Third Edition

Quantum Non-Locality & Relativity

Tim Maudlin

WILEY-BLACKWELL

Quantum Non-Locality and Relativity

Metaphysical Intimations of Modern Physics

Third Edition

Tim Maudlin

WILEY-BLACKWELL
A John Wiley & Sons, Ltd., Publication

This edition first published 2011

Blackwell Publishing was acquired by John Wiley & Sons in February 2007. Blackwell's publishing program has been merged with Wiley's global Scientific, Technical, and Medical business to form Wiley-Blackwell.

Registered Office
John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, United Kingdom

Editorial Offices
350 Main Street, Malden, MA 02148-5020, USA
9600 Garsington Road, Oxford, OX4 2DQ, UK
The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, UK

For details of our global editorial offices, for customer services, and for information about how to apply for permission to reuse the copyright material in this book please see our website at www.wiley.com/wiley-blackwell.

Library of Congress Cataloging-in-Publication Data

Maudlin, Tim.
Quantum non-locality and relativity : metaphysical intimations of modern physics / Tim Maudlin. – 3rd ed.
p. cm.
Includes bibliographical references and index.
ISBN 978-1-4443-3126-4 (hardback) – ISBN 978-1-4443-3127-1 (paperback)
1 Physics–Philosophy. 2. Quantum theory. 3. Relativity (Physics) I. title.
QC6.M3612 2011
530.01–dc22
2011001784

A catalogue record for this book is available from the British Library.

This book is published in the following electronic formats: ePDFs 9781444396959; Wiley Online Library 9781444396973; ePub 9781444396966

Set in 10/12pt Sabon by SPi Publisher Services, Pondicherry, India
Printed in Singapore

1 2011

Contents

Preface to First Edition

I state my case, even though I know it is only part of the truth, and I would state it just the same if I knew it was false, because certain errors are stations on the road to the truth. I am doing all that is possible on a definite job at hand.

Robert Musil

If the introductory chapter of a book is the overture to the ensuing score, a brief, undeveloped melange of themes and leitmotifs destined to appear again and again, the preface serves as program notes. Here may one find some small account of the events which propelled the project; some acknowledgment of the many friends who encouraged and nourished it; and some explanation of idiosyncratic elements which arise from the author's own peculiarities. And as an intriguing introduction may encourage the reader to warm to the subject, so the successful preface may inspire some sympathy and understanding in the reader for the author's plight, for the many compromises, lapses, and errors that attend the writing of a book. So how did this book come about?

In October 1990, John Stewart Bell succumbed quite suddenly and unexpectedly to a hemorrhage of the brain. Anyone who had studied Bell's works mourned the passing of an incisive intellect; those who had had the pleasure of discussing the knotty problems of quantum theory with him felt even more sharply the loss of a figure of inspiring integrity, clarity, and humor. Here at Rutgers, Renée Weber suggested that we honor Dr Bell's memory with a symposium on his work. David Mermin treated us to a non-technical exposition of Bell's Theorem, Shelly Goldstein spoke of the relationship between Bell's work and that of David Bohm, and Professor Weber recounted some parts of her recent interview with Bell. My part was to be a short discussion of the compatibility between Relativity theory and the violation of Bell's inequality.

When I originally agreed to the assignment, I thought that I knew just what I was going to say: Relativity has been interpreted in two quite different ways, as forbidding superluminal effects and as demanding Lorentz invariance, and one must sort out how to construe Relativity before one can address the question of compatibility with quantum theory. But after a few days I realized that another construal of Relativity was available (no superluminal signals), then another (no superluminal energy transmission), then yet another (no superluminal information transmission). Since all of these interpretations of Relativity were provably non-equivalent, this situation posed a straightforward analytical task: how do the various interpretations relate to one another and how does each fare if Bell's inequality is violated? This manuscript is my attempt to work through that analytical problem.

In writing the book I have been constantly surprised by the variety and beauty of the interconnections between these various questions. But I have been even more impressed by Bell's deep and steady understanding of the problematic. Over and over I found some terse passage in Bell's work to contain exactly what needed to be said on a subject, the decisive pronouncement. I have often felt that whatever is of value in this book could be found in Bell's "The Theory of Local Beables" (1987, ch. 7), and have consoled myself that this book will have served a great purpose if it does no more than encourage people to read Bell with the care and attention he deserves.

My foremost goal in composing the book has been to make it comprehensible to the non-specialist. The sparks which fly when quantum theory collides with Relativity ignite conceptual brushfires of particular interest to philosophers, problems about causation, time, and holism, among others. Unfortunately, much of the work done by philosophers presupposes a considerable amount of familiarity with the physics. This is particularly sad since the physics is not, in most cases, very complicated. I fear that many readers may be frightened off from the topic by unnecessary formalization, so I have tried to keep the mathematical complexity of the discussion to a minimum. But on the other hand, I have not wished to drop to the level of vague metaphor which sometime infects popularizations. Every compromise between rigor and simplicity is a bargain with the devil, and I have struck mine as follows. The presentation of Bell's inequality needs no more than some algebra, and is quite rigorous. Understanding Relativity also requires no more than algebraic manipulation, but enough that a purely mathematical account would tax the patience of the average reader. So I have tried to present Relativity pictorially, so far as possible. The figures in the book present the concepts of Relativity accurately, but demand of the reader some skill in interpretation. Pictures of space-time look misleadingly like pictures of space, and the novice must unlearn some of the conventions of normal pictorial representation to avoid being misled. Newcomers should therefore take great care with the pictures in chapter 2: if those are properly understood, the sequel will be easy.

Quantum theory itself has been another matter. Most of the content before chapter 7 can be understood without much discussion of quantum formalism. That formalism itself also uses no more than linear algebra and vector spaces. Interested neophytes can find enough technical detail in any standard introductory text. A particularly nice and accessible presentation of the requisite mathematics is provided in David Albert's *Quantum Mechanics and Experience* (1992, ch. 2).

Just as professional physics scares off the uninitiated, so does professional philosophy. Philosophers have developed many languages of technical analysis which permit concise communication among the cognoscenti but which make amateurs feel like unwelcome guests. But most clear philosophical ideas can be presented intuitively, shorn of the manifold qualifications, appendices, and terminological innovations that grow like weeds in academic soil. I have been very selective in my discussions of the philosophical corpus, usually focusing on a single proposal which illuminates a region of logical space. I do not pretend to comprehensiveness in my review of the philosophical literature, and can only plead for understanding that my decisions reflect a desire for a short, provocative text.

Finally, I feel I should explain the "metaphysical intimations" of my subtitle. Metaphysics has acquired rather a bad reputation in this century, following the insistence of Kant that all metaphysical speculations must be pursued *a priori*. It was not always so. The fount of metaphysics as a philosophical pursuit is the treatise on First Philosophy by Aristotle which has come down to us as the *Metaphysics*. Aristotle was concerned with analysis of what there is into its most generic categories: substance, quality, quantity, etc. I see no reason to believe that Aristotle thought such an examination could not be informed by experience. At its most fundamental level, physics tells us about what there is, about the categories of being. And modern physics tells us that what there is ain't nothing like what we thought there is.

I have used "intimations" rather than "implications" because we still do not know how this story ends. Quantum theory and Relativity have not yet been reconciled, and so we can now at best only guess what picture of the world will prevail. But we do know enough to make some guesses.

This book would not have come to be without help of all sorts. David Albert, Nick Huggett, Martin Jones, Bert Sweet, Paul Teller, and Robert Weingard all devoted their own time and insight reviewing the manuscript and generously shared their views with me. Abner Shimony pushed me to clarify the models in chapter 6, and thereby saved me from repeating some errors in print. Steve Stich expended considerable effort finding the manuscript a home, and always had a word of encouragement. The National Endowment for the Humanities graciously provided financial support in the form of Summer Stipend FT-36726-92 (money = time). And the atmosphere in which the book was completed was lightened by Clio Maudlin, who also improvised some emendations with her feet.

