1.05

PROCESS CAUSATION AND QUANTUM PHYSICS

Cynthia Kwai Wah Ma

SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
AT
THE DEPARTMENT OF PHILOSOPHY, LOGIC AND SCIENTIFIC METHOD
THE LONDON SCHOOL OF ECONOMICS AND POLITICAL SCIENCE
UNIVERSITY OF LONDON
LONDON, ENGLAND
JUNE 2001

THE LONDON SCHOOL OF ECONOMICS AND POLITICAL
SCIENCE
UNIVERSITY OF LONDON

The undersigned hereby certify that they have recommend to the Faculty of Graduate School for acceptance a thesis entitled "**Process Causation and Quantum Physics**" by **Cynthia Kwai Wah Ma** in partial fulfillment of the requirements for the degree of **Doctor of Philosophy**.

Dated: July 2001

External Examiner: ______________________

Research Supervisor: ______________________
Professor Nancy Cartwright (LSE)

Examing Committee: ______________________
Professor Phil Dowe (Tasmania)

____
Dr. Hasok Chang (UCL)

THE LONDON SCHOOL OF ECONOMICS AND POLITICAL SCIENCE
UNIVERSITY OF LONDON

Date: **July 2001**

Author: **Cynthia Kwai Wah Ma**

Title: **Process Causation and Quantum Physics**

Department: **Department of Philosophy, Logic and Scientific Method**

Degree: **Ph.D.** Convocation: **July** Year: **2001**

Permission is herewith granted to The London School of Economics and Political Science University of London to circulate and to have copied for non-commercial purposes, at its discretion, the above title upon the request of individuals or institutions.

Signature of Author

For my father, a chemical engineer,
a mathematician, a banker and
a philosopher in its truest sense...

Contents

Abstract

Philosophical analyses of causation take many forms but one major difficulty they all aim to address is that of the spatiotemporal continuity between causes and their effects. Bertrand Russell in 1913 brought the problem to its most transparent form and made it a case against the notion of causation in physics. The issue highlighted in Russell's argument is that of temporal contiguity between cause and effect. This tension arises from the imposition of a spectrum of discrete events occupying spacetime points upon a background of spacetime continuum. An immediate and natural solution is to superpose instead spatiotemporally continuous entities, or *processes*, on the spacetime continuum. This is indeed the process view of physical causation advocated by Wesley Salmon and Phil Dowe. This view takes the continuous trajectories of physical objects (worldlines) as the causal connection whereby causal influences in the form of conserved quantities are transported amongst events. Because of their reliance on spatiotemporal continuity, these theories have difficulty when confronted with the discontinuous processes in the quantum domain.

This thesis is concerned with process theories. It has two parts. The first part introduces and criticizes these theories, which leads to my proposal of the *History Conserved Quantity Theory with Transmission.* The second part considers the extension of the idea of causal processes to quantum physics. I show how a probability distribution generated by the Schrödinger wavefunction can be regarded as a conserved quantity that makes the spacetime evolution of the wavefunction a quantum causal process. However, there are conceptual problems in the interpretation of the wavefunction, chiefly to do, as I shall argue, with the difference in the behaviours of probabilistic potentials between quantum and classical physics. I propose in the final chapter, the Feynman Path Integral formulation of quantum mechanics (with the Feynman histories) as an alternative approach to incorporating the probabilistic potentials in quantum physics. This account of how to introduce causal processes in quantum mechanics fares better, I claim, than the previous one in dealing with the situational aspect of quantum phenomena that requires the consideration of events at more than one time.

Acknowledgements

First and foremost, I would like to express my deep gratitude to my thesis supervisor, Professor Nancy Cartwright for introducing me to this research and her inspiration. In particular, her inexorable insistence on my careful study of Russell's argument has sown the seeds of many of the thoughts on the subject of this work. I also thank her for granting me the intellectual freedom to pursue a philosophical topic I find truly interesting. Her support (both professionally and on a personal level) and permission to become a research associate on the project "Modeling in Physics and Economics" - which she directed at the Center for Philosophy of the Natural and Social Science (CPNSS) at the LSE - is gratefully acknowledged. From her, I have learnt the art of articulation.

I would like to show my appreciation to Dr. Carl Hoefer for his guidance and patience, his philosophical insights and un-selfish sharing of the secrets of thesis writing, and kind encouragement throughout the past two years when he acted as co-supervisor of my thesis. As I put the final touches to this thesis, Dr. Hoefer has been named the next Director of CPNSS - an appointment for which he is well deserved. I extend to him my hearty congratulations and good wishes.

There are many individuals who have contributed in various ways during my research. Three special people deserve specific mention. Professor Miles Blencowe has always been generous with his insights in theoretical physics and foundational issues and has encouraged me to think about physics cleverly and creatively. Professor Chris Isham is held responsible (favorably, of course) for initiating my interest in the foundations of physics when I attended his lectures on quantum mechanics as a second-year physics undergraduate at Imperial College in the Fall of 1984. He has since been a continuing source of insights and inspiration. I have learnt much and grown to appreciate real laboratory experimental metaphysics from Professor Yanhua Shih. All have taken great pains to carefully prove-read parts of what probably seems to them a rather obscure and cumbersome philosophical manuscript and I hope I have responded positively towards their challenging comments by expressing my ideas in a more articulate manner. I take this opportunity to thank them all for the numerous enlightening discussions. More importantly, their encouragement, infinite patience and friendship have seen an unconventional student of philosophy through a most difficult and uncertain period of her research.

Others, too, have in their own special ways inspired and influenced me with their ways of thinking, insights, discussions and conversations. The following list is by no means complete and exhaustive: Dr. Joseph Berkovitz, Dr. Harvey Brown, Professor Jeffrey Bub, Professor Jeremy Butterfield, Dr. Bob Coecke, Professor Raymond Chiao, Professor Francesco De Martini, Dr. Marco del Seta, Professor Lajos Diosi, Dr. Fay Dowker, Professor Augusto Garuccio, Professor Donald Gillies, Dr. Jonathan Halliwell, Dr. Etienne Hofsetter, Professor Bei-Lok Hu, Professor H.I.G. Hughes, Dr. Makoto Itoh, Dr. Adrian Kent, Dr.Yoon-Ho Kim, Professor Willis E. Lamb Jr., Dr. Fotini Markopolou, Professor David Malament, Professor Sarben Sarkar, Professor Marlan Scully, Dr. Orly Shenker, Professor Lee Smolin, Professor Rafael Sorkin, Professor Abner Shimony, Dr. Towfic Shomar, Professor Jan Von Plato, Dr. Ting Yu.

Thanks also go to the following individuals for friendship and encouragement: Ms. Klementina Balint, Dr. Sebastiano Bavetta, Dr. Vincenzo Berardi, Mr. Sze-Ting Chen, Dr. Michael Dash, Miss Sadaf Hafiz, Dr. Makiko Ito Harrison, Mr. William Irvine, Professor Josephine Lai, Dr. James McAllister, Mr. Robert Northcott, Ms. Antigone Nounou, Dr. Miya Paczuski, Miss Dorota Rejman, Dr. Michael Selby, Mr. Neeshmal Shah, Dr. Antony Valentini, Mr. Rafaelo Vardavas, Mr. Vincenzo Vitelli, Mrs. Kate Workman and Mr. George Zouros.

The staff members at the Imperial College Central Library have gone out of their ways to provide a haven during the research and writing phases of this thesis. My warm appreciation goes especially to: Bert, Clive, Emma, Ken, Kenny, Monica, Rita and Ronnie; who have all made me feel welcomed and at home.

Simon, the "big-bear" has occupied an episode of my life, with his human strengths and weaknesses. Strangely enough, our interactions have led me closer to discovering my true self, which perhaps, one day, he will develop the sensitivity to understand.

Last, but not least, I wish to express my love and gratitude to my family. Dad has imbued in me a life-long interest in physics and a passion for academic pursuits; since he first introduced to me the delights of algebra when I was at the tender age of nine. Mum is very much loved and thanked for showing endless interest, care and support in every single way through my early school years (and ever since) that laid the foundation for my intellectual developments. My sister Iris and brother Kenneth both have an extraordinary understanding of my character and have provided their unfailing support all through the years. Their constant looking up to me as the "big sister" figure has granted me strength to carry on when times are hard. Kenneth showed his unique way of support by racing me to the "Dr." title, which he has so admirably won last June (though, however, his title is of the medical profession!).

Whilst writing, I received news of the untimely passing of Professor Wesley Salmon.

As the philosophical community laments the loss of a rare talent, I dedicate Chapters 5-7 to his memory.

Cynthia K.W. Ma

London, June 2001

Chapter 1

Prelude[O]

1.1 Introduction

Causation is an active area of philosophical research and it is one notion that is deeply entwined with the foundational aspects of both classical and quantum physics.

When asked, "What is Causation? What do we mean when we think that one object is the *cause* of another or one event causes another to happen?", no doubt different connotations would immediately come to mind. And indeed we ought to ask "What are the connotations of causation?" These basic connotations, being largely empirical in character in the sense that they come from our experience and interactions with the physical world, generally receive different treatments by physicists and philosophers. The distinction is a matter of the difference in practice. Physicists accept these basic intuitions as facts about causation and physical theories are constructed to conform to these "conditions of causality", which are not to be violated. A good example would be that of the "past-future" directed Minkowski light-cone structure defined in terms of "cause-effect" relations[1]. Philosophers, on the other hand, approach the subject from a different angle; they conduct conceptual analyses of these causal connotations and see whether they do make good logical sense or are infected with grave inconsistencies. With the advent of relativity theories, physics has added important items to the stock of causal facts. Perhaps the most significant one is that special relativity places a limit on the velocity of propagation of causal

[0] Adopted with modifications and expansion from the paper by C.K.W. Ma (2000) p.631-641.
[1] See for example, *Taylor and Wheeler* (1966), p.39.

influences. For the serious philosophical minds, these results cannot afford to go unnoticed and it would indeed be of considerable interest to investigate the extent of the possible interplay between the findings from the respective disciplines.

With this motivation in mind, the plan of this chapter goes as follows. In Section 1.2, I shall consider a number of basic causal connotations and the various aspects that may be deduced from these considerations. We are then in a position that leads naturally to the main ideas of David Hume's theory of causation. Hume's theory is the start of the empiricist philosophical analysis. The Humean account and its more modern variants collectively still represent the predominant philosophical view on causation among anglophone philosophers. However, in a clever paper in 1913, Bertrand Russell was able to show that the Humean view is not free from inconsistencies given an important assumption on the nature of time. Russell's argument will be examined in Section 1.3. Section 1.4 focuses on the issue of causal continuity and the recent physicalist approaches to causation that attempt to resolve some of the more pressing difficulties associated with this issue.

It is my modest aim to bring into focus, in the following pages, some major philosophical worries on the subject of causation, which may be fairly regarded as one of the underlying puzzles encountered in the foundations of both classical and quantum physics.

1.2 Causal Connotations and David Hume's Theory of Causation

When one event (or something[2]) is regarded as the *cause* of another, we have "postulated"[3] the existence of a special relation or *connection* between the two events. How special is it? We may want to emphasize the importance of such a connection by the expression that the event we have chosen to call the *cause* and the one that is called

[2]What kind of entities does the causal relation relate is an important aspect of philosophical analyses of causation. Some argue that the "*relata*" should be events, while others insist that they may be facts, processes or states of affairs. However, as both Hume and Russell take events as the proper causal relata, we may therefore focus on events in the present work.

[3]Notice I have deliberately used the word "*postulated*" because it is only correct to remain philosophically neutral and avoid making undue assumptions from the outset.

the *effect* are *necessarily* connected to each other so that *had the cause not happened, the effect would not have happened either.* Put slightly differently, this counterfactual mode of representation of the special connection refers to an element of necessity in the sense that given the cause, the effect "must" follow and any other situations just simply cannot and would not happen.

How then are we to discover this necessary connection, whatever it may be? One useful place to look is to start from our observations of how causes and effects behave generally. An obvious observation is that "*causes precede their effects*"; namely, causes occur earlier in time than their effects. One realizes of course that not every pair of events happening at the same two respective instants of time are to be thought of as causes and effects. A concrete everyday example illustrates (Figure 1-1). Suppose we have two individuals, Angelo and Bianca, standing side-by-side in a room and Angelo, who is nearer to the light switch, turns it on and at the *same moment in time*, Bianca starts to sing. Even though Angelo's action and Bianca's singing are *both* events happening *prior to* the lamp being lit up, we would deem it appropriate to attribute the cause of the lamp lighting to the switching action provided by Angelo but not to Bianca's singing. Why? It is in part because a *continuous* physical connection is envisaged between the light switch and the lamp and there is in general no obvious and direct correlation between the processes of singing and the lamp lighting.

So temporal succession between two events alone is not a sufficient condition for causation. To place the matter in a more scientific perspective, consider the Minkowski lightcone in Figure 1-2. Events E_j and E_k[4] both lie in the future lightcone of the event E_i and hence are *causally connectible* to E_i since signals sent from E_i can reach either E_j or E_k. However, neither E_j nor E_k is *necessarily connected* to E_i for there *need not* be an existing connection after all. Whether there is in fact an actual connection depends upon the existence of actual physical processes linking E_i to E_j and/or E_k. Therefore, ***temporal succession and the***

[4]These are treated as simultaneous events that lie on the same hyperplane, but the argument would remain valid even if they do not.

<u>spatio-temporal continuity</u> between the cause and effect provided by physical processes together seem necessary of causation.

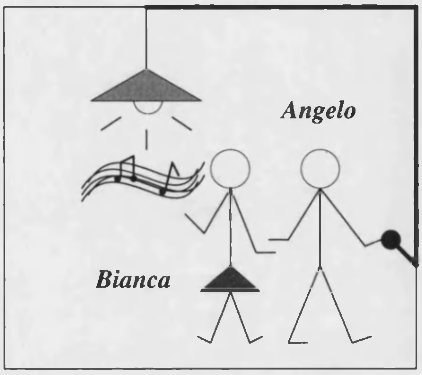

Figure 1.1:

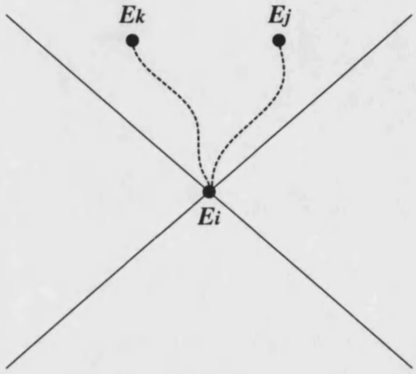

Figure 1.2:

Granted the physical connection between the light switch and the lamp, is this connection really necessary, in the sense that contrary situations are precluded from occurring? Unfortunately, the answer is in the negative. Although it takes an awfully short time for the electrical signal to travel from the light switch to the lamp, it is, however, always conceivable that an accident like a power cut may occur within this short time interval, as a result of which the lamp would fail to light up. It is therefore *not* necessarily the case that *whenever* the switch is turned on, the lamp *must* light up; this is only true if no other factor interferes.

