After PHYSICS

.... A Physical Model

Andrew Laidlaw

v3: July 8, 2021
v2: July 27, 2020
v1: July 29, 2019

Contents

Chapter 1

Preface

Modern Physics provides empirical results in a broad range of applications, but the interpretation is in disarray, littered with mystery and paradox. There is widespread recognition, both within the Physics community and without, that many issues in the foundations of Physics have never been satisfactorily resolved.

For a start, there are three disparate foundation theories, which seems like two too many. While each one is fruitful enough as far as testable predictions go, each of them entails conclusions that mystify. Moreover, after many decades of intense effort by the most highly skilled experts, Physics remains unable satisfactorily to unite any two of its three orphan theories. Quantum gravity, in particular, is heard but not seen.

If the objective is to unify Physics, then perhaps, instead of the usual, high level, mathematical abstraction, it would make more sense to begin with unity in the foundations.

Instead of the usual cluster of postulates, principles, axioms and various other disparate asseverations that motivate the foundations, Special Relativity, General Relativity and Quantum Mechanics will all be developed here from a single, already widely accepted, physical idea. When taken as a foundational concept, this idea is explicitly Metaphysical and so, while most experts do see it as a necessary consequence of Special Relativity, they would not countenance using it as a point of departure. The street is supposed to run in the other direction, from "sound" logical and / or observational principles to confounding Metaphysical conclusions.

Modern Physics eschews physical models, motivates its mathematics from such principles and then follows the formalism wherever it leads, and this has led us to a *status quo* where what will be shown to be commonplace wave phenomena are described with adjectives like "weird", "mysterious" and "counterintuitive".

The trend towards focussing on formalism began long before Special Relativity, but that was the tipping point where physical intuition came to be explicitly rejected as a valid criterion in Physics and physical models fell into disrepute. Not only must we routinely accept that our formalisms are not to be understood in any common sense way, but unintelligibility has been elevated to a virtue: One of Physics' vaunted achievements nowadays is to transcend merely human, intuitive ways of thinking.

The retreat into this formal mode of "understanding" is a deal with the devil. He provided some early material successes, but nowadays Physics is paying with its soul. The further it transcends human understanding, the more the devil dances: Physics' mission is to unravel the phenomena, but he has us addicted to mystery. We find ourselves in an ecosystem of oxymorons, inhabited by many universes, backwards in time causation, many space dimensions, space is time and time is space, singularities, paradoxes, information as the root of physical stuff, and so on. And then there are ideas that are outside the nine dots.

The purpose of this book is to redevelop the foundations of Physics in an accessible, intelligible and, above all, physical manner. I have shown that the formalisms we have, as we have them, can be understood physically, thereby ending the Metaphysical crisis that Physics finds itself in. The case is that Physics, as we have it, already has a physical wave basis, but it is just not recognised as such. This work provides one simple physical model for all three foundation theories, placing Special Relativity, General Relativity and Quantum Mechanics in the same conceptual framework. Lorentz symmetry is fully retained but the model nonetheless provides objective simultaneity and objective temporal ordering of events, eliminating the host of paradoxes. Finally, it is shown to be compatible with EPR / Bell instant causal correlations at a distance. The motivating idea is that energy is a wave phenomenon, specifically:

"Whether it appears as matter or as radiation, energy-momentum always propagates at the characteristic velocity, c, in a physical, existential medium."

This fundamental assumption will be referred to throughout as "the Wave Postulate". As all practising physicists know, the idea that energy propagates is one of the central conclusions drawn from Special Relativity, so the main idea is uncontroversial[1].

The great physicist, Richard Feynman, was once asked by a journalist what one message he would convey to the next generation in the hypothetical event that all of Physics had been obliterated. He replied with a familiar, basic and strongly Metaphysical proposition: "Everything is made of atoms". In fact, the notion of point particles as existential entities causes more than its fair share of difficulties in Physics, both in the Mathematics and interpretively. Let us reject that premise, replace it with what we know with the highest degree of certainty, "Everything is made of energy", and follow the reasoning where it leads us.

First it will lead us, unfortunately, to relativistic Mechanics and Lorentz covariance. I say unfortunately because Special Relativity is regarded by almost all physicists as the one and only pristine area of Physics that is so fully resolved that it requires no new insights. And yet, it has been the subject of such controversy over the last century that for many physicists, the worst part of the job is dealing with outsiders attempting to revisit, refute, criticize or otherwise provide any new insight whatsoever into the Special Theory.

Within the Physics community, ever since Einstein's "principled" approach there has been widespread recognition by prominent figures - from Einstein himself to Bell - of the need for a (satisfactory) "constructive" approach [1]. Chapter 4 of this work provides a fully satisfactory constructive theory of Lorentz Invariance, so it might be said that it replaces Lorentz, not Einstein. This con-

[1]Nor is the idea *per se* of a physical medium genuinely controversial these days. Although Physicists still find themselves using heavily qualified phrases like "some kind of a relativistic medium" (so as not to contravene Special Relativity's "No Preferred Frame"dictum), the fact is that the idea of space as a physically real medium resurfaced long ago, in both Gravity and Quantum Physics.

structive approach addresses Brown and Timpson's central argument in [1] that Special Relativity should not be considered as a model for a principled foundation for Quantum Mechanics because it "represents a victory of pragmatism over explanatory depth". The wave approach provides the long sought after explanatory depth.

Let me be clear. We are not going to revisit, refute or criticize the Special Relativity formalism, only redevelop it, exactly as is, from a pure field theory perspective. Many readers may have noticed that all the inherently relativistic theories are field theories. It turns out to be a singularly useful result to prove.

Theorists are already familiar with the argument from the relativity principle (and Einstein clock synchronisation) to $E = mc^2$ and the usual conclusion that the energy that constitutes matter must also be propagating at c, a conclusion that is largely confirmed by the constant modulus, equal to c, of the Dirac velocity operator. All that will be done in Chapter 4 is to develop an argument in the opposite direction, from constant speed wave propagation at c to the relativity principle. I shall not touch or vary the usual formalism, just derive it differently, in a manner that will allow one to bring all three foundation theories under the same roof, to remove the seemingly paradoxical consequences and to deliver foundations that restore physical intuition.

Each of the leading relativistic phenomena (the relativistic mass, length contraction and time dilation) will be shown for wave systems, independent of Special Relativity. Given Einstein clock synchronisation, the Lorentz Transformations then follow, and we arrive back at the relativity principle as a conclusion rather than as an assumption. What changes is that each of the phenomena can now be understood as a real physical effect. An important feature of the new theory is that, unlike Lorentzian relativity, the wave approach to Lorentz Invariance covers the massive particles.

Second, the idea of energy as a propagative phenomenon will lead us directly to the Dirac Equation, de Broglie waves and Quantum Mechanics (Chapters 5 and 6). Finally, it will lead us to gravity and general covariance (Chapters 10 and 11). These Chapters develop a physical model that is mathematically isomorphic to the sum total of three foundation theories, with two caveats:

1. The model does not cover non-relativistic theories. Ordinary Quantum Mechanics in particular will be seen as a non-relativistic approximation to the relativistic Dirac Theory, which is modelled without approximation. It will be clearly seen that taking the non-relativistic approximation frustrates any attempt at a physical wave interpretation of Ordinary Quantum Mechanics.

2. As far as gravity is concerned, this is a pure field approach and point sources cannot be put in by hand (see Section 2.3, criterion 8). The usual source term in the Einstein Field Equations is not consistent with that requirement and it has also been argued persuasively elsewhere that they do not reduce to Special Relativity in the correspondence limit [2], [3]. The target theory, which is modelled without approximation, is the Einstein-Yilmaz variation [4], [5]. It is an empirically indistinguishable, generally covariant theory of gravity that may be more amenable to unification with Quantum Field Theory. Yilmaz is essentially the same as Einstein, but with an additional, field source term on the right hand side of the Field Equations.

Note the distinction that has been drawn in the last paragraph. The physical model is explicitly not a model of the phenomena, but of the theories themselves. The reasons for drawing this distinction are outlined in Chapter 2, which identifies an appropriate philosophical basis for physical models, namely that they belong... "*After* Physics". Rather than a question of choosing between constructive approaches to Physics, *i.e.* physical models, and principled approaches / empirical models, the view to be developed here is that a mature, complete Physics requires both.

Physical models offer "explanatory depth", but one typically requires insight into the mathematical structure of phenomena to construct such models in the first place and the fact is that no model should ever be thought of as "true". The principled approach has been successful in the past at making predictions, but fails to meet humanity's desire to understand the world. More important, in recent decades, the search for new theories has tended to become unmotivated, reduced to a mathematical fishing expedition as it were.

However, the main target of this work is Quantum Mechanics, in particular the interpretation of quantum nonlocality as a wave phenomenon. Once we are given the d'Alembert wave equation, the passage to the Schroedinger Equation is at least well trodden, if not trivial. As easy as the transition is, it is anything but natural: Imposing the nonrelativistic energy-momentum relation on a perfectly good, relativistic wave equation results in a pseudo-wave equation that could never represent physical reality, and any attempt to envisage Schroedinger waves as physical wave packets is destined to mislead.

This is precisely what has confounded every effort at a physical model interpretation of Ordinary Quantum Mechanics. Furthermore, the usual (but physically wrong) conclusion from the Schroedinger Equation, namely that the particle velocity is the group velocity of a dispersive wave packet, given by $v_g = d\omega/dk$, is routinely carried over into the relativistic context where it makes no sense at all. Meanwhile, if energy propagates at c, there is no real dispersion and we require a different explanation for the group or particle velocity. That is why it is necessary to deal first with relativistic Mechanics, then relativistic quantum mechanics' Dirac Equation and only then step back to Ordinary Quantum Mechanics.

With the quantum mechanical equations in hand, the remainder of the basic formalism presents no great obstacles. These are linear equations, so methods for decomposing a general solution into eigensolutions are automatically available. There is also a singular beauty involved with the usual formalism which shows, once and for all, the power of a principled approach to Physics as opposed to physical models: The wave function concept beautifully cuts a gordion knot in the formal interpretation[2] of the quantum mechanical wave equations.

Having motivated the formalism and placed due emphasis on the virtue of the principled approach in Chapter 5, it is the question of making sense of the phenomena, as opposed to the formalism, that will be of most concern. We shall consider, in Chapters 6 and 7, some of the archetypical quantum phenomena from the point of view of the physical model.

I'd love to say that quantum nonlocality was the main impetus for rejecting "Everything is made of atoms" in favour of "Everything is made of energy" but

[2]Essentially the process of extracting predictions by connecting math symbols to observable quantities.

the basic stimulus for that shift came from a different context, the Two Body problem in Classical Electrodynamics, with which I was preoccupied well before the Aspect experiments of the early 1980's. We'll discuss Classical Electrodynamics' failure to deal satisfactorily with actual dynamics problems in Chapter 8. Nonetheless, it is more appropriate in this work on the foundations of Modern Physics to emphasise quantum nonlocality.

To approach the question, it will first be shown, in Chapter 6, how superluminal de Broglie waves appear in luminal wave systems, and the implications this has for the wavefield microstructure of any massive particle. This analysis establishes a visualisable physical context within which violations of Bell Inequalities can occur, and nonlocal relations between observables will be induced, in Chapter 7, from the cellular structure identified in Chapter 6 without requiring any part of the model reality to move faster than the characteristic velocity, thereby violating the Bell Inequalities without violating "local realism"[3].

We shall turn at last to gravity, confident that the rest of Physics is consistent with energy propagating at, and only at, c. An obvious mechanism for gravity then recommends itself, namely that the characteristic velocity should depend on the energy density in a given region of space. This general kind of "refractive medium" approach to gravity has been explored many times in the literature. Borrowing heavily from the well known Yilmaz theory of gravity, it is straightforward to develop a generally covariant physical model that places gravity in the same conceptual framework as relativistic quantum mechanics. At the heart of this model we shall encounter a simple equation, reminiscent of Gauss's Law:

$$\nabla^2 c(r) = \kappa \rho_E \ .$$

Mother nature has rightly earned a reputation for such simplicity. Perhaps most interesting are the various mathematical consequences that separate this generally covariant physical model from the usual (Einsteinian) form of General Relativity. Just to mention a couple of examples, the N-body metric[4] will be shown to be a product of the N 1-body metrics. Since these 1-body metrics have an exponential form, the N-body solution is formed simply by adding the exponents together. Also, there are no singularities, and no black holes (of course we routinely observe very dark grey objects - indistinguishable from black holes - where the speed of light falls very close to, but never reaches, zero). Yilmaz theory accounts for all the observables without the usual pathologies.

There are a couple of pertinent comments to be made up front.

This work could hardly have been accomplished if it were not for the important contributions by generations of physicists. There are too many to name here everyone who contributed to developing these ideas, so this is at best a partial list of the major physicists who played, for me, an important role: (On Mechanics) Newton, Hamilton, La Grange, Maxwell, Lamor, Fitzgerald, Poincaré,

[3] Once this term has been appropriately defined.

[4] The metric (tensor) is basically a function of the coordinates that determines the "distance" (not necessarily a physical distance) between nearby points in a curved space with generalised coordinates, analogous to how the Pythagoras theorem operates in flat spaces with orthogonal coordinate systems. For example, the familiar 2-space Pythagorean relation, $ds^2 = dx^2 + dy^2$, generalises to $ds^2 = \sum_{\alpha,\beta=1}^{4} g_{\alpha\beta} dx^\alpha dx^\beta$ in a curved 4-space. The 16 weights, $g_{\alpha\beta}$, are the metric.

Lorentz, Planck and Einstein; (On Quantum Mechanics) Schroedinger, Pauli, Bohr, Born, Dirac, de Broglie, Bohm, Hamilton, Bell and Einstein; (On gravity) La Grange, Dicke, Puthoff, Yilmaz and, of course, Einstein.

There are also several practicising physicists who provided considerable direct assistance, especially in the areas of science communication and identifying the right questions to research. I am sufficiently grateful that I refrain from potentially damaging any of their reputations by naming them here!

I shall be speaking throughout not as a Physicist, but as a design engineer. Consequently, what you are reading is not a “Physics Theory” *per se*; it is a physical model of Modern Physics, designed specifically to shed some much needed light on how that system works. As discussed in Chapter 2, the question whether this particular model is “true” is essentially meaningless. Maybe the basic idea here, that all energy propagates at c, is “true” and maybe it is not. I don’t care, and nor should you, because we’ll never know for sure.

Therefore physical models ought not be proposed as models of the physical reality itself, only as models of the theories: In principle as well as in practice, the model here could only have been designed literally *After* the Physics was already known.

The model is, however, complete, coherent and internally consistent. It has finally been shown that at least one intelligible, intuitable physical model exists that is commensurate with Modern Physics, inviting us to understand the foundations as a set of theories about nonlinear physical wave systems. This possibility, that Physics is about an intelligible physical system, has all but been dismissed for most of the last century.

The objective was not to do any new Physics, merely to design a constructive, physical model that instantiates the Physics we already have, as we already have it. My efforts to avoid committing Physics *per se* were largely successful but there are a few occasions where I fell short of that goal in a minor way, where I was unable not to modify (be it ever so slightly) a relevant empirical relationship. Any observable effects are tiny indeed, and it remains to be seen whether experiments could actually decide between the model and the target Theory. In any event, whether the defects are in the model or in the Physics, I apologise in advance. Fortunately, none of these cases goes to the basic premise.

In addition to developing the formal physical model for the foundations of Modern Physics, there are a few important detours along the way (in Chapters 6 and 7, 9 and 12), to provide model explanations on topics of general interest that would not usually be considered issues in the foundations of Physics: Why is angular momentum quantised? How can spacelike causal correlations be understood? What can be said about the physical structure of the electron? and so on. Hopefully, the reader will bear in mind the distinction between the formal physical model itself and some general talk about the insights it provides in various cases.

Finally, unlike most texts intended for a general science audience, I have not shied away from including the mathematics. This is *because* there is nothing in it that cannot be understood by a precocious high school student. In fact, visualisable mathematics is one of the key criteria established in Section 2.3 (see criterion 14) for the development of this model, and there is every reason to highlight this aspect of the model.

Chapter 2

Philosophical Basis of Physical Models

2.1 Truth, Realism and Physical Models

Can we ever know the truth about physical reality? Can we be certain even about the premise of the question, that the physical reality exists? Is anything knowable? For the purposes of the present work, let us take the answers to these three questions to be respectively, “No”, “It doesn’t matter”, and “Yes”. This Section considers how these questions are relevant to the role of physical models in Physics.

To see why we are choosing to answer them this way, let us begin with the last one, “Is anything knowable?”.

We shall take the view that some things are knowable with a degree of certainty that is (for the present purposes) indisputable. It is very difficult to construct affirmative statements of such a kind. Perhaps it is all an illusion. Perhaps I did not have breakfast this morning. Perhaps Descartes was not thinking after all. It is of course true beyond a shadow of a doubt that Descartes *might* have been thinking when he said, “I think therefore I am”, but this is a statement about uncertainty.

To doubt that I had breakfast this morning, or that any of us exists may seem unreasonable, but the fact is that everything we know is derived from sense experience. Directly or indirectly we rely on sense experience as a fundamental premise in any assertion that we make about the world.

At the same time we have come to understand that sense experience deludes us routinely, and fundamentally. The mountains and trees that we find have been reduced, under closer investigation, to a composite of subatomic particles that has dissolved, in its turn into... What? We have given it names, but at the level of quantum fuzziness, regardless of the name the meaning is not so clear as we thought it was with mountains and trees.

Even if we were, with more centuries of investigation, to develop a clearer understanding, we can never say that it might not dissolve again into a new, more fundamental abstraction away from the familiar world, and again.

Unreasonable doubts about the validity of sense experience and the status of our knowledge are therefore very much on the table.

However, I am sure beyond even the most unreasonable doubt that I do NOT know the precise date when work commenced on the Sphinx. Even in the face of the most abstruse of unreasonable doubts - past lives, the unity of an atemporal collective consciousness of which we are each but one facet, the world is a simulation on some alien's computer - such sentences are, in my view, undeniably correct by construction[1].

Statements in the negative can frequently be considered true beyond even an unreasonable doubt. When it comes to physical reality, such a statement recommends itself: What we know for sure is that we cannot know it for sure. I acknowledge the contrary views of mystics claiming direct experience, but let us put that aside. The reader may add the words "from experimental evidence" if he or she wishes. The statement is correct, in my view, because any experimental evidence about the system must be gathered from some apparatus that is itself of the system. The resolution of a distance measurement is limited by the "fuzziness" (to use a technical term) of the matter that constitutes whatever we may use as a "ruler", and so on.

This limitation is no cause for despair. On the contrary, one can relax about ever knowing the ultimate Metaphysical truth about reality, safe in the knowledge that it is definitely the wrong problem. It is better to consider problems where there is a chance of success.

Framing a more appropriate problem begins with one's motivation. There are many reasons why we humans study the world, which divide into two main categories, utility and curiosity.

As for utility, we have invented walls and wheels, buttons and spectacles, swords and guns, telephones and refrigerators, aeroplanes and rockets, integrated circuits, lasers, computers and the internet, robots, surveillance and artificial intelligence. We humans like our stuff, and the development of ever more sophisticated stuff needs science and in particular it needs a pragmatic, instrumentalist rendition in Physics.

On the other hand we are also curious in a distinctive way that goes far beyond merely collecting information. We want to understand. We want to organise information in ways that simplify and explain. We want to figure things out.

At first glance, these two objectives may not seem to be at odds with each other, but there is a definite tension - an inherent tradeoff is involved.

You wouldn't know it today, but once upon a time Physics recognised that dealing with the fundamental questions of Life, the Universe and Everything might be counterproductive for a scientific discipline. Where there is no testability there is only debate without progress, heat without light. Testability, of course, would mean dealing only with observables. That is to say, Physics would deal in a formal sense with testable, reproducible experimental facts about the world, leaving the questions of what it "really" is for others to argue about. By understanding the systematisation of the facts, one might over time come to better and better inferences about the truth. The arguments of philosophers could be better focussed.

That was the plan. In practice, things were never quite so cleanly divided. Physics began, inevitably, with physical models. The cannon ball would be

[1]The definition of "I" in the sentence having been made within the context of whatever delusion one might prefer to entertain.

modelled as a point endowed with observable physical quantities, like mass and velocity. Any such physical model is a deliberate abstraction from the system under consideration. It is an intentional simplification designed to focus on relevant features of the system. However, as Korzybski famously said, the map is not the territory [6]. The model is not the reality.

Naturally enough, people took the physical models of Mechanics and Optics more or less literally. The map and the territory were conflated but for hundreds of years the belief that the physical objects of our everyday experience "really" do consist of pointlike particles never did much harm.

The original Newtonian Mechanics was reformulated in some rather more abstract forms, especially Hamiltonian and Lagrangian Mechanics (indispensible methods today). Simple physical models for the Electric and Magnetic phenomena were developed leading to the separate concepts of Electric and Magnetic fields, and when these phenomena were unified in Maxwell's Laws and the Lorentz Force Law, naturally enough, people took the model literally.

Let us pause for a moment and consider this intermediate step. Although it had originated as a map for certain observable facts, physicists in the second half of the nineteenth century took it as given that the Electromagnetic field was real, perhaps even that it was fundamental. This was despite the fact that we could no longer see it. In fact, this is the point where the map began to form the territory.

By the late nineteenth century, hints of an absolute mismatch between the Laws of Mechanics and the Laws of Electromagnetics had begun to surface and efforts were ongoing to reconcile the two paradigms. Although it may not have been recognised as such at the time because the dissonances were not yet fully explicated, this was a project to bring Mechanics into the framework of Electromagnetics. It was, in any event, a project destined to fail.

The Relativity Principle had been widely discussed for aeons, but with Einstein's Special Theory it suddenly became the fulcrum for a new perspective. The mismatch between Classical Mechanics and Classical Electromagnetics was clearly exposed in a simple and immediately fruitful theory based on sound observational principles alone, not some "Metaphysical" model. Minkowski soon enough found the geometrical properties of the new theory and spacetime was born[2].

Strangely, almost incredibly, everyone took the new mathematical model literally, by which I mean that its heavily Metaphysical connotations were rapidly adopted. Time and space were now the "same thing" - spacetime, and any difficulties one might experience processing this new oxymoron came to be viewed as a lack of sophistication. This word establishes an underlying, Metaphysical equivalence between the everyday concepts of extension and change[3]. I say that this rapid adoption of such strongly Metaphysical notions is strange because the particular virtue of the new approach lay in removing the Metaphysical content of earlier physical models. Instead of saying "Well done Albert, this truly is a remarkable tool.", Physics immediately dived out of its new boat back into the murky waters of "What this really means is....".

Things had become qualitatively different. The new Theory had induced a

[2]The fact that this all happened so rapidly is highly pertinent.

[3]This will be identified in the physical model as an equivalence between path lengths of wave trajectory segments in matter and light, traversed in the same time interval for any observer.

fundamental oxymoron - extension is change - that could never be understood in the usual way. It was necessary to give up physical intuition resulting in an explicit tension between the two key objectives above, technology development and curiosity's desire to understand the world.

Over the next century there were three consequences.

First, the quest for understanding was abandoned in favour of instrumentalism. Physics today is primarily an integral part of the technological society. It's business first.

Second, at a deeper theoretical level, the door was opened to increasingly exotic "ideas" until today we have serious physicists developing theories in which all of the future affects all of the past all of the time and vice versa and the reality is an 8 dimensional crystal.

Third, curiosity came to be supplanted by mystery: the strings, loops, branes and 8-D crystals are all the more fascinating *because* they are not to be understood. It would become especially necessary to forego intuition in order to communicate what Modern Physics really means. I often wonder what the point is of informing an audience about the real meaning of Theory X while asserting, in the same breath, that it cannot be understood (although this seems to be working quite well for the new industry of science media).

By contrast, I take the view that ideas that plainly cannot be understood cannot be accepted. I am all the more reluctant to drink the Koolaid because I understand precisely where we went wrong: We have been playing only one part of the game. In order to play the whole game, bald, empirical theory must be complemented by physical / constructive models that provide physical, intuitive insight into theory, inform us which amongst competing theories can be regarded as "physically reasonable" and, ideally, suggest new avenues for empirical theory to pursue.

The motivation for my work has always been to understand. Not to understand reality - that's too much to ask - but to understand Physics, and the way to do that is with a properly motivated physical model.

Overall then, it is vacuous to construct physical models of the reality. The standard against which to evaluate a physical model is physical theory not physical reality, and therefore, one deliberately models the theory, and maintains silence on the reality.

When Theory abdicates understanding, as it has done this last century, the appropriate objective is to understand Theory. This places the physical model concept on a new footing. The scope is far less ambitious than the discredited models of Physics history, but the foundations are correspondingly solid - they now lie in the physical theory itself, not some speculation about the true nature of physical reality.

We can now return to the second of the three questions posed above, which had remained unanswered. The whole question of the existence of a physical reality is (appropriately) rendered moot: It doesn't matter. One takes it as a plausible assumption for the purpose of modelling the system under consideration, which is Modern Physics not reality. One is then free to imagine whatever reality seems fit and, provided one obtains physical insight into the theories under consideration, the truth value of the working assumption here, of a unique, objective, existential physical reality, is irrelevant.

2.2 Ontology and Epistemology

Philosophy is divided into various branches and sub-branches, most of which have little to do with Physics. In Physics, one often hears about the need to avoid "Metaphysics", in the sense of *untestable* propositions about the true nature and functioning of reality. In Philosophy, Metaphysics is the branch that deals with the entire range of possible approaches to questions in the realm of existence - what it is, how it works and even why. Realism, the implicit assumption in Physics throughout most of its history, is but a stream in Metaphysics.

There is a sub-branch of Metaphysics, Ontology, that is devoted to the more limited question of "what it is" as opposed to how it works and so on. The ontology is whatever stuff is actually considered to exist, in some wider Metaphysical setting.

Epistemology is another entire branch of Philosophy, supposedly distinct from Metaphysics. It is concerned with knowing. The scope is, again, far broader than Physics, but it has long been recognised explicitly in Physics that observations lie in the knowledge camp. They are in the philosophical category of Epistemology, while things that are held actually to exist are in the philosophical category of Ontology.

Physics uses these terms with gay abandon and an astounding lack of clarity. For example, taking Classical Physics as a physical model, is the Electromagnetic field part of the epistemology or part of the ontology of the model, or both? What about the point particles? Surely they used to be part of the ontology, right? We shall see in due course that it is completely inappropriate to afford an ontological status to the point phenomena that we encounter.

There is more than a mere lack of clarity in Modern Physics. Instead, there are black and white category errors. We shall see in Chapter 5 that Quantum Mechanics runs by almost deliberately scrambling the philosophical categories. This has resulted in a curious discussion in the modern literature that is deeply flawed on all sides. It is taken as given that the observations we make are unequivocally "real"[4]. The modern discussion focusses on whether the results of observations that we don't actually make should be afforded the same status.

This all began with the valid proposition that we are limited to our knowledge of the system and *therefore* to a Physics expressed as relations between observables, which is to say Physics must be limited to epistemological models.

Yet somehow, the firm conclusion has been reached that "local realism" has failed. That's a Metaphysical conclusion, and as we shall see in Chapter 7, it turns on having taken an epistemic approach to the definition of realism, which is plainly a wrongheaded move. In both relativity theory and Quantum Mechanics, one begins by excluding Metaphysics and ends by drawing firm Metaphysical conclusions that cannot be understood.

2.3 Modelling Criteria

The reader should now have firmly in mind the idea of a physical model as a tool for understanding physical theory rather than any kind of an attempt to describe physical reality as it "really is". However, one gains more confidence in any insights that a given model may provide to the extent that it seems

[4]Epistemology is now the basis for the definition of realism in Quantum Mechanics!

"physically reasonable", which is to say that a *credible* physical reality might comport with the model.

We should prefer models that make "good sense", but these notions of "physically reasonable" and "good sense" are not well defined. What one reader may consider physically reasonable, another may not. The process of building consensus about what constitutes "physically reasonable" needs to be seen as part and parcel of evolving and improving physical models. For these reasons any author formulating a physical model should explicate how the model handles various issues that go to the idea of "physically reasonable".

The construction of a physical model is then the construction of an imaginary reality that meets certain criteria which can be freely chosen. Again, one person's criteria may not be the same as another's, but these criteria need to be explicated because they are part of the basis for evaluating any physical model.

For me, this is a leading problem both with Classical Physics and with far too many exotic modern proposals in the literature: The metaphysical content is neither identified nor kept separate from the epistemology.

With that said, the present model has been developed in order to satisfy the following criteria:

1. A clear distinction between the philosophical categories of Ontology and Epistemology will be maintained.

2. Simplicity / minimalism. A physical model must assert the existence of an ontology, but less is more. The fewer the number of kinds of entity one puts in by hand and the less detail one asserts about properties, the better. Asserting only a generic wave ontology, without referring to any details as to the kind of waves, facilitates demonstrating just how much of the usual physics is common to the entire spectrum of possible physical wave models[5].

3. Neither the subjective nor the epistemic approach to the definition of realism is allowed. This is objective realism: The model reality is considered to exist independently of observation. When an observer in the model performs a measurement, his result is a fact that is part of the epistemology of the model, not the ontology.

4. The fact that the events that occur in the world do not depend on who observes them is taken as a valid demand for coordinate independence in the Laws of Physics.

5. 3D+t. The ontology will always be defined on a 3-space, and it will always evolve in time, subject to the next criterion below, but these objective quantities must also be operationally well-defined for all observers: Consequently, the model will provide both objectivity and covariance.

6. No superluminality: No part of the ontology may exceed the characteristic velocity.

7. No infinities: In particular, integrals over all space must converge.

[5]The attentive reader may be surprised to see just how much can be shown from so little.

8. No infinitessimals (nor can point particles be put in by hand: if one needs that concept, and we will, it has to be modelled in finite ways.).

9. No action at a distance: If two parts of the Ontology interact, they must be colocated.

10. No retrocausality. The future may not affect the past. Time reversal may be valid in Mathematics, but not in a Metaphysics constructed to satisfy the dictates of good sense.

11. Local Conservation Laws will be strictly enforced. Energy-momentum must be conserved, locally and globally, at all t. No "virtual" borrowing allowed.

12. Changes in momentum and changes in energy are connected by the work integral.

13. Continuity. Both space and time are continuous in the model. Space is not pixellated and time does not proceed in discrete steps, like the frames of a movie. *"There are real movements."* - Henri Bergson [7][6]. Discretised quanta will emerge naturally from within continuity.

14. Visualisable mathematics: No tensors. No Hilbert spaces etc. In order to consider it a physical description, the model ontology will be specified on a physical space.

15. No oxymorons. The model should be described in English, as well as in Mathematics, without violence to the language.

These common sense criteria are onerous. Classical Physics obeys most but not all. Modern Physics may comply with some of these points, but it eschews the framework as a whole. Nonetheless, it will be shown that we can keep the Physics without compromising on sensemaking.

[6]Bergson's brilliant dissertation on the nature of reality led him to two short, sharp conclusions: "There are real movements" and "Matter is movement".

Chapter 3

Wave Trajectories Methods

3.1 The Wave Trajectories Representation

After finding a simple analytical tool for studying Mechanics in wave systems, which focusses on energy flows as opposed to force fields, I eventually stumbled across the same idea in Physics, where it is called the "wave trajectories representation".

Although the physicists already know the main idea, this tool is a *sine qua non* in all that follows and the meaning and relevance of wave trajectories analysis needs to be clearly understood by all readers right from the start.

The term "field energy-momentum density" is often used in Physics, and to understand this most basic concept let us go for the first of several trips to the beach. Specifically, this is an Hawiian beach in the late 1950's, with waves coming in one at a time. The reader may have seen some of the classic video footage of long board surfers riding such single, isolated wavefronts for up to a kilometre at a time. Clearly, the surfer has linear momentum, and since he picked it up from the wave, the wave carries linear momentum. As it is arbitrarily extended, it is less helpful to talk about the momentum of the wave than it is to talk about the momentum per unit of length (along the wavefront), which is to say its momentum density. The momentum density is a vector that points in the direction of propagation of the wave.

The water wave under consideration is widely distributed in one-dimension, along the wavefront, and somewhat distributed in a second dimension normal to the wavefront. It can be modelled quite well by a 1-dimensional field variable[1]. The most common choice for this field variable is the height of the wave, h, and we would write $h = f(x, y-vt)$ as the "field variable representation" of the wave, where the x-direction is parallel to the beach, the y-direction is normal to the beach and the wave crest is located at $y = vt$. Alternatively, we could describe the wave in more detail with a 2-dimensional field variable, $h = f(x, y, t)$. This allows to describe the details of the profile of the wave in the y-direction as well as how its amplitude varies from place to place in the x-direction along the wavefront.

This "field variables representation" is the standard approach to describing

[1] A function of 1 space coordinate and time. The word "field" has a different definition in Mathematics but the formal definition makes no difference.

waves and it will be familiar to most readers from Classical Electromagnetics, where two vector field variables in 3 dimensions, $\mathbf{E}(x, y, z, t)$ and $\mathbf{H}(x, y, z, t)$, are typically used to describe the Electromagnetic field. We are going to be interested in describing Mechanics problems, so the question immediately arises "How are the field variables related to the field energy-momentum density?". There are simple enough answers for both water waves (under linear conditions) and Electromagnetic waves (also a linear theory), but what about nonlinear conditions?

This is an important consideration because there are no massive particles in linear field theories because there are no field-field interactions in linear theory. If wavefields don't interact with each other, any superposition will dissipate, so field-field interactions are necessary for forming bounded soliton[2] solutions that represent matter. Any constructive model of Lorentz Transformations must therefore cover nonlinear fields, and this is a key reason why Lorentzian relativity - based on the linear Electromagnetics field theory - is regarded as unsatisfactory in comparison with Special Relativity.

Both of these examples, simple water waves and Electromagnetics, also admit a different kind of representation. Instead of describing a physical wave by some set of field variable functions on a 1, 2, or 3 dimensional space plus time, we describe the wave by a complete set of "wave trajectories".

Consider the surfer on his long board. This is old school 1950's surfing, not the modern darting and weaving about on the face of the wave. He is just following the direction of propagation of the wave at the point where he is. His velocity is the wave velocity, which is always[3] in the direction normal to the wavefront in any kind of wave theory. Over time, his path traces out a trajectory, which is the wave trajectory of a single point in the wave, as it evolves in space and time.

Now let us give the surfer a small sign on which there is written a number equal to the magnitude of the linear momentum density of the wave at the place and time where he is. As he moves along the trajectory, the local momentum density might change, so the number written on the sign becomes a function of time. Now consider a large number of surfers at a myriad places all along the wavefront. This system of wave trajectories, associated to magnitudes of the local momentum density, provides an alternate method for representing the wave.

In Electromagnetics, the field momentum density extends naturally to a field energy-momentum density because there is a particularly simple relationship, $\rho_E = \rho_p\, c$, between the field energy density (energy per unit volume), ρ_E, and the field momentum density (linear momentum per unit volume), ρ_p, which is the (scalar) magnitude of the (vector) linear momentum density, $\vec{\rho}_{\mathbf{p}}$. As always, c is the characteristic velocity, the speed at which Electromagnetic waves propagate.

These two kinds of representation are equivalent in the sense that, in a given theory, each method provides a complete description of the waves, one in terms of field variables, the other in terms of energy-momentum densities.

[2] A solitary wave, localised in 3 dimensions, which retains a self similar form as it evolves in space and time.

[3] There are exceptions to this in dispersive media, but the medium here is not dispersive.

3.2 Advantages of Wave Trajectories Methods

The wave trajectories representation has two important advantages over the field variables representation for the purposes of exploring the foundations of Physics.

First, wave trajectories are expressed directly in the Mechanics quantities, energy and momentum, so this method focusses directly on energy flows within the system. One can then study the mechanics of a wave system without knowing the formal details of the connection between field variables and energy-momentum densities. One does need to know the relationship between energy density and momentum density, but the above relationship, $\rho_E = \rho_p\, c$, will easily be shown, in Section 4.2, to hold for any kind of wave that propagates at c.

Second, energy and linear momentum are conserved quantities. Even in non-linear systems where field variables don't superpose linearly, the total energy is always a space integral over a sum of the energy densities of component waves, and the total linear momentum is an integral over their momentum densities[4].

Since there are no structural models for the massive particles in any linear field theory, generations of physicists have studied a wide variety of non-linear field models, and soliton solutions that resemble massive particles have been found, see for example [8]-[14], but acceptable solutions with all the right properties remain elusive.

Non-linear mathematics is complex, the number of possible nonlinearities is infinite and we can make only scant progress at high cost in a tiny subset of the infinity of physical models that Mathematics can provide. Furthermore, the first two foundations theories - relativistic Mechanics and Quantum Mechanics - are, in the relevant sense, linear theories[5]. One tends to focus on the colourful adjectives, "relativistic" and "quantum", but the important consideration here is that they are both Mechanics theories, expressed in the usual Mechanics variables - mass, energy, momentum, position, time - and linearity is the direct result of the conservation Laws: if we have two isolated systems of energy E_1 and E_2 respectively, then the total energy is $E_{1+2} = E_1 + E_2$ for all time, regardless of how they interact, and the same applies to the linear momentum.

One simply cannot make corresponding statements about the field variables in a non-linear field theory. Any viable candidate nonlinear field Theory would have to conserve energy and momentum, or else it is discarded, but this has to be proved every time on a case by case basis and going from the field variables to the Mechanics can be tortured. By contrast, wave trajectories analysis deals directly with Mechanics and the results apply quite generally to all field theories, linear and nonlinear.

Both Quantum Mechanics and Electromagnetics run on linear[6] wave equations, but they make very different kinds of statement. Electromagnetic wave equations are equations in field variables. They are not *a priori* Mechanics

---

[4]This will break down when we come to the N-body problem in gravity. The total energy is always a space integral over the energy density, and we will still be able to separate the system into its component waves, but we will no longer be able to talk about the energy density of each particular wave, only of the system as a whole.

[5]Of course, $\mathbf{p} = \gamma m \mathbf{v}$ is a nonlinear relationship, but the coordinate transformations are linear in the coordinates and energy-momentum superposition applies.

[6]Meaning essentially that if A and B are solutions, then so is $\alpha A + \beta B$ where α and β are constants.

equations. Quantum Mechanical wave equations are Mechanics equations, but they are written in a "wavefunction" which is not, *a priori*, a field variable in the sense of a representation of a physical field. The relationship, if any, to physical field variables is indeterminate[7].

Quantum Mechanics is a highly effective workaround. By expressing relationships directly in terms of the Mechanics, it avoids the difficulties inherent in dealing explicitly with the nonlinear physical structures of particles, but it also violates many of the criteria in Section 2.3 (which is not any criticism of the Physics).

The method used here is also a workaround. It allows us to circumvent the nonlinear details inherent in the field variables representation of any viable candidate field model. One simply applies the Laws of Mechanics to the field energy-momentum density in wave systems where propagation is at the characteristic velocity. It reproduces all of the basic Physics theories without compromising on the criteria and, quite separately, it makes considerable progress with the physical structure. One does not need to know any wave equation. One does not need to know the relationship between field variables and energy-momentum. One does not need to know the details of nonlinearities. The results apply to all wave systems, linear and non-linear, that share the characteristic velocity.

The next Chapter provides the first two important results, relativistic Mechanics and Lorentz Invariance, which are simply the consequences of Conservation Laws in systems of luminal waves (*i.e.* waves that propagate at c).

[7] Chapter 5 argues that no such relationship exists in Ordinary Quantum Mechanics.

Chapter 4

Mechanics

4.1 Why Bother with Special Relativity?

My daughter once asked me what I was doing in Physics. She was 13 at the time. Rather than talking about nonlocality, I told her that people had been arguing about Special Relativity for over a century. There were two sides to the argument but they could never agree because they are arguing about different things. One side argues about the mathematics, you know the beauty and symmetry of the formalism, its empirical power and so on. The other side argues about the philosophical consequences, you know like paradoxes, the idea of "now" and so on.

At this point she interrupted me, saying "But surely we need to have both.".

Most physicists presume that any non-professional providing any new insight into relativistic Mechanics is automatically a barbarian set to destroy Western civilization. Cognisant of his colleagues' extreme sensitivity, one Physics professor suggested to me that I should assume Special Relativity and just use the luminal wave mechanics to explain all the relativistic phenomena. He said the content would be more palatable, and it would still be important. The situation is so fraught with preconceptions on all sides, that one might well ask "Why Bother?". There are two reasons.

First, the orthodoxy has been unable for the last 85 years to deal with the well known tensions between Ordinary Quantum Mechanics and Special Relativity that grew out of the EPR paper [15]. These tensions can only be removed by developing the theory of Lorentz Invariance without appealing directly to the relativity principle. By developing a constructive theory of Lorentz Invariance from the wave postulate, the relativity principle is recovered as a result for observers in the model (*i.e.* us). It remains as a sound epistemological principle but its problematic Metaphysical connotations are removed. On the present understanding of relativity of simultaneity in particular, when observers in relative motion disagree about the facts, for example the rates of their clocks or the temporal order of spacelike separated events, there is no underlying fact of the matter. The wave approach to Lorentz Invariance allows us to remove that specific dictum without altering any of the predictions or modifying any of the valuable mathematical symmetries. Since it will no longer be incumbent upon us to think of relativity of simultaneity as some kind of Metaphysical truth, the conflicts between Ordinary Quantum Mechanics and relativity theory will dissolve.

Second, one typically thinks of Special Relativity as a constraint to be imposed on the other foundation theories. For example, after physicists developed Ordinary Quantum Mechanics as a standalone theory, it then had to be rendered compatible with Special Relativity. By contrast, once Lorentz Invariance is developed directly from the wave postulate one is immediately led to all the usual quantum mechanical wave equations on one hand and to the physical implementation of a curved 4-space in gravity on the other.

Nonetheless, let me be totally clear on this point. The re-derivation is not about "disproving" Einstein. That way of looking at my work is wholly inappropriate because Einstein's principled development of the formalism remains valid. The constructive development here replaces Lorentzian relativity, not Einstein.

The present Chapter is first about ensuring that we may have our cake (objectively) and eat it (relativistically) because, as my daughter so astutely pointed out, we are indeed entitled to both. Then it is about building the platform for the other theories to be modelled in later Chapters.

Rather than beginning with postulates, developing coordinate transformation and then finally a mechanics of particles, the path pursued here is to study the basic Mechanics of waves propagating in a medium at a characteristic speed, c[1]. This subject is as simple and elegant as Special Relativity. It delivers a mathematical simplicity where the traditional approach is fully included, but which goes much further. It unites Newton and Einstein[2], it reconciles objectivity and covariance and it will take us to Quantum Mechanics and Gravity.

It is also inherently equipped with a preferred frame, hidden from Physics but revealed to philosophers. Experiments obey the Relativity Principle and the usual Einsteinian analysis remains valid but the idea of an observer independent moment, the idea of "now", is recovered, and all the relativistic paradoxes are removed.

4.1.1 The Relativistic Paradoxes

Before proceeding, it is appropriate here to survey briefly the reasons why the Special Theory has caused so much angst. How did we get into a century long debate, and exactly where lies the mathematical root of the problem?

It is assumed that the reader knows well enough how to derive Lorentz Transformations from two postulates, the first being the relativity principle - all experiments work the same for any inertial observer, the second being the constant (and isotropic) speed of light for all observers.

Note that the second postulate appears, on the face of it, to be manifestly incorrect - if the light is moving relative to you at c, and I am moving relative to you then surely (at least it might seem self-evident) the light cannot also be moving at c relative to me.

What most people have not focussed on is the third ingredient in Einstein's derivation, namely the Einstein clock synchronisation protocol. It is sufficiently

[1] Waves that propagate at c will be referred to throughout as luminal waves.

[2] This Chapter has the title "Mechanics" rather than "Relativistic Mechanics", because there is no non-relativistic version - the only Mechanics in wave systems is relativistic Mechanics.

obvious that observers who use the Einstein clock synchronisation protocol will all find the same speed of light so the second postulate is actually redundant in Einstein's derivation, because it is guaranteed by virtue of the Einstein protocol.

If this is not obvious, consider a light source at the midpoint between two clocks that emits a short flash of light in all directions. Using that light flash to synchronise the two clocks means by definition that the light took the same time to travel from the source to each of the clocks. Since the distance was the same in both directions, we have the same distance travelled in the same time and the result is that we just defined the velocity of light to be the same in both directions.

Many people may also be unaware that the Special Theory can also be derived (up to the value of the constant, c) from the first postulate alone [16], although it's quite a tedious mathematical exercise, which still involves making an assumption about clock synchronisation.

Personally, I prefer to reason as follows. If one takes the relativity principle literally then observers in all inertial frames must use the same procedure for synchronising their respective clocks. Einstein synchronisation is one procedure that satisfies the constraint, and the rest of the Special Theory follows as usual. However, in the absence of evidence for anything that moves faster than light, it turns out that all known protocols that satisfy this criterion are functionally equivalent to the Einstein protocol. The reader will be familiar with the usual analysis of slow transport as a synchronisation method. The other main possibility, using subluminal projectiles, is considered in Appendix A. Consequently, if one wants a relativity principle, which one most certainly does, one inevitably adopts Einstein synchronisation and the Special Theory then follows from the relativity principle alone.

Einstein would surely have known that his clock synchronisation protocol implies his second postulate. Perhaps he considered that a Theory based on a clock synchronisation protocol would seem to be based on a free human choice of a convention regarding coordinate systems. This would not look good in a Theory of coordinate independence! Alternatively, he could simply appeal to the then recent Michelson-Morley experiment showing the invariance of c, and put that in as a second postulate. The decision whether to tie the Theory to the choice of a convention or an experimental result would not have been difficult[3], but it has generated more than a little confusion.

Unfortunately, there is a well known loophole in the experimental evidence: The Michelson-Morley experiment only covers the two-way speed of light, from a given point out to a mirror and back to the same point. There's no clock synchronisation involved and while the two-way velocity of light does not depend on clock synchronisation, measuring the one-way velocity (obviously) does.

Putting on the philosophy hat for a moment, it is equally obvious that, unless the second postulate is thought of as an "objective truth", the Einstein protocol would have to be interpreted as introducing space dependent simultaneity errors between different observers' sets of clocks: By setting the clocks "wrong", the 2nd postulate is enforced, but relativistic simultaneity is then a mere appearance only for observers who use the Einstein protocol, not a fact for all observers in general and certainly not an objective fact in nature.

[3]It is hard to read Einstein's paper, [17], without thinking that this issue was very much on his mind.

This dilemma is what lies at the root of every (thoughtful) criticism of Special Relativity. Pushing forward to develop the Theory without resolving the circular reasoning involved at the very first step induced paradoxes. I shall mention just a couple of the classical paradoxes shortly.

Switching hats, for theorists the only relevant consideration in Physics is that which is observed, so they come naturally to the idea that the only simultaneity that "exists" is the simultaneity that appears to a certain observer. If two observers, considering the same pair of events, disagree about the temporal ordering this is not a theoretical problem as long as each observer's individual reality is internally consistent.

Switching hats again, from the philosophical point of view this is known as "subjective realism". What you see, what you measure corresponds to the "real" (for you). Note that in subjective realism, the realism is Epistemological, not Ontological. Meanwhile, the Special Theory is explicitly based on the notion that any observation is made relative to the observer's coordinate system. As one rather senior (but anonymous) physicist put it to me:

"The problem with Special Relativity is that it assigns an objective status to observations that are inherently relative."

While that may have philosophical appeal, the question on the Physics side is: "Is the Theory internally consistent, or is it not?". It is not. To be precise, the orthodox interpretation of Special Relativity is internally inconsistent, not the Theory itself.

To see that this subjective approach to realism is internally inconsistent, even when considered entirely from a Physics perspective, let us examine the various levels of paradox in the Theory. There are many to choose from but, for the purpose here, it will be sufficient to consider just three.

Paradox I: Observers in different conditions of motion both find that the other's clock is running slow.

A philosophical dilemma, yes, but let us put the Physics hat on, and leave it on: This is no problem at all. They are both right.

Paradox II: I assume all readers know about the twins "paradox"[4], but consider the related situation where the twin who left Earth in a spaceship, after making a long journey at some extreme relativistic velocity, finds himself at a distant planet that happens to be comoving with the Earth. He begins to orbit the planet in his spaceship. Sometimes he is moving towards the Earth, at others he is moving away from it. Although his velocity relative to Earth is non-relativistic (it can be made arbitrarily small if the distance to Earth is correspondingly large), he considers that his brother is dead some of the time but at other times his brother is still alive. This is a result of the space dependence of time in the Lorentz Transformation, $\gamma vx/c^2$, where the sign of v is changing as he orbits the planet, but not the sign of x.

The Physics point of view is that they are in spacelike separated regions[5] and any notions the distant twin may entertain about the life and death of his twin on Earth cannot be observed. At this point, even with my Physics hat

[4] There is no paradox in it at all: for all Lorentz observers the twin who goes away and comes back travels faster on average, so he experiences less total elapsed time.

[5] Regions that are sufficiently far apart that events (like "brother dies") at one place and time cannot be connected by light signals to events (like "thinking brother is dead") in another place and time.

securely on, I begin to wonder. The theorist is telling me that while the reality that this single observer can infer from the Theory is inconsistent, it is not a problem provided that he is not confronted with it.

So, let us confront him with it.

Paradox III: The double Bell paradox. I shall describe this paradox in some detail after the discussion of spacelike causal correlations in Chapter 7. Suffice it to say that, with the experimentally verified instant causal correlations at a distance as predicted by Quantum Mechanics, the twins in paradox II can indeed be connected. The standard escape clause is then: Yes, but we can't send information between them. This escape clause is voided by the double Bell paradox.

Without transmitting any information *per se* between the twins, the relativity formalism implies the existence of forbidden causal temporal loops in which an actually observed fact causes a chain of events that changes the same fact.

In the double Bell paradox, a spin that is observed to be in the "UP" state is driven by the loop into the "DOWN" state at the same instant in which it is observed. This is more than "counterintuitive", it is forbidden. We could connect this signal to the poison vial in the Schroedinger's cat thought experiment and experience a cat that is both alive and dead *while we are all looking at it.*

Thus, we now have cases where an individual observer participates in an impossible contradiction. There is no wiggle room left. Shimony's famous "peaceful coexistence" word salad[6] [18] cannot avoid the black and white fact that subjective realism is internally inconsistent (although we shall observe in Section 7.10 that there is no issue for Special Relativity itself.).

Whichever side of the debate one favours, all of us can benefit from considering how the basic Mechanics of waves propagating in a medium generates Lorentz symmetry, without provoking internal inconsistencies and paradoxes.

4.2 The Basic Mechanics of Wave Systems

The usual classical field approach to Mechanics in wave systems begins by choosing wave or field equations. Any analysis is immediately limited to the mechanics of one particular kind of wave system. We would identify various field variable solutions to the chosen equations. Field energy and momentum densities must be induced from these field variables, and after evaluating the spatial integrals of the energy and momentum densities we would arrive at expressions for the momenta and energies of the wave solutions and we could begin the mechanics.

However, on the one hand, we do not usually know the right classical wave equation for a given physical situation and working one by one with individual nonlinear wave equations is massively challenging while, on the other, there is a great variety of Lorentz covariant wave equations and relativistic classical field theories, all of which feature the characteristic velocity. This suggests that relativistic Mechanics is a general feature that many luminal wave systems have in common. It is one thing to notice this anecdotally, as many have, but quite another to prove that, when the field energy-momentum in a wave system is constrained to propagate at c (*i.e.* luminally), the system necessarily displays the usual relativistic Mechanics.

[6]It should be acknowledged that Shimony eventually recanted this anthropocentric line of thinking. [19].

Therefore, instead of taking the usual fields approach to mechanics let us take a *Mechanics* approach to *fields*, applying the basic Principles of Mechanics directly to a field energy-momentum density that propagates at c.

The following basic framework for Mechanics is common to Newtonian, relativistic and Quantum Theories of Mechanics. This framework is regarded as universally accepted:

1. The momentum of an object is defined as the product of its inertia times its velocity. Similarly, the field momentum density is defined as the product of inertia density and velocity.
2. Momentum is conserved. Field momentum is locally conserved.
3. The force acting on an object is, by definition, equal to its rate of change of momentum.
4. The resulting change in the energy of the object is given by the work integral.
5. Energy is conserved. Field energy is locally conserved.

The "objects" under consideration throughout the development of the physical model will be luminal wave objects, defined as follows:

A luminal wave object is some set of functions on a 3-space plus time[7], which together induce a field momentum density, $\vec{\rho_{\mathbf{p}}}\,(x, y, z, t)$, that propagates luminally according to a unique unit wave vector, $\hat{\mathbf{k}}\,(x, y, z, t)$ and whose spatial integral, $\int\int\int_{-\infty}^{+\infty} \vec{\rho_{\mathbf{p}}}(x, y, z, t)\, dxdydz$, is finite.[8,9,10] The unit wave vector, $\hat{\mathbf{k}}\,(x, y, z, t)$, is just a unit vector in the direction of propagation. The relationship between the wave momentum and the corresponding wave vector, $\mathbf{k}$, will be developed, not assumed.

An additional principle will be required for later Chapters, which it is appropriate to state here: The Principle of Local Action means that wave objects, as defined above, may interact with each other only in regions of space where they overlap: if two objects' momentum densities are non-zero at some place and time, they can interact at that point.

Given the definition of momentum, the field momentum density can be written as follows:

$$\vec{\rho_{\mathbf{p}}} = \rho_m c\, \hat{\mathbf{k}}\,, \tag{4.1}$$

where ρ_m is the wave inertia density, not to be confused with the mass density.

We are interested in the mechanics of systems that comprise multiple wave objects, and the subscript m is used for inertia density mainly because the subscript i will be used to label the wave objects in multi-object systems, but

[7] *i.e.* the spatial distributions of field variables.

[8] In addition to inducing the field momentum density, the space functions that define wave objects in a nonlinear field theory may also act as sufficient causes for any interactions that there may be.

[9] Note that infinite plane waves are not wave objects. See criterion 7 in Section 2.3.

[10] Neither the propagation of the space functions nor their relation to the linear momentum density are specified here. This allows for wave objects with intrinsic field angular momentum and, more generally, the definition accommodates two kinds of internal evolution, via the internal movements of an otherwise invariant set of functions and via their individual time evolutions.

also because, for any subluminally moving system, the space integral over all the inertia densities in a wave system will be found to correspond to the usual relativistic mass.

Let us begin with non-interacting systems, where the wave objects are not interacting with each other. The next Section focusses on the special case where each object's unit wave vector, $\hat{\mathbf{k}}\,(x, y, z, t)$, is a constant vector, independent of x, y, z and t. The momentum density distribution of such wave objects moves through space in a self-similar form at c. We shall refer to this special kind of wave object as a light flash.

4.3 Application to Light

Consider a source that simultaneously emits a set of N light flashes in various directions. The development here can be applied to any kind of light flashes, including individual photons, short segments of laser beams, or collimated beams in general, monochromatic or not. We require only that each flash propagates at c, carrying a finite linear momentum in a well-defined direction in space.

Let the i^{th} light flash carry linear momentum $\mathbf{p}_i = \iiint \vec{\rho}_{\mathbf{p}i}\, dx\, dy\, dz$. The total wave inertia of the i^{th} light flash is $m_i = \iiint \rho_{mi}\, dx\, dy\, dz = p_i/c$, where $p_i = |\mathbf{p}_i|$ is the magnitude of the momentum of the i^{th} light flash, also called the 'scalar momentum':

$$p_i = m_i\, c\,. \tag{4.2}$$

Along with the definition of wave momentum, (4.1), this is the basis for the entire physical model, which is essentially a consistent application of the basic Mechanics principles, using (4.2) in place of the familiar $p = mv$, where the speed v is a variable[11]. Note again that the inertia, m_i, of a wave propagating in a well-defined direction in space has, *prima facie*, nothing to do with the mass of a particle and we use the symbol m_i for the wave inertia because, unless they ALL propagate in the same direction, the total inertia of a set of N waves will be found to correspond to the usual, relativistic particle mass.

The time differential of (4.2) is:

$$\frac{dp_i}{dt} = c\,\frac{dm_i}{dt}\,. \tag{4.3}$$

Having fixed the propagation speed, c, changes of the scalar momentum are thus associated with changes of the wave inertia. It will become clear in Section 4.7 that the inertia changes we will be discussing are in fact frequency changes. Such changes may be due to a change of observer or they may be physical changes due to any forces that are acting on the wavefield.

If a force acts on a light flash then, since (4.3) is the force component parallel to the light flash's motion, the work integral is:

$$W = \int_{p_\mathrm{s}}^{p_\mathrm{f}} \frac{d\mathbf{p}}{dt}\cdot d\mathbf{s} = \int_{m_\mathrm{s}}^{m_\mathrm{f}} c\,\frac{dm}{dt}\,cdt = (m_\mathrm{f} - m_\mathrm{s})\,c^2\,, \tag{4.4}$$

[11] While (4.2) only applies to light flashes, where the unit wave vector is constant, this form provides the simplest notation conceivable, which it is hoped no reader will find intimidating. A simple trick will allow us to retain the same notation when we move on to more complicated wavefields.

where subscripts s and f refer to the words 'start' and 'finish'. The radiation reaction force that acts on a light flash reflected by a moving mirror is an example that highlights the role of the work integral in a basic Mechanics calculation (see below). According to the fourth basic principle, the work done equals the energy change, and we may assume that a light flash that has zero momentum requires zero energy, so the energy of the i^{th} flash is:

$$E_i = m_i\, c^2 = cp_i\,. \tag{4.5}$$

According to the second basic principle, momentum is conserved so the total momentum of a set of N wave objects is given by the vector sum over their momenta:

$$\mathbf{P} = \sum_{i=1}^{N} \mathbf{p}_i\,. \tag{4.6}$$

As far as possible, the notation used for the Mechanics quantities will be as follows: Upper case letters (*eg* $\mathbf{P}$) are used to refer to whole wave systems, lower case letters (*eg* $\mathbf{p}$) refer to individual waves, the greek ρ with a lowercase subscript refers to a scalar density, and $\vec{\rho}$ for a vector density. This notation distinguishes the relationships in the model from the usual Physics equations, which will be written in the usual way, typically with lowercase letters for the Mechanics quantities. The main exception to this is the energy, which is always an uppercase E, both in the model and in the usual Physics.

Suppressing the summation range henceforth, we write the total inertia as $M_e = \sum_i m_i$. The total energy of the set is then:

$$E = \sum_i cp_i = M_e c^2\,. \tag{4.7}$$

According to the (first and second) basic principles, the velocity of the centre of inertia of a system of objects is the inertia weighted average velocity, $\mathbf{V} = \sum_i m_i \mathbf{v}_i / \sum_i m_i$, so that:

$$\mathbf{V} = \frac{\sum_i \mathbf{p}_i}{M_e} \Rightarrow \mathbf{P} = M_e \mathbf{V}\,. \tag{4.8}$$

Let us pause a moment. Although we have barely begun, we have already taken steps - in the last three equations - that might be disconcerting to many an alert physicist. No one would have a problem with these equations in the context of light flashes, but I have already pointed out above that Mechanics is inherently linear even in nonlinear systems, and we will be using these same equations for matter as well as for light in the following Sections. Let us consider what such equations will mean when we turn our attention to matter.

Imagine that you have in front of you a golf ball. It is made of energy, as you already know, which propagates at c. The fundamental proposition (4.6) is that the momentum of the golf ball is given by a (vector) sum over all the constituent fields' momenta, which is to say that we evaluate, for each constituent field, the integral over all space of its field momentum density and then take the sum over all constituent fields (pertaining to every electron, proton, neutron and so on in the golf ball) to arrive at the momentum of the golf ball.

If you went to high school, this may all seem self-evident, but what it amounts to is the conservation of momentum between wave components and wave groups, an idea that many physicists would find problematic. What they have in mind is a well known standard analysis in Ordinary Quantum Mechanics of the relationship between the components of a wave group and the group velocity. It comes from a dispersion relation, where the propagation velocity depends on the frequency. I have seen a professor of Physics blithely import this non-relativistic dispersion model into the relativistic case, which is problematic to me because there is no dispersion in the usual relativistic Theory.

In the dispersive model, the wave components propagate in the direction of the group velocity at different speeds and they do not interact with each other. Since each component wave is moving at a different speed to the particle it is forever sliding backwards or forwards past the particle and the relation between components' momenta and the particle momentum is immediately unclear.

We shall return to this way of thinking about a physical wave interpretation, and why it does not work, in the next Chapter but for now please note that it could never produce a viable physical model of a massive particle that satisfies the criteria in Section 2.3. The wave components are continuously moving with respect to the particle and one therefore requires physical waves that extend to infinity, violating criterion 7. That's not a satisfactory physical model.

In our golf ball, at any instant of time every bit of energy is jiggling about very rapidly relative to the centre of inertia of the golf ball, however every wave trajectory remains within the immediate vicinity of the observable golf ball. One cannot fail to demand conservation of momentum between the underlying wave energy and the golf ball.

As far as nonlinearity is concerned, if the momenta of two disjoint waves are $\mathbf{p}_1$ and $\mathbf{p}_2$ respectively then no theorist would object to the idea that the momentum of the 2-wave system is forever $\mathbf{P} = \mathbf{p}_1 + \mathbf{p}_2$.

Let these two disjoint waves propagate towards each other and collide. Either they interact with each other or they don't. Whatever happens, after the collision the linear momentum of the resulting system is still the same $\mathbf{P}(after) = \mathbf{p}_1(before) + \mathbf{p}_2(before) = \mathbf{P}(before)$.

Suppose that they did interact. Long after the collision, there may be n disjoint waves, propagating away from each other, in which case:

$$\mathbf{P}(after) = \mathbf{p}_1(before) + \mathbf{p}_2(before) = \sum_{i=1}^{n} \mathbf{p}_i(after)$$

Alternatively, the collision may result in a bounded subluminally moving system, like our golf ball. For sure, the fields cannot be the same as they were before. For one thing, they are no longer propagating in straight lines. According to Newton, an object will continue moving in a straight line unless acted on by a force. Therefore, the energy in a subluminally moving system that persists in a self-similar form, like our golf ball, is constantly being acted upon by internal forces. There are no field - field interactions in a linear theory to provide any such internal forces, and it is well known that there can be no subluminal particles in linear pure field models. Therefore, any pure field model of the massive particles will necessarily be a nonlinear field model.

Force is defined as the rate of change of momentum, so forces in nonlinear field models are by definition exchanges of linear momentum between wavefields: When a force acts on a luminal wavefield it modifies the field momentum density, which is to say it modifies the field itself. This is not the usual notion of what

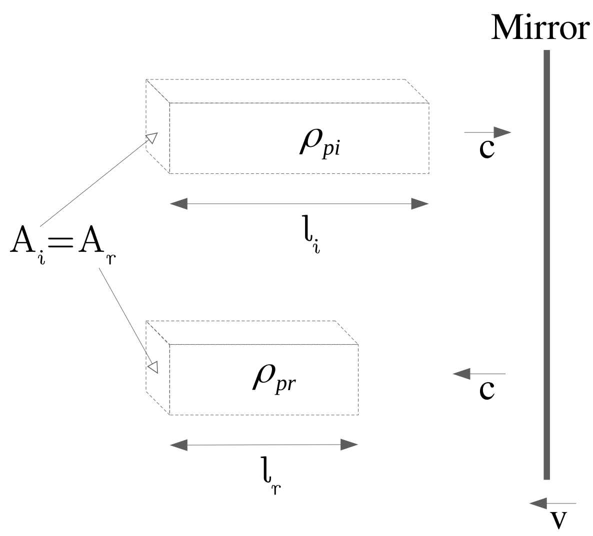

Figure 4.1: *Light reflected by a moving mirror*

happens when a force acts on a golf ball. We tend to think that the force changes the condition of motion of the golf ball - the golf ball accelerates - but it does not change the golf ball itself. However, if our golf ball is composed entirely of fields then the only way to act on the whole golf ball it is to act on the constituent fields, which materially changes the fields, which obviously materially changes the golf ball itself.