Preface to Second Edition

Publication of the second edition of this tome affords the opportunity, beside typographical corrections, for two more substantial changes. The first is a new derivation of the Relativistic mass increase formula, to be found in chapter 3. The new derivation is somewhat simpler than that in the first edition, and has the advantage of allowing the exact formula to be obtained by means of a few lines of algebra. There are many methods for deriving the formula, but to my knowledge this one is novel. The second is the addition of an Overview of Quantum Mechanics. The overview contains just the bare mathematical bones of the theory, but that is enough to explain how violations of Bell's inequality are implied by the theory. It is hoped that the overview, while not a complete account of quantum theory, helps make this study more self-sufficient.

Beyond providing the chance for small improvements, the issuing of the second edition invites reflection, at some years' remove, on the plan of the original. Perhaps the most vexing question confronting any study of Bell's inequality is how the role of quantum theory ought to be treated. On the one hand, there is little doubt that Bell's inequality, and the experimental observation of violations of that inequality, would never have been discovered if not for the existence of the quantum formalism. On the other hand, the inequality itself is derived without any mention of quantum theory and the violations are matters of plain experimental fact. So the explication and analysis of the importance of Bell's work can in principle proceed without mentioning quantum mechanics at all. Should an account of Bell's inequality emphasize its historical roots in the great mysteries of quantum mechanics or rather sever those ties in the interest of logical clarity?

In composing this book, I chose the second option, playing down the role of quantum theory in favor of pure experimental results. In retrospect, I stand by that decision: the interpretation of quantum theory is troublesome enough in its own right to overshadow and confuse the relatively

The reason that this development is so cheering is that it deflects attention from other less important aspects of quantum theory. For example, it has been repeated *ad nauseam* that Einstein's main objection to quantum theory was its lack of determinism: Einstein could not abide a God who plays dice. But what annoyed Einstein was not lack of determinism, it was the apparent failure of *locality* in the theory on account of entanglement. Einstein recognized that, given the predictions of quantum theory, only a deterministic theory could eliminate this non-locality, and so he realized that a local theory must be deterministic. But it was the locality that mattered to him, not the determinism. We now understand, due to the work of Bell, that Einstein's quest for a local theory was bound to fail.

Schrödinger, in his famous "cat" paper, remarked on the "*entanglement* of our knowledge of [...] two bodies" (1935, p. 161) found in the quantum-mechanical formalism, but denied, as Einstein did, that this could reflect any real physical connection between separated systems: "Measurements on separated systems cannot directly influence each other – that would be magic" (ibid., p. 164). Bell's work has shown that the magic is real, and physicists who study entanglement have accepted the non-locality that Einstein and Schrödinger could not abide. I have briefly adverted to recent work on the information-theoretic implications of quantum theory in chapter 6 of this new edition.

When I first wrote *Quantum Non-Locality and Relativity*, I tried to keep discussion of the foundations of quantum theory to a minimum. All that is relevant to Bell's theorem are the predictions of quantum theory, not how the theory itself is understood. Separating these issues was especially important at the time because discussions of quantum theory *per se* contained a wealth of distractions and confusion. Perhaps the time is at last ripe to open up the dialog again, and to recover an understanding of Einstein's and Schrödinger's real concerns through the lens of Bell's theorem. I hope that chapter 10 provides a small step in the direction of a clearer understanding of what a comprehensible presentation of any physical theory (and hence a comprehensible presentation of quantum theory) demands.

Introduction

In the 1930s, Otto Neurath was one among many philosophers engaged in the project of purifying scientific language of its ambiguities, its vagueness, and its "metaphysical" contents. One might hope to accomplish this task by an act of radical innovation, building anew from elements of perfect clarity and precision. Neurath realized that such hopes are unattainable, that at best we can only successively improve the language we have, always retaining some of its deficiencies. He illustrated our situation with a resonant image:

> No *tabula rasa* exists. We are like sailors who must rebuild their ship on the open sea, never able to dismantle it in dry-dock and reconstruct it there out of the best materials. (Neurath 1959, p. 201)

The physical sciences themselves suffer the same fate. Fundamental conceptual changes occur, but they are always modifications of a previously existing structure. The entire edifice is not reconstituted anew; rather, tactical adjustments are made in order to render the whole consistent. The *ad hoc* nature of this procedure may leave us with lingering doubts as to whether the whole really is consistent.

During the past century our physical picture of the world has undergone two revolutionary modifications. The Theory of Relativity has overthrown classical presumptions about the structure of space and time. The quantum theory has provided us with intimations of a new conception of physical reality. Classical notions of causality, of actuality, and of the role of the observer in the universe have all come under attack. The ultimate outcome of the revolutions is now but dimly seen, at best. The final reconciliation of

Quantum Non-Locality and Relativity: Metaphysical Intimations of Modern Physics,
Third Edition. Tim Maudlin.

quantum theory and Relativity is a theoretical problem of the first magnitude. No quantum version of General Relativity exists, and the prospects for one are murky. But even apart from that hurdle, problems about the consistency of our two fundamental physical theories may appear.

The problem that will concern us here is easily stated. It arises from the remarkable results derived by John Stewart Bell in 1964 concerning the behavior of certain pairs of particles that are governed by quantum laws. Bell showed that observable correlations between the particles could not be accounted for by any theory which attributes only locally defined physical states to them. The particles appear to remain "connected" or "in communication" no matter how distantly separated they may become. The outcome of experiments performed on one member of the pair appears to depend not just on that member's own intrinsic physical state but also on the result of experiments carried out on its twin.

Many features of this quantum connection are puzzling. It is, for example, entirely undiminished by distance. This distinguishes it from any connection mediated by a classical force, such as gravity or electromagnetism. But even more amazingly, the connection exists even when the observations carried out occupy positions in space and time which cannot be connected by light rays. The particles communicate faster than light.

It is this last feature which raises questions about the consistency of our fundamental theories. Relativity is commonly taken to prohibit anything from traveling faster than light. But if nothing can go faster than light, how can the particles continue to display the requisite correlations even when greatly separated? The two pillars of modern physics seem to contradict one another.

The predicted correlations have been experimentally confirmed. Indeed, they have been seen even in conditions where the communication between the particles would require superluminal velocities. So we are presented with the problem of determining whether Relativity has been violated, and, if so, whether our present account of space-time structure must be modified or abandoned.

The question of whether the quantum correlations are consistent with Relativity seems precise enough to admit a decisive answer, but on closer examination this appearance of clarity dissolves. Exactly what sort of constraints Relativity imposes on physical processes is a matter of much dispute. Many physicists and philosophers would agree that Relativity prohibits *something* from going faster than light but disagree over just what that something is. Among the candidates we may distinguish:

> Matter or energy cannot be transported faster than light.
> Signals cannot be sent faster than light.
> Causal processes cannot propagate faster than light.
> Information cannot be transmitted faster than light.

Most of these prohibitions are easily seen to be non-equivalent. For example, signals could in principle be sent without any accompanying transmission of matter or energy. Or again, superluminal causal processes could exist which, due to their uncontrollability, could not be used to send signals.

Yet another interpretation holds that Relativity requires only that

Theories must be Lorentz invariant.

This requirement is compatible with the violation of every one of the prohibitions listed above.

Not surprisingly, the various prohibitions are justified by different considerations. In one case it is claimed that a violation of the prohibition would require an infinite amount of energy, in another than it would engender paradox, in yet another that some relativity principle would be abrogated. We are therefore left with a rather tangled thicket of problems. We must consider each of the proposed prohibitions and ask whether it is violated by the quantum connection. We must ask how each prohibition is justified and how it connects with the formalism of the Theory of Relativity. We would also like to see how the prohibitions relate to one another. Until this work is done we cannot begin to evaluate the implications of the quantum correlations for our picture of the world.

This problematic directly dictates the structure of our inquiry. Chapter 1 presents Bell's results with a minimum of technical machinery. Chapter 2 is a short intuitive account of Special Relativity. The following four chapters examine the four prohibitions listed above, tracing their connection with Special Relativity on the one hand and their compatibility with quantum non-locality on the other. Chapter 7 delves into the technical requirement of Lorentz invariance and its implications. Chapter 8 touches on the difficulties involved in passing from the space-time of Special Relativity to that of General Relativity.

