$

where σ is a constant. This describes the large distance behavior between two quarks. The confining potential ensures that quarks do not break out of their bound state, which means that effectively free quarks cannot be observed.

The second interesting feature is asymptotic freedom. This is realized by a Coulumbic potential

$$V_c(r) \approx -\frac{\propto (r)}{r} (\text{small}\, r)$$

$$\text{where } \propto (r) \sim \frac{1}{ln(1/\lambda^2 r^2)}$$

The constant σ is called the string tension, because there are string models which yield $V(r)$. This is because, at large distances the inter-quark field is string like with the energy content per unit length becoming constant. Use of the angular momentum−−mass relation indicates that $\sigma \sim (400MeV)^2$. Such considerations lead to strings which are governed by the equation [54–57]

$$\rho\ddot{y} - Ty'' = 0, \tag{1.18}$$

$$\omega = \frac{\pi}{2l}\sqrt{\frac{T}{\rho}}, \tag{1.19}$$

$$T = \frac{mc^2}{l}; \quad \rho = \frac{m}{l}, \tag{1.20}$$

$$\sqrt{T/\rho} = c, \tag{1.21}$$

T being the tension of the string, l its length and ρ the line density and ω in (1.19) the frequency. The identification (1.19),(1.20) gives (1.21), where c is the velocity of light, and (1.18) then goes over to the usual d'Alembertian or massless Klein-Gordon equation.

Further, if the above string is quantized canonically, we get

$$\langle \Delta x^2 \rangle \sim l^2. \tag{1.22}$$

Thus the string can be considered as an infinite collection of harmonic oscillators [55]. (Indeed, we will return to this model of a collection of Harmonic Oscillators, in later Chapters.) Further we can see, using equations (1.19) and (1.20) and the fact that

$$\hbar\omega = mc^2$$

that the extension l is of the order of the Compton wavelength in (1.22), a circumstance that was called one of the miracles of the string theory by Veneziano [52].

It must be mentioned that the above considerations describe a "Bosonic String", in the sense that there is no room for the Quantum Mechanical spin. This can be achieved by giving a rotation to the relativistic quantized string as was done by Ramond [16, 58]. In this case we recover (1.17) of the Regge trajectories. The particle is now an extended object, at the Compton scale, rotating with the velocity of light. Furthermore in superstring theory there is an additional term a_0, viz.,

$$J \leq (2\pi T)^{-1}M^2 + a_0\hbar, \text{ with } a_0 = +1(+2) \text{ for the open (closed) string.} \tag{1.23}$$

The term a_0 in (1.23) comes from the Zero Point Energy. Usual gauge bosons are described by $a_0 = 1$ and gravitons by $a_0 = 2$.

It is also well known that string theory has always had to deal with extra dimensions which reduce to the usual four dimensions of physical spacetime when we invoke the Kaluza-Klein approach at the Planck Scale [59]. Briefly, this means that the extra dimensions have infinitesimal dimensions, or more accurately, are wound up in infinitesimal cylinders.

All these considerations have been leading to more and more complex models, the latest version being the so called M-Theory. In this latest theory supersymmetry is broken so that the supersymmetric partner particles do not have the same mass as the known particles. Particles can now be described as soliton like branes, resembling the earlier Dirac membrane. M-Theory also gives an interface with Black Hole Physics. Further these new masses must be much too heavy to be detected by current accelerators. The advantage of Supersymmetry (SUSY) is that a framework is now available for the unification of all the interactions including Gravitation. It may be mentioned that under a SUSY transformation, the laws of physics are the same for all observers, which is the case in General Relativity (Gravitation) also. Under SUSY there can be a maximum of eleven dimensions, the extra dimensions being curled up as in Kaluza-Klein theories. In this case there can only be an integral number of waves around the circle, giving rise to particles with quantized energy. However for observers in the other four dimensions, it would be quantized charges, not energies. The unit of charge would depend on the radius of the circle, the Planck radius yielding the value e. This is the root of the unification of electromagnetism and gravitation in these theories.