One consequence is that the image of the golf ball accelerates, another is that its mass changes, and others are length contraction and time dilation. We need to analyse how this works, but the golf ball is rather complicated, so we begin by analysing what happens to light waves when acted on by a force.

Just as we can manipulate the momentum of a golf ball by hitting it with a golf club, we can also manipulate the momentum of light by hitting it with a moving mirror. Consider first the case of light impinging with normal incidence on a moving mirror, as shown in Figure 4.1.

If we were given Special Relativity, we could simply assume that the velocity of light in the observer's frame is c, but we are not assuming Special Relativity. What follows will all turn out to work relativistically, but for the moment all we have assumed is that light propagates at c in some medium. Let us begin with a particular observer, at rest in the medium, for whom the speed of light is c[12].

Choose coordinates such that the plane of the mirror is parallel to the yz plane. Let the mirror approach the observer with velocity $\mathbf{v} = -v \ \hat{\mathbf{x}}$. During the reflection, the linear momentum of the incident light field is absorbed by the mirror so the rate of change of momentum of the incident field is just the rate at which momentum impinges on the mirror, $dp_i/dt = -A(c+v)\rho_{pi}$, where A is the cross-sectional area of the volume that contains a momentum density

[12]This much is agreed on all sides of the debate: There exists at least one inertial frame in which the speed of light is c in all directions.

ρ_{pi} and $c + v$ is the closing speed of the light onto the mirror[13]. Similarly for the reflected beam.

The result of this calculation, as shown in the first part of Appendix B, is that the reflected momentum is enhanced by the factor $(c+v)/(c-v) = (1+\beta)/(1-\beta)$, where $\beta = v/c$ as usual. A similar method applies to the case of non-normal incidence[14] where we get the more general results:

$$\frac{p_r}{p_i} = \frac{c + v\ cos\theta_i}{c - v\ cos\theta_r} = \frac{1 + \beta\ cos\theta_i}{1 - \beta\ cos\theta_r} \tag{4.9}$$

and

$$\frac{p_r}{p_i} = \frac{sin\theta_i}{sin\theta_r}\,,$$

Where θ_i is the angle between the incident wave vector and the normal to the mirror, similarly θ_r for the reflected wave. These relations are shown in the second part of Appendix B. Three technical points are worth noting.

First, this is a Newtonian method, but these are the usual, relativistically correct results, as shown in the third part of Appendix B.

Second, in the Special Relativity calculation, light reflected from a moving mirror is dealt with using two coordinate transformations, from the observer's frame to the mirror frame and then another from the mirror frame back to the observer's frame. Each such transformation is described by a relativistic Doppler shift and aberration operation. Not only does (4.9) factor into the sequence of two relativistic Doppler shift and aberration operations, but also the factorisation is unique because these phenomena do not involve clock synchronisation [20].

Third, the rate of change of momentum of the light field (*i.e.* the force) while being acted on by the mirror is proportional to the energy density of the field.

4.4 The Relativistic Momentum Equation

Let us begin with a couple of illustrations that motivate the relativistic momentum equation, and then derive it rigorously from conservation of momentum alone.

Consider, just for the moment[15], a force field that acts on the energy density of a target field, so that the rate of change of the target field momentum density has the simple form:

$$\frac{d\vec{\rho}_{\mathbf{p}}}{dt} = \alpha \mathbf{F} \rho_p\,,$$

where α is some scalar constant, $\mathbf{F}$ is the (vector) force field, $\vec{\rho}_{\mathbf{p}}$ is the vector momentum density of the field and ρ_p is the scalar momentum density, which is proportional to the energy density via $\rho_E = c\rho_p$, as follows from (4.5) in the previous Section. Note that this force is independent of the direction of propagation of the target field. For light flashes, we could implement this kind of force with moving mirrors.

[13] Both c and v are referred to the same observer: this is not a case of relativistic velocity composition.

[14] In this case the cross-sectional areas of incident and reflected waves are not in general equal.

[15] This is an assumption that must be removed in due course.

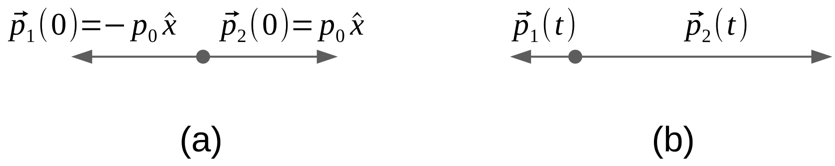

Figure 4.2: *Equal and opposite parallel momenta acted on by a force* $\mathbf{F} = \alpha p \hat{\mathbf{x}}$

For the first illustration, consider the situation shown in Figure 4.2a where a pair of waves of equal and opposite momenta is acted on by a force field in the x-direction. Then $d\mathbf{p}_1/dt = \alpha \mathbf{F} p_1$ and $d\mathbf{p}_2/dt = \alpha \mathbf{F} p_2$, so that $\mathbf{p}_1(t) = -p_0\ e^{-\alpha F t}\ \hat{\mathbf{x}}$ and $\mathbf{p}_2(t) = p_0\ e^{+\alpha F t}\ \hat{\mathbf{x}}$, as shown in Figure 4.2b. From these relations, the total energy and momentum of the 2-wave system and the velocity of its centre of inertia can be determined:

$$\begin{aligned} E_0 &= 2p_0 c\ , \\ E = c(p_1 + p_2) &= p_0 c\ (e^{-\alpha F t} + e^{+\alpha F t})\ , \\ \mathbf{P} = \mathbf{p}_1 + \mathbf{p}_2 &= (p_0 e^{\alpha F t} - p_0 e^{-\alpha F t})\ \hat{\mathbf{x}}\ . \end{aligned}$$

Define β and γ in the usual way:

$$\beta = \frac{V}{c} = \frac{P}{M_e c} = \frac{Pc}{E} = \frac{e^{+\alpha F t} - e^{-\alpha F t}}{e^{-\alpha F t} + e^{+\alpha F t}}\ ,$$

where we have used (4.8) and (4.7), and:

$$\gamma = \frac{1}{\sqrt{1-\beta^2}} = \frac{e^{-\alpha F t} + e^{+\alpha F t}}{2} = \frac{E}{E_0} = \frac{M_e}{M_0}\ ,$$

where M_0 is the value of M_e at $t = 0$, and finally:

$$\mathbf{P} = \gamma M_0 \mathbf{V}\ .$$

This is the relativistic momentum equation.

For the second illustration, consider the pair of waves shown in Figure 4.3a, where the initial condition is oriented transverse to the force field. After application of the force field for some time period, t, we arrive at the situation in Figure 4.3b. We shall see how to calculate $a(t)$ explicitly later on, but we don't need it for the moment due to the symmetry.

$$\begin{aligned} E_0 &= 2p_0 c\ , \\ E = c(p_1 + p_2) &= 2c\ p_0 \sqrt{1+a^2}\ , \end{aligned}$$

$$\begin{aligned} \mathbf{P} = \mathbf{p}_1 + \mathbf{p}_2 &= 2ap_0\ \hat{\mathbf{x}}\ , \\ \beta = \frac{V}{c} &= \frac{a}{\sqrt{1+a^2}}\ , \\ \gamma = \frac{1}{\sqrt{1-\beta^2}} &= \sqrt{1+a^2}\ , \end{aligned}$$

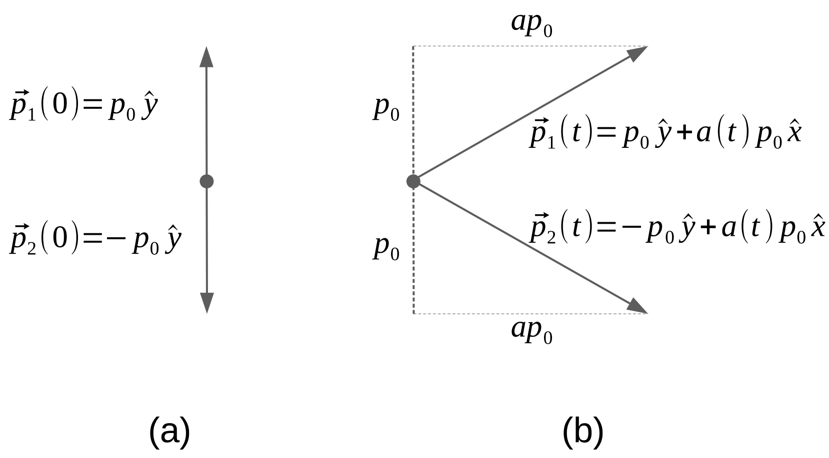

Figure 4.3: *Equal and opposite transverse momenta acted on by a force* $\mathbf{F} = \alpha p\hat{\mathbf{x}}$

and, using the same basic Mechanics relations, we arrive again at the relativistic momentum equation:

$$\mathbf{P} = \gamma M_0 \mathbf{V}\,.$$

Hopefully, working these special cases has provided the reader with an intuitive understanding of where the relativistic momentum equation comes from in systems of waves.

In order to derive the general case rigorously, we shall consider as initial condition a system comprising an arbitrary number of wave momenta of various magnitudes and directions, subject only to the condition that the total vector momentum sums to zero. This could be a system of light flashes or (in the most general case) a nonlinear field system executing bounded motion, like a golf ball initially at rest in our observer's frame of reference.

Both these cases are the same except that on the one hand we are analysing the momenta of light flashes while on the other we are analysing the field momentum densities of the golf ball's constituent fields.

In order to use the same notation for both, let us convert the momentum densities of the golf ball's constituent fields into momenta, as follows[16]: Divide the entire space into small regions of dimension $\delta x = \delta y = \delta z = \delta l$, where δl is sufficiently small that the field momentum density of any given wavefield of any given proton, electron, neutron *etc* can be regarded as constant in each incremental region. For the j^{th} field in the k^{th} region, we have $\mathbf{p}_{jk} = \vec{\rho}_{\mathbf{p}j}(\mathbf{r}_k)\delta l^3$, where $\mathbf{r}_k$ is the position vector to the centre of the k^{th} region. Finally relabel all these momenta such that $\mathbf{p}_{jk} \to \mathbf{p}_i$.

For the i^{th} wave momentum what we have been assuming up to now is:

$$\frac{d\mathbf{p}_i}{dt} = \alpha \mathbf{F} p_i\ ,$$

[16]This is the "simple trick" referred to in footnote 8 above.

so that for the system as a whole:

$$\frac{d\mathbf{P}}{dt} = \sum_{i=1}^{N} \frac{d\mathbf{p}_i}{dt} = \alpha \mathbf{F} \sum_{i=1}^{N} p_i \, ,$$

which implies:

$$\frac{d\mathbf{p}_i}{dt} = \frac{p_i}{\sum_{i=1}^{N} p_i} \frac{d\mathbf{P}}{dt} = \frac{p_i}{M_e \; c} \frac{d\mathbf{P}}{dt} \, .$$

There is a technical issue with this relationship, especially since we assumed that the same force field acted on all elements of the set. That's actually a very strong assumption. Moreover it is an assumption about Dynamics that was introduced into a discussion that is leading us toward coordinate transformations that are manifestly kinematic. We need to remove it. Therefore, let us write the above equation down a little differently:

$$d\mathbf{p}_i = \frac{p_i}{M_e \; c} d\mathbf{P} \, . \tag{4.10}$$

This has a distinct meaning. Without making any reference to forces or rates of change of momentum, it represents the connection between wave systems before and after the system as a whole is given an incremental momentum boost and transients have died out. Equation (4.10) will be referred to from here on as the incremental momentum boost generator. What the last step did was to remove all the preceding Dynamics content and replace it with the kinematic relation between systems before and after an incremental boost.

For the systems of light flashes in Section 4.3, an incremental momentum boost amounts to considering two otherwise identical light sources in slightly different conditions of motion, each of which emits a set of N light flashes. We don't need to consider how or why the light sources have different conditions of motion and there is no force involved.

For matter it amounts to asserting a one to one correspondence between the components of the before and after systems, such that each component transforms between frames independently of all the others, according to the same rules.

This is not a strong assumption at all, in fact it is implicit in the idea of having *any* form of coordinate transformation in the first place. In Special Relativity, for example, one routinely equates a Lorentz boost with a change of observer, which is to say that a single observer's descriptions of a particle before and after it is boosted are connected in the same way as are the descriptions of two observers in relative motion when considering either of the before or after particles.

Nor is the incremental momentum boost generator (4.10) in any way a new assumption. Although it was motivated above using a couple of simple illustrations, it is easily derived formally from the relativistic Doppler shift and aberration results (see Appendix C), but we have already derived these results[17] above, at (4.9). Moreover, these are observable facts that don't depend on

[17] The combination of two sequential Doppler shifts and aberrations was derived. As a technical point, when I wrote [21] I was not completely satisfied that the factorisation of the two sequential operations into the individual Doppler shifts and aberrations must be unique, but this follows from the results of Puccini [20] and Selleri [22]: These phenomena don't depend on clock synchronisation and are therefore independent of the coordinate transformations. An

any theory and there are also other coordinate transformations besides Lorentz Transformations that correctly predict the same Doppler shift and aberration relations.

Overall, all we require is that some kind of coordinate transformations exist, not that they are Lorentz Transformations. With the Dynamics content now removed[18], the relativistic momentum equation follows easily from the incremental momentum boost generator. The proof is both simple and pivotal.

Theorem: systems of luminal wave momenta that are connected by incremental momentum boosts obey the usual relativistic momentum equation for particles.

In the above, the incremental change in the scalar momentum of the i^{th} wave momentum, dp_i, is given by the component of the vector $\mathbf{dp}_i$ parallel to $\mathbf{p}_i$, namely:

$$dp_i = d\mathbf{p}_i \cdot \frac{\mathbf{p}_i}{p_i} \; .$$

Substituting (4.10) in this gives $M_e c \, dp_i = \mathbf{p}_i \cdot d\mathbf{P}$. Noting that $\sum dp_i = c \, dM_e$, summing over i gives:

$$c^2 M_e \, dM_e = \mathbf{P} \cdot d\mathbf{P} \; .$$

Noting that the right hand side is just $P \, dP$, this equation is easily integrated, resulting in the common expression for the invariance of the 4-momentum in Special Relativity:

$$M_e^2 c^2 = P^2 + M_0^2 c^2 \; , \tag{4.11}$$

where M_0 is the value of M_e for $P = 0$. Let $\beta = V/c$ as usual so β is a +ve real number in the interval $[0, 1]$. The basic equations of relativistic Mechanics, $\mathbf{P} = \gamma M_0 \mathbf{V}$ and $M_e = \gamma M_0$, where $\gamma = 1/\sqrt{1-\beta^2}$, follow upon substituting (4.8) into (4.11).

We have now derived the relativistic momentum equation for systems of luminal waves from the basic Mechanics principles. However, at this point, it has only been derived for a particular set of observers for whom the characteristic velocity is equal to c in all directions.

It was noted above that to alter the condition of motion of the golf ball is to alter the golf ball itself, by altering the directions of propagation and scalar momentum densities of all the internal wavefields. In order to find out what our coordinate transformations are, we first need to find out what impact such changes may have on the shape of the golf ball (and hence the dimensions of a ruler) and its rate of internal evolution (and hence the rate of a clock).

---

alternate factorisation would imply getting different intermediate results when using different coordinate transformations for the factorisation, which is not the case. Therefore, the mirror example constitutes a derivation of an individual Doppler shift and aberration operation and there is no new assumption here.

[18]Deriving Lorentz Transformations from the kinematic relation between systems will also be a vital consideration when we turn to the relationship between Lorentz Invariance and spacelike causal correlations.

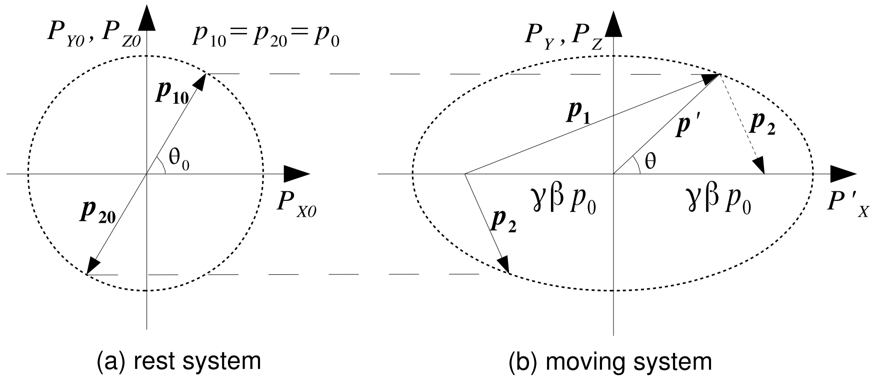

Figure 4.4: *Binary wave systems whose centers of inertia are (a) at rest and (b) moving at speed $V = \beta c$.*

4.5 Lorentz Fitzgerald Length Contraction

The first step in the analysis of length contraction is to determine the transformation rules for individual momenta in a wave system when the total system momentum undergoes a series of incremental boosts. Along the way we'll find out how to calculate vectors like $a\mathbf{p}$ in the second illustration of the relativistic momentum equation in the preceding Section. Please note that this first step is a momentum space analysis - the physical locations of the $\mathbf{p}_i$ aren't relevant.

The result will, of course, just be a return to the usual relativistic Doppler shift and aberration results, albeit in a component form more suitable for the second step, which is to show that as a wave system is boosted from rest[19] to speed V, it is physically compressed by the factor γ in the direction of its motion.

Since the individual momenta are going to end up transforming independently of each other, we can use any composite wave system to motivate the equation for the general case. The simplest case consists of two equal but opposite wave momenta. Consider the system shown in Figure 4.4, where a series of incremental momentum boosts transforms the rest system of Figure 4.4a, for which $P = 0$, into the moving system of Figure 4.4b, for which $P = \gamma M_0 V$.

Given that the total vector momentum in Figure 4.4 is $\mathbf{P} = \gamma m_0 \mathbf{V}$, it is a problem in high school trigonometry to show that each of the momenta obeys the following equation (see Appendix D):

$$\left(\frac{p_{ix} - \gamma\beta p_{i0}}{\gamma}\right)^2 + p_{iy0}^2 + p_{iz0}^2 = p_{i0}^2 \tag{4.12}$$

The result of applying this equation to an isotropically distributed wavefield is shown in Figure 4.5.

For the general case, consider as initial condition an arbitrary system with a number $N \geq 2$ of waves of various scalar momenta, p_{i0}, whose directions of propagation are distributed in space such that $\mathbf{P}_0 = \sum_i \mathbf{p}_{i0} = 0$ and $\sum_i p_{i0} =$

[19] "rest" is so far only defined for an observer who considers that the speed of light is equal to c in all directions.

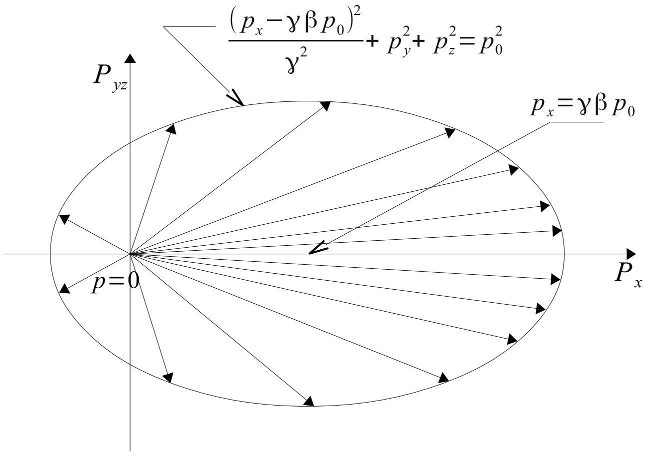

Figure 4.5: *Individual momenta in an isotropic wave system boosted to $V = \beta c$.*

$M_0\, c$. Each of the component momenta in this system is such that:

$$p_{ix0}^2 + p_{iy0}^2 + p_{iz0}^2 = p_{i0}^2 \,. \tag{4.13}$$

We can show that the i^{th} vector in this initial condition transforms into (4.12) by differentiating (4.12) with respect to β and then recovering (4.10):

$$\frac{dp_{ix}}{d\beta} = \gamma\left(p_{i0} + \gamma\beta p_{ix}\right). \tag{4.14}$$

Expanding the first term in (4.12) and using $\gamma^2\beta^2 = \gamma^2 - 1$ gives:

$$p_i = \frac{p_{i0} + \gamma\beta p_{ix}}{\gamma} \,. \tag{4.15}$$

From $P = P_x = \gamma M_0 V$, we also have $dP_x = \gamma^3 M_0\, dV$, so that:

$$dp_{ix} = \frac{dp_{ix}}{dV} dV = \frac{dp_{ix}}{d\beta}\frac{dP_x}{\gamma^3 M_0\, c} \,. \tag{4.16}$$

Finally, substituting (4.14) and (4.15) in (4.16):

$$dp_{ix} = \frac{p_i}{\gamma M_0\, c} dP_x \,,$$

which is the x-component of (4.10). Due to the choice of coordinates, the y and z components of momentum were unaffected, so the ellipsoidally modified distribution (4.12) is generated by the action of (4.10) on our arbitrary initial condition as expected. Comparing (4.12) and (4.13), the components of the moving system wave momenta are:

$$p_{ix} = \gamma(p_{ix0} + \beta p_{i0}) \ \ , \ \ p_{iy} = p_{iy0} \ \ , \ \ p_{iz} = p_{iz0} \,. \tag{4.17}$$

As shown in Appendix C, this is equivalent to the combination of relativistic Doppler shift and aberration operations acting on a wave in Special Relativity, written in component form, but recall again that we have obtained this result so far only for an observer for whom the velocity of light is equal to c in all directions.

Now, for the second step, consider the meaning of the wave velocity relative to the centre of inertia of the system to which it belongs. The complete set of wave trajectories defines both the shape of the energy distribution and its movement as a whole through our observer's frame of reference, which is to say the movement of the centre of inertia of the system. Consequently, the movements relative to the centre of inertia are the internal movements of the system, which determine the shape of the energy distribution in 3 dimensions. It will not be necessary directly to address the question what the shape is, because it will now be shown that, whatever shape it may form, the entire wavefield is compressed by the factor γ along the direction parallel to the velocity of the centre of inertia, $\mathbf{V}$.

Consider any individual wave trajectory in any luminal wave system, and let it be divided into N short segments. The relative velocity on the i^{th} segment is calculated as follows. The total velocity on the i^{th} segment, $\mathbf{v}_i$ has components $v_{ix} = cp_{ix}/p_i$, $v_{iy} = cp_{iy0}/p_i$, and $v_{iz} = cp_{iz0}/p_i$. The relative velocity is just[20] $\mathbf{v}_{\mathrm{r}i} = \mathbf{v}_i - \mathbf{V}$. Using (4.15) and (4.17) with $\gamma^2\beta^2 = \gamma^2 - 1$, it is readily shown that $\mathbf{v}_{\mathrm{r}i}$ has the components:

$$v_{\mathrm{r}ix} = v_{ix} - V = \frac{cp_{ix0}}{\gamma p_i}\ ;\ v_{\mathrm{r}iy} = v_{iy} = \frac{cp_{iy0}}{p_i}\ ;\ v_{\mathrm{r}iz} = v_{iz} = \frac{cp_{iz0}}{p_i}\ .$$

Finally, if $\mathbf{v}_{\mathrm{r}i}$ makes the angle ϑ_i with the x-axis, then:

$$\tan\vartheta_i = \sqrt{v_{\mathrm{r}iy}^2 + v_{\mathrm{r}iz}^2}/v_{\mathrm{r}ix} = \gamma\tan\vartheta_{i0}\,, \tag{4.18}$$

where ϑ_{i0} is the corresponding angle in the rest system.

This relationship means that the internal movements rotate towards the transverse plane, such that the ratio of parallel and transverse dimensions of the trajectory segment is reduced by the factor γ. This applies to every short segment of every wave trajectory in the system. Provided the transverse dimensions of whole trajectories don't change as the system velocity changes, the system as a whole is compressed in the direction of travel by the same factor, γ, which is to say that it undergoes a Lorentz-Fitzgerald length contraction.

For the transverse dimensions, consider a possible trajectory of an internal movement in the rest system that lies entirely in the plane transverse to $\mathbf{V}$. Consider the set of planes that are (a) tangent to this trajectory and (b) contain $\mathbf{V}$. Given that $V = V_x$, every such plane is parallel to the x-axis and since all the boosts are in the x direction, every incremental boost lies in one of these planes.

As the system is boosted, all along this trajectory the wave momentum begins in one of these tangent planes and it remains in the same tangent plane throughout the entire boost. The wave velocity simply rotates in a plane that is (a) tangent to the rest system trajectory and (b) contains $\mathbf{V}$. Since there is no component of the boost normal to any of these planes, none of them moves in an incremental boost, and therefore none of them moves throughout the entire boost process from 0 to V.

---

[20]Since $\mathbf{V}$, $\mathbf{v}_i$ and $\mathbf{v}_{\mathrm{r}i}$ are all referred to the same observer there is no relativistic composition of velocities involved here.

The transverse dimensions of a wave trajectory in the plane transverse to **V** are therefore preserved as the system is boosted to **V**.

Since that applies to one full trajectory, it also applies to the whole system of generalised wave trajectories (because, under the incremental momentum boost transformation, trajectory segments that are colocated at a given point remain colocated at the corresponding point in the transformed trajectory system after each incremental momentum boost).

That completes the derivation of the Lorentz-Fizgerald contraction for luminal wave systems.

However, amongst all the possible forms of wave trajectory system there is one particular kind that will be of special interest in the following Chapters, where every wave trajectory in the rest system exists on the surface of a sphere. This subset is of particular interest because of Special Relativity on the one hand and de Broglie waves on the other (as we shall see in Section 6.1).

In Special Relativity the full set of (homogeneous) Lorentz Transformations is named the Lorentz Group. There is a subgroup, known as the little group, that preserves the linear momentum of a particle[21][23, 24]. Inside the little group, there is a subsubgroup of transformations that preserves the linear momentum of a particle *in the comoving frame*, *i.e.* which preserves $P = 0$. This subsubgroup is the group of ordinary rotations in 3-Dimensions, SO(3).

What all of that means is that rest particles in Special Relativity evolve internally under rotations: The physical structure at any time t is connected to the physical structure at $t = 0$ by some, perhaps complicated, set of rotations, but the result is that any internal movement lies on the surface of a sphere.

As far as the little group for moving particles is concerned, what the general result above shows is that the internal movements of the moving particle lie on the surface of a moving ellipsoid, a sphere compressed by the factor γ in the direction of its motion. This special case is worked through in [21]. It is a visualisable example that provides physical insight into length contraction. Without going through the algebra here (see [21]), the physical space and momentum space transformations can be described as follows.

The top left hand corner of Figure 4.6 shows a sphere of radius r_0, which is represented in the figure as a circle. Consider an arbitrary point on the sphere such that the angle that the radius vector makes with the x-axis is ψ_0. Since the trajectories of the wave vector lie on the surface of a sphere, any wave momentum present at this point lies in the tangent plane. This tangent plane makes the angle θ_0 with the x-axis, and we construct for the purposes of analysis a plane parallel to it that passes through the center of the sphere. The diagram in the top left hand corner thus represents both the momenta at the chosen point in the rest system and their physical positions in the rest system, on the surface of the sphere.

The diagram in the bottom left hand corner shows how physical positions in the moving system correspond to positions in the rest system in accordance with the above transformation of the relative velocities and the rotation of trajectory segments, (4.18). This diagram shows that any sphere in the rest system is

[21]Recall that, when one combines two members of a group, one obtains another member of the same group, so the combination of two Lorentz Transformations is a Lorentz Transformation and likewise the combination of two little group transformations is another little group transformation.

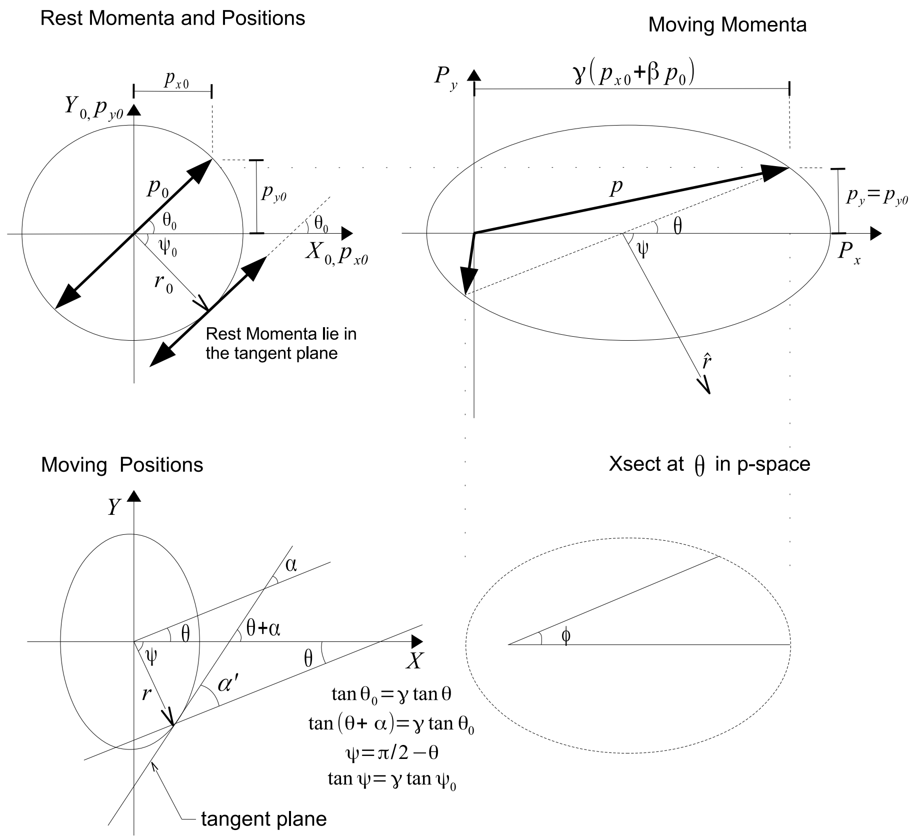

Figure 4.6: *Momenta and Positions in Rest and Moving Luminal Wave Particle Models evolving under the little group.*

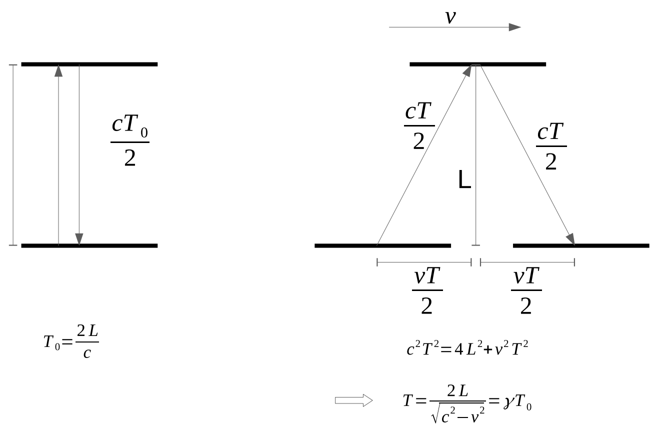

Figure 4.7: *The standard light clock.*

physically compressed by the factor γ in the moving system, which is to say that the system as a whole undergoes length contraction.

The diagram in the top right hand corner represents what happens to the momenta when the rest system is boosted in accordance with the incremental momentum boost generator, (4.10).

The key point is that, on one hand, the momentum distribution is stretched by the factor γ while, on the other, the physical shape of the sphere is compressed by γ. Overall, the wave trajectories that lay on the surface of some sphere in the rest system transform into trajectories that lie on the surface of an ellipsoid that is moving at speed V in the x-direction.

Corresponding to the wave momenta at any point in the rest system, moving system momentum vectors are drawn from a common origin in momentum coordinates. In the rest system, all these vectors lay in the plane at angle θ_0 with the x-axis. In the moving system, they do not lie in a plane. They lie on the surface of a cone whose vertex is at the origin of momentum coordinates and whose base is an ellipse lying in the plane, at an angle θ such that $tan\,\theta = tan\,\theta_0\,/\gamma$, through the centre of the momentum distribution (which is located at the point $p_x = \gamma\beta p_0$ in momentum space). The base of this cone is the ellipse shown in the bottom right corner of Figure 4.6, where ϕ is the angle into the page in the diagram above.

4.6 Time Dilation

Time dilation is perhaps the most intuitively obvious of the relativistic phenomena. When the golf ball is sitting in front of me, all the movements of all the energy inside it are internal movements, so it is changing on the inside very rapidly. However, if the same golf ball were moving at close to the speed of light then most of the movement of energy would be involved in transporting the golf ball through space and its internal evolution would necessarily slow down. Many readers would be familiar with the example of a light clock, as shown in Figure 4.7. For the moving clock, the light has to travel further to complete each period, so the periods are dilated.

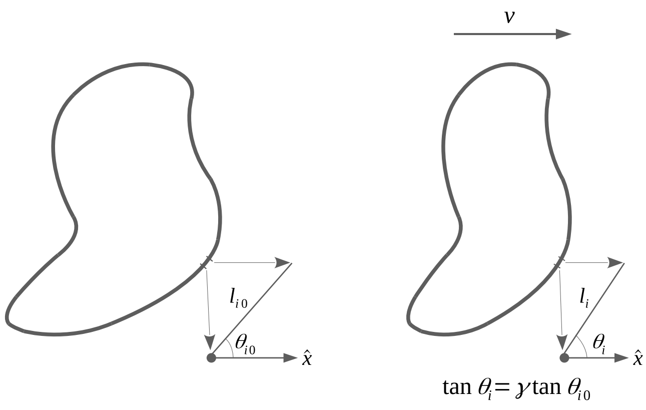

Figure 4.8: *Generalised light clock: Time dilates by γ on an arbitrary closed trajectory.*

In general, the energy that constitutes a physical object isn't just moving back and forth in one direction transverse to the comoving observer, as it is in the light clock. It could be moving in any direction at any given time, so the analysis of a generalised light clock, as shown in Figure 4.8, is required.

With respect to the rest system's wave trajectory system, consider any closed trajectory formed by n segments, where the i^{th} segment has length l_{i0} and makes the angle θ_{i0} with the x-axis. The speed on all segments is $v_0 = c$ so the period around the closed trajectory is:

$$T_0 = \frac{1}{c}\sum_{i=1}^{n} l_{i0} \ ,$$

where T_0 is the time elapsed on a clock in the rest frame to traverse the trajectory in the rest system. Lengths in the rest system may be written in component form such that:

$$l_{i0}^2 = l_{ix0}^2 + l_{iy0}^2 + l_{iz0}^2 \ .$$

Let the trajectory system now move in the x-direction at speed V. Given the length contraction shown in the previous Section, x-components contract by the factor γ and the corresponding relationship is:

$$l_i^2 = \frac{l_{ix0}^2}{\gamma^2} + l_{iy0}^2 + l_{iz0}^2 \ .$$

It is readily shown that:

$$l_i^2 = l_{i0}^2(1 - \beta^2 \cos^2\theta_{i0}) \ . \tag{4.19}$$

The moving and rest system angles are related by $\tan\theta_i = \gamma \tan\theta_{i0}$, from which it is also easily shown that:

$$\frac{\cos\theta_i}{\cos\theta_{i0}} = \sqrt{1 - \beta^2 \sin^2\theta_i} \ . \tag{4.20}$$

The relative velocity on the i^{th} segment in the moving system, $v_{\mathrm{r}i}$, is constrained by:

$$(v_{\mathrm{r}i}\cos\theta_i + V)^2 + v_{\mathrm{r}i}^2 \sin^2\theta_i = c^2 \ , \tag{4.21}$$

which leads to: $v_{\mathrm{r}i} + V\cos\theta_i = c\sqrt{1-\beta^2\sin^2\theta_i}$, from which, using (4.20):

$$v_{\mathrm{r}i} = \frac{l_{i0}\,c\,(1-\beta\cos\theta_{i0})}{\gamma l_i}\,.$$

The time taken to traverse the i^{th} segment in the moving system is $l_i/v_{\mathrm{r}i} = l_i^2/v_{\mathrm{r}i}l_i$, so, using (4.19), we may write the period elapsed on clocks in the rest system for traversals around the Lorentz contracted moving system trajectory as:

$$T_0^V = \sum_{i=1}^{n}\frac{l_i^2}{v_{\mathrm{r}i}l_i} = \frac{\gamma}{c}\sum_{i=1}^{n} l_{i0}(1+\beta\cos\theta_{i0})\,.$$

Since $\sum_i l_{i0}\cos\theta_{i0} = 0$ it follows that $T_0^V = \gamma T_0$. It might be argued that trajectories need not form closed loops, but a path that crosses a given plane transverse to $\mathbf{V}$ must eventually either recross the same plane or become confined to a smaller region, in which it must either routinely recross a transverse plane or become confined to an even smaller region and so on. In steady state, the trajectories can only be transverse or regularly recross a transverse plane. The analysis above also covers open paths between points in the same transverse plane, for which the condition $\sum_i l_{i0}\cos\theta_{i0} = 0$ is also fulfilled. The time between such crossing points dilates by γ. We conclude that the internal processes of a luminal wave system slow down by the factor γ. The argument from internal processes to real world clocks is well established [25], and tested [26, 27, 28], so moving clocks will run slow according to the usual relation $dt/dt_0 = 1/\gamma$.

4.7 Lorentz Transformations

So far, we have obtained the relativistic momentum equation, length contraction and time dilation from basic Mechanics and the idea of waves propagating at c in a medium for an observer at rest in the inertial frame of the medium. His mechanics is relativistic Mechanics. This is noteworthy in itself, but what about other observers? What about Lorentz Transformations? What about all the paradoxes? And what about the Metaphysical possibility of some other primordial form of energy that does not propagate at c?

Let us begin with the first two questions, where the result is already in the literature.

Selleri [22, 29] has studied in detail the question of how to identify all possible coordinate transformations that are consistent with these archetypical relativistic phenomena. He considered three assumptions, namely: length contraction, time dilation and constancy of the 2-way velocity of light[22] and showed that any two of these assumptions both implies the third and constrains the coordinate transformations between a preferred rest frame, $S_0 = (x_0, y_0, z_0, t_0)$ and a frame $S = (x, y, z, t)$ in standard configuration[23] moving with speed v to the following form:

$$x = \frac{(x_0 - \beta c t_0)}{\sqrt{1-\beta^2}}\;;\; y = y_0\;;\; z = z_0\;;$$

[22] *i.e.* A generalisation of the Michelson-Morley findings.

[23] *i.e.* The origins of the two observers' coordinate systems coincide at $t = t' = 0$, x, y and z are parallel to x', y' and z' respectively and their x axes are aligned parallel to the relative velocity between their inertial reference frames.

$$t = \sqrt{1-\beta^2}\, t_0 + e_1(x_0 - \beta c t_0)\,,$$

where $\beta = v/c$ and e_1 is a clock synchronisation parameter.

Setting $e_1 = -\beta/(c\sqrt{1-\beta^2})$ corresponds to the usual Einstein clock synchronisation convention and reduces this to the Lorentz Transformation.

Let us now formally adopt the Einstein clock synchronisation protocol. This is not a free choice. It is implicit in the logic of the relativity principle. If one wants covariant Laws, one must begin with a protocol that is consistent with the principle, *i.e.* All observers must use the same protocol. We shall see in Chapter 6 that this particular protocol is independently justified by the existence of de Broglie waves in nature. They would not exist if Mother Nature herself did not implement Einstein's protocol!

Our coordinate transformations are therefore Lorentz Transformations and the relativity principle and the constant speed of light for all observers are therefore results, not postulates.

The entire Theory of Special Relativity has now been recovered intact. It is also now finally clear that the wave inertia changes we have analysed are frequency changes corresponding to the relativistic Doppler shift[24], as opposed to, say, amplitude changes. In particular, "rest" is now defined for any observer, who considers moving clocks slow and moving rulers short. However, we now have objective length contraction and time dilation phenomena underlying the relativistic observer perspective. The next Section shows how this extends to objective simultaneity and objective temporal order, as is necessary to avoid the quantum before-before paradoxes [30, 31, 32].

4.8 The Preferred Frame and Objective Simultaneity

4.8.1 The Preferred Frame

Although the Einstein protocol is the only acceptable choice, its implications need to be understood and the whole question of clock synchronisation protocols bears thoughtful consideration and discussion. Although we have the optimal protocol to use for Physics, we are not close to being done with the wider question.

Is the Einstein protocol a mere human device or is it intrinsic to reality? Mother Nature has provided us with facts that argue both ways, depending on what one means by "intrinsic". In the case of de Broglie waves, it's intrinsic because, as we shall see in Section 6.1, when Mother Nature implements simultaneity at a distance, she is constrained by circumstance to implement relativistic simultaneity.

Nonetheless, there are important facts that argue that when we synchronise our clocks it is merely a practical choice and there has to be a deeper "fact of the matter" regarding the temporal order of spacelike separated events. The temporal order may appear to us to be observer dependent, but this cannot be the case. This Chapter began by considering the different flavours of relativistic paradoxes, leading to the conclusion that the present understanding of

[24]It should be noted that until this point, we have been writing these expressions in terms of momenta not frequencies. Now that we have the usual Theory intact we can take on board all of its results.

the Theory involves an internal inconsistency. The paradox that most directly exposes the inconsistency, at the level of Physics, is the double Bell paradox [30, 31]. It will be described in some detail at the end of Chapter 7, which deals with the instantaneous causal correlations at a distance that are the foundation for the paradox. However, it is not the only paradox in the literature where an individual observer is confronted face to face with a contradiction resulting from before-before timing with causally related events, for example [32].

Let us now consider how this model fixes all such problems and more without modifying the relativity principle, the clock synchronisation protocol or the covariance of the Physical Laws.

Where the theorists made the argument from the relativity principle, to the Lorentz Transformations, to relativistic Mechanics and the conclusion that energy propagates at c, we have made the case in the opposite direction. What is the difference?

There is no difference whatsoever in any of the predictions, but we did begin with the idea that energy propagates at *c in a medium*, whereas the Physics community long ago concluded that there is no medium. In the language of the early 20th century, there is not supposed to be an "aether". That was a strong, highly Metaphysical conclusion.

This conclusion was drawn, not because it was ever proved that there is no medium, but because they thought at the time that, even if there was, its inertial frame could never be found. This "no preferred frame" conclusion is based on an especially naive, positivist reading of the formalism which gained traction, and came to be regarded as a doctrine. It goes like this: Theoretical Physics ought not depend on something that can never be found. If it cannot be found, it has to be redundant and things that are redundant can, should and must be removed from the Theory.

That is all very well (in the limited domain of Physics), except for one tiny problem. It has been found. Repeatedly. We routinely observe the inertial frame of the medium, and then we routinely throw the observations away, because the observed facts must not be allowed to get in the way of the nice, philosophical reasoning in the last paragraph.

All that has been missing for the last six decades, and what this model provides, is the connection between the observations and the role that the medium plays in the Theory of Lorentz Invariance: Until now there has been no formal development of Lorentz Invariance that predicts BOTH the covariant Laws AND the preferred frame evidence.

The early 20th century position was thus understandable, but this is the 21st century and the Physics community is so tied up in quantum knots that it is unacceptable to compound their difficulties by choosing to see Lorentz Invariance as part of the problem when it is actually pivotal to the solution.

Most physicists today believe that space is real. They frequently proclaim that the quantum vacuum is real, but they cannot allow themselves to use the word "medium" without qualifying it. They have to say "relativistic medium" or more often "some kind of relativistic medium". They insist on talking about spacetime as a physically real entity composed, *inter alia*, of time, but time is not a thing. It is a measure of the changes in things. That's categorically different. My coffee cup may have been made in the 20th century, but it is not made of the 20th century.

What we have shown above is that a physical medium in which energy propagates at c generates Special Relativity for all observers. All the usual phenomena were easily understood.

The medium in this physical model seems hidden from observers by virtue of the Lorentz symmetry of luminal wave systems, but we can find the inertial frame of the medium. It will be helpful for this part of the discussion to reference a video presentation by Nobel prize winning physicist, Frank Wilczek, which is entitled "Frank Wilczek and Lawrence Krauss: The materiality of a vacuum." [33].

Professor Wilczek gives an excellent overview of the reasons why modern physicists accept that the vacuum is material substance. At the risk of doing a disservice to his full presentation, the essential idea is that, since the Theory contains particle pairs that emerge spontaneously from the vacuum, the vacuum must be considered existential. That's reasonable as well as Metaphysical. The key device that he uses to illustrate his case is an imaginary world called "Silicon world". I will tell the story a little differently, but consistent with the original.

Imagine an extended region of our space that contains a large block of silicon. Physics predicts, and experiments confirm, the existence of quantised vibrations of the silicon lattice, which are called phonons. Phonons propagate through the silicon lattice at a constant speed, the phonon velocity, analogous to our characteristic velocity. We can also easily imagine that the phononic energy in this system could form more complex subphononic entities analogous to our particles, atoms and molecules, that life would emerge and there could ultimately be physicists living in Silicon world.

They would not know about the silicon lattice. They would think of it as an empty space through which they can move at will. That is the essence of Professor Wilczek's analogy to Silicon World. With all that in mind, let us go a step further.

Would they have Special Relativity? Yes they would. The same analyses, both Einstein's and mine, apply to their world, with the phonon velocity replacing our characteristic velocity. Do they have a preferred frame? Of course they do. We would laugh at their physicists' preferred frame ban because we can see the lattice.

Although they can never see it, can they identify the inertial frame of the lattice? Indeed they can, as follows. As well as their phonon energy based material systems, they also find free phonons propagating this way and that through their space, analogous to our Cosmic Microwave Background Radiation (CMBR). However their cosmologists view its origins, they are all in agreement that such a radiation bath is *prima facie* isotropic. The expected value of the momentum density in any one direction should be the same as in any other direction. The radiation bath, as a whole, should not carry any linear momentum. However, when the Silicon world physicists make the necessary measurements they find that the observed momentum density depends strongly on the direction in space and the result they get depends on the condition of motion of the observer who is making the measurements. This is analogous to the velocity dipole that we observe in the CMBR [34, 35].

Readers may have seen whole sky images of the CMBR showing tiny variations in the CMBR temperature at the level of one part in a hundred thousand, as in Figure 4.9. These "ripples in Spacetime" images are obtained after post-processing the data to remove a much stronger velocity dipole, at the level of one part in a thousand. Although the velocity dipole actually dominates the results it is seldom mentioned in presentations to the general public. Figure 4.10 shows the whole sky CMBR image before processing, with the dipole magnitude and direction indicated.

The reason for the velocity dipole in Silicon world is that when the observer is moving relative to the lattice, his observations of phonon momenta are enhanced (by the relativistic Doppler effect) when he is looking in the direction of his motion relative to the lattice and diminished when he is looking in the opposite direction. There is one unique condition of motion for which the phonon radiation bath is observed to be isotropic, which is when the observer is comoving with the lattice.

When Silicon world physicists declare this effect to be an "anomaly", and insist on the preferred frame ban, our real world physicists are laughing at them because we can see the silicon lattice, we know its inertial frame and we recognise that this measurement reveals the preferred frame relevant for silicon world observers - as discussed below it is the unique frame in which their clocks objectively run fastest and their rulers are longest.

An interesting question arises about the word "objective". Their entire silicon lattice could be moving through our Universe. Professor Wilczek directly (and correctly) addresses this question in his talk. He states that Silicon world physicists study the physics of Silicon world and not some other, putative super universe with superbeings, like us. Their preferred frame is the inertial frame of their lattice, and similarly our preferred frame is the inertial frame of our medium, whether it is a lattice, a continuum, or whatever.

We are in the same position as the denizens of Silicon world. Although we cannot see the medium, as Wilczek points out, we do know that it exists. Unless someone wants to claim that our Universe, as a whole, is moving through some unobservable super universe, then our preferred frame is the unique frame of reference in which the velocity dipole of the CMBR vanishes [34, 35][25]. Note at this point how any claim that our Universe is moving through an unobservable super universe breaks all the rules of Physics. It appeals to something that cannot be observed in order to justify ignoring what actually is observed!

4.8.2 Objective Simultaneity

A more reasonable contrary view is that while the CMBR frame is all very well for light, it does not follow through for matter, including especially clocks and rulers, however, in this physical model, it does indeed follow, as follows:

CMBR photons are referenced to the CMBR frame: They really do propagate at c in that frame, regardless of anyone's choice of a clock synchronisation protocol. Since everything in the model is a wave phenomenon, there is no valid distinction between a CMBR photon and one that was emitted by some atom on Jupiter. All light and all radiation is therefore referenced to the CMBR frame.

In this physical model, and only in this model, all the component parts of a material system also propagate in the same medium at the same speed and so they are all referenced to the CMBR frame. That provides the link to clocks and rulers, and one can only form it when all the ontology is propagative. We saw

[25] See also Section 4.10. It's not just the CMBR. Any *a priori* isotropic radiation bath will do. Other such baths exist and, yes, they confirm the same preferred frame for the medium in this model.

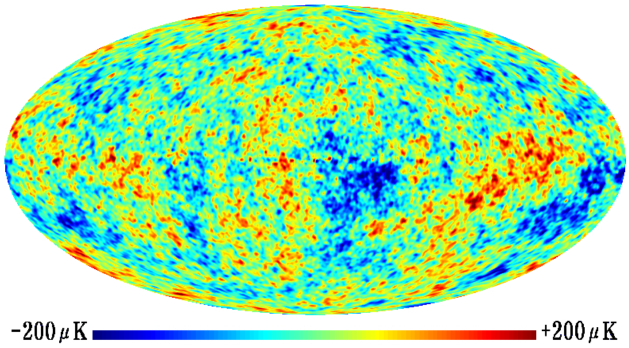

Figure 4.9: *The usual "ripples in time" image. Higher temperatures refer to higher momentum densities of the CMBR. Note that the colour scale here spans only* ±200 *micro Kelvin.*

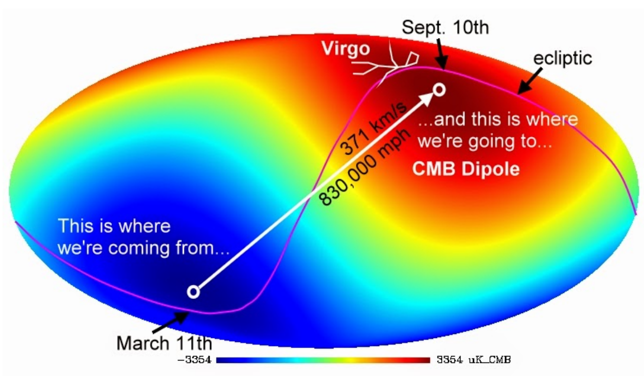

Figure 4.10: *An original whole sky image of the CMBR, before processing the data to remove the velocity dipole. Note that the colour scale here spans some* ±3350 *microkelvin.*

in Sections 4.5 and 4.6 that clocks made from matter that has an objectively isotropic momentum density distribution run faster, and rulers are longer.

That condition is uniquely fulfilled when the matter is at rest in the CMBR frame of reference. Only in this frame are clocks and rulers undistorted by their condition of motion. This completes the logical passage from the idea of a medium to identifying its frame of reference in practice, to the preferred status of CMBR clocks and rulers for the purpose of defining objective simultaneity at a distance.

In other frames, the combination of length contraction, time dilation and clock synchronisation cancels, and we still observe the same speed of light, c, but the simultaneity that appears to an observer is just that, an apparent simultaneity, warped by the velocity dependence of clocks and rulers. Only CMBR observers see the "unwarped" version, which is to say the objective simultaneity.

This does not undermine the relativity principle. The main idea in this principle is that, however an observer's clocks and rulers may be distorted, the Laws of Physics, expressed in his time and distance units, remain the same. In a world with velocity dependent clocks and rulers, the way to achieve that result is for all observers to follow the same procedure when synchronising clocks. The principle implies the protocol and the protocol yields Laws that obey the principle.

If one uses a protocol that violates the relativity principle, one arrives at frame dependent Laws, and, theoretically speaking, chaos ensues, as we shall now see.

4.9 Alternative Clock Synchronisation Protocols

I had heard of Franco Selleri and the inertial transformations, running on a different clock synchronisation protocol, but I'd never read his work until it came time for me to show whether the Lorentz Transformations do in fact follow from length contraction and time dilation alone or whether Einstein clock synchronisation has to be introduced as a separate assumption.

When this exercise turned out to be not quite as easy as it might seem on the face of it, it was time to look at Professor Selleri's work. It was immediately apparent that he had developed not just the proof cited in Section 4.7, but quite a bit more.

He had shown that once one assumes length contraction and time dilation, the general form of coordinate transformations is fixed up to a single parameter, named the "synchronisation parameter", e_1. For frames in standard configuration, the synchronisation parameter determines how the primed observer's time variable, t', depends on the unprimed observer's space variable, x.

The appropriate choice of the synchronisation parameter, e_1, in order to reduce Selleri's general form for the coordinate transformations to the Lorentz Transformations was given, so my task was done.

Professor Selleri, however, focusses on a different choice, $e_1 = 0$, for which the coordinate transformation looks like this:

$$x' = \gamma(x - vt) \quad ; \quad y' = y \quad ; \quad z' = z$$

$$t' = \frac{t}{\gamma} \; .$$

These are today known as the inertial transformations, and they were first identified by Tangherlini [36] in the 1960s. The space transformation is the same as under Lorentz Transformations, but note the difference in the time transformation rule. Under Lorentz Transformations it is $t' = \gamma(t - \beta x/c)$. Note that γ appears in the denominator for Selleri, but in the numerator in Lorentz Transformations. Since it is in the denominator for Selleri, $t' < t$, which is immediately consistent with a dilation of time.

For inertial transformations, clocks are synchronised as follows: One inertial frame is chosen to be the "preferred" frame. This choice is completely free. Clocks in the designated "preferred" frame are synchronised using the usual Einstein protocol. Clocks in other frames are synchronised by setting each of them equal at time $t = 0$ to a nearby clock in the "preferred" frame so that, at $t = t' = 0$, all clocks in all inertial frames begin with the same reading. Note that since observers in various frames do not all follow the same procedure, this is not a relativistic protocol.

Selleri (and others) had done a lot of work [20, 22, 29, 37, 38, 39, 40] which proved that all the usual empirical predictions can be made using inertial Transformations and their results are identical to the results from Lorentz Transformations.

All observers agree that clocks in moving frames run slow, and by how much. Although the time coordinates are different, they also all agree about simultaneity at a distance. And they agree about the lengths of rulers. There are thus no paradoxes, and this kind of synchronisation protocol looks good from a philosophical point of view, but that will be undermined shortly.

This approach to clock synchronisation is very effective at solving problems on a rotating platform, where Special Relativity is quite tedious because clocks at the perimeter of the platform aren't in inertial frames. However, the magnitude of the relative velocity of any two clocks on the platform does not change, so they always remain synchronised with inertial synchronisation. That is why the inertial approach is used in the GPS satellite system's clock synchronisation scheme. With Einstein synchronisation the satellites' onboard clocks would constantly be in need of resynchronisation.

Those are the main benefits of inertial transformations, but as good as it is on the rotating platform, it is really bad at everything else. The problem is the lack of symmetry.

Many nonphysicists would have heard of Maxwell's Laws, and know that every time a physicist mentions Maxwell's Laws, he also says the word "symmetry". Many may know that almost all the content of Maxwell's Laws was known before he added just one extra term in the name of symmetry, the displacement term. It was soon found experimentally. Most probably don't know that there are alternative formulations of the full set of Electromagnetic relations where the equations are written using full derivatives, like $d\mathbf{E}/dt = \partial\mathbf{E}/\partial t + v\,\partial\mathbf{E}/\partial x$, instead of partial derivatives. The reason this is not commonly known is not because it's empirically wrong to do it that way but because the resulting equations lack symmetry.

To nonphysicists, this word "symmetry" might seem like a magician's incantation. The fact is that symmetry matters (a lot) and the inertial transformations provide an excellent opportunity for non-specialists to understand why it is considered so very important.

For this purpose, solving the above equations to obtain x, y, z and t in terms of x', y', z' and t' gives the inverse transformation:

$$x = \frac{x'}{\gamma} + \gamma v t' \quad ; \quad y = y' \quad ; \quad z = z'$$

$$t = \gamma t' \ .$$

This is the form Selleri gives and the first thing to notice is that the inverse transformation is different from the forward transformation. It's asymmetrical.

Moreover, it is not properly speaking an inverse transformation (a transformation from the primed frame back to the unprimed frame), because it uses parameters, γ and v, that belong to the unprimed observer. Indeed, for the unprimed observer, the speed of the origin of the moving system, *i.e.* the point $x' = 0$, is v, while for the primed observer the speed of the origin of the unprimed system, *i.e.* the point $x = 0$, is $v' = -\gamma^2 v = -v/(1-\beta^2)$, where $\beta = v/c$, so $v' \neq -v$.

If Special Relativity's second postulate appears unreasonable at first glance, this is hardly an improvement: If you are moving relative to me at speed v, then I am moving relative to you, not at v, but at a different speed, v'. That completely undermines the intuitive appeal of the inertial transformations. In both cases, intuition is restored by recognizing that clocks and rulers are velocity dependent machines. In Special Relativity what one needs to understand is that these are objective phenomena, not some mere relativistic artefacts, and the formal symmetry is inseparable from the clock synchronisation protocol.

Now, if we want an inverse transformation that only uses primed frame values, as it should, then $\gamma v t'$ should be replaced by $v't'/\gamma$, but we would also need to replace γ, which is equal to $1/\sqrt{1 - v^2/c^2}$, with something that only involves primed frame values. This is problematic. Not only is $v' \neq -v$ but v' can be greater than c, which is also not a constant. It depends on the direction in space, as follows:

$$c'(\theta) = \frac{c}{1 + \beta cos\ \theta} \ ,$$

where θ is the angle between the x-axis and the direction in space in the unprimed frame. There is consequently no clean definition of a gamma-like parameter for the primed frame. That's enough math, the point is made. It is not that Physics problems *can't* be solved but that the lack of symmetry has immediately led us into an unduly complicated mess. At the very least we can all now understand why Selleri gave his inverse transformation in a form that mixes primed and unprimed variables.

There are a couple more unpalatable consequences for physicists. First, the inertial transformations don't form a group, which essentially means that they may not obey simple algebra rules, as are usually defined for group structures. For example, the "multiply" operation is defined for the Lorentz group such that the product of two Lorentz transformations is always another Lorentz transformation.

Second, the Physical Laws are different in different frames, which brings us to an important hypothetical question regarding the status of the Einstein clock synchronisation protocol. How would Mother Nature write down the Laws of

Physics? Like superobservers of Silicon world, let us try to put ourselves in her shoes.

She knows that the events that happen in the world are the same for anyone who observes them, but she requires no coordinate system to tell the physical state of affairs. For her it might be a basic rule that the quantity of any movement (*i.e.* momentum) is equal to the amount of stuff (*i.e.* energy) that is moving multiplied by the speed at which it moves. The only facts she has are the objective facts.

Her approach doesn't involve anyone looking, or how they measure things relative to some coordinate system. The usual presumption of physicists is that, if we want the simplest and best set of rules, then that's very likely to be the best way to formulate them. However, for us, an observation of the momentum of an object obviously does depend on the condition of motion of the observer.

In order to get away from the inherent coordinate dependence, what we do is to identify quantities in the theory that are the same for all observers. These quantities are the "invariants" of the theory. They are used to identify physical Laws, like $\mathbf{p} = \gamma m \mathbf{v}$, that have the same form for all observers. For example, the quantity $ds^2 = c^2dt^2 - dx^2 - dy^2 - dz^2$ is the same in all coordinate systems, so it is an invariant (called the invariant line element). It is the Einstein protocol that enables us to write Physics this way.

None of this means that Selleri's coordinate transformations are "wrong", but I'll always remember the feeling I had when reading the entire argument that he makes for inertial transformations in his book "Weak Relativity - the physics of space and time without paradoxes"[26] [41]. A line from a popular song by Joss Stone kept going round and round in my head (I paraphrase, with apologies): "If loving Einstein is wrong, I don't wanna be right".

4.10 A Preferred Frame Thought Experiment

There is a second way in which logic induces a preferred frame, independent of the CMBR radiation. According to the luminal wave model, both methods should identify the same frame. According to the relativist Metaphysic both are "anomalies", but if they both identify the same frame, one must ask if there is any word in English for the square of an anomaly?

Consider two observers, Alice and Bob, located near to each other in a remote region of space. Assume that when she looks out of her spaceship's windows, Alice sees an isotropically distributed Universe: For her, the angular number density and brightness of astronomical objects is random, but independent of the direction in space.

Let Bob's velocity relative to Alice be $0.9\,c$. Although the available light sources are the same for both observers, the Doppler shift of light from any given source is clearly very different. Consequently, when Bob looks out of his spaceship's window, the Universe he sees is decidedly anisotropic: In the direction of his velocity relative to Alice, the stars appear far brighter and their number density is enhanced while, in the opposite direction, the stars appear

[26] For anyone who may be interested, this book provides an excellent summary of the most important criticisms of Special Relativity (all of which are resolved here without destroying Western civilization).

dimmer and their number density is diminished. This is an experimentally confirmed fact [42].

So, who's moving (and who's not)?

Of course, only relative velocities can be measured. For example, the Sun's velocity can be measured relative to the Earth. It can also be measured relative to an abstract point, like the midpoint of a line joining the Earth and the Moon. In general, the velocity of one system of objects is measurable only relative to another. In this context, the entire observable Universe is a system of objects whose velocity is not measureable, even in principle (as there is nothing else available to measure it relative to).

Since the very concept of motion cannot be applied to the Universe, it must be considered motion-less. The same applies to the centre of inertia of the observable Universe. It is a motionless point that lies somewhere within the observable Universe.

The light that reaches an observer from distant stars and galaxies is not the only relevant global property of the Universe in this regard. As we saw in Section 4.7, there is also the Cosmic Microwave Background Radiation (CMBR), a *prima facie* isotropic radiation bath. It is well known [35] that CMBR observations show an anisotropic distribution with a velocity dipole analogous to the one in the present example.

In fact, Blake and Wall's result in [42] is that these two velocity dipoles agree: Earth's velocity relative to the millions of observable radio galaxies (an unbiased sample of all the visible matter in the Universe) coincides with its CMBR velocity. Consequently, while every CMBR photon is moving at c, the centre of inertia of all the CMBR radiation in a given region of space is also a motionless point, and we are moving at ~ 370 Km/sec relative to it[27].

In the example above, Alice's CMBR velocity is therefore zero, while Bob's is $0.9\,c$. She is not moving (relative to the centre of inertia of the Universe), while Bob is. The text in parentheses is actually redundant because the centre of inertia of the Universe is a motionless point. In summary: Bob is moving. Alice is not. The idea of an expanding space can also be brought into this discussion, leading to the conclusion that Alice is not moving in the local space, while Bob is.

How does this compare with the usual relativist assertion that there can be no preferred frame? Light is the common element in both kinds of observation, but the usual interpretation of relativity cannot avail itself of this fact to predict the consonance of these two "anomalies" because in neither case is the light considered to be propagating in a physical medium. Let us summarise the state of play:

First, the preferred frame as defined here neither plays any role in Dynamics nor breaks Lorentz symmetry.