Any book which attempts to deal with quantum theory, Special Relativity and General Relativity courts various forms of disaster. Technical and mathematical detail can easily push the discussion beyond the ready grasp of the general reader, and the philosophical interpretation of the mathematical formulae can be even more daunting. In this last respect an asymmetry regarding our two fundamental theories should be noted. Relativity is quite well understood. Although it employs ideas that depart radically from those of classical physics, the concepts are themselves unproblematic and become quite transparent with use. Quantum theory, in contrast, still presents deep and basic interpretational problems, the discussion of which could fill several volumes. Fortunately, our concerns will not draw us much into these controversies. Bell's theorem can be proven without so much as a mention of quantum theory, and although one uses quantum theory to

1

Bell's Theorem: The Price of Locality

According to our naive, everyday conception, and even according to most of our refined theories, the physical world is composed of separate individually existing objects. The book on my desk sits apart from the glass, each constituted separate from the other and with its own intrinsic properties. The book has its mass, shape, number of pages, the marks of its history engraved on it. It is made up of atoms, each with its own physical constitution, tied together by chemical bonds. The glass similarly exists on its own, constructed from a separate complement of particles. There are, of course, relations between the book and the glass. The book is heavier and occupies more volume; there is a certain definite distance between them. Spatial separation plays a unique role: as an external relation it is not determined by any facts about the book and the glass taken individually. But once we have taken into account their intrinsic properties and their situation in space we appear to have exhausted the facts about the pair. All other facts about them are determined by these.

Each of the pair may influence the other. The glass, full of steaming tea, raises the temperature of the book which is in its proximity. But this interaction is mediated by other localized bits of matter. Air molecules around the glass are made more energetic through interactions with the tea, some wander off and communicate their energy with the book, heating it. The book exerts a slight gravitational pull on the glass and vice versa. This is a subtle matter, but we come to think of this too as a mediated interaction, an effect of a gravitational field.

The fields of classical physics are not so familiar as books or atoms but they too are local entities. Although an electric field may spread out and

Quantum Non-Locality and Relativity: Metaphysical Intimations of Modern Physics, Third Edition. Tim Maudlin.

permeate the universe, the state of the field is determined entirely by its value at each point of space. Disturbances propagate through the field, but they do so by local interactions: changes in the field quantities induce other changes nearby and so ripple off to infinity. Like the transmission of heat, this process takes time as the vibrations of the field are passed along.

Einstein set great store by the idea that the physical state of the universe is determined by a set of locally defined physical magnitudes so that the state of any localized entity exists independently of all spatially separated systems. As he expressed it in a letter to Max Born:

> If one asks what, irrespective of quantum mechanics, is characteristic of the world of ideas of physics, one is first of all struck by the following: the concepts of physics relate to a real outside world, that is, ideas are established relating to things such as bodies, fields, etc., which claim "real existence" that is independent of the perceiving subject – ideas which, on the other hand, have been brought into as secure a relationship as possible with the sense-data. It is further characteristic of these physical objects that they are thought of as arranged in a space-time continuum. An essential aspect of this arrangement of things in physics is that they lay claim, at a certain time, to an existence independent of one another, provided these objects "are situated in different parts of space". Unless one makes this kind of assumption about the independence of the existence (the "being-thus") of objects which are far apart from one another in space – which stems in the first place from everyday thinking – physical thinking in the familiar sense would not be possible. It is also hard to see any way of formulating and testing the laws of physics unless one makes a clear distinction of this kind. This principle has been carried to extremes in the field theory by localizing the elementary objects on which it is based and which exist independently of each other, as well as the elementary laws which have been postulated for it, in the infinitely small (four-dimensional) elements of space.
>
> The following idea characterizes the relative independence of objects far apart in space (A and B): external influence on A has no direct influence on B; this is known as the "principle of contiguity," which is used consistently in the field theory. If this axiom were to be completely abolished, the idea of the existence of (quasi-) enclosed systems, and thereby the postulation of laws which can be checked empirically in the accepted sense, would become impossible. (Born 1971, pp. 170–1)

Bell's theorem addresses the implications, and ultimately the tenability, of this picture.

Given the extreme generality of the local conception of reality it is hard to imagine that it could, by itself, have any testable empirical consequences. No constraints have been put on the nature or complexity of the locally defined quantities. The locality condition allows, for example, that every particle in the universe could retain traces of every interaction it has ever undergone. It allows a system to be governed by laws which are deterministic or are probabilistic, placing no limit on the subtlety or sophistication of the laws. Nonetheless, Bell was able to show that some behavior of separated pairs of systems cannot be explained by *any* local physical theory if the systems do not interact. Although Bell's results can be derived in different ways and with great generality, we will begin by focusing on a singular fact about light.

Polarization

When one passes a beam of sunlight through a polarized filter, such as the material used in Polaroid sunglasses, two things happen. First, about half of the light is absorbed and half transmitted, as is immediately evident. Second, the light which is transmitted displays an entirely new and surprising characteristic: it shows a particular directionality. This directionality can be most easily observed if one passes the new beam through a second polarized filter. The effect of the second filter depends critically on its orientation with respect to the first. In one orientation the second polarizer will have no effect at all, allowing the entire beam to pass. But as it is rotated, the second filter allows less and less of the light through. By the time it has been turned 90°, it absorbs the beam entirely; as it is rotated further it permits ever more light to pass until, at 180°, the whole beam passes again.

The directionality that the sunlight acquires depends on the orientation of the first polarizer. When the first filter is rotated, the characteristic orientation at which the transmitted beam passes the second filter rotates with it. So light which has passed through a Polaroid filter acquires a new property, a polarization, which is associated with some direction perpendicular to its line of motion.

All that really concerns us is the behavior recounted above; the explanation of the phenomena will ultimately be irrelevant to our concerns. But to help fix our ideas it may help to recall the classical theory of polarization. The classical theory provides us with a simple picture of polarization which should, however, be taken *cum grano sails*, for it cannot be straightforwardly extended when quantum phenomena are taken into consideration.

According to classical physics, light is an electromagnetic wave, a propagating disturbance of the electric and magnetic fields. The fields that vary always point perpendicular to the direction of motion of the light. At any given moment the electric and magnetic fields are also perpendicular to each other, but as time goes on their direction and magnitude may change in any number of ways. For example, if we look at a ray of light head on as it comes toward us, the electric field may rotate in a circle, either clockwise or counterclockwise (circularly polarized light); or it may trace out an ellipse,

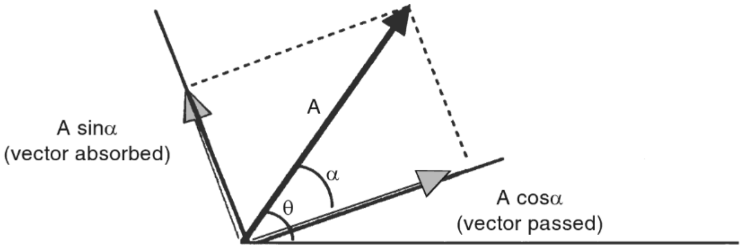

Figure 1.1 Resolving a Vector

rotating and varying in length; or it may simply oscillate back and forth without rotating, always remaining in a plane that points in a given direction. This last possibility, plane polarized light, is the case of interest to us. Plane polarized light has a characteristic direction, the direction of the plane in which the electric field vector always lies. Furthermore, for any direction θ we choose, light of any sort can be analyzed into a component plane polarized in that direction and a component polarized in the perpendicular (that is, θ + 90°) direction. Even circularly polarized light can be constructed from two such elements, if they are added together in the right way, with the right phase relations.