Let us now imagine instead the scenario where Bianca is the only person in the room and there is no light switch attached to the wall. We observe that when Bianca starts to sing the light comes on a split second later. On one mere instance of this observation, it would be reasonable to put it down as a case of sheer coincidence because we do not usually conceive of a possible (physical) connection between these two events. However, if such an observation is repeated many times and each and every time the same sequence obtains so that whenever Bianca starts to sing, the light comes on, then we would no doubt conclude that the occurrences of both events in close temporal succession are too regular to be discounted as pure chancy coincidences. And so from repeated observations of the regular succession of the two events, we find it proper and indeed justified to "infer" a special connection (Figure 1-3) between this pair of events and set out to search for the "hidden" mechanism that could have been responsible for giving rise to such a correlation.

The question remains: is such a connection we have so inferred upon repeated observations of regularities a *necessary* one? Although experience teaches us that frequent correlations are usually *prima facie* good indicators of causation, it is however well-known that correlations *do not have* to imply causation. Logic does not prohibit the apparent correlations we see as arising from pure chance. One may well imagine the world to be a chancy enterprise in such a way that "it so happens" that whenever E_1 occurs, E_2 follows later. Still, one may argue that quite unbeknown to us, there could exist some kind of a voice-recognition device which provides the physical connection between Bianca's singing and the lighting up of the lamp. But what gives us the impression and prompts us to look for this "unknown device" in the first place? Nothing other than the constant conjunction of the two events of the singing and the lamp being lit - *given the same cause, the same effect follows.* Once again, such a physical connection is by no means necessary as for instance, a power cut may occur and tamper with the normal functioning of the device, thus rendering the succession of the two events unattainable.

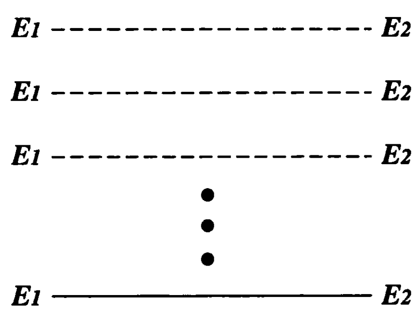

Figure 1.3:

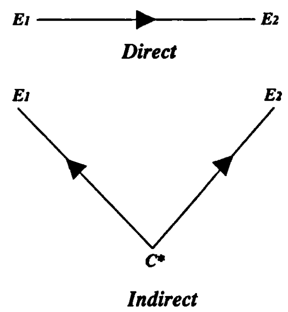

Figure 1.4:

Temporal succession of two events is not sufficient for causation. A case of causation is deemed to obtain when temporal succession is supplemented by the presence of a spatially and temporally continuous physical connection which provides the link between the two events in question. Similarly, the situation where the same cause has found to be followed by the same effect on a great number of occasions also gives us the feel of a causal correlation and induces us to look for an underlying connection. It is in this way that it is often thought that *spatio-temporal continuous physical processes are essential for making sure that "the same cause is to be followed by the same effect"*. Unforeseen circumstances may happen during the spatial and temporal course of this continuous physical process and frustrate the connection between the two events. Due to the absence of necessity, in the sense that unforeseen circumstances can always occur, it follows that there can never be any guarantee without fail that the same cause is to be followed by the same effect. This clearly indicates an incompatibility between these two causal connotations and that in reality, "*the same cause is always followed by the same effect*" is <u>not</u> warranted by "*spatio-temporal continuity*". This is quite contrary to our ordinary way of thinking.

Spatially and temporally continuous physical processes, though they are by no means necessary connections, do nevertheless impress upon us the idea of the existence of causal relations by way of the physical connections. Such physical connections between two events may, however, either be (i) *direct* or, (ii) *indirect*. In the latter case, we seek a third event C^* that existed in the overlap region of the respective past lightcones of the two events so that both events are direct consequences of C^*. In this model, both of the events are effects of the common cause C^* (Figure 1-4). This illustrates nicely the appeal of hidden variable programs in the efforts to provide an explanation of the Einstein-Podolsky-Rosen (EPR) paradox. Special relativity rules out a direct physical connection between the two space-like separated measurement events. However, the existence of the remarkable correlations of the results of measurements calls urgently for an explanation. It is thus natural to look for possible common causes (hidden variables) in the past that could have given rise to these correlations. But, while such a procedure may satisfy one's intellectual urge, logic permits a world in which these correlations are all there are; the empire of chance rules in such a manner that the correlations always obtain *even without* any underlying spatio-temporal continuous connection, be it direct or indirect. In fact, the significant achievement of Bell's 1964 inequalities[5] rests on their success in dispelling local hidden variable models in favour of quantum mechanics, as has since then been so forcibly confirmed by many sophisticated experiments[6].

In relation to the *spatio-temporal continuity between causes and effects*, another common connotation we have for causation is to suppose that a certain cause event has occurred but nevertheless the expected effect has somehow failed to materialize. The situation is usually explained by the presence of other events that must have got in-between them and inhibited the occurrence of the effect. In other words, the spatio-temporal continuity between the supposed cause and the supposed effect is disturbed. In order to diminish the opportunity for other events to go "in-between" and behave mischievously, it is desirable to make both the *spatial* and *temporal* intervals between the cause and the effect as short as possible. The shorter these intervals are, the smaller the probability generally for other factors to interfere. Causes and effects are expected to be spatially and temporally close to each other, so that they should exhibit a degree of spatio-temporal "nearness". For events happening at vastly separated spacetime locations to be causally connected, we look for events to provide the intermediate links between these two spacetime locales, and hence, we arrive at the idea of a "causal chain" to ensure causal continuity across spacetime regions.

The foregoing discussions now lead appropriately to the introduction of the major ideas of David Hume's theory of causation, which exerted tremendous influence over the logical positivists and their contemporaries such as Bertrand Russell.

Hume maintains that there are two basic elements to human understanding that form the pillars to his philosophical system: *impressions* and *ideas.* Impressions correspond to all "lively signals" we receive from the physical world through our senses, like perceptions, sensations, feelings etc. Ideas, on the other hand, consist in the formation of a conception of the impressions. The general principle he adopts for his philosophical analyses is that: "*all ideas originate from the association and combination of the different impressions*". So any idea we may possess for some entity

[5]Bell, J.S. (1964).

[6]See for example, Aspect et al. (1976) and Shih and Alley (1986).

has to come from our perceptual *experience* with it through our senses.

Armed with this principle, he then asks, from *which impressions* do we form the *idea of cause-and-effect* as some sort of a *necessary connection*? He is able to identify three such impressions from our empirical experience of two events behaving like causes and effects. These are: " ***priority in time***" of the cause , "***constant conjunction***" between the cause and the effect and " ***contiguity***" in space and time between causes and effects. And these should be of some familiarity to the reader since they refer to none other than the three causal connotations we have considered in the above: "*causes precede their effects*", "*given the same cause, the same effect follows*" and "*continuity*" respectively. But as we have already discussed, from these three properties and these three alone, one can never deduce the element of necessity. Hume argues that even though there may actually exist connections in the world which are necessary in the above sense, beyond this, the only real idea we can have of this connection is of the three properties above. Since these properties are not sufficient to entail necessity, philosophical prudence must now compel us to take a skeptical view of the idea of necessary connection between causes and effects.

Granted that our experience is incapable of furnishing us with the idea of a necessary causal connection, how are we able to associate the three impressions of causes and effects to arrive at the idea of a necessary connection between two events? Hume answers that *after many instances of observing* the behaviours of constant conjunction, priority in time and contiguity in time and space between the two events *without exception*, the mind has in the course grown accustomed to expect that the second to follow the first on any new occasion. This feeling of expectation thus leads to the impression from which our idea of connection is copied. Therefore, the idea of a necessary connection comes not from our experiences of the external world but rather orignates from our own response to it. In other words, the causal relation as a necessary connection is an idea "imposed" by the mind upon the unfailing, successive observations of these regular behaviours of causes and effects. The three impressions of priority in time, constant conjunction and contiguity in time and space can never provide us with the idea of a necessary connection.

It must again be emphasized that it is never Hume's intention to deny the existence of necessary connections in Nature. Rather, the three impressions are all we possess by way of a source for the idea of causal necessity. Since this evidence alone is not adequate to reveal to us such an element of necessity, it would be more reasonable not to impose its existence on Nature, leaving this instead as an open question. And Hume concludes[7],

> ***As to what may be said, that the operations of nature are independent of our thought and reasoning, I allow it; and accordingly have observed, that objects bear to each other the relations of contiguity and succession; that like objects may be observed in several instances to have like relations; and that all this is independent of, and antecedent to the operations of the understanding. But if we go any farther, and ascribe a power or necessary connection to these objects; this is what we can never observe in them, but must draw the idea from what we feel internally in contemplating them.***

[7]Hume D. (1888), p.168-9.

1.3 Russell's Objection to Hume's Temporal Contiguity Thesis

In 1912, Bertrand Russell made his presidential address[8] to the Aristotelian Society the occasion to cast doubt on the tenability of the Humean account of causation and to argue against the notion of cause in physics.

We have already taken pains to stress the inherent incompatibilities among the three causal impressions of *priority in time*, *contiguity in space and time* and *constant conjunction*. In particular, it has been indicated that in the absence of the ingredient of necessity, spatio-temporal continuity is not really capable of ensuring the constant conjunction of the causes and effects. The main reason for this is that even if there is a continuous spatio-temporal physical process connecting the cause and the effect, anything can still happen during the time interval while the causal influence is transmitted down the connection and this results in an uncertainty in the production of the effect. The light switch and the lamp in the last section form a good example. This is why events which are too removed from each other in both spatial and temporal dimensions are not considered as reliable causes and effects.

An immediate solution would be to require that both the spatial and temporal distances between the two events be decreased to such an extent for them to stand "*adjacent*" (or *contiguous*) to each other, so that other factors cannot impose themselves and thwart the occurrence of the effect. But what exactly does one mean by two events being "*adjacent*" to each other when they are embedded in a background of spacetime continuum? The criterion of spatial contiguity between two events is easily satisfied and in the limit it is met by the case where two events, happening at different times, may indeed occupy the *same* location in space. The notion of temporal contiguity is, however, more problematic since given that two events occur at the same spatial location with one after another, how are we to ensure that they are temporally contiguous to each other? This problem reduces to one that concerns temporal contiguity and this is indeed the important issue addressed by Russell in

[8]Delivered on 4 November 1912. The ensuing essay was published in the *Proceedings of the Aristotelian Society*, 13 (1912-13) and reprinted in Russell, B. (1917), p.180-208.

his paper.

Russell's argument begins with a statement of the temporal contiguity thesis (***TC***). The properties of *priority in time* and *temporal contiguity* between cause and effect can be summarised as follows:

> ***TC: Whenever the first event (cause) ceases to exist, the second comes into existence immediately after.***

To place ***TC*** in the correct perspective, Russell makes the major assumption that time is to be modeled as a mathematical continuum and is therefore considered as a dense series. A dense series has the distinctive feature that the notion of a "next point" does not make sense because between any two points there always exist others, no matter how close these two points are to each other. It is instructive to contrast the idea of a dense series such as the real number line with the discrete series of positive integers where the notion of consecutive (or "next") members does take on a well-posed meaning. Having specified how the temporal continuum is to be represented, we now consider two point events c and e occurring at two respective instants of time t_1 and t_2 ($t_1 < t_2$). Because time is a dense series, it follows that between any two instants (points) of time, there are always other instants (points) no matter how short we make the interval $t_2 - t_1$. That is, there exists always a temporal gap between c and e and so c cannot be regarded as contiguous *in time* to e. Furthermore, this temporal gap provides ample opportunities for other events to creep in between c and e and to interfere. Whilst these other factors may not prove harmful to the production of e at t_2, they may, however, also behave otherwise and hinder the occurrence of e (Figure 1-5). Under these circumstances, one cannot be certain that the same cause is always followed by the same effect since there can always be the chance of e not occurring *whenever there is to be this temporal gap between the two events.*

In order to be rid of unsolicited factors, one must devise a means to ensure that the temporal gap is filled. An obvious way to accomplish this is to suppose the cause event as having a temporal dimension (Figure 1-6). Russell argued that in this case the cause would have to be a *static, unchanging* event[9], occupying the half-open interval just previous to the effect. For imagine the cause changed. Then we would need to postulate a cause-effect relation between the part of the cause before the change and the part after, and we would be no further ahead. Thus the cause must be supposed to sit there from time t_1 to t_2, filling the temporal gap and all of a sudden, turns into e at t_2. However, Russell objects: he argues that it is not at all logical why, being unchanging and sitting there complacently, c has to turn into e at t_2 but not at any other moment, say the earlier t_0 or the later t_3?

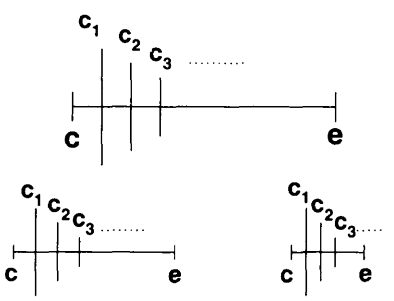

Figure 1.5:

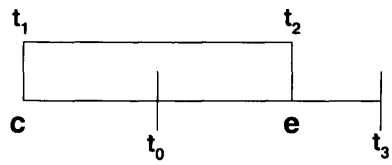

Figure 1.6:

And so static, unchanging events are dismissed outright by Russell as an impossibility. But these static, unchanging events seem to be the only means by which the temporal gap can be occupied. Since these are not plausible, one must draw the conclusion that there always exists a temporal gap between c and e as a result of which c cannot be contiguous to e. Our intuition about the temporal continuity of causes and effects comes under threat given the assumption of physical time as a mathematical continuum and as a consequence, constant conjunction cannot be guaranteed. Russell has succeeded in showing that there exist tremendous tensions between our usual connotations of the causal relation.

[9]For a non-static, changing event such as one composed of a causal chain of discrete events as in Figure 1-5, the problem of temporal gaps existing in-between these events within the causal chain remains.

1.4 Causal Continuity and Recent Physicalist Accounts of Causation

Despite the difficulty brought to light by Russell's critique of the Humean temporal contiguity thesis, one is, of course, entitled to argue that the major issue is really the feasibility of the definition of events as points occurring at *discrete* temporal instants within the temporal continuum. There is simply no place for the notion of discreteness within a temporal continuum. A more congenial approach would then be to "superpose" a continuum of events - *a continuous rope of events* - upon this temporal continuum, in the sense that we consider *all* the events that happen locally within this time interval. One finds a convincing example on causal continuity from Elizabeth Anscombe[10]: "*Find an object here and ask how it comes to be there?*" (Figure 1-7),

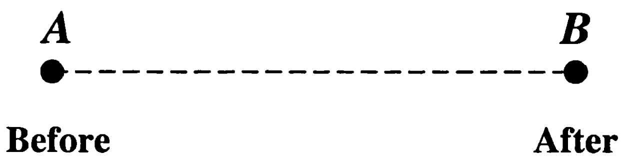

Figure 1.7:

A *causal* explanation, says Anscombe, would be "*it went along some path from A to B*". The locution "along some path" in fact entails more than the case where the object just turns up at location B after having been at A previously. It requires the object to occupy also *all* the intermediate positions between A and B. To satisfy *constant conjunction*, it is sufficient for the object to turn up at location B after having been at A some moments earlier and *without* having to assume the intermediate positions between the two locations. But this would not be deemed to be an adequate causal explanation[11]. And so to "explain causally", a path has to be imposed to provide the *connection* between the two events of the object being at the two respective

[10]Anscombe, E (1974), p.150.