Second, the preferred frame ban was philosophical, not evidence based. Philosophical positivists took the view that what cannot be observed cannot be said to exist but, if they'd had the evidence [34, 35, 42], it is inconceivable that the Physics community of the early 20^{th} century would have considered that the preferred frame cannot be observed. Surely, they'd have insisted that it must be an integral part of any acceptable theory of Lorentz Invariance.

Third, it has been found [30, 31, 32] that quantum nonlocality combined

[27] The dipole obviously varies during the year as the Earth orbits the sun.

with relativistic simultaneity leads individual observers to encounter manifest contradictions. As will be argued in Section 7.10, these paradoxes falsify the purist relativist Metaphysic, without touching the relativity principle itself.

Finally, the model here induces precisely this preferred frame, and no other [21, 43]. It predicts the consonance of the Blake and Wall result [42] with Smoot's [35]. It also makes testable predictions about velocity dipoles that have not yet been observed in other *a priori* isotropic radiation baths including Galactic redshifts and Cosmic rays. (Note that a Cosmic Ray bath is composed of matter not light, so this would be a distinct form of test). Bob and Alice continue to use the same well-known, covariant physical Laws, but her clocks and rulers have preferred status for defining objective simultaneity at a distance.

The combination of forbidden causal temporal loops in the relativist philosophy with an observable preferred frame takes us beyond Physics' subjectivism to a necessary and valid inference about the real, underlying, objective physical state of affairs: The necessity to draw such an inference is a consequence of the theories themselves, not merely some esoteric philosphical consideration.

We now know for sure that clocks at rest in the CMBR frame run fastest because they alone are undistorted by the velocity dependence, so when two observers disagree about whose clock is running faster, we can decide who is correct and who is incorrect. As far as forbidden causal temporal loops are concerned we know that relativistic simultaneity is apparent, before-before timing is an artefact, and all pairs of events, spacelike separated or not, have a definite temporal order and we know how to determine it. We shall return to the issue of forbidden causal temporal loops in more detail in Section 7.9.

What we don't know yet, but we will soon enough find, is that none of this makes any difference to any empirical prediction, neither by Special Relativity nor by Quantum Mechanics.

It might seem at the moment that Einstein synchronisation has been relegated to the "mere human device" category, but this is because we have been focussing in this Section on one side of the evidence. In Section 6.1 we shall see that Mother nature herself needs to synchronise her own "clocks", and she uses the Einstein protocol for this purpose.

Overall, on the one hand, objective simultaneity clearly exists, and the fact of the matter about the temporal order of spacelike separated events is knowable, the relativistic paradoxes are resolved and those philosophers who appeal to a universal "now" are vindicated. On the other, Physics' use of the relativity principle is fully justified while the observable facts also directly vindicate the Einstein clock synchronisation protocol because they show that Mother Nature herself does it the same way.

4.11 Against Non-luminal Energy

That leaves us to consider, briefly, the possible alternatives to this physical model, which is to say models in which the energy, $E = mc^2$, that matter is made of does not propagate at c.

This is not something that can be disproved, but it is easily discredited. I have repeatedly pointed to the fact that luminal propagation of energy is a standard conclusion drawn from Special Relativity. Everybody "knows" that it is so, and considers that it is part of the meaning of $E = mc^2$. However, there is actually no rigorous proof, and there can be no proof *because* Lorentz Transformations cover all physical systems. Therefore, in order to discredit the "inherently subluminal matter" idea, let us consider what happens when one does not impose luminal propagation in a physical model.

To begin with, any such model would seem to break the usual work integral connection between momentum and energy, so one would have to redefine momentum as being fundamentally velocity dependent. Of course, this is precisely what the usual Physics does (at the level of the particle as a whole). The redefinition in Physics, $\mathbf{p} = \gamma m_0 \mathbf{v}$, is the same result that we derived from conservation of momentum alone. In Physics, this redefinition is properly justified first by the relativistic invariance of the norm of the energy-momentum 4-vector and second by experimental verification. While the redefinition was justified by formalism and observation, that does not mean that it is not begging for an explanation.

This is "*After* Physics", and we are asking the "How does that work?" questions. The fact is that Physics simply does not know. The job of physicists (as defined by physicists) is to get results and no one can say that they have not performed their function well. For explanations you have to go to a different department, the department of Metaphysics. It's across the road in the Philosophy building.

Having said that, for a physical model of a particle of matter based on "nonluminal energy" one would have to redefine momentum not just for the particle as a whole but also for the internal structure. To be relativistically correct on the conservation of momentum and energy between the particle and its parts, one would have to use the same definition, $\mathbf{p}_i = \gamma_i m_{i0} \mathbf{v}_i$, for all the parts of the structure, which just shifts the problem to a lower level and explains nothing. Such a model is pointless.

Nonluminal energy also produces a bizarrely complicated internal structure due to the fact that, in Special Relativity, the total velocity in an observer's frame of any of the parts of the structure depends on his observed particle velocity in accordance with the relativistic composition of velocities:

$$\mathbf{v}_i = \frac{\mathbf{V} + \mathbf{v}_{i0\parallel} + \sqrt{1-\beta^2}\,\mathbf{v}_{i0\perp}}{1 + \frac{\mathbf{V}\cdot\mathbf{v}_{i0}}{c^2}}\,. \tag{4.22}$$

That's the result of reasoning from the Theory to the physical structure under the assumption of nonluminal energy. However, much to the dismay of physicists, the natural world actually works the other way: This structure is what is necessary to reproduce Lorentz Transformations from subluminal energy. It is an obscenely complicated way to arrive at simple and beautiful coordinate transformations. In Physics, the tradition is to seek to identify simple explanations for apparently complex behaviour, not the other way around.

On the other hand, if we put $v_{i0} = c$ into the above equation, then, of course, $v_i = c$ for all V, and the entire Theory is explained by the definition of momentum as inertia times velocity, which is to say that the simple but perplexing Special Theory has an even simpler underlying explanation that is not at all perplexing.

4.12 Against Lorentzian Relativity

This Chapter would not be complete without mentioning the way that relativity theory was shaping up pre-Einstein. Time dilation and length contraction had been shown within the Electromagnetics formalism and it had also been shown that the field energy of the electron would increase with velocity by the factor γ.

There are two parts to the reasoning why the Lorentzian physical model is not the usual theory today.

The first part is that it invokes an "unfindable" preferred frame. While we have partially refuted that argument by identifying a preferred frame, it remains correct that Lorentzian relativity cannot make the link that explains why the CMBR frame is "the" preferred frame to which electrons and other particles should be referenced for the purpose of defining absolute simultaneity at a distance. The aether frame remains unfindable in the Lorentzian approach.

The second part is that Lorentzian relativity, being an Electromagnetics based argument, does not cover the fermions (particles of matter like electrons and protons), for which there can be no explanation in linear Electromagnetics. They have to be put in by hand.

It is well known that Quantum Mechanics' instant influences at a distance are problematic for Special Relativity, and that one way to begin to address the problem is to revert to an aether Theory, along the lines of Lorentzian relativity, as was actually suggested by John Bell[28] himself.

However, the Lorentzian approach is *especially* unsuitable for the context of instant causal correlations. This is a third objection to the Lorentzian physical model, and, for me, it is the most damning of all. In that model, length contraction comes from the elliptical deformation of a moving charged particle's Electromagnetic field, which is calculated using the Lienard-Wiechert retarded potential. That means that it is calculated from the assumption of retarded interaction, where the Electromagnetic field (and any other way for one particle to influence another) propagates away from the body of a pointlike electron at speed c.

Once one makes that assumption, instantaneous causal correlations between point events at a distance are impossible. This is precisely the assumption that is falsified by EPR experiments. Lorentzian relativity cannot resolve the problem of the experimental existence of instant influences at a distance, because it is the problem!

Even putting aside the question of retardation, it should be noted that the Lorentzian derivation of length contraction is intrinsically dynamical, whereas the Lorentz coordinate transformations are manifestly kinematic. Using Dynamics to obtain a kinematical result is problematic because it ties what should be a general, theory independent, result to specific dynamical relations from a particular theory. Meanwhile, Einstein's entire derivation is as kinematic as mine.

Overall, the Einsteinian derivation does not block instant correlations at a distance but the Lorentzian approach does. Many Physicists would assume that it is the other way around, but we shall emphasise in Chapter 8 that the assumption of retarded interaction, which is used by Lorentz, is not used

[28] J.S. Bell, the originator of the Bell Inequalities.

by Einstein. Instead, it was put in by hand, in the form of light cone causal analysis, *after* the Special Theory had already been fully developed. There is nothing in the kinematical Special Theory that requires one to make any such dynamical presumption.

4.13 Spacetime is an Oxymoron

"Spacetime" is a dreadful word. There is space, and there are movements in space. Generally speaking, movements are either extended trajectories through space, or periodic vibrations or a combination of the two. Periodic movements are used to measure time, or more precisely to compare time intervals.

The word spacetime equates space with time. It equates extension with change and dimension with movement. I have no issue with the 4-vector algebra, but this idea is linguistically unintelligible, like saying that the length of a ruler and the falling of a ruler are the same thing. We can do better, as follows.

Consider our golf ball, and let us assume that we see it moving through space. What is actually happening is that the corresponding energy pattern is evolving in such a way that the golf ball-like image that we see changes position. The energy pattern is evolving in space.

Meanwhile, the golf ball is also evolving internally, with a forever changing underlying energy pattern evolving in time. Since the speed of energy movement is the same regardless of the golf ball's condition of motion, the total amount of evolution of the system is the same for all observers. This quantity is called the spacetime interval and it is a Lorentz Invariant scalar, which is most commonly written in the form of an invariant line element in a 4-dimensional mathematical space:

$$ds^2 = c^2 dt^2 - d\mathbf{x}^{\mathbf{2}} = c^2 d\tau^2 \,.$$

What this line element relation says is that if I see a clock move in space by $d\mathbf{x}$ on my rulers in time dt on my clock, then I see the time shown by the moving clock advances by $d\tau$. That is correct, but let us rewrite it as follows:

$$c^2 dt^2 = d\mathbf{x}^{\mathbf{2}} + c^2 d\tau^2 \,,$$

which says that any object in my frame of reference has the same fixed rate of spatiotemporal evolution. My own clock is at rest relative to me, and cdt on the left hand side is just the distance travelled in time dt by the energy that constitutes it. The energy in the moving clock moves the same total distance with the result that I see the clock move by $d\mathbf{x}$ in space and advance by $d\tau$.

It couldn't be any simpler. Rather than messing with the consciousness of the human race with the unnecessary, oxymoronic neologism that is "spacetime", a far better description would be "spatiotemporal evolution".

I hear you say: "What's in a name? That which we call a rose by any other name would smell as sweet." [44]. Does this actually matter? Yes it does. Spacetime has been a leading candidate for scientific realism (essentially the assertion that nature "really is just like" some mathematical formalism).

To give one of many examples of how this has played out, there are far too many physicists who have reasoned as follows: Space and time are the same thing, they are just dimensions. Things can move backwards as well as forwards in space. The same should apply to time, and so retrocausality (the notion that an event can causally influence events in its past) is on the table.

This is actually not what Special Relativity has to say on the subject! It is true that space and time enter Lorentz covariant Laws of Physics "on an equal footing", but it is not true that space and time are the "same" in relativity theory. That minus sign in the equation for ds^2 above makes a lot of difference in the details of the 4-vector algebra, as anyone who has studied Wigner's seminal 1939 paper "On unitary representations of the inhomogeneous Lorentz Group" [23] would confirm. I understand that Dirac's well known remark to the effect that "American physicists should study Special Relativity in greater depth" was an effort to point this out to Wheeler, but he was not listening.

Although the Theory itself does not support the "just another dimension" argument from covariant equations to backward in time action, aspiring young physicists have for generations been rendered susceptible to flights of fancy because of that one unintelligible word, "spacetime". It is part of a doctrine that must be embraced uncritically in order to get into the profession, and to stay in it[29].

What this word has done is to increase meaninglessly what engineers call the "parameter space" within which physicists conduct research. That can be a good thing, when well-motivated new ideas open promising new avenues for research, but spacetime is not a well-motivated idea. It has come to dominate the parameter space to the exclusion of all else, and once a fantastical, incomprehensible pseudo-explanation is preferred, it only leads to more of the same.

Having embraced the idea that time is just another dimension, why stop at four when eleven is such a nice round number? And since the physical dimensions are now the result of the Mathematics, and not *vice versa*, let us construct solutions where the curved spacetime curves back on itself so that we can worm our way across the Universe in a jiffy. Instead of being guided by Einstein's remarkable contributions, we have been derailed.

In the coming Chapters, we are slowly moving towards developing a physical understanding of EPR's instant causal correlations at a distance. The quest to understand this phenomenon has been compared to a man who has lost his keys in a dark alley, where there is just one dim streetlight. He searches in the small area that is illuminated by the light. He has no good reason to believe that his keys are in that part of the alley, but it is the only place where he might find them if they were there. We shall see that the original Einsteinian relativity is one of the main keys, and the mathematical light burns more brightly when we understand it. Unfortunately, spacetime has had us looking for EPR everywhere except in the alley.

4.14 Summary

The Special Theory is usually presented neither as a particle theory nor as a wave theory. It is a theory, first and foremost, about coordinate transformations, from which inferences as to the mechanics of both particles and waves followed.

[29] And it is the reason why the various practising physicists who I quote remain anonymous: Criticising spacetime is a career limiting move.

However, when one considers the list of inherently relativistic theories, they are all field theories, starting with Electromagnetics, the Dirac Theory and all the quantum field theories, but there is also an impressive variety of inherently relativistic, nonlinear field theories that feature the characteristic velocity [8] - [14]. Consistent with this circumstance, it made complete sense for us to approach Lorentz Invariance as a wave phenomenon.

If this seems like a step back from the total generality of the Special Theory, consider the benefits:

1. All the particle results are contained in the wave results: no redefinition of momentum is required and there is no step back from full generality.

2. The wave interpretation predicts the preferred frame that is observed. Unlike the usual interpretation, it is consistent with all the relevant evidence.

3. We do not conceptually abandon absolute space and physical intuition along with it. The wave interpretation contains both the relativist and the objectivist perspectives.

4. There are no paradoxes.

5. Quantum Mechanics runs on wave equations, and in the next Chapter it will be derived from conservation of energy-momentum alone.

6. When particles are ontology, Special Relativity blocks EPR correlations. When waves are ontology, Special Relativity becomes the key that will unlock the mystery in Chapters 6 and 7.

7. When particles are ontology, gravity is seen as an ugly duckling in Chapter 10, but when waves are ontology, it turns into a swan in Chapter 11.

Einstein always wanted a pure field Physics. His dream is just beginning.

Chapter 5

Quantum Formalisms

5.1 The Theory that Nobody Understands

Why can't Ordinary Quantum Mechanics be interpreted as a physical wave theory? The basic formalism makes predictions by way of three essential ingredients that will be of interest as we proceed to address this question:

1. The Schroedinger Equation, a linear "pseudo-wave" equation, with solutions that superpose *i.e.* If A and B are solutions, then so is $\alpha A + \beta B$ where α and β are constants.

2. The notion that, for any given physical state of affairs, the complete solution to the Schroedinger Equation can be represented as a weighted sum over the "eigensolutions" of a "complete" set of "commuting" observables[1].

3. The projection postulate: When a measurement is performed, and a measurement result is obtained, the above weighted sum of terms is reduced to just one of the terms, the one that corresponds to the obtained measurement result. The probability of obtaining a particular result is proportional to the squared magnitude of the corresponding weight in the superposition.

There are so many issues surrounding the interpretation of Quantum Mechanics in general, and Ordinary Quantum Mechanics in particular, that it is known as the Theory that "no one understands". In order to understand both why that is, and how it all relates to the idea of waves propagating at c in a medium, we'll consider each of the above ingredients in turn, beginning with the Schroedinger Equation, but to understand the issues surrounding that Equation it will be helpful first to develop a Theory that actually does make sense.

[1] These weights are, in general, complex numbers. The terms in quotation marks will be discussed below, but for the moment just note the correspondence between eigensolutions and observables: In the complete solution, the *physical* state of affairs is being represented as a superposition of *observables*. This is one of the most beautiful and fruitful ideas in Physics (but it's also another deal with the devil).

5.2 The Dirac Equation

The Dirac Equation is typically presented as a decidedly more advanced topic, not usually as part of the foundations. However, the physical model here is inherently relativistic and any direct "derivation" of a nonrelativistic theory could only be contrived *ad hoc*. The Dirac Equation was the first "proper" relativistic wave equation accepted into Quantum Physics and once we have seen how completely satisfactory the derivation is for physical waves, we'll be better equipped to deal with the Metaphysical chaos that Schroedinger's equation entails.

Dirac's Theory of the Electron was published several years after the famous Schroedinger Equation, but it is the place where Special Relativity meets Quantum Mechanics. We shall see how the wave postulate binds the Dirac Theory so tight to relativity theory that they are no longer to be seen as separate, but as one and the same Theory, fruit of the Conservation Laws.

In his two seminal articles, [45] and [46], Dirac developed an equation to represent the electron with a view to analyzing the fine details of Hydrogen atom spectra, but that turned out to be just the beginning. The Dirac Equation is the only equation we have that fully describes a real world particle in closed form, arguably the most perfect thing in Physics. Dirac's Equation for the electron does not stop at electrons. It has become the foundation for all the Quantum Field Theories.

However, after the Dirac Equation, as one climbs further up the mountain to Quantum Electrodynamics and the Standard Model, things really go downhill from the point of view of mathematical perfection. This Equation is the gold standard against which to compare and contrast the Schroedinger Equation. The famous story of how Dirac found it is also one of the most brilliant in Physics. For all those reasons, it makes sense here first to understand how Dirac's development of his Equation builds neatly on the platform constructed in the last Chapter.

To develop the Dirac Equation does require some mathematics that isn't taught in high school, but surprisingly little and there is nothing beyond the grasp of high school graduates who took an interest in Mathematics, which should include almost every reader. We will also need all of these same ideas for the Schroedinger equation.

5.2.1 Basic Mathematical Ideas for Wave Equations

The first new idea is the identity:

$$e^{i\theta} = cos\,\theta + i\ sin\,\theta\,,$$

where $i = \sqrt{-1}$. This identity is useful when manipulating expressions that contain sines and cosines of functions of space and time. It basically "automates" the most common trigonometric identities from high school. The identity is easily proved, but to understand it, just differentiate both sides with respect to θ twice.

The familiar high school expression for a plane wave propagating in the x-direction is then written as:

$$cos\ (kx - \omega t) = Re\,[e^{i(kx-\omega t)}]\,,$$

where "*Re*" means "the real part of". Conventionally, one drops the real part and just writes an elementary plane wave, propagating in the x-direction as $\psi = e^{i(kx-\omega t)}$. More generally, the symbol ψ is used for any function that is a solution to the wave equation, hence the generic term "wavefunction", used to refer to the solutions of all the various wave equations in Quantum Mechanics.

For the plane wave above, $\psi = \psi(x,t)$, but more generally, $\psi = \psi(x,y,z,t) = \psi(\mathbf{r},t) = e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)}$ throughout the following discussion.

Take the partial derivative of this expression with respect to t:

$$\frac{\partial\psi}{\partial t} = -i\omega e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)} = -i\omega\psi$$

Or take the partial derivative with respect to x:

$$\frac{\partial\psi}{\partial x} = ik_x e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)} = ik_x\psi$$

Where k_x is the x component of the wavevector, $\mathbf{k}$. Alternatively, we could take the space gradient, $\vec{\nabla} \equiv \partial/\partial x\ \hat{\mathbf{x}} + \partial/\partial y\ \hat{\mathbf{y}} + \partial/\partial z\ \hat{\mathbf{z}}$, and get:

$$\vec{\nabla}\psi = i\mathbf{k}\psi$$

Now, in a small departure from the normal course, let us limit our attention to the special case of luminal waves, which propagate at, and only at, the characteristic velocity, $c = \omega/k$.

We saw in Chapter 4 that the transformations of wave *momenta* are identical to the Doppler transformations of *frequencies*, and it follows that the magnitude of the momentum of a luminal wave, p, is proportional to its frequency, but since we also had $E = cp$, energy is also proportional to frequency. The proportionality constant has, of course, been determined for energy quanta to be Planck's constant, h and the all pervasive relation in Physics for both matter quanta and radiation quanta is the Planck relation:

$$E = \hbar\omega\ ,$$

where $\hbar = h/2\pi$. Using the expression for $\partial\psi/\partial t$ above,

$$E\psi = i\hbar\frac{\partial\psi}{\partial t} = \Big(i\hbar\frac{\partial}{\partial t}\Big)\psi\ ,$$

so that the energy, E, of a quantum of luminal wave energy can be identified with the mathematical operation, $i\hbar\,\partial/\partial t$ acting on a luminal wave. The velocity of the wave is the velocity of a point of constant phase, for which $\mathbf{k}\cdot\mathbf{r} - \omega t = constant$, and differentiation gives the propagation speed, c, as: $c = |d\mathbf{r}/dt| = \omega/k$. From $E = cp$, we get the de Broglie relation, $p = \hbar k$, and the x component of momentum is:

$$p_x = \hbar k_x \Rightarrow p_x\psi = \frac{\hbar}{i}\frac{\partial\psi}{\partial x} = \Big(\frac{\hbar}{i}\frac{\partial}{\partial x}\Big)\psi\ .$$

These expressions are routinely used to define energy and momentum "operators" as follows:

$$\hat{E} = i\hbar\frac{\partial}{\partial t} \quad , \quad \hat{p}_x = \frac{\hbar}{i}\frac{\partial}{\partial x} \quad , \quad \hat{p}_y = \frac{\hbar}{i}\frac{\partial}{\partial y} \quad , \quad \hat{p}_z = \frac{\hbar}{i}\frac{\partial}{\partial z}\ ,$$

where a carat over a symbol signifies that it's an operator. Although operators aren't usually taught at high school, the idea is very simple. They are the mathematical operations as separated from the functions to which they are applied. This can be very helpful because a lot of manipulations can be carried out amongst the operators without knowing the specifics of the function to which they are applied.

To illustrate this, consider the combination of the two operations "take the partial derivative with respect to x" and "multiply by x", applied to some completely unknown function $f(x)$. One might multiply f by x and then apply $\partial/\partial x$ or *vice versa.* The result is not the same, it depends on the order in which the operations are applied. In fact:

$$(\frac{\partial}{\partial x}x)f = f + x\frac{\partial f}{\partial x}$$

whereas with the order of operations reversed:

$$(x\frac{\partial}{\partial x})f = x\frac{\partial f}{\partial x}$$

so that:

$$(\frac{\partial}{\partial x}x - x\frac{\partial}{\partial x})f = f$$

Looking at that in terms of operators we can write:

$$\frac{\hat{\partial}}{\partial x}\hat{x} - \hat{x}\frac{\hat{\partial}}{\partial x} = \hat{1}$$

where $\hat{1}$ is the operation "multiply by 1". This particular example is rather important in Quantum Mechanics, because it is the basis of the Heisenberg Uncertainty Principle. It is an example of operators that don't commute which just means that the order we carry out the operations is important. This is already familiar from high school matrix algebra, where in general for the multiplication of two matrices, A and B, $AB \neq BA$.

If we were to treat those same matrices as operators, $\hat{A}$ and $\hat{B}$, then the "commutator" is defined as:

$$[\hat{A}\,, \hat{B}] = \hat{A}\hat{B} - \hat{B}\hat{A}$$

Some operators commute, like $\hat{E}$ and $\hat{p}_x$ above (or $\hat{p}_y$ or $\hat{p}_z$), in which case the commutator is zero, $\hat{E}\hat{p}_x - \hat{p}_x\hat{E} = 0$ and so on. Others don't, in which case the commutator is nonzero. Finally, there are also operators that "anticommute", for which:

$$\hat{C}\hat{D} + \hat{D}\hat{C} = 0\ .$$

Now, that's all the math we require, so let us look at the problem Dirac had to solve in order to reach the Dirac Equation.

5.2.2 Derivation of the Dirac Equation for Physical Waves

One starts with the relativistic energy-momentum relation for particles (which we derived in Chapter 4):

$$E^2 = P^2c^2 + (M_0c^2)^2$$

The first part of the question is how to turn this Mechanics equation into a relativistic wave equation, which is what the above energy and momentum operators do for us. We know that, for luminal waves, $E\psi = \hat{E}\psi$ so that $E^2\psi = \hat{E}^2\psi$ and similarly for the momentum (the square of an operator is obtained by applying it twice in a row). After multiplying the equation above by ψ, and replacing the energy and momentum variables with the corresponding operators, we get:

$$\{\hat{E}^2 - [\hat{\mathbf{p}}^2c^2 + (M_0c^2)^2]\}\psi = 0$$

This is the well known Klein-Gordon equation, and it is a proper, relativistic wave equation[2]. Unfortunately, it doesn't apply to any of the common sub-atomic particles. The Klein-Gordon equation contains second order derivatives with respect to space and time (because $\hat{E}^2 = -\hbar^2\partial^2/\partial^2 t$ and similarly for $\hat{\mathbf{p}}^2$). This led to some issues that were considered, at the time, to be problematic and the equation had all but been discarded[3].

Dirac was looking for a wave equation incorporating the relativistic energy-momentum relation that contains only first order derivatives, *i.e.* the operators $\hat{E}$ and $\hat{\mathbf{p}}$, not their squares. Now, if one could find an operator $\hat{O}$ such that:

$$\hat{O}^2 = \hat{\mathbf{p}}^2c^2 + (M_0c^2)^2\,,$$

then the Klein-Gordon equation would become:

$$\{\hat{E}^2 - \hat{O}^2\}\psi = 0\,,$$

and, as long as $\hat{E}$ commutes with $\hat{O}$, one could then write:

$$(\hat{E} + \hat{O})(\hat{E} - \hat{O})\psi = 0\,.$$

Since ψ cannot be assumed to vanish, one of the terms in brackets would have to vanish and the sought after equation would be either:

$$(\hat{E} - \hat{O})\psi = 0$$

or

$$(\hat{E} + \hat{O})\psi = 0\,,$$

depending on whether one takes the positive or negative "square root" of the operator $\hat{\mathbf{p}}^2c^2 + (M_0c^2)^2$. Because both the (magnitude of) momentum and the mass are positive numbers, it's somewhat more natural to take $\hat{O}$ as the positive "square root" and discard the second form above.

[2] Although it was not considered to be as "proper" in the 1920's as it is today.

[3] These issues have been largely resolved and nowadays the Klein-Gordon equation is considered to represent spin-0 bosons. Such particles were unknown at the time.

This brings us to the climax of the story, how Dirac found the square root of the operator $\hat{\mathbf{p}}^2c^2 + (M_0c^2)^2$. If we were dealing with real numbers, one might suggest that perhaps: $p^2c^2 + (M_0c^2)^2 = (apc + bM_0c^2)^2$??? That would imply:

$$(apc + bM_0c^2)^2 = (apc)^2 + 2(apc)(bM_0c^2) + (bM_0c^2)^2 \,,$$

and we would need $a^2 = b^2 = 1$, in which case we could never remove the cross term, because it is proportional to $ab = \pm 1$. However, if a and b were matrices, say $\hat{\alpha}$ and $\hat{\beta}$, we know that matrices don't necessarily commute and the cross term would have to be written out fully, not as $2(apc)(bM_0c^2)$, but as:

$$\hat{\alpha}\hat{p}c\;\hat{\beta}M_0c^2 + \hat{\beta}M_0c^2\;\hat{\alpha}\hat{p}c\,.$$

As long as $\hat{\alpha}$ commutes with $\hat{p}$, the crossterm would vanish if:

$$\hat{\alpha}\hat{\beta} + \hat{\beta}\hat{\alpha} = 0\,,$$

which is to say if the $\hat{\alpha}$ and $\hat{\beta}$ operators anticommute.

Things are actually just a little bit more complicated because there are three components of momentum and $\hat{p}^2 = \hat{p}_x^2 + \hat{p}_y^2 + \hat{p}_z^2$. Dirac needed to find four operators, $\hat{\alpha_x}$, $\hat{\alpha_y}$, $\hat{\alpha_z}$ and $\hat{\beta}$, all of which anticommute with each other, all of which commute with the momentum operators and all of which satisfy $\hat{\alpha_i}^2 = \hat{\beta}^2 = \hat{1}$.

He knew that a similar problem had been solved by Pauli, who had introduced a set of 2 x 2 matrices in order to modify the Schroedinger Equation to include an energy term related to a particle's angular momentum. In the resulting equation, the Pauli Equation, the 2 x 2 matrices act on a 2-component wave function, usually written like a column vector:

$$\psi = \begin{bmatrix} \psi_1 \\ \psi_2 \end{bmatrix}$$

There was no solution to Dirac's problem with 2 x 2 matrices, but Dirac soon identified a set of four 4 x 4 matrices with all the necessary properties. These matrices just contain some real and imaginary numbers. A new equation could finally be written down whose solution would be a 4-component wave function, also usually written like a column vector:

$$\psi = \begin{bmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \end{bmatrix}$$

We have now identified the operator $\hat{O}$ above. In the model notation it is:

$$\hat{O} = \overrightarrow{\alpha} \cdot \hat{\mathbf{p}}\, c + \hat{\beta}\, M_0\, c^2$$

Where $\overrightarrow{\alpha}$ is a vector of matrix operators, $(\hat{\alpha}_x, \hat{\alpha}_y, \hat{\alpha}_z)$, and $\overrightarrow{\alpha} \cdot \hat{\mathbf{p}} \equiv \hat{\alpha}_x\hat{p}_x + \hat{\alpha}_y\hat{p}_y + \hat{\alpha}_z\hat{p}_z$ is the usual inner (or dot) product of two vectors familiar in high school except that the components of these vectors are 4 x 4 matrices on the one hand and differential operators on the other. Dirac's specific choice for these matrix operators is not unique and there is nothing to be gained from reproducing them here.

The resulting equation is the Dirac Equation. In the present notation, it reads like this:

$$\{\hat{E} - \overrightarrow{\alpha} \cdot \hat{\mathbf{p}}\, c - \hat{\beta}\, M_0 c^2\}\psi = 0\,.$$

It will be recalled that we took a "small departure from the normal course" above, by defining the energy and momentum operators only for the special case of luminal waves, and the reader my wonder why I took what looks like a backward step, from the general case to the special case of luminal waves. After all, the same operators are also used with the Schroedinger Equation in non-relativistic Quantum Mechanics, with the more general meaning.

This book is about the existence of a physical model, with real, physical waves. Defining the energy and momentum operators as above shows that the Dirac Equation isn't merely a quantum mechanical equation. It is also an equation valid for physical luminal wave systems. In fact, we have just derived the Dirac Equation for physical waves[4], from conservation of linear momentum alone.

As far as the predictions of the Theory are concerned, physicists soon found that the usual probability interpretation of Ordinary Quantum Mechanics could not be sustained with the new equation. They were somewhat guided by Ordinary Quantum Mechanics, but the formal interpretation of the Dirac Equation - *i.e.* how one obtains predictions of observable quantities from it - had to be worked out as a standalone exercise. The linearity of the Dirac Equation brings into play all the machinery of linear algebra so, for the formal interpretation, it is more or less "straightforward" to introduce wavefunction solutions constructed from eigensolutions corresponding to all the possible observable outcomes upon measurement. There is nothing in the wavefunction concept introduced as part and parcel of that exercise that is incompatible with the physical model. Instead of reworking the formal interpretation of the Schroedinger Equation, they could equally well have started with an understanding of the Dirac Equation as a physical model with wavefield solutions. The Dirac wavefunction might then be reworked to fit Schroedinger's Equation[5] instead of vice versa.

All of the resulting Physics can now be thought of as fruit of the Conservation Laws alone. That is why we took the backward step.

There are, however, many ways to write down the Dirac Equation, the set of 4 x 4 matrices chosen by Dirac is not unique, and neither are the solutions. Altogether, what this implies is that one cannot blithely presume a one-to-one correspondence between the quantum mechanical eigensolutions and physical wave solutions.

Let us now consider, then, what *can* be gleaned from the quantum mechanical results.

The first thing to notice is that the spin of the particle is automatically included by virtue of the α matrices. The physical fields, ψ_1 - ψ_4, in the model are physically coupled together because the time evolution of any one of these

---

[4]Up to the value of the constant $\hbar$. As a technical point worth noting, in subluminal systems such as the electron, the global field is typically a superposition of lower level, pointlike, cellular solutions (See Chapter 6). The global angular momentum of the quantum, $h/2$, is a sum over the angular momenta of these parts, so we cannot say that the Dirac Equation applies locally to the lower level field elements exactly as written (*i.e.* including $\hbar$).

[5]In a relativistic world, this is arguably a more logical path to the development of Ordinary Quantum Mechanics.

fields is coupled to space derivatives of the others via the alpha matrices. Now, "physically coupled" simply means interacting. We have a multi-part wave system with fields that are interacting with each other such that they execute bounded motion, and we will be discussing the details of these internal field movements (of the electron) later.

This situation, with multiple interacting fields, immediately raises the important point first mentioned in Chapter 3. If the fields interact with each other, then surely one requires a non-linear theory, and if so, why is the Dirac Theory linear? It is linear because the equation is written in terms of the energy-momentum variables, which are Mechanics quantities subject to conservation Laws, rather than in terms of field variables, like **E** and **H** in Electromagnetics.

For a complete physical model, expressed in terms of field variables, one would indeed have to use a nonlinear field theory to specify the interactions between fields that cause them to spin around each other in just the right way. That would be a Dynamics theory. The Dirac Equation, as with other quantum equations, does not pretend to explain the internal forces that bind the field system together. It just makes an extremely pertinent and useful statement about energy-momentum conservation in the underlying particle dynamic system, independent of these nonlinear interactions.

The physicists are, technically speaking, very much aware of this. For instance, the subject is called quantum MECHANICS not Dynamics. However, I've never seen any official celebration of the fact that switching from a field variables description to an energy-momentum based formalism gets results *because* it circumvents underlying nonlinearities at the field level. It can't be celebrated because the very notion of physical structure has been rendered "unspeakable in quantum mechanics" [47]. What this analysis shows is that the energy-momentum density underlying the relativistic quantum mechanics formalisms is luminal. The formalism is, at its root, no more than the same simple, wave trajectories formalism that we used in Chapter 4.

The reader will now be completely unsurprised to learn that the formally derived velocity operator for the Dirac Equation is just $\overrightarrow{\alpha}\, c$, and that applying this operator to ψ returns a velocity whose modulus is c [48]. Of course, the electron is not moving at c. This velocity is a reference to the internal movements, as we might expect from an operator that couples the field momenta in a bounded motion system.

An equation has also been derived for the movement of the electron as a whole, the so-called "group velocity". I put that term in quotations because it has nothing to do with the usual conception in Ordinary Quantum Mechanics (and elsewhere) of the group velocity of a wave packet, and we'll return to this in due course. Here is the equation that governs translations of the electron system through space [49]:

$$\overrightarrow{\alpha}(\mathrm{t})\, c = \big(\overrightarrow{\alpha}(0)\, c - \frac{\mathbf{p}}{\mathrm{H}}\big)e^{-2i\mathrm{H}t} + \frac{\mathbf{p}}{\mathrm{H}} \tag{5.1}$$

Where the momentum, $\mathbf{p}$, and the Hamiltonian[6], H, are both constants and the group velocity is $\mathbf{v}_g = \mathbf{p}/\mathrm{H} =$ constant. The first term on the right hand side is routinely interpreted as the internal movements of the moving electron, the

[6]Just think of it as an expression for the total energy.

so-called 'zitterbewegung'. Its quantum mechanical expectation is:

$$\frac{< \psi \mid (\overrightarrow{\alpha}(0)\ c - \mathbf{v}_g) \mid \psi >}{< \psi \mid \psi >},$$

(where conjugation of the "ket" is implicit in the bra-ket notation). The term in the middle, $\overrightarrow{\alpha}(0)c - \mathbf{v}_g$, is the difference between a vector of matrix operators and an ordinary vector. It's expectation value (noting that all the $\hat{\alpha}_i$ have real eigenvalues) varies with v_g as $\sqrt{1 - v_g^2/c^2}$: The zitterbewegung slows down by a Lorentz factor as the group velocity increases. This is the basic cause of time dilation and it is completely unsurprising in a luminal wave system[7].

Overall, once Special Relativity was seen as a theory about luminal waves, we were able to identify a direct route from the definition of momentum, $p = mc$ or $\vec{\rho}_{\mathbf{p}} = \rho_m c\,\hat{\mathbf{k}}$, to Dirac's Theory of the electron. Unfortunately, the same cannot be said of Ordinary Quantum Mechanics. The only two things that inhibited physicists from clarifying all the above connections in the relativistic theory are the notion of spacetime (which we saw above is better thought of as a constant rate of spatiotemporal evolution) and the confusion over Ordinary Quantum Mechanics that begins (as we shall soon see) with the fact that it does not admit a physical wave interpretation.

While the energy and momentum operators do work for the Schroedinger Equation, the discussion in the next Section will emphasise the reasons why solutions to that equation generally cannot be interpreted as physical waves[8]. One cannot help but think that much of the ensuing confusion and angst over Ordinary Quantum Mechanics has been uncritically and inappropriately imported to the relativistic domain, on the basis that ("obviously") the latter would be far more complicated and mysterious than the former. In fact, the relativistic theory is the easy one.

Let us now turn our attention to the interpretive difficulties inherent in nonrelativistic Quantum Mechanics' Schroedinger "pseudo-wave" equation.

5.3 The Schroedinger Equation

The time dependent Schroedinger equation is:

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

where $\nabla^2 = \partial^2/\partial x^2 + \partial^2/\partial y^2 + \partial^2/\partial z^2$ is the Laplacian and $V(\mathbf{r}, t)$ is the potential energy the particle has by virtue of its position in an external field and (since we are not in the physical model here) the particle mass is the lowercase m. Recalling the energy and momentum operators from Section 5.2.2 this is also written as:

[7] The same result can easily be shown for the root mean square average of the speeds of internal movements in luminal wave systems [43].

[8] A singular exception to that is the case of de Broglie's matter waves, which do satisfy the (time independent) Schroedinger equation and which must be interpreted physically. They propagate at c^2/v, where v is the particle velocity. In fact, the de Broglie relation above was originally found from Special Relativity. It will be shown later that this exception to the non-physicality of Schroedinger solutions actually proves the rule: de Broglie waves will be derived directly by superposing luminal waves, and the above result for luminal waves $p = \hbar k$, carries over naturally to luminal wave superpositions that generate de Broglie waves.

$$\hat{E}\psi = \left[\frac{\hat{\mathbf{p}}^2}{2m} + V(\mathbf{r},t)\right]\psi = \hat{H}\psi \; .$$

Since the kinetic energy in nonrelativistic Newtonian Mechanics is $mv^2/2 = (mv)^2/2m = p^2/2m$, this equation just makes another straightforward Mechanics statement: Total energy equals the sum of kinetic and potential energies.

Another form of the Schroedinger Equation, used for "stationary states", where the energy E is constant, is the time independent form:

$$E\psi = \left[\frac{\hat{\mathbf{p}}^2}{2m} + V(\mathbf{r})\right]\psi = \hat{H}\psi \; ,$$

where the energy, E, is a constant rather than an operator, and the potential is time invariant. There are many ways to derive, motivate or otherwise induce these equations. We shall focus here on the compromises inherent in the construction of these non-relativistic wave equations.

With the Dirac Equation, the reader was dragged through the analysis step by step to make the point that there is no "heuristics" involved in the relativistic theory. Nothing was broken or compromised. We started with the already derived relativistic energy-momentum relation, multiplied it by a luminal wave function, ψ, for which the energy and momentum operators had been shown. When we put the operators in there could be doubt the result was fully implied by the line above.

Can we really say that about the corresponding steps above, from the time differentials version to the operators version and on to the physical interpretation of the operators as the momentum and energy of a particle? We can, but there is a price to pay! Consider a plane wave solution to the Schroedinger Equation, with the potential energy term, V, set to zero:

$$\psi = e^{i(kx-\omega t)} \; ,$$

where the propagation velocity of the wave is now $v = \omega/k$. $\hat{E}\psi = \hbar\omega\psi$ and $\hat{p}\psi = \hbar k\psi$ as before, so $E = \hbar\omega$ and $p = \hbar k$. However, we also have in the Schroedinger Equation the nonrelativistic energy-momentum relation:

$$E = \frac{p^2}{2m} \; ,$$

which implies a dispersion relation (*i.e.* where the wave propagation speed is frequency dependent):

$$\omega = \frac{\hbar k^2}{2m} \; ,$$

so the propagation velocity of the wave is:

$$v = \frac{\omega}{k} = \frac{\hbar k}{2m} \; .$$

The standard path to try a physical interpretation of the Schroedinger Equation is, of course, to say that the particle is represented by a wave packet, a superposition of plane wave solutions in a spectrum of frequencies, for which the group velocity of the wave packet is shown to be:

$$v_g = \frac{d\omega}{dk} = \frac{\hbar k}{m} = \frac{p}{m} = v_p \ ,$$

where v_p is the velocity of the particle. Note that the group velocity here has nothing to do with the group velocity for the Dirac Equation, which manifestly reflects the conservation of momentum between particle and wave components.

As is well known, such attempts to look at the Schroedinger Equation as a physical model make no sense. Every wave in the packet has a momentum in the vicinity of the particle's momentum. They propagate at various speeds in the vicinity of half the particle velocity. Why? Given that they don't interact with each other, they cannot be finitely distributed and also propagate at around half the particle speed. These waves extend to infinity. Say goodbye to the idea that the wave's momentum is the space integral of its momentum density, and say goodbye to the idea that the particle momentum is the sum of component's momenta. Find an all new way to think about Conservation Laws[9].

The bottom line here is that the wave components of the Schroedinger wave packet have no physical interpretation. Unlike the Dirac Theory, Ordinary Quantum Mechanics could never make sense from a high school perspective. A different kind of formalism was required by the very structure of the Schroedinger Equation. Physicists applied the craft and they identified various formal approaches that extract empirical information from the Schroedinger Equation.

The standard "Copenhagen" approach asserts that each and every one of the waves in the packet "could be" the particle. Each one corresponds to a possible result upon actually measuring the particle's momentum but any notion of its existence is at best probabilistic. Really, the whole idea of the wave packet forming a spatially well-defined image of a particle is then a red herring from the Copenhagen point of view because it is an essential part of the Theory that it can never be experienced.

There is an alternative way to make this same point. The time independent Schroedinger Equation can be derived, or at least induced, directly from a relativistic wave equation with the time dependence suppressed, namely the Helmholtz Equation. The vital step is to replace the wave number, k, in the Helmholtz Equation with the (nonrelativistic energy-momentum) form, $\sqrt{2mE}/\hbar$.

In other words, the Schroedinger Equation is what you get when you take a perfectly good relativistic wave equation and break it by enforcing the non-relativistic energy-momentum relation. One cannot crudely impose an energy-momentum relation that is known to be wrong with no consequences in the algebra. Something had to break, the consequence was dispersion and what got broken is the physical interpretation as waves propagating in a medium. I have been referring to this Equation as a "pseudo-wave" equation because it is an attempted cross breed, half way between (Newtonian) particle concepts and (inherently relativistic) wave concepts. Despite having been very fruitful empirically, the shotgun wedding of these two irreconcilable ontological concepts could never be fully satisfactory.

There is a very widespread tendency to think of the Schroedinger Equation as a non-relativisic, no spin approximation to the Dirac Equation, which is correct

[9]They did, and it's internally consistent, but one has to countenance non-physical ideas as the basis for the conservation of physical quantities.

enough. However, the routine presumption is that the broken physical interpretation aspect of Schroedinger's equation extends to the Dirac Equation. Nothing could be further from the truth: It's the act of taking the non-relativistic approximation that breaks the physical interpretation. In a Lorentz Invariant world, this is really no surprise.

5.4 The Wavefunction as a Superposition of Observables.

It is almost unnecessary in a discussion on physical wave interpretations to comment further on the formalism of Ordinary Quantum Mechanics, since there is no physical wave interpretation. However, the standard formal interpretation of the Schroedinger Equation doesn't just ignore the physical interpretation problem above, it ignores it in an astonishingly fruitful way that is worthy of consideration in its own right. The proposition, as mentioned above, is that, for any given physical state of affairs, the complete solution to the Schroedinger Equation can be represented as a wavefunction that is a weighted sum over the "eigensolutions" of a "complete" set of "commuting" observables.

In mathematics, eigensolutions (and/or eigenvectors) are just the solutions, ψ_i, to the eigenvalue problem:

$$\hat{A}\psi_i = \lambda_i \psi_i \ ,$$

where $\hat{A}$ is some operator, typically a matrix. The eigenvalues of $\hat{A}$ are the constants, λ_i, and every eigensolution is associated to one of the eigenvalues.

Observables are represented in Quantum Mechanics by operators, for example the momentum operators that we encountered in the discussions on Dirac and Schroedinger Equations. The full and complete quantum state is then represented by a sum of terms, like this:

$$\psi = \sum_{i=1}^{N} c_i \psi_i \ ,$$

where the sum is taken over a "complete" set of "commuting" observables, the ψ_i being the eigensolutions for the corresponding set of operators. Mathematically, "complete" essentially means that the eigensolutions / eigenvectors must form an "eigenbasis" for the set of all possible states[10]. Commuting, as we already saw, just means that the order of the mathematical operations is unimportant.

In the real world it means that the temporal order of the measurements is unimportant (a recurring consideration in Chapter 7 with respect to spin measurements on *different* particles). Physically speaking, the meaning of a complete set of commuting observables is a set of measurements that a) can be performed in any order without changing the probability distributions for the results and b) provides maximal information about the particle's properties.

When measurements don't commute, like the x-momentum and x-position operators or spin measurements along different axes of the *same* particle, the state can be represented in alternative ways, using different complete sets of

[10]Analogous to how the unit vectors $\hat{x}$, $\hat{y}$ and $\hat{z}$ form a basis for all the possible space points.

eigensolutions, corresponding to a different complete set of commuting observables that also provides maximal information (albeit a different kind of information), and forms an alternative eigenbasis, by far the most common being the momentum basis and the position basis. Such eigenbases provide alternative ways of representing the same space of states.

When a measurement is performed, the quantum state "collapses": all of the terms in the above sum are eliminated except one. The result of the measurement is the eigenvalue corresponding to the one surviving ψ_i term and the probability that it will be the one to survive is proportional to the squared magnitude of the (complex) weight, c_i, in the superposition prior to the measurement.

One can see that this standard formalism is an abstract mathematical procedure, divorced from physical reality, but those are the rules. As beautiful as it is as a simple and effective procedure for generating predictions, let us consider the issue that bedevils this formalism from the point of view of philosophical local realism.

Consider experiencing a piece of fruit that is initially hidden inside a box. Let's say that we know in advance that it is either an apple or an orange and suppose that you can either open the box, look at the fruit and weigh it (before or after looking) and only then eat it, or you can eat it with your eyes closed and weigh it (before or after eating), and only then look at it. The observables are colour, weight and taste and their values are determined respectively by the corresponding measurement operations, looking, weighing and eating.

If you look at it first, the colour and weight observables commute, because looking at the colour does not change the weight or vice versa. On the other hand, the taste and colour observables do not commute, because eating a green apple or an orange orange leaves behind respectively a white core or a white and orange peel. Finally, taste and weight do not commute because eating a piece of fruit changes the weight of what is left behind to be weighed[11]. So, the quantum mechanical state would be either a sum over all the possible combinations of colour and weight or a sum over all the possible tastes.

This general kind of approach to specifying the physical state of a physical system is a conceptual nightmare. One result of it is that Ordinary Quantum Mechanics equates the term "reality" with the existence of well defined properties. It ends up insisting that a piece of fruit "is" the colour and the weight that it has, while also insisting that it "is" its taste. The underlying physical reality has been deleted.

Does the wavefunction, ψ, relate to physical reality in the sense of an ontological status or is it "just" information, that which is knowable about that which is real, in which case ψ has an exclusively epistemological status? In particular, does the collapse of the wavefunction upon measurement correspond to a real change in the physical state of affairs, or merely a change in our knowledge?

The answer to the second question is straightforward. The collapse is always a change in our knowledge and usually, but not always, it also corresponds to a real change in the physical state of affairs. The only answer to the first question is "yes". "Yes" it is intrinsically 100% epistemological and "yes" it also has an intrinsic ontological significance. That is the problem.

[11]The analogy is less than perfect because weighing first does not alter taste, but this won't interfere with the conclusion.

The discussion on this point in the Physics literature is robust and ongoing. See, for example, the PBR Theorem [50, 51], criticisms thereof, and various other "psi-ontic" theorems. These issues are well expressed by an oft used analogy:

Trying to interpret the wavefunction is like trying to unscramble an omelette.

Expressing the quantum state as a superposition of observables scrambled Ontology and Epistemology. In Philosophy, it would be called a category error. In Chapter 7, we are going to encounter the corresponding poisoning of the literature as it relates to EPR experiments.

The point of all this is that the whole discussion was moot because the physical interpretation was already broken (by the Schroedinger Equation) before the formal interpretation could ever be considered, but this is of no concern for our physical model because the probability interpretation simply does not apply to the Dirac Equation. It admits both a clear physical interpretation and practicable formal interpretations for the observable physics.

The other point to take note of is that the probability interpretation of Ordinary Quantum Mechanics is an especially useful thing that facilitates all manner of quantitative predictions. This point bears further emphasis.

Imagine that you have a perfect physical model that provides a clear cut specification of exactly what the reality is. Perhaps it is a very complicated nonlinear field model. Perhaps it is Einstein's dream. Let us assume that one can calculate with arbitrary precision the space integral over fields' momentum densities in the model and the exact location of the centre of inertia. That information is not enough by itself to make predictions in relation to the observables, the results upon actually measuring a particle's position and momentum, which requires interacting with the measured system.

The vexed connection between the reality underneath and the facts we encounter is a gordion knot that any physical model would have difficulty untying. The probability interpretation cuts straight through it and the fact is that it works, although, to the devil's delight, we are left describing Ordinary Quantum Mechanics as the theory that nobody understands.

On the relativistic side, things are slightly better. The main reason why the probability interpretation of Ordinary Quantum Mechanics did not carry over directly to the relativistic Dirac and Klein-Gordon equations (where we have an *a priori* reason to believe in the possibility of a physical interpretation) is that the Dirac and Klein-Gordon currents are not positive definite, so they cannot be interpreted as probability currents since a probability can't be less than zero, at least it shouldn't[12].

With these relativistic equations, the interpretation of the solution to the equation, ψ, is nuanced. For the Dirac Equation, the "first" interpretation of ψ_1 - ψ_4 is as a superposition of an electron in the spin-up state, an electron in the spin-down state, a positron in the spin-up state and a positron in the spin-down state.

However, there are many ways to formulate, manipulate and decompose these equations pertinent to various physical contexts. Perhaps it is not reaching too far to characterise the *status quo* by saying that some interpretations of both Klein-Gordon and Dirac equations lean in the direction of a charge density or

[12] Physicists have tried negative probability interpretations. Vive La Nouvelle Cuisine! [47].

even a signed energy density.

Overall, in Ordinary Quantum Mechanics, the projection postulate is an unexplainable *"Force Majeure"*, but with the relativistic equations there are at least some physicalist aspirations and projection looks physically very reasonable: The chance for something to happen is driven by the proportion of the system energy corresponding to an observable result. Hopefully the existence of an unequivocally physical, luminal wave interpretation of the Dirac Equation will further assist.

I took the view above that the mysterian preconceptions from Ordinary Quantum Mechanics have been uncritically exported to the relativistic theory on the basis that, surely, the latter would be even more mysterious than the former. In fact, only the relativistic theory makes physical sense, which is as it should be in a Lorentz Invariant world.

Nonetheless, up until now the result of all the quantum machinery is twofold: Everybody knows the answers, nobody understands the phenomena. After commenting briefly in the next Section on the quantum measurement problem, we shall put the quantum theory aside in the next two Chapters and consider some of the most archetypical quantum phenomena from the perspective of luminal wave superpositions.

5.5 The Quantum Measurement Problem

Assuming that many readers would already be familiar with this well known problem in the Foundations of Ordinary Quantum Mechanics, while others can easily google it, only the briefest overview of the problem is provided here.

It will immediately become apparent why the quantum measurement problem does not arise in relation to nonlinear field models.

Consider a quantum system, A, which can be in one of two states, $|A1>$ and $|A2>$, which interacts with a measurement device, B, in the initial state $|B_{ready}>$ such that when A is in state $|A1>$, the measurement device transitions from $|B_{ready}>$ to the classically observable state $|B1>$ corresponding to the result 1 and when A is in the state $|A2>$, B transitions to $|B2>$, corresponding to the classical result 2. According to the usual Quantum formalism, when A is in the superposition state[13] $(|A1>+|A2>)$ then B necessarily transitions to $(|B1>+|B2>)$, which is a superposition of classically observable states. However, in practice we only ever observe unique outcomes, either 1 or 2, and not superpositions.

This is resolved in the formalism by means of the collapse postulate which selects one or the other outcome randomly upon observation, but the question is raised "At what point does the collapse occur?" On its face, the formalism suggests that there is no collapse until a conscious observer experiences the result, but this is clearly such an unpalatable consequence that Schroedinger designed his "Schroedinger's Cat" thought experiment as a *"reductio ad absurdum"* to discredit that notion as ridiculous.

Overall then, collapse in the usual formalism is revealed as an *ad hoc* procedure that is not properly defined, which breaks the linearity of the Schroedinger Equation and the unitary evolution of the quantum state under it. In order to

[13] These states are unnormalised here, as the factor of $\sqrt{2}$ brings nothing pertinent.

remedy this shortcoming, at least one of three conceptually appealing notions in relation to the quantum formalism must be rejected. They are[14]:

1. COMPLETENESS: The wavefunction ascribed to a system by the usual quantum mechanical formalism provides a complete physical description.

2. LINEARITY: The wavefunction of a closed system always evolves in accord with Schrödinger's equation, which crucially is a linear equation. Clearly, the entire universe constitutes a "closed" system since by definition there is nothing outside it.

3. UNIQUE OUTCOMES: In everyday experiments, such as the experiment on the cat described by Schrödinger, there is a single, unique physical outcome that corresponds to the outcome we experience.

Clearly, the second of these propositions is immediately inappropriate for the context of nonlinear field models for the massive particles. Since we fundamentally require nonlinearity to form particles in the first place, any interactions that there may be between quantum systems must also be based in nonlinearity. The linearity of a Mechanics equation does not imply linear dynamics. Rejecting Universal linearity removes the measurement problem.

A real, physical collapse of the superposition occurred when the quantum system initially interacted at the microscopic level with the measuring equipment, but the measurement state, $(< A1B1| + < A2B2|)$, is retained in the algebra until the point of actual observation. This is not, however, a real superposition; the system is in either $|A1B1 >$ or $|A2B2 >$, we just don't know which. When the collapse postulate is applied upon observation there is a mere change in our state of knowledge.

This does not, however, close off the subject altogether.

Instead of the measurement problem, the question becomes "When and why are we able to use linear equations to chart the evolution between preparation and measurement for quantum systems despite the fact that such systems exist only by virtue of nonlinear internal interactions?".

A partial answer to this for individual, isolated quantum systems has already been provided by observing that the equations we use are Mechanics equations, subject to Conservation Laws. A Classical analogy sheds some more light on this still only partially satisfactory answer.

Consider the bolas, a system of two massive balls joined by an inextensible cord. It is prepared by flinging it in the air, where it rotates smoothly as long as it is in flight and the internal state displays a unitary evolution that can be described with a simple mechanics equation. However, we can say very little about the nature of the interaction that maintains a fixed distance between the two balls. Until the bolas interacts with some other system, we cannot tell the difference between an inextensible cord, a rod, an elastic cord or a spring. Only when it does interact do the nonlinearities become manifest.

This explanation shows that there is nothing untoward about unitary evolution in a nonlinear system under steady state conditions, but it is, again, only partial because many quantum mechanical analyses involve quantum systems that do interact with other systems between preparation and measurement, often creating real superpositions in the process. Quantum Mechanics retains the

[14]Thanks to Tim Maudlin for providing this nice summary on the point.

linear Schroedinger evolution and keeps track of all the possibilities until such time as we are able, even in principle[15], to obtain the relevant information.

In summary, although the nonlinear field model concept removes the measurement problem *per se*, the more salient underlying question about whether information, which is to say epistemology, is playing a direct role in Physical Theory remains unresolved. We shall return to this issue in Chapter 14.

[15] In a double slit experiment, for example, if we have detectors at the slits that *can* provide "which way" information the interference pattern disappears whether or not we actually avail ourselves of those results.

Chapter 6

Quantum Phenomena I: de Broglie Waves

6.1 de Broglie Waves

There are two Louis de Broglies. First is the young de Broglie whose famous 1924 Phd thesis [52] proposed the matter waves that nowadays bear his name. He is also the physicist who proposed the pilot wave model at the 1927 Solvay conference, which is the origin of de Broglie-Bohm pilot wave theory and Bohmian Mechanics.

The pilot wave is also commonly known as the "guiding" wave. The basic idea is that of a pointlike particle guided by an external wave analogous to the original de Broglie waves. This conceptually distinct approach will be discussed later, in Section 6.3. For the moment, let us address de Broglie's original idea in the way that he originally proposed it. It was an explicitly relativistic endeavour for de Broglie and the same applies here.

Considering the relationship $E_0 = \hbar\omega_0$ for a massive particle as seen from the comoving frame, de Broglie proposed that there should be, associated to the system, a "periodic phenomenon" of radian frequency ω_0 [52]. He had noticed a curious consequence of Special Relativity. On the one hand, when considered from a frame in which the particle is moving, its energy is increased by the factor γ and one can write $E = \gamma\hbar\omega_0$, implying that the frequency is higher. On the other hand, time dilation implies a reduced frequency.

This is readily understood from the perspective of Chapter 4: the periods of internal movements involve velocities *relative* to the centre of inertia, whereas the energy is found by integrating an energy density of fields propagating at the full speed, c.

De Broglie analyzed this using a method that he called the "harmony of phases", which we don't need to consider for the present purpose. Two ideas emerged. The first is the famous de Broglie wave. It is important to be quite precise about what was proposed and what was not. De Broglie proposed a correspondence, $p = \hbar k$, between the linear momentum of the particle, p, and the wavenumber, k, of an associated, superluminal phase wave. He did not propose that this was a wave of momentum p. In fact, he explicitly rejects that idea on the basis that a superluminal wave could not transport energy-momentum superluminally:

> *"Here we must focus on the nature of the wave we imagine to exist. The fact that its velocity $v = c/\beta$ is necessarily greater than the velocity of light c, (β is always less than 1, except when mass is infinite or imaginary), shows that it can not represent transport of energy. Our theorem teaches us, moreover, that this wave represents a spacial distribution of phase, that is to say, it is a "phase wave"."* — Louis de Broglie [52].

The second idea was to recognize that in order for the phase wave to be spatially distributed the underlying "periodic phenomenon" of frequency ω would have to be similarly distributed in space:

> *"Must we suppose that this periodic phenomenon occurs in the interior of energy packets? This is not at all necessary; the results of §1.3 will show that it is spread out over an extended space. Moreover, what must we understand by the interior of a parcel of energy? An electron is for us the archetype of an isolated parcel of energy, which we believe, perhaps incorrectly, to know well; but, by received wisdom, the energy of an electron is spread over all space with a strong concentration in a very small region, but otherwise whose properties are very poorly known. That which makes an electron an atom of energy is not its small volume that it occupies in space, I repeat: it occupies all space, but the fact that it is undividable, that it constitutes a unit."* — Louis de Broglie [52].

The de Broglie wave does not transport energy, but in order for it to exist, an underlying "periodic phenomenon", had to exist. In any physical wave model, if it exists, it has to have an energy density. Finally, it must have the same phase at every space point. De Broglie's idea of a space independent phase gets little attention, because when one thinks of a wave model one thinks initially of superpositions of plane waves or spherical waves and the fact is that no such system has a space independent phase. Recognizing that this was not an easy thing to visualise, de Broglie gave the following mechanical analogy:

> *"To make the last point more precise, consider a mechanical comparison, perhaps a bit crude, but that speaks to one's imagination. Consider a large, horizontal circular disk, from which identical weights are suspended on springs. Let the number of such systems per unit area, i.e. their density, diminish rapidly as one moves out from the centre of the disk, so that there is a high concentration at the centre. All the weights on springs have the same period; let us set them in motion with identical amplitudes and phases. The surface passing through the centre of gravity of the weights would be a plane oscillating up and down. This ensemble of systems is a crude analogue to a parcel of energy as we imagine it to be."* — Louis de Broglie [52].

De Broglie waves arise because of the dephasing of these systems when considered from the point of view of a moving observer. The time transformation for an observer moving at speed v in the +ve x-direction is:

$$t' = \gamma\Big(t - \frac{v}{c^2}x\Big) ,$$

where it is most notable that v/c^2 is the inverse of the de Broglie wave speed, c/β. What this relationship means is that, if all the springs have the same phase for the comoving observer, then they do not all have the same phase for an observer moving relative to the system. To see this, consider frames in standard configuration and assume that at $t = 0$ in the comoving frame, all the springs are simultaneously at a maximum. In the moving frame, the spring at $x = x' = 0$ is also at a maximum at $t' = 0$, but at every other location, $t = 0$ gives a different value for t' so the springs reach their maxima at different times in the moving frame. The de Broglie wave stems from this dephasing in moving frames.

There is an interesting question to bear in mind regarding de Broglie's mechanical analogy: What happens if, instead of changing observer, an identical acceleration is provided to each of the springs for an identical period of time, thereby boosting the system to the speed v in the frame of the original observer? The standard answer in Special Relativity [41] would be that the springs are no longer Einstein synchronised for a new comoving observer: They still appear "synchronous" to the original observer.

In order for the original observer to see the de Broglie waves, as he in fact does in the real world, the spring system would need to be resynchronised in the new frame according to the Einstein protocol. It will be shown later in this Section how this happens in luminal wave models.

Using the information de Broglie developed, we can now write the equation for a de Broglie plane wave propagating in the x-direction in the present notation, consistent with the usual energy and momentum operators.

The particle momentum is $P = \gamma M_0 V = \gamma \hbar \omega_0 V/c^2 = \hbar(\gamma\beta k_0)$, where ω_0 and k_0 are respectively the radian frequency and the wave number of the underlying luminal energy density. The wavenumber of the de Broglie wave is just $\gamma\beta k_0$. Note that the wavelength, $2\pi/k$, is infinite for $V = 0$.

The energy of the moving particle is $E = \hbar(\gamma\omega_0)$, so the basic plane wave form of a de Broglie wave is therefore:

$$\psi = Re\left\{e^{i(\gamma\beta k_0 x - \gamma\omega_0 t)}\right\} ,$$

where $Re\{\,...\,\}$ means "the real part of". Note that the propagation speed is $\gamma\omega_0/\gamma\beta k_0 = c/\beta$.