The phenomena recounted above are now easily explained. A Polaroid filter in effect analyzes all incoming light waves into two parts: one plane polarized in the direction of the filter's polarization, the other perpendicular to that direction. It then absorbs the perpendicular component, allowing only the plane polarized remainder to pass through. If the incoming light is unpolarized, this means that on average half of it will pass through and half be absorbed. The effect of the second filter then depends crucially on its orientation relative to the first. If they are perfectly aligned, the light which passes the first is already polarized in the direction of the second and so all gets through. If the second is misaligned by 90°, then exactly the component which passes the first will be absorbed by the second, and none will get through.

What if the two filters are misaligned by some angle α between 0° and 90°? We can represent the light coming through the first filter by a vector pointing in the θ direction whose length A represents the maximum amplitude of the electric field. The second filter resolves this vector into two components, one parallel to θ – α, the other perpendicular (see figure 1.1). The perpendicular component is absorbed by the filter, so the amplitude of the transmitted light is A cos α.

We now must appeal to a seemingly minor but highly significant fact. The energy of plane polarized light is proportional to the *square* of the amplitude of its electric field vector. So if we measure the amount of light

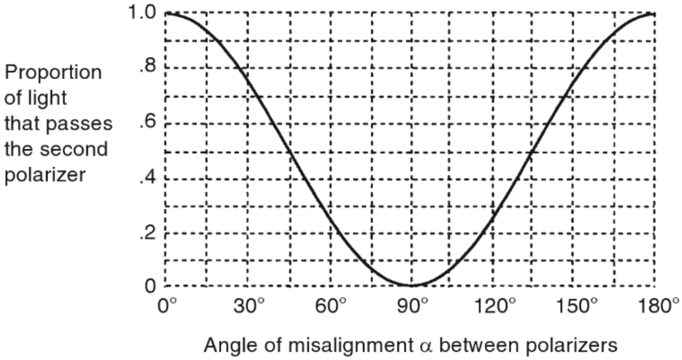

Figure 1.2 Proportion of Light Passing Second Polarizer

which passes the second polarizer by the energy of the beam, the proportion of the beam that gets through is $A^2 \cos^2 \alpha / A^2 = \cos^2\alpha$. Figure 1.2 shows the proportion of the beam which passes the second filter as a function of the angle of misalignment α.

As expected, when $\alpha = 0°$ and the filters are aligned, all the beam is transmitted. When the filters are misaligned by 90° none of the light gets through. But the most significant behavior is found between these extremes. For the moment we need only note that when $\alpha = 30°$, $\cos^2 30° = \left(\sqrt{3}/2\right)^2 = \frac{3}{4}$ of the beam gets through, while when $\alpha = 60°$, $\cos^2 60° = \left(\frac{1}{2}\right)^2 = \frac{1}{4}$ of the light is transmitted.

Light Quanta

According to the classical conception light is a wave, spread out in space. Whenever a plane polarized beam impinges on a filter oriented at, say, 30° off of the polarization plane of the incoming beam the same thing happens: $\frac{3}{4}$ of the beam passes, and what gets through is polarized in the direction of the filter. A beam always comes out with its amplitude and energy reduced by a fixed proportion.

But as Einstein observed in 1905, light does not always behave like a wave. For example, when light falls on certain metals it can knock out electrons causing a current to flow, the so-called photoelectric effect. When one measures the energy of the electrons so liberated one finds that the energy of the incident light is not delivered uniformly over the surface of the metal as one would expect. The energy rather comes in small but discrete packets.

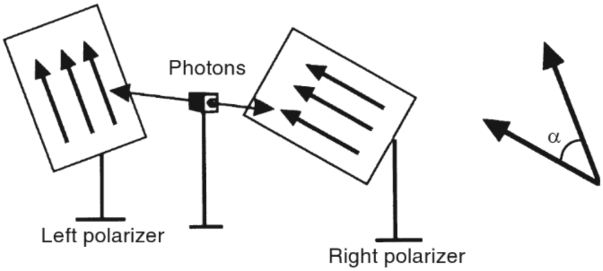

Figure 1.3 Experimental Set-up

half jointly absorbed. If the polarizers are misaligned by 90° the photons always disagree, one being absorbed, the other not. And for any other angle of misalignment α the percentage of pairs which agree (in the long term) is $\cos^2 \alpha$, as shown in figure 1.2.

Note that when we set up a particular experiment we have two choices to make. First we must choose the angle θ of the right-hand polarizer. Then we choose the degree of misalignment α of the left-hand polarizer. If we decide to examine a case of perfect alignment ($\alpha = 0°$) we are still at liberty to set the pair of filters in any direction θ we choose. The fact that we have two free variables, θ and α, is just a reflection of the fact that we have two decisions to make: the angle of the right polarizer and the angle of the left. But no matter how we set the two, the only relevant parameter for calculating the probability of agreement is α, the degree of misalignment (see figure 1.3). If $\alpha = 30°$ the photons will agree $\frac{3}{4}$ of the time; if $\alpha = 60°$ they will agree one time out of four. These simple facts about pairs of photons emitted by calcium vapor are enough to destroy any theory according to which physical reality is local.

How Do They Do It?

Suppose that you and a friend are set the task of reproducing the behavior of the photons: one of you will play photon L, the other photon R. These are the rules of the game: you and your friend start out together in a room (the "calcium atom"). You know that each of you will leave the room by a different door, and after some period of time you will each be asked a question. The question will consist of a number between 0 and 180 written on a piece of paper. Your answer must be either the word "passed" or "absorbed." Before you leave the room, you have no idea which question either of you will be asked. However, while in the room you and your friend are permitted to devise any strategy you please in order to coordinate your answers. Your aim is to ensure that after many repetitions of the game (you are permitted to adopt an entirely new strategy each time) your answers display exactly the same sorts of correlations as the photons show. That is, your strategies must ensure that, in the long run, when the question asked you differs from that asked your friend by an amount α, your answers agree $\cos^2 \alpha$ of the time.

For the moment, we will simplify your task even further. Unlike the photons, which have no information at all about which question will be $\cos^2 \alpha$ asked, you and your friend can know that only one of three possible questions, "0?", "30?" or "60?", will be asked. (We will eventually simplify the task even more, but it is easiest to begin here.) Of course, while you are in the room you still have no idea which of the three questions either of you will be asked. And once you leave the room, we suppose *you have no way of knowing what question has been asked (or will be asked) of your partner*. Your behavior may be determined by your agreed upon strategy and by the question you are asked, but not by the question which your friend happens to be asked. Once again, you must each respond either "Passed" or "Absorbed" when a question is asked.

Over a long run of this game you are aiming to reproduce the behavior of the photons in similar circumstances. That is, after a long series of plays you want to ensure that

Fact 1: When you and your friend happen to be asked the same question you always give the same answer.

Fact 2: When your questions differ by 30, that is, when one is asked "0?" and the other "30?" or one is asked "30?" and the other "60?", you and your friend agree $\frac{3}{4}$ of the time.

Fact 3: When your questions differ by 60, that is, when one of you is asked "0?" and the other "60?", your answers agree $\frac{1}{4}$ of the time.

After all, this is what the photons manage to do.

You and your friend are free to agree on any strategy you like, and you are free to vary your strategy from experiment to experiment. We may suppose that the questions to be asked are chosen at random, so that the pair of the questions "0?" to R and "30?" to L, for example, occurs $\frac{1}{9}$ of the time. It is not, however, important that the questions be asked equal amounts of the time, only that the choices be made at random, so that you can have no idea what is to come. How might you go about settling on a strategy?

The first obvious point is that there is no advantage, and much disadvantage, to using any sort of random element after you have left the room. For suppose your strategy demands that if asked the question "0?" you will decide your answer by a flip of the coin. Since you are unable to communicate with your partner there would be no way for your friend to know how you have answered the question, and so no way to be sure of matching your answer if asked the same question. In general, there is no possible way of satisfying Fact 1 above without deciding in the room *how each of you will answer each question if asked*. For without the knowledge of how your partner would answer a question you cannot act so as to ensure that your answers will match if you happened to be asked the same question.