[11]One reason is that the two locations A and B may be situated very far apart. Given that reliable causes-and-effects are thought to be events that occur locally to each other in space and time, so although it may be the case that there is constant conjunction between the two events of the object occupying the respective "far-away" locations; they would certainly not be deemed to be genuine cause-and-effect in accordance with what is generally perceived as a causal situation.

spacetime locations. For the purpose of explanation, it is therefore proper to consider spatio-temporal continuous connections when thinking about causation.

Consider the simple case of one mass in motion colliding with another that is at rest (with both masses being apart initially), and subsequently setting the second into motion (Figure 1-8),

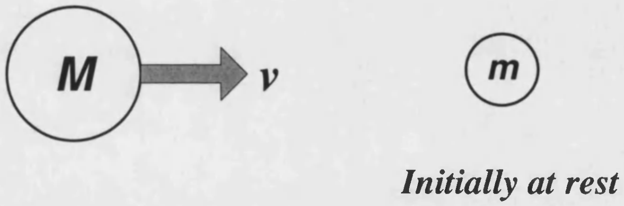

Figure 1.8:

Taking causes and effects as events, one may speak about two kinds of "events" in this circumstance. One would say that the collision event between the masses M and m is responsible for bringing about the dynamical changes in both masses. However, one may also extrapolate further and assert that it is the *motion of* M which causes the collision in the first place. It would therefore seem that "the motion of M" is the "*cause*" of the collision. It is the motion of M through spacetime that brings its closer to and triggers the collision with mass m that is at rest. It is the motion of M that provides sufficient momentum and energy required for the collision to occur (two masses at rest in contact would not give rise to a collision, even though they are in close encounter in a localized region of spacetime). But may we speak of the *motion of* M as an event?

An event is often thought to occupy a localized position in spacetime. That is, events are thought of as occurrences at spacetime points. Here, however, the motion of M is an entity that spans both a spatial and a temporal extent, one which is composed of all the point-events that correspond to the different positions taken up by M at different times during its motion. Any one point-event on its own is unable to achieve the effect - the event of collision. But each and every one of them contributes to the production of the effect; each subsequent event brings M closer to m. It would therefore not be sound reasoning to single out any one such point-event, any one specific position of M at a certain instant of time and ask how it might relate to the collision.

The point is that each of these point-events derives its significance from the contribution[12] it makes to the overall motion of M. Each of these events are prior in time to the final event of collision but would only be deemed meaningful as a part of the cause (Figure 1-9),

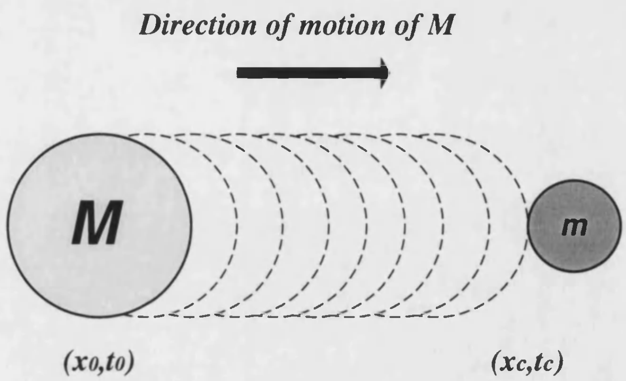

Figure 1.9:

To emphasize the fact that it is indeed the entire collection of these point-events that is to be considered as the cause and that such a cause is an extended entity in both the spatial and temporal sense, we introduce the term "***process***" to signify this series of point-events.

Now suppose that an impulse is applied to M at position x_0 at time t_0 and set it into motion. The motion carries M through all the intermediate spacetime points and arrives at x_c at time t_c when it collides with m, which is at rest. In the spirit of Russell's temporal gap, one may now ask: given the earlier event E_0 that corresponds to the application of the impulse at (x_0, t_0), how is this one event to *bring about* the later E_c that corresponds to the collision between the masses M and m? Our reply would be: the two events of E_0 and E_c are connected by the *continuous* spatiotemporal motion of M - a "continuous rope of point-events". It is by virtue of this motion that energy and momentum supplied by the impulse in the first place get transported from the point of application to the location of collision. The motion of M - the *causal*

[12]These may be regarded as equal contributions.

process - provides the missing link to restore spatiotemporal continuity

In physics, the motion of a system is usually represented either by spatially and temporally continuous paths or by trajectories in phase space. These spacetime paths and phase space trajectories are the solutions of differential and integral equations - the expressions of spatiotemporal continuity. Whilst these differential and integral equations guarantee continuity, it is, however, a fact that they lack the crucial causal aspect of an explicit temporal order for cause and effect.

Consider once more the example of mass M colliding with mass m and setting the latter into motion from an initial state of rest. The same state of affairs can be described by two different causal stories. In the rest frame of m the moving mass M travelling with velocity v appears to be the *earlier* event - the cause which is responsible for the change of states of both masses. On the other hand, in the rest frame of M, the *earlier* event of m moving with velocity -v is now regarded as the cause giving rise to the subsequent change of motion of both masses. Hence, we find ourselves confronted by two different causal stories with accuracies depending on the frame of reference in which the same state of affairs is viewed (Figure 1-10),

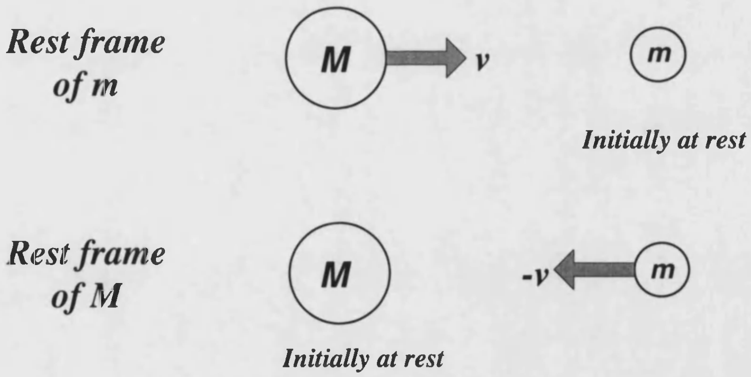

Figure 1.10:

The *objective* matter-of-fact is, however, that for *all* inertial frames of reference, the "collision" between M and m produces the subsequent "changes" in the motion of each of the masses. The *collision* does indeed occur in all frames of reference, after which is followed by changes in motions of these masses. It is indeed potentially with respect to this event that a temporal ordering may be established. Because of this "causal interaction" between the masses M and m, their respective energies and momenta are modified accordingly. This interaction may be thought of as an encounter between the two masses as their respective path crosses. Both masses, having interacted, will carry the causally modified dynamical properties via their continuous spatio-temporal trajectories and may participate in further interactions.

The idea of a continuous spatiotemporal path hence forms the backbone of the physicalist theories of causation. Although there are variations amongst these theories of physical causation that have been put forward, they nevertheless share one basic underlying idea: *causal continuity is guaranteed by the transmission of causal influences (objective physical quantities) along continuous space-time paths governed by physical laws.* The "objective physical quantities" being transferred refer usually to either momentum or energy in the group of approaches subsumed under the title of "transference theories of causation"[13]. A more sophisticated version is the class of the so-called "process-theories" of causation as pioneered by Wesley Salmon in the early 1980's, which has subsequently undergone substantial further developments by both Salmon himself and Phil Dowe during the 1990's[14].

"*Casual processes*" and "*causal interactions*" thus form two fundamental causal notions of process theories. These approaches take ***processes***, entities construed as spatio-temporally continuous paths of physical objects, as the basic unit of causal analysis. Events - the interactions - are fixed by the crossings of processes on these accounts. When two processes meet so that changes are brought to at least one of them as a result, then it is said that the processes "*interact*". Interactions are thus of prime importance for they are responsible for the production and modifications of the participating processes by introducing changes in their dynamical properties and structures. The two notions of process and interaction lie at the very heart of process theories of causation.

[13]Dieks, D. (1981), Ehring, D. (1986), Fair, D. (1979).

[14]The original idea of process causation was advanced by Wesley Salmon in a series of papers during the 1970s (see Salmon, W.C. (1977) and (1980a)), which later received a more complete and systematic treatment in his book, Salmon, W.C. (1984). More recent modifications and improvements of Salmon's 1984 account are: Dowe, P. (1992a) and (1992b), Salmon, W.C. (1994) and (1997).

Processes are the vehicles by which causal influences, such as physical information and signals, are propagated. Any changes in the make-up of a process are due to its interactions with other processes. Perhaps the analogy with the special relativistic representation of events may be too plain a correspondence to deserve any attention. On Minkowski spacetime diagrams, all the encounters experienced by a physical object, that is, a history of events which involves the object, are connected by a spatio-temporally path, the *worldline* of the object. When the object interacts with another object, the interaction is an event common to the histories of both objects and is hence represented as the intersection of both worldlines. So if we take the analogy at face value, the processes may then be identified with the worldlines of objects and the interactions given by the crossings of these worldlines. This is the current view held by both Salmon and Dowe.

One of the most significant impacts of the theory of special relativity on our worldview of nature is the existence of an upper limit on the velocity of propagation of information and signals. This serves as a limit to facilitate the decision as to which events can be physically connected and which cannot, on the understanding that a light signal may be sent from one event and received by another after a lapse of time. However, there are also events which are so vastly separated in space that both happen within an extremely short time-interval; too short for a physical signal to travel over from one event to the other and to have influenced it. In such cases, successful information or signal transfer would require a speed of propagation to exceed c, the speed of light, which is blatantly prohibited by special relativity. So if in the first instance, to be "physically connected" is a prerequisite condition of being "causally connected" for two events, then the process that provides this connection have to carry the physical information or signal at a speed less than or equal to c. Indeed, this is one important reason why Salmon subdivides the set of processes in the world into *causal* and *non-causal* ones with respect to the propagation of physical influences. This is, however, not all there is to the story, for there exists yet another strong reason to segregate the two kinds of processes. This is to do with two events being "*causally connected*" and their being "*causally connectible*". We refer the reader to the discussion in Section 1.2 and Figure 1-2.

To cast the discussion in a more concrete context, imagine the case of a large circular building with a spotlight mounted on a rotating platform at its center, which Salmon used in first introducing his account of causal processes. Once the spotlight is switched on and set into rotation a spot of light will sweep across the wall of the building in circular motion. Although the moving spot of light appears to be spatio-temporally continuous and travelling at a speed that is below c, it nevertheless cannot provide any sort of connection between two events due to its inability to transfer information and signals. If the spot of light is intercepted by a red lens at one single spot on its path, it would become red at that very point but then such a change (the change from being a white light spot to one being red) would not be transmitted beyond this point: the information that it has *interacted* with the red lens is not permanently registered. The travelling spot of light "reverts" back to its original state of being a white spot once it has moved beyond the lens.

Why is this process not capable of carrying or transmitting this information? Contrast this spot of light with the light ray that travels from the spotlight at the center and impinges on the wall. The ray of light, being spatially and temporally continuous, is also a good paradigm of a process. Suppose a red lens is placed somewhere along its path, the light ray will be coloured red at this point and would remain red from then on. In this case, the information that it has been intersected and interacted with a red lens is registered and propagated down the process itself.

Looking carefully at each of the two situations, one notices that the moving spot of light does not have an independent existence of its own; it owes its existence to the impingement of the light rays radiating from the center spotlight on the wall. In other words, one point within the moving light spot is not the *source* of the next and hence any intervention with this "process" only affects the very point where the intervention is introduced but not others. So the basic intuition is that though the moving light spot does appear to be a uniformly continuous process in space and time, it is in reality constituted of <u>*a series of unlinked interactions*</u> between the light rays and the wall of circular building; this is quite unlike a self-sustaining process such as a light ray, with its every stage being physically connected to the previous (and also to the next) via the transfer of energy (and hence information) between the various stages of the process. It is by virtue of transmission of causal influence that the successive stages of a proper causal process are linked. It is thus of paramount importance that only processes that are endowed with the ability to *transmit* the results of changes introduced by their interacting with other processes are regarded as *causal* ones. In short, the capability of some processes to transmit causal influences like energy and information is the criterion for being causal processes.

Both the theories of Salmon and of Dowe strive to capture how causal influences are carried across spacetime regions in genuine causal processes by imposing their respective criterion for when a transference has occurred. Dowe maintains that the "*possession of a conserved quantity*" by an object along its worldline is a sufficient condition to ensure the transference of causal influences. Salmon, however, argues that Dowe's condition of "*mere possession*" does not capture the fact the the possession of a conserved quantity at each of the spacetime points must not arise from interactions.

The foregoing serves only as a concise outline of the contents of Chapters 3 to 4 of this thesis, where we shall follow the development of the process causal theories closely. Salmon's 1984 theory of process causation will be introduced in Chapter 3 and the key problematic issue of counterfactuals that plagues Salmon's program will be discussed in detail. The subsequent developments form the main subject of Chapter 4, which leads to my proposal of the "*history view*" that takes the place of the worldline view of process causation in order to pave the way for the consideration of discrete processes in the quantum domain.

But first, in Chapter 2, we shall present Russell's argument against the temporal contiguity between cause and effect, to which the process theories have provided a solution.

Chapter 2

Russell on the Notion of Cause

"Then I must discover the truth about Causality - in the paper I read the other day, I only showed that all current views are wrong, and I am at a loss as to what is right."

(Bertrand Russell - from a letter to Lady Ottoline Morrell, 23 February 1913)

2.1 Russell's 1913 Paper

In the essay "On the Notion of Cause", his presidential address to the Aristotelian Society in 1912[1], Bertrand Russell attempts to show how the philosophical and the scientific concepts of causality differ from each other. Today, there are two important inter-related reasons for one who desires to investigate the concept of causality in physics to begin with Russell's 1913 paper. First, in this paper Russell masterfully cast the Humean viewpoint in a precise form that makes it possible for him to analyse more fully, and to subsequently raise the important issue of the *temporal contiguity between cause and effect* with this standard philosophical picture of causation. Second, troubled by this worry with the standard picture, Russell provided an argument against the notion of cause in physics. He concludes that the strict, certain and universal doctrine of causation - *the same cause always produces the same effect* -

[1]Delivered on 4 November 1912. The ensuing essay was published in the Proceedings of the Aristotelian Society, 13 (1912-13) and reprinted in Russell, B. (1917), p.180-208.