To obtain this superluminal form of wave, consider the general case of the superposition of two luminal waves, which is to say that we are superposing waves that are not simple plane waves, but a more generalised waveform (with space and time dependent parameters) as follows:

$$\psi_i = Re\left\{A_i(\mathbf{r},t)e^{i(\mathbf{k}_i(\mathbf{r},t)\cdot\mathbf{r} - \omega_i(\mathbf{r},t)t)}\right\} ,$$

where $i = 1, 2$, $\mathbf{r} = x\hat{x} + y\hat{y} + z\hat{z}$, and the main constraint is the constant propagation speed, $\omega_i(\mathbf{r},t)/k_i(\mathbf{r},t) = c$. In the most general case, the wave amplitudes, A_i, could be almost anything, complex numbers, vectors and so on, but let us assume here that they are real scalars, and that $A_1(\mathbf{r},t) = A_2(\mathbf{r},t) = A(r)$. This generalised kind of wave is still locally planar: the wave vector at any point is normal to a tangent plane to the wavefront at the same point. Note that the wave number, and hence the field momentum density, may change as it propagates along the (in general curved) field lines of the wave vector. These are waves in an interacting system.

Let us also assume that there exists one observer for whom the wave number and wave frequency are constants, $k_1(\mathbf{r},t) = k_2(\mathbf{r},t) = k_0$ and $\omega_1(\mathbf{r},t) =$

$\omega_2(\mathbf{r}, t) = \omega_0$ independent of x, y, z, and t. It will be found below that this observer is a comoving observer. For him, the superposition of ψ_1 and ψ_2 is the real part of:

$$\psi(\mathbf{r}, t) = A(r)\big[e^{i(\mathbf{k}_1\cdot\mathbf{r}-\omega_0 t)} + e^{i(\mathbf{k}_2\cdot\mathbf{r}-\omega_0 t)}\big] = 2A(r)\,cos\Big(\frac{\phi_1 - \phi_2}{2}\Big)e^{i\left(\frac{1}{2}(\phi_1+\phi_2)\right)}\,, \tag{6.1}$$

where $\psi = \psi_1 + \psi_2$, $\mathbf{k}_i = k_0\,\hat{\mathbf{k}}_i(\mathbf{r}, t)$, $\hat{\mathbf{k}}_i$ is the unit wavevector of the i^{th} field and $\phi_i = \mathbf{k}_i \cdot \mathbf{r} - \omega_0\, t$. The result is obtained using the definition of $e^{i\theta}$ and the usual trigonometry rules for summing sines and cosines.

Now, consider the same system from the point of view of an observer moving at speed v in the -ve x direction, whose coordinates are (x', y', z', t') with the frames in standard configuration.

The frequencies and wave vectors for each wave in the primed frame can be evaluated using the relativistic Doppler shift and aberration results or (more efficiently) using the wave transformation shown in Chapter 4.

$$\omega_i' = \gamma\omega_0(1 + \beta\, cos\,\theta_i) = \gamma\omega_0\big(1 + \beta\frac{k_{ix}}{k_0}\big) \tag{6.2}$$

$$\mathbf{k'}_i = \gamma(\beta k_0 + k_{ix})\,\hat{x}' + k_{iy}\,\hat{y}' + k_{iz}\,\hat{z}' \tag{6.3}$$

Where $\beta = v/c$, $\gamma = \sqrt{1/(1-\beta^2)}$, $\mathbf{k}_i = k_{ix}\,\hat{x} + k_{iy}\,\hat{y} + k_{iz}\,\hat{z}$, θ_i is the angle between $\mathbf{k}_i$ and the x-axis and we have used the fact that $c = \omega_1'/k_1' = \omega_2'/k_2' = \omega/k$.

Using those equations with the inverse Lorentz Transformation: $x = \gamma(x' - vt')$, $y = y'$, $z = z'$ (frames in standard configuration) and the relation $\omega\beta = kv$ gives expressions for $\phi_1' - \phi_2' = (\mathbf{k}_1 - \mathbf{k}_2)\cdot\mathbf{r}$ and $\phi_1' + \phi_2' = 2\gamma\beta k_0 x' - 2\gamma\omega_0 t' + (\mathbf{k}_1 + \mathbf{k}_2)\cdot\mathbf{r}$, and substituting these in (6.1) gives the superposition for the primed observer, which is the real part of:

$$\psi' = 2A\; cos\Big[\frac{(\mathbf{k}_1 - \mathbf{k}_2)\cdot\mathbf{r}}{2}\Big]\;\; e^{i(\gamma\beta k_0 x' - \gamma\omega_0 t')}\; e^{i\left[\frac{1}{2}((\mathbf{k}_1+\mathbf{k}_2)\cdot\mathbf{r})\right]}\,. \tag{6.4}$$

Note that this is written in a “mixed frame” form in order to facilitate determining conditions on the comoving system wavefield that lead to de Broglie waves, and matter beam interference phenomena, in the moving system. The term in the middle, $e^{i(\gamma\beta k_0 x' - \gamma\omega_0 t')}$, is the de Broglie plane wave propagating in the x-direction identified above. It is multiplied by two extra terms that are cosines and/or sines of $1/2\,(\mathbf{k}_1 - \mathbf{k}_2)\cdot\mathbf{r}$ and $1/2\,(\mathbf{k}_1 + \mathbf{k}_2)\cdot\mathbf{r}$. They are written here in the unprimed variables, so even for the unprimed observer, for whom the middle term is $e^{-i\omega_0 t}$, these two new terms describe the space dependence of the phase. If either of them is not equal to unity, the de Broglie wave interference pattern in a matter beam interferometry experiment would be washed out because the phases of waves from different, spatially separated slits would be effectively random.

To illustrate this, consider the case of spherical waves which propagate radially into and out from the centre of the system, as in the usual Classical model [53]. In this case, $\mathbf{k}_1 = -\mathbf{k}_2$, which makes the second term equal to unity, but not the first. The superposition for such balanced spherical waves is the real part of:

$$\psi = 2A\ cos(k_0 r)\ \ e^{i(\gamma\beta k_0 x' - \gamma\omega_0 t')}\ . \tag{6.5}$$

With reference to Figure 6.1, it is readily seen that this does not produce the usual interference pattern at the screen in, say, a 2-slit interference experiment. For example, if we consider screen regions where the path difference is half the de Broglie wavelength, so that the (exponential) de Broglie term by itself would interfere destructively, (6.5) may exhibit anything between constructive and destructive interference, depending on the much shorter wavelength of the cosine term. Furthermore, Figure 6.1 shows the space dependence of the phase in the x-direction, but there is also a space dependent phase in the y and z directions, which are in the same plane as the slit system. The result is that different slits are illuminated with a different amplitude and phase, which again washes out any interference phenomena. Finally, de Broglie waves are also solutions to the Klein Gordon and time independent Schroedinger Equations, whereas (6.5) is not. The conclusion here is that wave models constructed from radially propagating waves are inconsistent with the observed phenomena.

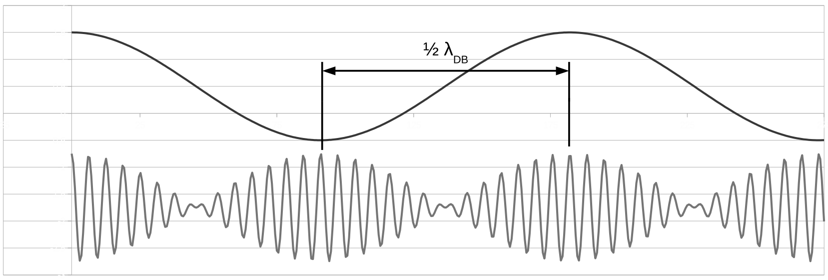

Figure 6.1: *The de Broglie wave (top) and the real part of (6.5) plotted against* x' *for* $t' = 0$, $y' = z' = 0$ *and* $\beta = 0.0345$, *an example where the spherical wave model leads to constructive interference in screen regions where the de Broglie wave alone interferes destructively.*

To make the offending cosine term above equal to unity, we require $\mathbf{k_1} \cdot \mathbf{r} = \mathbf{k_2} \cdot \mathbf{r}$, rather than $\mathbf{k_1} \cdot \mathbf{r} = -\mathbf{k_2} \cdot \mathbf{r}$. The second new term then becomes $e^{i\mathbf{k}_1 \cdot \mathbf{r}}$, and after taking the real part, the full superposition is:

$$2A\left(cos\left(\gamma\beta k_0 x' - \gamma\omega_0 t'\right) cos\left(\mathbf{k}_1 \cdot \mathbf{r}\right) -\ sin\left(\gamma\beta k_0 x' - \gamma\omega_0 t'\right) sin(\mathbf{k}_1 \cdot \mathbf{r})\right)\ ,$$

which is:

$$2A\, cos\left(\gamma\beta k_0 x' - \gamma\omega_0 t' + \alpha(\mathbf{r})\right)\ ,$$

where $\alpha(\mathbf{r}) = \mathbf{k}_1 \cdot \mathbf{r}$. This is another planar wave with a space dependent phase. Different slits are illuminated with different phases and there is no interference pattern. Although it is unlikely he was thinking of luminal wave superpositions, de Broglie would certainly have been aware of the general issue because it is the reason for his mechanical analogy. To obtain de Broglie's space independent phase in a luminal wave model we require $\mathbf{k_1} \cdot \mathbf{r} = \mathbf{k_2} \cdot \mathbf{r} = 0$.

This provides critical information about the wavefield structure of any luminal wave model of a massive particle. For every wave trajectory in the unprimed system, $\mathbf{k} \cdot \mathbf{r} = \mathbf{0}$. What this means is that the field lines of the wavevector, *i.e.* the wave trajectories, must lie on the surface of a sphere. Such a field is not moving through the unprimed observer's inertial frame, so his reference frame, the one where ω and k are the constants ω_0 and k_0, is the comoving frame of the particle. Upon substituting $\mathbf{k_1}\cdot\mathbf{r} = \mathbf{k_2}\cdot\mathbf{r} = 0$ into equations (6.1) and (6.4), this now precisely implements de Broglie's original formulation where a distributed, periodic phenomenon of frequency ω_0 with a space independent phase in the comoving frame, exhibits de Broglie phase waves in moving frames according to:

$$\psi = 2Ae^{-i\omega_0 t} \qquad \rightarrow \qquad \psi' = 2Ae^{i(\gamma\beta k_0 x' - \gamma\omega_0 t')} \,. \tag{6.6}$$

This constraint, $\mathbf{k}\cdot\mathbf{r} = \mathbf{0}$, has a direct counterpart in Special Relativity, where there is a subgroup of the Lorentz Group, known as the "little group", which (as mentioned in Section 4.5) is the group of transformations that preserve the linear momentum of a particle. As far as a particle whose momentum is $P = 0$ is concerned, the little group further reduces to the ordinary rotations group, $SO(3)$, which means that in its comoving frame a particle evolves under rotations, which in turn means that internal movements lie on the surface of a sphere, and so:

Luminal wave models that satisfy Special Relativity's little group generate de Broglie waves.

Thus far, what we have is an electron field structure underlying de Broglie waves that is constituted by a physically widely distributed "periodic phenomenon" of frequency ω_0, in which any field line of the wave vector exists on the surface of a sphere. But how widely is this field distributed, and what is / are the size(s) of spheres?

In matter interferometry experiments the slits in typical grids are spaced on the order of 500nm, so the de Broglie wave must be planar in an extended region of at least several microns, which is at least tens of thousands of times larger than the classical Compton radius of the electron. It will be argued below that in typical interferometry setups the planar region of the de Broglie wave covers the full width of the electron beam.

This raises a challenging question: How can such a widely distributed field oscillate at a single frequency, when the wavefield is subject to all of the following constraints?

1. In the comoving system, any field line of the wave vector evolves at c on the surface of a sphere.
2. The wave frequency, and the wavelength along the trajectory, are constrained by $E = \hbar\omega = 2\pi\hbar c/\lambda = mc^2$.
3. The phase along any trajectory must satisfy a resonance condition.
4. The space independent phase applies to the standing wave as a whole, not just a particular trajectory or subset of trajectories.

Consider all the possible curvilinear paths on the surface of a sphere, assuming closed trajectories subject to a resonance condition, such that the trajectory length must be equal to an integral number of wavelengths: $L = n\lambda_0$, where λ_0 is the wavelength in the comoving frame, corresponding to k_0. Any other assumption would conflict with everything we know about standing waves. In order to illustrate the issue, consider the specific case $n = 1$. In the comoving system:

$$E = \hbar\omega_0 = 2\pi\hbar c/\lambda_0 = M_0 c^2$$

For the purpose of illustration, let us also assume for the moment a circumferential trajectory. In this case, the radius of the sphere is $\lambda_0/2\pi$:

$$r_s = \frac{\hbar}{M_0\, c} = \frac{\lambda_c}{2\pi}\ , \tag{6.7}$$

where $\lambda_c = \lambda_0 = 2.426 \text{ x } 10^{-12}$ metres is the usual Compton wavelength, and r_s is the usual Compton radius.

This particular trajectory choice is not plausible. Looking back at (5.1), for the transport of the particle through space, the exponential in the first term (for the internal movements) is e^{-2iHt}. For the free particle in the comoving frame, this is just $e^{-2i\hbar\omega_0 t}$, which means that the zitterbewegung frequency is twice the particle frequency. A circumferential trajectory, however, has only one frequency associated to it.

In general, closed curvilinear paths on the surface of a sphere can be associated to at least two frequencies: the energy frequency, corresponding to the Compton wavelength, and one (or more) "geometrical" frequencies as the trajectory weaves around on its spherical surface.

Closed trajectories do exist that have a geometrical frequency (analogous to the zitterbewegung frequency) that is exactly twice the energy frequency associated to a "round trip" around the whole sphere. The seam on a tennis ball is an obvious example.

Solving for suitable wave trajectories for the electron in closed form is nontrivial (probably involves solving unknown nonlinear field equations as opposed to "merely" the Dirac equation). It is also unimportant for the present purpose[1].

What is relevant here is that, for any candidate $n = 1$ trajectory, $r_s \leq r_c$. The size of such a sphere is many orders of magnitude smaller than the known minimum physical extent of de Broglie waves, not to mention other quantum phenomena, especially EPR correlations that can involve distances measured in kilometres. We also require a space independent phase, which is to say that the phase on the sphere has to be the same at every space point all over the field. If we were to think about the electron from the classical viewpoint, where the particle "is" the sphere, it would clearly not be possible to resolve this puzzle.

Instead, the only solution is to take de Broglie's original concept, communicated with his weight and spring mechanical oscillator analogy, literally. The widely distributed field then consists of a myriad spherical resonators whose centres are distributed throughout the space and r_s is the radius of an individual resonator.

Once it is accepted that energy is a propagative phenomenon and that de Broglie waves exist, this kind of field microstructure is essentially inevitable. We then have an underlying field energy density distribution for the de Broglie wave, as follows:

[1]Although it is worth mentioning that there are closed trajectories with a geometrical frequency exactly twice the round trip frequency, and where the associated angular momentum is exactly half that of the corresponding circumferential trajectory. We shall return to this point in Chapter 12.

$$\rho_E(\mathbf{R}) = \rho_N(\mathbf{R}) E_{res}(\mathbf{R}) \ ,$$

where $\mathbf{R}$ is a global radial coordinate for the electron field system as a whole (with $\mathbf{r}$ being a local coordinate for any given resonator), $\rho_N(\mathbf{R})$ is the number density of resonators, and we will see later on why the energy per resonator, $E_{res}(\mathbf{R})$, generally depends on its position in the field system.

One might think of every resonator as a tiny clock, and all them proceeding in lockstep. As the particle encounters the tos and fros of existence in a busy Universe, every part of it is constantly being jostled about. How can they all be in phase? To maintain the synchronisation implicit in the space independent phase, the resonators would have to interact with each other. They must be coupled together by internal self-interactions of the electron field to form a single distributed entity. If the phase is disturbed in some part of the system then the phase of the whole system adjusts.

Any interactions that communicate, say, a pointlike phase disturbance to the rest of the system are limited by the speed of light. Mother Nature is in effect using the Einstein synchronisation protocol to establish simultaneity at a distance. Perhaps, like the superobserver of Silicon world, she knows about absolute simultaneity but there is no available mechanism she can use to implement it (at least there is no such mechanism in any physical model that satisfies the criteria in Section 2.3).

Concerning the interactions that couple the system together, what was found above is that the energy of each resonator is confined in a small region of space, on the order of the Compton radius. There can be no direct interaction between two resonators in different parts of the field. This is the local action principle as given in Section 4.2. In order for a pair of resonators to interact, they must overlap with each other. This field structure can be visualised as a 3-dimensional chain mail of interlocking trajectories, which together mediate any disturbances at a given point to other places in the field. It also corresponds to the "mattress model" that theorists often refer to in order to explain the idea of quantum fields [54].

This is not the usual idea of a smooth and continuous wavefield, like a water wave, and nor should it be. The water wave is a high level model of a complicated system of particles, and similarly a continuous global field solution of the Dirac equation is an amalgamation of pointlike subsystems that have been induced here from the de Broglie wave mathematics, not put in by hand.

Nonetheless, it is obvious from the way that we derived it that the basic form of the Dirac Equation applies both at the level of trajectories and at the global level. It is a linear equation, so solutions superpose. If a group of trajectories can form a cellular resonator solution located in a small, pointlike space region, then in a global field consisting of many such cells, the mechanics quantities for the whole quantum are found by summing over the corresponding mechanics quantities at the cell level.

The next question is whether sufficiently well-localised, pointlike solutions to the Dirac Equation exist. They do.

A vast number of pointlike models of the electron has been reported in the literature, for example [55, 56], usually with energy distributed on, or in, a sphere whose radius is related to the Compton radius[2]. While many properties

[2]The cell energy density here cannot be *perfectly* localised on, and only on, the surface of a

of the electron have been accurately reproduced, such models are not usually seen as pertinent because they don't begin to address the basic quantum issues of nonlocality, wave particle duality and so on. These issues manifest themselves in many ways, but perhaps the two most characteristic, quintessentially quantum phenomena are matter beam interferometry, see Section 6.2, and spacelike causal correlations, see Chapter 7. The classical idea of a particle as a pointlike corpuscle provides no insight into such cases.

However, if one thinks of such physical models not as models of the electron but as models of the spherical resonator component parts of an electron then a "classical-like" picture of the electron as a dynamic system becomes available in which all the known phenomena can be seen as physically reasonable in the sense of the modelling criteria of Section 2.3.

Such classical models are not generally solutions to the Dirac Equation, at least it is not usually shown. Pointlike solutions to the Dirac Equation have however been found by numerical analysis [57]. The method used was to set up an asymptotic boundary condition, as $r \to \infty$ in the far field, such that the magnitude of the wavefunction varies as e^{-r}/r in the region $r > r_s$. The key mathematical property of this asymptotic function to take note of is that it solves the equation $\nabla^2 F(r) = F(r)$.

This mathematical property is relevant because there is a group of equations in the Quantum Mechanics literature that goes directly to the question how the internal interactions within a wavefield determine the global form of solutions. They are the Quantum Hamilton-Jacobi equations, so named because of the similarity with the Hamilton-Jacobi Equation of Classical Mechanics.

In essence, any of the main quantum mechanical wave equations[3] can be decomposed into a pair of equations. This is done by writing the complex valued solution $\psi = a + ib$ in the polar form $\psi = Fe^{iS}$, where both F and S are real-valued functions of the coordinates. After substituting this into the wave equation and then collecting real and imaginary parts, two equations are obtained, both of which are satisfied by any solution to the original wave equation. One of these equations is a continuity equation and the other is a Quantum Hamilton-Jacobi equation.

The most common Quantum Hamilton-Jacobi equation is the one derived from the Schroedinger Equation:

$$\frac{\partial S}{\partial t} = -\Big[\frac{|\nabla^2 S|}{2m} + V - \frac{\hbar^2}{2m}\frac{\nabla^2 F}{F}\Big] .$$

The phase, S, is identified with the particle momentum, but the term of interest is the last term in brackets on the right hand side, which is known as the quantum potential, because it plays a similar role in the equation as the usual (Electric) potential, V. Since it is a potential energy term, the gradient of the quantum potential describes an underlying "quantum force" that is operating within the wavefunction, entirely distinct from the Electromagnetic force.

---

sphere: Continuity demands an asymptotic region outside the sphere and the notion of curved wave trajectories implies binding interactions, which implies multiple overlapping interacting fields inside the sphere. However, the resonator is characterised by the specific trajectories that lie on the particular spherical surface where the resonance condition is fulfilled and it is far easier to visualise just those characteristic trajectories.

[3] and also the Helmholtz equation.

Analogous terms show up in a wide variety of wavefields in general, not just in Quantum Mechanics.

Note that the quantum potential depends on $\nabla^2 F/F$. All of the various quantum potential terms stemming from the different wave equations have this feature in common. We saw above that in the region $r > r_s$, pointlike solutions to the Dirac Equation satisfied $\nabla^2 F = F$, so the quantum potential of these solutions is constant in the asymptotic region, with the implication that the quantum force of any one of the resonators above is zero for $r > r_s$. Their energy is confined within $r \leq r_s$, the quantum force is short range, and they don't interact with each other at a distance. This is a "good thing" for Pakula's analysis.

Instead, at the level of the global field, the distributed field structure of the quantum force is built up as the superposition of the myriad short range force fields of the individual resonators.

6.2 Matter Beam Interferometry

The role of the quantum potential terms is perhaps most clearly seen in the classical context of Electromagnetic waves propagating in a medium of variable refractive index. This is singularly appropriate because Hamilton-Jacobi Theory in Classical Mechanics had its roots in Optics, which is basically the same problem.

Consider a collimated Electromagnetic beam. The governing equation is the Helmholtz Equation. After a polar decomposition, one of the two resulting equations contains a "wave potential" term of the same mathematical form as the quantum potential [58]. In collimated beams, the gradient of the wave potential acts transverse to the wave propagation, and its action generates all the usual diffraction and interference phenomena.

What the wave potential does is to manage the physical distribution of the wavefield. As the conditions on the wavefield change - from the beam collimation apparatus, to the free field, to encounters with regions where the refractive index is inhomogeneous, the Helmholtz wave potential adjusts individual wave trajectories in response to the changing boundary conditions as determined by the environment as a whole. Since the wave system is fully represented by the complete set of trajectories, and the wave potential is itself calculated from the trajectories, [58] is a coupled wave trajectory formalism.

As shown in [58], essentially the same formalism also applies to the time independent Schroedinger Equation (which is appropriate for particle beams of fixed kinetic energy per particle). The de Broglie waves associated to such particles are, of course, solutions to this equation, but recall that the energy operator has been replaced by a constant in this time independent equation - there is no dispersion, the de Broglie waves have a definite speed, c/β, and the physical interpretation as a luminal wave superposition isn't blocked.

In field theories, where problems are routinely specified and solved on whole regions, the question of solving a wave equation inevitably involves boundary conditions. If a wave system obeys a given wave equation, for example the Dirac Equation, then the global fields must self-adjust, reconfiguring the field energy density distribution to fit whatever boundary conditions are being applied from time to time.

In a matter beam experiment, before the beam reaches any slit system what one has is a set of free electrons of some fixed kinetic energy (determined by the plate voltage of the gun) that emerged from a collimation apparatus. At this stage, the wavefield has only been constrained to a region of dimension equal to the beam width.

Under these conditions, it is only natural to conceive of each individual particle as a planar de Broglie wave that occupies the entire beam width, where the underlying energy density is a field of cellular oscillators drifting through space at speed v. This corresponds to the momentum basis in Quantum Mechanics: It corresponds to boundary conditions where particles are prepared with a well defined momentum and a weakly defined position. Of course, nothing prevents the mathematics from also describing them in the position basis (after a Fourier Transformation) but the particles are, more or less, in the momentum basis in the sense that the momentum basis is a natural choice for describing the physical state of affairs.

For another example, if we consider an electron in a Hydrogen atom, then it has a well enough defined position, but the vector momentum is all but completely undefined. The central boundary condition in this case has put it, (so to speak, more or less) in the position basis. If the position were *perfectly* localised then the position basis would be the natural choice for describing the physical state of affairs.

If we consider a matter beam interferometry experiment where the beam intensity is reduced so that there is only one electron in the apparatus at any time, then until it reaches the slit system, each electron "is" a planar wave system that occupies the entire beam. There is no question of using the wave description out of a lack of knowledge about the precise location of the particle. There is no particle. It is a widely distributed wavefield. When the wave impinges on the slit system, new boundary conditions are imposed, with two possible results: Recalling that the quantum as a whole is bound together by self-interactions at the resonator level, either the whole of the quantum passes through the slits, or it is absorbed. The ones that don't pass through don't get counted. For the ones that do, each transmitted electron emerges from all of the slits.

The de Broglie phase waves that are characteristic of the moving electron wave system then lead to a wave intensity interference pattern at the screen, where new boundary conditions are imposed in the process of the particle being absorbed somewhere on the screen. During that process, the electron wave interacts strongly with some elements in the screen, whose positions are well-defined, and it obtains a well-defined position as a result.

The wavefield has been reconfigured continuously, subject to boundary conditions. It has made a transition from conditions most naturally described in the momentum basis, to conditions most naturally described in the position basis.

On the other hand, if we contrive to identify the specific slit where the electron passes through then the very nature of any such intervention is to take it out of the momentum basis and put it in the position basis. The energy density distribution accommodates and the system as a whole then emerges (predominantly[4]) from just the one slit, and the interference pattern is washed

[4]Particles that are well located still have their far fields, but the energy density is already

out.

If there is a mystery in this, it is not wave-particle duality. Rather it is how the quantum potential "manages" the field during the absorption/slit location process. The quantum mechanical wavefunction collapses "instantaneously" in such cases, but this seems like an over idealisation. To the extent that the only limit on the operation of the quantum potential in reconfiguring the field energy distribution is the speed of light, we can safely assert that communication across the physical extent of a 10 micron wide beam can be achieved in around 3×10^{-14} seconds, which seems near enough to instantaneous collapse in the context of the velocities of electrons in typical matter beam interferometry setups. Furthermore, there is no good reason to think that the electron wavefield must have finished redistributing itself completely - with all transients decayed - before the spot appears on the screen.

6.3 The Problem with Guiding Waves

There are dozens of alternative interpretations of Ordinary Quantum Mechanics, which is to say alternate theories that produce either identical or experimentally indistinguishable results. One of the most popular is Bohmian Mechanics [59], which lays claim to being a "physicalist" interpretation because a purportedly real particle follows a definite trajectory in the theory. It has its roots in two ideas, namely de Broglie's "double solution" and the Quantum Hamilton Jacobi Equation.

De Broglie's double solution idea [60, 61] was essentially that the whole field-particle system should be thought of as a nonlinear field solution in two parts, a high energy density core (the particle) and a low intensity pilot wave that surrounds this corpuscular phenomenon and guides it along whatever specific trajectory it actually takes. In the basic Bohmian Mechanics formalism, the idea is similar except that the particle is now just a Classical point particle of mass m, put in by hand. The trajectories that it might take (depending on its initial position) are essentially the wave trajectories of the applicable solution to the Schroedinger Equation, which is to say the standard quantum mechanical wavefunction. These trajectories are evaluated using the Quantum Hamilton-Jacobi Equation in conjunction with the guidance condition $\mathbf{p} = \nabla S$, where it will be recalled from above that S is the phase of the wavefunction when written in the polar form used to decompose the Schroedinger Equation into continuity and Quantum Hamilton-Jacobi Equations.

We thus have a wave trajectory representation of a guiding field and a particle somewhere in it that is being pushed around by virtue of its interaction with the guiding field. Probability enters the picture because the initial position of the particle is unknown, so one begins with a probability distribution for the starting position of "the particle", which then has a notionally definite but functionally unknown position throughout the evolution of the system between preparation and measurement.

The reader may have seen pictures of the Bohmian electron trajectories in the solution for the Hydrogen atom, which were calculated by choosing some definite initial condition for the electron and then calculating the corresponding trajectory. The full set of trajectories is constructed by choosing a sufficiently

greatly reduced at the first neighbouring slit.

large number of initial conditions and performing the calculations in every case. This procedure raises an initial objection to the theory, namely that in none of the individual cases does "the" electron actually occupy the whole orbit.

In the ground state, it can be argued that in the real world the electron is constantly being jostled by the environment leading to effectively random transitions between neighbouring trajectories such that, over time, it occupies the whole orbit. However, this does not explain many of the excited states, like the p-orbitals, where the Bohmian trajectories of a single state are found to lie in disjoint physical regions. One really requires all of the trajectories to be occupied all of the time.

Another common objection to the theory is that it is explicitly nonlocal, as is clearly revealed by the fact that, in a system of N-particles, the wavefunction at any point and time depends on the positions of all N particles at the same time. This is not such a strong objection as it might appear at first blush. It can be argued that the N-particle positions are "really" being determined by the wavefunction, not the other way around, and that the mathematical form of the wavefunction corresponds to a logical inference rather than a causal impact.

Whereas the usual objection is that a (hypothetical) change in the position of any one of the particles instantly changes the wavefunction everywhere, one can equally take the view that it was the change in the global wavefunction that actually caused the change in the position of the particle in question and such changes are locally determined by the interaction between particle and guiding wavefield.

However, this kind of objection points to a different and more salient problem from the present perspective of physical models in the sense of Chapter 2. Where is the energy of the guiding field? Bohm obviously knew about this issue because he artfully described it as an "information field".

That is a category error. Either the guiding field is a physical field or it isn't. In fact, the role of the guiding field is very physical in the Bohmian theory. It transfers linear momentum to the particle, which is a "muscular" thing to be doing. It implies that the guiding field must contain energy-momentum. Of course, Bohm does not give us an energy density for the guiding field for the good reason that all of the system energy and momentum is already, by assumption, in the particles: The particle momentum is $\mathbf{p} = \nabla S$, its energy is $p^2/2m$ (in the non-relativistic version) and so on, and correspondingly in relativistic Bohmian Mechanics. There simply is no energy budget for the guiding field.

This is no problem from the point of view of Physics. Call it an information field, construct the Mathematics. From the present perspective, however, the whole theory runs on the essentially non-physical concept of a field that acts mechanically on a particle without being the physical stuff necessary to do the work.

Chapter 7

Quantum Phenomena II: EPR Correlations

7.1 Historical Background

It is well known that Einstein was not exactly onboard with the Copenhagen interpretation that emerged from the 1927 Solvay conference. He and Niels Bohr conducted a famous debate that continued for years, with Einstein proposing various thought experiments (trying to expose some error or internal inconsistency in the new Theory of Quantum Mechanics), and Bohr refuting them one by one, often with reference to his notion of "complementarity" (closely related to non-commuting observables and the proposition that one cannot look at an orange and eat it).

In 1935, two of Einstein's colleagues, Podolsky and Rosen, teamed up with him to deliver what they considered to be the fatal blow in an article entitled "Can Quantum-Mechanical description of physical reality be considered complete?". Today, this paper is just called by the initials of its authors: EPR [15].

It considered the case of a pair of particles, A and B in what follows, with position-momentum entanglement[1]. In particular, what was known about the system is that both particles had started at the origin with equal and opposite linear momenta. They had then become well separated from each other.

It will be recalled from Section 5.2 that position and momentum operators don't commute, and most readers already know the consequence - the Heisenberg Uncertainty Principle, which states that a particle's position and momentum cannot both simultaneously be known with arbitrary precision.

The EPR reasoning is like this. If we measure the position of particle A, the wavefunction of the two particle system collapses and we now know the exact position of particle B. Therefore B is in a definite position state. Since the particles are (arbitrarily) well separated, it was presumed that there is no possibility of any influence on particle B when we perform a measurement on particle A. Therefore, particle B always had a definite position, regardless of whether or not we actually perform the measurement on A.

[1] An entangled wavefunction includes terms relating to the combination of both particles rather than to each of them separately.

Similarly, we might have chosen instead to measure the momentum of particle A, in which case we would know the definite momentum of particle B. Since it was presumed there could be no influence, it follows that particle B always had both a definite momentum and a definite position. That violates the Heisenberg Uncertainty Principle, and therefore Quantum Mechanics is internally inconsistent.

Bohr's reply is famous, but for all the wrong reasons. He produced a word salad, served with a complementarity dressing, about the "very conditions of measurement". To this day, physicists continue to strive to decipher the meaning of his prose although most tend to agree with it anyway.

On the kindest interpretation, Bohr's argument could only make sense if he was asserting a physical influence of some kind on particle B when we perform one or the other measurement on particle A, but neither Bohr nor Einstein could countenance such a possibility: Surely, such instant interaction at a distance would not just be "spooky", but far worse: It would violate Special Relativity[2].

Here is the core of Bohr's response. It begins by stating that the "criterion of reality" in the EPR paper:

> *"contains an ambiguity as regards the meaning of the expression 'without in any way disturbing a system'."* — Niels Bohr [62],

and then explains that, when we perform a measurement on particle A:

> *"there is no question of a mechanical disturbance"* (of particle B) — Niels Bohr [62],

however, there is, according to Bohr's salad:

> *"an influence on the very conditions which define the possible types of predictions regarding the future behavior of"* (particle B) — Niels Bohr [62].

No one understands that sentence, however if we delete a few words he is talking about an influence on the possible types of predictions in relation to particle B. That's nonsense. Why would measuring particle A limit the possible predictions about B or measurements we could perform on it? Even in Quantum Mechanics, measurements on different particles always commute, and Bohr's notion of complementarity is all about when measurements don't commute.

For the record, here is the "Reality Criterion" that Bohr was targetting:

> *"If, without in any way disturbing a system, we can predict with certainty (i.e., with probability equal to unity) the value of a physical quantity, then there exists an element of reality corresponding to that quantity."* — Einstein, Podolsky and Rosen [15].

I understand that the EPR article was actually written by Podolsky, and Einstein didn't like the reality criterion because he felt that it muddied the water. Nonetheless, Podolsky was correct to recognise that an orange should

[2]We shall see that there is nothing spooky (nonlocal) involved and nor does it violate the Special Relativity in Einstein's 1905 article. What it violates is a later extension to the Special Theory, light cone causal analysis, made under the Metaphysical presumption that the ontology consists of atomist particles.

not be confused with the colour orange and he asserted that when we know for sure what the value of an observation will be if we measure it, then we are entitled to infer the existence of something physically real that CORRESPONDS to that value. He didn't say "causes" or "is colocated with" or any other strong formulation. The word he chose, "corresponds", goes as far as one needs to go in order to respect the category difference between Epistemology and Ontology, and no further. For me, this is neither unclear nor does it open any ambiguity door for Bohr.

It has been astutely observed that Bohr won the Einstein-Bohr debate mainly because what Einstein was saying could be understood. Here is what John Bell had to say about the passages above:

> *"While imagining that I understand the position of Einstein, as regards the EPR correlations, I have little understanding of the position of his principal opponent, Bohr. Yet most contemporary theorists have the impression that Bohr got the better of Einstein in the argument and are under the impression that they themselves share Bohr's view."* — J.S. Bell [47].

All of that is long ago and far away. The Bell Inequalities were published in 1965 [63]. They provide limits on the strengths of statistical correlations amongst the results of alternative measurements in a world constrained by Bohr and Einstein's common assumption of "no influence". The quantum mechanical predictions involve correlations that are stronger than what is allowed under the "no influence" assumption. This implies a causal influence from the intervention made when measuring one particle to the result of measuring the other, even when the measurements are spacelike separated.

The first experiments were reported early in the 1980s [64] and today many, many experiments on well separated entangled particles have been performed. Today, the experiments have verified the quantum mechanical predictions in excruciating detail and the question of a "mechanical influence" is very much on the table because the experiments routinely violate the Bell Inequalities. They show correlations between the results of spacelike separated measurements that cannot be replicated in the absence of a causal influence that affects particle B when we measure particle A.

7.2 The Definition of Local Realism

That last sentence will cause consternation in some parts of the Quantum Mechanics community [65, 66], but not others [67]. The discussion of EPR correlations has left Physics in a semantic muddle. Every nuanced detail has been dissected and argued, to the point where it is no longer possible to say anything without others finding fault in it. In this case, the modern conclusion *vis a vis* the meaning of Bell Inequality violations is that either "locality" or "reality" must fail. My crime in that sentence was to ignore the "reality fails" point of view. We shall return to discredit that argument in due course.

For now, what the "reality fails" case questions is the reality of an unobserved observable, even when it can be predicted with certainty. This has virtually nothing to do with the idea of realism in the present work, namely the idea of an objectively existing ontology.

The reasons for Quantum Mechanics' curious approach to such definitions begin with historical traditions that date back to Niels Bohr and beyond: Physics does not deal with physical reality, it deals with observables; the wavefunction is a superposition of observables so, when quantum physicists talk about realism, what they mean is the reality of properties, the reality of the colour and the taste, not the reality of the orange. If I have two boxes containing an apple and an orange, and I open the first box and see a green fruit, the reality fails argument is that I cannot conclude that the other one is orange because its colour does not exist until I actually look, even when I have performed the experiment a million times and every time I look at the second box after finding a green fruit in the first, it is always orange.

The key insight here is that the common sense idea of "reality" has been undermined in the EPR literature by starting from the Theory and ending at the definition. Similarly, where reality has been redefined with quantum mechanical observables in mind, locality was redefined to reflect Special Relativity's light cone causal analysis, so let us reflect on that.

Having gone to such admirable lengths, both in Quantum Mechanics and relativity theory, to reduce the world to point events, when the EPR community considers locality, it transitions seamlessly from the fact that we observe a point event to the notion of existence as pointlike properties to the unspoken presumption that the underlying physical reality - the ontological support for all our observables - is *fundamentally* pointlike. This is in spite of the fact that every quantum mechanical equation is a wave equation.

This Metaphysical presumption is much stronger than Podolsky's "corresponds". "Corresponds" does not imply that the ontological support for a point event is pointlike any more than a point event in a wave theory implies that the corresponding wave ontology is pointlike. Having just examined de Broglie waves, where the evidence was clear that they are physically much more widely distributed than the dot that we see on the screen, it should be equally clear to every reader that the mere fact of us finding a pointlike property at some place does not justify the Metaphysical inference that its ontological support (all the physical reality that CORRESPONDS to the observed property) exists exclusively at the same place, but that is the basis for local realism in the EPR literature.

The concept of local action is one of the oldest ideas in Physics, and it is a physical idea. It is the idea that for one physical (*i.e.* ontological) thing to affect, influence or interact with another, they must be touching. This is the original Principle of Local Action, or "near working" as Einstein had it. We can see that at its root it is explicitly Metaphysical - it belongs in the physical reality camp as opposed to the knowledge camp.

However, the usual classical understanding, consistent with the point particles of the Classical physical model, had always been that any interaction between two well separated particles must have been mediated by a field that had travelled from one to the other, so the interaction must be retarded, or delayed, by the light time between the particles. That "retarded interaction" idea very obviously involves an additional Metaphysical presumption about the nature of the ontology, namely that massive particles, like the electron, really are the tiny little corpuscles of our imagination, rather than, say, the widely distributed correlated fields of the present luminal wave approach.

The usual light cone analysis of causal relations between events in Special Relativity simply instantiates the retarded interaction worldview to the exclusion of every other kind of ontological support for our observations of massive particles except well-localised, atomist point particles. As we shall see in Chapters 8 and 9, retarded interaction is not implied by the Principle of Local Action alone. It requires both the Principle of Local Action AND a point particle ontology.

Finally, it has to be emphasised that retarded interaction is an important assumption in Lorentzian relativity but it does not appear at all in Einstein's derivation of Lorentz Transformations (or mine for that matter). It was put into Special Relativity by hand years after the Theory was published.

The result is that in all the EPR-Bell literature, including especially every derivation of a Bell Inequality, "local" ends up having two strictly unnecessary connotations. First, it means that any influence between two point events has to travel from one point to the other at or below the speed of light and second it means that all the ontological support for a pointlike event is colocated with the event.

When John Bell introduced the concept of a "Local Hidden Variables" (LHV) model, what he had in mind was to cover every possible kind of physical model where the physical reality of the particle is well-localised in the vicinity of the place where we find it. The result of the last forty years of experimental investigation is that all such physical models are excluded.

In other Physics communities, notably in both the Classical Field Theory and the Quantum Field Theory literature, there is a very different understanding of local action, as follows: Fields at different points in space do not interact with each other.

That is the definition of local action that we are using here. It is restricted to the level of Ontology and so it makes perfect sense. Adding the idea that no field may propagate faster than c, completes an appropriate definition of local realism for the Metaphysical presumption here, namely that physical reality comes in widely distributed wavefields that propagate at c. Our task will be to show that this ontological approach to local realism permits instant causal influences between point events at the level of Epistemology, which is to say that EPR correlations are permitted in Physics.

7.3 EPR Correlations as a Wave Phenomenon

7.3.1 An Analogy to Water Waves

The EPR paradox cannot be satisfactorily addressed with mere mathematics because we already have the mathematics that correctly predicts the observables. The question raised is "How can that be so?", but Mathematics cannot begin to answer that. Therefore, let us first discuss qualitatively, step by step, how to build up EPR correlations in a wave ontology: a model reality where all the physical "stuff" comes in widely distributed waves. The subsequent construction of an empirical model in the next Section will then be relatively straightforward.

This discussion begins with another trip to the beach. Imagine that Bob and Alice (the two famous observers from EPR experiments) have taken the day off and gone to the beach. We find them standing at the shoreline, several hundred metres apart, each one looking at his or her feet.

A wave comes in, and breaks at once all along the shore. Bob and Alice get their feet wet. They look at their watches and each one records the time. After a few dozen such experiences, they get together for coffee and a quick chat, and consider the data. It shows that they always both got their feet wet at the same time. They conclude that they have found nonlocal correlations, which indeed they have. These are not causal correlations[3], but we have taken the first step:

Waves are systems of distantly correlated movements.

The well known fact is that any kind of wave system involves instant correlations at a distance that are sustained by strictly local interactions. The space independent phase of the de Broglie wave model in Section 6.1 is an example of instant correlations at a distance, maintained over time by local interactions between neighbouring resonators. This feature will prove to be highly relevant to the discussion here.

After coffee, Bob and Alice take a walk along the coast, below the cliffs, where they find an unusual geological feature. Time and circumstance had carved a wide rectangular channel into the rock, with parallel sides and a sheer cliff face at the end. Bob and Alice take up their positions on opposite sides of the channel and each one observes the water below his or her feet.

Waves entering the channel propagate to the cliff face where they are reflected and propagate back out into the ocean. The channel isn't very deep. Occasionally, two waves come in from the ocean one after the other. The first, smaller wave propagates along the channel, hits the wall and is reflected. On its way out, it encounters the second, larger wave coming in. Because the water is shallow, the superposition of two such waves can be highly nonlinear. Instead of passing through each other, as two small waves would, they interact strongly - in a sense, the waves are both breaking at the same time. The wave system collapses, literally, so that all that is left is foam and white water. This collapse happens at the same moment everywhere across the entire span of the channel. Bob and Alice each see the same phenomenon at the same time: The clear water below each observer rises and then disintegrates into white water. It happens in different places at the same time.

The two waves in the channel above are an analogy, not to entanglement, but to the physical ontology corresponding to a measurement made by either observer. Entanglement between the two measured particles will be brought into the picture shortly.

In the analogy, the small wave is Alice's particle, A, and the larger wave is her measuring device, M_A. What this analogy already shows is that nothing prevents a measurement by Alice[4] from having ontological consequences in Bob's vicinity provided that a) both of her measured and measuring systems have ontological support in the region near Bob and b) both of these systems feature suitable correlations at a distance.

Such consequences still don't amount to causal correlations between events, but we have taken the second small step:

The distributed interaction between two distributed wavefields can have distributed consequences, that occur in different places at the same time.

[3] They have a "common cause". The term "causal correlations" refers to correlations that cannot be explained by common causes.

[4] Without loss of generality, it will be assumed throughout that Alice makes the first measurement.

This analogy is an example of an interaction between one-dimensional waves. Let us consider the corresponding scenario with three dimensional waves.

For a particle of matter like the electron to have a finite total energy, its field energy density must go to zero as the distance from the centre of the particle goes to infinity: $\rho_E \to 0$ as $r \to \infty$. To keep the total space integral of the energy density finite, the field energy density must go to zero at least as rapidly as $1/r^4$. Now, what we find in the world is that force field strengths vary as $1/r^2$ in the asymptotic region. Meanwhile, the field energy density varies as the square of the force field strength. Thus, the image of the particle in a wave model will be as a spherically symmetric field energy density distribution with a central peak and a far field where the energy density, $\rho_E \propto 1/r^4$. The central part of this distribution can be a broad plateau, as with de Broglie waves, or it can be sharp, when the particle has a well defined position.

We can envisage two such systems interacting with each other, like the waves in the channel above. A strong interaction occurs as the two systems align globally and the peaks of the two energy density distributions intersect. The far fields of Alice's measured particle and her measuring device provide the requisite ontological support in Bob's vicinity. As we saw in Section 6.1, the near and far fields are instantly correlated, both for the measured particle and the measuring device.

7.3.2 Entangled States

There are two categories of instantaneous EPR correlations at a distance in the literature. The first of these is exemplified by the EPR paper, which considered particles entangled in position-momentum. Regardless of the completeness or integrity of the quantum mechanical formalism, we know for sure today that such correlations do not in fact require any physical influence on Bob's particle, B, when Alice performs her measurement on particle A. The two particle wavefunction collapses instantaneously and there may be a physical influence involved, but the collapse can also be explained as a mere change in our state of knowledge. We know this because the correlations in this case do not violate any Bell Inequality. Therefore, an LHV model can be constructed. These correlations do not necessitate a "causal" explanation, in the sense that the intervention Alice makes on particle A causes a change in the physical state of affairs of particle B.

The second category is the case of instantaneous causal correlations at a distance, or "spacelike causal correlations", where "spacelike" means that the time interval between Alice and Bob's measurements is not sufficient to allow any luminal or subluminal signal emitted from Alice's location to reach Bob's location before his measurement has been done, and *vice versa*. These are the cases studied in the vast majority of EPR experiments. Rather than position-momentum entanglement, the two particles are either spin entangled (for matter particles like the electron) or helicity/polarisation entangled (for experiments with photons).

A similar argument applies in both cases, but let us focus on two spin entangled electrons, in the so called "singlet" state.

What we know about the two-electron system in the singlet state is that the total angular momentum vanishes. In the cellular resonator model above all the cells in a given electron have the same phase, which is continuously maintained by inter-cellular interactions between nearby resonators. However, the resonators must also be equipped with angular momentum, because otherwise the electron itself would not have angular momentum. We know that angular momenta align or counteralign with an applied magnetic field. In the case of an individual cell, this field is supplied by the other resonators in its vicinity. Finally, the Ising model [68] shows that an individual electron will settle into its lowest energy state, with all of its resonators aligned in one direction.

This applies to each electron separately, but when two electrons are in the entangled singlet state the total angular momentum of the system vanishes. Therefore, in the singlet state all the resonators in one electron system are counteraligned with all the resonators in the other.

These states don't just come and go. They have quite long lifetimes, measured in tens of milliseconds. The Compton wavelength is $\lambda_c = 2.4 \text{x} 10^{-12}$ metres so the period of its oscillations is $1/f = \lambda_c/c \sim 10^{-20}$ seconds, so a millisecond is an aeon in the timescale of an electron. Recall that with de Broglie waves, the continuing existence of the space independent phase in a noisy environment implied interactions between field resonators to maintain the phase locally and globally. Similarly, the observed persistence of entanglement implies that the overlapping, distributed fields of the entangled particles must be interacting with each other in order to maintain the anti-correlation of the entangled particles' angular momentum densities between preparation and measurement.

With both electrons constantly exposed to random interactions with a noisy environment, it would be unreasonable not to consider that the two wavefields are interacting with each other. The physical basis for such an interaction is straightforward: being counteraligned is a lower energy state than some random alignment, part way between aligned and counteraligned, so the entangled state has a binding energy.

The 2-particle system, having been prepared with angular momenta in one subsystem counteraligned with those in the other, persists in the same condition until a sufficiently energetic disturbance, for example a measurement interaction, overcomes the binding energy of the entangled state.

7.3.3 Bell Measurements

Let us now consider how the above ontological context for the singlet state relates to the observables in EPR experiments and the way they are measured.

Quite separate from entanglement, the basic proposal here is that when we measure the spin of a particle all of the measured particle, both near fields and far, interacts with all of the measurement device. The near field of the measured particle interacts with the near field of the measurement device and the far field of the measured particle interacts with the far field of the measurement device. There is no implication of any interaction at a distance, where the near field of the measuring device would interact with the far field of the measured particle or *vice versa*. That would violate local realism.

In Quantum Mechanics, the spin observable is closely related to a particle's angular momentum, but a spin measurement does not measure a particle's preexisting angular momentum. If we could measure it the result would be a vector, the magnitude of the angular momentum and a direction in space. Since the angular momentum is quantised, for all intents and purposes the result of measuring the angular momentum would be a direction in space, which is an analogue quantity. Unfortunately, we cannot perform such a measurement. Instead, the measurement device has to be set up to test the angular momentum along a particular space direction, and the result obtained is called the spin observable. It is a binary quantity: either "UP" or "DOWN" along the chosen measurement axis. Let us assume that the chosen axis is a. The possible results of a spin measurement along a are "UP ALONG a", which will be written as $< +_a >$, and "DOWN ALONG a", to be written as $< -_a >$.

One of the waves in the analogy above represents the measured particle, whose internal physical state is distributed across the entire channel (*i.e.* it occupies all of space, at least in principle). The other wave represents the measuring device. Having been set up to measure along a certain direction in space, its internal state includes the specification of the chosen measurement axis, which is therefore also distributed throughout the whole space (at least in principle). When the measurement interaction occurs, it occurs in parallel, throughout all of space at the same moment.

Spin measurements are distributed. They happen in parallel throughout all of space.

Spin measurements do not tell us about the spin of the particle before the measurement. Instead, the internal state is "projected" onto the chosen axis and the measurement result obtained is the post-measurement spin state of the particle, which corresponds to the post-measurement physical state of the particle's field angular momentum *density*, which is also distributed throughout all of space (at least in principle).

These measurements are carried out with magnetic fields, and the quantum mechanical projection of the spin state onto the chosen axis corresponds to a rotation of the measured particle's field angular momentum density until it aligns (spin UP) or counteraligns (spin DOWN) with the externally applied magnetic field.

With that ontology-epistemology relationship in mind, let us now turn to the measurements in a typical EPR-Bohm experiment involving spin measurements on two electrons in the singlet state (maximal spin entanglement). Assume that Alice measures her electron, A, along axis a. Let us also assume (for the time being) that she chose her measurement axis, a, in the distant past[5]. When she measures the spin, the physical angular momentum density of electron A is instantaneously modified across all space, including the space in Bob's vicinity. Immediately following the measurement her field angular momentum density is aligned either parallel or anti-parallel to her measurement axis a.

In the cellular resonator model, all the cells become either aligned or counteraligned with the measurement field.

There are a couple of issues still to resolve with this measurement interactions scenario, which will be addressed in the next Section, but let us now bring the entanglement back and bring this qualitative discussion to its conclusion.

Prior to Alice's measurement, electron A was spin entangled with electron B. As discussed above, all the resonator cells of Alice's electron were counteraligned with all the cells of Bob's electron, and there is an interaction between the two field systems that maintains the *status quo*.

This interaction is, of course, also a distributed interaction. Consequently,

[5]Delayed choice experiments will be addressed in due course.

when Alice's measurement causes electron A's angular momentum density to rotate towards the a axis, the distributed interaction between the A and B particles causes B's angular momentum density to rotate in the opposite direction.

After Alice's measurement (but before Bob's), if A ends up in the state $< +_a >$, which will be written as $< A = +_a >$ (NB: A is a measurement result), we can infer that B must be in the state $< -_a >$ so let us write $< \lambda_B = -_a >$ (NB: λ_B is an internal state, which is to say that it is a hidden variable in the Bell vernacular). Conversely, whenever Alice's measurement result is $< A = -_a >$, Bob's electron ends up in the internal state $< \lambda_B = +_a >$.

Of course, this is precisely what Quantum Mechanics predicts and it is experimentally verifiable if we now measure Bob's electron along the same axis, a.

However, in EPR experiments we do not typically measure Bob's electron along axis a. We actually measure it along some other axis, b, where $b \neq a$.

When Alice performs her measurement along axis a, the results that she obtains are for a particle in an unknown spin state, so she obtains the measurement outcomes $< A = +_a >$ in 50% of trials and $< A = -_a >$ in the other 50% of trials. Bob, on the other hand, is measuring a particle that has been prepared in a well defined spin state. The quantum mechanical prediction for measuring a particle prepared in the $< +_a >$ state along the b axis is as follows:

$$P(+_b|\theta) = cos^2\left(\frac{\theta}{2}\right) = \frac{1 + cos\,\theta}{2}\,, \tag{7.1}$$

where $\theta = a - b$ is the angle between the preparation and measurement axes, a and b, There is no hint of nonlocality or anything mysterious about this relationship (7.1) and there is no shortage of arbitrary looking LHV toy models for it in the literature, beginning with Bell's own illustration [63].

Of course, Bob's measurement is spacelike separated from Alice's and he doesn't know in which of his trials his particle was prepared as $< +_a >$ and which as $< -_a >$. In fact, he does not know along which axis Alice chose to measure. The way the experiment works is to bring the data from each trial to a central station and calculate the relevant statistics as follows.

Bob and Alice each report which of two alternative axes they chose to measure along on each trial, a or a' for Alice, b or b' for Bob, and the result they obtained. Alice's possible records are thus $A = +_a$, $A = -_a$, $A = +_{a'}$, or $A = -_{a'}$ and Bob's are $B = +_b$, $B = -_b$, $B = +_{b'}$, or $B = -_{b'}$. The full ensemble of trials is then divided into four subensembles according to Alice's and Bob's chosen axes. One subensemble contains all the trials where Alice measured along a and Bob measured along b. Another contains all the trials where Alice measured along a' and Bob measured along b' and so on.

For each subensemble, a correlation function is then calculated. For example, for the subensemble where Alice chose a and Bob chose b, the correlation function is defined as:

$$C(a, b) \equiv \frac{P_{++} + P_{--} - P_{+-} - P_{-+}}{P_{++} + P_{--} + P_{+-} + P_{-+}} \tag{7.2}$$

where P_{++} is the probability of finding both $A =< +_a >$ and $B =< +_b >$ in a given trial and so on. These probabilities can be calculated using (7.1) above, or they can be determined directly from the datasets. It is easy to verify that the correlation function predicted by Quantum Mechanics' equation (7.1) is:

$$C(a,b) = -\hat{\mathbf{a}} \cdot \hat{\mathbf{b}} \tag{7.3}$$

Bell and others derived various inequalities that LHV models would have to obey involving the correlation functions for different measurement axis settings. Collectively, such relationships are known as the Bell Inequalities. The most common of these is the CHSH Inequality:

$$|C(a,b) + C(a,b') + C(a',b) - C(a',b')| \leq 2$$

To understand the meaning of this relationship, recall the Bohr-Einstein debate. Both protagonists had one assumption in common, namely that Alice's measurement could not influence the state of Bob's electron. The essence of the matter is that if the "no influence" assumption is correct then the above inequality holds for any choices of Alice's and Bob's measurement axes.

Quantum Mechanics, however, says that the inequality does not always hold. Consider, for example, the following specific measurement axis choices: $a = 0$, $a' = \pi/2$, $b = \pi/4$ and $b' = -\pi/4$. The first three terms in the CHSH inequality are all equal to $-1/\sqrt{2}$, but the last one, which has to be subtracted from the others, equals $+1/\sqrt{2}$ so the total result on the left hand side is $|-2\sqrt{2}| = 2\sqrt{2} \simeq 2.82$ which is greater than 2. For appropriate choices of the angles between the measurement axes, the quantum mechanical correlation functions calculated from (7.3) can and do violate the inequality and the entire body of experimental work conducted over the last 40 years confirms the quantum mechanical predictions.

Millions of pages have been written about the subtleties and loopholes involved in Bell Inequalities, ranging from conspiracies between measuring devices and the particles that they are going to measure, to the question of Bob and Alice's free will in making their choices. We'll address these issues later, but the qualitative discussion above already points to the fact that it is entirely possible for Alice's measurement to influence Bob's electron in a wave ontology. Let us now construct a quantitative model that prepares Bob's electron in a known state using only strictly local interactions that are consistent with the above definition of local realism *i.e.* the combination of "Fields at different places do not interact" and "no superluminal movements".

7.4 A Physical Model of EPR Correlations

The following physical model for the quantum predictions in EPR experiments [69] constitutes a proof that instant influences between point events are possible in local realism. Indeed, such phenomena are only to be expected in luminal wave systems, where the interactions that occur are distributed throughout all of space and not just at the places where we observe pointlike events[6].

The model also highlights a couple of lacunae in the logic of the interacting water waves analogy when applied to real world EPR experiments. These issues

[6] A related model will be developed in Chapter 9 for the Classical Electromagnetic interaction to drive home the key point that there is nothing inherently problematic about instant interaction at a distance in point models of distributed wave ontologies. The present context of spin measurements avoids certain ancillary phenomena (brehmstrahlung radiation etc) that arise in the case of nonuniform, distributed exchanges of linear momentum between fields, so it provides a good introduction to the basic concept of distributed interaction.

will be identified and resolved by giving toy model[7] solutions to establish the fact that there is no problem for local realism *per se*.

The usual convention will be followed, where the experimenters in each arm of the apparatus are named Alice and Bob; their respective measurement outcomes are A and B; their electrons are electron A and electron B; and their respective measurement axes are labelled a and b. Spin measurements on different particles always commute, so it will be assumed without loss of generality that Alice makes the first measurement.

In order to focus on the logical structure, the model begins by replacing the 3 dimensional cellular resonator field microstructure developed in Section 6.1 with a one dimensional system of cellular automata [70, 71] that spans the space between Alice and Bob[8]. Each of the four subsystems (the 2 entangled particles and the 2 measuring devices) is represented by one of the rows in Figure 7.1.

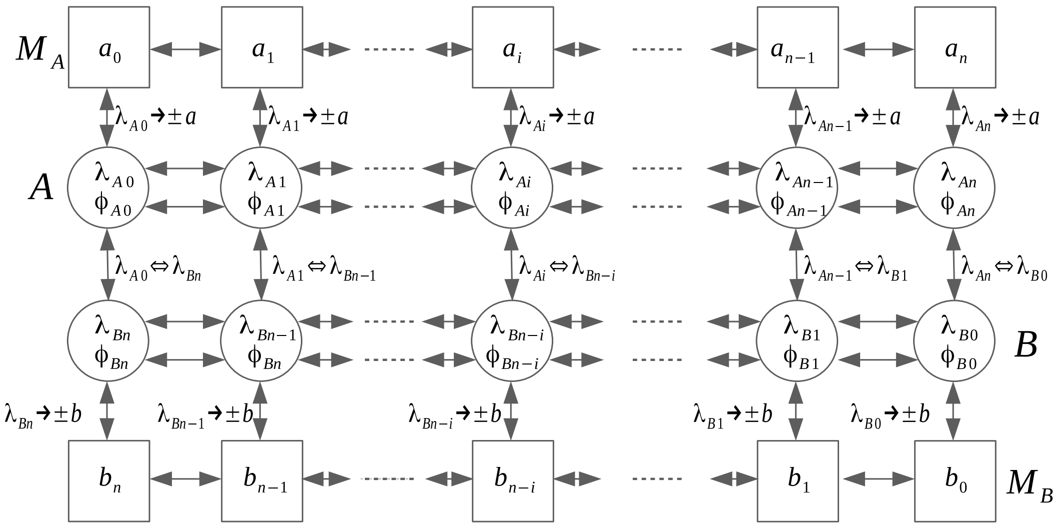

Figure 7.1: *1D cellular automata schematic of a Bell measurement. The measurement devices are M_A and M_B (first and last rows). The electrons are A and B (middle rows). Circles represent electron cells, squares represent measurement device cells. Horizontal arrows show the nearest neighbour interactions within each subsystem that regulate its distributed variables. Vertical arrows between circles are the entanglement bonds, ensuring that, at the first measurement (Alice's), $\lambda_A = -\lambda_B$. Measurement interactions (vertical arrows between squares and circles) are synchronised by the range variable in each cell, represented here by the subscripts, i for Alice's electron and $j = n-i$ for Bob's electron. At Alice's measurement, $\lambda_{Ai} \rightarrow \pm a$ for all i, $\lambda_{Bj} \rightarrow \mp a$ for all j, and then the entanglement bonds are broken. At Bob's measurement, $\lambda_{Bj} \rightarrow \pm b$ for all j.*

The Principle of Local Action is implemented as follows. Within each sub-

[7]In the EPR literature, a toy model is some typically artificial and over-simplified device introduced to establish a logical possibility. This Section is a proof that spacelike causal correlations are available in principle in local realism. It is not a Physics theory and there is no suggestion that physical reality would actually work like the model here.

[8]The cell contents are continuous variables here. Note that the cell lattice, constructed from energy that propagates at c, is inherently Lorentz covariant.

system, which is to say within each row, nearest neighbour interactions between adjacent cells are represented by the horizontal doubleheaded arrows. The interacting entities are near to each other, but not exactly colocated. Consequently, these interactions are limited by the speed of light. Since Bob's measurement is spacelike separated from Alice's, there is no possibility for such interactions to communicate any influence from Alice's location to Bob's in the time between the measurements[9].

Alice's measurement result depends entirely on the ontology in her vicinity. To be more specific, it depends only on the local parameters for her electron and her measurement device, which are shown as a_0, λ_{a0} and ϕ_{a0} in the schematic diagram. Similarly, when Bob performs his measurement just after Alice's, his result depends only on the parameters b_0, λ_{b0} and ϕ_{b0}. Note that the space labels run in opposite directions, so that the cell containing λ_{b0} is at Bob's location, while the cell containing λ_{a0} is at Alice's and so on.

Interactions between the subsystems, *i.e.* the measurement interactions and the entanglement between the particles, are represented by doubleheaded vertical arrows between colocated cells in adjacent rows.

As for the measurements, when Alice performs her measurement, each of the cells of her measurement device, the M_{Ai}, interacts with the corresponding cell of her electron. Similarly, when Bob performs his measurement, each of the cells of his measurement device, the M_{Bj}, interacts with the corresponding cell of his electron. There is never any interaction between any part of Alice's measurement device and any part of either Bob's measurement device or his electron, regardless of colocation and similarly there is never any interaction between any part of Bob's measurement device and any part of either Alice's measurement device or her electron. These last two conditions remove any possibility of pre-measurement contextuality, in the sense that the physical states of the particles before Alice's measurement do not depend on the measurement context.

There is, however, a further interaction shown between colocated cells of the entangled particles, which constantly communicates the hidden variable λ between the entangled particles throughout the time between preparation and Alice's measurement. These "entanglement bonds" are essentially what explains the spacelike causal correlations in the model. Their existence is implied by the extremely long lifetime of entanglement in a noisy environment. It is justified in Physics by the low energy of counteraligned angular momentum densities.

Finally, and most importantly, unlike previous attempts to explain EPR correlations, this feature is testable, as discussed in Section 7.8.

There are three main steps to discuss: distributed variables; entanglement as interaction; and distributed Bell measurements.

7.4.1 Distributed Variables

The space independent phase in Section 6.1 is the archetypal distributed variable. Given that the angular momentum of each resonator is proportional to its rate of change of phase, the field angular momentum (density) is also distributed, as shown in Figure 7.1 by the distributed variable, λ, that represents the internal state associated with the electron's angular momentum. It will be

[9] As seen by any Lorentz observer.

sufficient to let λ be an angle in the range $[-\pi, \pi]$, such that λ_{Ai} is the value recorded in the i^{th} cell of Alice's electron. Similarly for Bob's. A second distributed variable, ϕ, will be introduced in connection with the measurement model below.