In M-Theory, the position coordinates become matrices and this leads to, as we will see, a noncommutative geometry or fuzzy spacetime in which spacetime points are no longer well defined [60]

$$[x, y] \neq 0$$

From this point of view the mysterious M in M-Theory could stand for Matrix, rather than Membrane or magic as some suppose.

So M-Theory is the new avatar of Quantum Superstring Theory. Nevertheless it is anything but the last word. There are still any number of routes for compressing ten dimensions to our four dimensions. There is still no contact with experiment. It also appears that these theories lead to an unacceptably high cosmological constant and so on and so on. This has prompted some String theorists to invoke an Anthropic Principle approach.

According to this, the universe is really a landscape of some 10^{500} universes, each with its own fundamental laws and constants. It so happens that we are inhabiting a universe with the observed laws and physical constants. This could well be the death knell of the theory.

All this along with the non-verifiability of the above considerations and the fact that the Planck scale $\sim 10^{20} GeV$ is also beyond forseeable attainment in collidors has lead to much criticism even though it is generally accepted that the mathematical ideas have been rich and promising. Indeed it is becoming increasingly clear that this intellectual tour de force, touted as The Theory of Everything has perhaps reached its limits. Not just String Theory, but also all so called “reductionist” theories are under the scanner.

1.11 End of the Road?

Reductionism has been at the heart of twentieth century Theoretical Physics. Beginning with the atomism of ancient India and Greek thinkers, it was reborn in the nineteenth century. This spirit is very much evident in Einstein’s concept of locality in which an arbitrarily small part of the universe can be studied without reference to other parts of it. Indeed it is this philosophy of reductionism which has propelled the most recent studies such as String Theory or other Quantum Gravity approaches.

Against this backdrop, the first salvo was fired by Nobel Laureate R.B.Laughlin. “A Different Universe”, his recent book, would come as a shock because he debunks reductionism in favour of what is these days called emergence. That is his central theme [61]. The fundamental laws of nature emerge through collective self organization and do not require knowledge of their component parts, that is microscopic rules, in order that we comprehend or exploit them. The distinction between fundamental laws and laws descending from them is a myth. In his words, “... I must openly discuss some shocking ideas: the vacuum of space-time is ’matter’, the possibility that relativity is not fundamental...” He argues that all fundamental constants require an environmental context to make sense. This is contrary to the reductionist view that basic bricks build up structures.

Laughlin takes pain to bring to our notice that there is now a paradigm shift from the older reductionist view to a view of emergence. For him Klitzing’s beautiful experiment bringing out the Quantum Hall effect is symbolic of the new ethos. The Quantum of Hall resistance is a combination of fundamental constants viz., the indivisible quantum of electric charge e, the Planck constant h and the speed of light c. This means that these supposedly basic building blocks of the universe can be measured with breathtaking accuracy, without dealing with the building blocks themselves. Though, from one point of view, this resembles the fact that bulk properties emerge from underlying and more fundamental microscopic properties, he argues that this latter effect reveals that supposedly indivisible quanta like the electric charge e can be broken into pieces through self organization of phases. That is, the supposedly fundamental things are not necessarily fundamental. Furthermore, for example, in superconductivity, many of the so called minor details are actually inessential−−the exactness of the Meissner and Josephson effects does not require the rest of the finer detail to be true.

Admittedly there are a number of grey areas in modern theoretical physics which are generally glossed over. For instance Dirac's Hole Theory of anti matter. However in silicon, there are many electrons locked up in the chemical bonds and it is possible to pull an electron out of a chemical bond. This makes a hole which is mobile and acts in every way like an extra electron with opposite charge added to the silicon. This idea however requires the analogue of a solid's bond length. In Particle Physics such a length conflicts fundamentally with the Principle of Relativity as we have seen, unless, as we have argued, it breaks down at the Compton scale. On the contrary, Laughlin laments, "... instead, physicists have developed clever semantic techniques for papering it over... Thus instead of Holes one speaks of anti particles. Instead of bond length one speaks of an abstraction called the ultra violet cutoff, a tiny length scale introduced into the problem to regulate−−which is to say, to cause it to make sense. Below this scale one simply aborts one's calculations... Much of Quantum Electrodynamics, the mathematical description of how light communicates with the ocean of electrons... boils down to demonstrating the unmeasurableness of the ultra violet cutoff... The potential of overcoming the ultra violet problem is also the deeper reason for the allure of String Theory, a microscopic model for the vacuum that has failed to account for any measured thing... The properties of empty space relevant to our lives show all the signs of being emergent phenomena characteristic of a phase of matter. They are simple, exact, model insensitive, and universal. This is what insensitivity to the ultra violet cutoff means physically."