Besides, no possible advantage can be gained by the introduction of random elements. If one of you may have to flip a coin when asked a question, why not flip it beforehand in the room and share the result with your partner? Or flip it three times, one for each possible contingency. Your partner would then have more information than would be available if you only appeal to the random element when actually asked the question. That excess information cannot possibly *degrade* your performance since, in the worst case, the information can just be ignored. Thus we have the simple result that *any strategy which involves local stochastic elements can do no better than a corresponding strategy where the random choices are made at the source*. A "local stochastic element" is a random process which takes place outside the room and whose outcome cannot be communicated to one's partner. In your case, "at the source" means "in the room"; for the photons it means "in the calcium atom." So in the first place, strategies utilizing local stochastic elements cannot ensure the perfect correlation when identical questions are asked, and in the second place, for every strategy using such elements an equally effective strategy which eschews them exists. If you have a penchant for flipping coins, you may as well flip them in the room. Given these two facts we may now narrow our search to strategies which involve no local stochastic elements. This means that when you leave the room each of you knows exactly what the other will do in each possible situation.

Furthermore, not any such deterministic strategy will do. Since you always run the risk of being asked identical questions, you and your friend must resolve to give the same answer as each other to each question. Only in this way can you assure that when answering identical questions your answers will tally.

So our situation has been greatly simplified. Only eight possible strategies are available, corresponding to the possible ways of answering the three questions. You might, for example, decide to answer "passed" no matter which of the three questions is asked. We will represent that strategy as

⟨P, P, P⟩, where the first slot represents the answer to "0?", the second to "30?" and the third to "60?". The eight possible strategies are then:

(1) ⟨P, P, P⟩	(2) ⟨A, A, A⟩	(A)
(3) ⟨A, P, P⟩	(4) ⟨P, A, A⟩	(B)
(5) ⟨P, A, P⟩	(6) ⟨A, P, A⟩	(C)
(7) ⟨P, P, A⟩	(8) ⟨A, A, P⟩	(D)

Since we are only interested in whether the answers given by you and your friend agree or differ, we can regard each of the corresponding mirror-image strategies above as equivalent. That is, if you choose either strategy (1) or strategy (2) you will agree no matter what pair of questions is asked, if you choose (3) or (4) you will disagree if exactly one person is asked "0?" and agree otherwise, and so on. (Of course there are *other* facts, such as that in the long run approximately half the photons pass and half are absorbed, that would demand a judicious choice between the strategies in the right column and those in the left, but those facts have been omitted from our list.) So we may lump together strategies (1) and (2) calling each "strategy (A)," either (3) or (4) will be "strategy (B)," (5) or (6) "strategy (C)," (7) or (8) "strategy (D)." In order to ensure the strict correlations of Fact 1, you and your friend must choose among strategies (A), (B), (C), and (D) every time a new experiment is run. The only real option that is left open to you, then, is what proportion of the time each strategy will be chosen.

Let us suppose that your decisions over the long run result in choosing strategy (A) a proportion α of the time, strategy (B) β of the time, strategy (C) γ of the time, and strategy (D) δ of the time. α, β, γ, and δ must all be positive numbers (or zero), and of course $\alpha + \beta + \gamma + \delta$ must equal unity.

You and your friend must make your choice of which strategy to adopt in complete ignorance of what questions you are to be asked. Further, we may assume that the choice of questions is determined by a process which is random with respect to your choice of strategy. The experimenters, however they decide which questions to ask, do not do so by predicating their choice on your predetermined strategy. In these circumstances, the long-run results of many repetitions of these experiments will depend solely on the values of α, β, γ, and δ. For example, suppose we wish to know how often the pair of questions "0?", "60?" will receive answers which disagree. They will do so exactly when you have chosen strategy (B) or strategy (D), as can be verified by inspection. In the long run, you choose those strategies $\beta + \delta$ proportion of the time. And since the selection of experiments in which that pair of questions is asked constitutes a random selection from the sequence of strategies you choose, in the long run that pair of questions will receive disagreeing answers $\beta + \delta$ of the time.

By only selecting among the eight strategies we have ensured that Fact 1 will be satisfied. What of the other Facts? Fact 2 states that when the pair of questions "0?" and "30?" or "30?" and "60?" are asked, your answers will agree $\frac{3}{4}$ of the time. Another way of putting this is that your answers will disagree $\frac{1}{4}$ of the time. Similarly, Fact 3 states that when asked the pair of questions "0?" and "60?" you must disagree $\frac{1}{4}$ of the time. We have already seen that the proportion of the "0?" – "60?" experiments which yield disagreeing answers is $\beta + \delta$. By similar reasoning the proportion of "0?" – "30?" experiments which yield disagreement is $\beta + \gamma$, and the proportion of "30?" – "60?" experiments which yield disagreements is $\gamma + \delta$. To recover the correlations of the photons, then, you and your friend must arrange things so that

$$\gamma + \delta = 0.25$$
$$\beta + \gamma = 0.25$$
$$\beta + \delta = 0.75.$$

But now the rub becomes apparent. For on the one hand, the first two equations together imply that $(\beta + \gamma) + (\gamma + \delta) = 0.25 + 0.25 = 0.5$. But on the other hand, $(\beta + \gamma) + (\gamma + \delta) = 2\gamma + (\beta + \delta) = 2\gamma + 0.75$ (by the last equation). These results together imply that $0.5 = 2\gamma + 0.75$ or $2\gamma = -0.25$, so that $\gamma = -0.125$. But γ must be a positive number: it is not among your options to choose strategy (C) –12.5 percent of the time. In sum, there is no possible long-term selection of strategies that you and your friend can adopt which will ensure that your answers will display the same correlations as those of the photons.

Bell's Theorem(s)

The result just obtained can be generalized in many ways, all of which address themselves to variations of the question: given collections of two or more particles and a choice of observations that can be carried out on each, what sorts of constraints on the correlations among results can be derived if the observation carried out on one particle cannot influence the result of observations carried out on the others? We have just seen that in the case of two particles with three possible observations on each particle, if the results when the same experiments are carried out on both wings are perfectly correlated then (proportion of disagreement when experiments 1 and 2 are chosen) + (proportion of disagreements when experiments 2 and 3 are chosen) ≥ (proportion of disagreements when 1 and 3 are chosen). We can abstract from the exact nature of the experiments which are carried out, for in any such case the same reasoning leads to $(\gamma + \delta) + (\beta + \gamma) \geq (\beta + \delta)$.

just as before, with exactly the same strategies available. Strategy ⟨P, A, A⟩ represents the decision that the right-hand photon will pass if measured at 0°, both will be absorbed if measured at 30°, and the left-hand photon will be absorbed if measured at 60°.

Indeed, the impossibility of satisfying these conditions is even more obvious now. The photons must agree on how they will both act if measured at 30°. In order to achieve the 75 percent agreement rate for the 0°–30° possibility, 25 percent of the time the right-hand photon must choose a strategy in which 0° differs from 30°. To achieve the 75 percent agreement for 30°–60°, the left-hand photon can only allow its 60° value to deviate from the common 30° value 25 percent of the time. But if the right-hand photon lets the 0° value deviate from the 30° value only 25 percent of the time, and the left-hand photon only allows 60° to deviate from 30° 25 percent of the time, then at least 50 percent of the time neither will so deviate. But then at least 50 percent of the time 0° and 60° will agree with each other (since neither deviates from the common 30° value) and the observed 0°–60° disagreement rate of 75 percent cannot be recovered.

The analysis again depends on the primary assumption: the setting of the polarizer on one side cannot be communicated to or have an effect on the photon on the other side. Experimentally we can try to ensure this condition in two ways. First, we want to separate the two analyzers from one another in space. Second, we want to choose the setting of the polarizer at the last possible moment. If the setting is chosen just before the measurement is made then the second photon could not adjust its strategy on the basis of prior knowledge of the question to be asked its partner. That is, if while you and your partner are in the room it has not yet even been decided which questions will be asked, then you cannot agree on a successful strategy while still in the room. Hence an ideal experimental condition will have two polarizers each of which can be set to one of two settings, well separated in space, with quick choices of the settings being made.