Two manifestly local realist proposals regarding the evolution of distributed variables between preparation and measurement are asserted. The first of these is that distributed variables are regulated by nearest neighbour interactions, corresponding to the retarded self-interactions of Section 6.1.

This provides noise immunity in case individual cells are perturbed by interactions with the environment. For example, the following local interaction model implements a proportional control mechanism that smooths out perturbations across space[10]:

$$\frac{d\lambda_{Ai}}{dt} = \alpha((\lambda_{Ai+1}(t_r) - \lambda_{Ai}(t)) + (\lambda_{Ai-1}(t_r) - \lambda_{Ai}(t)))\,. \tag{7.4}$$

Where $t_r = t - r_s/c$ is the retarded time, r_s is the cell spacing and c is the characteristic velocity. Similarly for the λ_{Bi}. Of course, the speed of any such algorithm in smoothing out perturbations across space is inevitably limited by the speed of light, far slower than would be required for perturbations near Alice to influence affairs near Bob within the time window of a Bell measurement [72].

In an imperfect real world, such a control mechanism might also have a role in measurement interactions, as discussed below. However, the measurement model, (7.8), will be assumed to work perfectly, creating the same measurement outcome in every cell, which will render this possibility moot.

Thus, all that is being asserted here is the reasonable notion, already well known in field theory, that correlations at a distance, existing within each electron at preparation, will persist until measurement.

7.4.2 Entanglement Bonds

While every electron has its own distributed variables, as discussed above, what distinguishes the singlet state is that, at preparation, these variables are common to both members of the entangled pair.

The second local realist proposal is that this *status quo* is also maintained between preparation and the first measurement by a further interaction that operates between colocated cells of Alice and Bob's electrons, see Figure 7.1. These "entanglement bonds" implement a second control mechanism, like (7.4) but operating between the entangled particles to ensure that the internal angular momentum state, λ, is equal but opposite: $\lambda_A = -\lambda_B$. This results in a 2-particle system with zero angular momentum, where the entanglement displays noise immunity.

Again, there is no question of any violation of locality involved.

[10]Or, for a discrete time version: $\lambda_{Ai}(t_{n+1}) = \lambda_{Ai}(t_n) + \beta\Big((\lambda_{Ai+1}(t_n) - \lambda_{Ai}(t_n)) + (\lambda_{Ai-1}(t_n) - \lambda_{Ai}(t_n))\Big)$, where the timestep, $(t_{n+1} - t_n)$, is of the order r_s/c. It is straightforward to verify the stability of this for small β, but the details of the algorithm will be of no consequence in what follows.

7.4.3 A Distributed Measurement Model

In the de Broglie-Bohm EPR experiment, as soon as Alice has measured the spin component of her electron, A, Quantum Mechanics makes a prediction with certainty about the result upon measuring the same spin component of Bob's electron, B. Using the notation from the preceding Section, it follows that if Alice's result is $< A = +_a >$, *i.e.* "UP along a", then as far as Quantum Mechanics is concerned *after* her measurement B is in the pure state $< \lambda_B = -_a >$, for which $\lambda_B = a - \pi$. Similarly, if her result is $< A = -_a >$, then $\lambda_B = a$. This is what needs to be shown in a physical model[11].

The measurement model here deals with the general case, where the B electron spin is actually measured along $\hat{\mathbf{b}}$, where in general $\hat{\mathbf{b}} \neq \hat{\mathbf{a}}$. However, there is no doubt that this counterfactual is valid[12] as this is an experimentally verified prediction of Quantum Mechanics concerning spin observables of different particles (which always commute).

If we can specify a local interaction measurement model, using the system of interactions depicted in Figure 7.1, such that $< A = +_a > \Rightarrow \lambda_B = -_a$ and $< A = -_a > \Rightarrow \lambda_B = +_a$, then Bob's measurement becomes a case of a spin measurement on a single electron prepared in a known state. The usual quantum probability for this, given in the previous Section, can be written slightly differently as:

$$P(+_b|\theta_B) = cos^2\left(\frac{\theta_B}{2}\right) = \frac{1 + cos\,\theta_B}{2}\,, \tag{7.5}$$

where $\theta_B = \lambda_B - b$. As already mentioned, there is nothing mysterious or nonlocal about this relationship. However, in addition to explaining the causal influence $< A = +_a > \Rightarrow \lambda_B = -_a$ we also need to specify a *distributed* model of this relationship. This isn't so important for Bob's measurement, but it is for Alice's: Alice's possible results are actually equiprobable, $P(A = +_a) = P(A = -_a) = 0.5$, but we need ALL of Alice's electron's cells to respond to the measurement in parallel (at the same time). Then, and only then, can we take advantage of the specific causal influence in Bob's vicinity, $< A_n = +_a > \Rightarrow \lambda_{B0} = -_a$, where A_n is the result in the n^{th} cell of Alice's electron - the one that is colocated with Bob. This is the only influence that is causally related to Bob's measurement result. Having prepared that particular cell of Bob's electron, the part of Bob's measurement that causes his measurement result is written as:

$$P(+_b|\theta_{B0}) = cos^2\left(\frac{\theta_{B0}}{2}\right) = \frac{1 + cos\,\theta_{B0}}{2}\,, \tag{7.6}$$

where $\theta_{B0} = \lambda_{B0} - b = \mp a - b$. The correlation function for the model can then calculated as:

$$C(a,b) = \frac{P_{++} + P_{--} - P_{+-} - P_{-+}}{P_{++} + P_{--} + P_{+-} + P_{-+}} = -\hat{\mathbf{a}} \cdot \hat{\mathbf{b}}\,, \tag{7.7}$$

where P_{++} is the probability of finding both $A =< +_a >$ and $B =< +_b >$ in a given trial and so on. The physical model then violates all the usual Bell Inequalities.

[11] Note that the question of what state either electron was in *before* Alice's measurement is irrelevant.

[12] The "reality fails" argument will be discredited in due course.

The entire mystery is how distributed action enables Alice's measurement to prepare cell 0 of Bob's electron in a known state. Given the proposed entanglement bonds, this may seem straightforward enough at first glance: At Alice's measurement, λ_A rotates onto either $\hat{\mathbf{a}}$ or $-\hat{\mathbf{a}}$. λ_B, being coupled to λ_A, rotates onto $-\hat{\mathbf{a}}$ or $\hat{\mathbf{a}}$ respectively. Once this has occurred, and Alice's measurement outcome has been determined, the entanglement bonds are broken because the interaction energy exceeds the binding energy of the entangled state.

While that covers Bob's electron's cells in Alice's vicinity, according to the Principle of Local Action any impact on those cells has no immediate effect on Bob's electron's cells in his own vicinity. To make this work, the measurement interactions have to be distributed. Since the measurement field is a magnetic field related to angular momenta, and each measurement device is constructed from quantum systems that have their own sets of distributed variables, there is nothing new in requiring that the measurement axes, a and b, are also distributed variables, as shown in Figure 7.1. But this is still not enough. There are two further issues involved.

First, not only must Alice's measurement be distributed, but also it has to be synchronised so that all her measurement interactions happen at the same instant, t_{mA}, in Bob's vicinity as well as in her own. This task of locally detecting the global alignment of Alice's electron's field in relation to Alice's measurement field can be accomplished in various ways without resorting to any nonlocality.

The toy model method used here is to add a new variable to each cell that encodes its range to the center of the subsystem to which it belongs[13]. Alice's measurement interaction, a physical model of the random process (7.5), is then implemented in both Bob and Alice's spacelike separated vicinities at the moment when corresponding range variables match, which is to say when Alice's measuring and measured systems become colocated. This is represented in Figure 7.1 by using subscripts that increase in the direction away from the relevant experimenter: Alice's measurement device and electron subscripts start at 0 at A and increase towards n at B and vice versa for Bob's. This implements a "distributed collapse" upon measurement.

Second, if one simply applied the *random* process, (7.1), in every cell, the result would be that some A cells rotate toward the $< +_a >$ state while others go toward the $< -_a >$ state. It might be thought that the control algorithm, (7.4), would *eventually* sort things out, but what if there were more cells in the $< +_a >$ state in Alice's vicinity and more in the $< -_a >$ state in Bob's vicinity?

This is actually not the problem. Recalling that the cell dimension, r_s, is on the order of the reduced Compton wavelength, a region of dimension 1 nanolightsecond is 35 orders of magnitude larger than the cell size, and the cells in subsection 6.1 overlap, so there is a vast number of cells in any small space region. Unless θ in (7.1) is exactly equal to $\pi/2$ (which has probability 0), the control algorithm, (7.4), operating in parallel with the measurement interaction, might rapidly ensure the same decision independently in every space region, on a timescale many orders of magnitude faster than the light time from Alice to

[13] This local realist protocol is unlikely, but it is logically speaking possible: While $E = \hbar\omega$ means the energy of the whole is quantised and all resonator frequencies in Section 6.1 are identical, their amplitudes and their number density will be range dependent. Each cell can then use locally available information to identify the direction of the energy density gradient, the range variable in the adjacent cell, and the distance to it.

Bob.

The somewhat more subtle problem is that such a mechanism is unacceptable. If $P_{cell}(+) > 0.5$ in the *random* process (7.1), the control mechanism, (7.4), that regulates distributed variables within an electron immediately drives the whole measurement result into the $< + >$ state with probability 1. That's inconsistent with the quantum predictions.

In order to model the quantum predictions correctly, $P(+|\theta)$ in (7.1) must be interpreted as the probability that a majority of cells produce the $< + >$ result, and the simplest way to do that is with a model that produces the same result in every cell. Any role for (7.4) in the measurement process is then moot, so there is no need for any analysis here of the control algorithm.

As suggested by Jaynes [65], hidden variables may also obey equations of motion, and the probabilistic outcome of (7.1) can be implemented by a *deterministic*, but time dependent, local process that produces the same result in every cell. If the system were measured at a given instant, there would be a definite measurement result, the same in every cell, but the value of the result depends on the moment in time when the system is measured.

A simple toy model for the measurement interaction shows this to be entirely feasible. Let the measurement outcome in the i^{th} cell be:

$$A_i = SGN[sin\,(\omega_{\sigma i} t_{mA}) - f(\theta_{Ai})]\,, \tag{7.8}$$

where $\theta_{Ai} = \lambda_{Ai} - a_i = \lambda_A - a = \theta_A$ (this follows from the control algorithm having maintained the distributed variables λ_A and a constant across the system prior to Alice's measurement); t_{mA} is the measurement time, the instant when range labels align in the measured and measuring systems; ω_σ is a frequency associated with the time evolution of the spin, common to all cells with the space independent phase, $\phi_A = \omega_\sigma t$, which is the remaining distributed variable shown in Figure 7.1.

This measurement interaction model, (7.8), provides the same result in every cell and it is easy to show that it also implies the statistical distribution of results, $P(+|\theta_{Ai}) = (\pi - 2\,sin^{-1} f(\theta_{Ai}))/2\pi$. Substituting (7.1), the choice $f(\theta_{Ai}) = sin(-\frac{\pi}{2} cos\,\theta_{Ai})$ yields a measurement outcome with the appropriate statistics in accordance with the usual quantum mechanical result, (7.1).

Let us briefly review the complete process for a Bell measurement in the model.

Since the initial state, λ, is unknown, in Alice's measurement, (7.8) combines with the entanglement bonds to project the two particle entangled state onto one or other of the states $< A = +_a, \lambda_B = -_a >$ or $< A = -_a, \lambda_B = +_a >$ with equal probablility. In Bob's measurement, (7.8) then projects the B electron state onto either $< B = +_b >$ or $< B = -_b >$ in accordance with (7.6), where $\theta_{B0} = a - \pi - b$ or $\theta_{B0} = a - b$ respectively. A straightforward calculation using (7.2) shows that this replicates the complete set of quantum predictions for any choice of the measurement axis settings, a and b.

We have been assuming all along that the measurement axis choices, a and b, were made in the distant past. As far as the Bell literature is concerned, this implies the possibility of causal influences from Alice's choice to (say) the physical state of Bob's electron. The design of delayed choice experiments has made an essential contribution by eliminating this possibility. From the present perspective, which is not an LHV model, delaying the choices of measurement axes until just before the measurements raises no new issue because implementing any such choice is itself an inevitably distributed process that happens in parallel across all the space at the moment when the switch is thrown (which is, in its turn, yet another physically distributed process).

The Bell Inequalities have now been violated without invoking any form of action at a distance. This proof is a physical instance of Redhead's old result: *Ontological locality does not of necessity imply Epistemological locality.* [73].

The next part of this discussion focusses on an experiment with entangled electron spins separated by over a kilometre. That's 4×10^{14} times the Compton radius. Far fields are attenuated in the order of 10^{-30} and energy densities on the order of 10^{-60} relative to the centre of a field distribution for a particle of mass 10^{-30} Kg. Experiments even provide evidence for entanglement between a particle on board a satellite and one on the ground, which is to say at a range of hundreds of kilometres.

These are large exponents, but very large and very small numbers are the norm in Physics. We detect light from sources billions of light-years away. When an individual atom a billion light years away emits photons, there is a finite probability for each of them to reach us here on Earth. The Electric force between two charged particles is 10^{43} times the strength of gravity. That's a large number too, but gravity rules the Cosmos. The fact is that the far fields matter at any distance.

The far fields here are obviously quite weak, but the question is "Weak relative to what?" How strong does the coupling between entangled particles need to be to maintain the state over long periods of time in the presence of environmental noise? We are in no position, theoretically speaking, to assess the efficacy of *coherent* quantum entanglement bonds as against *incoherent* Electromagnetic noise. Moreover, if the Bell measurements were far apart, but not spacelike separated, it would be taken as given that the far field, no matter how weak, mediates the influence. The question whether interactions between entangled particles are strong enough would not even be raised.

It is well known that the noise is an important consideration. For example, the Delft experiment in the next Section is conducted with both electrons held at a temperature near absolute zero in order to minimise noise and obtain those very long entanglement lifetimes on the order of 10 milliseconds.

The good news is that the above model's reliance on weak far fields to mediate the interaction between members of an entangled pair makes the model testable, as discussed in Section 7.8. This distinguishes it from every other proposed resolution to the EPR crisis that has not already been excluded by performing appropriate experiments.

7.5 The Delft Experiment

Physics has gone to extreme lengths to close down even the slightest possibility of a non-causal explanation for the observations.

For example, in one recent experiment [74], light from two different quasars, millions of light years away from Earth and from each other was used to generate random numbers that select Alice's and Bob's measurement axes after the soon to be measured photons were already in flight. This is done to exclude any logical possibility of a conspiracy between the measured particles, the local environment and the selection of measurement axes such that the "random" settings end up with a nonrandom dependence on the physical state of affairs pertaining to the entangled system.

Even this is criticized: Most of the light from the quasars is absorbed in transit, and it could be logically possible that it is absorbed in just such a way that what remains is correlated with the entangled system of photons already in flight. This kind of argumentation might seem to go well beyond anything that could be considered reasonable. The fact is that spacelike correlations have been established beyond a reasonable doubt, but this fact has been considered so radically unpalatable that even the most extreme of unreasonable doubts is countenanced.

There used to be other doubts, so called "loopholes" in the experimental evidence that were not at all unreasonable. The two most important loopholes were the detection loophole and the locality loophole.

Way back in the 1990's, there were two main classes of experiment, one with the polarisation observables of photons, the other with the spin observables of massive particles, usually electrons. In both cases, particle pairs were prepared at a common origin with a vanishing total angular momentum and became well separated as they travelled towards the experimenters, Alice and Bob, who carried out the measurements.

With the photon experiments, the photons could be separated to a sufficiently large distance that we could be sure that the measurements were spacelike separated: For any observer, there was no chance for any light signal to reach from one arm of the apparatus to the other in the time between the measurements. These experiments closed the "locality" loophole.

Unfortunately, only a tiny fraction of the particle pairs generated was ever detected. Many would be going in the wrong direction, and never be coupled into the fibre optic cables that would route them to Alice and Bob and many more were absorbed en route.

These experiments opened a sampling loophole: the sample size, or subensemble, on which the experimental conclusion is based is far smaller than the original full ensemble of pairs generated. There might be a sampling bias. The particular form of sampling bias of most concern is known as the detection loophole. The idea is that for any particle to reach a given detector, it would have had to "make a decision" at the polarising beam splitter that divided the incoming photons into those that are Horizontally polarised and those that are Vertically polarised. Given a random distribution of polarisation for the incoming photons, some of them would be close to the 45 degree decision threshold between horizontal and vertical polarisations when they impinge upon the beam splitter.

The detection loophole is the very real possibility of a dead zone in the apparatus such that photons whose internal state is very close to the threshold are more likely to be absorbed than those for which the decision was clear cut.

With such a dead zone, the detection probabilities depend on the internal state of measured particles and it was straightforward to construct models that replicate the quantum predictions and violate Bell Inequalities.

On the other hand, in the experiments with electrons, almost 100% of the generated particle pairs could be detected, but they could only be separated by short distances and it could not be guaranteed that a signal had not gone from one to the other. Closing the detection loophole had opened the locality loophole.

Many physicists considered that closing a loophole in one experiment closed it, period. Both loopholes had been closed in different experiments so the case for nonlocality in nature was made. But the standard of proof is high because the stakes are enormous, and many others did not agree.

In particular, I recall exchanging correspondence with Ian Percival (originator of the double Bell paradox) in 2010, or thereabouts. When I asked the professor how he thought his paradox should be resolved, he pointed me to an as yet unpublished article of his arguing the case for a new Law of nature that precludes closing all the loopholes in a single experiment. In due course, I was able to read that article when it came out. A couple of years later, the detection and locality loopholes were indeed closed in a single experiment, the Delft experiment [72].

The Delft experiment involves two pieces of diamond that are in different laboratories, 1.3 Km apart, and held at cryogenic temperature close to absolute zero. In each diamond, one of the carbon atoms is replaced by a nitrogen atom (which has an extra electron). The spins and orbits of these Nitrogen Valence (NV) electrons can be manipulated and measured in various ways using microwave pulses.

First, each member of the pair is prepared into an unknown spin state with the result that they may have either equal spins or opposite spins, with an equal probability in each trial. They are then subjected to an "entanglement swapping" protocol (which will be discussed in Section 7.7) the result of which is to identify specific trials in which the particle pair is entangled. Spacelike separated spin measurements are then conducted on 100% of these pre-identified entangled electron pairs and the quantum predictions are tested. This is also a "delayed choice" experiment, where the selection of Alice's and Bob's measurement axes is also spacelike separated.

In practice, the delayed choice spin measurements are actually carried out on all trials, and the results for all those trials where the pair had not been pre-identified as being entangled are deleted. This is not a concern because the event ready signal, which decides which specific trials to include, is spacelike separated from the choices of measurement axes. It is also either spacelike separated or before the actual Bell measurements are performed. The result of a trial that has already been pre-excluded from the experiment before it takes place is of no consequence.

This experiment is not unique, and there are now several other experiments where all the loopholes have been closed in the same experiment. It is safe to say that the orthodox conclusion in relation to the entire body of experimental evidence is finally secure: Spacelike causal correlations are real.

7.6 Discrediting the "Reality Fails" Argument

Are there still physicists who deny this reality? Yes, of course there are, and they come in two main categories: those who question the evidence and those who question the assumptions involved in the Bell Inequalities and/or their meaning.

I mention these two categories in the interests of completeness, but to discuss either of them further here would be to venture, along with Alice, down a rabbithole and into Wonderland. For example, interested readers may wish to develop for themselves a clear understanding of "non-traditional" probability distributions [65].

In deference to this line of thinking, however, the orthodox conclusion in relation to spacelike causal correlations is usually stated this way: The combination of "reality" and "locality" must fail. Recall that "reality" refers to the objective existence of unmeasured properties and "locality" refers to the idea that the only way for a measurement "at" point A to causally influence the result of a measurement "at" point B is by sending a thing from A to B, at or below the characteristic velocity c.

This conclusion is, in my view, valid (under quantum mechanics' epistemic definition of "local realism", locality fails), but let us analyze it a little further. We actually do know about the objective existence of (the relevant) unmeasured properties, because we can predict them with certainty, because we have verified those predictions experimentally, and because they are predictions in relation to commuting observables[14].

In particular, immediately AFTER Alice measures her particle, A, along axis a, and obtains the result $< UP >$ along a, we know with certainty that Bob's particle B is $< DOWN >$ along the same axis a. We can, and do, separately verify that part of the theory experimentally. Therefore we do not need actually to perform the measurement along a to know that it is true.

In an EPR experiment, one measures particle B along a different axis, $b \neq a$ and the experimental results show that, whenever Alice's result is $< UP >$ along a the result of this measurement by Bob along b is statistically consistent with a particle that had been prepared in the $< DOWN >$ along a state.

Any discussion about the failure of the quantum mechanical definition of "reality" is ultimately a discussion about the validity of the minor counterfactual element in the above: We may know for sure that a certain value of a property (in this case the spin of particle B along a) will be displayed if measured, but we don't actually measure it, we measure a different property, namely the spin along b.

The "reality fails" argument is also commonly referred to as the failure of "counterfactual definiteness" [66, 75]. Essentially, as far as spacelike causal correlations are concerned it amounts to the proposition that, even when we know what a property is, it does not "exist" unless it is observed. No one has ever understood what this means (observed facts being epistemological, "exists" being ontological). The notion that until an observed fact is actually crystallised we are not entitled to draw inferences regarding its ontological support flaunts the basis for every prediction in science.

The appeal of the "reality fails" argument lies in the twin facts: it is unintelligible, and it can never be disproved or even tested. Since it can never be tested, "reality fails" is an unscientific point of view. That's why it should be discredited in the community: physicists need to stop appealing to untestable arguments in order to reject experimental findings.

In a very real sense nothing has changed here since the Bohr-Einstein debate and Podolsky's prescient reality criterion. Bohr's argument prevailed back then because it could not be understood. The corollary today is a case of choosing the lesser of two evils: Some physicists are less uncomfortable with something that

[14] Spin measurements on *different* particles always commute.

is unscientific, untestable and unintelligible, "reality fails", than they are with something that is clearly understood, "locality fails", but which they presume to be impossible ostensibly *"because of Special Relativity"*.

However, we have already seen that the failure of the EPR community's definition of "locality" doesn't imply any superluminal interaction at a distance and that neither Special Relativity nor the Principle of Local Action is falsified in EPR experiments. What has been falsified are two closely related, long standing Metaphysical prejudices:

Prejudice 1. When two golf balls interact with each other at a distance, the interaction is a physical thing that travels from one golf ball to the other, at or below c.

Prejudice 2. All of the physical reality corresponding to the image of a golf ball that we see is inside the image.

Note that if the second prejudice were true, the first would also have to be true. But nowadays, we know for sure (eg from quantum fields and the existence of de Broglie waves) that the second one is false. Nonetheless, the first continues to be universally and uncritically accepted, *"because of Special Relativity"* and despite the fact that Einstein's Theory has nothing to do with it, as discussed in Chapter 8.

7.7 Barrett-Kok Entanglement Swapping

An important difference between the Delft experiment and the experiments of the 1990s is that the two electrons that form the entangled pair in the Delft experiment are always well separated. They do not begin at a common origin and proceed to Alice and Bob's measurement locations.

The entangled state is induced by an "entanglement swapping protocol", the Barrett-Kok protocol [76]. According to the usual understanding, the spin-entangled electron pair never interact directly with each other. Instead, they become entangled as the result of a direct interaction between a "pair" of photons with which each electron had interacted in the past. The quotation marks around "pair" are necessary because, in all cases where the Barrett-Kok protocol succeeds, the pair has only one member.

This might seem to be an arcane topic to cover in a book about the foundations, but it is pivotal to understanding the Delft experiment, it is interesting in its own right and it provides a clear cut instance of a wavefunction collapse that can only be explained as a change in our state of knowledge because it definitely does not involve any relevant change in the physical state of affairs of the involved electrons.

Schroedinger's cat is the most famous example of a "mere change in our state of knowledge" wavefunction collapse. Schroedinger's thought experiment involves a cat in a box with a poison vial that would be broken upon the decay of a radioactive particle. Since the particle may or may not have decayed, as long as the box is closed it is unknowable whether the cat is alive or dead. According to Quantum Mechanics, it is in a superposition of "alive" and "dead" states.

Schroedinger's motivation for coming up with such a gory example was to provide a *"reductio ad absurdum"*: a cat could not reasonably be considered both dead and alive at the same time and this wavefunction collapse had to be a mere change in our state of knowledge and quantum superposition would then not be a physical superposition in this case.

Unfortunately, "cat states" do exist experimentally (not with real cats), and there seems no way to prove Schroedinger's point.

The Barrett-Kok protocol provides an instance where there is a compelling argument for a wavefunction collapse that corresponds to a mere change in our state of knowledge. In the Delft experiment, the protocol works as follows:

Each of the two electrons is prepared separately into an unknown spin state, a superposition of $< UP >$ and $< DOWN >$ along some axis.

Then, an Electromagnetic pulse is applied to each of the electrons. This pulse is tuned to a particular transition between energy levels such that any electron that is in (say) the $< UP >$ state is stimulated to transition into an excited state, while there is no effect on any electron that is in the $< DOWN >$ state.

The excited state then decays, emitting exactly one photon and any electron that was in that state returns to its original, spin $< UP >$, state.

Note that the total number of photons emitted by both electrons is either 0 (both electrons were in the $< DOWN >$ state), or 1 (one $< UP >$ and one $< DOWN >$) or 2 (both $< UP >$).

These photon(s) are collected (inefficiently) into fibre optic cables and routed (with further losses) to a central station where there is a polariser, beam splitter and a pair of detectors.

On any given trial the possible results at this stage of the protocol are:

No photon is detected: The protocol fails - entanglement of the electron pair cannot be shown.

One photon is detected: Either the electrons are one $< UP >$ and one $< DOWN >$ or they are both $< UP >$, but one of the two emitted photons was never detected.

Two photons are detected: The protocol fails - they are both $< UP >$ (i.e. ignoring dark counts).

After applying this pulse and waiting a sufficient interval for any excited state to decay with high probability, the next step in the protocol is to flip the spin of both electrons. Any electron that was $< UP >$ is now $< DOWN >$ and *vice versa*. A second round of the protocol begins.

The stimulating pulse is applied again. Any photons emitted by $< UP >$ electrons in the excited state are collected, routed to the central site and the number of detections recorded.

If there is exactly one photon detected on the first round AND exactly one photon detected on the second round, then, and only then can we be certain that we have one electron $< UP >$ and one $< DOWN >$.

In this case, the pair is deemed to be in the entangled state. Note that it is critically important that we do not know, indeed we cannot know, which electron is $< UP >$ and which is $< DOWN >$. In fact, the notion that one is definitely $< UP >$ and the other $< DOWN >$ is incorrect: The total spin is zero.

There are two pertinent considerations.

First, every time the protocol succeeds, we can be certain retrodictively that the first round was a case of one $< UP >$ and one $< DOWN >$ and not a case of two $< UP >$ where one of the emitted photons was not detected. Therefore, when the protocol succeeds, we can be certain that the pair was already entangled after the first round. All that changes on the second round is our state of knowledge. Schroedinger's point is made: wavefunction collapse does not always correspond to a change in the physical state of affairs. Sometimes the quantum superposition is not physical, but only represents a lack of knowledge.

Second, every time the protocol succeeds, on each round of the protocol there was a total of one photon in the apparatus. This is noteworthy because the usual understanding is that the interference at the beam splitter is crucial, but this is an interference, not between two actual photons, but between the two possible ways for one photon to have been generated.

While there is no reason to doubt the reliability of the quantum formalism, we have seen that it involves blurring together the philosophically clear category distinction between physical influences and states of knowledge. It runs on a category error.

As much as I am reluctant to disagree with the physical interpretations of experts on this point, we know for sure that entanglement is not physically swapped by the second round photon to the electron pair, because it was already there.

There is then the question whether entanglement is physically swapped on the first round, or whether it already existed even before the first round and the Barrett-Kok protocol is not "swapping" entanglement at all. It is merely detecting instances where the two-particle system "fell" into entanglement as a consequence of the preparation.

Of course, in that case they would have to have interacted with each other at preparation, and they would need to be interacting with each other thereafter. All of which is, by assumption, considered impossible. The deep irony here is that this is the very same assumption that is falsified in the very same experiment!

Is it really better to interpret entanglement swapping as the causal consequence of physical interference between two entities only one of which exists?

However, let us admit for the purpose of argument the idea of a single "split reality" photon, that interferes with itself at the beam splitter. There is nothing about this notion that is so different in principle to the different fields emerging from different slits in a de Broglie wave scenario. These fields could have a real, physical interference at the beam splitter, but what cannot rationally be maintained is that their interference causes the entanglement: they could only have existed in a split reality form in the first place as a consequence of the system already being entangled.

Either way, we come back to the one and only rational physical interpretation: The electrons fall into entanglement and Barrett-Kok merely detects entangled states.

But why would two electrons, 1.3 Km apart, fall into either an entangled state, or a coaligned state? According to the Theory, 50 percent of all trials go one way and 50 percent go the other way. The explanation lies in the engineering carried out in the name of ensuring that the photons generated in either arm of the apparatus are "indistinguishable"[15].

[15] *i.e.* We cannot obtain *information* about which side the detected photon came from. On the face of it, not being able to obtain information seems crucial to what happens in the physical state of affairs *vis a vis* entanglement. Like the measurement problem, the question is raised whether Epistemology is playing a role in Ontology.

According to the paper [72], "The optical frequencies of the NV are tuned by a d.c. electric field applied to on-chip gate electrodes". The physical environment for each electron is manipulated to ensure both run at, for all intents and purposes, *exactly* the same frequency. As anyone who has ever tried to run plesiochronous (almost the same frequency) oscillators in the same laboratory would understand, the issue is not why do they lock onto each other but how could it be prevented: When the natural frequencies are sufficiently close to each other, whether the interaction mechanism is Electromagnetic or quantum and however weak it may be, the low energy state is always for the spins to align or counteralign.

For me, this example highlights the intrinsic difficulty with the quantum mechanical superposition of observables:

First, we destroy the clear category distinction in Philosophy between physical reality (Ontology) and knowledge about it (Epistemology), then we interpret our results in terms of those same categories. Having scrambled the omelette, when we interpret the analysis as a real physical interference, we are behaving as if it were never scrambled. As discussed in Chapter 5, the quantum mechanical superposition of observables is at once a beautiful thing for predictions, and a deal with the devil for interpretation.

7.8 Testing the Physical Model

There are some obvious ways to test the proposed distributed interaction interpretation. Since the entangled quanta have finite energy, the strength of their interaction is range dependent, most likely with a $1/r^2$ asymptotic behaviour.

One needs only to test the range dependence of the noise immunity of the experimental visibilities (a measure of how close the observed correlations are to the theoretical ideal). This can be done, for example, by injecting noise [77, 78] after the entangled state has been established, by increasing the temperature, by a suboptimal tuning of the NV optical frequencies or simply by introducing a programmable delay between the event ready signal and the Bell measurements.

It is noteworthy comparing [79] with [72], that the protocol success rate in two experiments by the same group at different separations was not strongly range dependent. (The difference in the reported protocol success probabilities is explained by the extra attenuation due to the long fibre optic cables in the latter case.) This does not imply that the capacity for creating entanglement is not range dependent. This is a go - no go situation and the experiments may not be close to the range limit beyond which entanglement does not occur. Given that the protocol success rate is already very low, it is not practical to test for range effects simply by further increasing the range - one needs to test the range dependence of the correlations against a known and controllable imperfection in the experimental setup.

The proposed experiment does not require spacelike separation and can be performed in a laboratory, with separations up to say 30 metres. That will avoid the fibre optic cable losses, significantly speeding up the experiment. Finally, the Delft TU group has made substantial progress towards using cavities to collect the photons emitted at the NVs, which promises to improve the protocol success rate dramatically [80].

Any one of the above methods can reduce the lifetime of the entangled state until it becomes comparable with the duration required to conduct the measurements, at which point the correlations begin to wash out. Does the amount of the temperature rise, detuning, noise or delay that reduces the visibilities by (say) 3 dB depend on the distance between Alice and Bob?

If it does, then EPR correlations can reasonably be interpreted as just another wave phenomenon.

Note how this proposal is different from Many Worlds [81], retrocausality [82], superdeterminism [70, 71], superluminality and every other proposed resolution of the EPR crisis: It can be excluded by experiments that we already know how to perform. Finally, quantum nonlocality is in the process of being industrialised. Regardless of interpretation, isn't it important to understand the noise immunity better than we presently do?

7.9 The Double Bell Paradox

Consider the meaning of the term "causal correlations" at the level of individual trials.

It describes a causal relationship between Alice and Bob's datasets that has been experimentally demonstrated at a statistical level, by considering correlations between the datasets in a large population of individual trials. Let's assume again that the experimenter who makes the first measurement is assigned the name Alice. The finding is that certain correlations are stronger than anything possible in the absence of a causal influence from the intervention by Alice to the result by Bob.

Now, in order for there to be a causal influence at the population level, there must be causal influences on at least some individual trials.

On individual trials, the Bell measurements are binary input / binary output: Each experimenter has a binary input, the choice of measurement axis, a or a' for Alice, b or b' for Bob; each experimenter obtains a binary output: $+1$ or -1, where $+1$ means spin $< UP >$ and -1 means spin $< DOWN >$ along the chosen measurement axis etc.

Let us assume further that Bob's measurement axis is fixed along b. The existence of a causal relation on an individual trial then means that Bob's result depends on Alice's choice: If Alice chooses a, Bob gets (let us say) $+1$; If she chooses a', Bob gets the opposite result, (say) -1.

It is of course true that Alice's result has a role in the correlations, but it isn't relevant here. The above could be written as follows: If Alice chooses a and gets result A, then Bob gets (let us say) $+1$; If she chooses a', and gets result A', Bob gets the opposite result, (say) -1. At the level of individual trials "causal influence" can only mean that if she makes a different intervention, he gets a different result.

Now, we can unfix Bob's choice, and divide the population into two subensembles, corresponding to Bob chose b and Bob chose b'. In each subensemble, what the quantum predictions mean is that, on at least some occasions, Bob's result depends on Alice's choice.

This is strictly implied in the logic, but for those who prefer a formal analytical development, Professor Percival developed a specific formalism just for this purpose, which describes EPR experiments in terms of transfer functions between the parameters on one side and those on the other [30, 31].

Let us refer to those individual trials that feature causal dependence as causal dependence trials. We don't know which specific trials they are, but we know that they constitute a substantial fraction of all trials. Rather than delving into the algebra to quantify that, Percival notes that there is a "nonzero probability" of causal dependency on any given trial.

He then considers the case of two identical EPR experimental setups in relative motion with respect to each other along the x-axis. Percival called the experiments "Experiment 1" and "Experiment 2", but in the description here, they will be referred to as the "Red" experiment and the "Green" experiment. The experimenters will be Red Alice and Red Bob, who conduct the Red experiment, and Green Alice and Green Bob, who conduct the Green experiment.

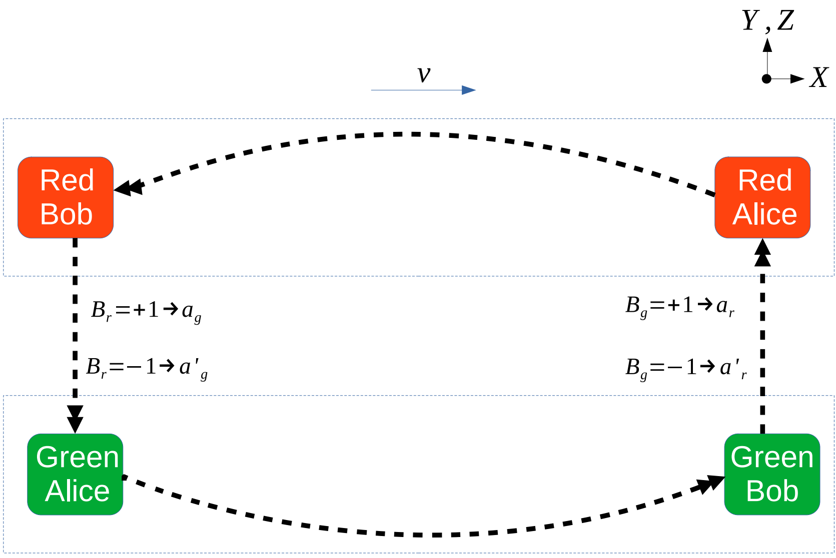

Figure 7.2: *The double Bell paradox. The red frame is moving to the right at speed v relative to the green frame. According to relativity of simultaneity, there is a temporal loop: Green Bob's measurement result selects Red Alice's measurement axis which causally influences Red Bob's measurement result, which selects Green Alice's measurement axis BEFORE her measurement causally influences Green Bob's original measurement result. On at least some occasions, the loop is forbidden.*

Red Alice's measurement axis choices are a_r and a'_r. Similarly, Green Alice's axis choices are a_g and a'_g. Red Bob and Green Bob both use fixed measurement axes, respectively b_r and b_g. To provide a concrete example, specific choices for the various measurement axes can be chosen relative to some arbitrarily chosen direction in the transverse plane, for example: $a_r = 0$, $a'_r = \pi/2$, $a_g = 0$, $a'_g = \pi/2$, $b_r = 0$, $b_g = 0$.

With reference to Figure 7.2, the basic idea is as follows. Green Bob performs a measurement, the output of which is used to select Red Alice's measurement axis. If he gets a $+1$ she chooses a_r but if he gets a -1 she chooses a'_r. Red Alice then performs her measurement, which influences the state of the photon received by Red Bob. On causal dependence trials in the red experiment, Red Bob's result is causally dependent on Red Alice's choice, which was determined by Green Bob's result.

Red Bob's result is then used to choose Green Alice's measurement axis. She then performs her measurement, and according to Special Relativity (as we shall see below), this happens BEFORE Green Bob's original measurement. Meanwhile, on causal dependence trials in the green experiment, Green Bob's result depends on Green Alice's choice.

Let us refer to those occasions where both the red and green experiments have causal dependence trials together as "double causal dependence trials". On such trials, Green Bob's result causally depends on Green Alice's choice, which is determined by Red Bob's result, which causally depends on Red Alice's choice, which was determined by Green Bob's result. Thus, there is a causal temporal loop in which Green Bob's result depends upon itself.

There are two possibilities: either a $+1$ causes a $+1$ and a -1 causes a -1, which is a permitted loop, or a $+1$ causes a -1 and a -1 causes a $+1$, which is a forbidden loop. As we shall see below, what Percival showed is that forbidden causal loops are predicted on some double causal dependence trials.

Experimental setups like this one are so routinely analysed that there is a name for it: "before-before" timing. While the construction of before-before timing scenarios is mundane in Special Relativity, let us run through the details anyway, for the benefit of non physicists. The only new element, which leads to the paradox, is the abovementioned spacelike causal dependence between events in the same frame.

The Red and Green frames of reference are in the standard configuration, where their x-axes are parallel, coincident with the direction of the relative velocity between the Red and Green frames. Red Alice and Green Bob are located at $x_r = 0$ and $x_g = 0$ respectively, and the origin is chosen so that, as the experiments slide past each other, they are colocated at $t_r = t_g = 0$. The y and z coordinates are not involved.

Red Bob and Green Alice are located at $x_r = -L$ and $x_g = -L$ respectively, where L is the distance from Alice to Bob in each experiment (as measured in the inertial frame of the experiment).

This is all as shown in Figure 7.2. All clocks are Einstein synchronised.

In the interest of clarity, there are a few corners cut in the math here. In particular, the Lorentz-Fitzgerald contraction of each apparatus in the frame of the other experiment will be ignored. It is easy to show that the light time introduced between (for example) Red Bob and Green Alice can be absorbed into the parameter ϵ without changing the conclusions.

From the point of view of the Green frame, the relative velocity of the Red frame is $+v$ in the x-direction so that at $t_g = 0$ we have, from the Lorentz Transformation:

$$t_r(x_g = 0) = 0$$

(this is the time showing on Red Alice's clock when Green Bob's clock reads 0.)

$$t_r(x_g = -L) = \gamma \frac{vL}{c^2}$$

(this is the time showing on Red Bob's clock when Green Alice's clock reads 0.)

From the point of view of the Red frame, the relative velocity of the Green frame is $-v$ in the x-direction so that at $t_r = 0$ we have:

$$t_g(x_r = 0) = 0$$

(this is the time showing on Green Bob's clock when Red Alice's clock reads 0.) Note that Green Bob and Red Alice agree on the readings of each other's clocks.

$$t_g(x_r = -L) = \gamma \frac{-vL}{c^2}$$

(this is the time showing on Green Alice's clock when Red Bob's clock reads 0.) Note that Green Alice and Red Bob both agree that his clock is advanced by $\gamma vL/c^2$ relative to hers. Since v can be assumed to be non-relativistic, $\gamma \sim 1$ throughout.

Now, suppose that in the Green frame, Green Bob conducts a measurement at $t_g = 0$ and suppose that he gets the result $+1$. He immediately passes the result to Red Alice, who uses it to select her measurement axis according to the rule: if Green Bob gets $+1$ choose a_r; if Green Bob gets -1 choose a'_r.

Very shortly thereafter, Red Alice and Red Bob then perform near simultaneous measurements in the Red frame. Let us assume that Red Alice's measurement is at $t_r = \epsilon$ and Red Bob's measurement is at $t_r = \epsilon + \delta$, where $\epsilon << vL/c^2$ is the short delay between the measurements by Green Bob and Red Alice and $\delta << vL/c^2$ is the short delay between the measurements by Red Alice and Red Bob. Note that L is a free parameter: even for small values of v, the quantity vL/c^2 can be made arbitrarily large.

Red Bob now immediately passes his result to Green Alice, who chooses her measurement axis according to the corresponding rule: if Red Bob gets $+1$ Green Alice chooses a_g, and if Red Bob gets -1 she chooses a'_g. At the moment when he passes the information to Green Alice, Red Bob's clock, simultaneous with Red Alice's, reads $\epsilon + \delta$, so Green Alice's reads $\gamma(\epsilon + \delta) - \gamma vL/c^2$. After another short delay, ϵ in her frame, she is ready to perform her measurement at $t_g = \gamma(\epsilon + \delta) - \gamma vL/c^2 + \epsilon$. Green Alice then makes her measurement at $t_g = -\delta$, just before Green Bob's measurement at $t_g = 0$.

As long as $(\gamma+1)(\epsilon+\delta) < vL/c^2$, there is a temporal loop as described above, where Green Bob's measurement is before Red Alice's which is before Red Bob's which is before Green Alice's, which is before Green Bob's. On double causal dependence trials, the temporal loop is causal and Green Bob's measurement result depends on his own measurement result in the same trial.

The dependence of Green Bob's result on itself in this causal temporal loop may be permitted or forbidden, as follows. Assume Green Bob's "originating" result is $B_g = +1$. There are four possibilities for the causal links in the loop, which are as follows:

$$B_g = +1 \to a_r \to B_r = +1 \to a_g \to B_g = +1$$

This is permitted.

$$B_g = +1 \to a_r \to B_r = -1 \to a'_g \to B_g = +1$$

This is also permitted.

$$B_g = +1 \to a_r \to B_r = +1 \to a_g \to B_g = -1$$

This is forbidden.

$$B_g = +1 \to a_r \to B_r = -1 \to a'_g \to B_g = -1$$

This is also forbidden.

Note that a permitted loop can be converted into a forbidden loop simply by inverting the rule that Alice uses to select her measurement axis in either the Red or the Green experiment (but not both). We cannot say in a particular causal dependence trial which choice of axis causes what measurement result, but we can say that one choice causes a $+1$ result and the other causes a -1 result.

We can also say that the mapping on causal dependence trials from a or a' to $B = +1$ or -1 can only be either fixed (the same on all causal dependence trials) or variable (sometimes $a \to B = +1$ and sometimes $a \to B = -1$).

In the first case, half of the ways to choose the measurement axes are associated with forbidden loops on all double causal dependence trials and in the second case all ways to choose the axes are associated with forbidden loops in some double causal dependence trials.

Hence we arrive at Percival's conclusion that, one way or another, forbidden causal temporal loops are predicted.

This is not like EPR, where the Theory predicts something we did not believe could happen, but does. We might use this self-inverting measurement result to break the poison vial in the Schroedinger's cat thought experiment. This is now a case of a cat that is both dead and alive *while we are all looking at it.* This does not happen. Forbidden causal temporal loops do not exist. The notion of a classical signal, a measurement result that has been recorded and read out to a human being, being the cause of its own inversion is not possible, and it follows that there is something wrong with the assumptions.

7.10 Special Relativity is not "Wrong"

This is not a case of "maybe" the assumptions are wrong. We can be certain that at least one of the assumptions is wrong, so what are the assumptions? There are three assumptions in the double Bell paradox:

1. The quantum predictions for individual EPR experiments.
2. Independence of the Red and Green experiments.
3. Relativity of simultaneity. Specifically, when observers disagree about the temporal order of spacelike separated events, there is no objective fact of the matter. The only simultaneity is that which appears to observers.

The choice is clear. Either (1) the experimental facts are wrong, (2) the experiments collude with each other, or (3) choices made in the present do not depend on facts that lie in their own future (as relativity of simultaneity suggests can happen with before-before timing).

Now that Quantum Mechanics has made the temporal loops causal, in order to retain the already counterintuitive relativity of simultaneity assumption, one has to compound the affront to common sense: Either facts are no longer facts or experiments collude.

Presumably, the Physics community will either go that way, or it will continue to do what was done with this and other similar paradoxes [32] for the two decades before the EPR loopholes were finally closed: sweep them under the carpet. The relativist Metaphysic is far too important to let experimental facts get in the way.

Let us consider the more sensible approach, the approach that physicists who were actually doing Physics (as opposed to dabbling in Metaphysics) would take: The fact that an observer sees events in a particular temporal order does not necessarily imply that they really are in that order. Relativistic simultaneity is a matter of appearance, not a matter of fact, but there is a matter of fact.

Is Special Relativity still valid for making predictions? Yes it is. This change in the status of relativity of simultaneity has no impact on the Special Theory *as a Physics Theory.* After all, Physics Theories are *supposed* to be about appearances, not Metaphysical truths. At least that's how it was before Einstein developed Special Relativity.

The known facts in relation to quantum before-before timing paradoxes do force any physicist who is doing Physics to acknowledge that there is an objective fact of the matter about temporal order. This in turn implies the objective existence of physical reality, independent of observers.

What was shown in Chapter 4 is precisely what is required: An objective reality, overtly Metaphysical, equipped with a preferred frame but characterised by Lorentz covariant observables that are fully compatible with Special Relativity.

Moreover, Special Relativity is not wrong in these cases. On the contrary.

The Theory *correctly* predicts that observers using the protocol will sometimes disagree about the temporal ordering of events. Since the underlying temporal order is objective, what the Special Theory is making clear with this prediction is that an observers' facts do not necessarily coincide with the objective truth. That's good Physics. The whole point of the Theory is that the objective truth value of an observation is not a relevant consideration.

By contrast, the traditional assertion that "both observers are correct", is not Physics, but Metaphysics. Such misguided defenses of the Theory always involved drawing conclusions that are outside the domain of Physics.

Observables are inherently relative. Relative, that is, to the observer. An observer's facts are not nature's underlying, objective facts. Instead they are the way that her objective facts appear to an observer whose condition of motion is, *according to the Special Theory*, an essential part of every measurement she makes.

Even as it correctly predicts what each and every observer does, in fact, see, the theory also correctly predicts that sometimes what the observer sees is wrong. Therefore:

It is not the Special Theory that is wrong, but the observer.

Even without an observed preferred frame as part and parcel of the theory of Lorentz Invariance, we would still know that there is an underlying fact of the matter about temporal order because, if it were not so, if the only truth is relative to the observer, then a cat can be alive and dead while we are looking at it.

The vital thing about the Theory is that these errors in an observer's assessment of the temporal ordering of events never lead to errors in the predictions that he or she makes when using the Theory[16]. Again, it has never been the role of Physics to interconnect Nature's underlying, objective facts. The role is to interrelate the observables.

We also saw, in Section 6.1, that Mother Nature herself uses Einstein clock synchronisation when she establishes simultaneity at a distance in the form of the space independent phase of a quantum. There can be no firmer basis than that for us to accept the same kind of protocol.

Consequently, it should not be thought that relativity of simultaneity is false *per se*, only that we need to understand what a clock synchronisation protocol does and what it does not do. The existence of de Broglie waves illustrates this perfectly: The objective truth here is that Mother Nature herself cannot, in practice, establish a space independent phase independent of the condition of motion of the quantum. Correspondingly, when we use the same method it does not synchronise clocks independently of their condition of motion, a conclusion that is staggeringly irrelevant in practice.

The problem here is with the physicists, not the Physics. Having established for themselves the appropriately limited domain of relations between observables, the Physics community routinely flaunts the boundary. In particular, with the doctrinal exclusion of the preferred frame concept, they have painted themselves into a corner on what is no more than a belief system: It was insisted that mother nature be limited in the same way that Physics (incorrectly) once thought we were.

There is a clean and simple path forward - acknowledge the logic error of the early 20^{th} century, accept the observed facts in relation to the CMBR preferred frame and adopt its consequences for objective temporal order of events.

Nothing changes in the Physics, but it may take a few more generations of physicists for the paint to dry.

[16]Note in this context that spin measurements on different particles always commute so the *statistical* quantum predictions for EPR experiments similarly don't depend on the temporal order of the Bell measurements.

Chapter 8

The Worst Mistake in the History of Physics

In the Bohr-Einstein debate, both sides agreed on one thing: instantaneous action at a distance between point particles could not happen. This belief has been falsified in the last Chapter, and it will be falsified again in the next. We are reminded of another, even more famous, debate between intellectual giants, when the great mathematician and philosopher, Leibniz criticised Newtonian gravity on the grounds that the new Theory was describing instantaneous action at a distance.

Newton innoculated himself with two of the best known remarks in all the history of Physics. He initially responded to the critique with this:

> *"I frame no hypotheses; for whatever is not deduced from the phenomena is to be called an hypothesis; and hypotheses .... have no place in experimental philosophy."* — Sir Isaac Newton [83].

Perhaps he should have left it at that, but he could not let it go. Fifteen years went by until, finally, he framed an artful Hypothesis, in the negative:

> *"that one body may act upon another at a distance thro' a Vacuum, without the Mediation of any thing else, by and through which their Action and Force may be conveyed from one to another, is to me so great an Absurdity that I believe no Man who has in philosophical Matters a competent Faculty of thinking can ever fall into it."* — Sir Isaac Newton [84].

These remarks left an indelible mark on the consciousness of every physicist who ever lived. The first one focusses the scientific discipline his beautiful work created on finding relations between observables, as opposed to explaining the phenomena. It is a sufficient reason for the title of this book. There is no doubt in my mind that he was right, but it is also the first paving stone on the path to modern mysterianism, what J.S Bell calls "la nouvelle cuisine"[47].

I am confident that if he were around today, Newton would have defended his Mechanics, and his absolute space, against the relativist Metaphysic: fix the speed, vary the inertia and Modern Physics fits neatly within his originating paradigm.

The second remark has had an awesome negative power. The father of Physics gave the one and only true path for any theory of interaction mechanisms. Anything else but interaction as a thing that travels between primary participants must be the product of incompetence. Ouch!

When Maxwell's Laws and the d'Alembert wave equation were found, Newton had long ago put the possibility that interactions do not travel between particles beyond consideration so it was immediately recognised that the speed of such interactions as they travel between two massive particles would have to be constrained by the speed of light, and the retarded interaction paradigm was born.

I was blissfully ignorant of all this when, soon after my sixteenth birthday, I found myself musing about the question that has stimulated all my interest in Physics. It was my final year of high school and the subject was Coulomb's Law. I wondered how that works:

How do two charged particles interact with each other at a distance?

Of course, the physicists would have it all sorted out, or so I thought, but there was no internet in those days. In any case, I wanted to see if I could figure it out for myself. After all, this is the most basic question in Physics, how hard could it be?

The first idea that came into my mind was, of course, that they would interact via an intermediary transmitted from one particle to the other. After just a few minutes reflection, it became obvious that this idea must be wrong. Little did I know!

I knew that I was wrong because I could only write the first line of the play. The intermediating agent travels from one toward the other.... and then?

How does particle A know where particle B is going to be in the future, when the intermediating agent that A transmits finally interacts with B? I would need agents to be radiated in all directions. Where does the energy for all that radiation come from?

How does particle A know when and where the intermediating agents are absorbed: either I just violated Newton's third Law or I had not resolved the problem of action at a distance at all, merely shifted it to a different level of knowledge at a distance.

The linear momentum of my intermediating agents could only be in the direction they are moving, radially away from the source. How is that consistent with the existence of attractive forces?

It was not that these particular issues couldn't be resolved in principle but that the whole thing was a can of worms. The closer you look the worse it gets.

Over the next several years, I became absorbed by the problem, and I got precisely nowhere with it. I was both dreading and looking forward to the day when all would be revealed in an appropriate university setting, and I would feel like a fool for having missed the obvious.

At long last, the big day came in a third year Electromagnetics course, which I suppose means the year must have been 1978. The answer given was, of course, retarded interaction. The lecture material was for me inadequate because it failed to address any of the issues above. After the lecture I went straight to the textbook to look up the questions at the end of the relevant chapter. There were questions about calculating the fields of the moving particle, and questions about calculating the response of the moving particle to a given field, but the one I was looking for was not there. There was nothing about the Dynamical interaction between two moving charged particles. How is that not what this course is supposed to be about?

I tried to formulate it anyway, but (as is well known) even the first step doesn't work. The forces in the undergraduate theory don't conserve momentum.

Further research soon showed the bottom line. Classical Electrodynamics is a wonderful theory in all respects but one: It does not cope with actual Dynamics problems. In particular, it does not cope with the Two Body problem. Every attempt to apply Classical Electrodynamics to real Dynamics problems opens.... a can of worms[1]. Moreover, this Mathematical can of worms has its roots in the very same issues that had led me to reject retarded interaction as a sixteen year old, long before I knew its name.

Physicists continue to work on Classical Electrodynamics and other theories to this day, to no avail. The summary of this conference report [87] claims that there is now a real prospect of one day being able to "formulate the problem". That was written in the 1970's, and we are still waiting. Many of the biggest names in 20^{th} century physics presented papers to this conference. Anyone who thinks that Quantum Field Theory is highly mathematical should look at this report to see the mathematical heights that have been scaled - without success - trying to solve the most basic problem in Physics.

At this time the Aspect experiments had not yet been reported, and in any case I was blissfully unaware of the Bell Inequalities. I did however know that Classical Theory had long been superseded. If there was an answer to my question, it was only to be found in Quantum Electrodynamics, but I also knew by reputation that any Theory with the word "quantum" in it does not answer the interesting questions.

It turns out that Quantum Electrodynamics is not at all silent on this issue. The most usual approach for interpreting the interaction mechanism looks like this:

The interaction is mediated by "virtual photons", whose lifetime is limited by the time-energy version of the Heisenberg Uncertainty Principle. The "virtual" concept specifically addresses the conservation of energy issue above. The energy for the virtual particles is "borrowed" from the vacuum, and only made permanent when a particle is absorbed. At the time that seemed to me like a cheat, but in retrospect (with a better understanding of what Physics is and what it is not) the mathematics does not seem at all unreasonable.

The objection that remains is to the verbal description that has been developed in an unsuccessful attempt to shoehorn the formalism to fit the usual retarded preconception. This part of the exercise is remarkably deficient:

The virtual particles propagate at the characteristic velocity, but they do NOT (in general) propagate from A to B. Nor are they particles, in the sense of being well-localised. The "virtual" photon is instead to be thought of as a momentum eigenstate that occupies all of space, which might be propagating in any direction.

[1]The real problems begin with the pre-acceleration of the Abraham-Lorentz force [85], where the particle begins to respond *before* an external force is applied. Konopinski [86] goes further into the issues, but these textbooks barely scratch the surface of the research.

For an attractive force, there is a preponderance of absorbed virtual photons that are propagating in the direction from target to source. They approach the target from the wrong side.

That's clearly not a retarded interaction mechanism. The verbal description fails to reconcile the Theory with our presumption that the primary participants in a long range interaction are tiny corpuscles. Although there is a "conveying of force and action from one to the other" it isn't travelling from one to the other.

In fact, upon closer inspection there are no retarded interaction mechanisms anywhere in Modern Physics.

The idea that the interaction occupies all of space does however begin to answer my original question. The "virtual photon" is not conceived of as originating at the place where we "find" the pointlike source of the interaction, and nor can it be conceived of as interacting with a pointlike target at some other place.

Self evidently, if the interaction mediating agent is distributed then so must be its cause and so must be its effect. The mathematics can only be describing a distributed interaction between a distributed source field and a distributed target field.

The "virtual particle" device is then redundant. It's *raison d'être* was to give the interaction some semblance of physical reality in the time between leaving one particle and arriving at the other. The discussion never worked anyway, but once it is recognised that the interaction is between overlapping distributed fields, there is no need to introduce any time interval during which the interaction is in between the particles. The "virtual particles" of the Electromagnetic interaction never have to exist at all as distinct entities in their own right. They need be no more than a quantification of the direct transfer of luminal energy-momentum from a source field into a target field; more of an accounting entry than a thing in itself.

Furthermore, it is not at all clear if the widely distributed book keeping entries of the Theory are correspondingly distributed in the world. They could equally well be representing a sum over a myriad point-like interactions that occur independently of each other at all points where the source and target fields overlap.

There are thus two distinct Metaphysical concepts of distributed interactions that can be considered in the ontological context of the cellular resonator model.

In the first case, an individual resonator in one field interacts with an overlapping resonator in the other. Such interactions take place independently of each other throughout the entire space.

In the second case, more closely aligned with the usual momentum eigenstate interpretation, the interaction is a distributed quantum of energy-momentum that is transferred at once, as a single whole, between the source and target field systems. This whole, *fundamentally distributed*, interaction quantum is physically instantiated by a distributed system of synchronised, pointlike interactions between groups of colocated resonators in each of the source and target field systems.

In this case, we find ourselves confronted by the same issue that arose with spin measurements in Sections 7.3 and 7.4: How are these parallel occurrences synchronised without introducing nonlocality?

We saw that it *can* be done in a local realist manner, so there is no fundamental logical roadblock, but is the toy model synchronisation mechanism given in Section 7.4 satisfactory? I do not think so and if I had a better explanation, I would give it.

There is a wide range of facts that supports the idea of energy quanta as widely distributed wholes that interact as wholes via distributed interactions that are themselves widely distributed wholes. We shall return to this issue in the final Chapter but the present mystery is the ongoing universal commitment to retarded interaction, especially *vis a vis* spacelike causal correlations and the definition of locality in the EPR community.

Retarded interaction doesn't make sense in the physical logic. It didn't work in Classical Theory. It isn't present in Modern Physics. It has long ago been understood that not all the energy of the golf ball is inside the image of it. Quantum Fields have replaced point particles. We know that de Broglie waves are distributed. Modern EPR experiments prove beyond any reasonable doubt that interactions between different particles are not retarded. Finally, observations show that the acceleration of the Earth is directed towards the Sun's instantaneous position, not its retarded position [88].

Surely, this deeply flawed, redundant, Metaphysical idea is the worst mistake in the history of Physics. So why the ongoing commitment?

One might assume that it is due to the lack of a viable alternative Metaphysic. Perhaps, but in all my experience physicists barely consider such questions. What seems more reasonable is that Physics is imprisoned by its own history. Instead of asking "How does this work?" it asks "How can we fit this into the framework that has been handed down for generations?".

The empirical evidence for this is strong.

Relativistic causality is one of the best known concepts in relativity theory. For many decades, it was formally defined as the observer independence of the notion that causes must precede effects. Accordingly, any effect of a point event must lie in its forward light cone, any cause of a point event must lie in its backward light cone, and there can be no causal influences between spacelike separated events. Spacelike causal correlations exposed the failure of this definition.

Instead of addressing the question why causal relations aren't constrained by light cones, Physics has adjusted itself to the now accepted fact of EPR correlations simply by changing the definition of causality. Relativistic causality is no longer about causality. It is now about "signalling". There is a "no signalling" theorem [89] to the effect that EPR correlations cannot be used to transmit information, so the new position is effectively that only signals *that transmit information* count as causal. Seriously.

The present generation of physicists has already been inculcated with the new definition. Faced with a black and white conflict between (the orthodox understanding of) the Theory and the facts in the world, Physics has chosen to retain the orthodoxy and reject the world. The new definition of causality keeps the retarded part (light cones) and deletes the causal part. Why throw out the baby and keep the bathwater if not to preserve an existing paradigm?

Let us briefly consider the difference between the Special Theory and the orthodox understanding thereof. One of the fastest, cleanest ways to derive the Special Theory is to consider two observers in relative motion who are colocated at $t = 0$ with a light flash that propagates at c in all directions radially away from some point event. From the second postulate, the receding wavefront is spherical for both observers, which establishes a connection between their respective coordinate systems. One arrives at the Lorentz Transformations in a few lines of algebra. It's very nice. Once one has the coordinate transformation, the 4-vector algebra follows. There's a step or two more to get the full mechanics, which is not relevant here.

What is important to notice is that the usual development refers to movements (of light waves), but there is no reference to causality, interaction (retarded or otherwise), or any other similarly dynamical concept. The passage from the Special Theory to the orthodox understanding in relation to causality was intended to address the question how the Theory constrains causal relations, however the singular advantage of Special Relativity is that it is silent on the question.

Instead, what actually happened is that the traditional understanding of a causal influence (as a thing that travels from A to B) was crudely and uncritically passed down from Newton to Lorentz and then imposed upon the Special Theory after the fact. That traditional Newtonian understanding was, however, from the very start predicated on the ontological status of point particles.

Although relativistic quantum fields have destroyed the ontological status of point particles, retarded interaction is still considered synonomous with Special Relativity. It remains beneath suspicion because it is universally presumed to be an integral part of the Special Theory. This conflation of retarded interaction with Special Relativity is completely wrong. Special Relativity never had any need for it.

In fact, once it is recognised that all the inherently relativistic physics theories are field theories then Special Relativity can finally be seen as the foundation stone for EPR correlations.

Chapter 9

Distributed Interaction Mechanisms

9.1 Introduction

The concept of distributed action raises the question how to derive or induce the usual empirical Force Law relations from Classical Physics, which are written in point form.

Once we are given relativity theory and the Coulomb Law from Electrostatics the basic Electrodynamics (for particles moving at constant velocity) follows very easily. Einstein showed how the Electric field of a stationary particle transforms into the Lorentz contracted Electric and Magnetic fields of the moving particle. Alternatively, the Coulomb Force can be seen as the first three components of a 4-force, which is defined from $\mathbf{F} = d\mathbf{P}/dt$, as the time derivative of the energy-momentum 4-vector. Since energy-momentum is a 4-vector, its modulus is invariant - the same for all observers, so the 4-force must be normal to the energy-momentum 4-vector. This constraint similarly introduces the magnetic field components [90].

All of this is well known and presumably appears in many textbooks. The same applies to the close relationship between Maxwell's Laws and the Special Theory. There is no need to go over all that ground again here. We have Special Relativity so it all follows in the usual way. Consequently, this Chapter can be limited to the question of how to cover the Coulomb Law in a distributed, luminal wave particle model that is also applicable to the Dirac Theory, matter beam interferometry, spacelike causal correlations and gravity.

Just as the basic mechanics quantities attributable to "point particles" are given as integrals over all space of the corresponding mechanics variables at the field level, so must the total force acting on a notional "point particle" be calculated as an integral over all space of the rate of change of its field momentum density. According to the Principle of Local Action, the interaction between two fields at a given point can depend only on the fields local to that point.