"which philosophers advocate is an ideal, possibly true, but not <u>*known*</u> *to be true in virtue of any available evidence. What is actually known , as a matter of empirical science, is that certain constant relations are observed to hold between the members of a group of events at certain times"*[2].

The difficulty with the temporal contiguity between cause and effect arises mainly because of the fact that events defined in physics are conceived of as occurrences at spacetime points - spatial points at specific instants. Two events being temporally contiguous means that one event must happen immediately after the other. But since instants of time are described by the mathematical continuum in which all the points are subjected to a dense ordering; it follows therefore that between any two events happened at two respective instants, there always exist other instants in-between, which could have accommodated the occurrences of other events. And the existence of these other events implies that, after all, the two events in question cannot be occurring immediately one after another. The spatiotemporal continuity between two events cannot be attained. Here lies Russell's challenge: given that there is a temporal gap between the cause and effect, how are we to fix it?

In fact, we note that there always exists a continuum of these other events in-between any two specified events because there is a continuum of instants between any two specified instants. Any one of these events can be considered as contributing to the production of the effect, and is thus deemed a cause of the effect; although not the sole cause. An appropriate view is to take all these "causes" into account and not isolating any one of them. This then forms the key to a solution of Russell's temporal gap problem.

We need to take a spatiotemporally continuous sequence of events as the "cause" that is to be met immediately by the effect. In physics, these continuous sequences are provided by the continuous trajectories of physical objects in spacetime. Indeed, in the simple case of two balls colliding, we speak about the motion of one ball bringing it close to and eventually into contact with another. These spatiotemporally continuous trajectories of physical objects take care of Russell's temporal gap by encompassing

[2]A quote from Russell, B. (1914), and reprinted (1993), p.230.

a continuous sequence of events and it is the way to causation in physics. Moreover, since the "causal connection" between one instance of cause and effect is now given by a physical connection linking the two, one finds a meaning for singular causation in physics. These continuous trajectories of physical objects form the basic units of analysis for the process view of causation advocated by Wesley Salmon and Phil Dowe, which are discussed in Chapters 3 and 4 that follow. The process view of causation therefore provides a solution to the temporal gap problem.

In the present chapter, I first present a study of Russell's argument against the problem of the temporal contiguity between cause and effect (Section 2.2) and then show how his analysis poses an seemingly insuperable challenge for the advocates of singularist theories of causation (Sections 2.3 and 2.4). I will also indicate the possible ways to resolve this difficulty (Sections 2.5-2.6).

2.2 A Definition of Causality

Russell begins his investigation in the 1913 paper by first establishing what was at his time of writing the philosophical "received notion" of causality. With this aim in mind, he turns to the definition of causality as given in James Baldwin's 1902 edition of the Dictionary of Philosophy and Psychology[3]. There, this definition of causality is given as: ***the necessary connection of events in the time series***[4]. Being "necessary" implies that the "connection of events in the time series" must hold under all circumstances. First to notice is that the definition reveals no details about the nature of the *"connection"* except that those events which stand in such a connection must obey certain *temporal* relations, simply because events are taken as temporally ordered. This is consistent with our usual empirical observation of events that we see as causes and effects since they all bear a "before-after" relation to each other. That is to say, *causes precede their effects*[5]. Any two events standing in a cause-effect

[3]Baldwin, J.M. (ed.) (1902).

[4]As an interesting historical note, I should remark that this entry on causation is a contribution by G.E. Moore who provides what is essentially a Humean account on causation.

[5]Of course, the "before-after" relation is not the only temporal relation that can exist between two events. There is also the possibility of simultaneity - with two events happening at the same time. But as far as our causal intuition permits, an event we call the "cause" is one that happens

	C	E
Yesterday	7.00am	10.00am
Today	5.00am	8.00am
Tomorrow	3.00pm	6.00pm
Day after Tomorrow	9.00pm	12.00am

Table 2.1:

relation must satisfy this temporal ordering. This is so in both common sense and in physics. Put differently, whatever the nature of this "connection" of the events may be, the events so connected would have to be found to be organized into a temporal schema of "before-after" as an observational consequence.

Now we must probe deeper and ask in what sense is such a connection "necessary"? We have already remarked that for something[6] to be necessary, it ought to hold under all circumstances. To be precise, some feature(s) of it has to hold in all cases. In respect to the relation of cause-and-effect, this would mean, at the minimum and without assuming any particulars about the nature of the "connection", that the "before-after" temporal ordering between the cause and its effect must hold under all circumstances. In the first instance, a specific temporal ordering holds under all circumstances if these "circumstances" refer to *all times.*

A simple example illustrates. Let C and E be two events that we refer to as the cause and effect respectively. On each of the four days in Table 2.1, C occurs earlier than E. This "earlier than" ordering is a prerequisite for C and E to qualify as cause and effect. In addition, the table brings to our attention another important point: not only does the temporal-order between C and E remain invariant through a translation in time, so does the temporal-interval between the two; there is a three-hour delay of the occurrence of E in relation to that of C. To help the reader to appreciate the point, the scenario depicted in Table 2.1 is compared with that in Table 2.2.

In the scenario given in Table 2.2, although the temporal ordering of "earlier

earlier than the one we call the "effect".

[6]Strictly, this "something" ought to be a proposition for only propositions are capable of being true or false.

	E_1	E_2
Yesterday	6.00am	9.00am
Today	3.00pm	4.00pm
Tomorrow	7.00pm	11.00pm
Day after Tomorrow	12.10pm	15.45pm

Table 2.2:

than" is also obeyed by the same two events E_1 and E_2 in each of the four instances, one would be reluctant to infer a cause-effect relation between them as the temporal intervals between the occurrences of E_1 and E_2 are not constant. Insofar as our causal intuition[7] is concerned, they fail to become causes and effects.

So for a "connection" between two events to be necessary, two primary conditions must be satisfied,

(1) the temporal-ordering remains invariant under a translation in time *and*,

(2) the interval between the times of occurrences of both events stays constant under a temporal translation.

The definition of causality given as an entry in Baldwin's Dictionary may now be stated with more accuracy as Russell suggests[8],

> ***Given any event e_1, there is an event e_2 and a time-interval τ such that, whenever e_1 occurs, e_2 follows after an interval τ.***

To cast the definition into the vocabulary of *cause-and-effect*, the principle of causality amounts to the statement that given the same *cause* (e_1), there is *always* the same *effect* (e_2) but only after the lapse of an (constant) interval τ.

[7]I should point out that this intuition holds provided all the conditions under which the events C and E occur remain the same. Take the simple example of an individual urinating after the consumption of some liquid. Of course, the time interval between the occurrences of the two events varies since in hot weather, perspiration would have helped to lengthen the interval. On the other hand, when perspiration is less likely in cold weather, the event of urination would be brought closer in time to that of the consumption of the liquid. In both cases, then, the conditions under which the events occur differ.

[8]Russell, B. (1917), p.183.

So far we have only indicated that the interval τ must be constant and nothing has been said about its duration. It is indeed legitimate to inquire about how long or short this interval should be. For the longer the interval is, the greater the possibility there is for other events to occur within the interval that may frustrate the production of the effect. And so ideally, it is desirable to keep this interval as short as possible, which in the limit would mean that there is no lag of time between the cause and the effect. Hence, according to the Baldwin dictionary's definition of cause and effect,

> ***Cause and effect are correlative terms denoting any two distinguishable things, phases, or aspects of reality, which are so related to each other that whenever the first ceases to exist the second comes into existence immediately after, and whenever the second comes into existence, the first has ceased to exist immediately before.***

Thus, the cessation of the cause is to be followed *immediately* by the corresponding creation of the effect. However, it is relatively simple to reject this statement. Take for example, one may well attribute the cause of the triggering of the smoke-alarm to a fire; but no doubt it most probably would not be the case that once the alarm is set off, the fire vanishes accordingly. Everyday life is flooded with numerous examples of such kind. This situation, however, does in no way invalidate our definition of the cause as an event that happens before the effect. For even though it does not cease after the effect comes into existence, the cause has existed *before* the commencement of the effect. Counterexamples are easy to find because events which are regarded as cause and effect occupy extended regions of spacetime and may overlap. This concerns the issue of temporal contiguity between cause and effect and we shall see how Russell acknowledges the related difficulties.

2.3 The Issue of Temporal Contiguity between Cause-and-Effect

The thesis of temporal contiguity between two events can be stated as: ***whenever the first event ceases to exist the second event comes into existence immediately after.***

Let us first be clear about one thing. In this discussion, the causal relata[9] are taken to be events. The reason is because these are the entities that stand in causal relations to each other in physical formulations. And the focus on events as causal relata then allows us to pave the way towards an investigation into causation in physics.

Events happen in spacetime and are therefore characterized by their *spatial locations* and *times of happening*. Events in real life have both spatial and temporal extent; they occur in a region of space and over a finite interval of time. However, unlike real events, events in physics are idealized mathematical points. They are supposed to occur at single spacetime points[10].

Conceiving physical events as *point events* occurring at abstract spacetime points constitutes indeed the root of the problem of temporal contiguity between cause-and effect. We shall now explain why.

There exist essentially four possible kinds of spatial and temporal relation that any two events may bear to each other,

(1) Events located at the same spatial location but occur at different times *or*,

(2) Events located at different spatial locations and happen at different times *or*,

(3) Events located at different spatial locations and happen at the same time *or*,

(4) Events located at the same location and happening at the same time.

[9]Discussions on causation may also center around the possible relata of a causal relation of which events are but amongst one of the candidates.

[10]Here, we refrain from the debate on the reality of spacetime points but treat them merely as a mathematical abstraction used in physics.

These four categories provide the basis for discussing both spatial and temporal contiguities. We first show that spatial contiguity[11] is less of a problem than temporal contiguity and this will help us to appreciate why indeed Russell takes issue with the latter and not the former.

Being spatially continuous means that two events must be so closely adjacent to each other in space that there is overlap between them to some extent. In the limit, this is trivially attained by the two events happening at the *same* spatial location to achieve maximum overlap. Events that occur at the same location may happen at different times (as in (1)) *or* at the same time (as in (4)). It is easy to visualize events of the first kind but events happening at the same location at the same time requires some clarification.

Suppose a piece of coal, initially kept at room temperature, is to be heated in a furnace. We may speak of the following two events as occurring simultaneously at the very same location where the piece of coal is situated: *the supplying of heat to the piece of coal by the fire* as *the piece of coal becoming red-hot.*

Two interesting observations arise. First, it is possible for distinct events to occupy the same spatial location at the same time. The supply of heat by the furnace fire and the piece of coal reaching its red-hot temperatures are distinct, though related events[12]. Second, both events have *temporal* extent as implied by the usage of words like "supplying" and "becoming". No doubt the supply of heat to the piece of coal by the furnace fire is an event which precedes that of the piece of coal becoming red-hot. But these events are not point events in the sense of the supply of heat at an *instant* being followed by *the piece of coal becoming red-hot*, happening also at an *instant.* The supply of heat is a gradual undertaking and the temperature of the piece of coal is raised from its initial value to that of red-hot through a series of intermediate temperatures. It is emphatically not the case that the temperature jumps from one instant at room value to its red-hot value at the next instant.

[11]As we shall see later in the chapter, a solution of the temporal contiguity problem in terms of spatiotemporally continuous (physical) entities, would also make good causal connections for events described in (2) and (3) above (given that events described in (3) are not spacelike separated).

[12]It is indeed part of the aim of study of causation to elucidate this relation. This stage of the discussion, however, addresses only the spatiotemporal relation.

The underlined phrases reveal a subtle difficulty. The treatment of events as point-events is grossly unrealistic for each of the above takes a finite amount of time to occur and does not just occur at an instant. This is most easily visualized by the following diagram (Figure 2-1),

Figure 2.1: Two point-events occurring at two instants of time.

Such an idealization misses out on what happens in the intermediate times[13] t between t_1 and t_2 with $t_1 < t < t_2$. A more realistic picture is afforded by the consideration of events having temporal extent.

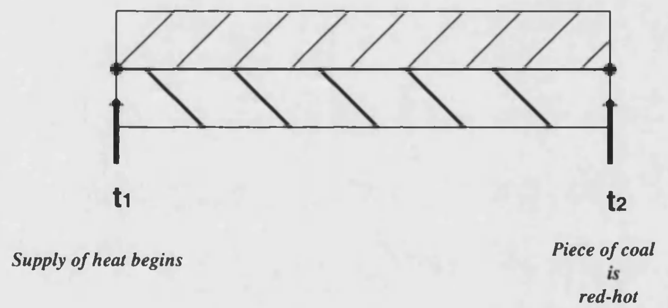

Figure 2.2: Two temporally-extended events overlapping.

Here the "continuous" supply of heat to the piece of coal raises its temperature from its room value to that of red-hot in a continuous and gradual manner. The physical process that take place in the intermediate times are thus accounted for. In this particular example, the two events overlap in the interval $t_1 < t < t_2$ (Figure 2-2).

So it would seem that the reply to the question of whether two events are temporally contiguous to each other is parasitic upon the definition of an event with respect to its temporal extent. For, if we bring the times t_1 and t_2 very close to each other

[13]Here, of course, time is taken to be endowed with the structure of a mathematical continuum.

so that in the limit, t_1 and t_2 are so adjacently close to each other that the temporal duration of both events are diminished to such a degree that their endpoints stay very close to each other, see Figure 2-3.

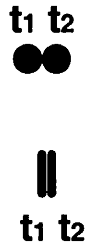

Figure 2.3: Temporal Contiguity?

This situation can, in principle, be achieved by the supply of a large quantity of heat as a consequence of which the temperature of the piece of coal is raised to its red-hot level in the "next instant" so-to-speak. But can t_2 be brought so close to t_1 in order for it to be considered as a next instant? If the answer is yes, then it would seem that we may indeed focus on just the two endpoints as two point events and ask in what sense are they contiguous to each other.

Russell argues that there cannot be a notion of contiguity between any two events occurring at two respective instants. This is his well-known objection against the temporal contiguity between events as causes-and-effects. His argument is based on the contention that physical time is described by a dense series in a mathematical sense; that is, the mathematical continuum.

A mathematical series is regarded as *dense* when no terms in the series are consecutive, but between any two there are always others. When this idea is applied to the time-series, it implies that no two consecutive instants of time are contiguous; the assertion of a next instant simply does not make sense because in-between any two instants, there exists, in a mathematical sense, a continuum of others. Now suppose we halve an interval, and then halve the half. We can indeed continue the process for as long as we please, and the longer we continue it, the smaller the resulting interval becomes. One hopes that such a procedure would eventually bring the two instants so close together that any lapse of time between them may be considered as negligible;

that is, the interval only spans an infinitesimal duration (as in Figure 2-3).