Such an experimental situation was realized in 1982 by Alain Aspect and his collaborators (Aspect *et al.* 1982). Since physically rotating a polarizing filter cannot be done quickly, Aspect hit on a clever means of choosing between the two possible experiments on each side. Two polarizers and detectors were set up on each side of the experiment, with a very fast optical switch which could send the photons to either one (figure 1.4). Each of the optical switches alternated the beam between the two polarizers every 10^{-8} seconds. The right-hand apparatus was about 12 meters away from the left-hand one.[2] (These details will be of some importance in the coming chapters.) For the moment we can only note that it is in no way obvious how the result on the right-hand side could depend on which detector the left-hand photon is sent to. And if no such dependence exists then Bell's inequality cannot be violated: the quantum correlations could not be reliably

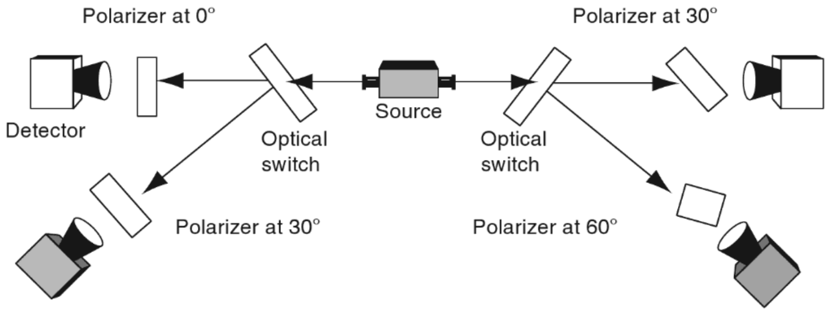

Figure 1.4 Aspect's Experiment

produced.[3] In Aspect's experiment the quantum predictions were confirmed and Bell's inequality was violated.

What Is Weird About the Quantum Connection?

Aspect's experiment and other such experiments have produced observable data which cannot be predicted by any theory which disallows influence of the career of one particle on the behavior of the other once they separate. Somehow the particles must remain in communication, the observable behavior of one being determined, in part, by the nature of the observations carried out on its twin. After being created together the pair of particles remain interconnected.

This interaction among distantly separated particles presents profound interpretive difficulties. But one might initially be surprised that this behavior should elicit any concern at all. After all, classical physics is shot through with such causal connections among distant particles. Newtonian gravitational theory, for example, postulates that every massive particle in the universe exerts a gravitational force on every other, a force of magnitude Gm_1m_2/r^2. When a sparrow falls in Yugoslavia it has effects in New Brunswick and on Saturn and in the most distant galaxy. Some small gravitational tug will register in the smallest parts of the most far-flung stars. In the face of this sort of interconnectedness the quantum connection looks rather modest.

But there are at least three features of the quantum connection which deserve our close attention. All of them are, to some extent, surprising. The first two prevent our assimilation of these quantum effects to those of a force like gravitation. The last presents problems for reconciling the results of experiments like that of Aspect with the rest of our physical picture.

1. The quantum connection is unattenuated

The fall of a sparrow in Yugoslavia may have its effects in New Brunswick and on Saturn and beyond, but the effect becomes progressively smaller the farther away one goes. Since the gravitational force drops off as the square of the distance it eventually becomes negligible if one is concerned with observable effects. The gravitational pull of the sparrow plays no noticeable role in affairs in New Brunswick, much less in the affairs of extra-galactic societies.

The quantum connection, in contrast, appears to be unaffected by distance. Quantum theory predicts that exactly the same correlations will continue unchanged no matter how far apart the two wings of the experiment are. If Aspect had put one wing of his experiment on the moon he would have obtained precisely the same results. No classical force displays this behavior.

2. The quantum connection is discriminating

When the sparrow falls in Yugoslavia I feel a slight gravitational tug in New Brunswick. So does the computer on my desk, and the cat asleep on the bed. Every inhabitant of Princeton is jostled slightly, and to nearly the same extent as the population here. The effects of the sparrow's fall ripple outward, diminishing as distance increases, jiggling every massive object in its way. Equally massive objects situated the same distance from the sparrow feel identical tugs. Gravitational forces affect similarly situated objects in the same way.

The quantum connection, however, is a private arrangement between our two photons. When one is measured its twin is affected, but no other particle in the universe need be. If we create a thousand such correlated pairs and send the right-hand members all off in a group, each particle still retains its proprietary connection with its partner. A measurement carried out on one member of the right-moving hoard will influence only one member of the left-moving group, one particle situated in the midst of a thousand seemingly identical comrades.

The quantum connection depends on history. Only particles which have interacted with each other in the past seem to retain this power of private communication. No classical force exhibits this kind of exclusivity.

3. The quantum connection is faster than light (Instantaneous)

Of all the peculiarities of the particle communication, this might seem to be the most benign. For although no classical forces are unattenuated or discriminating, all were at least originally described as instantaneous. Classical gravitational and electrical forces were described as being determined by the contemporaneous global distributions of matter or of electric charge. Any change in that global distribution would therefore immediately have effects on the forces felt everywhere.

Although instantaneousness was a feature of the first theories of gravitation and electricity, it was not an essential feature. Newton thought that gravitation must be the effect of some subtle particles, about which he famously framed no hypotheses. He would therefore have expected a perfected theory of gravitation to take the speed of these particles into account. In such a final theory one would expect some delay to intervene between the sparrow's fall and the slight jostle it causes in New Brunswick. Of course, in the classical regime no *a priori* constraint could be put on the velocity of the gravitational disturbance, but one might reasonably expect it not to be infinite.[4]

But the modern theory of space and time differs radically from the classical view. The revolution has come in two stages, both initiated by Einstein: the Special and General Theories of Relativity. The Special Theory confers upon light, or rather upon the speed of light in a vacuum, a unique role in the space-time structure. It is often said that this speed constitutes an absolute physical limit which cannot be broached. If so, then no relativistic theory can permit instantaneous effects or causal processes. We must therefore regard with grave suspicion anything thought to outpace light.

The quantum connection appears to violate this fundamental law. Aspect's experiment was so contrived that the setting of the equipment at one side could not be communicated, even by light, in time to influence the other side. All three of these weird aspects of the quantum connection are related to spatial structure. Classical forces depend on spatial separation while the quantum connection does not. The effects of classical forces are determined by spatial dispositions: two electrons near one another will be (nearly) identically affected by distant gravitational or electrical sources. The quantum connection discriminates even among identical sorts of particles which are in close proximity to one another. Finally, the speed of the quantum communication appears to be incompatible with relativistic space-time structure.

Our concern will almost entirely be with the last of these three features. It is surprising that the communication between particles is unattenuated and discriminating, but often our best counsel is simply to accept the surprising things our theories tell us. The speed of the communication is another matter. We cannot simply accept the pronouncements of our best theories, no matter how strange, if those pronouncements contradict each other. The two foundation stones of modern physics, Relativity and quantum theory, appear to be telling us quite different things about the world. To understand, and perhaps resolve, that conflict we must consider carefully just what Relativity tells us about space and time.

Appendix A: The GHZ Scheme

If both you and your partner are uninformed about the question being asked the other, there is no strategy for answering questions which will reliably reproduce the quantum correlations in the long run. But this trouble matching the behavior of photons only appears in the long run: in every individual "experiment" you and your partner can be assured that your responses *in that particular experiment* are responses which the photons might have given. If, for example you decide during one particular game that both players will answer "passed" no matter which question is asked, you can be assured, no matter which questions are asked, that your responses will not in themselves violate any quantum-mechanical predictions. For the only ironclad constraint imposed by the quantum correlations is that both partners give the same answer if they are both asked the same question. So long as you have agreed to a common response to each question you are safe on that run: the answers you give will certainly be quantum-mechanically permissible. Only after many games will your failure to match the target correlations emerge.