For an infinite range (*i.e.* $1/r^2$) force, this interaction will be modelled by a first order nonlinearity. In that case, the rate of transfer of linear momentum from a source field into a target field will be proportional to a *suitable* product of the two fields.

Consider then the integral over all space of the product of two $1/r^2$ scalar fields, which has a general form like this:

$$\int\int\int \frac{1}{|\mathbf{r}-\mathbf{r_A}|^2}\frac{1}{|\mathbf{r}-\mathbf{r_B}|^2}dxdydz$$

It is easily shown that the result of such a calculation is not proportional to $1/R^2$, where $R = |\mathbf{r_A} - \mathbf{r_B}|$ is the distance between our notional point particles. Also, both the force between particles and the rate of change of any field momentum density are vector quantities so, for comoving particles, the local interaction must be a vector whose space integral is a vector proportional to $\hat{\mathbf{R}}/R^2$. Once we have the result for comoving particles, Special Relativity generalises the result to particles in relative motion[1]

There may be many ways to achieve that result, but it is not the purpose of this Chapter to recommend any particular approach to nonlinear field dynamics. On the contrary, the entire reconstruction of the foundation theories throughout this work is intentionally carried out at the level of Mechanics, limiting the physical assumptions involved to the general notion of a wave ontology and avoiding nonlinear dynamics entirely.

Were it not for the rather basic problem above, I'd have been inclined to forget about Classical Physics. When considered from the wave perspective, the idea of real corpuscles is just too messy. The mechanics of wave systems is naturally Lorentz covariant, and we shall see in the next Chapter that the most natural approach to gravity leads to a curved space, generally covariant model. The non-relativistic Schroedinger Equation only causes conceptual confusion in comparison to the elegance and perfection of Dirac's true wave equation, and so on.

The original purpose of the model in this Chapter was therefore simply to show that distributed field-field interaction models, which reduce to the usual Classical model, *i.e.* Coulomb's Law, *can* be constructed *in principle* and no more. The treatment is unapologetically *ad hoc* because the entire Classical system is *ad hoc*: The point particles were always put in by hand, and interaction at a distance between them is something that only exists in the mind, not in reality. We have already seen that pathologies emerge as soon as the Classical, particle based approach engages with real Dynamics problems.

Nonetheless, there are several other good reasons to include a distributed action model for the Coulomb Law:

First, readers will be interested how an underlying wave explanation of modern Physics relates to the basic undergraduate Physics.

Second, the Coulomb Law is still used in Quantum Mechanics, where one routinely writes in the scalar and vector potentials for a point particle in an applied field.

Third, in my own personal experience, even after identifying the logical necessity for it, distributed action was difficult to get used to. I experienced the deeply ingrained mental habit of visualising points: point particles and / or point events. Discussing it with several of my physicist contacts, the questions that came back showed that they were doing the same. This one is here, that one is there, the interaction has to go from here to there, right? Hopefully, going

[1] At uniform velocities. After that, things go downhill.

through this exercise will help readers to understand how distributed interaction mechanisms work in general. It may even help people to notice that, actually, they never understood how retarded interaction works.

Fourth, corresponding to the logical difficulties with retarded interaction, there is a major glitch in the way that the Classical model "balances energy accounts", as follows:

Distributed action models have the salient virtue that they automatically respect local conservation Laws, whereas retarded models don't. Consider for a moment the energy fluxes in the Classical model, where the field is produced by the "body" of the particle and then propagates away from it. The field energy density is proportional to the square of the field amplitude: $\rho_E \propto \mathbf{E}^2$, where E is for energy, and boldface $\mathbf{E}$ is for the Electric field strength. So the field energy density varies as $1/r^4$. In the Classical model, this induces a total power radiated through any spherical surface (whose total area is $4\pi r^2$) that is proportional to $1/r^2$. In that case, the total energy inside any given spherical annulus must increase with time. Meanwhile, the space integral of the field energy density is finite[2]. This basic contradiction is inherent in the idea of retarded interaction.

Of course, when one talks about fluxes in the Classical model, one only talks about Gauss's Law, which concerns the flux of the field strength rather than the energy. The point here is that the Classical model does not give a sensible result on energy fluxes, but it should. While this is never mentioned in textbooks (which focus on the things that work) this issue has always been known. My view is that, if there is to be a classical-like model to teach the children, then it should make proper sense both ways: energy fluxes as well as field strength fluxes, and of the two, the energy flux is more important.

The contradiction is removed if the asymptotic fields do not propagate away from the particle. The fields in the physical model of the electron here are, of course, propagating at c, but they are instantiated in well-localised resonators vibrating on the spot as opposed to moving away from the centre.

9.2 Choosing the Source and Target Fields

It will be recalled from Chapter 4 how the incremental momentum boost generator, (4.10), was initially motivated by considering a vector force acting on the (scalar) energy density of a target field. Given Einstein synchronisation, boosting a particle was equivalent to a change of observer and the incremental change in the momentum density of a target field was proportional to its magnitude but independent of its direction of propagation. Although it was important in Chapter 4 not to rely on Special Relativity to get that result, Appendix C also shows that this result from Special Relativity applies to any relativistic field theory.

Consequently, it makes good sense for the present Chapter to consider the first order nonlinear interaction between a vector source field and a scalar target field which is spatially distributed in proportion to the field energy density of the target particle. (Note that the target's field energy density is obviously not a viable choice because the model must apply to charged particles of very different masses.)

[2]At least it does not diverge as $r \to \infty$.

Let us refer to the two interacting particles as particle A and particle B. In order to calculate the action of particle A on particle B, we shall associate a vector field with particle A (as source) and a scalar field with particle B (as target). The product of these two fields at any point in space will equal the local rate of change of the target's linear momentum density. The space integral of this local interaction equals the total rate of transfer of linear momentum from the A fields into the B fields, which will be designated $\mathbf{F}_{AB}$. Clearly, any linear momentum transferred from A into B is given up by A, so there is an equal and opposite reaction force acting on A, which is the space integral of the reaction force in any incremental volume. Global conservation of energy-momentum thus results from local conservation.

As for the action of particle B on particle A, the roles are reversed and the result of the space integral is $\mathbf{F}_{BA}$.

The total force acting on the B particle is then equal to the sum of the action of A on B, $\mathbf{F}_{AB}$, plus the reaction to $\mathbf{F}_{BA}$. This is written as $\mathbf{F}_B = \mathbf{F}_{AB} - \mathbf{F}_{BA}$.

Since force is by definition a rate of change of linear momentum, the above vector field associated with a charged particle is, literally, a force field.

This is conceptually distinct from the Electric field associated with the Coulomb Law, $\mathbf{F} = q_1 q_2 \hat{\mathbf{r}}/4\pi\epsilon_0 r^2$. The electric field begins with the force between particles, which is used to define a field, $\mathbf{E}(\mathbf{r})$, that would act on a standard, 1 coulomb point charge if there was one present at the point $\mathbf{r}$. The Electric field was originally just a map of the Coulomb Law and only later came to be afforded an ontological status. In contrast, the force field here is a physically existing property of the ontology. It begins as an ontological force field from which to induce the usual relationships between observables. It will turn out not to be empirically identical to the Electric field, so let us use a different symbol, $\vec{\mathcal{F}}$.

As already discussed, the relationship between a force field and the corresponding force field energy density, $\rho_E \propto \mathcal{F}^2$, follows from basic principles: In order to increase the field strength in a given region, work must be done to move field carriers against the pre-existing field, for which the work increment is $d\rho_E \propto \mathcal{F}$ and integrating this gives $\rho_E \propto \mathcal{F}^2$. Since the space integral of the force field energy density must be finite, the force field asymptote can be no stronger than $1/r^2$ and we shall write the force field of particle A as:

$$\vec{\mathcal{F}_A}(\mathbf{r}) = \frac{KQ\,\hat{\mathbf{a}}}{a^2}\ ,$$

where K is a suitable natural constant, Q is some parameter that represents the strength of the field source, a is the distance from the source to the point $\mathbf{r}$, and $\hat{\mathbf{a}}$ is a unit vector that points radially away from the source. It will be clear from the integrals below that we can disregard the near field region of the source[3].

As for the target, the field energy density of the particle itself, as distinct from its force field, is unsuitable since we must include, for example, both protons and electrons in the model. It is tempting to use the energy density of the force field, but when we bring charged particles together, work has to be done which is associated to the energy density of the resultant force field so the

[3]The integrands don't diverge near the source mainly because $\hat{\mathbf{a}}$ changes rapidly in the vicinity of $a = 0$.

force field energy density ends up being context dependent. One could resolve the resultant force field energy density into a base energy part for each particle, which determines the interaction, and an interaction energy part that doesn't. Or one might consider using the field angular momentum density as the basis for the interaction.

We have identified no good reason for any specific choice of the target field. Therefore, in order to minimize the inevitable Metaphysical/Dynamical content, let us simply follow the classical model and refer to the target field by the name "charge density" and use the symbol ρ_q. To keep the total charge finite, the asymptotic behaviour must be no stronger than $1/r^4$.

In the near field region of the target, it will be clear from the integrals below that the result is not sensitive to the details of how ρ_q is distributed, so let us consider a simple charge density distribution with some finite, spherically symmetric near field distribution and a $1/r^4$ asymptote, in proportion with the particle's field energy density and note that this also automatically corresponds to the $1/r^2$ force field strength asymptote. This charge density profile is shown in Figure 9.1. The radius r_0 is a parameter that depends on boundary conditions to be discussed in Section 9.6. Under conditions where the particle is very well located it may be on the order of the Compton radius (although there is some experimental evidence suggesting an even smaller value).

$$\begin{aligned}\rho_q(r) &= \rho_q^{NF}(r) && r < r_0 && (9.1)\\ &= \rho_q^{FF}(r) = \rho_{q0}\left(\frac{r_0}{r}\right)^4 && r \geq r_0\,, && (9.2)\end{aligned}$$

Where $\rho_q^{NF}(r)$ and $\rho_q^{FF}(r)$ are the near field and far field charge densities respectively. The space integral of this function does not diverge either as $r \to \infty$ or $r \to 0$, so a finite total charge, q, is obtained:

$$q = q^{NF} + q^{FF} = \int_0^\infty 4\pi r^2 \rho_q(r)\, dr = \int_0^{r_0} 4\pi r^2 \rho_q^{NF}(r) dr + 4\pi\, \rho_{q0}\, r_0^3 \quad (9.3)$$

9.3 Coulomb's Law

With reference to Figure 9.2, the interaction between two charged particles at any given point in space can then be written as follows:

$$d\mathbf{F}_{AB}(\mathbf{r}) = \vec{\mathcal{F}}(\mathbf{r})\, \rho_{qB}(\mathbf{r})\, dV \;= \frac{KQ_A}{a^2}\, \rho_{qB}(\mathbf{r})\, dV\, \hat{\mathbf{a}}\,, \quad (9.4)$$

Where dV is an incremental volume located at point $\mathbf{r}$ and $d\mathbf{F}_{AB}(\mathbf{r})$ is the force of particle A on particle B in the incremental region dV located at $\mathbf{r}$. Note that a can go to zero in this expression for points near the source. Note also that $d\mathbf{F}_{AB} \neq d\mathbf{F}_{BA}$, but there is local conservation of energy-momentum due to the reaction force that acts on particle A. Global conservation will be restored following (9.10).

When the action, (9.4), is integrated over all space, components transverse to the direction from source to target, $\hat{\mathbf{R}}$, cancel due to rotational symmetry. The remaining component of the action of A on B in the incremental volume

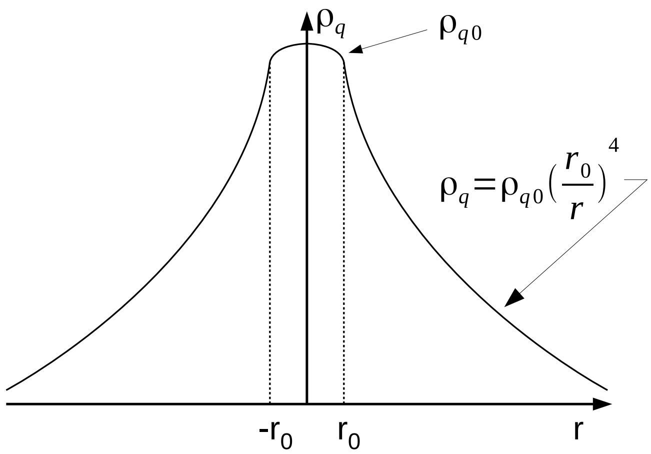

Figure 9.1: *The modelled charge density profile, (9.1) & (9.2). (Not to scale.)*

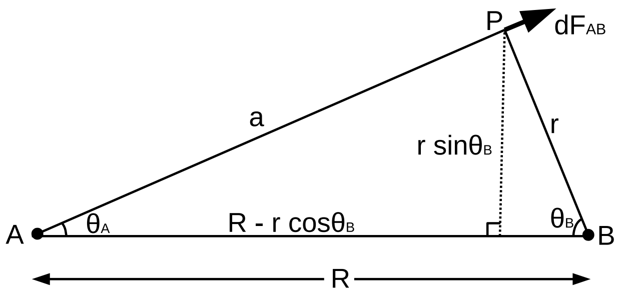

Figure 9.2: *Shows dF_{AB} in Eqn. (9.4), and the geometry and notation used in the space integral thereof, Eqn. (9.7).*

element is $d\mathbf{F}_{AB}\, cos\,\theta_A$. For the near field, or "body", of the target, $r < r_{0B}$. Assuming only that $R > r_{0A} + r_{0B}$, it is easily shown that:

$$\int_0^{r_{0B}} d\mathbf{F}_{AB} = \frac{KQ_A\, q_B^{NF}}{R^2}\,\hat{\mathbf{R}} \tag{9.5}$$

For any spherically symmetric near field charge density distribution. This is equivalent to a corresponding result in Electrostatics, but the result is also clear from the far field integral, which is also spherically symmetric with a counterpart in Electrostatics (where a spherical charge distribution, considered from points outside the distribution, is equivalent to a point charge at the centre of the distribution [91]).

For the asymptotic region, $r > r_{0B}$. Using $cos\,\theta_A = (R - rcos\,\theta_B)/a$ gives:

$$d\mathbf{F}_{AB}\, cos\,\theta_A = \frac{K'(R - rcos\,\theta_B)}{a^3 r^4} dV\,\hat{\mathbf{R}}\,, \tag{9.6}$$

where $K' = KQ_A\,\rho_{qB0}\, r_{0B}^4$. The action of the source, A, on the target, B, is given by integrating this:

$$\int_{r_{0B}}^{\infty} d\mathbf{F}_{AB} = 2\pi K' \int_{r_{0B}}^{\infty} \int_{0}^{\pi} \frac{(R - r\cos\theta_B) \sin\theta_B}{a^3 r^2} d\theta_B dr\, \hat{\mathbf{R}}\,. \tag{9.7}$$

The published article goes straight to the result of this double integral, which is:

$$\int_{r_{0B}}^{\infty} d\mathbf{F}_{AB} = \frac{K Q_A\, q_B^{FF}}{R^2} \left(1 - \frac{r_{0B}}{R}\right) \hat{\mathbf{R}} \tag{9.8}$$

This result contains a term in r_{0B}/R, which is ignored on the basis that $r_{0B} \ll R$. The goal here was to model the existing Physics without changing anything, but this term is new physics and, although it is very small, it has to be addressed.

This new term in the force law implies an adjustment in the interpretation of Gauss's Law. To understand its meaning and the complete experimental situation, it will be helpful to consider the first step in the integral above. We shall return to that below, but for the moment note that r_{0B} is small, on the order of 10^{-13} metres, so the term is essentially unobservable in experiments that directly measure the Coulomb force operating between particles as a function of the distance between them.

Ignoring the new term and adding the near and far field parts together, (9.5) + (9.8), gives the total action of A on B:

$$\mathbf{F}_{AB} \simeq \frac{K Q_A\, q_B}{R^2} \hat{\mathbf{R}}\,. \tag{9.9}$$

According to Newton's third Law, the total resultant force on the target, $\mathbf{F}_B$, is the sum of the action of A on B plus the reaction to the action of B on A, so:

$$\mathbf{F}_B = \mathbf{F}_{AB} - \mathbf{F}_{BA} = \frac{K}{R^2}(q_B Q_A + q_A Q_B)\, \hat{\mathbf{R}}\,. \tag{9.10}$$

This is actually the integral of $d\mathbf{F}_B = d\mathbf{F}_{AB} - d\mathbf{F}_{BA} = -d\mathbf{F}_A$, so the model respects global conservation laws because it respects local conservation laws in every incremental volume, as mentioned above. If the source parameter, Q, of a charged particle is proportional to the space integral, q, of its target charge density, ρ_q, which seems reasonable, then (9.10) takes the usual form of Coulomb's Law: Rather than equating the source and target parameters, let $q = kQ$ for any charged particle, then

$$\mathbf{F}_B = 2Kk\frac{Q_A Q_B}{R^2} \hat{\mathbf{R}}\,. \tag{9.11}$$

With $2Kk = 1/4\pi\epsilon_0$, this is now identical to Coulomb's Law.

That shows how a distributed action mechanism works in the static case. Let us now focus on how this all relates to action at a distance in the case with moving particles.

9.4 Action at a Distance with Moving Particles

This interaction mechanism models an instantaneous distributed transfer of energy-momentum from one field system into another and so the point model force law, (9.11), is, *prima facie*, an instant action at a distance force law. Of course, the same applies to Electrostatics, where there is no distinction between instantaneous and retarded positions. The distinction only arises in the case of moving particles.

Let us consider, to begin with, particles that are moving at constant velocity in an observer's inertial frame.

In Electromagnetics, the velocity field of the moving particle is most easily calculated directly from the Lorentz Transformation of the coordinate system, leading to an instant action at a distance relation between particles for any observer. It is also well known that a retarded time analysis happens to give the same result: When the particles are in inertial conditions of motion, the Electromagnetic interaction is calculated on an instantaneous basis.

Here, the far field is not propagating away from the particle as it is in the retarded paradigm. It comoves with the observed location of the moving particle. Consequently, the retarded time approach is deprecated because it does not represent the physical structure of the model. Every field element here is a luminal wave resonator that is Lorentz contracted and time dilated by virtue of its motion relative to the observer. The shape of the field distribution changes, and the "magnetic" forces are introduced, directly as a consequence of the Lorentz Invariance. This is all as it is in Special Relativity, as distinct from Lorentzian relativity.

Now, consider the interaction in some incremental region of space at point P and time t. Regardless of where P may be, the momenta of both the source and target particles - as calculated from a space integral over each particle's field momentum density - change instantaneously at time t. The momentum *observable* is conceptually different. Assume for example that the particle momentum is observed by measuring the velocity of the peak of the energy density distribution. It would not respond instantly to a single hypothetical pointlike interaction at some remote point P. However, if we consider a globally distributed system of pointlike interactions at the same time, t, which corresponds perfectly to a transformation between inertial states then, yes, the momentum observable would indeed have responded instantly.

The conclusion is that distributed action models can lead to instantaneous action at a distance as between notional point particles without violating either local action or the restriction that nothing propagates superluminally. Furthermore, unlike the Lorentz Force Law, the force acting between notional point particles automatically conserves momentum.

While an interaction that instantly boosts the particle from one inertial state to another might be conceivable in the case of a uniform externally applied force field, it is certainly not possible in the Two Body problem, for example the interaction between two electrons or two charged golf balls. First, the applied field is not uniform. Second, the momentum exchanged in any given incremental region includes transverse components that do not cancel locally even in the static case. Third, we know from the relation $E^2 = P^2c^2 + M^2c^4$ that $dE/dP = 0$ for an observer in the comoving frame, so there must be bremsstrahlung radiations [92].

Overall, following the interaction at a given instant of time, there is an internal process where the field resonators within each electron interact locally with each other to spread the absorbed linear momentum evenly across the system and to radiate off the excess energy that is involved. These internal self-interactions can only be communicated throughout the distributed electron system at or below c. They are retarded. Similarly, the radiated fields of the particle are, by definition, retarded. However long it may take for the system to settle into the new inertial state, that does not change the fact that the primary interaction between the particles is always instantaneous in a distributed interaction mechanism.

In the most general case of rapidly accelerating particles that are not in inertial states, the primary interaction is still an instantaneous distributed transfer of energy-momentum between the involved particle's field systems, and it is still accompanied by retarded internal self-interactions and radiations. However, at this point the very idea of a "particle position" becomes unclear in a luminal wave interpretation. Do we mean the centre of inertia or do we mean the peak of the energy density distribution? In general, they are not the same. Even if one ignores the radiation field (on the basis that it is no longer part of the particle), the relation between the comoving field of the particle and its now ill-defined "position" is not as clear as the usual Classical Theory would have it. These issues go to the question of what is "really" going on at the physical level, and it is fortunate indeed in Modern Physics that Quantum Electrodynamics has found a way not to deal with them.

As far as the Classical Two Body problem is concerned, what we now know is that each part of the system, both electrons and their respective radiation fields, obeys an applicable wave equation. To resolve the two body problem in a physical model, one has to deal with the multi-electron, multi-photon wave system as a whole.

The result of all this is that the 2-body problem is not a problem in the domain of Classical Physics. The question of the dynamic interaction between two charged point particles involves an essential misstatement of fact: They are not point particles. They are distributed wave systems and that is an indispensible part of the problem.

Where this whole question of distributed versus retarded interaction has the greatest observable impact is, of course, with spacelike causal correlations. When it comes to spin measurements, there are no messy internal transients to muddy the water. The resonators of the measured particle align or counteralign with the measurement field. There may or may not be some minor, short-range internal spin-flipping at the resonator level (in the sense of the Ising model [68]) but these details are not pertinent. The spin measurements in this model are well described by the nonlocal collapse of the spin part of the system wavefunction in Quantum Mechanics. This phenomenon now has a clear, local realist, physical interpretation as a distributed interaction in the luminal wave model.

9.5 Gauss's Law and the Cavendish Experiments

In the above model with the $1/r^2$ force field, $\vec{\mathcal{F}}$, the action of the source, A on the target, B, in the asymptotic region was found to be:

$$F_{AB}^{FF} = \frac{KQ_A\, q_B^{FF}}{R^2}\left(1 - \frac{r_{0B}}{R}\right)\hat{\mathbf{R}} \tag{9.12}$$

To understand where the r_{0B}/R term comes from, consider the intermediate step, (where $a = \sqrt{R^2 + r^2 - 2R\,r\,cos\,\theta_B}$ is the distance from the source to point

$P)$:

$$\int_0^\pi \frac{(R - r cos\,\theta_B) sin\,\theta_B}{a^3 r^2} d\theta_B = \frac{1}{r^2}\left[\frac{r - R\,cos\theta_B}{aR^2}\right]_0^\pi = \frac{1}{R^2 r^2}\left[1 - \frac{r-R}{\sqrt{((\pm(r-R))^2}}\right]$$

Since a is positive real, the square root term is always positive for all r. For $r < R$ the total result at this step is $2/r^2R^2$, but for $r > R$ it vanishes[4]. Substituting this back into the above, the complete far field integral is now, as above:

$$\int_{r_{0B}}^{\infty} d\mathbf{F}_{AB} = \frac{4\pi K'}{R^2}\int_{r_{0B}}^{R}\frac{1}{r^2}dr\,\hat{\mathbf{R}} = \frac{KQ_A\,q_B^{FF}}{R^2}\left(1 - \frac{r_{0B}}{R}\right)\hat{\mathbf{R}}\,, \tag{9.13}$$

where the space integral has been truncated at the radius $r = R$. The term in r_{0B}/R corresponds to part of the target charge being excluded from the integral. When the reaction forces are included to obtain the full expression for the force law between two electrons there is also a similar term in r_{0A}/R.

These terms *can* both be removed by "tweaking" the charge density distribution, for example with an exponential form for the asymptotic field distribution, or one could use a model where the near field charge far exceeds the far field charge. Not only would that be artificial, but it may go on to interfere with the physical model of gravity in Chapter 11. Instead, let us reinterpret Gauss's Law. It applies perfectly to the ontological force field that underlies the Coulomb Law in this model, but with the observable based "Electric" field there is a minor approximation involved.

Gauss's Law is, of course, a stalwart of the Classical Theory, but the fact is that it does not apply perfectly in Modern Physics. There are a few other modifications at very close range, the best known of which is a vacuum polarisation effect that leads to charge masking, so that the bare charge is slightly greater than the charge observed from a distance. The effect here is somewhat similar, only operating to reduce the observed charge as one approaches it. The quantum mechanical effects lie well beyond the scope of the foundations, but they make it clear that the charge observable is not the perfectly conserved quantity it was taken to be in the Classical Theory.

The question of interest is whether this departure from the usual Coulomb Law can be detected experimentally when the far field charge, q^{FF}, is comparable to the near field charge, q^{NF}. With the value of r_0 on the order of the Compton radius of a charged particle, these terms are far below the threshold of even the most precise experiments that directly measure the force between macroscopically separated charges.

On the other hand, there are null experiments, beginning with the Cavendish experiment, designed to test whether the Coulomb Law is exactly a $1/r^2$ law or not. These experiments are amazingly precise: Modern versions can identify departures of one part in 10^{16} or less in the Coulomb Law exponent [93].

However, this kind of experiment does not test for the r_0/R terms above.

Consider a pointlike charge located at the centre of a metal sphere of radius R. Outside of the sphere, the Electric field has the same value as the field of the

[4]It is also notable here that the integral does not diverge at the source, so there's no need to switch to a near field model at that point.

free charge, but it is not the Electric field of the charge that is inside the sphere. It is the Electric field of the induced charge on the outside surface of the sphere. All of the force field of the charge inside the sphere is now inside the sphere, so space integrals from $r = 0$ to $r = R$ include all of the force field of the particle. For this particular charge in this particular situation, the values of Q and q may be very slightly different from their free space values, but the applicable force Law that governs interactions with surface charges on the interior surface of the sphere is exactly a $1/r^2$ law.

The Cavendish experiments involve spherically symmetric charge distributions on spheres nested concentrically inside other spheres. So consider instead a spherically distributed classical charge on the surface of an interior metal sphere, inside the metal sphere of radius R above. Again, the entire force field of this classical surface charge is in the region between the two nested metal spheres, and the relevant force Law in the model above is, again, exactly $1/R^2$. A particle's charge may have a different value from what it would be in the absence of the exterior sphere, but the Cavendish experiments only measure the potential difference between the spheres. They only measure whether the forces cancel, not their absolute values.

There is really no coincidence here because the Cavendish experiment is based on an integral in Electrostatics that establishes the constant value of the Electric potential inside a metal sphere, which is essentially the same integral as the intermediate step discussed above.

There may or may not be a distinction between the model here and Electrostatics with non-concentric spheres, but in this case there are well known second order effects when the classical charge distributions lack spherical symmetry, which are found in the usual Electrostatics theory, [91]. When the charge distributions are asymmetrical, the potential inside the metal sphere is not perfectly constant and the null experiment loses sensitivity.

9.6 The Parameter r_0

The sharp separation between near and far field regions is quite artificial but the asymptotic field cannot extend to $r = 0$, so some kind of transition between a near field region and an asymptotic region is necessary in a model of notional point particles.

For maximum simplicity, it could be asserted that the size of the near field region, r_0, is a natural constant. The proton and electron near fields would then be identical, despite the different masses and different Compton radii, which is not reasonable. Moreover, such a physical model would not be compatible with the experimental evidence for de Broglie waves obtained in matter beam interferometry experiments. r_0 must be a parameter determined by the context (*i.e.* the boundary conditions) not a constant, and one should expect it to be different both in different contexts and in different species of particle. Broadly speaking, the more the charged quantum interacts, the smaller r_0 becomes.

As the plateau width parameter, r_0, varies, the field distribution in this *ad hoc* model is constrained as follows. From (9.3) and the assumption $q = kQ$:

$$q = q^{NF} + q^{FF} = q^{NF} + 4\pi\,\rho_{q0}\,r_0^3 = kQ\ , \tag{9.14}$$

Because the force field source, $\vec{\mathcal{F}}(\mathbf{r})$, is proportional to $\sqrt{\rho_E(r)}$ it is also proportional to $\sqrt{\rho_q(r)}$, so $\rho_q(r)$ cannot depend on r_0 for $r > r_0$. Otherwise, the asymptotic force field strength would depend on r_0. Equally, the force between well separated particles does not depend on r_0, so again, $\rho_q(r)$ - this time as target - cannot depend on r_0 for $r > r_0$. Consequently, as shown in Figure 9.3, $\rho_{q0}\ r_0^4 = D$, where D is a constant, independent of r_0.

The far field term in the preceding equation can then be written as $q^{FF} = 4\pi D/r_0$ and the condition on the near field charge distribution is:

$$q^{NF} = \int_0^{r_0} 4\pi r^2 \rho_q^{NF}(r)dr = kQ - \frac{4\pi D}{r_0}\ .$$

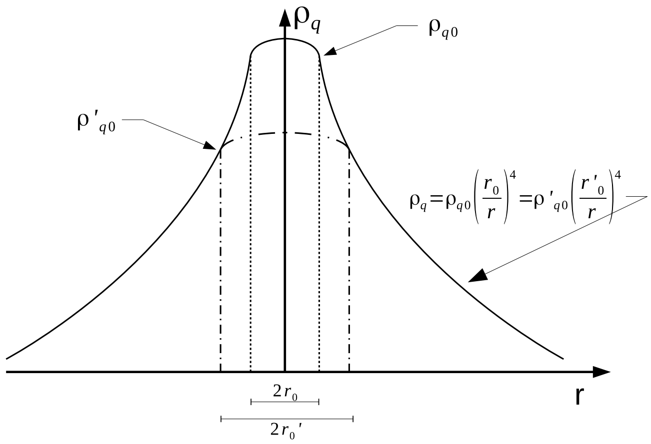

Figure 9.3: *As the plateau width parameter changes from r_0 to r_0', the asymptotic field, $\rho_q(r)$, is invariant, so $\rho_0\ (r_0)^4 = \rho_0'\ (r_0')^4 = D$, where D is a constant.*

When this condition is satisfied, and the target and source near fields do not overlap, the Coulomb force Law is reproduced for whatever value of r_0 results from the boundary conditions. For electrons in an electron beam, r_0 could be measured in microns. For Cosmic radiation, it could be measured in kilometres and for well-located particles it could be on the order of the Compton radius.

This highly artificial, strange looking condition, the difference between one constant and another divided by r_0, is the direct result of artificially dividing the field system into near and far field regions, as one must in the absence of a satisfactory field solution for the field system as a whole. In this case, it is readily associated with introducing the idea of notional point particles.

We are about to encounter a very similar form that arises in the Schwarzchild solution to the Einstein Field Equations. The strangeness of the Schwarzchild formulae will be traced to essentially the same issue: The usual source term in Gravity is either a point particle or it is absolutely confined to the region of space where we see the planet. This enforces an absolute discontinuity on the continuously differentiable curvature field on the left hand side of the equation. However in a pure field theory (and Einsteinian Gravity is considered to be a pure field theory), a source term should never be absolutely confined inside a well-defined region.

Chapter 10

Prelude to Gravity

10.1 Introduction

I became involved with gravity by mistake and succeeded with it by serendipity.

The mistake was to mention to a physicist friend of mine that, if I were right about Special Relativity, then Gravity should be easy (my real meaning being "It should be easy... for you."). Without the slightest hesitation, the response came straight back: "Ah, you should show that." I was hoist by my own petard.

However, my assertion that it would be easy turned out to be correct, not because it's inherently easy, but because the work had already been done. Most of the mathematics necessary to implement an idea I'd long had in mind was already in the literature.

The idea is as follows. Energy propagates in a physical medium. Since nothing is perfect, the medium cannot have an infinite capacity to support energy waves (Section 2.3, criterion 7). When we put a lot of energy, like a planet or a star, into some region of space, the properties of the medium should be expected to change. What are the properties of the medium? The only property that has been proposed is the characteristic velocity, so the model in previous Chapters immediately places one in the domain of refractive medium interpretations of gravity.

One of the most basic ideas on which the model here runs was also in the literature: The (locally measured) velocity of light is constant for all observers, and gravity had already been derived from it using the Lagrangian method.

In order to construct the physical model, we are thus looking, right from the outset, for two quite specific things.

1. A set of relationships describing the impact of variations in the characteristic velocity on the rates of clocks and the lengths of rulers, these devices being inseparable from the idea of a "metric".

2. A relationship between the energy density in the medium and variations in the characteristic velocity, specifically a relationship with a distributed source term (the energy-momentum density) that varies as $1/r^4$ in the asymptotic region.

The energy present in a region of space would then change the rules of geometry. To put that sentence in language that all readers can understand, consider a right angle triangle where the ruler used to measure the hypotonuse is shorter than the ones used to measure the other two sides. Pythagoras Theorem would not work. The sines and cosines would all be wrong, and if one calculated the internal angles from the lengths and the sines and cosines, they would not sum to 180 degrees.

The appropriate mathematics for this context is Riemannian (or curved space) geometry rather than the usual Euclidean (or flat space) geometry and (as is well known) a refractive medium is a physical instantiation of a mathematical curved space. The third part of all curved space approaches to gravity, namely the d'Alembert equation of motion, is taken as given here, because it is as Theory independent as Newton's first Law: bodies move along geodesic paths in the 4-space of the coordinate system.

I was well aware that there were several refractive medium approaches to gravity to be found in the literature but there was no good reason to anticipate that the specific physics needed to fit Special Relativity, Electromagnetics and General Relativity together in my specific model already existed, albeit in parts scattered across several articles by different authors, and expressed in mathematics that conceals the physical interpretation: That's more than just serendipitous. We are going to see three different, independent ways to approach the same mathematics, which is pertinent because there are often several ways to approach good results in Physics.

Riemannian geometry is, of course, also the basic platform for Einstein's General Theory of Relativity. Before embarking on my own work, I first learned Riemannian geometry and studied the General Theory. The first of these two exercises was entirely satisfactory.

On the other hand, I had assumed that the General Theory itself would be every bit as correct, final and inviolable as the Special Theory. A majority of Physicists seem to see it that way. Unfortunately, upon arriving at the Schwarzchild solution, the line element was an immediate cause for concern from the point of view of designing a physical model for the General Theory.

It contains terms in $1 - r_s/r$ and its inverse $r/(r - r_s)$, where $r_s = 2GM/c^2$ is the "Schwarzchild" radius. This is written in "spherically symmetric" coordinates where the "radius" of a sphere is defined from its circumference, $r = C/2\pi$, (because the rulers on any given spherical surface all have the same physical length in this coordinate system[1]).

For comparison, in "isotropic spherical" coordinates, the abovementioned terms in the Schwarzchild radius are replaced using:

$$1 - \frac{2GM}{c^2 r} = \left[1 - \frac{GM}{2r_1 c^2}\right]^2 \Big/ \left[1 + \frac{GM}{2r_1 c^2}\right]^2 ,$$

where r is the radial coordinate in the spherically symmetric scheme and r_1 is the radial coordinate in the isotropic spherical scheme.

[1]It is noteworthy that if we measure the radius of the sphere with spherically symmetric rulers stretching from the centre of the sphere to its surface, then we get a different value (because the length of a radial ruler depends on its position and the radial rulers are not the same length as the tangential rulers). This rough characterisation of the relation between the math symbols and physical rulers in spherically symmetric coordinates is not simply a consequence of the physical effect of gravity on rulers. It is a mixture of the choice of a coordinate scheme (also known as a chart) and the physical effects of gravity.

The distinction between Special Relativity and General Relativity *vis a vis* coordinate systems and coordinate independence, which is to say the distinction between general covariance and Lorentz covariance, needs to be clearly understood. This will be the first preliminary topic below.

However, my concern with these terms was about the underlying physics that they represent. From the point of view of modelling the Theory, they are mathematically ugly terms. I knew immediately that one could never specify a credible physical mechanism that would lead to such terms. We need to understand exactly where these terms come from, and why they are so ugly. This will be the second preliminary topic to discuss.

To confirm that view, I next went looking for refractive medium interpretations of the General Theory, with the result that, yes, other authors (eg: [94]) had invariably found themselves putting in by hand an ungainly, artificial functional relationship between the mass distribution and the refractive index that could never correspond to any credible physical mechanism. For me, this was now more than just a concern: if professionals could not reproduce those ugly terms, I was not inclined to waste time on it.

It was necessary to consider whether a physical model of Einstein's Theory was an appropriate objective. There are quite a few nagging little doubts about the General Theory that have been raised over the years, not by me but from within the gravity community.

First, the General Theory is a test particle theory. This is a valid concern about the Theory. It means that it is unable to formulate (never mind solve) any actual problems involving pointlike masses - planets and stars - in closed form. The field of computational relativity emerged and in practice we use numerical methods to make quantitative predictions. That is also the case in many fields, not just gravity, but the problem is that, even in principle, Einsteinian gravity cannot formulate any non-trivial problems in its own domain.

Second, General Relativity is (for the same reason) often referred to by specialists as an "effective theory". Now, the common usage of the words "effective theory" in Physics is to denote a theory that is conceptually compromised in some way. (In the standard model, for example, renormalisation is an "effective theory", one that gives good answers despite us having done "bad things" in the Mathematics.)

Third, the General Theory contains singularities - physical conditions where the Theory breaks down, and a galaxy can be compressed into a mathematical point.

Fourth, despite Herculean efforts over many decades, there is no acceptable unification with Quantum Fields.

Fifth, the local observed speed of light is not always identically equal to c. That's not in the spirit of relativity.

Sixth, the General Theory does not reduce to the Special Theory in the weak field limit [2].

Seventh, As discussed in Sections 10.4 and 11.1, the field equations do not properly satisfy important 4-space mathematical identities, namely the Bianchi and Freud identities, in N-body problems.

Not only was changing the objective a matter of pragmatism, it was justified by the totality of the circumstances, not just my own aesthetic difficulties with certain terms in the Schwarzchild solution.

The question now was what to model. An obvious (wrong) choice would be to focus on reproducing the observables. I would not then be making a physical model of a Theory, I would be making a Physics Theory. That's not my interest. Building new physics theories is the work product of professional physicists. My work is to design physical models "*After* Physics": It needs to be done under the umbrella of pre-existing Theory.

In all my experience it is easier to make a physical model for a beautiful theory than an effective one. Whereas one does not find empirically distinct alternatives to the Special Theory or Dirac's Equation for the electron in the literature, in gravity physicists have developed a plethora of plausible alternatives to the Einstein Theory. There is even a "Parameterised Post Newtonian" theory of metrics [95] with the effect that the number of possible curved space theories that fit all the observed data is infinite.

There was no question in my mind about Einstein's line of reasoning and there can be no question about the algebra. The only place where a reasonable doubt exists in the development of the usual Theory is in the choice of the Einstein Field Equations, specifically with respect to the choice of the source term on the right hand side. These equations were chosen to be as short and simple as possible. To be fair, one might say that they are somewhat motivated by physical and formal considerations, but they are certainly never derived.

I needed to find a professionally respectable, existing alternative curved space theory of gravity with field equations that are more conducive to physical models in general, and mine in particular.

My literature searches on refractive medium interpretations had uncovered a subset of the plethora of alternative theories of gravity, known collectively as the exponential metric theories. Amongst this group of theories, one stands out from the crowd, although it is not, at first blush, a refractive medium interpretation: the Yilmaz Theory of Gravity [4, 5, 96], developed by the Turkish physicist Huseyin Yilmaz, originally while taking his Phd at MIT in the 1950s. His theory is also known as the Einstein-Yilmaz variation of General Relativity, like some alternate move in a well known chess opening. The gravity group at Maryland University, USA, focusses on Yilmaz theory. Debate over the Yilmaz theory continues within the community, and we shall touch later on a couple of those discussions.

Yilmaz's theory has nice mathematics from the model design perspective. The local speed of light measures the same for all observers, and one of the La Grangian derivations Yilmaz gives is based on that principle. Corresponding to the above terms from the Schwarzchild solution, the Yilmaz theory has e^{-2GM/rc^2} and its inverse e^{2GM/rc^2}. (This is in an isotropic coordinate scheme, using the same rulers in all directions.) Of special interest, in the N-body case, the metric is determined by the product of the N 1-body metrics, so one just adds the exponents from the N 1-body expressions together.

In essence, the difference between the Yilmaz and Einstein Theories is a new source term on the right hand side of the field equations. The third and final preliminary topic discussed below is the rather obvious physical and formal motivations for the new term and why it turns Schwarzchild's ugly duckling into a swan.

From my early perspective, the exponential terms in Yilmaz's covariant expressions looked promising because it is very easy to generate exponential functions with physical mechanisms. There was every prospect for a physical model, and I soon enough found that one such model already existed, namely Hal Puthoff's polarisable vacuum model [97]. It had been developed along a different strand in the literature. Rather than coming from the Riemannian geometry, it is based on prior work by RH Dicke [98] using a standard Lagrangian method in a fixed, underlying coordinate system.

Puthoff also gets his results from a Lagrangian, but in my experience, applying the formal procedure frequently does not provide results in an intuitively meaningful form. I soon enough found that the most pivotal differential equation that came out of the Dicke/Puthoff Lagrangian is actually a member of a family of empirically equivalent equations in different powers of the refractive index. From the point of view of a physical model one of those equations could be rewritten in a very appealing way:

$$\nabla^2 \frac{1}{\epsilon(\mathbf{r})} = \frac{4A^2}{\epsilon r^4}$$

or equivalently,

$$\nabla^2 c(\mathbf{r}) = \kappa \rho_E(\mathbf{r}) \tag{10.1}$$

Where ϵ is the dielectric constant (which equals the refractive index in what follows here), $2A$ corresponds to $2GM/c^2$, κ is a natural constant, $c(\mathbf{r})$ is the (position dependent) velocity of light (*i.e.* in the unit system of observers outside the gravitational field - see Section 10.2) and ρ_E is the energy density in the refractive medium. Huseyin and Hal had done all the hard work for me, although neither of them had recognised that gravity could be written in such a simple (dare I say neat?) form.

The main content in Chapter 11 covers the extension of the resonator physical model from Chapter 6 to include exponential metric gravity. The outrageous claim at the end of the last paragraph is then discussed and two specific mismatches between Puthoff's analysis and Yilmaz's will be identified in order to explain why Puthoff's theory, in its present form, does not work in strong fields while Yilmaz's, in its present form, cannot be unified with Electromagnetics (and hence relativistic quantum mechanics). If there is one desired outcome from this book, it would be that the professionals involved with each of these theories would attend to these issues.

10.2 Coordinate Independence

The events that occur in the world do not depend on the coordinate systems we use to specify their locations in space and in time[2].

That's a fact, but it does not strictly follow from it that the Physical Laws *must* therefore be coordinate independent. Indeed Selleri gave the Laws of Mechanics in an empirically equivalent, coordinate dependent form. Selleri's version of the Laws was rejected in Chapter 4, not because it is wrong, but because it is complicated. On the other hand, it works better than the coordinate

[2] Recall that we have already seen that spacetime geomtery is mathematically justified by luminal wave models, but the usual interpretation of scientific realists, that space and time are a single entity, has been replaced here by the notion of a combined spatiotemporal evolution.

independent form on a rotating platform and in the GPS system, where the accelerating reference frames still have constant speed.

Although it cannot be proved, let us take it as a valid inference that Physics should be written in a coordinate independent form (Section 2.3, Criterion 4).

Einstein's step from Special Relativity to General Relativity began with the intention to address gravity on the basis of the (weak) equivalence principle, which is the proposition that an observer cannot distinguish between being in a gravitational field and being in an accelerated reference frame. The motivating idea then was that the Laws of Physics should have the same form for all observers, including especially observers in accelerated reference frames. A branch of Mathematics that was useful to formulate the problem is the mathematics of curved spaces, but note that in the originating mathematics these spaces don't necessarily correspond to anything physical.

The mathematics of curved spaces has a much broader scope than just handling accelerated reference frames. Problems are formulated in "generalised coordinates", which is to say that the four numbers used to specify the location of a point in a 4-space do not necessarily correspond to measurements made using three rulers and a clock, or one ruler, two protractors and a clock. Instead of labelling the coordinates as x, y, z, t they are usually labelled as $0, 1, 2, 3$. With respect to an ostensibly physical 4-space, as long as every point is uniquely identified, nothing prevents any given coordinate from comprising, for example, a mixture of time, angle and distance.

The central idea is to write equations that hold regardless of how the coordinates are chosen. This mathematical property is called general covariance. The coordinates that appear in these equations clearly do not have any specific *a priori* physical interpretation unless and until a particular coordinate scheme (a "chart") has been chosen, and even then the physical interpretation may be non-trivial.

This whole context is qualitatively different from Special Relativity where the coordinates (usually)[3] have a direct physical interpretation in the sense that the t coordinate is measured with a clock, the x-coordinate is measured with a ruler and we use the same ruler to measure the y and z coordinates. Generalised coordinates involve a level of abstraction that is, in my opinion, well motivated and well justified by the first sentence in this Section, but when Einstein chose this path to develop his theory of gravity, it was more of a superhighway.

If one is doing Physics, then the superhighway is the road to take. However, we are not doing Physics, we are constructing a physical model of a theory in Physics and the abstraction inherent in manifest general covariance precludes visualisability. Moreover, we did not construct a physical model of Lorentz covariance by adopting a covariant formalism. The idea was to show how the Lorentz covariant formalism arises in an objective, visualisable, intelligible setting that fits everyday intuition. In the process we resolved all the mysteries and paradoxes that Minkowski's mathematical abstraction had thrown up[4].

It is a basic premise in the construction of an acceptable physical model that it should use the objective concepts of distance and time that we have in our minds when we visualise the world. The distance between two points is then an objective fact regardless of how a physicist measures it. The time between

[3]Nothing prevents using generalised coordinates in a flat space, but it is not usually helpful.

[4]I understand that Einstein's initial reaction to Minkowski's 4-space was negative.

two point events is an objective fact that has nothing to do either with how a physicist measures it or the objective distance between the events. The essential purpose of a physical model is to connect that intuitable framework with the Physics. The task here is to take the reader from his or her ordinary notions about space and time to generally covariant Physical Laws without sacrificing any of those common sense notions along the way.

The ideal would be to adopt the perspective that we had as the superobservers of Silicon world, but we are not outside of this world and our common sense notions of separation and interval as objective facts do require an operational implementation. The basic intuitions that we have can also be understood from the perspective of a certain class of real observers. These observers are located in a region of space that is far away from any stars, planets or other massive objects. Their instruments, (rulers, clocks and protractors) are too light to have any impact on the space that they are in. Finally, they are at rest in the CMBR preferred frame[5] identified in Chapter 4.

We are interested in their description, in their system of units, of gravitational phenomena far away in the Universe, where they can see the stars, the planets and the galaxies. If it can be shown that the equations these observers get from the model in the preceding Chapters are isomorphic to the equations in General Relativity, in the sense of giving mathematically identical results, then the model has succeeded.

Now, the principle object of study in General Relativity (Einstein or Yilmaz) is the metric tensor. It is to be thought of as a 4x4 matrix, with up to 16 independent components in the most general case. It is a mathematical object that describes a mathematical space with no *a priori* physical interpretation. In some specific cases in the Einstein Theory (and in a wider range of cases in the Yilmaz theory), there exist choices of the coordinate system that diagonalise the metric tensor. In such cases the invariant line element has a somewhat direct physical interpretation such that the physical effect of gravity on clocks and rulers corresponds to the square root of the time or space component in the line element. For example, with respect to the Schwarzchild solution, the standard relationship for the gravitational time dilation of clocks is:

$$t_r = t_\infty \sqrt{1 - \frac{r_s}{r}}$$

Where $r_s = 2GM/c^2$ is the Schwarzchild radius, r is the radial coordinate in Schwarzchild coordinates, t_r is the time elapsed on a clock located at r from the gravitational source and t_∞ is the time elapsed on our observers' clocks, far outside the field in the limit as $r \to \infty$. Note that time stops at $r = r_s$. The line element in this coordinate scheme is:

$$ds^2 = -c^2\Big(1 - \frac{r_s}{r}\Big)dt^2 + \Big(1 - \frac{r_s}{r}\Big)^{-1} dr^2 + r^2(d\theta^2 + sin^2\theta\, d\phi^2)$$

Local Lorentz Invariance is included (via the minus sign in the first term) but this line element is qualitatively different in an important respect. Although they are being expressed within the context of a certain coordinate scheme, these terms are describing objective facts in the Theory, not relative perceptions.

[5] *i.e.* the dipole component of the Cosmic Microwave Background Radiation momentum density vanishes for these observers.

Observers inside and outside of the field can look at a clock in a gravitational field and compare it with a clock in free space. Comoving observers at different places in a gravitational field do not disagree about the rates of their clocks and the lengths of their rulers. Everyone agrees that the clocks and rulers of observers at infinity are undistorted by gravity.

Corresponding to this objective status, the invariance of the line element is preserved in this coordinate scheme with a space term mulitplying dr^2 that is the inverse of the time term multiplying dt^2. Compare that with the Lorentz Transformations, where both the space coordinate and the time coordinate are multiplied by the same factor γ, and the line element is preserved, in spite of this, because of simultaneity differences between the frames that are introduced by the Einstein clock synchronisation protocol.

It has been emphasised in the above that the passage from the metric tensor to the line element to the physical impacts on clocks and rulers is highly conditional. However, we are going to be arguing the other way, from objective facts about clocks and rulers to the line element and then finally to the inherent general covariance of the model. There are no such conditionalities when making the case in this direction (provided the target theory is manifestly covariant and the model is isomorphic to it).

10.3 What's Wrong with $1 - r_s/r$?

The problem with these terms in $1 - r_s/r$ is that the kinds of physical processes that one wants use to build a physical model generally produce terms that are simple functions of the radius, not the difference between a constant and a function of the radius. The reader's attention was drawn to a similar term that arose when we artificially imposed a separation between the near and far field regions in the distributed action model of the Coulomb Law. The issue here is of a somewhat similar nature.

On the left hand side of the Einstein Field Equations there are terms related to the curvature of the space. On the right hand side there is the source term related to the matter distribution (more generally the energy-momentum distribution). In the Static Spherically Symmetric (SSS) case, or any N-body problem for that matter, the Einstein source term is treated as a point source, and the right hand side of the field equations vanishes in the region of most interest, which is the space around the planet. The left hand side must therefore also vanish and the whole system reduces to the deceptively simple "vacuum field equations":

$$R_{\alpha\beta} = 0$$

Where $R_{\alpha\beta}$ is the Ricci tensor. This is actually 16 homogeneous equations each of which has several terms in various multi-index, non-tensor, Christoffel symbols and their partial derivatives, each of which contains multiple terms that are nontrivial functions of the 16 elements of the metric tensor. Even the simplest problem in the Theory is kind of complicated, but the issue here is not at all complicated.

Here is a rather nice quote from Einstein on a different day:

> *"But the division into matter and field is, after the recognition of the equivalence of mass and energy, something artificial and not clearly defined. Could we not reject the concept of matter and build a pure field physics? What impresses our senses as matter is really a great concentration of energy into a comparatively small space. We could regard matter as the regions in space where the field is extremely strong. In this way a new philosophical background could be created."* — Einstein & Infeld [99].

I couldn't agree more. In fact, this is the quotation I used to introduce the article on Special Relativity [21]. The gravity field equations are indeed a pure field Theory, but in the SSS problem the source term is absolutely confined to a sharply defined region, be it an infinitessimal point or be it a planet. That does not happen in pure field theories. The source term has been put in by hand in the same, traditional way that point(like) particles were put in by hand in Electrostatics. On the left hand side there is a field - which really is a field - which has to obey differential equations in functions that are continuously differentiable. The absence of any source term in the space outside the planet imposes an absolute discontinuity on a continuous field at the space / planet boundary, resulting in terms that say: 1 minus a departure from oneness that depends on how far you are from the discontinuity.

Although one may indeed think of matter as the regions in space where the field is extremely strong, in Mathematics that does not justify excluding from the right hand side of the field equations those regions in space where the source field is rather weak. In particular, in our physical model we obviously need an energy density on the right hand side of the SSS problem that varies as $1/r^4$ in the asymptotic region, which is to say that our source term cannot vanish in the space around the planet.

10.4 The Yilmaz Theory of Gravity

The critical difference between the Yilmaz Theory [5, 96] and the Einstein Theory is an extra source term on the right hand side which does not vanish in the space outside the planet. In addition to the tensor that corresponds to the matter distribution - the pointlike sources - this new tensor term represents a "field stress energy" in the space surrounding the planet. Continuity is restored and a line element emerges in simple functions of the radius.

There are two main criticisms of the Yilmaz Theory that have been raised by advocates of the orthodox set of field equations.

First, this new tensor is considered to be "not well-defined". This kind of broad brush technical critique often reflects a lack of understanding of the details. The covariant mathematics is complicated and the professionals often find themselves arguing over abstruse details like 15^{th} century philosophers debating the number of angels that will fit on a pinhead. It was sufficiently well-defined for Yilmaz to get answers.

I certainly do not understand the Yilmaz theory well enough to presume to respond on his behalf, but what I can say for sure is that it will be well-defined in what follows here. What is not so well-defined is, as always, the part inside those pointlike particles. Fortunately, the Yilmaz theory also renders that problem moot. It is able to take the limit as $r \to 0$ and associate a finite mass to a point, and the model here benefits in the same way in Section 11.3.

The other criticism [100] was raised by Charles Misner, in an article entitled "Yilmaz cancels Newton". Misner is famous for being one of the authors of the bible on gravity, "Gravitation", usually referred to by the authors' names, Misner, Thorne and Wheeler. The essence of the criticism is:

> *"For matter described as a perfect fluid, and with Yilmaz's choice of signs when introducing these quadratic terms, we find that the Euler hydrodynamic equation in the Newtonian limit is modified to remove all gravitational forces."* [100].

The refutation, [101], is comprehensive. It shows most clearly that Misner was operating under a total misunderstanding of the theory. He was applying assumptions from one context, where he is a leading expert, to a different context, where he is not. Nonetheless, if you look this up at Wikipedia the focus is on the criticism of the famous gravity expert rather than the refutation by experts in the Yilmaz theory.

As a necessary part of doing some quality assurance on Yilmaz theory, I sat in on a discussion between experts on both sides of the fence. Would the orthodox expertise expose any flaws in Yilmaz? Over the course of several months I saw a sequence of detailed formal arguments set out by the Yilmaz expert, with no substantive responses from the Einstein expert, only handwaving and repeated appeals to authority (especially Misner, Thorne and Wheeler [95]).

Much of this handwaving took the form of repeated references to the phrase *"One wing of marble, the other of wood"*, arguing that the two sides of the Einstein Field Equations are very different, but both were solid, and you can't mess with it. I didn't know the origins of this famous remark until I read [101], which contains a paragraph that makes the true meaning of Einstein's marble and wood analogy clear:

> *"A most important feature of the new theory is that the Bianchi identity is satisfied both by the left-hand side and the right-hand side of the field equations as an identity, whereas in general relativity the left-hand side satisfies the Bianchi identity identically but the right-hand side does not. We are told that Einstein himself was aware of this and that is why he many times said "My equation is like a house with two wings; the left-hand side is made of fine marble, but the right-hand side is perishable wood". It is said that it was his "dream" to find a right-hand side that also satisfies the Bianchi identity, but this was judged to be too difficult or impossible, and it was given up. Instead, the divergence of the right-hand side is forced to zero. But then this becomes a constraint on matter or on the field (or both), making the theory mathematically overdetermined."* — Alley, Ascham and Yilmaz [101].

In other words, the "marble and wood" remark goes to a weakness, an imperfection in the field equations, and Einstein was very much aware of it. Although he saw it as a math problem rather than in terms of the physical discontinuity that a spatially confined source term imposes on a continuously differentiable field, this is why the Einstein Field Equations break the curved space identities.

Now, the experimental facts on gravity are in two categories. There is high precision evidence from observations on the solar system, but the gravitational fields here are very weak: the departure from flat space on the surface of the sun is no more than 1 part in 10^7. These are nonlinear theories, and when we compare the orthodox Theory with the Yilmaz variation, the leading terms in the Taylor expansion are identical, while the next term is unobservable under weak field conditions. To distinguish between the theories, one needs precise measurements under strong field conditions. Unfortunately for Physics (but fortunately for the human race), all the strong gravitational fields in the Universe are hundreds of light years away from us, and there are essentially no precision measurements possible.

There is a partial exception with binary pulsars. We can measure the decay rate (corresponding to gravitational radiation of energy from the system) with great precision, but neither the masses nor the eccentricity of the orbit can be known with certainty, independent of the Theory. The General Theory provides a reasonable, but not especially convincing (at the 1 percent level of accuracy), agreement with observation in this case.

To distinguish experimentally between Einstein's Field Equations and Yilmaz's would require a much larger sample size of suitable binary pulsars than is presently available or the development of new techniques for obtaining precise observations on phenomena that are (at least) hundreds of light years away from us. For the time being, these two theories are experimentally indistinguishable.

Another important consequence of the Yilmaz Theory which should be mentioned in passing is that there are no singularities, which is to say no black holes *per se*. The routinely identified celestial objects that we regard as black holes are merely very dark grey in Yilmaz: Time slows down and approaches zero as the mass increases, but it never reaches zero. The mass density can increase without limit but, as that happens, its efficacy in causing gravitational effects is reduced in strong fields. The solutions don't runaway. Entire galaxies do not get compressed into a single mathematical point. This is a good thing because all the observational evidence can be explained with a theory that does not break down or produce infinities. We shall see in Chapter 11 that the same applies in the physical model.

10.5 Why Does an Apple Fall?

The geodesic equation of motion is common to all curved space approaches, and will be taken as given. Bodies follow 4-geodesics. In the case of an apple falling from a tree, the space part of the geodesic describes the physical path it takes as it falls, and the time part describes the evolution of the time coordinate along the path. Progress along the path is determined by the d'Alembertian equation of motion. Nothing could be simpler, but why does the apple fall? To say that the apple falls because the space is curved merely refers the question back to an abstract formalism.

Refractive medium interpretations in general are intended to provide an intuitive explanation, but in most cases this only goes as far as light. We are comfortable with the idea that light is refracted when it encounters a change in the refractive index, which is to say a change in the speed of light. The physics is very simple and will be familiar to all readers from high school.

If we now say that the effect of gravity is to increase the refractive index in the region surrounding the gravitational source then one can immediately understand why light bends as it passes close to the sun, for example. One now needs to provide a credible explanation for why the refractive index increases, which the model will provide, but bending light does not properly explain why the apple falls.

In this particular refractive medium model, the explanation for massive bodies is sufficiently clear. The apple is composed of luminal wave energy that propagates at c. In a refractive medium, the speed on any given trajectory depends on its distance from the centre of the earth. Those parts of the energy-momentum distribution that are further from earth travel further in a given interval of time than the parts closer to earth.

The time has come for a final trip to the beach. You may have noticed that the waves always arrive more or less parallel to the beach. However they did not necessarily start out propagating directly toward the beach. Indeed, if you go to another nearby beach on the same day, where the coastline is oriented in a different direction, the waves still come in parallel to the beach. What happens for waves with non-normal incidence is that the part of the wave closest to the beach reaches the shallows first, and its propagation speed slows down because the propagation velocity of water waves depends on the depth of the water. For non-normal incidence waves there is a speed variation across the wavefront, so that the parts of the wave that are in deeper water catch up with the parts in shallower water. The entire wave rotates until the wavefront becomes parallel to the beach.

This is an example with 1-dimensional wavefronts in a 2-dimensional space with a refractive index that is constant for points the same distance from the beach. With the apple we have 2-dimensional "wavefronts" (planes transverse to the momentum density) in a 3-dimensional space, with a refractive index that is constant for points the same distance from the centre of the earth. We also have waves that are propagating in every direction. All the momentum densities rotate towards the centre of the earth, and so the apple falls.

There is a somewhat subtle point to this verbal description in the field resonator model. The effect is not driven so much by the variation in the refractive index across an individual resonator as it is by the different refractive indices in resonators at different space points, all of which are coupled together (at the level of subatomic particles) to form a widely distributed whole.

As intuitively appealing as the qualitative description may be, it is far easier, analytically, to construct the curved space as follows in the next Chapter.

Chapter 11

The Gravity Model

The model here covers the time independent N-body case, where time independent means that the N massive bodies are not moving at relativistic velocities. This covers most of the gravitational phenomena in the Universe. For example, the falling apple, the motions of the planets in the solar system, the stars in the Milky Way and so on, are all covered by the time independent case. An important reason for making the restriction to the time independent case can be seen from Figure 4.5: The source term will inevitably be anisotropic in the relativistic case, which would necessitate (to begin with) using a tensor energy-momentum density and tensor equations that are not readily visualisable.

On a more technical note, in the N-body time independent case, there is always a transformation to an isotropic coordinate system, and the line elements here will correspond to the isotropic coordinate scheme. Yilmaz published two versions of his theory, a time independent version based on a scalar "potential" in the 1950's [4], and beginning in the 1970's, a time dependent theory with a tensor "potential" [5, 96].

The scalar "potential" in the time independent theory corresponds to the scalar refractive index that will be used here. However, using a scalar parameter for the refractive index is only viable if there is a mathematical transformation to an isotropic coordinate system (i.e. one in which the velocity of light is the same in all directions), as discussed in Section 11.7. Finally, there is a theorem that states that force and work methods are available without approximation in curved spaces [102], but it also applies only when there is a transformation to an isotropic coordinate system. That does not imply that force and work methods are not generally valid in curved spaces, but I understand that it remains to be shown. Consequently, the refractive medium physical model is best restricted to the time independent case, at least at the outset.

Having said that, this model is isomorphic to the time independent Yilmaz Theory [4]: it produces the same results for any mass distribution, as will be shown by finding the N-body line element. It also inherently features local Lorentz Invariance. Therefore, we have both the transformations between comoving observers at different places in a gravitational field and the Lorentz Transformations between colocated observers in relative motion. Mathematically speaking, the availability of an extension to the time dependent case is all but implied, as discussed in Section 11.7.

Refractive medium models have their roots in Electromagnetics, rather than

Optics, and the standard approach is to consider a medium of variable dielectric constant ϵ. This also assumes that the impedance of the medium is constant, so that the ratio of the dielectric constant to the permeability, ϵ/μ, is constant. The Electromagnetic approach is then a single parameter model where the dielectric constant plays exactly the same role as the refractive index in Optics: $\epsilon/\epsilon_0 = n/n_0$, where n is the refractive index and ϵ_0 and n_0 are the free space values, far away from gravitational fields. This is the usual basis for a refractive medium approach, and it is adopted here.

There are four main steps in the development:

1. Find the impacts on clocks and rulers of changes in the characteristic velocity, *i.e.* when $\epsilon \neq \epsilon_0$. Using $E = \hbar\omega = Mc^2$, these length and time transformations have knock on effects on the self-energy and the mass of a particle. We shall identify the full set of metric transformations.

2. Identify and solve the differential equation that relates a spherically symmetric energy density source term, $\rho_E(\epsilon, r)$, to the impact on the dielectric constant, $\epsilon(r)$.

3. Find the energy density distribution of an elementary massive particle in a medium of variable but spherically symmetric dielectric constant, $\epsilon = \epsilon(r)$, and from the result develop a suitable source term to use with planets or stars. Given $\epsilon(r)$ we obtain $\rho_E(\epsilon, r)$, thereby closing the loop such that the energy density profile generates the correct dielectric constant profile to produce the correct energy density profile for the SSS case.