At first sight, the infinite divisibility seems to imply that there are infinitesimal intervals, i.e. intervals so small that any finite fraction of, say, a second would be greater. This, however, is an error. The continued bisection of the interval, though it gives us continually smaller intervals, gives us always finite intervals. If our original was a second, we reach successively half a second, a quarter of a second, an eighth, a sixteenth, and so on; but every one of this infinite series of diminishing intervals is *finite*[14]. We then find ourselves justified in treating this series as an arithmetic progression the sum of which is finite and this corresponds, of course, to our original finite interval between the two instants. In other words, there exists always a finite interval between the two respective instants of time[15].

Now if the definition for the temporal contiguity between two events - *whenever the first ceases to exist the second comes into existence immediately after* - is, in fact correct, and the time series is to be considered as the mathematical continuum, the finite interval that exists between the *instant of cessation* of the first event and the *instant of commencement* of the second must have to be somehow taken into account. This can readily achieved by the interval being absorbed into the temporal dimension of the events.

And so Russell suggests that the finite interval between the two instants is to be accounted for by supposing either the cause *or* the effect *or* both to endure for a finite time. However, he at once notes that this supposition runs into serious difficulties.

The only options open to us with respect to the temporal dimension of an event are that the event,

(1) behaves as a point which is to be understood as occupying a region of infinitesimal spacetime *or*,

(2) exists for a finite period of time.

[14]Any series which exhibits the feature of denseness, there always exists a *finite interval* between any two points of this series.

[15]Because of this finite interval, temporal contiguity cannot be attained, despite that there is spatial contiguity.

With respect to (2), we may either think of an event as existing for a finite period of time during which it changes or during which it remains unchanged (a static event). Events that change may then be analyzed further as either a sequence of point events *or* a sequence of events that stay unchanged within their finite temporal durations *or* a sequence of combinations of both.

We now consider each type of events in turn. First, we have the events corresponding to both the cause and the effect as point events. In this scenario, each of the two point events is to occupy an instant in the time-series and because of the dense nature of the time-series, there exist always a finite, albeit extremely short perhaps, interval to separate the two. Hence, the two point events in question cannot be temporally contiguous to each other. Second, we have the situation where one *or* both events are considered to occupy a finite duration involving "change within itself". Imagine the cause changed. Then we would need to postulate a cause-effect relation between the part of the cause before the change and the part after. Each of these "parts"[16] is taken to be either as a point *or* as one with a finite duration and *static* (in that they are not decomposable into further distinct events that are causally related). There are four exhaustive possibilities[17] (Figure 2-4),

(i) the first is of a finite duration with the second as a point-event *or*,

(ii) the first as a point-event with the second event having a finite duration *or*,

(iii) both the first and the second events span finite durations *or*,

(iv) both the first and the second events span finite durations and they overlap to some extent.

Given the situation in (i), argues Russell, it would seem that only the later (point) events within the causal process E_1 can be relevant to bringing about the effect E_2 since the earlier ones are not contiguous to the effect and therefore cannot directly influence it.

[16]Following Russell, we may loosely speak of an event with a finite duration as a "process" to signify the presence of temporal parts.

[17]For the case of an event of finite duration the "parts" of which are considered as points.

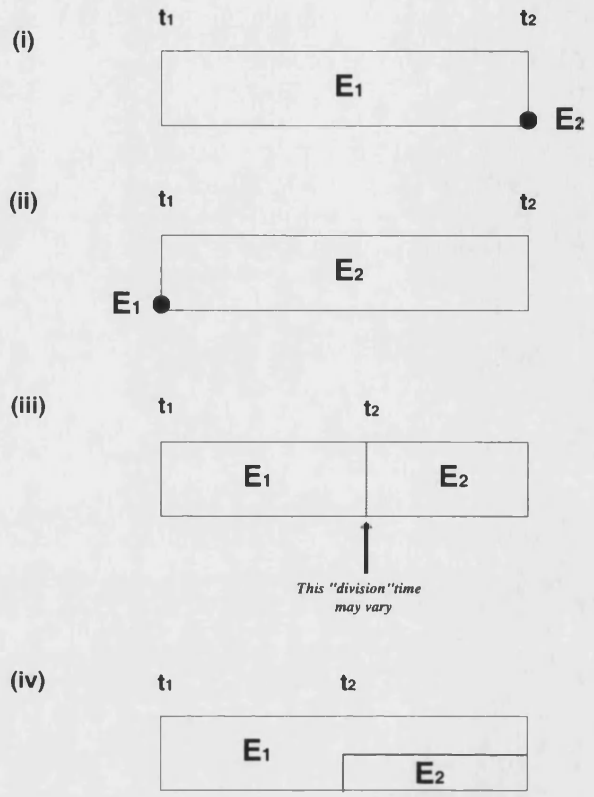

Figure 2.4: The temporal gap issue between cause and effect.

In order words, all the subsequent point events that occur at all times t such that $t_1 < t < t_2$ are temporally closer to E_2 than the "initial" cause as a "point event" commencing at time t_1. But how *close* can these later parts of E_1 be to E_2? If these later parts are themselves treated as points and since E_2 is supposed to occur at a point instant, then there always exists a finite interval to separate them and destroys temporal contiguity.

Russell also explores the possibility of diminishing the duration of the process without limit; in the hope to bringing the beginnings of both E_1 and E_2 closer together. It is, nonetheless, immediate obvious that, even in the limiting case, we shall again arrive at the familiar situation of finding always a finite interval between two instants, due to, of course, the denseness of the time-series. Hence there will still remain earlier events which do not directly influence the effect! The "true cause" event (the beginning of E_1), so to speak, will never have been reached. By symmetry, similar arguments can be made against the respective cases of (ii) and (iii)[18] above. The same argument can again be iterated for the scenario in (iv) where there is an overlap of the two events E_1 and E_2, since we are interested in the commencement points of both events and in this case, there is still a finite interval between t_1 and t_2, in spite of the overlap after t_2.

The foregoing arguments all apply to events being decomposable into parts. Now what if, the extended event, or process E_1, is static and not decomposable into parts but rather, it *persists through time without change*? Under such a circumstance, E_1 can be considered as "one and the same" cause and we do seem to have a cause (E_1) being temporally contiguous to the effect (E_2). Russell, however, strongly protests against the existence of such kinds of events[19], "*...it seems strange - too strange to be accepted, in spite of bare logical possibility - that the cause, after existing placidly for some time, should suddenly explode into the effect, when it might just as well have done so at any earlier time, or have gone on unchanged without producing its effect.*"

To understand Russell's objection, take an event that has a finite duration during which it does not change (Figure 2-5).

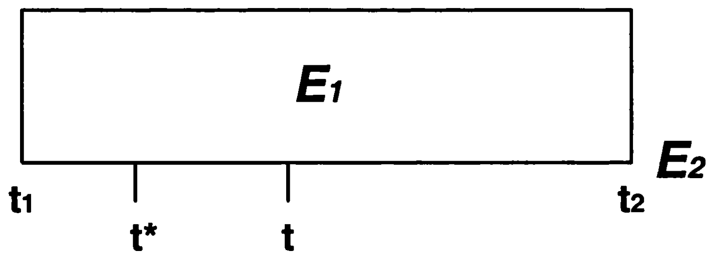

Figure 2.5: A static event.

Here, the static cause E_1 comes into existence at time t_1 and remains unchanged until time t_2. $(t_2 - t_1)$ denotes the time lapse between the occurrence of E_1 and the occurrence of the effect E_2. Russell's point is: isn't it strange that E_1 occurs at t_1 and sits there unchanged, then all of a sudden, E_2 bursts into existence at t_2? Why has this suddenly taken place at t_2 while logically, it might have done so at t^* or some

[18]As we are interested in the commencements of both the events E_1 and E_2 at times t_1 and t_2 respectively.

[19]ibid., p.184.

other time t during the interval $(t_2 - t_1)$?

Static events persisting over an interval without change is thus not acceptable and the cause cannot be made contiguous to the effect by the appeal to static events.

In summary, for two events to be temporally contiguous to each other, the finite interval between the times of commencement of both events must be accounted for. This may be achieved by demanding that one *or* both of the events to have temporal extent and of finite duration, in order for the temporal gap to be filled. Events of finite duration can either considered to be capable of changing and decomposable into other events *or* to be non-decomposable and with itself persisting without changing into other events. For events that change, the temporal gap remains because the "initial" cause cannot be reached. For non-changing events, the hope to close off the temporal gap is shattered by the logical unacceptability of such a kind of event. As a consequence, temporal contiguity between the cause and effect cannot obtain in both cases.

It is useful at this point to sum up this important argument by a schematic representation.

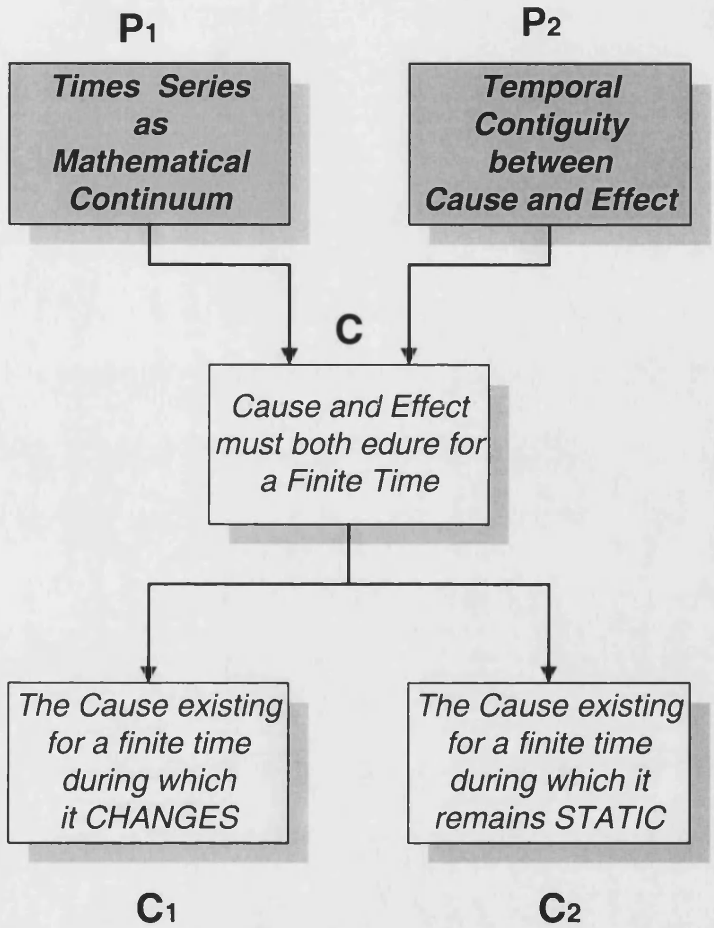

Figure 2.6: The logical structure of Russell's argument against the thesis of temporal contiguity between causes and effects.

Written in symbolic form, the logical structure of Russell's argument against the temporal contiguity between cause-and-effect goes as follows,

$$P_1 \wedge P_2 \rightarrow C$$

$$C \rightarrow C_1 \vee C_2$$

$$C_1 \rightarrow \neg P_2$$

$$\neg C_2$$

P_1

Therefore, $\neg P_2$

It is now obvious that the validity of Russell's argument depends on the truth or falsehood of the premise P_1, namely that whether it is in fact correct in representing the time-series by the mathematical continuum. This is in its own right a very profound philosophical question that deserves a detailed and careful investigation, the scope of which is well beyond the present study. Suffice it to say that Russell's argument is valid insofar as the temporal series is represented by the mathematical continuum.

2.4 Is There a Notion of Cause in Physics?

The second part of Russell's 1913 paper is devoted to an assessment of the applicability of the principle of causality in a theoretical science like physics, with specific reference, of course, to the temporal gap problem[20].

The principle of causality, as has been established in the preceding section, stipulates that the *same cause is always followed by the same effect* (but only) *after a finite interval of time.* By considering physical time as the mathematical continuum makes it impossible to close off the temporal gap between the cause and effect. But the temporal gap presents a problem because it enhances the opportunity for other events to "creep in" during this interval and frustrate the production of the effect.

In physical situations, "causes" are to be identified with the so-called "initial (or antecedent) conditions" that represent occurrences ahead in time of other events. Having the *same* cause *always* followed by the *same* effect implies that the causes, the initial conditions, are repeatable. Russell maintains that on a *practical* front, if an event is ever to *recur*, it must not be defined too narrowly and not take into account too precise details. Once the initial conditions become too complex, the likelihood of recurrence of each of these fine details is greatly diminished. Thus Russell tells us[21], *"An "event"...is a universal defined sufficiently widely to admit of many particular occurrences in time being instances of it."*

One obvious way of narrowing the scope of an event is to restrict the temporal definition of its duration. In the limit, the event is defined to occur at an instant and as we have seen, such point events suffer miserably from the temporal gap problem. The presence of the temporal gap provides ample opportunities for other events to happen during that interval that may interfere with the occurrence of the effect.

The only reliable means to dispose of the temporal gap is to take into account also those events that occur in the gap which would affect the occurrence of the effect in some way. So the cause event and all the ensuing events[22] happened during the

[20]It highlights how the temporal gap destroys our much cherished causal intuition of "the same cause is always followed by the same effect".

[21]ibid., p.187.

[22]All these, including the cause event, are considered as points.

times running up to the time of commencement of the effect are considered as *one* initial condition (Figure 2-7).

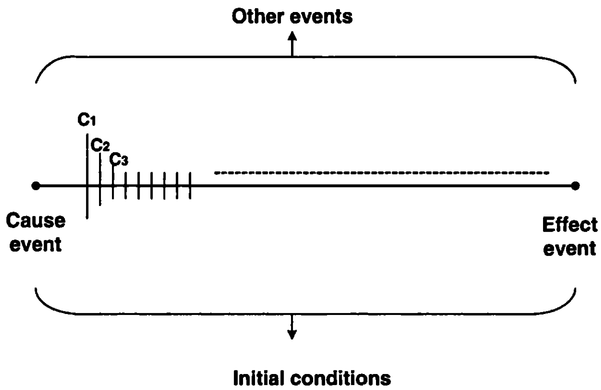

Figure 2.7:

The intricate details of the initial condition renders it highly unlikely, if not impossible, to be repeatable and thus making the phrase of the "same cause" vacuous. If we cannot have the same cause, then there is simply no meaning for the expression "the same cause is always followed by the same effect."

One recognizes, of course, that "the same cause is always followed by the same effect" is one of the famous doctrines of Humean causation - that of constant conjunction between two events as cause-and-effect. It is undeniable that the constant co-joining and the temporal-ordering between two events together establish for us the intuition that the cause is regularly followed by the effect after a definite time interval.

Russell has shown that this Humean doctrine is not applicable in physics. First, this is because the temporal gap presents the possibility for other events to intervene and thus one can no longer be sure of the occurrence of the effect. Second, even if the temporal gap is "closed" by encompassing all the intermediate events with the original cause, the complicated initial condition thus results is unlikely to find itself being repeatable. Either way, the doctrine of "same cause, same effect" is made inapplicable in physics. There is therefore no notion of causality in physics in the sense of "same cause, same effect".