Moreover as we will see particularly in Chapter 4, quantized sound waves or phonons have an exact parallel with photons−−in fact their quantum properties are identical to those of light. However sound is a collective motion of elastic matter, while in our understanding, light is not. This means that quantization of sound may be deduced from the underlying laws of Quantum Mechanics obeyed by the atoms, whereas in the case of light this is postulated. This is a logical loose end and ultimately we bring in the gauge effect to cover this. But unfortunately, "there is also a fundamental incompatibility of the gauge effect with the principle of relativity, which one must sweep under the rug by manipulating the cutoff." Laughlin complains that in spite of the evidence against reductionism, sub nuclear experiments are generally described in reductionist terms. These points will be considered in detail, in the following Chapters.

Turning to General Relativity, Laughlin points out that it is a speculative post Newtonian Theory of Gravity, an invention of the mind, "it is just controversial and largely beyond the reach of experiment", unlike Special Relativity which was a discovery of the behavior of nature. He then points out the contradiction between Special and General Relativity−−in the former Einstein did away with the concept of the Ether. But this reenters the latter theory in the form of the fabric of space. Touching upon the skeletons in the closet of General Relativity, Laughlin discusses the embarrassment that is caused by a non zero cosmological constant.

He concludes that if Einstein were alive today, he would be horrified at this state of affairs and would conclude that his beloved principle of Relativity was not fundamental at all but emergent.

Laughlin takes a critical look at renormalizability, a pillar of modern theoretical physics, and cosmology. Indeed, we have already commented on this. "If renormalizability of the vacuum is caused by proximity to phase transitions, then the search for an ultimate theory would be doomed on two counts: It would not predict anything even if you found it, and it could not be falsified...

"The political nature of cosmological theories explains how they could so easily amalgamate String Theory, a body of mathematics with which they actually have very little in common... (String Theory) has no practical utility however, other than to sustain the myth of the ultimate theory. There is no experimental evidence for the existence of Strings in nature... String Theory is, in fact a textbook case of ... a beautiful set of ideas that will always remain just barely out of reach. Far from a wonderful technological hope for a greater tomorrow, it is instead a tragic consequence of an absolute belief system in which emergence plays no role..."

Laughlin has captured the mood of pessimism that prevails in the minds of several high energy physicists. He goes on to cite the famous joke that the

Physical Review is now so voluminous that stacking up successive issues would generate a surface travelling faster than the speed of light, although without violating Relativity, because the Physical Review contains no information anyway.

However before proceeding further it must be mentioned, as we will see in the following Chapters, that the author's own work during the last decade has borne out the spirit of these ideas, that the iron clad law of physics is more thermodynamic and stochastic in nature, that the velocity of light or the gravitational constant can be deduced from such considerations rather than be taken as fundamental inputs; how, it is possible to have schemes that bypass the awkward questions raised in the book, without brushing them away below the carpet, by considering an a priori Quantum vacuum in which fluctuations take place.