4. Formulate and solve the N-body problem by removing the assumption of spherical symmetry from the above.

Once we have the model in place, its various ramifications will be discussed in the later Sections.

With respect to the published article on this model [103], I had not considered de Broglie waves before developing the gravity model in 2002 - 2004, so the cellular microstructure field concept was unknown to me. This concept is very well suited for the physical model, but the original article had to be written knowing that I did not know the details of the field structure. In fact, I'd been presuming a standard soliton kind of model with trajectories that orbit the centre of the particle, but I was very much aware of the difficulties with respect to satisfying both the little group and the resonance condition at the same time, so care was taken to avoid any such structural assumption. Consequently, none of the analysis changes here, but there are several parts of the original text that are revisited in this updated presentation of the 2004 model.

11.1 The Metric Transformations

The transformations developed here are referred to as metric transformations because that is what they are: Transformations of measurement results between observers in a gravitational field and observers outside the field using identical measuring instruments to measure the same quantity. Some of these transformations constitute the line element, so there is a connection to the metric tensor when expressed in certain coordinate schemes.

The set of transformations is wider than just clocks and rulers. We require the physics to be the same for all observers so that all the usual results, like $E = Mc^2 = \hbar\omega$ and $E^2 = P^2c^2 + M^2c^4$, hold good for observers both inside and outside the field, and the transformations of lengths and time that provide the line element will also necessitate transformations of energy, momentum and mass.

Puthoff [97] has developed the same complete set of metric transformations on the basis of several more or less satisfactory heuristic arguments. However, they are easily derived from the proposition that energy propagates at the characteristic velocity in a medium with variable dielectric constant. Thus, the characteristic velocity becomes a variable, $c(\mathbf{r}) = c_0\,\epsilon_0/\epsilon(\mathbf{r})$. Of course, the local observed value will always be $c_0 = c$ for all observers[1].

Amongst all of the interrelated Mechanics equations, there is one quantity that involves neither c nor r nor ω nor P nor E nor M, namely the angular momentum of a particle. Whether it is an electron, a proton or a photon, the angular momentum is fixed. It is independent of the particle's energy and it is independent of the gravitational field, or, in this model, the dielectric constant of space.

Consider some wave element of momentum, $p = mc = E/c$ that is moving in a circle[2] of radius r in the xy plane. The angular momentum is $\mathbf{L} = \mathbf{p} \,\mathrm{x}\, \mathbf{r} = pr\,\hat{z}$, because $\mathbf{p}$ is transverse to $\mathbf{r}$. As far as the usual energy independence of a system's angular momentum is concerned, the product $pr = Er/c = E/\omega$ is fixed, so one may write $E = H\omega$ for some constant H, where H is the angular momentum. The relationship between H and Planck's constant h is of no concern for the gravity model, but the difference between the photon angular momentum, h, and the electron angular momentum, $h/2$, will be discussed in Chapter 12.

Now, this momentum is moving in a circle, so there must be a force acting on it to account for the acceleration, $a = v^2/r = c^2/r$. Since the inertia is m, the force is $d\mathbf{p}/dt = -mc^2\,\hat{\mathbf{r}}/r = -E\,\hat{\mathbf{r}}/r = -H\omega\,\hat{\mathbf{r}}/r = -Hc\,\hat{\mathbf{r}}/r^2$. This is a $1/r^2$ force which is also proportional to the characteristic velocity. In a dielectric medium where $c(\epsilon) = c_0\epsilon_0/\epsilon$, this force actually has the same form as Coulomb's Law, but note that r here is just the radius of a circle, not the distance between charges, so this is of no particular significance.

Consider two instances of this system in media of different dielectric constants, ϵ_0 and ϵ respectively. The product of the radius and the angular frequency is reduced in the latter case as: $r\omega = c(\epsilon) = c_0\epsilon_0/\epsilon = r_0\omega_0\epsilon_0/\epsilon$. If the radius of the system did not change, the interaction $d\mathbf{p}/dt = -Hc(\epsilon)\hat{\mathbf{r}}/r^2$ is reduced by the factor ϵ/ϵ_0 but the acceleration, $c(\epsilon)^2/r$, necessary to maintain the circular trajectory is reduced by the factor $(\epsilon/\epsilon_0)^2$. Clearly, under a step change, $\epsilon_0 \to \epsilon$, the value of r in the new system will be smaller, $r < r_0$. The central force will have done work, changing the value of p, and the corresponding work integral is:

$$\Delta E = \int_{r_0}^{r} \frac{Hc_0\epsilon_0}{\epsilon r^2} dr \tag{11.1}$$

[1] Wherever the characteristic velocity is a variable, the functional dependence will be written out explicitly, as in $c(\epsilon)$ or $c(\mathbf{r})$.

[2] This could equally well be any movement on the surface of a sphere. The geometry isn't critical for the scaling properties developed here.

We cannot, however use this integral (written for a step change in ϵ) because the initial condition is wrong. As written it answers the question how an ϵ_0 entity would change if it were to find itself, unaltered, in an ϵ medium, but that cannot happen. We must bring the system from the ϵ_0 medium into the ϵ medium. Consider, therefore, adiabatic changes of the dielectric constant, and let us try to find a solution such that:

$$r_0 \to r = r_0\big(\frac{\epsilon}{\epsilon_0}\big)^{-m}$$

and

$$\omega_0 \to \omega = \omega_0\big(\frac{\epsilon}{\epsilon_0}\big)^{-n} \ .$$

Clearly we require $m + n = 1$, and it remains to solve for m and n. The work integral can now be written out with the dielectric constant as a function of the radius,

$$\epsilon = \epsilon_0\big(\frac{r_0}{r}\big)^{\frac{1}{m}} ,$$

which gives:

$$\Delta E = \int_{r_0}^{r} \frac{Hc_0\epsilon_0}{\epsilon_0(\frac{r_0}{r})^{\frac{1}{m}} r^2} dr = Hc_0 r_0^{-\frac{1}{m}} \left[\frac{r^{\frac{1}{m}-1}}{\frac{1}{m}-1}\right]_{r_0}^{r} = \frac{m}{1-m}(H\omega - H\omega_0) \quad (11.2)$$

Since $\Delta E = H\omega - H\omega_0$, $\frac{m}{1-m} = 1$ and so $m = n = 1/2$. The transformations above are therefore given by:

$$r = r_0\sqrt{\frac{\epsilon_0}{\epsilon}} \qquad \omega = \omega_0\sqrt{\frac{\epsilon_0}{\epsilon}} \quad (11.3)$$

These metric transformations then fall through to the dimensions and frequencies of all physical systems, including especially clocks and rulers. Since we write $E = H\omega$, the transformation of ω enforces a similar transformation of the self-energy, so the energy of the entity is now:

$$E = E_0\sqrt{\frac{\epsilon_0}{\epsilon}} \quad (11.4)$$

This in turn requires a transformation of the wave inertia, m, which we can deduce either from $E = mc^2$, or by writing $L = mr^2\omega =$ constant so that:

$$m = m_0\Big(\frac{\epsilon}{\epsilon_0}\Big)^{3/2} \quad (11.5)$$

And similarly:

$$p = p_0\sqrt{\frac{\epsilon}{\epsilon_0}} \quad (11.6)$$

These last two transformations are of particular interest in obtaining the right field equations. The system energy decreased, but the wave inertia and momentum increased (in the units system of an observer far outside the field).

When this observer formulates a problem in gravity he does not use the variables that pertain to observers on the distant planet under consideration. In particular, he uses the rest mass of the planet as he observes it, not as it might be observed locally. When we combine two gravitational masses, the sum of their rest masses increases, since each one is now in a stronger gravitational field with a greater value of ϵ. Therefore, this set of transformations does not conserve the total rest mass of a system of N-bodies. Nor does it conserve the self energy of a particular body, but that does not prevent overall conservation of energy-momentum.

The first thing to notice in passing is that the Einstein Field Equations do conserve the rest mass, and this is one of Huseyin Yilmaz's most salient criticisms [2]: If the rest mass is conserved, then energy-momentum cannot also be conserved. Indeed, there is an identity in the curved space mathematics, known as the Freud identity, that must be satisfied in order to conserve energy-momentum. This essential identity was not discovered until many years after Einstein had already published on General Relativity, and the Physics had been set in stone:

> *"Freud identified in 1939 a fourth identity that, unfortunately, was not aligned with Einstein's doctrines and, as such, the identity was ignored in virtually the entire literature on gravitation of the* 20^{th} *century."* E. A. Notte-Cuello and W. A. Rodrigues Jr. [104]

Einstein's Theory doesn't satisfy the Freud Identity in the general case, so it does not conserve energy-momentum.

The second thing to notice is the question that the metric transformations raise about the source term in gravity. In the field equations, it is not strictly speaking a mass tensor on the right hand side, but an energy-momentum tensor. The mass of the planet is ultimately used as the source term, but this involves the presumption that mass and energy are interchangeable, which is valid enough in the curved space theories of Einstein and Yilmaz. However, with refractive medium interpretations, the question arises whether one should use mass or energy as the source term. When the characteristic velocity is a variable, they are not interchangeable.

Whether or not either Yilmaz or Puthoff confronted this particular issue explicitly is unclear, but they ended up using different source terms. The source term in Yilmaz is equal to a constant, M^0, the "non-interacting mass", divided by the square root of a term equal to the refractive index here, so Yilmaz's source term is the same as my self-energy. Puthoff's model uses mass as the source term. Like mine, his equations are written in the unit system of the observer at infinity so the source term for any given mass increases when it is placed in a gravitational field. Consequently, his line element is different from Yilmaz's: The results are similar in weak fields but they diverge in strong fields.

I have asked Hal about this, and he informs me that he is committed by the algebra to using the mass. I don't agree, because we shall see in the next Section that the equation that comes out of his Lagrangian is a member of a family of equivalent equations, open to multiple interpretations corresponding to different source terms. However, it is his theory, he understands it best, and I take him at his word despite being unable to identify any place in his algebra that fixes the source term.

In the present model, this family of differential equations does not enforce the choice of the source term, but there is another place in the algebra that does. We shall see why the refractive medium model has to use the self-energy, the same source term as Yilmaz.

The metric transformations already hint as to the reason: if we use as source term a variable - mass - that increases in gravitational fields, the source term may runaway to infinity in strong fields. On the other hand, with a source term that decreases in strong fields, the model may be self-limiting so that no matter how much mass is gathered together in one place, it can never lead to a singularity.

11.2 The Chacteristic Velocity Profile

The role of the characteristic velocity profile in a refractive medium approach is analogous to the field equations in general relativity. On one side of the equation, there should be a source term, and on the other the resulting space dependence of the characteristic velocity, which falls straight through to the curved space line element in isotropic coordinates as a consequence of the metric transformations identified in Section 11.1 above.

The differential equation used for this part of the model comes from a Lagrangian written down by R.H. Dicke, and later included in Hal Puthoff's mixed Electromagnetic and gravity Lagrangian approach. In constructing the Lagrangian, there is a free choice of an arbitrary function that multiplies the term for the scalar wave Lagrangian density. The function freely chosen at that stage, $g(\epsilon) = 1/\epsilon^2$ ([97] equation 31), determines the exponential metric result. Parameters are then chosen to match up with the evidence, and the theory becomes as viable as the results, but there is really no physical motivation for the choice of this function.

Both Yilmaz and Puthoff talk about including an extra, distributed field source in addition to the usual mass term.

For Yilmaz, it is the gravitational field stress-energy [4], which is more a name than an interpretation. There is an analogy to the Newtonian gravitational potential, which is actually a work function in a field. None of this is unsatisfactory, but the abovementioned criticism remains that this stress energy field is "not well-defined".

Puthoff does offer a clear physical interpretation. He proposes that space is polarisable, which is to say that it is a dielectric medium, with a variable characteristic velocity, but then he uses a Lagrangian method with the above freely chosen term to determine the distribution of the polarisation field. Again, there is nothing unsatisfactory about that in Physics, but here, "*After* Physics", we are looking for physical motivations behind target theories. Motivations, in particular, that take us back to the core physical concepts of energy and momentum.

The interpretation of the mathematics that emerges below has nothing to do with a polarisation energy, which is a secondary energy induced by some primary applied field. It is an interpretation directly in terms of the field energy that we must associate with the particle itself in order to obtain a finite total energy for the particle as a space integral over its energy density. What we shall find is that the seemingly arbitrary choice of a function multiplying the scalar wave Lagrangian density corresponds to a very well known physical mechanism. This seems to strike an appropriate balance in the sense that we obtain a meaningful physical explanation that is more independent of any Metaphysical inferences about the polarisable nature of the medium.

Upon limiting Puthoff's general mixed gravity-EM equations, equation (34) of [97], to the general time independent gravity case, the Lagrangian method provides the following equation for the dielectric constant:

$$\nabla^2\sqrt{\epsilon} = \frac{\sqrt{\epsilon}}{4}\left(\frac{\nabla\epsilon}{\epsilon}\right)^2 , \tag{11.7}$$

the solution of which for the SSS case is an exponential profile, given by Puthoff in the general form:

$$\epsilon(r) = e^{2A/r} , \tag{11.8}$$

where the constant, A is chosen (by Puthoff) as $A = GM/c^2$ to match the weak field evidence. However, the Lagrangian method provides at best limited insight into the physical basis for this outcome, relies on a formalism with point charges written in, and masks what we shall argue is a flaw in Hal Puthoff's choice of the constant, A. Using (11.7) with the identity $\nabla^2(\epsilon^n) \equiv n(n-1)\epsilon^{n-2}(\nabla\epsilon)^2 + n\epsilon^{n-1}\nabla^2\epsilon$ gives a family of equivalent equations for different values of the exponent on the left hand side of (11.7) :

$$\nabla^2\epsilon^n = n^2\epsilon^n\left(\frac{\nabla\epsilon}{\epsilon}\right)^2 \tag{11.9}$$

All of these equations have the same meaning and (11.7) is the equation for $n = 1/2$. In the SSS case, after substituting $F(r) = \epsilon(r)^n$, using $\nabla\epsilon/\epsilon = \nabla F/nF$ and writing out the Laplacian and the gradient, this becomes:

$$\nabla^2F(r) = \frac{2}{r}\frac{dF}{dr} + \frac{d^2F}{dr^2} = 1/F\left(\frac{dF}{dr}\right)^2 . \tag{11.10}$$

The solution of (11.10) is $F = e^{a/r}$, for some constant a. The LHS is $\nabla^2F(r)$, so this general form of equation should be open to an interpretation by analogy to Gauss's Law, which states that the volume rate of production (consumption) of the flux of the gradient of the function, *i.e.* ∇^2, equals a source (sink) density. We might anticipate finding a recognisable source term on the RHS, but this is not yet obvious.

Substitute $F = e^f$ so that $dF/dr = Fdf/dr$. Then the RHS is equal to $F(\frac{df}{dr})^2$, rendering the solution, $f = a/r \Rightarrow F = e^{a/r} \Rightarrow \epsilon = e^{a/nr}$ more obvious. Given that solution, the RHS of (11.10) is $a^2e^{a/r}/r^4$ and the SSS equation in ϵ^n, (11.9), is:

$$\nabla^2\epsilon^n(r) = \frac{a^2\epsilon(r)^n}{r^4} \tag{11.11}$$

This is now in the same form as Gauss's Law for a $1/r^4$ distributed source term. Note that we obtained this equation by knowing the solution, but it is nonetheless valid for the SSS case, and it will become obvious below that its extension to the general time independent case is also valid. We are free, at this point, to choose any value of n, but the correct choice, $n = -1$, will be identified in the next Section from the condition that the space integral of the source term must remain finite. This equation now takes on its penultimate form in the SSS case:

$$\nabla^2\epsilon^{-1}(r) = \frac{a^2}{\epsilon r^4} = \frac{4A^2}{\epsilon r^4} \tag{11.12}$$

Where $a = -2A$, so that the solution is the same as Puthoff: $\epsilon = e^{-a/r} = e^{2A/r}$ and we can directly compare the choice of the constant. It is shown below that the energy density at radius r in a medium with spherically symmetric ϵ is $\rho_E \propto 1/\epsilon r^4$, so the right hand side of the above equation is proportional to the energy density, and it becomes[3]:

$$\nabla^2 c(r) = \kappa \rho_E \tag{11.13}$$

For an appropriately chosen natural constant, κ, to be identified in the next Section, along with the right choice for the constant, A, in this model. This is now manifestly of the same form as Gauss's Law.

11.3 The Field Energy Density of a Celestial Body

If we consider the far field, say out near the Moon, of an individual electron or proton that is part of a ruler here on earth, it has very little to do with the space that is spanned by the image of the ruler. The length of the ruler is entirely driven by the fields in the immediate vicinity of the ruler. Therefore, when calculating the metric transformations in Section 11.1 above, it was sufficient to consider how dimensions and frequencies scale for entities in a medium of constant ϵ. On the other hand, when it comes to considering the total gravitational field of billions upon billions of elementary massive particles, the far field energy density matters.

To include the far fields appropriately, it is necessary to find the energy density of an individual particle in the far field, at least in the spherically symmetric ϵ case. We begin with the idea, already familiar from Chapter 9, that any changes in the force field strength of underlying fields must ultimately be implemented by separating fundamental entities across the field, a process which, according to the Newtonian paradigm, requires that work be done in proportion to the pre-existing field strength, $\mathbf{E}_i$, and the magnitude, $d\mathbf{E}_i$, and physical extent, Δx, of the change:

$$dW \propto \mathbf{E}_i \Delta x d\mathbf{E}_i \tag{11.14}$$

Where the field strength, $\mathbf{E}_i$ should not be confused with the system energy, E. Integrating this Newtonian conception of the field increment leads to correspondence (up to a constant) with the usual result from Electrostatics for the relationship between field strength and energy density:

$$\rho_E = \frac{1}{2}\epsilon \mathbf{E}^2 = \frac{\mathbf{D}^2}{2\epsilon} \tag{11.15}$$

Where $\mathbf{D} = Q\mathbf{r}/4\pi r^3$ is a true $1/r^2$ field that is independent of ϵ. It will now be shown from within the model, *i.e.* without assuming Electrostatics or the

[3] To be fair to the reader, I actually came to this equation in a different way, by thinking of the space as an elastic continuum and the impact of an energy density as the removal of part of the material. The surrounding elastic material would be stretched leading to a Gauss Law impact on properties such as density, stress etc. Such devices are not presented due to the difference between models of reality and models of theories: the idea of space as an elastic medium introduces redundant Metaphysical content (Section 2.3, criterion 2.).

above equation, that the same ϵ dependency, $\rho_E \propto 1/\epsilon$, applies to field energy densities of massive particles in general.

This can be seen from at least two distinct considerations. First, let us write the relationship between energy density and the self-energy of a quantum system in a region of constant ϵ as:

$$E = \int_{r_c}^{\infty} \frac{4\pi r^2 K}{r^4} f(\epsilon) dr = \frac{4\pi K}{r_c} f(\epsilon) \Rightarrow f(\epsilon) = \frac{E r_c}{4\pi K} \ , \qquad (11.16)$$

where K is a constant, ϵ is not a function of r, and r_c is a cutoff radius, which will be dispensed with below. For the moment, we may ignore the integral in the region below r_c on the basis that this system, in a flat space, scales as a whole according to the transformations in Section 11.1, so that, whether it is significant or not, the integral in the region below r_c remains in proportion to the integral above r_c, and can be absorbed into the constant, K. This form of relationship applies to individual, isolated, non-interacting quantum systems at rest in a flat space of non-unity ϵ. We know from the metric transformations that both the system self-energy, E, and the lower limit of the integration region, r_c, vary as $1/\sqrt{\epsilon}$, so the energy density varies with ϵ as $f(\epsilon) = 1/\epsilon$.

Whilst this argument is strictly available only in a space with constant ϵ, the free space value, r_{c0}, is already so small relative to the dimensions of massive objects in all situations of interest (strong and weak fields alike) that, for any practical purpose, the energy density integral can be truncated within a region where ϵ is constant. Note that the energy density of an individual quantum system is such that $\rho_E/\rho_{E_0} = (E/E_0)^2$ in a space of constant ϵ.

Second, in this model, the energy is instantiated in tiny spherical cellular resonators that obey the little group. Consider, in flat space, $\epsilon_0 = 1$, similar annular regions in the far field, which is to say regions of radius R and thickness ΔR, where $\Delta R \propto R$. We expect to find similar fields in similar regions, which is to say that the field in one region is a scaled version of the field in a different region. In particular, if the asymptotic energy density varies as $1/r^4$, the number density of resonators will be $1/r^3$ and the energy per resonator will vary as $1/r$ (which corresponds to its potential energy in a $1/r^2$ force field system).

Now, if we compare a quantum system in free space with the same system after being moved to the centre of a spherically symmetric inhomogeneity in the medium (*i.e.* such that $\epsilon = \epsilon(r)$) the resonators that are associated with the radius r' in the second case were previously associated, in the free space system, with the radius $r = r'\sqrt{\epsilon(r')}$. This is because they are sitting on a surface of constant ϵ, and all dimensions on that surface are reduced in accordance with the metric transformation. In the free space system, the self energy per resonator at r is $1/\sqrt{\epsilon(r')}$ times the self energy per resonator at r'. When the quantum is moved to the centre of the inhomogeneity, the corresponding resonator number density increases by the factor $\epsilon^{3/2}$, but the result of this is just that the resonator number density at radius r' is the same in both systems.

In the course of moving from free space to a position centred on the gravitational inhomogeneity, the self-energy of each resonator was further reduced by a factor $\sqrt{\epsilon(r')}$, so the overall field energy density at r' (recalling that the number density is invariant) is reduced by the factor $\epsilon(r')$, and we may now write the SSS energy density as $\rho_E = \rho_{E_0}/\epsilon = K/\epsilon r^4$. This result is independent of the functional relationship between ϵ and r, but still restricted to spherical symmetry, although we shall see later that even this restriction can be removed.

To calculate the energy density distribution for a large "point-like" gravitational body of self-energy E, we now substitute the solution for ϵ identified in Section 11.2 for the spherical case, which is in the general form $\epsilon(r) = e^{2A/r}$. Since the massive object is actually distributed, and the value of ϵ experienced by a given quantum system is overwhelmingly due to other nearby systems, it is meaningless to set the lower limit of integration equal to r_c, and, following Yilmaz [4], we shall consider the limiting behaviour as $r \to 0$. The system energy for quanta is:

$$E = \int_{r\to 0}^{\infty} \frac{4\pi r^2 K e^{-2A/r}}{r^4} dr = \frac{4\pi K}{2A}\left[e^{-2A/r}\right]_{r\to 0}^{\infty} = \frac{4\pi K}{2A} \tag{11.17}$$

Which fortunately does not diverge as $r \to 0$, but note that this depends on the choice $n = -1$. As discussed below, if we use mass as the source term, the corresponding integral diverges.

Having associated a finite energy with a point, the N-body solutions identified below can be used to model any mass distribution. Since each quantum system is embedded in a region of effectively constant, but non-unity ϵ, the self-energy of the massive object is $E \simeq m_0 c^2/\sqrt{\epsilon_{max}}$, where $m_0 = \Sigma_i m_{0i}$ and the m_{0i} are the free space values of the individual particles constituting the gravitational mass. However, we are not usually interested in what the mass of a celestial object would have been if it were divided into elementary quantum systems, so we shall take $K = AE/2\pi$ and the energy density as:

$$\rho_E = \frac{AE}{2\pi}\frac{1}{\epsilon r^4} \tag{11.18}$$

Now, as far as the system constant, A, and the natural constant, κ, are concerned, we also have the equations from Section 11.2 above:

$$\nabla^2 \epsilon^{-1}(r) = \frac{4A^2}{\epsilon r^4} \tag{11.19}$$

and

$$\nabla^2 c(r) = \kappa \rho_E \ . \tag{11.20}$$

If we write $c_0 = 1$, $\epsilon_0 = 1$, then $c(r) = 1/\epsilon(r)$, and we have:

$$\kappa = \frac{8\pi A}{E}$$

Which is the result given in the published article [103]. A/E obviously has to be a system energy independent constant. Therefore, where Puthoff chose $A = GM/c^2$, we must choose $A = GE/c^4$, where E is the self energy of the body for observers outside the field and c is a constant. This choice is enforced by the choice of $n = -1$ in (11.9). Other choices could be made, with other source terms. In particular, choosing $n = +1$ corresponds to choosing the mass density as the source term, and then one would have to choose $A = GM/c^2$. The problem is that, in this case, the above integral (11.17) for the total quantity of the source diverges and we could not take the limit as $r \to 0$. A related consequence of choosing mass as the source term is to produce a theory that is pathological in strong fields. The source term used here varies as $1/\sqrt{\epsilon}$, not $\epsilon^{3/2}$.

Yilmaz uses the same source term as the present model, as can be seen from [4], equation (37'), which reads:

$$M^j = M^{0j} exp\left(-\sum_{k\neq j} GM^k/c^2|\mathbf{r}_k - \mathbf{r}_j|\right) \tag{11.21}$$

Where M^j is the interacting mass of the j^{th} body and M^{0j} is its non-interacting mass. Of course the non-interacting mass already takes into account the decomposition of the j^{th} body into its myriad parts, each of which is sitting in the gravitational field of all the others, so Yilmaz's source term is the same as in this model.

Finally, if we write out the characteristic velocity in the normal way, $c(r) = c_0\epsilon_0/\epsilon(r)$, then the value of κ is:

$$\kappa = \frac{8\pi\epsilon_0 G}{c_0^3}$$

Where c_0 is the usual background value of the characteristic velocity.

The treatment above is very much analogous to [4], Equations 14, 21, 38 and 39.

The result is only as good as the use of a singularity to represent a celestial object, and does not apply to individual sub-atomic particles in a curved space. However, since this assumption is common to all present theories of gravity, and since the intention is to model the theories as opposed to merely the observables, we shall adopt it without further comment except to mention that the form of (11.18) (i.e. with $\rho_E \propto E^2$ for a singularity) is common to both PV (Dicke and Puthoff) and the Yilmaz theory. It is a consequence for the static limit of the usual demand for scalar gravitational waves, which is the usual assumption made when writing down a Lagrangian density for the scalar field in any form akin to Puthoff's $L_d^\epsilon = \lambda g(\epsilon)((\nabla\epsilon)^2 - \epsilon^2(\partial\epsilon/\partial t)^2)$ (and recall that, in the Dicke / Puthoff Lagrangian, $g(\epsilon) = 1/\epsilon^2$).

In Section 11.1 we established an association between quantised angular momenta and $1/\epsilon r^2$ force fields. In Section 11.2, we established the association between a $1/\epsilon r^4$ energy density and the exponential profile of the characteristic velocity and here we have closed the loop with the association between $1/\epsilon r^2$ force fields and the $1/\epsilon r^4$ field energy density. We are now ready to extend the above SSS model to the N-body problem in gravity.

11.4 The N-body Problem

Although the equation $\nabla^2 c(r) = k\rho_E$ above arose in the context of an SSS problem, it is a local equation analogous to Gauss's Law and one should expect it to be similarly independent of the source configuration. Hence we might anticipate using the same equation in the general case of N bodies. However, the energy density on the right hand side of, for example, (11.12) (which is a function of $1/\epsilon r^4$), does not superpose linearly because it depends on a property of the medium, $\epsilon(\mathbf{r})$ that depends in turn on all of the massive bodies involved in an N-body problem. Consequently, we need to consider the general equation, not limited to the SSS case, for the case $n = -1$ in (11.9), which is:

$$\nabla^2 \frac{1}{\epsilon} = \frac{1}{\epsilon}\Big(\frac{\nabla \epsilon}{\epsilon}\Big)^2 \ . \qquad (11.22)$$

The N-body energy density will emerge from the solution to this equation. Yilmaz's solution for the N-body time independent case corresponds to:

$$\epsilon(\mathbf{r}) = Exp\left[\frac{2G}{c^4}\Big(\Sigma_{i=1}^{N}\frac{E_i}{r_i}\Big)\right] = \Pi_i \epsilon_i \ , \qquad (11.23)$$

where $\epsilon_i = e^{2GE_i/r_i\,c^4}$, the E_i are the observed self energies of the various gravitating bodies, r_i is the distance[4] from the i^{th} body to the point $\mathbf{r}$, and $\epsilon_0 = 1$. The term E_i/c^4 is identical to Yilmaz's interacting mass term, M^i/c^2, as discussed above at (11.21). After substituting this trial solution into (11.22), and noting that $\nabla^2[\Sigma_i(E_i/r_i)] = 0$ because $\nabla^2[1/r_i] = 0$, both sides reduce to:

$$\nabla^2 \frac{1}{\epsilon} = \frac{1}{\epsilon}\Big(\frac{\nabla \epsilon}{\epsilon}\Big)^2 = \Big(\frac{2G}{c^4}\Big)^2 \frac{1}{\epsilon}\Big[\Big(\Sigma_i E_i \frac{\partial r_i^{-1}}{\partial x}\Big)^2 + \Big(\Sigma_i E_i \frac{\partial r_i^{-1}}{\partial y}\Big)^2 + \Big(\Sigma_i E_i \frac{\partial r_i^{-1}}{\partial z}\Big)^2\Big]$$

where $r_i = \sqrt{(x-x_i)^2 + (y-y_i)^2 + (z-z_i)^2}$. Therefore, (11.23) is the solution to (11.22), which is the main result of this Section. Note, however, that the right hand side is equal to:

$$\Big(\frac{2G}{c^4}\Big)^2 \frac{1}{\epsilon}\Big(\Sigma_i \frac{E_i}{r_i^3}\mathbf{r}_i \cdot \Sigma_j \frac{E_j}{r_j^3}\mathbf{r}_j\Big) = \Big(\frac{2G}{c^4}\Big)^2 \frac{1}{\epsilon}\Big(\Sigma_i \frac{E_i}{r_i^3}\mathbf{r}_i\Big)^2 \ , \qquad (11.24)$$

In gravity, where the theories are nonlinear, the mere fact of having replicated the observational evidence is essentially trivial because there are an infinite number of similar nonlinearities that can reproduce the same observations within experimental errors. This applies to the present model as well as to every theory of gravity that has ever been written, including Einstein's and Yilmaz's. The model here, however, is not initially a model of the phenomena, it is a model of the generally covariant Yilmaz theory. Either it matches the theory perfectly or it does not.

In both the Yilmaz theory and the present model, a finite energy can be associated with a point because the limit as $r \to 0$ in (11.17) was finite. In this case, the N-body solutions can then be used to solve arbitrary matter distributions. The fact that both the theory and the physical model share the same N-body solutions shows that they are fully isomorphic to each other. Since the Yilmaz theory is generally covariant, so is the physical model.

The N-body line element in this model for isotropic coordinates is then induced by the metric transformations in the usual way:

$$ds^2 = g_{ij}dx^i dx^j = \frac{1}{\epsilon(\mathbf{r})}c^2 dt^2 - \epsilon(\mathbf{r})(dr^2 + r^2 d\theta^2 + r^2 sin^2\theta d\phi^2) \ , \qquad (11.25)$$

where $\epsilon(\mathbf{r})$ is given in (11.23).

[4] In Yilmaz's notation this is written as $|\mathbf{r}_i - \mathbf{r}|$.

Showing the N-body solution for (11.22), also shows the N-body solution for Puthoff's Polarisable Vacuum Theory (since both (11.22) and Puthoff's equation, (11.7), are members of the same family of equivalent equations, (11.9), all of which share the same solutions). To the best of my knowledge, Puthoff's group had not identified their N-body solution when this result was first published. Without a clear interpretation of the meaning of the SSS source term, (11.12), it would be a more challenging task.

11.5 Nonseparability of the Energy Density

As far as the N-body energy density is concerned, consider the Electrostatics relationship mentioned at the start of Section 11.3.

$$\rho_E = \frac{1}{2}\epsilon\mathbf{E}^2 = \frac{\mathbf{D}^2}{2\epsilon} \tag{11.26}$$

Where the displacement, $\mathbf{D} = Q\mathbf{r}/4\pi r^3$ is a true $1/r^2$ vector field, independent of ϵ. For a medium with a known $\epsilon(\mathbf{r})$, we could construct the Electric field energy density at a given point by first superposing the displacement fields, taking the squared magnitude of the resultant field and then dividing by $\epsilon(\mathbf{r})$. Although ϵ might itself depend on the charge distribution (as it depends on the mass/energy distribution in gravity), the vector superposition of displacement fields is ϵ independent.

With that in mind, the intermediate result above, (11.24), can now be interpreted as the N-body energy density, so that the equation $\nabla^2 c(r) = \kappa\rho_E$ applies to the general case as expected.

Calculating this expression begins by taking the vector sum of N *medium independent* vector fields, in the general form $\hat{\mathbf{r_i}}/r_i^2$, thus forming a global vector field analogous to the global displacement field in Electrostatics. The squared magnitude of this vector field is a medium independent scalar field and the N-body energy density at any given point is the product of this field times the medium dependent factor $1/\epsilon(\mathbf{r})$.

Let us try to take this analogy between the Electrostatic field energy density and the mass field energy density a step further. The suggestion seems to be that all of a particles' field energy density should ultimately be associated with the energies of underlying force fields, Electromagnetic or otherwise, to which we can associate a force field energy density.

In general, uncharged entities, the neutron for example, must be considered to have non-vanishing far field energy densities. For example EPR experiments are also conducted with neutrons, while the planets and stars here clearly have non-vanishing far field energy densities.

One way to put this analogy into practice is to consider the field of a neutral entity, like a planet, to consist of equal densities of force field constituents that are signed positive and negative. Both kinds of field would be considered to exist, even when their superposition vanishes, as in the case of a neutron or a planet that carries no electric charge.

The energy density for an uncharged entity would then be in the form $\rho_E = \kappa'(\mathbf{d}_+^2 + \mathbf{d}_-^2)/\epsilon$, where the $\mathbf{d}_+$ relate to the positively signed fields and similarly for the $\mathbf{d}_-$.

There are then two main possibilities for the energy density in an N-body problem in gravity, namely:

$$\rho_E = \frac{\kappa'}{\epsilon} \Sigma_{i=1}^{N} [(\mathbf{d}_{i+}^2) + (\mathbf{d}_{i-}^2)] \,, \tag{11.27}$$

where one squares the fields of individual bodies and then takes the scalar sum, or:

$$\rho_E = \frac{\kappa'}{\epsilon} [(\Sigma_{i=1}^{N} \mathbf{d}_{i+})^2 + (\Sigma_{i=1}^{N} \mathbf{d}_{i-})^2] \,, \tag{11.28}$$

where one takes the vector sum first and then takes the square.

The intermediate result above, (11.24), corresponds to the latter possibility. This expression is non-separable in the sense that one cannot strictly speaking talk about the energy density "of the i^{th} body" at some point, $\mathbf{r}$, and one cannot strictly define the self-energy of any one body in terms of an integral over all space of "its" field energy density. Only the self-energy of the entire system of N bodies is well-defined.

We can, however, talk about the abovementioned medium independent fields, on a particle by particle basis, but the energy that any given body's field contains depends on all the other bodies in the vicinity. We shall return to the interesting Metaphysical question this raises in the final Chapter.

From a practical perspective, we saw in the discussion following (11.16), that the total self-energy of each sub-atomic particle is for all intents and purposes equal to its value in a region of constant ϵ. This is also implicit in Yilmaz's equation 37', quoted above as (11.21). The approximation involved in connecting the observed self-energy of a body with the integral over "its" field energy density is certainly good enough for any and all practical purposes.

Furthermore, all theories of gravity operate on the basis of presumed separability: Each planet has its own mass-energy and so on. While we can identify invariants in the model that might allow us to redefine the source terms in a separable way, the physical model would go beyond its function, namely to model the Yilmaz theory, which it achieves without approximation.

11.6 Equivalence Principles

There is no good *a priori* reason to expect a strong equivalence principle[5] in a physical model where the mass observable is not a fundamental quantity. It is an effect that results from the inertia density of underlying propagative fields, and there is no surprise that the source term in gravity refers to the energy of underlying fields as opposed to their inertias.

The source term is driven by what things actually are in the model, in an ontological sense, rather than by an epistemic property, like mass. The metric transformations of Section 11.1 show clearly that the self energy of the body (or its "interacting mass" in Yilmaz [4]), is not the same as the value of M that is to be used in Mechanics equations by observers outside the gravitational field. Strong equivalence refers to the left hand side of $E = Mc^2$, not one of the parameters on the right.

[5]Strong equivalence means that active mass (the source term) = passive mass = inertial mass.

The usual weak equivalence principle that comes from equating a gravitational field with an acceleration, which is the equivalence of the passive mass and the inertial mass, applies in the usual way.

By contrast, Yilmaz induces a strong equivalence principle between his active and passive masses as part of justifying the usual d'Alembertian equations of motion. However, his result is simply that the bodies follow geodesics in the 4-space, which just means that the equations of motion involve neither inertial nor passive masses, so the question is moot. While Yilmaz's result is no doubt correct in his gravity formalism, the problem is that this formalism lacks the flexibility necessary in order to combine gravity and Electromagnetics (where dynamical interactions scale with the characteristic velocity) in the same model.

To see this, consider the equation $F = ma$, which must hold for all observers in the time independent limit. The situation in the present model, and in Puthoff and Dicke, is as follows.

Let observers outside the gravitational field consider a certain object that is outside the field and ascribe to it the mass M Kilograms. Write this as: $M_{OO} = M$, where the first subscript refers to the position of the observer and the second to the position of the object. M_{OO} is the mass value found by observers outside the field for the object outside the field. When the same object is placed inside the field, the value for observers outside the field changes to $M_{OI} = M\epsilon^{3/2}$. For observers inside the gravitational field, the value of the characteristic velocity inside the field is the same as the value of the characteristic velocity outside the field is for observers outside the field: $c_{II} = c_{OO} = c$. There are no transformations to apply, so the mass value they ascribe to the object, when it is inside the field, is $M_{II} = M_{OO} = M$. When it is outside the field, the value they ascribe to the mass is $M_{IO} = M\epsilon^{-3/2}$. Both observers agree that the object is lighter outside the field than inside, by the same factor.

Next, let observers outside the field consider the acceleration of an object that is outside the field, and let us say they find the result $a_{OO} = a$. Observers inside the field considering the same situation with their slow clocks and short rulers find the result $a_{IO} = a\epsilon^{3/2}$. Both observers agree on the product ma.

Now, let observers inside the field consider the acceleration of an object inside the field and find the result $a_{II} = a$. Observers outside the field find $a_{OI} = a\epsilon^{-3/2}$, and both observers again agree on the product ma. The right hand side of $F = ma$ is invariant under a change of observer.

On the left hand side, both observers agree that the interaction field is stronger outside the gravitational field by the factor ϵ, and they both agree that the dimensions of objects are larger outside the gravitational field by the factor $\sqrt{\epsilon}$, so the force that operates between two entities, $F \propto 1/\epsilon r^2$ is also invariant under a change of observer and so both sides of the equation $F = ma$ are invariant under a change of observer:

The equation doesn't just hold for both observers in the sense that $F = ma$ and $F' = m'a'$, but also $F = F'$ and $ma = (ma)'$.

This contrasts with the situation in Yilmaz, where the product ma is not invariant under a change of observer. Regardless of units, observers do agree about the facts. They agree, in particular, that the clocks and rulers of the observer inside the field are slower and shorter than the clocks and rulers of observers outside the field. To avoid confusion with the refractive medium interpretation, let us write Yilmaz's $e^{2GM/r} = \eta$. Recall that his source term M is the same as the self energy source term here. Observers in the Yilmaz theory have the corresponding relations: $a_{IO} = a\eta^{3/2}$, for an observer inside the field considering the acceleration of an object outside the field that observers outside the field measure as a; and $a_{OI} = a\eta^{-3/2}$ for an observer outside the field considering the acceleration of an object inside the field that observers inside the field measure as a.

It is clear that the relationship $c = c_{II} = c_{OO}$ also holds, because the locally measured value of the characteristic velocity is invariant for Yilmaz. It is also clear from Yilmaz's equation, cited above as (11.21), that the relationship $M_{IO}/M_{II} = M_{OO}/M_{OI} = \sqrt{\eta}$ holds: In Yilmaz the mass of the body increases outside the field. Finally, with respect to the self-energy the relationship $E_{II} = M_{II}c^2$ holds, but is this $E_{II} = E_{OO} = M_{II}c^2 = M_{OO}c^2$ or is it $E_{II} = E_{OO}/\sqrt{\eta} = M_{II}c^2 = M_{OO}c^2/\sqrt{\eta}$? Does $M_{II} = M_{OO} = M$ or does $M_{II} = M_{OO}/\sqrt{\eta} = M/\sqrt{\eta}$? It is not clear to me which way Yilmaz would have it, because he had no need to consider the question.

In the first case, $M_{IO} = M\sqrt{\eta}$. When the observer outside the field considers an accelerating body outside the field, and obtains the product ma, the observer inside the field, considering the same situation, obtains the product $M\sqrt{\eta}a\eta^{3/2} = Ma\eta^2$. To preserve the equation $F = ma$ he needs a theory such that the Electromagnetic field strength varies as η^{-3}. In the second case, $M_{IO} = M$. The observer inside the field obtains the product $Ma\eta^{3/2}$ and requires a theory where the Electromagnetic field strength varies as $\eta^{-5/2}$. Either way, treating the characteristic velocity as a constant for the purposes of evaluating the metric transformations has an unpleasant consequence.

This is precisely the difference between the Yilmaz theory and a refractive medium interpretation. In one case you have a strong equivalence principle for mass but Electromagnetics (and hence Quantum Field Theory) is incompatible. In the other you have a strong equivalence principle for energy and Electromagnetics is automatically included by virtue of the first motivating assumption, $c(\mathbf{r}) = c_0\epsilon_0/\epsilon(\mathbf{r})$. Puthoff deals with mixed Electromagnetic and gravity cases using the complete set of metric transformations. His approach is the one that is consistent with the present model in this respect, and his basic formalism unifies gravity with Electromagnetics - a necessary precursor to unifying gravity and Quantum Field Theory.

We have seen above that Puthoff's theory is not empirically equivalent to the manifestly generally covariant Yilmaz theory, even in the time independent limit, because the source terms are different. Puthoff's theory is not written in a manifestly covariant form. It is not isomorphic to Yilmaz's manifestly covariant theory and since the limit as $r \to 0$ in (11.17) does not exist for Puthoff, sharing the same functional form of N-body solutions doesn't help. Yilmaz's approach is the one that is consistent with the present model as regards gravity.

11.7 The Time Dependent Case in Gravity

11.7.1 The 1-Body Case

We have both the Lorentz Transformations, $dx^{a'} = L_a^{a'}dx^a$, that connect the coordinate systems of colocated inertial observers in different conditions of motion and the time independent limit, [4], of the Yilmaz Theory [5, 96] fixes the coordinate transformation $dx^{b''} = A_{b'}^{b''}dx^{b'}$ between comoving observers outside and inside a gravitational field. The model here obviously shares this transformation.

For the 1-body time dependent case these transformations can be combined in the usual way, so that the transformation $dx^{b''} = A^{b''}_{a'} L^{a'}_{a} dx^a$, connects the coordinate system of our "reference" observers (outside the moving field, at rest in the CMBR), with that of an observer inside the field and comoving with it.

The one-body time dependent problem is then straightforward. Consider a massive object moving at relativistic speed relative to an observer at rest in the CMBR frame, far outside the gravitational field. What does he find for the characteristic velocity at some point inside the field?

To begin with, we know that a second, colocated observer (outside the field but comoving with the massive object) finds an isotropic characteristic velocity that is less than the background value, in accordance with $c(\mathbf{r}) = c_0\epsilon_0/\epsilon(\mathbf{r})$, where $\epsilon(\mathbf{r})$ is a scalar. His velocity observations are connected to velocity observations by the first observer in accordance with the relativistic composition of velocities, (4.22). It can be seen by inspection that the characteristic velocity inside the field of a rapidly moving massive object is anisotropic.

We also know that the field energy-momentum density of a moving object exhibits a particular elliptical anisotropicity, distended by the factor γ in the direction of the relative velocity. This is shown in Figure 4.5, where the transformations of field momenta are identical with Special Relativity. An observer needs to use a tensor energy-momentum density to describe the moving source and he must also use a tensor to describe the dielectric constant because a scalar dielectric constant cannot describe an anisotropic characteristic velocity. Puthoff's model, with a scalar refractive index for the time dependent case, is inconsistent with the physics of his own model.

For the present model, the construction of the appropriate tensor equation is "straightforward" in the sense that it is a merely mathematical exercise. The scalar energy-density used here maps onto an energy-momentum tensor in the usual way, so the source term is given. We also know all the results of this equation, since they can be obtained by Lorentz Transformations without using the tensor equation directly. Since the mapping on the right hand side is the same as in Yilmaz, and the static case is the same as in Yilmaz, the one body time dependent results are the same as long as his theory has local Lorentz Invariance and reduces to the static limit in the time independent case (which it does). The mapping on the left hand side of the field equations here must then correspond, so the one-body time dependent model is the same as in Yilmaz, and the tensor potential terms that appear in his line element correspond to terms in a tensor refractive index in the physical model.

11.7.2 The N-Body Case

The question not answered here is whether this correspondence will carry over to the N-body time dependent case, where we cannot obtain the results with the above procedure because there is no observer who is comoving with all of the sources. When the sources are all in relative motion, a given Lorentz observer can only be comoving with at most one of the them.

The claim was made in the introduction to the previous Chapter, in Section 10.1, that gravity could be represented by the very simple equation $\nabla^2 c(\mathbf{r}) = \kappa\rho_E$. This is the loophole in that seemingly outrageous claim. This equation covers the N-body time independent case and, combined with the inherent Lorentz Invariance of the model, it contains all the predictions for the one body time dependent case[6].

The extension from the 1-body time dependent case to the N-body time dependent case is not 100% guaranteed but the existence of the Yilmaz theory [5, 96] is a strong argument that it should exist, with an N-body tensor refractive index formed by summing exponents from the N-one body tensor refractive indices induced by the discussion in Section 11.7.1. This loophole is at a rather theoretical level, and I leave it for interested theorists to close.

Of greater interest here is the physical meaning of a tensor refractive index in a physical model. The implication is that the impact of energy on the characteristic velocity depends (somehow) on the direction and speed at which the field is moving *as a whole*. This is not the same concept as the instantaneous direction of propagation of a field, *i.e.* the unit wave vector.

Recall that, in the comoving frame, the wave vector exists on the surface of a sphere. For the moving particle the wave vector exists on the surface of an ellipsoid of revolution (a sphere compressed by γ in the direction of motion) moving at speed v in the observer's frame of reference. At any given space point in the particle system, the effect of this is that the wave momenta form a cone, as shown in the top right hand side of Figure 4.6.

A more complete physical model should begin with the (anisotropic) impact of wave momentum on the characteristic velocity as a function of the angle between the wave vector and the direction in space. A tensor source field can then be induced by averaging over such vector fields and so on. But at the bottom of it all, the angular dependence of the characteristic velocity on this vector source term is unknown. In order to test whatever relationship is considered, the model should be built up (taking into account the nonlinear effects of gravity all along the way) before finally arriving at a pre-existing theory that validates the assumed angular dependence.

Meanwhile, at the top of it all, the target theory is less than fully clear because the distinction between the Yilmaz time dependent theory [5, 96] and a refractive medium interpretation has not been fully resolved, as we saw in Section 11.6. There is plenty of work for theoreticians still to do to put the theory of gravity on solid conceptual foundations commensurate with achieving the long sought after unification with Quantum Theory.

[6] Note that, unlike all the usual theories of gravity, we are not assuming local Lorentz Invariance. It is built into the model.

Chapter 12

Spin 1 and Spin 1/2 Particles

While the development of the physical model of the foundations of Physics is now complete, there are some outstanding qualitative questions that the model can shed some light on that may be of interest to readers. Specifically, why is angular momentum quantised, or "digital", in an analogue field model where it is comprised by myriads of pointlike field resonators? The reader may also be interested in physical models of the resonators themselves. Finally, in the development of the metric transformations, Section 11.1, angular momentum quantisation was assumed and discussion of the relationship between the photon energy-angular momentum relations $E = \hbar\omega$, $L = h$ and the corresponding electron relations, $E = \hbar\omega$, $L = h/2$, was postponed until now because it was not relevant for gravity.

Given the wave postulate, the energy independence of a system's angular momentum, be it light or matter, is readily understood: Systems are described by a wave trajectory structure and, for any given shape / topology, luminal propagation implies an inverse proportional relation between energy and dimension. Consider any given closed wave trajectory in a system and let the frequency on the trajectory be ω. Choose any two points on the trajectory, and let the distance between them be r. Clearly, $r \propto c/\omega$, so that if we now construct a scaled version of this trajectory in which $r \to r/n$ then $\omega \to n\omega$.

Since we showed in Chapter 4 that $E = cp$ the product $pr = Er/c \propto E/\omega$. For wave systems of the same kind, the consequence that the angular momentum is independent of the system energy is noteworthy. The usual vector cross-product in $\mathbf{L} = \mathbf{p} \text{ x } \mathbf{r}$ is irrelevant because the geometry is the same at any scale. We also saw that the energy-momentum transformation between frames is the same as the relativistic Doppler shift for frequencies, so that energy, momentum and frequency all vary together. The overall result is that, as we transform any given system from one frame to another, the angular momentum remains independent of the energy because adding linear momentum to the system does not change its angular momentum.

If we now consider a physical model of one particular photon from the perspectives of many different observers, we can induce the physical models of any photon of any energy. The energy of such a system is not in any way quantised, but regardless of the energy, the photon's angular momentum is invariant. As long as radiation comes in packets, it is only to be expected that they should all have the same angular momentum.

As for the idea that radiation comes in packets, the necessity to satisfy a resonance condition when forming a 3-dimensional closed trajectory system for any particle of matter implies a discrete set of solutions with discrete energy levels for atoms. Transitions between those levels would then come in packets of radiation that are emitted or absorbed as wholes, and the fact is that our experience of radiation is limited to such pointlike events.

It is worth mentioning in passing here a nice proof in Konopinski's Classical Electrodynamics textbook related to this point. Regardless of quanta, he shows that any free Electromagnetic field must contain angular momentum [86].

While such considerations don't prove conclusively that it must be so, they do show that angular momentum quantisation makes good sense for both matter and radiation in an (analogue) wave ontology subject to a characteristic velocity. The analogue system resolves itself in a digital way. This contrasts with ontologies where things don't move at fixed speed where there is no correspondingly simple explanation for the phenomenon.

The question of the different angular momentum quanta, h for photons and $h/2$ for fermions, can be approached from the point of view of the topological constraints on the corresponding wave trajectory systems. Without developing as yet unknown nonlinear field equations, or Quantum Field Theory, it is not possible to predict the discrete spectrum of matter particles directly from the wave postulate alone. However, the existence of a discrete spectrum of massive "fundamental" particles makes complete sense from the point of view of resonance conditions. The notion that they go on to form a discrete set of subatomic particles which goes on to form a discrete system of atoms with discrete energy levels all makes sense as the consequence of satisfying various resonance conditions at different stages.

On that basis, we can draw some qualitative pictures which may shed some light on the question why photons are spin-1 and fermions are spin-1/2.

With respect to the photon first, there is a little group for photons, similar to the little group for subluminal solutions, like the electron, except that the evolutions in the comoving frame under SO(3) (rotations in 3 dimensions) are replaced by evolutions under SO(2) (rotations in a plane). Since light has been recognised as a transverse wave phenomenon for centuries, one begins with the idea of rotations in the plane transverse to the direction of propagation. Light has a wave length and a periodicity. As it moves forward in space by one wavelength, the oscillation goes through one period, which corresponds to an SO(2) movement around a circular path in the transverse plane.

We thus have a helical "internal trajectory" that exists on the surface of a cylinder, but such an individual trajectory could not exist in isolation: there must be forces acting on it to provide the central acceleration. Although more complicated structures are possible, the simplest system comprises two such "trajectories" executing a bounded, internal circular motion, spinning about each other as the system moves through space. The basic metaphor for the photon structure is then a double helix formed by a binary field system. This kind of physical structure has been proposed so many times, by so many authors that one might say this is the usual idea of the photon structure.

Of all the many ways to motivate the usual structure, Donev and Tashkova's approach [105] most closely parallels the present model. They enquire about nonlinear bivector fields (*i.e.* by analogy to the Electromagnetic bivector) in general on the basis that energy and momentum are conserved, and therefore linear, even when field variables are not. Forces between wavefields are defined in the same way, as distributed transfers of linear momentum, and the double helix structure is induced without making any dynamical assumptions about the nonlinearity.

The difference here is how we are to interpret the double helix model. For Donev and Tashkova and all previous authors, the double helix structure "is" the photon. One well-localised helix = one photon $\Rightarrow$ no EPR correlations.

If we were developing a photon model here, we would have begun with the double slit experiment for light and it would have led us to a coupled field resonator model, where each resonator corresponds to the usual double helix model and the photon is a widely distributed system of resonators coupled to each other so that they all share the same space independent phase.

It is pertinent to note that these internal trajectories of the double helix are not *luminal* wave trajectories (which correspond to the wave vector of the photon field and represent the linear momentum density of the *composite* system - a field resonator that is moving at c in one particular direction, not helically). The internal trajectories are at the level of the photon substructure where the speed on the helical trajectory necessarily exceeds the characteristic velocity. Some degree of superluminality in the substructure is more or less implicit in the (luminal) wave postulate to the extent that one needs a balanced binary field system to generate luminal trajectories.

Let us add a further assumption, namely that the helix pitch in the substructure is 45 degrees, so that the path length around the circle is then equal to the wavelength, with the result that it is also traversed at the characteristic velocity c. This assumption is not at all justified, but nor are the "pictures" here relevant to any of the derivations (for Special Relativity, the Dirac Equation, or gravity) that constitute the physical model for the foundations of Physics.

Adopting this assumption simplifies the pictures and leads to a reasonable looking structure for spin-1/2 wave trajectories.

There are two available photon helicities, left handed and right handed, which correspond to the usual circular photon polarisations. Other kinds of polarisation can be constructed accordingly. For example, by splitting each internal trajectory into left and right handed parts, linear polarisation (Horizontal / Vertical) is obtained in the usual way.

The fact that all such systems have the same angular momentum implies that the binding interaction, expressed at the level of the internal trajectories, is proportional to the total energy and inversely proportional to the radius and since the radius is inversely proportional to the energy, the interaction is $1/r^2$, independent of the system energy (as already discussed in Section 11.1).

That is not to say that the internal trajectories interact with each other at a distance via a $1/r^2$ force. The internal trajectory pair is a visual metaphor for a binary field system where the fields overlap, interpenetrate and interact with each other throughout the resonator, and decay rapidly outside of it. The "real" interaction is in the form of distributed transfers of energy/momentum between distributed fields, just as in Donev and Tashkova. The field resonator is represented by a pair of distinct trajectories separated in space by the distance r for the present purposes of visualisation and discussing how the system scales: when the field resonator scales, so does any given pair of internal trajectories, and similarly the rate of internal momentum exchanges scales with the characteristic dimension r of the system as $1/r^2$.

Let us now consider the field resonators for an electron model. We have internal trajectories that exist on the surface of a sphere rather than a cylinder. In this case, the characteristic trajectories that meet the resonance condition perfectly are luminal wave trajectories (as opposed to the superluminal internal trajectories of an underlying, helically propagating substructure). We know this from the fact that the Dirac equation is a Mechanics equation that describes luminal wave systems and the field movements in that equation come from the constant modulus of the velocity operator, $\overrightarrow{\alpha}$, which is equal to c not $\sqrt{2}\,c$. We know it also from Special Relativity, where the little group is driven by the Lorentz factor, which involves the characteristic velocity, which is the velocity of field linear momentum in the model as opposed to the velocity of substructural fields that only exist at a lower hierarchical level ("beneath" the luminal wave ontology of the physical model here).

In the photon model, as far as angular momentum is concerned, we can ignore the movement of a field resonator through space at c, which produces no angular momentum. Instead, the circular internal binary field movements produce the angular momentum quantum, spin 1, and these are movements of wave inertia on a circle at speed c. The magnitude of this rotating momentum component is not relevant because the product pr is invariant.

On the other hand, in the electron model, we have movements of wave inertia on the surface of a sphere at speed c. If they were movements on a circle, which is to say circumferential movements, then given the scaling properties, one would anticipate the same angular momentum, spin 1, for the electron, but such a particle would not be properly 3-Dimensional and does not in fact exist.

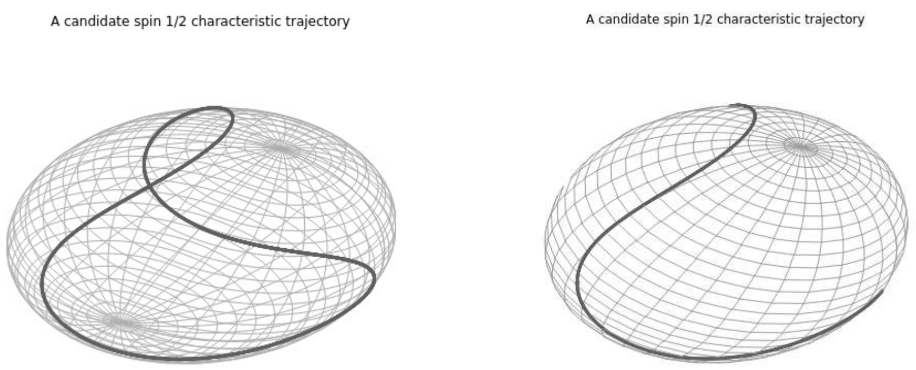

Figure 12.1: *Two views, with and without hidden lines showing, of a trajectory on a sphere for which the angular momentum is half that for a circumferential trajectory. It somewhat resembles the seam on a tennis ball.*

Instead, the trajectory should be curvilinear so as to maximally occupy the surface of the whole sphere, and not just be confined to a circumference. The characteristic trajectories will be curved on the locally flat surface of the sphere.

Any such curvature obviously reduces the angular momentum associated to the trajectory, because (a) the directions of angular momentum contributions from all the various trajectory segments are not aligned in the same direction and (b) in order to satisfy the same resonance condition, the dimension of the sphere, r_s, must be reduced. Such curvature also introduces a new frequency into the system, which comes with a new resonance condition. It must be an integer multiple of the base frequency that corresponds to full transits around the sphere.

As discussed in Section 6.1, this new frequency shows up in the Dirac Equation as the zitterbewegung frequency, which is twice the usual particle frequency from $E = \hbar\omega$. Instead of writing $E = \hbar\omega$ one might consider writing $E = (\hbar/2)\omega_z$, where $\omega_z = 2\omega$ is the zitterbewegung frequency (at least for the comoving frame). The fact that the angular momentum associated to the luminal wave trajectory is divided by an integer, 2, when the context is changed from movements on a circle to curvilinear movements on a sphere now appears plausible.

The question remains whether it can be shown that closed trajectories do exist that satisfy the double resonance condition and also reduce the angular momentum of a circumferential trajectory by a factor of 2. Even without detailed physical field solutions to the Dirac Equation, this limited question can be addressed kinematically, with a simple numerical model that draws trajectories on the surface of a unit sphere.

At each step in this model, the trajectory moves a fixed incremental distance on the surface of the sphere, and the direction of the trajectory is then rotated in the tangent plane in accordance with a sinusoidal transverse acceleration of variable amplitude and frequency.

Various closed, repeating paths that reduce the angular momentum by a factor of 2 and also exhibit 2 geometric cycles for each transit around the sphere were identified. A trajectory that satisfies the additional constraint that the path should not cross itself is shown in Figure 12.1. It resembles the seam on a tennis ball.

I make no representation whatsoever that this picture has anything to do with the physical structure of an electron's field resonators. On the other hand, it is encouraging that there are at least some visualisable possibilities for the internal energy flows in a particle of matter. We can begin to picture an entire physical reality that reproduces the observables as we find them and shares all the same Physical Laws.

Chapter 13

Reviewing The Foundations

13.1 Results, Assumptions and Lacunae

The foundations of Modern Physics have been reconstructed from a single physical idea, the wave postulate:

"Whether it appears as matter or as radiation, energy-momentum always propagates at the characteristic velocity, c, in a physical, existential medium".

Rather than three different theories based on three different sets of asseverations - principles, axioms, postulates, whatever - the entire foundations have been induced from the definition of momentum as inertia times velocity.

The central motivating assumption, energy propagates at c, is not some wild, counterintuitive, oxymoronic neologism. It was already a widely accepted consequence of the existing Theory, although its explanative power was perhaps lost in a mathematical fog. Where previously there were paradoxes and pathologies, with the wave postulate, there are none. Where there were tensions between any two of the foundation theories, there is unity. Where oxymorons had banished understanding, there is a new clarity that invites us all to seek to understand Modern Physics within a visualisable, intuitively accessible framework. The mathematics is as basic as the idea from which it sprang, as is appropriate for foundations.

As for results and assumptions, relativistic Mechanics, Lorentz covariance, the Dirac Equation, de Broglie waves, the time independent and the 1-body time dependent cases of the Yilmaz theory of gravity have all been derived without approximation from the definition of momentum for luminal waves: momentum density equals inertia density times velocity. $\vec{\rho}_{\mathbf{p}} = \rho_m c\,\hat{\mathbf{k}}$. This was the main assumption.

The work integral relation between momentum and energy changes was also assumed. Quantisation of angular momentum was easily explained but it was not quite completely shown and conservation of the angular momentum in a gravitational field was assumed. The existence of some convenient relationship between the energy density and variations in the characteristic velocity was a guess, but the actual relationship, in the form of Gauss's Law, was shown from a well known Lagrangian that involves the free choice of the function, $g(\epsilon) = 1/\epsilon^2$, multiplying the scalar wave term.

Those are the only assumptions involved. There are then three *apparent* lacunae:

First, only the passage to the quantum mechanical wave equations was shown; the remainder of the quantum formalism was not developed but it was observed that these are linear equations and the power of linear algebra facilitates the formal interpretation by connecting observables with eigensolutions. It was shown that the Schroedinger Equation has no physical wave interpretation and the idea of wavefunction collapse must be considered as a utilitarian device for making calculations.

Importantly, when physicists came to the Dirac Equation, the original probability interpretation of the wavefunction was found inapplicable, and the projection postulate seems intuitively reasonable in the context of an underlying luminal energy density. Although they were guided by prior art, physicists developed a standalone new formal interpretation to extract observables from the Dirac equation, and the wavefunction representation of observables is part of that. This is not a lacuna. The work has already been done so there was no need to reinvent the wheel here.

Second, none of the above results explicitly covers Classical Physics while (at the foundation level) Modern Physics is not a complete departure from Classical Physics. In particular, Quantum Mechanics routinely writes in a Classical 4-potential. Given Lorentz Invariance, and all the preexisting analyses on transformations of the fields of the moving particle, this residual presumption in Modern Physics was reduced to the question of modelling Coulomb's Law, and the result is not perfect. There is a small, unobservable approximation involved and we cannot say if it is in the model or in Classical Physics. This adjustment stems directly from the use of the point particle assumption in Classical Physics. As has repeatedly been pointed out, this is a deeply flawed assumption that causes many, far more serious problems in Classical Electrodynamics.

Upon removing the point particle assumption, the *ad hoc* Classical model here did however address a major lacuna in the usual retarded field model with respect to energy fluxes, as discussed in Section 9.1. We no longer need to sweep this serious problem under the carpet.

Coulomb's Law is the only part of Classical Physics where the point particle Metaphysic impinges on Modern Physics, which is otherwise "clean" in this respect: there are no point particles in Special Relativity or the Dirac equation, and Yilmaz gets a clean result because one can take the limit as $r \to 0$ without causing an infinite divergence, as was shown at (11.17).