Russell brings compromise to this seemingly pessimistic state of affairs by proposing an operational meaning of causality[23], "*...any case of sufficiently frequent sequence will be causal in our present case.*" One should *prima facie* give up the hope of finding "causal laws", in the strict sense of "same cause, same effect" in physics. These frequent sequences referred to by Russell are often generalized into the so-called laws of physics. But how about the "laws" of physics that we have come to be so familiar with, like Newton's force laws or the law of universal gravitation? With these laws, Russell argues[24], "*there is nothing that can be called a cause, and nothing that can be called an effect; there is merely a formula. Certain differential equations can be found, which hold at every instant for every particle of the system, ...render the configuration at any other earlier or later instant theoretically calculable. That is to say, the configuration at any instant is a function of that instant and the configurations at two given instants...But there is nothing that could be properly called "cause" and nothing that could be called "effect" in such a system.*" And again[25], "*...it is not in any sameness of causes and effects that the constancy of scientific laws consists, but in the sameness of relations... "sameness of differential equations"...*".

How does this "functional"[26] view differ from the philosophical principle of causality? First, this view of the principle of causality does not have an *a priori* feature of necessity but mere empirical generalizations. Second, the functional view makes no difference between the past and the present. Given data from the past, the future state of a physical system can be readily computed and vice versa. Third, although it makes no demand for an element of necessity, the functional view does, however, require some sort of "uniformity of nature". The uniformity of nature ensures that it is a feature of our universe for a law, which has been found to hold throughout the observable past, to be expected to hold also in the future.

[23]ibid., p.193.

[24]ibid. p.194.

[25]ibid., p.195.

[26]These functional laws are the equations of motion of physical systems and we shall see in the later chapters how causal processes - the trajectories of physical objects in spacetime - of the Salmon and Dowe varieties adheres to this view.

2.5 Possible Solutions to the Temporal Gap Problem

In his famous treatise[27] on causation, David Hume came to identify three main impressions from which our idea of a necessary (causal) connection between two events arise:

(1) *the priority in time of the first event*

(2) *the constant conjunction of both events (same cause, same effect)*

(3) *the spatiotemporal continuity of the two events*

The three impressions are related to each other in the following manner. Our intuition tells us that for any pair of events to be deemed the cause-and-effect, one of the two events has to commence before the other and that whenever one occurs, so does the other. These observations are subsumed under (1) and (2) above. However, as it has been illustrated by the scenarios depicted in Tables 1-1 and 1-2, the "regular" observation of the cause-and-effect encompasses the fact that there exists a temporal order with a more or less constant time interval between them.

In particular, such an interval ought not be too long or else it would provide ample opportunities for other factors to act and affect this particular "causal" relation. The most effective way to ensure that the first event, and no other, is the one which *causes* the second, is to bring the cause and effect as close as possible both spatially and temporally. The cause and effect are deemed to be closest spatiotemporally when the first event draws to a close as and when the second begins. This is how (3) that concerns spatiotemporal continuity between the cause-and-effect gets connected with that of (1) and (2). And it is Hume's most remarkable achievement to have shown that from these three impressions alone, one never gets to discover the element of necessity in what is supposed a connection between two events regarded as cause and effect. Russell had, of course, written in this Humean spirit and advocated the regularity account as he opted for the functional view of causality in physics.

[27]Hume, D. (1888).

With reference to this standard, Russell has shown us the incoherence between (2) and (3) (Section 2.4 above). In order to achieve condition (2), events ought to be defined with some restriction on its temporal extent in order to maximize their chances of re-occurrence. In the limit, one considers the cause and effect as point-events. With two point-events standing at two instants in the temporal continuum, there always exists a finite interval - a temporal gap - between the two events. And so the "cause" (the event that is prior in time) cannot be temporally contiguous to the "effect" (the latter event). In this case, "same cause, same effect" (2) can only be attained at the expense of temporal continuity between the two events as in (3).

This temporal gap may be reduced, so that the cause and effect are brought closer together in time, if we consider, say, the cause as an extended temporal entity[28]. A cause having temporal parts may consist of a great deal of intricate details, which in turn diminishes its chance of being repeatable. Here, temporal continuity between the cause and effect may be achieved but only at the expense of "same cause, same effect".

This much being said, the thesis of "same cause, same effect" may, however, be abandoned in the consideration of singular causation where the condition of repeatability is relaxed. There are temporal sequences of events which we consider as causal but are not repeatable (in principle) by their very nature. An extreme example is the evolution of the universe[29]. We would like to be able to speak about the notion of causation in the context of such singular sequences. Generally speaking, in situations where there exist spatiotemporally physical connections, we speak of the causes and effects as linked by these physical connections in question. After a light switch is thrown, the lamp is lit. We say that the lamp being lit is *caused* by the switch being turned on. The switch being turned on is the *cause* with the lamp being lit as the consequent *effect*. This is envisaged to be so because there is a physical connection - an electric wire that acts as the medium for a current to travel from one respective spacetime locale to another - between the light switch and the lamp.

[28]The same argument applies if the effect is considered as an extended temporal entity instead.
[29]Here, we refrain from considering physical models that describe cyclic universes.

If it is indeed legitimate to think of causation on a singular level, the problem of the temporal gap may then be duly overcome by focusing on the issue of temporal dimensions of causes and effects as has been indicated earlier in the discussion.

Here, we shall briefly discuss the attempt of the resolution of the temporal gap problem by C.J. Ducasse[30]. His proposal is of particular relevance to our discussion as it centers around the temporal dimensions of events.

Ducasse maintains that it is essential to distinguish clearly "a time" as an instant represented by a *cut* in the time series, from "a time segment" that is defined by *two cuts* (or instants). An event[31], he argues, cannot be said to occur at a time but only during a time segment. So, an event, for Ducasse, is properly described by a section (which itself is continuous) of the temporal continuum. Thus viewed, it seems that spatiotemporal continuity of the cause and effect may now be achieved: the same cut in the time series marks both the end of the cause and the beginning of the effect. The cut itself have no spacetime dimension at all. Ducasse thus concludes[32],

> ***With cause and effect and their spacetime relation so conceived, there is no possibility that, as Russell contended, some other event should creep in between the cause and the effect and thwart the production of the effect.***

It is one matter to assert that the same instant - the one cut in the time series - should serve as both the cessation point of the cause and the commencement point of the effect, but it is quite a different matter to decide where in the temporal continuum ought this cut be situated. Ducasse seems to think that this is a pseudo-problem as he writes[33],

> ***Nor are we compelled, as he (Russell) also contended, to trim down indefinitely the beginning part of the cause (and, mutatis mutandis, the end part of the effect) on the ground that the***

[30]Ducasse, C.J. (1926). Reprinted in Sosa, E. and Tooley, M. (eds.) (1993), p.125-136.
[31]Here, he means real happenings that take place over finite durations of time.
[32]ibid., p.129.
[33]ibid.

early part of the cause is not necessary to the effect so long as the end part of the cause occurs. For, once more, the cause means something which was sufficiently necessary to the effect. Thus the spacetime limit of the cause process at the outer end is as elastic as we please, and varies with the spacetime scope of the particular description of the cause that we give in each concrete case. And the same is true of the outer end of the effect process.

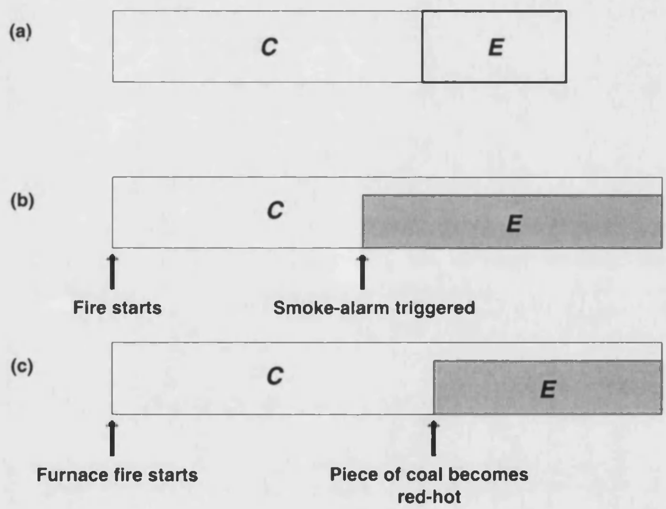

Figure 2.8: Where should the temporal cut be situated?

Ducasse's argument, of course, presupposes that the cause and effect are not concurrent so that the cause must cease as soon as the effect begins. The previous examples of the fire triggering the smoke-alarm and the furnace heating up a piece of coal are both evidence to the contrary. In both of these situations, the cause and effect can undoubtedly co-exist in time after the commencement of the effect (Figure 2-8).

It is therefore not at all a vacuous exercise to dwell on the debate as to at which location the cut should be introduced. Some method must be sought to help fixing the position of the cut.

Careful reflection of both examples lends us an important idea. In the first scenario, the fire starts[34] at a small region of spacetime and this event subsequently generates a sufficient amount of smoke which triggers the smoke-alarm when it reaches it. The two events - "the starting of the fire" and "the trigger of the smoke-alarm" - being relatively localized in their respective regions of occurrences, are *physically connected* by the smoke that traverses the intermediate region. The "spreading smoke" forms the crucial factor to link the two cause and effect events in question, while at the same time it serves the vital purposes of closing the temporal gap.

Entities which have relatively large spatiotemporal extent and stable in structure (so that the different parts of the entity differ only in their relative positions in spacetime and not in any other characteristics) are ones that must be sought after to fill the temporal gap.

In the example, at each of the two localized points where the two events are supposed to occur, there are at least two entities involved. At the vicinity of the cause, there are the fire and the smoke and at that of the effect, there are the smoke and the smoke-alarm. In each instance, the two entities concerned somehow "physically interact" to result in the cause and the effect respectively. The place where the cut that signifies the commencement of the effect (without necessarily the cessation of the cause) may now less arbitrarily and more meaningfully defined by reference to these physical interactions[35].

Similar reasoning applied to the second scenario leads to a similar picture. Here, the event that corresponds to the heating of the piece of coal by the furnace fire and the subsequent one where the piece of coal becomes red-hot are not separated in space but only in time. Furthermore, at the very same location of both events, there are the same two entities involved - the *furnace fire* and the *piece of coal.* Under this circumstance, it is the piece of coal itself - its very spatiotemporal continuous existence[36] - that forms the physical connection between the event of its being heated by the furnace fire at room temperature and the event corresponding to it reaching

[34]Assuming that it is not an explosion that takes up a relatively substantial chunk of spacetime.
[35]See later discussions in Chapter 3.
[36]Here we make the basic assumption that the piece of coal has an identity over time.

the red-hot temperature. It is the spatiotemporally continuous existence of the piece of coal that functions to close the temporal gap.

It is also noted that despite a continuous string of interactions, in terms of a continuous exchange of energy between the fire and the piece of coal, it is still reasonable to pinpoint two specific events and ask about how they are to be connected temporally.

The foregoing examples have suggested the consideration of physical entities of spatiotemporally continuous extent as a realistic means to resolve the temporal gap issue. This also seems to be Russell's idea when he had the occasion to reflect on the notion of causation more than three decades later after the publication of his seminal 1913 essay.

2.6 Russell's Causal Lines

In the preceding sections of the chapter, we have detailed Russell's view on causation at the turn of the twentieth century. Three and half decades later in 1948, he renewed his opinion on the subject in a more optimistic maneuver with the introduction of the notion of "causal lines", which can be seen as the precursor to the more recent process theories of causation.

With the good intention of filling the hazardous temporal gap, we have appealed to physical entities that span an extended spatiotemporal existence, during which they exhibit stability in their physical characteristics and makeup. This is indeed what Russell had in mind when he introduced the notion of *causal lines*[37],

> ***The concept of more or less permanent physical objects, in its common-sense form, involves "substances" and when "substance" is rejected we have to find some other way of defining the identity of a physical object at different times. I think this must be done by means of the concept "causal line". I call a series of events a "causal line" if given some of them, we can infer something about the others without having to know anything about the environment... When two events belong to one causal line, the earlier may be said to "cause" the latter. In this way laws of the form "A causes B" may preserve a certain validity. They are important in connection both with perception and with the persistence of material objects.***

And again[38],

> ***A "causal line", as I wish to define the term, is <u>a temporal series of events</u> so related that, given some of them, something can***

[37]Russell, B. (1948), p.333
[38]ibid., p.477.

> ***be inferred about the others whatever may be happening elsewhere. A causal line may always be regarded as the persistence of something - a person, a table, a photon, or what not. Throughout a given causal line, there may be constancy of structure, or gradual change in either, but not sudden change of any considerable magnitude.***

Russell intends for the concept of causal lines to define the identity of the time of physical objects. Stated simply and naively, the identity over time of an entity concerns the question of what makes us think the object we perceive now is the "same" one as we perceive at some previous time? The demand for this "sameness" has to do with what Russell has referred to in the foregoing passage: *the persistence over time of something*[39].

Think of the simple scenario where we have a tennis ball sitting on the table. The ball is recorded as *one and the same object* between a "continuous" time interval, say, t_1 and t_N, provided that at each of the instants within the interval, there is an entity having the same physical qualities and characteristics as the one that existed at t_1. Logically speaking, the entities at each of these instants need not be the *same* object but different objects with identical physical qualities and characteristics will do. One can always imagine a universe where objects with identical properties popping-in-and-out of existence at each and every instant.

In that world, one would not be able to tell the difference as to whether the tennis ball sitting on the table is indeed the one and the same or if it is consisted of a sequence of events, with each corresponding to the individual existence of a different object.

However, during any point, say, t_M, within the interval, if the tennis ball or a different one that shares the identical properties with the original fails to exist, then this discontinuity in existence would be a sufficiently strong basis for our reasoning to reject the hypothesis that the tennis ball in existence at some times after t_M is

[39] At a minimum level, one would argue that persistence over time entails the continual existence of the entity in space and time.

necessarily the same one which existed before.

The concept of causal lines is brought in to provide what one may call an identity criterion, which stipulates the conditions for an entity to be considered as one and the same that persists over time. The causal line does not make a distinction between the cases of having in reality the same object throughout the interval and those having different objects that share identical properties existing in turn during the interval. In both cases, one considers the object as one and the same object and this is *the one* object that provides the causal connection. The front most part of the smoke cloud that reaches and triggers the smoke-alarm is the part that has emerged first from the burning fire. Likewise, it is one and the same piece of coal that is subjected to the continuous combustion of the furnace fire.

These causal lines, made up of a sequence of events, each corresponding to the existence of a physical object at each point in spacetime, allow us to speak of the same object persisting in time. Its ability to persist in time sanctions one to speak about the object being the link - the *causal connection* - between the two events at different points in time in which it participates.