Returning now to String Theory there is no doubt that it has straddled the past two decades and more as the only contender for the Theory of Everything. Some years ago Nobel Laureate Sheldon Glashow described it, sarcastically, as the only game in town. In recent months though the theory is not only being debunked, it is facing a lot of flak, particularly in the worldwide media. Laughlin may be faulted on the grounds that he is not a String Theorist or a Particle Physicist. The decisive tilt has come from Nobel Laureate David Gross, very much an insider, who as it were, spilt the beans at the recent 23rd Solvay Conference in Physics held in Brussels, Belgium, in late 2005. He stated "We don't know what we are talking about!" He then went on to say, "Many of us believed that string theory was a very dramatic break with out previous notions of quantum theory. But now we learn that string theory, well, is not that much of a break." And that physics is in "a period of utter confusion."

At this meeting Gross compared the state of Physics today to that during the first Solvay Conference in 1911 "They were missing something absolutely fundamental," he said. "We are missing perhaps something as profound as they were back then."

Thus the Time Magazine, August 14, 2006 issue notes:

"By now, just about everyone has heard of string theory. Even those who don't really understand it-which is to say, just about everyone-know that it's the hottest thing in theoretical physics. Any university that doesn't have at least one string theorist on the payroll is considered a scientific backwater. The public, meanwhile, has been regaled for years with magazine articles

"But despite its extraordinary popularity among some of the smartest people on the planet, string theory hasn't been embraced by everyone-and now, nearly 30 years after it made its initial splash, some of the doubters are becoming more vocal. Skeptical bloggers have become increasingly critical of the theory, and next month two books will be hitting the shelves to make the point in greater detail. Not Even Wrong, by Columbia University mathematician Peter Woit, and The Trouble With Physics, by Lee Smolin at the Perimeter Institute for Theoretical Physics in Waterloo, Ont., both argue that string theory (or superstring theory, as it is also known) is largely a fad propped up by practitioners who tend to be arrogantly dismissive of anyone who dare suggest that the emperor has no clothes

"Bizarre as it seemed, this scheme appeared on first blush to explain why particles have the characteristics they do. As a side benefit, it also included a quantum version of gravity and thus of relativity. Just as important, nobody had a better idea. So lots of physicists, including Woit and Smolin, began working on it.

"Since then, however, superstrings have proved a lot more complex than anyone expected. The mathematics is excruciatingly tough, and when problems arise, the solutions often introduce yet another layer of complexity. Indeed, one of the theory's proponents calls the latest of many string-theory refinements "a Rube Goldberg contraption." Complexity isn't necessarily the kiss of death in physics, but in this case the new, improved theory posits a nearly infinite number of different possible universes, with no way of showing that ours is more likely than any of the others.

"That lack of specificity hasn't slowed down the string folks. Maybe, they've argued, there really are an infinite number of universes-an idea that's currently in vogue among some astronomers as well-and some version of the theory describes each of them. That means any prediction, however outlandish, has a chance of being valid for at least one universe, and no prediction, however sensible, might be valid for all of them.

"That sort of reasoning drives critics up the wall. It was bad enough, they say, when string theorists treated nonbelievers as though they were a little slow-witted. Now, it seems, at least some superstring advocates are ready to abandon the essential definition of science itself on the basis that string theory is too important to be hampered by old-fashioned notions of experimental proof

"And it is that absence of proof that is perhaps most damning."

"It's fine to propose speculative ideas," says Woit, "but if they can't be tested they're not science." To borrow the withering dismissal coined by the great physicist Wolfgang Pauli, they don't even rise to the level of being wrong. That, says Sean Carroll of the University of Chicago, who has worked on strings, is unfortunate. "I wish string theorists would take the goal of connecting to experiment more seriously," he says.

According to the August 27, 2006 issue of the Scientific American,

"With a tweak to the algorithms and a different database, the Website could probably be made to spit out what appear to be abstracts about superstring theory: "Frobenius transformation, mirror map and instanton numbers" or "Fractional two-branes, toric orbifolds and the quantum McKay correspondence." Those are actually titles of papers recently posted to arXiv.org repository of preprints in theoretical physics, and they may well be of scientific worth-if, that is, superstring theory really is a science. Two new books suggest otherwise: that the frenzy of research into strings and branes and curled-up dimensions is a case of surface without depth, a solipsistic shuffling of symbols as relevant to understanding the universe as randomly generated dadaist prose.