Overall, if there is a lacuna here, it is not in the physical model. There is a residual gap in the foundations themselves, although it is satisfactorily addressed in Quantum Electrodynamics.

Third, no model of the N-body time dependent case in gravity was provided. The physical model becomes unduly complicated, whereas the idea is to simplify. That said, there is every reason to anticipate a manifestly covariant extension of the mathematics, as is all but guaranteed by the existence of the time dependent Yilmaz theory. The distinction between Yilmaz theory and a luminal wave refractive medium model was identified: Yilmaz theory has a strong equivalence principle for mass, and does not avail itself of the full set of metric transformations implicit in the wave postulate. We showed that this probably causes problems for Yilmaz in mixed Electromagnetic-gravity cases.

That is the extent of the lacuna in gravity. Since gravity theories run on geodesics anyway, there seems to be no good reason why Yilmaz's mathematics should not be able to take advantage of the full set of metric transformations. While this definitely remains as a substantial gap both in the physics and in the model, I trust that it will be seen as an opportunity, not a problem!

13.2 The Einstein Field Equations

Meanwhile, the dominant theory of gravity, using the usual Einstein Field Equations, has no luminal wave model. Although it can solve for metric solutions with notionally continuous Cosmological mass distributions, it cannot begin to formulate an N-body problem because it is a test particle theory. Yilmaz, by contrast, does formulate N-body dynamics in closed form.

In a pure field theory, the source term cannot be absolutely confined, either to a point or to a bounded region, without causing problems. Identities are broken, inelegant results emerge and the present Theory exhibits pathologies.

The lesson in this is the same as with the failure of Classical Electrodynamics in the Two Body problem: When we put a point source into a field theory by hand, chaos ensues.

Einstein's only mistake in gravity was to have done the physics too early, before the mathematics was fully mature. He was thirty years ahead of the Freud identity. Yilmaz's was to do the work too late, too long after Einstein's theory had come to be taken as fact[1]. These are the only conceivable reasons why the Yilmaz theory, which is clearly superior on the merits, is not the orthodox model at the present time. Having said that, these are not really different theories, but the same theory with Yilmaz making a small correction to the source term.

13.3 "Massless" Particles

While there are no substantive empirical conflicts between the physical model and the Physics, there are semantic tensions that stem from the very different Metaphysical frameworks - explicit here but unspoken in Physics.

One of these concerns the idea of "massless" particles. In the usual reasoning, it comes straight from the Lorentz factor, γ, and the relation $p = \gamma m_0 v$. If m_0 is a finite constant, then since $\gamma \to \infty$ as $v \to c$, the particle momentum, p, must also go to infinity. Since photons have finite momentum they must have zero mass, and hence the curious notion of massless particles. That is the usual reasoning and note that it moves from a relationship for matter particles to one for photons.

The suggestion is that 0 mass times ∞ Lorentz factor gives us a finite momentum, but the end result is that a physical object with zero inertia moving at finite speed is given a finite momentum. This should be regarded as unsatisfactory.

We did not however *deduce* or *induce* the luminal wave momentum relation, $p = mc$ (more generally $\vec{\rho}_{\mathbf{p}} = \rho_m c\hat{\mathbf{k}}$), from the matter relation $p = \gamma m_0 v$. Instead, we *derived* the latter from the former. The rest mass in the model, M_0, is then an epistemological construct, based on an ontological concept of

[1]Einstein published in 1915, the Freud identity was found in 1939 and Yilmaz only began to publish in 1958.

wave inertia, ρ_m. Naturally, this ontological m also shows up as an observable in the case of photons, when we measure the momentum and divide by the characteristic velocity, c as the definition of momentum implies. There is no conflict in the idea that a photon has inertia $m = p/c$, the model is self-consistent and the distinction with respect to the orthodoxy is merely semantical.

13.4 Curved Spacetime

The purpose of this entire exercise with physical models was to show that modern physics can be comprehended in terms of simple $3D + t$ physical models. The point has been made, well and truly.

The curved spacetime Metaphysic only came into the culture because certain pieces of mathematics happened to provide suitable techniques for implementing coordinate independence, especially the relativity principle and the weak equivalence principle. Like taking a sledge hammer to a peanut, this mathematics far exceeded its purpose, resulting in a plethora of spurious consequences from disagreements about the rates of clocks to forbidden temporal loops and from black holes to worm holes and warp drives (all you need is negative energy).

One need not read so much into coordinate independence, not when there is a simpler alternative explanation, one that can actually be understood.

The approach to flat space here was to consider waves propagating at c in a physical medium. The lengths of rulers and the rates of clocks were shown to be variables that depend on their conditions of motion. It was then straightforward to induce the usual Lorentz covariant formalism.

The approach to curved space here was to treat the defining property of this $3D$ physical medium, the characteristic velocity, as position dependent. The lengths of rulers and the rates of clocks then also depend on their positions in relation to massive objects. Geodesic paths in Physics' curved 4-space correspond to refraction phenomena in ordinary space and so on. This approach covered all the phenomena without predicting unicorns[2].

This is not to say that we should not use the 4-space approach in Physics, only that we should do so with a modicum of discretion. Physicists should use whatever approach is most effective for making calculations. On the other hand, the Metaphysical connotations that have the entire world bemused can, and should, all be dispensed with forthwith.

The word "spacetime" is an oxymoron where the idea of extension is equated with the idea of change. Inherently, such words do not make sense. Spacetime is literally non-sense, and the idea of a curved spacetime is nonsense on stilts.

This is not just a matter of taste. Once we began to engage with ideas that could not be understood, tangible errors crept in. They had to be allowed in to preserve the unintelligible ideas that had already been accepted. The ability to discriminate between science, science fiction and mere fantasy was undermined.

13.5 Interactions don't Travel between Particles

The single most important case where the lack of clear language has been inhibiting progress is retarded interaction. Of all the mysteries, this one was the

[2]Could that be the weakness?

most important to unravel because it seems so indispensible and so independent of the usual Physics. Let us review the development of the worst mistake in Physics.

We see one object here and another there. That's Epistemology, but if we conclude that the physical reality corresponding to the observation is similarly limited in its spatial extent, that's a Metaphysical conjecture. When we proceeded to a Classical Physics based on the point particle concept, the conjecture came to be taken for fact.

It might be said that the point particle was the sustaining myth for Classical Physics. It is not that we should not have brought it in, for it has been useful, but that we should have done so with our eyes open.

We see that the two objects interact with each other at a distance. If the myth were true it would follow that something had to travel from one to the other. We did examine that idea critically long ago, well before Special Relativity. It was always found wanting, but we accepted it anyway because the myth had long been taken for fact.

A very nice Physics theory, Special Relativity, is then developed on the basis of an epistemic principle that we do understand quite well - the relativity principle. The Theory is simple, it works very well and the postulates are exclusively in the domain of Physics, Epistemology. The Theory blurs our attempts to understand why the observed speed of light is constant, but that kind of question was now placed outside the domain.

We no longer understand the Theory. It contains obvious paradoxes. For many decades, the mathematics shows that the paradoxes do not refute the Theory, so we move all our considerations to a formal domain, where the reason for the Mechanics of particles is now the mathematical isomorphism of the Minkowski metric to spacetime rather than *vice versa*[3]. There is no longer any territory, only the map.

The Theory contains a speed limit. What do we do with it? We have just been stripped of every possibility to think outside the formalism because common sense no longer makes sense. The only thing left is the original Metaphysical conjecture of a world spatially separable into distinct particles, one that is here and not there, the other that is there and not here. With no proper reflection, we invent relativistic causality and light cone causal analysis, thereby imposing our Metaphysical prejudice on a nice Physics Theory that had been independent of it.

More decades will pass until a new Theory is developed to the point where it removes the point particle Metaphysic altogether — Quantum Fields become well established. What had been particles are now distributed fields. Whenever we find them, of course we find them somewhere specific but how is that not the nature of the act of finding as opposed to the nature of the entities themselves? Clearly, the point particles belong to Epistemology, not Ontology. Importantly, the interactions in the Quantum Field Theories have nothing to do with retarded interaction or light cone causal analysis.

Do we go straight back to Special Relativity and remove the Metaphysical baggage that we put in by hand? We could, but of course we do not. The entire history prevents it.

---

[3]This idea is so obviously preposterous that it must be pointed out that it comes, verbatim, from a prominent physicist.

We know for sure that there is no Metaphysical conjecture in Special Relativity. The formalism is no longer a model of the physical world, we stopped thinking that long ago. Instead, the territory is the map, and the map is as good as Mathematics, which is to say it is perfect. To discriminate between the core Theory and the lightcone causal analysis that we bolted onto it requires an act of physical reasoning, a tool long ago discredited.

The idea of Quantum Fields preceded John Bell by decades, but when he developed the Bell Inequalities, and the experiments were performed, did we identify the Metaphysical conjecture in the light cones bolted on to Special Relativity for causal analysis? Of course not. We talked about non-Kolmogorov probability distributions, the reality of unobserved observables, many Universes, causes that operate into the past and our lack of free will. Testability? No longer relevant.

Finally, we deleted the question altogether by redefining causation. We had to keep the retarded part, because that is the part that Special Relativity provides, isn't it? (It isn't). Instead, we dispensed with causation itself: We cannot talk about cause and effect, that's Metaphysical. Who are we to say? We must limit ourselves to talk about signalling. That much is in our domain.

The insane result, which is the *status quo*, is that any interaction between two particles is retarded, but cause and effect relations between point events are outside the domain. Meanwhile, the retarded interaction between charged particles of opposite sign moves between them in the wrong direction, and comes in point particles that occupy all of space.

All the information has been available for decades to identify the source of this Metaphysical nervous breakdown with Newton, not Einstein: He told us that Metaphysics is no good and then he told us that Action and Force must be conveyed *from* one "body" *to* the other. That's Metaphysics. The father of Physics broke his own rule, and (unsurprisingly for the late 17^{th} century), he got it wrong.

Physics has succeeded so well in the goal to remove itself from Metaphysical speculations that it has lost the ability to process physical information. Their heads are spinning so fast defending Special Relativity that none can see why it has no need for lawyers - it's a field theory.

The question of cause and effect, with or without signalling, is not outside the domain, as the redefinition of relativistic causality would have it. Physics may have put it in a box labelled "mysteries", but it refused to stay in the box:

Quantum before-before paradoxes, especially the double Bell paradox, were the final step in this centuries long drama. They have been known for over twenty years.

When the theory predicts that a cat can be alive and dead *while we are looking at it* (see Section 7.9), something is wrong. Peaceful coexistence? Relativistic causality as signalling? Complementarity? No semantic artefice can avoid the conclusion. When it comes to the single most important foundational issue of all, Physics has failed.

This is not a mathematical failure. We saw how to interpret the mathematics. It is a failure of the program as a whole. When Newton said "I frame no Hypotheses", he was not just deluding himself, but generations to come. Metaphysical content is inevitable in Physics and our attempts to suppress it are the leading cause of the current dilemma. It is precisely the exclusion of Metaphysics that has led us, time and time again, to unintelligible Metaphysical conclusions.

We make Hypotheses. It's what we do, and the hypotheses into which Physics is descending in modern times are growing increasingly absurd precisely because we fail explicitly to consider Metaphysical content. We have, in fact, advanced to the point where there can be a useful interplay between Metaphysics and Physics. If only we do it with our eyes open, we are able to discriminate between alternative explanations on the basis of the Metaphysical "costs" involved: which intuitively reasonable ideas are being sacrificed? This is the part of the game that we have not been playing.

Chapter 14

Epilogue: After Reductionism

> *"There are more things in heaven and earth, Horatio, than are dreamt of in your philosophy."* - Hamlet to Horatio [106].

Having taken a strictly reductionist approach throughout, moving from Ontology to Epistemology, in this final brief chapter let us consider what the physical model has to say about the realist philosophy in general, but also about several unresolved questions that were raised in various places throughout the text.

Of course, a model is just a model. One should not take its Metaphysical consequences too seriously, but nor should one ignore, in particular, a core issue that permeates the interpretation of Quantum Mechanics, namely the role that "in principle available information" plays in Quantum Physics. The discussion on this issue below is largely conjectural, but perhaps the Metaphysical problem that it raises can at least be stated plainly, and the possibility to resolve it within the framework of the present model noted.

The origins of the physical model lay in the failure of the retarded interaction paradigm with the Two Body problem in Classical Electrodynamics. In Chapter 8, we saw that this paradigm fails qualitatively to provide a credible explanation for interaction at a distance; that it fails quantitatively to provide a satisfactory mathematical account; and that it has failed experimentally, not just with respect to spacelike causal correlations but also with precision gravity measurements showing that the Earth accelerates towards the instantaneous position of the Sun, not the retarded position [88]. It fails on every measure.

However, if one accepts an ontology of point particles, there can be no intelligible alternative to retarded interaction. Rather than accepting the modern consequences that can never be understood, and in any case fail to address the question of interaction at a distance, we asked what, if anything, we can say with confidence about physical reality. Two facts seem pertinent. First, there are entities that propagate at c. Second, particles of matter have the energy content $E = mc^2$.

The additional observed fact that each of these forms, matter and radiation, is convertible into the other questions the need to think of radiation, which propagates at c, and matter, which moves subluminally, as distinct Ontological classes. After all, if the luminal/subluminal distinction is intrinsic, what is the nature of the transformation?

From the point of view of Metaphysics, one must at a minimum assert an ontology that provides movements at the fixed speed, c. Then either one must address the question of what it means for the intrinsically luminal to transform into the intrinsically subluminal or one must explain subluminally moving phenomena within a framework where all the ontology propagates at c.

These considerations all point to the idea of a distributed wave ontology, in which case an alternative, distributed approach to interaction at a distance becomes available. This path led on to the complete mathematical foundations of Physics and the physical idea of quanta as inherently distributed systems that evolve under distributed interaction.

Unlike the retarded version of interaction at a distance, distributed interaction can be understood: The whole story of the interaction between two quanta was told without violating the Principle of Local Action, the Conservation Laws, or the language, and we emerged understanding that instant causal relations between point events are entirely plausible in "local realism", once this term is properly defined. To the extent that the mathematics of Quantum Electrodynamics can be interpreted, the Electromagnetic interaction is compatible with the distributed interaction framework, while QED can never be reconciled with retarded interaction.

The model, as presented, obeys every rule of common sense proposed in Section 2.3. It was as materialist and as reductionist as Classical Physics. We arrived at an internally consistent, bottom up construction from energy that propagates at c, to subluminal matter and on to the observable reality in which we find ourselves.

However, along the way, we encountered a number of consequences that raise many philosophical issues already familiar in Modern Physics, including:

1. In Chapter 4, it was shown that the momentum density distribution, the mass/energy, the rate of internal evolution and the physical shape of a quantum depends on its condition of motion.

2. Section 5.5 finished with an open question regarding the role that information is playing in Physical Theory.

3. In Section 6.2, on matter beam interferometry, we encountered the context dependence of the physical shape of a quantum's energy density distribution.

4. In Section 7.4, when Alice measures the spin of her electron, we required local interactions between resonator cells of the measured and measuring systems to occur at Bob's location at the same moment when they occur at Alice's location. An example of a local realist protocol that produces the result was given, where the global alignment between two quanta is detected locally. It was a toy model, not to be regarded as satisfactory, but it shows that local realism is no logical barrier to a distributed action interpretation of spacelike causal correlations.

   The fact remains however that, in all our Physics theories, the quanta come in wholes and they interact as wholes.

5. In Section 7.7, we questioned the idea of a real physical interference between two photons only one of which exists. Nonetheless, for the protocol to succeed, we cannot know, or be able to know, which side the detected photon came from. Photons from the two sides of the apparatus must be "indistinguishable".

6. In Chapter 8, we noted that in Quantum Electrodynamics the usual interpretation of an interaction (as a momentum eigenstate that occupies all of space) has two physical interpretations in a wave ontology. It can be interpreted as a process involving individual point like interactions at the resonator level, which occur independently of each other. On the other hand, taking the mathematics literally, the interaction can be interpreted as a single, widely distributed, whole.

7. In Section 11.4, we encountered a non-separable N-body energy density that was a property of the system as a whole. We could identify all the parts of the system as medium independent fields, one for every quantum in the system, but we could not assign an energy content to any of these fields on its own. Although there were distinct, identifiable parts, the physical instantiation as energy was an indivisible whole.

   In principle, the source terms in an N-body problem could only be defined "perfectly" from the point of view of Epistemology, as opposed to Ontology. This is ontological non-separability coexisting with epistemological separability.

Again, the map is not the territory and a model is just a model, never to be taken literally, but these consequences of the model do seem to inform two deep questions running through all the observed quantum phenomena and our attempts to interpret the modern Theory.

First is the question of primary versus secondary properties.

As far as primary, or intrinsic, properties are concerned, the model has dispensed with all the usual suspects. Every physical property of the quantum, except its angular momentum[1], was mutable. All this plasticity was nonetheless compatible with the fact that indivisible quanta persist throughout a wide range of changing contexts.

Of course, there are high speed collisions, absorptions and various other transformations from one set of quanta to another, so the quanta are ultimately impermanent but they persist over extended periods while their physical instantiations are relatively ephemeral: The actual physical implementation of a given quantum depends on its condition of motion, boundary conditions / dynamical interactions, the mere fact of other quanta being in the vicinity and the physical state of the local space, the "quantum vacuum", itself.

The physical instantiation of a quantum must therefore be regarded as incidental.

Second is the question of the relationship between parts and wholes. Since physical instantiation is incidental, the whole (quantum) is as fundamental as

[1]Angular momentum invariance was understood in two ways: (a) within each of the classes of fermions and photons and (b) with respect to changes in the physical instantiation of any given quantum.

its parts. Moreover, since all of the ontology is to be found amongst the parts, the quantum cannot have a primitive ontological status in this model[2].

If the quanta have a primary status that is not ontological, what kind of status is it? First, from the point of view of Ontology, they are systems. Second, from the point of view of Epistemology, they are findable. They are the basic elements of Epistemology in the sense of being associated to observable quantities. On one hand that which is epistemologically primitive is already a system of underlying ontological activity while, on the other, these indivisible, observable wholes clearly play a vital role in the underlying physical structure.

The role cannot be limited merely to what a quantum does. Even as the context in which a quantum exists is physically constructed from the bottom up, the context as a whole has a role in determining the physical structures of individual quanta from the top down. The main question here concerns the nature of that role.

Rather than the usual bottom up construction assumed throughout, where Ontology produces Epistemology, let us begin by recognising the dialectic relation. In a sense, observable wholes are on an equal footing with the underlying physical implementation, each one giving rise to the other. Of course, this corresponds to the situation in Quantum Mechanics, where the question of in principle available information - which is always information about wholes - comes into so many calculations.

Physics is traditionally reductionist, and the foundations were constructed here in a reductionist manner. Every quantum is constructed from its myriad resonator parts. The parts construct wholes, and the global behaviour of the whole *can* be analyzed as an aggregate of the behaviours of well-localised parts, each one acting locally in accordance with local Physical Laws. The entire hierarchy from field resonators to galaxies *can* be constructed from the bottom up. It seems logically essential that a *sufficent* account of the entire reality must exist along such reductionist lines.

However, we might complete the reductionist program, a "bottom up" theory of everything if you will, without ever engaging with the question "how does that work?", especially in relation to the role that information is playing in Modern Physics. We could have a *sufficient* explanation for everything as opposed to a *satisfactory* one.

Here is a remark from a highly accomplished zen buddhist monk of the twentieth century that has always stood out to me as capturing the essence of the matter. Although I can no longer find the exact source, I believe that this is sufficiently close to his exact wording[3]:

> *"The first time I had satori[4], I knew that the mountains are not mountains and the trees are not trees, but now, after many decades and thousands of satori, I know that the mountains really are mountains and the trees really are trees."*

We are reminded of one of Wheeler's oxymorons [107]: "It from Bit". That's not a dialectic, he went too far; The quantum is observable, but no reason has

[2]This is not as model dependent as it might seem: If one attempts a model where the quanta are the ontological primitives, the observable facts (corresponding to the above list) lead one to introduce substructure, and hence to the conclusion that the quanta are not ontologically fundamental.

[3]It's actually a fairly common zen metaphor.

[4]Enlightenment, awakening, understanding, the direct experience of reality.

been identified to think that its existence results from being observed. So let us rephrase:

$$\text{It} \Leftrightarrow \text{Bit}.$$

On the face of it, the notion that information *per se* impacts physical reality would imply a non-physical mechanism, with all the attendant Metaphysical costs: "Common sense" demands that only physical mechanisms should operate in physical systems, so let us consider what can be introduced physically in order to avoid non-physical mechanisms that clearly lie outside the spirit of the modelling criteria in Section 2.3.

Consider three kinds of facts: Observed facts, observable facts and objective facts. Observed facts are the facts that some observer has actually seen. They are a subset of the observable facts, which could, in principle, be seen were an observer to look. These in turn are a subset of all objective facts that exist in nature, including many that cannot be observed.

To illustrate this, consider the measurement problem. When the quantum to be measured first interacts with the detector, there is a nonlinear interaction at the microscopic level which has a definite outcome, which produces an objective fact in the present model that cannot as yet be observed. There is then a cascade of interactions that amplify the signal leading to an observable fact, which finally becomes an observed fact when someone actually looks.

There is little distinction between observable and observed facts in the application of the quantum formalism. For example, if "which way" information is available in a two slit experiment, the interference fringes disappear whether or not an observer reads the information. Consequently, the focus here can be on the distinction between observable and objective facts.

Consider next two electrons in the singlet state. To the extent that the ontology is objective, as it is in the model, there are objective facts in relation to the angular momenta of the members of the entangled pair. These objective facts are not observable. This is distinct from the measurement problem, where the objective fact about the result of a measurement interaction is in the process of becoming observable. With the singlet state, whatever objective facts exist in nature about the angular momenta of the members of the entangled pair, they can never be observed. Say the internal state contains information about what the result of a spin measurement by Alice at a given moment along axis a would be. Then it also contains information about the counterfactual result, if she were to measure along a' at the same moment.

Clearly, there is information instantiated in the physical reality that we cannot obtain and there is a distinction between the objective facts that can be observed and those that cannot. Recall that it is the former category that modifies wavefunctions and washes out interference patterns in Quantum Mechanics.

As for the physical basis for this distinction, what we observe are point events. In order to count as an observable fact, an objective fact must be findable, and to be findable, it must be associated to a point-like physical instantiation. Findability requires a sufficiently well-localised wavefield to trigger an irreversible nonlinear interaction with some element of a detector. However, all objective facts (including observable facts) are physically instantiated by widely distributed field systems so, at least in principle, the objective facts are everywhere although the observable facts are, for us, localised.

In the Barrett-Kok case of interference between two possible ways for a single photon to be emitted, it was noted that the photons on each side had to be "indistinguishable" for the entanglement to form in the first place, but this is not enough. They must remain indistinguishable at detection, because otherwise the photon detection would constitute a spin measurement of the electron that emitted it. According to the Theory, creating that information would destroy the entanglement.

This is exactly the kind of apparently non-physical action one wants avoid if at all possible. Perhaps Redhead's nonseparability hypothesis [73] can avoid it, as follows.

Let us take a step back and consider again the notion of the split reality photon in Section 7.7. Under Redhead's hypothesis, the 2-particle singlet state is a distributed, nonseparable whole. When the microwave pulses that stimulate any spin $< UP >$ electron into the excited state are applied, they are applied to the system as a whole, which absorbs and then later emits, a single photon. The excited state is then a state of the system as a whole rather than of the individual electrons, so the state space of the system is modified by entanglement. Given the usual Hamiltonian time evolution of quantum systems in general, there is no basis to infer the existence of a stable objective fact of the matter regarding the spin state of either electron upon detecting the emitted photon.

On the other hand, if the apparatus does identify which side the photon came from, then we can infer that the stimulated state, *i.e.* before the photon was emitted, was a state of an electron on one side, rather than a state of the system as a whole. To see the physical basis for this distinction, consider an apparatus with a separate photon detector for each side. Such an apparatus cannot respond to states of the entangled system as a whole. It can only respond to physical states of one or other of the parts considered separately. We can now begin to think of the role that information plays in the usual interpretation as an artefact of a theory that runs on information as opposed to its distributed physical instantiation.

Thinking of the 2-particle entangled system as an indivisible whole can thus avoid the non-physical implications of the quantum formalism. When quantum mechanical experiments confront us with cases where the availability of information has observable consequences, this notion that physical wholes exist whose parts may seem, from a macroscopic perspective, to be spatially separate, but which are in fact distributed systems that are in physical contact with each other, may well be the only way to avoid the dilemma that a non-physical mechanism would pose.

A final, macroscopic example is perhaps as confronting as the Barrett-Kok photon. Consider the facts regarding crystal polymorphisms.

Many compounds can form multiple kinds of crystal lattices, which is to say they have polymorphisms. A new atom bonding to the lattice can be facing identical *local* conditions in which it might bond in one of two ways, corresponding to two different kinds of lattice structure. It chooses whatever bond corresponds to the pre-existing lattice. The whole matters and so the construction is as much top down as it is bottom up.

Strangely, polymorphisms can also disappear. Sometimes new, more stable, crystalline forms of a compound arise in nature for the first time, and *the original crystal form becomes harder to grow* [108, 109]. The new form takes over. This has had practical consequences in the pharmaceutical industry. The HIV drug, Ritonavir, is a very well known example [110] where the new, less soluble, crystal structure was less therapeutically effective and since they were no longer able to grow the old structure, the drug had to be taken off the market!

We saw above how taking a "system as a whole" viewpoint in the context of experiments at the quantum level avoided making recourse to non-physical mechanisms. These polymorphism cases point to system level fields: Nonlinear, interacting field superpositions that reflect the lattice structure as a whole. The global field has physical properties that are not to be found in any of the parts and the physical instantiation of the system level information is inherently widely distributed.

While a tree may not be what it seems, perhaps it really is a tree after all.

Chapter 15

Appendices

15.1 Appendix A: Clock Synchronisation with Subluminal Projectiles.

Consider the synchronisation protocol shown in Figure 15.1, where projectiles are fired from the midpoint between two clocks at the speed w, as seen in the inertial frame of the clocks. Lorentz observers in other inertial frames don't agree that these projectiles reach the clocks at the same time. In particular, an observer for whom the relative velocity is v sees the projectile speeds as:

$$w'_{-} = \frac{v-w}{1-vw/c^2} \qquad ; \qquad w'_{+} = \frac{v+w}{1+vw/c^2}$$

in the directions antiparallel and parallel to the relative velocity respectively. This is the relativistic composition of velocities result. Since each of the clocks is moving at speed v in his frame, the closing speeds are respectively:

$$cs'_{-} = v - \frac{v-w}{1-vw/c^2} = w\frac{c^2-v^2}{c^2-vw} \qquad ; \qquad cs'_{+} = \frac{v+w}{1+vw/c^2} - v = w\frac{c^2-v^2}{c^2+vw}$$

The length of the ruler in his frame is L/γ, where the factor of gamma is due to the Lorentz contraction of the moving apparatus, so that the difference in the arrival times of the projectiles in his frame is:

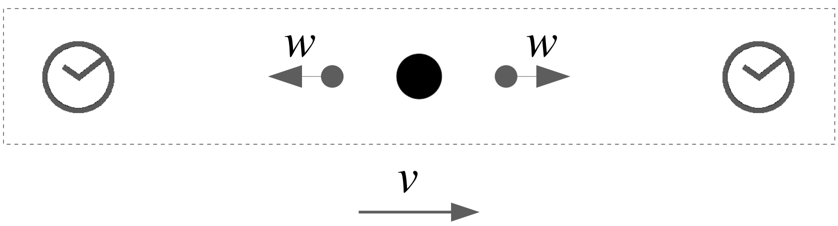

Figure 15.1: *Clocks are synchronised in one inertial frame by sending projectiles from a central station at speed w. The entire setup is moving at speed v in the primed reference frame.*

$$\Delta t = \frac{L}{2\gamma w}\left[\frac{(c^2+vw)-(c^2-vw)}{c^2-v^2}\right] = \frac{vL}{\gamma}\frac{1}{c^2-v^2} = \frac{\gamma vL}{c^2}\ ,$$

which is independent of w. Meanwhile, according to the Lorentz Transformation the usual transformation of time is:

$$t' = \gamma t - \frac{\gamma vx}{c^2}$$

and we can see from the last term on the right hand side that the above result, $\gamma vL/c^2$, is identical to the space dependence of time at points separated by $\Delta x = L$. Therefore, *according to Special Relativity*, synchronising clocks with projectiles is equivalent to the usual Einstein protocol with light signals.

However, we can go further and drop the words "according to Special Relativity". As discussed in Sections 4.7 and 4.9, the same conclusion follows from the Michelson-Morely result for the 2-way velocity of light and time dilation. Similarly, it follows from length contraction and time dilation, which were derived independent of both Special Relativity and clock synchronisation in Sections 4.5 and 4.6 respectively.

Synchronising clocks with subluminal projectiles is therefore equivalent to the Einstein protocol.

15.2 Appendix B: Light Reflected by a Moving Mirror

15.2.1 Normal incidence, Work Method

Consider a constant momentum density $\vec{\rho}_{\mathrm{pi}}$ in a region of transverse crossectional area A and length l_{i} as shown in Figure 4.1. The total momentum is $\mathbf{p}_{\mathrm{i}} = Al_{\mathrm{i}}\rho_{\mathrm{pi}}\hat{\mathbf{k}}$. Let this be normally incident on a mirror that is moving with velocity $\mathbf{v} = -v\hat{\mathbf{k}}$. Let the reflection begin at $t = 0$. It then ends at $\Delta t = l_{\mathrm{i}}/(c+v)$, after which there is a reflected wave with momentum density $\vec{\rho}_{p\mathrm{r}}$ that occupies a region of length $l_{\mathrm{r}} = (c-v)\Delta t$ and crossectional area A, so the momentum of the reflected light flash is $\mathbf{p}_{\mathrm{r}} = -Al_{\mathrm{r}}\rho_{p\mathrm{r}}\hat{\mathbf{k}}$.

During the reflection, the rates of change of momentum for the incident and reflected waves are $\dot{\mathbf{p}}_i = -(c+v)A\vec{\rho}_{\mathrm{pi}}$ and $\dot{\mathbf{p}}_{\mathrm{r}} = (c-v)A\vec{\rho}_{\mathrm{pr}}$ respectively, where a dot over a variable indicates the time differential. The total rate of change of momentum is:

$$\dot{\mathbf{p}} = \dot{\mathbf{p}}_{\mathrm{i}} + \dot{\mathbf{p}}_{\mathrm{r}} = -A((c+v)\,\rho_{p\mathrm{i}} + (c-v)\,\rho_{p\mathrm{r}})\hat{\mathbf{k}}\,,$$

where $\rho_{p\mathrm{i}} = |\vec{\rho_{\mathrm{pi}}}|$ and $\rho_{p\mathrm{r}} = |\vec{\rho_{\mathrm{pr}}}|$. As far as scalar momentum is concerned, for the incident wave $\dot{p}_{\mathrm{i}} = c\dot{m}_{\mathrm{i}} = -A(c+v)\,\rho_{p\mathrm{i}}$, for the reflected wave $\dot{p}_{\mathrm{r}} = c\dot{m}_{\mathrm{r}} = A(c-v)\,\rho_{p\mathrm{r}}$ and the total is:

$$\dot{p} = c\dot{m} = c\dot{m}_{\mathrm{r}} + c\dot{m}_{\mathrm{i}} = A((c-v)\,\rho_{p\mathrm{r}} - (c+v)\,\rho_{p\mathrm{i}})\,.$$

The work done by the mirror on the incident and reflected waves is: $\int \dot{\mathbf{p}}_{\mathrm{i}} \cdot d\mathbf{s}_{\mathrm{i}} = -\int_0^{\Delta t} A(c+v)\,\rho_{p\mathrm{i}}\; cdt$ and $\int \dot{\mathbf{p}}_{\mathrm{r}} \cdot d\mathbf{s}_{\mathrm{r}} = \int_0^{\Delta t} A(c-v)\,\rho_{p\mathrm{r}}\; cdt$ respectively, where $d\mathbf{s}_{\mathrm{i}}$ and $d\mathbf{s}_{\mathrm{r}}$ are the incremental movements of the incident and reflected waves, in the directions $\hat{\mathbf{k}}$ and $-\hat{\mathbf{k}}$ respectively. The total work done is just $W = \int_0^{\Delta t} c\,\dot{m}c\,dt = (m_{\mathrm{r}} - m_{\mathrm{i}})c^2$.

The energy change of the light flash is of course equal and opposite to the work done by the radiation pressure force on the mirror, so $(m_{\mathrm{r}} - m_{\mathrm{i}})c^2 = -(-\dot{\mathbf{p}})(-v)\Delta t$, and it is easily shown that $p_{\mathrm{r}}/p_{\mathrm{i}} = (c+v)/(c-v)$, from which we may infer the momentum shift factor for light emitted by a source moving towards an observer as $\sqrt{(c+v)/(c-v)}$, because there are two identical Doppler shifts involved in the moving mirror example, first from the source frame to the mirror frame and then back from the mirror frame to the source frame.

The result of the work method thus agrees with the usual relativistic Doppler shift

15.2.2 Non-Normal incidence, Work Method

In the case of non-normal incidence, with the angle of incidence θ_i and the angle of reflection θ_r, the analysis above is modified as follows:

The closing speed between the mirror and the light flash is $c\ cos\theta_{\mathrm{i}}+v$ (instead of $c+v$). The leaving speed of the reflected light flash is $c\ cos\theta_{\mathrm{r}} - v$ (instead of $c-v$). The duration of the reflection is then:

$$\Delta t = l_{\mathrm{i}}cos\theta_{\mathrm{i}}/(c\ cos\theta_{\mathrm{i}} + v) = l_{\mathrm{r}}cos\theta_{\mathrm{r}}/(c\ cos\theta_{\mathrm{r}} - v)\ .$$

During the reflection, the total rate of change of momentum (in the $\hat{\mathbf{k}}$ direction) is:

$$\dot{\mathbf{p}} = \dot{\mathbf{p}}_{\rm i} + \dot{\mathbf{p}}_{\rm r} = -A_{\rm i}\rho_{p{\rm i}}(c\ cos\ \theta_{\rm i} + v) - A_{\rm r}\rho_{p{\rm r}}(c\ cos\ \theta_{\rm r} - v)\ ,$$

where $\rho_{p{\rm i}} = |\vec{\rho_{p{\rm i}}}|$ and $\rho_{p{\rm r}} = |\vec{\rho_{p{\rm r}}}|$. The work done on the mirror equals the energy change, $\Delta E = cp_{\rm r} - cp_{\rm i}$:

$$-(-\dot{\mathbf{p}})(-v)\Delta t = cp_{\rm r} - cp_{\rm i} = v\big(A_{\rm i}\ \rho_{p{\rm i}}\ l_{\rm i}\ cos\ \theta_{\rm i} + A_{\rm r}\ \rho_{p{\rm r}}\ l_{\rm r}\ cos\ \theta_{\rm r}\big)$$
$$\Rightarrow p_{\rm r} - p_{\rm i} = \beta\big(p_{\rm i}\ cos\ \theta_{\rm i} + p_{\rm r}\ cos\ \theta_{\rm r}\big)$$
$$\Rightarrow \frac{p_{\rm r}}{p_{\rm i}} = \frac{1+\beta\ cos\ \theta_{\rm i}}{1-\beta\ cos\ \theta_{\rm r}}\ .$$

Meanwhile, the component of momentum parallel to the mirror is unchanged, which implies:

$$\frac{p_{\rm r}}{p_{\rm i}} = \frac{sin\ \theta_{\rm i}}{sin\ \theta_{\rm r}}\ .$$

15.2.3 Normal incidence, Special Relativity Method

Introduce a new subscript, M for mirror, so that the angle of incidence / reflection for an observer on the mirror is θ_M, his momentum is p_M, and so on. The standard, relativistic Doppler shift and aberration formulae are:

$$p = p_0\gamma\,(1+\beta\cos\theta_0) \text{ and } \cos\theta = \frac{\cos\theta_0+\beta}{1+\beta\cos\theta_0}\ ,$$

where the 0 subscript refers to the source frame. Thus:

$$p_M = p_{\rm i}\gamma\,(1+\beta\cos\theta_{\rm i}) \text{ and } \cos\theta_M = \frac{\cos\theta_{\rm i}+\beta}{1+\beta\cos\theta_{\rm i}}\ .$$

Similarly,

$$p_{\rm r} = p_M\gamma\,(1+\beta\cos\theta_M) \text{ and } \cos\theta_{\rm r} = \frac{\cos\theta_M+\beta}{1+\beta\cos\theta_M}\ .$$

$$\Rightarrow p_{\rm r} = p_{\rm i}\gamma\,(1+\beta\cos\theta_{\rm i})\,\gamma\left(1+\beta\frac{\cos\theta_{\rm i}+\beta}{1+\beta\cos\theta_{\rm i}}\right) = p_{\rm i}\gamma^2(1+2\beta cos\ \theta_{\rm i}+\beta^2)$$

and

$$\cos\theta_{\rm r} = \frac{\frac{\cos\theta_{\rm i}+\beta}{1+\beta\cos\theta_{\rm i}}+\beta}{1+\beta\frac{\cos\theta_{\rm i}+\beta}{1+\beta\cos\theta_{\rm i}}} = \frac{(1+\beta^2)\cos\theta_{\rm i}+2\beta}{1+2\beta cos\ \theta_{\rm i}+\beta^2}\ .$$

From the second last expression above:

$$\frac{p_{\rm r}}{p_{\rm i}} = \frac{1+2\beta cos\ \theta_{\rm i}+\beta^2}{1-\beta^2} = \frac{(1+\beta cos\ \theta_{\rm i})[1+2\beta cos\ \theta_{\rm i}+\beta^2]}{(1+\beta cos\ \theta_{\rm i})(1-\beta^2)}$$
$$= \frac{(1+\beta cos\ \theta_{\rm i})[1+2\beta cos\ \theta_{\rm i}+\beta^2]}{1+\beta cos\ \theta_{\rm i}-\beta^3 cos\ \theta_{\rm i}-\beta^2} = \frac{(1+\beta cos\ \theta_{\rm i})[1+2\beta cos\ \theta_{\rm i}+\beta^2]}{[1+2\beta cos\ \theta_{\rm i}+\beta^2]-\beta((1+\beta^2)cos\ \theta_{\rm i}+2\beta)}\ ,$$

so that, finally:

$$\frac{p_{\rm r}}{p_{\rm i}} = \frac{1+\beta cos\ \theta_{\rm i}}{1-\beta\frac{(1+\beta^2)cos\ \theta_{\rm i}+2\beta}{1+2\beta cos\ \theta_{\rm i}+\beta^2}} = \frac{1+\beta cos\ \theta_{\rm i}}{1-\beta cos\ \theta_{\rm r}}\ ,$$

which is identical to the result obtained above using the work method.

15.3 Appendix C: Derivation of the Incremental Momentum Boost Generator

With respect to the system of light flashes in Section 4.3, let us impose the condition in some inertial frame that $\mathbf{P}_0 = \sum_i \mathbf{p}_{i0} = 0$. The momentum of the i^{th} light flash, referred to this frame, is then:

$$\mathbf{p}_{i0} = p_{i0}\left(\cos\theta_{i0}\,\hat{\mathbf{i}} + \sin\theta_{i0}\cos\phi_{i0}\,\hat{\mathbf{j}} + \sin\theta_{i0}\sin\phi_{i0}\,\hat{\mathbf{k}}\right),$$

where θ_{i0} is the angle with the x-axis and $\sum_i p_{i0}\cos\theta_{i0} = \sum_i p_{i0}\sin\theta_{i0}\cos\phi_{i0} = \sum_i p_{i0}\sin\theta_{i0}\sin\phi_{i0} = 0$.

Let an observer move relative to this frame with velocity $\mathbf{v} = -\beta c\hat{\mathbf{i}}$. Since $p_i/p_{i0} = f_i/f_{i0}$, the standard relativistic Doppler shift and aberration formulae (with the observer moving towards the source at speed v) give, respectively:

$$p_i = p_{i0}\gamma\left(1 + \frac{v}{c}\cos\theta_{i0}\right) \text{ and } \cos\theta_i = \frac{\cos\theta_{i0} + \frac{v}{c}}{1 + \frac{v}{c}\cos\theta_{i0}}.$$

Note that the same result also holds for non-monochromatic light flashes. The scalar momentum of the i^{th} flash in the observer frame is:

$$p_i = p_{i0}\gamma(1 + \beta\cos\theta_{i0}).$$

Summing over i, the total energy, $m_e c^2 = \gamma c \sum_i p_{i0} = \gamma m_0 c^2$, where m_e and m_0 are as defined in Section 4.3 and Section 4.4 respectively. Using the aberration formula with the above equation, the (vector) momentum of the i^{th} flash is:

$$\mathbf{p}_i = p_{i0}(\gamma(\beta + \cos\theta_{i0})\,\hat{\mathbf{i}} + \sin\theta_{i0}\cos\phi_{i0}\,\hat{\mathbf{j}} + \sin\theta_{i0}\sin\phi_{i0}\,\hat{\mathbf{k}}).$$

Note that this is the same as the transformation found in Section 4.5 at (4.17). Summing over i, the total momentum is $\mathbf{P} = \gamma\beta\sum_i p_{i0}\,\hat{\mathbf{i}}$. Differentiating each of the two previous equations with respect to β, we get $d\mathbf{p}_i/d\beta = \gamma^2 p_i\,\hat{\mathbf{i}}$ and $d\mathbf{P}/d\beta = \gamma^2 m_e c\,\hat{\mathbf{i}}$, so that:

$$\frac{d\mathbf{p}_i}{d\beta} = \frac{d\mathbf{P}}{d\beta}\frac{p_i}{\sum_j p_j} = \frac{d\mathbf{P}}{d\beta}\frac{p_i}{m_e c}.$$

Finally, since the above expressions for $\mathbf{p}_i$ and $\mathbf{P}$ are functions of β alone, the incremental changes can be written as $d\mathbf{p}_i = (d\mathbf{p}_i/d\beta)\,d\beta$ and $d\mathbf{P} = (d\mathbf{P}/d\beta)\,d\beta$, upon which:

$$d\mathbf{p}_i = \frac{p_i}{m_e c}d\mathbf{P}.$$

Therefore (4.10) holds for a collinear incremental boost. For transverse boosts, consider as initial condition a system with a centre of inertia that is moving in the y-direction at speed V, so $m_e = \gamma(V)m_0$. We may repeat the above analysis for an observer moving at speed v_x in the x-direction with $\beta = v_x/c$ and $\sum_i p_{i0}\sin\theta_{i0}\cos\phi_{i0} \neq 0$. Evaluating the resulting expression for $d\mathbf{P}/d\beta$ at $v_x = 0$, then yields the same result, $d\mathbf{p}_i = p_i\,d\mathbf{P}/m_e c$, for an incremental transverse boost. In Special Relativity, the general boost decomposes into a collinear boost, a transverse boost and a rotation (a Thomas precession). As the latter has no impact on linear momenta, (4.10) is generally valid for incremental boosts of systems of luminal wave momenta.

15.4 Appendix D: Derivation of Eqn. 4.12 for N=2

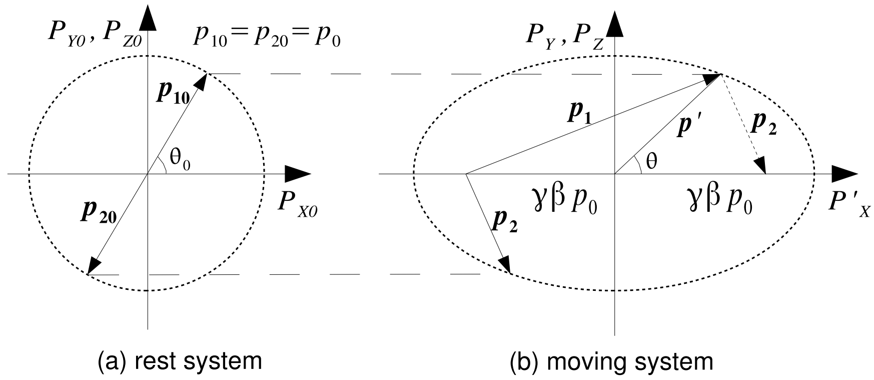

Figure 15.2: *Binary wave systems whose centers of inertia are (a) at rest and (b) moving at speed $V = \beta c$.*

In Figure 15.4a, $M_0 = (p_{10} + p_{20})/c = 2p_0/c$. Recalling from Section 4.4 that $M_e = \gamma M_0$, the sum of scalar momenta in the moving system of Fig. 15.4b is:

$$p_1 + p_2 = M_e c = 2\gamma p_0 \,, \tag{15.1}$$

whilst the total momentum, $\mathbf{P} = M_e \mathbf{V}$, is the vector sum of momenta:

$$\mathbf{P} = \mathbf{p}_1 + \mathbf{p}_2 = \frac{2\gamma p_0}{c}\mathbf{V} = 2\gamma\beta p_0\,\hat{\mathbf{i}}\,.$$

Consider the vector $\mathbf{p}'$ in Fig. 15.4b, where $\mathbf{p}_1 = \mathbf{P}/2 + \mathbf{p}'$ and $\mathbf{p}_2 = \mathbf{P}/2 - \mathbf{p}'$. Using the law of cosines, its magnitude, p', is such that:

$$p_1^2 = p'^2 + (\gamma\beta p_0)^2 + 2\gamma\beta p_0 p' \cos\theta \tag{15.2}$$

$$p_2^2 = p'^2 + (\gamma\beta p_0)^2 - 2\gamma\beta p_0 p' \cos\theta \,, \tag{15.3}$$

where θ is the angle $\mathbf{p}'$ makes with the x-axis. Upon eliminating p_1 and p_2 from (15.1)-(15.3) we find that $p' = p'(\theta)$ is the ellipsoid:

$$p'(\theta) = \frac{p_0}{\sqrt{1 - \beta^2\cos^2\theta}}\,. \tag{15.4}$$

Writing the momenta in component form as $\{p_{ij}\}_{i=1,2\,;\,j=x,y,z}$, (15.4) is then the ellipsoid:

$$(p'_x/\gamma)^2 + p_{iy0}^2 + p_{iz0}^2 = p_{i0}^2\,,$$

where $p'_x = p_{1x} - \gamma\beta p_{10} = -(p_{2x} - \gamma\beta p_{20})$, so that the moving system momenta satisfy the following equation:

$$\left(\frac{p_{ix} - \gamma\beta p_{i0}}{\gamma}\right)^2 + p_{iy0}^2 + p_{iz0}^2 = p_{i0}^2\,. \tag{15.5}$$

Which is equation 4.12.

Bibliography

[1] Brown H.R., Timpson C.G. "Why Special Relativity Should Not Be a Template for a Fundamental Reformulation of Quantum Mechanics." In: Demopoulos W., Pitowsky I. (eds) Physical Theory and its Interpretation. The Western Ontario Series in Philosophy of Science (2006) vol **72**. Springer, Dordrecht.

[2] Yilmaz H. "Correspondence Paradox in General Relativity" *Leterre Al Nuovo Cimento,* (1973). vol. **7**, N. **9**

[3] Yilmaz H. "On the "Derivation" of Einstein's Field Equations" *American Journal of Physics* (1975). Vol. **43**, No. 4, April 1975.

[4] Yilmaz H. "New Approach to General Relativity." (1958). *Phys. Rev.* **111** 1417.

[5] Yilmaz H. "Toward a Field Theory of Gravitation" *Nuovo Cimento*, (1992).

[6] Korzybski A. "Science and Sanity. An Introduction to Non-Aristotelian Systems and General Semantics." The International Non-Aristotelian Library Pub. Co. (1933). **107** 8 747–61.

[7] Bergson H. *Matter and Memory* (1899). Zone Books.

[8] Bellazini J., Benci V., Bonanno C., Micheletti A.M. "Solitons for the nonlinear Klein-Gordon equation." (2007), arXiv: math.AP/0712.1103v1.

[9] Finkelstein R.J. "A Field Theory of knotted solitons." (2007). arXiv: hep-th/0701124v2.

[10] Radu E., Volkov M.S., "Existence of stationary, non-radiating ring solitons in field theory: knots and vortons." (2008). arXiv: hep-th/0804.1357v1.

[11] Moret-Bailly J. "Electromagnetic solitons and de Broglie's double solution." *J. of Theoretics*, (2003). v. **5-5**, 11. arXiv: math-ph/0201002v1.

[12] Borisyuk D., Faber M., Kobushkin A. "Electromagnetic waves within a model for charged solitons." *J. Phys. A: Math. Gen.*, (2007). v. **40**, 525–531. arXiv: hep-th/0708.3173v1.

[13] Diaz-Alonso J., Rubiera-Garcia D. "Generalized gauge field theories with non-topological soliton solutions." (2007). arXiv: hep-th/0708.0636v1.

[14] Blas H., Carrion H.L. "Solitons, kinks and extended hadron model based on the generalized sine-Gordon theory." *JHEP*, (2007). **0701**, 027. arXiv: hep-th/0610107v2, 2006.

[15] Einstein A., Podolsky B., Rosen N. "Can Quantum-mechanical description of physical reality be considered complete?" *Phys. Rev.*, (1935). **47**, 777.

[16] Feigenbaum M.J. "The Theory Of Relativity - Galileo's Child", (2008). arXiv:0806.1234v1

[17] Einstein A. "On the Electrodynamics of Moving Bodies". (1905). Annalen Phys. **17**, 891-921.

[18] Shimony A. "Reflections on the philosophy of Bohr, Heisenberg, and Schrödinger" in *Physics, Philosophy, and Psychoanalysis*, (1983). R. Cohen and L. Laudan (eds.), Dordrecht: D. Reidel, 209–222.

[19] Shimony A. "Unfinished work: a bequest." in *Quantum Reality, Relativistic Causality, and Closing the Epistemic Circle.* (2009)., W. C. Myrvold and J. Christian (eds.). Berlin, Springer, 479–492.

[20] Puccini G., Selleri F. "Doppler effect and aberration from the point of view of absolute motion." *Nuovo Cim. B* (2002). v. **117**, 283.

[21] Laidlaw A. "Relativity and the luminal structure of matter." *Progress in Physics.* (2017), **13**, Issue 1.

[22] Selleri F. "Remarks on the transformations of space and time." *Apeiron*, (1997). v. **9(4)**, 116–120.

[23] Wigner E. "On unitary representations of the inhomogeneous Lorentz Group" *Annals of Mathematics,* (1939). **40**, 39.

[24] Kim Y.S. "Internal space-time symmetries of massive and massless particles and their unification." *Nucl.Phys.Proc.Suppl.*, (2001), v. **102,** 369–376. arXiv: hep-th/0104051.

[25] Bohm D. The Special Theory of Relativity. (1965), 23–25, W.A. Benjamin New York.

[26] Hafele J., Keating R. "Around-the-world atomic clocks: predicted relativistic time gains." *Science*, (1972), v. **177**, 166.

[27] Kundig W. "Measurement of the transverse Doppler effect in an accelerated system." *Phys Rev.*, (1963), v. **129**, 2371.

[28] Ives H., Stillwell G. "An experimental study of the rate of a moving atomic clock." *Journal of the Optical Society of America*, (1938). v. **28**, 215–226 and (1941), v. **31**, 369.

[29] Selleri F. "Non-invariant one-way speed of light." *Found. Phys.*, 1996, v. **26**, 641.

[30] Percival I. "Quantum Transfer Functions, weak nonlocality and relativity." *Physical Letters A*, (1998), **244: (6)**, 495-501. arXiv: quant-ph/9803044.

[31] Percival I. "Quantum measurement breaks Lorentz symmetry." (1999). arXiv: quant-ph/9906005.

[32] Hardy L. "Quantum Mechanics, local realistic theories, and Lorentz-invariant realistic theories." *Phys. Rev. Lett.*, (1992). **68** 2981-4.

[33] Wilczek F. and Krauss L. "Frank Wilczek & Lawrence Krauss: Materiality of a vacuum" https://www.youtube.com/watch?v=BBXDrNn6pUg

[34] Peebles P.J.E., Wilkinson D.T. "Comment on the Anisotropy of the Primeval Fireball." *Phys. Rev.*, (1968). **174**, 2168.

[35] Smoot G. "Detection of Anisotropy in the Cosmic Blackbody Radiation." *Physical Review Letters*, (1977). **39**, 14, 898.

[36] Tangherlini F.R. "On energy momentum tensor of gravitational field." *Supl. Nuov. Cim.*, (1961). v. **20**, 351–367, (2^{nd} Trimestre 1961).

[37] Rizzi G., Ruggiero M.L., Serafini A. "Synchronization gauges and the principles of Special Relativity." *Found. Phys.*, (2005). v. **34**, 1885.

[38] Selleri F. "Superluminal signals and causality." *Annales de la Fondation Louis de Broglie*, (2003). v. **28**, 3–4.

[39] Goy F. "Derivation of three-dimensional inertial transformations." (1997). arXiv: gr-qc/9707004.

[40] Goy F., Selleri F. "Time on a rotating platform." (1997). arXiv: gr-qc/9702055v2.

[41] Selleri F. "Weak relativity - The Physics of space and time without paradoxes." (2009). C. Roy Keys Inc.

[42] Blake C., Wall J. "A Velocity Dipole in the Distribution of Radio Galaxies." *Nature*, (2002). **416**, 180-182.

[43] Laidlaw A. "Some Advantages of a Local Realist Wave Soliton Approach to EPR." *Apeiron*, (2002). Vol. **9** No. 1. quant-ph/0110160 v1 (2001)

[44] Shakespeare W. *Romeo and Juliet.*

[45] Dirac P.A.M. "The theory of the electron (part 1)." *Proceedings of the Royal Society in London*, (1928). **A117**, p610.

[46] Dirac P.A.M. "The theory of the electron (part 2)." *Proceedings of the Royal Society in London*, (1928). **A118**, p351.

[47] Bell J.S. "Speakable and Unspeakable in Quantum Mechanics", Cambridge: Cambridge University Press. (1987).

[48] Breit G. "An interpretation of Dirac's Theory of the electron." *National Academy of Sciences USA,* (1928). **14** p553.

[49] Messiah A. "Quantum Mechanics." Vol 2, Ch XX, (1965). 922-925, North Holland Publishing Company.

[50] Pusey M.F., Barrett J., Rudolph T. "The quantum state cannot be interpreted statistically." (2011). arXiv:1111.3328v1.

[51] Pusey M.F., Barrett J., Rudolph T. "On the reality of the quantum state". *Nature Physics.* (2012). **8 (6)**: 475–478.

[52] de Broglie L. "Recherches sur la Théorie des Quanta." *Annales de Physique*, (1925). 10^e Serie Tome III (Janvier - Fevrier 1925).

[53] Shanahan D. *Found. Phys.*, 2014. **44** pp. 349-367 arXiv:1401.4534v3.

[54] Skinner B. "A children's picture book introduction to Quantum Field Theory." (2015). https://www.ribbonfarm.com/2015/08/20/qft/

[55] Barut O.A. "A theory of particles of spin one-half." *Ann. Phys. (N.Y.)*, (1958). **5**, 095.

[56] Williamson J.G., van der Mark M.B. "Is the electron a photon with toroidal topology?" *Annales de la Fondation Louis de Broglie.* (1997). **22**, no.2, 133.

[57] Pakula R. Solitons and Quantum Behaviour. (2016). arXiv: 1612.00110.

[58] Orefice A., Giovanelli R., Ditto D. "Complete Hamiltonian Description of wave-like features in Classical and Quantum Physics." *Found. Phys.*, (2009). **39**, 256.

[59] Bohm D. "A Suggested Interpretation of the Quantum Theory in Terms of Hidden Variables I". *Physical Review,* (1952). **85** (2) 166–179.

[60] de Broglie L. "La mécanique ondulatoire et la structure atomique de la matière et du rayonnement." *J. Phys. Radium*, (1927). **8(5)**:225–241.

[61] Colin S., Durt T., Willox R. "de Broglie's double solution program: 90 years later." (2017). arXiv:1703.06158v1

[62] Bohr, N. "Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?" *Physical Review*, (1935). **48**: 696–702.

[63] Bell J.S. "On the Einstein Podolsky Rosen Paradox." *Physics.* (1965). **1**, 195–200.

[64] Aspect A., Grangier P., Roger B. "Experimental realization of EPR-Bohm gedankenexperiment: A new violation of Bell's inequalities." *Phys. Rev. Lett.*, (1982). v. **49**, 1804–1807.

[65] Jaynes E.T. "Clearing up mysteries, the original goal." *Proceedings Maximum Entropy and Bayesian Methods.* (1989). J. Skilling, Editor, Kluwer Academic Publishers, pp. 1-27.

[66] Aharonov Y., *et al.* "Revisiting Hardy's Paradox: Counterfactual Statements, Real Measurements, Entanglement and Weak Values." (2001). arXiv:quant-ph/0104062v1.

[67] Norsen T. "J.S. Bell's Concept of local causality." (2007). arXiv: 0707.0401v3.

[68] Ising, E. "Beitrag zur Theorie des Ferromagnetismus." *Z. Phys.*,(1925). **31** (1): 253–258

[69] Laidlaw A. "An Unspeakable Mechanism". *Intl. J. Adv. Res. Phys. Sci.*. (2018), **5**, 6, 10-28.

[70] 't Hooft G. "The cellular automaton interpretation of quantum mechanics." (2014). arXiv: 1405.1548.

[71] 't Hooft G. *Found. Phys.*, (2013). **43**:597-614. arXiv: 1205.4107.

[72] Hensen B. *et al.* "Experimental loophole-free violation of a Bell inequality using entangled electron spins separated by 1.3 km." *Nature*, (2015). **526**, 682–686. arXiv: 1508.05949.

[73] Redhead M. Incompleteness, nonlocality and realism. 1987, Oxford University Press.

[74] Rauch D. *et al* "Cosmic Bell Test Using Random Measurement Settings from High-Redshift Quasars." *Physical Review Letters*, (2018). **121** 080403 arXiv: 1808.05966.

[75] Peres A. "Unperformed experiments have no results". *American Journal of Physics.* (1978). **46** (7): 745–747.

[76] Barrett S. D.. Kok, P. "Efficient high-fidelity quantum computation using matter qubits and linear optics." *Phys. Rev. A*, (2005). **71**, 060310.

[77] Szańkowski P. et al. "The dynamics of two entangled qubits exposed to classical noise: role of spatial and temporal noise correlations." *QIP*. 2015, **14**, 3367-3397. arXiv:1408.4254.

[78] Throckmorton R., Barnes E., Das Sarma S. "Environmental noise effects on entanglement fidelity of exchange-coupled semiconductor spin qubits." *Phys. Rev. B.* 2017, **95**, 085405.

[79] Hensen B. et al. "Heralded entanglement between solid-state qubits separated by 3 meters." *Nature.* 2013, **497**, 86–90. Arxiv: 1212.6136.

[80] Bogdanovic S. et al. "Design and low-temperature characterization of a tunable microcavity for diamond-based quantum networks" (2016). arXiv:1612.02164v2.

[81] Everett H. "Relative State Formulation of Quantum Mechanics." *Rev. Mod. Phys.*. (1957), **29** 454–462.

[82] Wharton K., Price H. "Dispelling the quantum spooks – a clue that Einstein missed?" (2013), arXiv:1307.7744.

[83] Newton I. Letter to Robert Hooke, 5 February 1676.

[84] Newton I. Isaac Newton, Letters to Bentley, 1692/3.

[85] Jackson J.D. *Classical Electrodynamics*, (1962). Chapter 17. John Wiley and Sons.

[86] Konopinski E. J. *Electromagnetic Fields and Relativistic Particles* (1981). Mcgraw Hill pp 441-454

[87] Kerner E.H. "The Theory of action at a distance in relativistic particle dynamics. A reprint collection." *International science review series.* (1972). **v. 11.** New York, Gordon and Breach.

[88] van Flandern T."The Speed of Gravity - What the Experiments Say." *Physics Letters A*, (1998). **250** 1-11

[89] Peres A., Terno, D. "Quantum Information and Relativity Theory". *Rev. Mod. Phys.*, (2004). **76** (1): 93–123. arXiv:quant-ph/0212023

[90] Cui H.Y. "Direction adaptation nature of Coulomb's force and gravitational force in 4-Dimensional spacetime." (2001), arXiv: physics/0102073.

[91] Jackson J.D. *Classical Electrodynamics*, (1962). Chapter 2. John Wiley and Sons.

[92] Jackson J.D. *Classical Electrodynamics*, (1962). Chapter 15. John Wiley and Sons.

[93] Crandall R E *Am. J. Phys.* (1983). **51** 698 - 702.

[94] de Felice F. "On the Gravitational Field Acting as an Optical Medium" Gen. Rel. Grav. (2), pp 347-357 (1971)

[95] Misner C. W., Thorne K. S., Wheeler, J. A. *Gravitation*, (1973). W. H. Freeman and Co.

[96] Yilmaz H. "New approach to relativity and gravitation." *Annals of Physics*, (1973). Volume **81**, Issue 1, November 1973, 179-200.

[97] Puthoff H. E. "Polarizable-Vacuum (PV) Representation of General Relativity" *Found. Phys.* (2002). **32** 927-943 gr-qc/9909037

[98] Dicke R. H. "Gravitation without a Principle of Equivalence," *Rev. Mod. Phys.*, (1957). **29**, 363-376

[99] Einstein A., Infeld L. "Evolution of physics: The growth of ideas from early concepts to relativity and quanta." (1961). Simon and Schuster, New York, 242-243.

[100] Misner C.W. "Yilmaz cancels Newton." (1995). arXiv: gr-qc 9504040

[101] Alley C.O., Aschan P.K., Yilmaz H. "Refutation of C. W. Misner's claims in his article "Yilmaz Cancels Newton"." (1995). arXiv: gr-qc/9506082.

[102] Evans J., Nandi K.K. and Islam A. "The Optical Mechanical Analogy in General Relativity: New Methods for the Paths of Light and of the Planets". *American Journal of Physics*, (1996). **64** (11), November 1996.

[103] Laidlaw A. "On the Electromagnetic basis for gravity." *Apeiron*, (2004). **11**, 3.

[104] Notte-Cuello A., Rodrigues W. A. Jr. "Freud's Identity of Differential Geometry, the Einstein-Hilbert Equations and the Vexatious Problem of the Energy-Momentum Conservation in GR." *Adv. appl. Clifford alg.* (2009). **19** 113–145

[105] Donev S., Tashkova M. "On the Structure of the Nonlinear Vacuum Solutions in Extended Electrodynamics" hep-th/0204217 (2002)

[106] Shakespeare W. *Hamlet.*

[107] Wheeler J. A. "Information, physics, quantum: The search for links". *Complexity, Entropy, and the Physics of Information*, (1990). Addison-Wesley. ISBN

[108] Lancaster R. W., Harris L. D., Pearson D. "Fifty-year old samples of progesterone demonstrate the complex role of synthetic impurities in stabilizing a metastable polymorph" (2011). *CrystEngComm* **13**.

[109] Dunitz J. D. and Bernstein J. "Disappearing Polymorphs." (1995). Acc. Chem. Res. **28**, 4, 193–200. https://doi.org/10.1021/ar00052a005

[110] Bauer, J., Spanton, S., Henry, R. et al. "Ritonavir: An Extraordinary Example of Conformational Polymorphism." (2001) Pharm Res **18**, 859–866 . https://doi.org/10.1023/A:1011052932607