The major demand for events making up the causal line is that they should display a certain degree of constancy and sameness in their physical qualities and attributes. In the example of the tennis ball, all the events exhibit constancy and sameness in qualities in that each corresponds to the existence of *a* tennis ball and *any one of these* balls at each time shares identical (or extremely similar) attributes with the others.

Any sudden change disrupts the smooth flow of a causal line. Again, this is easily illustrated with the example of the tennis ball when its career in spacetime is punctuated by the sudden disappearance of the ball at some point in time and re-appearing at another point later. Since a causal line does not obtain, hence the ball cannot be said to enjoy an identity over time. We are no longer able to speak about the ball that has re-appeared as the one that vanishes some time before, as further evidence must be gained to support the contrary.

With the introduction of the causal line concept, Russell has made a deciding step towards resolving the temporal gap problem. However, in order to develop the idea into a proper causal concept, further work needs to be accomplished and this is achieved through the work on process causality by Salmon and Dowe (Chapters 3 and 4).

2.7 Reflections and Afterthoughts

In his 1913 essay, Russell has provided a sophisticated argument, essentially Humean in spirit, against the orthodox view that the philosophical (Humean) notion of causation is the one that is employed as the concept of causality in physics. Based upon the two major assumptions: that events are conceived as occurrences at individual spacetime points *and* that time is a dense mathematical continuum, Russell's analysis reveals to us the inevitable existence of a gap between any two point-events situated in the temporal continuum. This temporal gap has far-reaching implications, both on the levels of the regularity and the singular account of causation.

With respect to the regularity account, the observation of the constant conjunction between two events, from which the idea of "necessary connection" arises, cannot be maintained in the presence of the temporal gap. The temporal gap makes it possible for other events to intrude into the temporal gap and interfere with the production of the effect. Such a possibility defeats the notion of a necessary connection for the fact that we cannot be sure whether the effect (the latter event) is to follow in the next observation.

Since events are indeed treated as point-like entities in physics, it follows that the temporal gap renders the idea of causation, as a necessary causation arising from a pair of constantly conjoining events, sterile in physics.

The issue of the temporal gap has also an important impact on singular causation. Two events, with one preceding the other in time, can bear some sort of relation over and above this mere temporal relation that results in this time-ordering. Or, the "apparent" time-order of their occurrences has arisen as a sheer accidental fact. They may occur in a random manner such that they *just happen* to be standing in this particular temporal order on this particular occasion. What does it take for two events, one preceding the other in time, to be regarded as cause and effect? The regularist and the singularist have opposite recipes to decide between the two possibilities of a mere temporal ordering or a specific relation from which this ordering arise as a consequence.

The regularist maintains that it is not sufficient to draw any conclusion on the observation of one instance of these happenings and as such, it is thus meaningless to speak about causation on a singular level.

Contrary to the regularist, the singular causal theorist takes the view that there is a meaning to call this pair of successive events cause and effect, provided one can specify a relation in addition to the temporal ordering. The pair of events are considered as cause and effect by virtue of this special relation. The regularity we see of cause and effect is then supported by the fact that every pair of like events possesses this relation.

Because of the temporal gap problem, the singular theorist is confronted with the challenge to use their stipulated relation to close up the temporal gap. The issue of temporal gap emerges essentially as a problem for two events that stands in a temporal relation to each other. And so it is relevant to the consideration of singular causation, which focuses on the causal relation as one that exists between an instance of the occurrence of two events.

Even if the causal relation is a mere temporal relation between two events, that is, if E_1 is an earlier event than E_2, then E_1 is deemed the cause of E_2 (and E_2 is deemed the effect of E_1), the presence of the temporal gap makes it possible for other events to occur in the interval and hence there could be events taking place later than E_1 but earlier than E_2. And these other events, rather than E_2, might be deemed the effect of E_1. Similarly, the same events might be deemed the proper cause of E_2 rather than E_1. Under such circumstances, the "cause-effect" relation of two events as a temporal relation is thereby utterly destroyed.

In order to maintain the "cause-effect" relation, one must devise a means to close off the temporal gap. The most natural and obvious way is to include and take into account all the intervening events during the interval and consider an "event" as an entity having a temporal extent instead of as one that occurs at a single spacetime point. Russell has proposed such a structure as a causal line. However, such "extended" events themselves are in actual fact composed of a continuum of point events. To make our distinction explicit, we shall refer to these temporally-extended entities as "*processes*" to signify a continuous series of events in spacetime.

Serving well the function of filling the temporal gap, such a continuous process forms also the physical connection between the events E_1 and E_2 and hence making the notion of causality intelligible on a singular level, for now a physical relation is specified to be the causal connection.

It is the major task of this thesis to elucidate the nature of these spatiotemporally continuous processes, especially in relation to the extent that they may be considered as causal in both the contexts of classical (Chapters 3-4) and quantum physics (Chapters 5-7).

Before we embark on this philosophical project, there is one last point that deserves our attention. This has to do with the assumption that physical time forms a mathematical continuum. Indeed, physical time being continuous is the deciding premise for Russell's argument on the temporal contiguity between cause and effect to remain valid.

The temporal gap arises because of our desire to forge the integration of the concepts of the continuous and the discrete, and our wish to "divide" the temporal continuum into dimensionless parts - the durationless instants. The notion of the continuum, although explained with reference to the concept of divisibility, is itself not supposed to be a divisible entity in the ordinary sense of the word. In particular, it is conceivable that the continuum is a description of "indivisible parts".

If this is indeed the case, Kline[40] argues, then Russell's argument cannot be successfully held against Hume's temporal contiguity thesis because in Russell's exposition, time is supposed to be continuous and divisible into instants while in Hume's treatment, time comes in "indivisible parts". At a minimum, Kline maintains, that the idea of indivisible parts entails the discreteness of time.

[40]Kline, A.D. (1985).

So if time can take on discrete units, with each spanning a finite interval, it would then be capable of accommodating the presence of static events whose very existence Russell has so strongly condemned. Of course, the possibility of static events raises the hope to foreclose the temporal gap, as a consequence of which the temporal contiguity thesis would remain intact.

Kline himself has realized the seriousness of this claim and suggests that we must ultimately look to the best physics to see what it teaches about the nature of time. In both classical and quantum physics, in the theories of Euler-Lagrange and Schrödinger, time enters as a continuous parameter. The whole of physics has been founded on a smooth manifold of spacetime. In more recent decades, research in the field of quantum gravity - the attempt to unify quantum mechanics with Einstein's theory of gravitation - has brought the difficulty of reconciling the continuous with the discrete under scrutiny. Several programs[41] within the domain of quantum gravity research entertain serious proposals of an underlying discrete structure of spacetime. Such proposals are pretty much exploratory in nature but represent a positive step forward and we await further advances in this direction.

[41]See Bombelli, L. et al. (1987), Rovelli, C. and Smolin, L. (1995) and Penrose, R. (1971).

Chapter 3

Process Theories of Causation

"A common, if loosely worded, statement of an important consequence of special relativity is: "No signal can travel faster than light". The more sweeping statement, "Nothing can travel faster than light", is contradicted by the familiar example of the spot of light thrown on a sufficiently distant screen by a rotating beacon. The apparent velocity of the spot of light can exceed c, but this does not contradict special relativity since there is no causal relation between successive appearances of the spot."

(*J.C.Garrison, M.W.Mitchell, R.Y. Chiao & E.L. Bolda , 1998*[1])

"A rotating lighthouse beacon impinges on a distant wall. The resulting spot is not a violation of relativity; one spot is not the source of the next spot."

(*Rolf Laudauer, 1993*[2])

[1]Garrison, J.G. et al., (1998).
[2]Laudauer, R. (1993).

3.1 Introduction

In Chapter 2 we learnt from Russell's analysis of the notion of cause the problem of the temporal continuity between cause and effect. The problem arises from the consideration of events as specified by dimensionless spacetime points situated in the temporal continuum. In the continuum, there is no meaning to speak of a next point and so events thought of as occurring at temporal points - the instants - cannot be "next to" each other. Therefore, the supposed temporal continuity of cause and effect (both as point-events) is thwarted as there always exists a *temporal gap* between any two such events. The trouble stems from the misconception that the temporal continuum is made up of discrete instants. This poses tremendous difficulties in thinking about causal chains as sequences of point-events linked together by causal relations. In particular, one asks how one would set out to define the various "components" of these sequences.

To amend the temporal gap, to ensure temporal continuity between two events, we may either think about (i) a continuum of point-events between the two as providing a physical connection or, (ii) physical time as coming in discrete quanta so that "static" events occurring in discrete and yet finite intervals are made possible to provide a connection. Between these two tentative solutions, (ii) seems too controversial a conjecture to proceed with at this point in time. Instead, we pursue a solution to the temporal gap difficulty in the direction of (i).

In relation to (i), an immediate question comes to mind: are we permitted to conceive of a spatiotemporally extended entity as something which is not amenable to be reduced to a sequence of dimensionless point-events? It is to this very idea that we should address ourselves in this chapter and the next.

As a serious acknowledgment to the well-known problems deeply rooted in analyses of causation based on an event ontology (that is, where the causal relation is taken as one that obtains between two events), Wesley Salmon has proposed an alternative theory of causation which takes *spatio-temporal continuous processes* as the basic unit of analysis. These processes, claims Salmon, are themselves the "causal connections" that Humeans have given up looking for. Wesley Salmon's original account, expounded in details in his 1984 book "*Scientific Explanation and Causal Structure of the World*"[3] has sowed the seeds of a series of insightful exchanges between himself and Phil Dowe in the last decade. Inspired and motivated by Salmon's seminal work, the ***Conserved Quantity Theory*** (CQ) was put forward by Dowe[4] in 1992 whose main intention was to modify Salmon's account and to free the process theory from the haunting spell of counterfactuals that was seen to be the main drawback of that theory. This consequently led Salmon to abandon much of his original formulation and to replace the older account by his 1994 ***Invariant Quantity Theory***[5] (IQ). Subsequently, in response chiefly to forceful arguments from Dowe[6] and Hitchcock[7] in 1995, Salmon is understood to adopt a version of ***Conserved Quantity Theory with Transmission***[8] (CQT). It should be emphasized that although now there is much convergence between their respective views on the subject, Salmon and Dowe still have one obstacle to overcome - namely, the place of transmission within the causal analysis. Salmon thinks the concept is vital, Dowe says it is not required. Since a thorough understanding of Salmon's original 1984 theory is essential in order to appreciate the developments of the subsequent philosophical ideas, we shall begin with an exposition of his 1984 account in this chapter and then continue with the more up-to-date philosophical scene on the topic in Chapter 4.

3.2 Salmon's 1984 Process Theory of Causation

It is undeniable that causality plays a vital role in scientific explanation. However, one has to admit that not all scientific laws express causal relations. Take for instance the ideal gas law, $PV = nRT$. As it stands, the law provides only the inter-relationships amongst three quantities, pressure (P), volume (V) and temperature (T) of a gas.

[3]Salmon, W.C. (1984).
[4]Dowe, P. (1992a).
[5]Salmon, W.C. (1994).
[6]Dowe, P. (1995).
[7]Hitchcock, C.R. (1995).
[8]Salmon, W.C. (1997).

It reveals to us only the behaviours of the other variables as and when one or more is varied. It divulges to us no information as regards the temporal order in which each of these variables is to occur. This temporal aspect of causality aside, it still seems possible to give a *causal explanation* by appealing to the actual *physical process* involved. In a situation where a gas is heated, according to the ideal gas law the resulting increase in its temperature is to be accompanied by a corresponding increase in its volume and at the same time, a decrease in the pressure. In another situation, we may find that the temperature of the gas is raised instead by a moving piston as a consequence of which would also have a rise in pressure joined by a fall in volume. Supplying heat to the gas or compressing it by means of a moving piston represent two different physical processes although they lead to the same end results. Although the ideal gas law does not express an explicit "temporal" relation[9], it is capable of being explained *casually* if supplemented by the information of the underlying physical processes. In other words, causal explanations encompass more meaning than the mere temporal orderings amongst events. In our example, two causal explanations are available by references to the two physical processes above. It was Salmon's aim to furnish an account of causation which places due emphasis on these physical aspects.

In the standard view on causality that we inherit from Hume, considerable controversy has centered round the nature of the causal relation and also the relata that this relation is supposed to relate. In this Humean picture, "events" are the relata connected by the cause-effect relations. As for the causal relation, it has been argued to be some kind of statistical relations *or* some rather complicated combinations of sufficient and necessary conditions[10]. However, we have been duly cautioned by Russell on the difficulties inherent in the Humean view when events are defined at spacetime points. With a strong appeal to physical processes, Wesley Salmon presents an alternative account to rival the standard picture, which downplays the status of events in favour of processes that are taken as extended spatio-temporal continuous entities, as the basis units of analysis.

[9]At least it neither reveals nor depicts the time-ordering among the various variables.

[10]A good general introduction can be found in Sosa, E. and Tooley, M. (eds.) (1993), p.1-32.

The development of Salmon's process theory of causation began as early as in the 1970's and has been subsequently organized into a systematic form in his 1984 book "Scientific Explanation and Causal Structure of the World". In this "1984 program", the cause-effect relation is analyzed in terms of three components - an event that constitutes the cause, another event that constitutes the effect and a causal process that connects the two. Within this framework, the basic unit of analysis is a *process* and in this regard two notions have assumed fundamental importance: ***propagation*** (or ***transmission***) and the ***production*** of causal influences. Causal processes are the vehicles by which causal influences are propagated and the production of these influences are brought about by the interactions of two causal processes. The main reason for focusing on causal processes is that in many situations where we talk about the causal relations between pairs of events separated in spacetime, continuous causal processes seem just the right kind of entities to provide the connections to ensure spatiotemporal continuity between them. Also, we iterate that it emphasizes the physical aspect of causation.

This picture can easily be visualized with the aid of a concrete physical example. An electron gun produces an electron that travels towards a fluorescent screen some distance away and impinges on it, resulting in the darkening of a patch of the screen. There are two events representing two changes at two spacetime regions: the electron is brought into being at the source (the electron gun) and the impingement of the electron that brings about the darkening of the screen. The electron in flight provides the connection between these two events, which enables one to entertain a causal story linking the happenings in those two separate spacetime regions. In Salmon's terminology, the causal influence is produced at the electron gun and this influence is carried through space and time over to the fluorescent screen by the electron, which is the casual process. On its arrival at the screen, the electron produces the darkening of the screen by interacting with it. In this example, the electron *is* the causal connection. Consider also the case of an ox pulling a plough as a more congenial everyday example. The changes in the motion of the ox is be to followed by the changes in the motion of the plough via the connection of a belt in-between. There is a causal interaction between the ox and the belt where a force is exerted on the latter. This force eventually reaches the plough and pulls it. The belt is the causal process making a connection between the two events of the ox exerting a pulling force on the belt and the plough being pulled by the belt. Both examples drive home the point that a causal explanation is obtained by appealing to the underlying physical process in each situation.