This state of affairs is frustratingly equivocal. Non-communicating partners can be certain that in no particular game will they diverge from what the photons might do, but can be equally certain that over time they must diverge in their cumulative behavior. There is something a bit ephemeral or ghostly about the problem, in that it lies entirely in long-term averages. Is this an indication of something deep about the nature of the quantum predictions?

A discovery by Daniel Greenberger, Michael Horne, and Anton Zeilinger (1989) dispels this suspicion. Greenberger, Horne, and Zeilinger (GHZ) found that in some instances quantum theory makes predictions about correlations between particles which are so strong that no local (i.e. non-communicating) strategy can be assured of matching the quantum predictions on any single run of the experiment. Although the GHZ scheme has not been tested in any actual experiment, it merits our attention as an indication of the strongly non-local character of quantum mechanics.[5]

The GHZ scheme uses three particles rather than two, and measures spin rather than polarization. The three particles can be created and allowed to separate to arbitrary distance, at which time a spin measurement is made on each. As in the Bell case, we can model the situation as a game. In this one, you and two partners begin together in a room. Some time after you depart the room, traveling in different directions, each of you will be asked one of two questions. We will denominate the questions "X?" and "Y?"; they correspond to measuring the spin of the particles in either of two orthogonal directions. To each question you must answer either "up" or "down" (in the literature, the responses of the particles are also often represented as 1 and −1). Since each particle can be asked one of two questions, there are eight distinct possible experimental arrangements, but of these only four will ever be used. Either two of the players will be asked "Y?" and the last one "X?" or all the players

2

Relativity and Space-time Structure

In the last chapter we saw that no collection of systems, no matter how constructed or governed by law, could reliably produce violations of the Bell inequalities provided that one condition be satisfied: the question asked or experiment done on one side of the apparatus can have no influence on the result at the other side. We denominated that condition "locality," a decision which demands some justification.

If we want to isolate the two wings of the experiment from one another, a natural thing to do is to move them far apart. In the classical regime this tactic seems reasonable for several reasons. Since the known classical forces decay with distance one might hope, first of all, to reduce any effects to a negligible magnitude. Further, if the causal connection is mediated by a process which takes time, e.g. by particles or waves which travel at a finite speed, then at a sufficiently great distance the setting of one detector cannot be communicated to the other wing in time to influence its outcome. The further apart the two wings are, the longer messages would take to get from one side to the other.

In the classical regime these considerations yield at best plausibility arguments, not proofs. Gravitation and electrical forces may decay with distance, but nothing in the classical world-view requires this of all forces. And no fixed upper bound exists for the speed of propagation of classical influences. Although spatial separation may suggest causal isolation to the classical mind, it is only suggestion, not implication.

In the relativistic regime the situation looks rather different. The speed of light acquires a central role, serving in at least some contexts as an absolute physical limit. If it is a limit on all causal influences then the inference

Quantum Non-Locality and Relativity: Metaphysical Intimations of Modern Physics,
Third Edition. Tim Maudlin.

from "spatially separated" to "causally isolated" will become, in appropriate circumstances, valid. Our choice of "locality" as a name for the crucial isolation condition would then be vindicated: Einstein's systems A and B, if sufficiently far apart, have their own physical states and "external influence on A has no direct influence on B." Given a finite limit to the speed of causal processes, nor could external influences on A have an indirect (mediated) influence on B.

Our main order of business, then, is to determine whether Relativity really does forbid influences between sufficiently separated events. If so, then Aspect's experiments constitute both a vindication of quantum theory and a refutation of the Theory of Relativity. But even if Special Relativity does not strictly forbid superluminal processes, the situation is very different from that of classical physics. Due to the role of the speed of light in Relativity, superluminal processes must at the least have very exotic properties, signaling a radical change from the more familiar mechanisms of nature.[1]

Our goal in this chapter is to present the Special Theory of Relativity and its account of space-time structure. But as a propaedeutic to this task we begin with a review of some familiar facts about the maps we draw on space and time.

Coordinate Systems: Euclidean Space

We begin by considering the standard two-dimensional Euclidean plane. The plane contains an infinite multitude of points, so in order to describe figures in the plane we need some method for assigning names to the points. We could, in principle, simply invent names for various points, such as "the Grange" or "Chesney Wold," but the inconveniences of such a technique are patent. Arbitrary names convey no information about the relative locations of points, about their distances apart or the directions between them.

The solution to this problem lies in using numbers as names for the points and in assigning the names in a systematic way. At a *minimum* we would like the numbers to be assigned to the points in such a way that nearby points are assigned nearby numbers. More exactly, we would like to arrange things so that as we travel along a continuous trajectory from any point P to any other point Q the numbers assigned to the points we visit vary continuously from those assigned to P to those assigned to Q.

In a two-dimensional plane the only way to satisfy this continuity constraint is to assign (at least) a pair of numbers as coordinates to each point.[2] To coordinatize the plane we lay down two families of curves, the members of each family being parameterized by a real number. Let's call the first family of curves the x-family and the second the y-family. If we choose our families judiciously every point on the plane will lie at the unique intersection

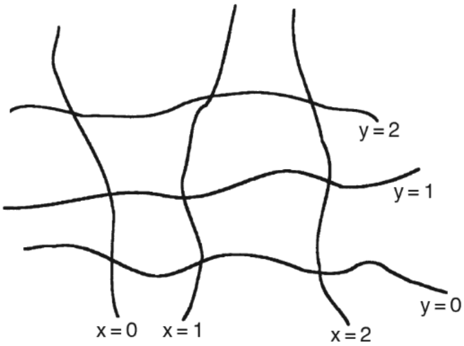

Figure 2.1 A Coordinate System

of one member of the x-family with one member of the y-family. Thus each point gets assigned a unique pair of numbers (x, y), its x-coordinate and y-coordinate.

There are infinitely many ways of laying down families of coordinate curves on the Euclidean plane. One such possibility is depicted in figure 2.1. The coordinate system of figure 2.1 does satisfy the continuity constraint, and so is significantly more convenient than assigning a random proper name to each point. When we get to the space-times of General Relativity in chapter 8 we will be quite satisfied with a system like that of figure 2.1. But given that we now have the Euclidean plane, curvilinear systems like the one depicted have several drawbacks. The amount of information we can infer given only the coordinates of two points is quite limited, since the curves in each family are so dissimilar from one another. We cannot infer, for example, how far the point (0, 0) is from (1, 2), or in exactly what direction it lies. Clearly we can do better than this.

First, it is convenient to use *rectilinear* coordinate curves rather than curvilinear ones so we know that so long as we keep to one such curve we will travel in a straight line. Since no two members of a family can intersect each other, this means we must choose families of parallel straight lines. Parallel lines always retain a constant distance from one another, so it will be helpful to have the numerical values assigned to the members of each family reflect the distance between them. Finally, calculations are simplified if we require the two families to lie at right angles to each other so we can use the Pythagorean theorem to compute distances. In this way we finally arrive at the familiar system of Cartesian coordinates, a rectangular grid.

We have gone such a long way round to a familiar goal to emphasize that Cartesian coordinates are just one of an infinite number of ways to assign numbers to the points in the plane. Further, in order for a Cartesian system to be possible at all, the surface being coordinatized must have several

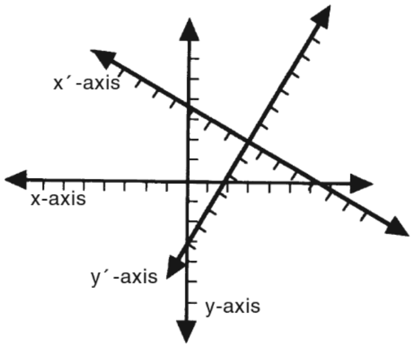

Figure 2.2 Two Cartesian Systems

special symmetries. Such a system cannot, for example, cover the whole surface of a sphere, or any finite region of it.