"In this grim assessment, string theory-an attempt to weave together general relativity and quantum mechanics−−is not just untested but untestable, incapable of ever making predictions that can be experimentally checked. With no means to verify its truth, superstring theory, in the words of Burton Richter, director emeritus of the Stanford Linear Accelerator Center, may turn out to be "a kind of metaphysical wonderland." Yet it is being pursued as vigorously as ever, its critics complain, treated as the only game in town.

"String theory now has such a dominant position in the academy that it is practically career suicide for young theoretical physicists not to join the field," writes Lee Smolin, a physicist at the Perimeter Institute for Theoretical Physics, in The Trouble with Physics: The Rise of String Theory, the Fall of a Science, and What Comes Next. "Some young string theorists have told me that they feel constrained to work on string theory whether or not they believe in it, because it is perceived as the ticket to a professorship at a university."

"Neither of these books can be dismissed as a diatribe. Both Smolin and Woit acknowledge that some important mathematics has come from contemplating superstrings. But with no proper theory in sight, they assert, it is time to move on. "The one thing everyone who cares about fundamental physics seems to agree on is that new ideas are needed," Smolin writes. "We are missing something big."

"The story of how a backwater of theoretical physics became not just the rage but the establishment has all the booms and busts of an Old West min-ing town. Unable to fit the four forces of nature under the same roof, a few theorists in the 1970s began adding extra rooms−−the seven dimensions of additional closet space that unification seemed to demand. With some mathematical sleight of hand, these unseen dimensions could be curled up (“compactified”) and hidden inside the cracks of the theory, but there were an infinite number of ways to do this. One of the arrangements might describe this universe, but which?

“The despair turned to excitement when the possibilities were reduced to five and to exhilaration when, in the mid-1990s, the five were funneled into something called M Theory, which promised to be the one true way. There were even hopes of experimental verification

“That was six years ago, and to hear Smolin and Woit tell it, the field is back to square one: recent research suggests that there are, in fact, some 10^{500} perfectly good M theories, each describing a different physics. The theory of everything, as Smolin puts it, has become a theory of anything.

“Faced with this free-for-all, some string theorists have concluded that there is no unique theory, that the universe is not elegant but accidental. If so, trying to explain the value of the cosmological constant would make as much sense as seeking a deep mathematical reason for why stop signs are octagonal or why there are 33 human vertebrae”

An article in the Financial Times (London) in June 2006 by Physicist Robert Mathews noted:

“They call their leader The Pope, insist theirs is the only path to enlightenment and attract a steady stream of young acolytes to their cause. A crackpot religious cult? No, something far scarier: a scientific community that has completely lost touch with reality and is robbing us of some of our most brilliant minds.

“Yet if you listened to its cheerleaders-or read one of their best-selling books or watched their television mini-series-you, too, might fall under their spell. You, too, might come to believe they really are close to revealing the ultimate universal truths, in the form of a set of equations describing the cosmos and everything in it. Or, as they modestly put it, a “theory of everything”.

“This is not a truth universally acknowledged. For years there has been concern within the rest of the scientific community that the quest for the theory of everything is an exercise in self-delusion. This is based on the simple fact that, in spite of decades of effort, the quest has failed to produce a single testable prediction, let alone one that has been confirmed.

“For many scientists, that makes the whole enterprise worse than a theory that proves to be wrong. It puts it in the worst category of a scientific theories, identified by the Nobel Prize-winning physicist Wolfgang Pauli: it is not even wrong. By failing to make any predictions, it is impossible to tell if it is a turkey, let alone a triumph.

"It is this loss of contact with reality that has prompted so much concern among scientists-at least, those who are not intimidated by all the talk of multidimensional superstrings and Calabi-Yau manifolds that goes with the territory. But now one of them has decided the outside world should be told about this scientific charade. As a mathematician at Columbia University, Peter Woit has followed the quest for the theory of everything for more than 20 years. In his new book "Not Even Wrong" he charts how a once-promising approach to the deepest mysteries in science has mutated into something worryingly close to a religious cult."