Even having settled on a rectangular, rectilinear coordinate system we are left with several choices to make. Many distinct Cartesian grids can be put down on a Euclidean plane. Two such systems are depicted in figure 2.2. The x–y coordinates differ from the x′–y′ coordinates in two ways: they have different origins and are oriented in different directions. As a result, every point is assigned a different set of coordinates in each system.[3] But there are equations which allow us to determine the coordinates of a point in one system given its coordinates in the other. For the example depicted in figure 2.2 the transformation equations are:

$$x = \frac{\sqrt{3}}{2}x' + \frac{1}{2}y' + 3$$

$$y = -\frac{1}{2}x' + \frac{\sqrt{3}}{2}y' + 2.$$

Thus the point (0, 0) in the x′–y′ system is called (3, 2) in the unprimed system and the point (3, 4) in the primed system is (7.598, 3.964) in the unprimed.

Let's simplify matters yet further by considering only Cartesian co-ordinates which share the same origin. In this case the systems can differ only in the orientation of their axes (figure 2.3). If the angle between the x and x′ axes is θ then the transformation equations are:

$$x' = x\cos\theta + y\sin\theta$$
$$y' = -x\sin\theta + y\cos\theta.$$

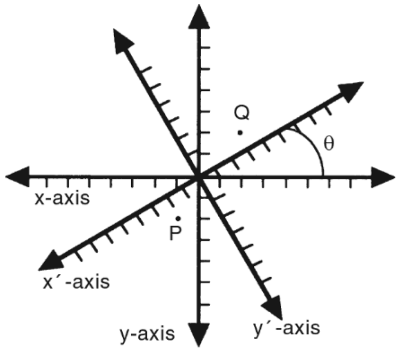

Figure 2.3 Rotated Coordinate Systems

Invariant Quantities

For the moment, consider two Cartesian coordinate systems which share an origin but whose axes are inclined at 30° to one another. The coordinate transformation becomes:

$$x' = \frac{\sqrt{3}}{2}x + \frac{1}{2}y$$

$$y' = -\frac{1}{2}x + \frac{\sqrt{3}}{2}y.$$

The point P in figure 2.3 would then be assigned coordinates (–1, –2) in the x–y system and $(-(\sqrt{3}/2+1), \frac{1}{2}-\sqrt{3})$ in the x′–y′ system. Similarly, the point Q is called (2, 2) in the unprimed coordinates and $(\sqrt{3}+1, \sqrt{3}-1)$ in the primed.

There are few things on which the two coordinate systems agree. Except for the origin, they assign different coordinates to all the points in the plane. Nor do they agree on coordinate differences. For example, the difference in x-coordinates between Q and P, Δx_{QP}, is 2 – (–1) = 3 in the x–y system but $\sqrt{3}+1-(-(\sqrt{3}/2+1)) = 3\sqrt{3}/2+2$ in the x′–y′ frame. Δy_{QP} = 4 while $\Delta y'_{QP} = 2\sqrt{3} - \frac{3}{2}$. Despite these differences, certain quantities remain unchanged when we pass from one coordinate system into the other. In particular, although both Δx_{qp} and Δy_{QP} change, the quantity $\sqrt{(\Delta x_{qp}{}^2 + \Delta y_{qp}{}^2)} = 5$ remains the same. This is, of course, just the distance between the points as calculated using the Pythagorean theorem.

We can now make a fundamental distinction between two sorts of quantities. *Frame dependent* quantities such a Δx and Δy change from reference frame to reference frame. *Invariant* quantities, such as the

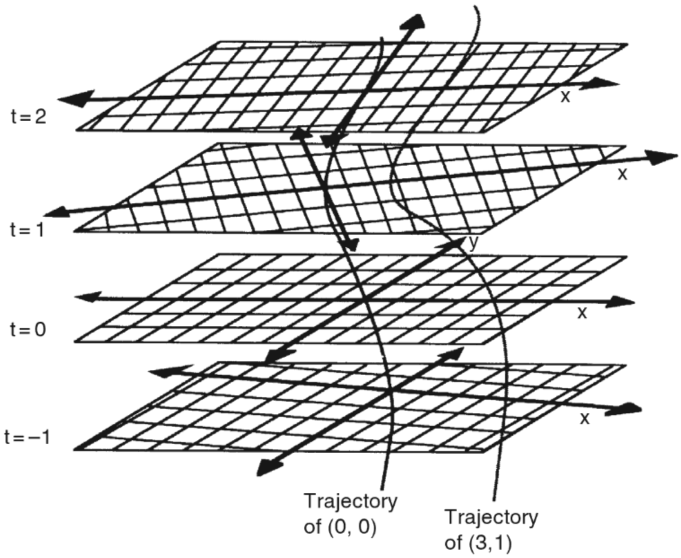

Figure 2.5 A Coordinatization of Space-time

such choice is depicted in figure 2.5. The implicit claim of such liberality in choice of coordinate systems is that the objective structure of space-time consists solely in the absolute division into simultaneity slices plus the spatial geometrical structure on each of those slices (plus continuity). This very weak objective structure is called *Leibnizian space-time*. It does not contain enough structure for classical physics.[4]

Consider the t axis, the trajectory of the point x = 0, y = 0 through time. In Leibnizian space-time, any trajectory which is always future-directed and which passes through the point (0, 0, 0) can serve as the time axis of a coordinate system (by convention, we will take the first coordinate to be time). This is in striking contrast to our coordinatization of three dimensional Euclidean space. In Cartesian coordinates the z axis in such a space must be a straight line orthogonal to the z = 0 plane. Cartesian coordinates are rectilinear and rectangular, and this greatly reduces the reference frames available to us.

So why not simply demand that the t axis of our reference system be a line which is both straight and orthogonal to the t = 0 hyperplane? Here we must proceed with extreme caution. Given our graphical presentation of the three-dimensional space-time it is easy to presuppose that there must be a distinction between curved and straight trajectories through space-time. But does such a distinction really exist? Are there any physical facts which would allow us to distinguish a straight from a curved trajectory, or to calculate the "angle" between such a trajectory and the $t = 0$ hyperplane? Each of these notions, that of a straight trajectory and that of an orthogonal trajectory, carry along substantive assumptions and cannot be admitted as legitimate without scrutiny.

Does nature make a distinction between straight and curved paths through space-time? The centerpiece of classical physics can be construed as answering in the affirmative. Newton's first Law of Motion reads:

> Every body continues in its state of rest, or of uniform motion in a right line, unless it is compelled to change that state by forces impressed on it.

In order for Newton's Law to *make sense*, for it to *make any claim at all*, there must be a distinction in nature between the trajectories of particles which are at rest or in uniform motion and those which are not. So there must be enough objective structure in space-time to found such a distinction.

Our first conclusion, then, is that there *is* a distinction in nature between uniform and accelerated motions. We further postulate that uniform motions trace out straight trajectories through space-time and accelerated motions follow curved paths. (This identification of uniform motion with straight trajectories is most natural since any two bodies each in uniform motion do not accelerate relative to one another.) Newton's first Law simply states that any body not subject to a force travels along a straight path through space-time. A body subject to a force will occupy a curved trajectory, and Newton's second Law states exactly how the path will curve. *Acceleration* of a body is nothing more than the curvature of its trajectory through space-time.

A corollary of Newton's first Law, then, is that we can legitimately demand that the time axis of our coordinate system be a straight line in space-time. Indeed, we can demand that every spatial coordinate sweep out a straight trajectory. Coordinate systems which satisfy this constraint are precisely the inertial frames of classical physics. The postulation of global inertial frames is a bold and risky hypothesis. There is no guarantee that nature has been so kind as to provide us with a space-time which can be coordinatized so that (1) the spatial coordinates trace out straight trajectories and (2) the spatial coordinates maintain the same distances from one another through time. It is a fundamental postulate of classical physics that such global inertial frames exist.

This leaves us still with the question of the orthogonality of the time axis. Does anything in nature distinguish a trajectory which is at right angles to a simultaneity hyperplane from one which is inclined at some other angle?

Again, Newton's first Law provides his answer. The first Law mentions a distinction between bodies in a state of rest and those in a state of uniform motion in a right (that is, straight) line. Both sorts of bodies travel along