A review in the December 2006 issue of Physics Today notes: "Noted theoretical physicist Sheldon Glashow has famously likened string theory to medieval theology because he believes both are speculations that cannot be tested. Yet if readers believe Lee Smolin and Peter Woit, they might conclude that the more apt comparison is to the Great disappointment of 1844, when followers of the Baptist preacher William Miller gave up all their worldly possessions and waited for the Second Coming. The empirical inadequacy of that prediction led to apostasy and schisms among thousands of Miller's followers. At least one of the branches claimed that the event had in fact occurred, but in a heavenly landscape linked to the world of experience through only the weak but all-pervasive spiritual interaction. Yet irritating differences exist between Miller's followers and the "disappointed" of the 1984 coming of the theory of everything; a majority of the latter seem to have preserved their faith and gained worldly fortune in the form of funding, jobs, and luxurious conferences at exotic locales."

In all this confusion, we should not forget two important points. The first is that String Theory still remains a mathematically beautiful self consistent system of thought. Perhaps the flak that String Theory is receiving is more due to reasons in the domain of the sociology of science. To elaborate, String Theory has been touted as a theory, which, in the strict sense it is not. If it had been promoted as a hypothesis, one of a few possible, perhaps, there would have been much less criticism.

Indeed, some years ago Nobel Laureate 't Hooft had noted [62] "Actually, I would not even be prepared to call string theory a "theory" but rather a "model" or not even that: just a hunch. After all, a theory should come together with instructions on how to deal with it to identify the things one wishes to describe, in our case the elementary particles, and one should, at least in principle, be able to formulate the rules for calculating the properties of these particles, and how to make new predictions for them."

Moreover in the process String Theory adopted strong arm facist type tactics including marketing through the media, while at the same time making not too covert attempts to suppress other ideas. The backlash was therefore inevitable.

The second point is even more important and is expressed in David Gross's statement that perhaps we are missing something very profound. This throws up a great challenge and makes for very exciting times.

Science has been described as a quest for the how and why of nature. Over the centuries it has been guided by some principles which have crystallized into a methodology. Thus observation leads to the framing of hypotheses. It is expected that the hypotheses would have maximum simplicity and maximal economy. This means that, apart from being simple, the maximum number of observations are explained by a minimum number of hypotheses. Further tests would then confirm or disprove the hypotheses, if the hypotheses are found to be consistent with experiment. The richness of a hypothesis is judged by the predictions it can make. These predictions must be either provable or disprovable as stressed by Sir Karl Popper.

As far as Fundamental Physics is concerned, starting from the early days of Indian and later Greek Atomism, through the Atomic Theory of the 19th century, and subsequently the developments in the 20th century, and the early part of the 21st century, the route followed has been one of a descending-in-size cascade. We have been propelled by the belief that the universe could be understood by a study of its ultimate subconstituents. This spirit as noted is very much evident in Einstein's concepts of locality in which an arbitrarily small part of the universe can be studied without reference to other parts of it. A few decades later Wheeler observed that our studies of the inaccessible Planck scale of $10^{-33} cms$ were really like an understanding of bulk properties of matter by studying the subconstituent molecules [45].

Indeed it is this philosophy of reductionism which has propelled the most recent studies such as String Theory or other Quantum Gravity approaches. Decades of labour has gone into these endeavours and the research output has been enormous.

Nevertheless we seem to have reached an impasse of a type that is all too familiar from the past. There are minor discrepancies or corrections, which nevertheless would point to, not just an incremental change of our concepts, but rather to a paradigm shift itself of the type we witnessed with Kepler's ellipses or Einstein's Relativity or Quantum Theory. Any theory can go only so far as its inherent limitations or constraints permit. At that stage, as Thomas Kuhn [63] notes, there would be a revolution, an overturn of concepts and the old way of looking at things. It is no longer an incremental improvement. It is proposed in this book that the so far one way street of reductionism has reached such a limit, and now has to be tempered in the spirit of Thermodynamics or emergence.