Given that causal processes are conceived to be continuous spatio-temporal *physical* entities, it is therefore crucial for Salmon to make a distinction between causal and non-causal processes in order to rule out instantaneous causation between two space-like separated regions as prohibited by the theory of special relativity (**SRel**). This is to be achieved by invoking the criterion of mark transmission, as was previously suggested by Reichenbach, where a causal process is defined to be one that is capable of transmitting a mark whereas a pseudo-process is not. A mark, following Reichenbach, is the result of an intervention by means of an irreversible process. One may now pause to ask what does a mark have to do with restricting the propagation of causal processes to sub-luminal velocities? If the mark itself is thought of as some kind of information or signal and since influences in the form of physical information or signals are known to propagate at a velocity which is less than that of light, it follows that a causal process is restrained to propagate at a velocity less than the speed of light in order to be able to carry and transmit a mark.

It is essential to grasp a good understanding of how these causal concepts of causal processes, marks, propagation and interaction all tie together. So we shall now examine each notion in greater details in the ensuing sections.

3.2.1 Causal processes

Signals in the form of electromagnetic waves are propagated from the transmitter at the broadcasting station to the receiver in our television set at home. This is possible because casual influence can be propagated through space and time. Causal processes are the devices by which these influences are propagated or transmitted. Now let us look more closely at the notion of a process in the context of Salmon's theory. While no formal definition is attempted, a process is supposed to be something which has much greater temporal duration and probably of greater spatial extent than an event which is, in comparison, more localised in space and time. In the Minkowski spacetime diagram, events are represented by spacetime points while processes are represented by lines. An object at rest and yet persists through time therefore qualifies as a process. When seen in this light, familiar examples of processes include any object such as a car (whether stationary or moving), a light pulse or even a shadow of an aircraft moving across the landscape. Such objects are considered as processes in the sense that they are represented by their respective worldlines on the Minkowski diagram. A car travelling from town A to town B is a process and so is a stationary car that stays parked in the garage from Monday to Friday. Activation of a photocell is an event but a pulse of light travelling from a distant star and a shadow of an aircraft moving over the landscape would both count as processes.

In general, when one speaks of a process, it is intuitively conceived to be constituted out of a sequence of events. However, Salmon contends that this way of thinking about a process is not essential. For him, an essential feature of a process lies in the degree of constancy of of its own structure[11],

> ***A given process, whether causal or pseudo, has a certain degree of uniformity - we may say, somewhat loosely, that it exhibits a certain degree of structure.***

Without committing to a definition of what is considered to be "a constancy of structure", Salmon regards a process as the persistence in spacetime of an entity, which invariably manifests itself in the constancy of quality or structure as similar to Russell's "causal lines" as we have already met in Section 2.6.

Intended for as the foundation of a theory of physical causation, Salmon finds it necessary to lend the concept of processes scientific legitimacy. In **SRel**, an "event" is a frame-independent concept. And the spacetime manifold is regarded as a collection of events that bear space-time relations to one another. That is, **SRel** has been built

[11]Salmon, W.C. (1984), p.144.

upon an "event ontology". An alternative approach to **SRel**, originated by Alfred A. Robb[12], employs solely the idea of light paths to determine the fundamental relation of "before and after" between time instants out of which the geometrical properties of space can be constructed. Furthermore, he also shows that it is possible to recover Minkowski's purely analytical treatment by the introduction of coordinates in his spacetime system. In particular, the "before" and "after" relation enables him to establish the invariance of the interval as in Minkowski geometry. Robb carried the deep conviction that a description for spacetime relations by abstract geometry should correspond with physical facts and he thought such a correspondence ought to be established by the physical properties of light, because the experimental observations at that time all led to the conclusion that the speed of light represents an upper limit, a physical bound, at which causal influences can travel. The "causal element" enters into Robb's theory via the definition of the primitive "after" relation by reference to the ability of producing an effect at one instant by another. Causality is thus reduced to a temporal relation between events. Paths of light are the basic entities used in his theory and Robb especially calls attention to their *continuous* character[13],

> ***The types of geometry with which we are specially concerned when we attempt to map out time and space involve an infinite set of elements forming what is called a "Continuum". The classes of these elements, such as lines, planes, etc., with which we are concerned, are defined by means of certain relations among the elements involved.***

The continuous character of these basic units of analysis seems just what Salmon has intended for his processes. Based on Robb's seminal work, Salmon argues that one may equally well adopt a "process ontology" for **SRel**. **SRel** places a constraint on the upper limit of propagation of signals and information. It is because of this that a process, as a carrier of causal influences in the form of signals or information, must therefore be restricted to travelling at less than or equal to the speed of light c. In

[12]Robb, A.A. (1936).

[13]ibid., p.6.

the Minkowskian language, this is tantamount to the statement that causal processes are coextensive with the lightcone - a conical region of spacetime surrounding an event which can be reached by a velocity less than or equal to the light speed. Only those processes found lying within or on the lightcone are capable as carriers of causal influences. This serves to rule out processes that travel at arbitrarily high velocities (those greater than c); that is, those which lie in the spacetime region outside the lightcone and hence are deemed incapable of transmitting causal influences and regarded as unphysical. Experience shows that damage is brought to a causal process like a car when it collides with a lamp post. On the other hand, a process like a shadow cannot be causal because processes as such are unable to bring about a genuine physical change, like an actual structural damage in the case of the car. The shadow of a car may be momentarily distorted as it crosses a lamp post, but it "regains" its original shape immediately after the encounter.

To cast these ideas in a better perspective, consider Salmon's paradigm example of the astrodome. We are invited to think of a large circular building with a spotlight mounted at its center. Once the spotlight is switched on, a ray of light travels from the source to fall on a certain point on the far wall and this ray constitutes a causal process. Now suppose the spotlight is mounted on some sort of rotating mechanism like a turntable and is set into rotation. Although the spot of light it casts upon the wall revolves in a highly regular fashion - thereby possesses the essential feature of being a process - Salmon wants to rule out this revolving spot of light as a genuine causal process.

To understand how he makes the distinction, it is advisable to first take stock of the similarity and difference in the two cases. The ray of light and the spot are similar in that both appear to be spatio-temporal continuous processes exhibiting a high degree of uniformity. The subtle difference lies in the fact that the ray of light is capable to sustain itself without an external aid once it departs from the source. On the contrary, the light spot on the wall will not persist without the source, quite independent of its prior history. Put another way, the continuous appearance of the moving spot of light is constituted out of a series of "unlinked" units of *different* tiny spots of light at each position on the wall each spot occupies. These spots are created by the series of positions on which the rotating beam falls. We ought then to devise a method, a criterion to distinguish this sort of disconnected regularities from those that arise because of a cause-effect relation. The chosen criterion must capture the ability of causal processes to produce physical changes and these resulting changes are to be "carried along", or in technical jargon, transmitted by the process itself through space and time. In Salmon's framework, the essential features of a causal process are: (i) that it would persist even if it was isolated from external causal influences (that is, regardless of whatever happens elsewhere), (ii) it has the capacity to produce physical changes (that is, to "mark" or being "marked" when coming into contact and interacting with other processes) and (iii) its ability to transmit causal influences (that is, transmitting the mark that results from its physical interaction with another process). These ideas are summarized in the following quote[14],

> ***A causal process is one that is self-determined and not parasitic upon other causal influences. A causal process is one that transmits energy, as well as information and causal influence. The fundamental criterion for distinguishing self-determined energy transmitting processes from pseudo-processes is the capability of such processes of transmitting marks.***

The notion of a causal process is thus intuitively bound up with the idea of the transmission of marks, since it is defined in terms of the latter. Because of the central role played by the notion of transmission, it deserves a detailed examination in its own right. For heuristic reasons it would perhaps seem more appropriate to consider first the production of marks (or physical changes) before dealing with the concept of transmission. However, as the business of mark transmission is so much intertwined with the nature of causal processes, we shall now move on to consider the topic of causal transmission in the next section.

[14]ibid., p.146.

3.2.2 Causal Transmission

Causal processes are identified as so by their abilities to participate in interactions that produce physical changes and to *carry along* or *propagate* these changes across spacetime regions. Speaking in the language of marks as physical changes undergone by processes, we say that causal processes can be marked and that they have the potential to transmit these marks. On reflection, pseudo-processes may also be marked but the crucial difference is that the marks or changes so acquired are not genuinely physical in the sense that they cannot be transmitted beyond the point of interaction. A shadow may be distorted momentarily when it passes along a rough patch of the wall but would "spring back" to its original shape once it moves beyond.

Let us now return to the astrodome and concentrate on the light ray travelling from the rotating beacon to the wall. One may "intervene" by placing a red filter somewhere in its path. The intervention constitutes the production of a mark and at that point, the light ray is marked - it experiences a physical change as it becomes red. This mark is a genuine physical change as the light ray stays red all the way until it reaches the wall. In a like manner, the moving spot of light may also be marked by placing a piece of red glass at a particular point on the wall. The process - the light spot - will be modified and becomes red only at that one point but continues on beyond that point *as if* no interaction had occurred previously. These examples serve to illustrate a very important condition for causal transmission, namely that, ***"a single intervention at one point in the process transform it in a way that persists from that point on"***[15].

In the case of the moving spot of light, one may choose to make the spot red at other places by the installation of a red lens in the source. This would, however, not constitute a local intervention at an isolated point in the process (here the spot) itself. One may also think of putting red glasses at places along the wall but again this fails to satisfy the above condition since it involves many interventions rather than just a single one. The condition thus, looks as though it is *prima facie* able to eliminate pseudo-processes. However, a clever counterexample exploited by Nancy Cartwright

[15]ibid., p.142.

seems to have circumvented the above condition. If a red glass is inserted at the source a fraction of a second before the red lens is placed at the position of the wall on which the spot of light falls, argues Cartwright, then the light spot becomes red the moment it reaches the wall and will stay red (but only because of the red glass at the source) from that point on. This would make it appear like a single intervention at one point on the wall in the process of the moving spot has transformed it in such a way that persists beyond that point. To remedy this shortcoming, Salmon was forced to introduce a counterfactual claim into the above condition for causal transmission. He has to stipulate in the condition that the moving spot would have turned red in any case because of the red lens inserted at the source, even given that no marking activity had occurred locally at the wall. We are now in a position to state the modified criterion for mark transmission (**MT**)[16],

> ***Let P be a process that, in the absence of interactions with other processes, would remain uniform with respect to a characteristic Q, which it would manifest consistently over an interval that includes both of the space-time points A and B (A ≠ B). Then a mark (consisting of a modification of Q into Q'), which has been introduced into process P by means of a single local interaction at point A, is transmitted to point B, if P manifests the modification Q' at B and all stages of the process between A and B without additional interventions.***

Let us apply this criterion to the respective processes of the light ray and the spot of light and see if it fulfills its promise to make a distinction between the two. The light ray (process P), in the absence of interactions with other processes and being left on its own, would remain uniform with respect to the characteristic (Q) of being "white", which it manifests constantly over an interval including the two spacetime points from the source (A) to the wall (B). Then a mark that consists of a modification of the characteristic of being white (Q) into one being red (Q'),

[16]ibid., p.148.

that has been introduced into the light ray by means of the insertion of a red glass at the source (A), is said to be "transmitted" to the wall (B) if the light ray (P) manifests the modification of being red (Q') at the wall (B) and at all stages between A and B without additional interventions. The question of whether process P is to maintain the modification all the way from the source to the wall is to be settled by observations and here it seems to be just what is expected.

Next we consider the case of the moving spot of light cast upon the wall. In the absence of interactions with other processes, this process (P) would also remain uniform with respect to the characteristic (Q) of being a "white" spot and continue to manifest this characteristic uniformly over an interval between any two positions (A and B) along the wall. Once again, a mark can be introduced into P by way of a red lens placed at A, which results in the characteristic of being a white spot (Q) changed into that of being red (Q'). Here, observations show that this mark (Q') would not be transmitted to point B since once the spot moves beyond A, it reverts to being white once more. So, *prima facie*, **MT** can indeed be regarded as a sufficient criterion to make the distinction between causal and pseudo-processes.

The real test, however, remains with the Cartwright type counterexamples. The magic touch comes from the assurance levied upon the counterfactual claim that the process P of the light spot would maintain the characteristic of being white in the absence of interactions with other processes. In Cartwright's argument, even though the process itself has not gone into any sort of interaction with other processes (in this case, a red lens has not been placed at the point A), observations show that the spot would still become red; the spot would still suffer the modification Q'. In this instance, the counterfactual claim is proved to be wrong and hence the conclusion that P is not a causal process follows. The falsehood of the counterfactual comes about because of the dependence of this particular process on an external source and it highlights the inability of the process to transmit the mark within itself. The introduction of a counterfactual then seems inevitable for the effective formulation of **MT**. Salmon seems to be much disturbed by this issue and fears that counterfactualness would undermine the objectivity of **MT**, as the determination of the truths of counterfactuals is by no means a straightforward philosophical matter. Since, as we shall see, that counterfactuals are also required in Salmon's definition of causal interaction, we shall delay our exposition of counterfactuals until Section 3.3, which this issue will be examined in some detail.

We now turn to the core of **MT** - the notion of transmission. **MT** stipulates that a mark (which has been introduced into the process by means of a single interaction) at A, is "transmitted" to point B provided that the process P manifests the modification "*at*" B and "*at*" all stages of P between A and B. For this reason, **MT** is also widely known as the "*at-at*" theory of causal transmission.

The "*at-at*" theory was first proposed by Russell as a solution to the paradox of the flying arrow, one of the well-known Zeno paradoxes of motion. Intuitively, transmission involves something that moves across space and time and that is precisely why it is so closely associated with the notion of motion. In particular, it captures the fact that, for an entity to be transmitted from one point to another in the spacetime continuum, it must traverse "continuously" all the intermediate points. However, since there is no concept of a next point in a continuum, and in-between any two points, there always exists a continuum of others, one must therefore not consider separately the individual spacetime points within the motion except for the correlation of the different places with different times[17],

> ***If there is to be motion, we must not analyse the relation into occupation of a place and occupation of a time. For a moving particle occupies many places, and the essence of motion lies in the fact that they are occupied at different times.***

And again[18],

> ***Motion consists in the fact that, by occupation of a place at a time, a correlation is established between places and times; when different times, throughout any period however short, are***

[17]Russell, B. (1903), p.472.
[18]ibid., p.473.

correlated with different places, there is motion; when different times, throughout some period however short, are all correlated with the same place, there is rest.

We may now proceed to state our doctrine of motion in abstract logical terms, remembering that material particles are replaced by many-one relations of all times to some places, or of all terms of a continuous one-dimensional series t to some terms of a continuous three-dimensional series s. Motion consists broadly in the correlation of different terms of t with different terms of s. A relation R which has a single term of s for its converse domain corresponds to a material particle which is at rest throughout all time...The motion is continuous if the correlating relation R defines a continuous function. It is to be taken as part of the definition of motion that it is continuous, and that further it has first and second differential coefficients.

In classical mechanics, such continuous functions are the equations of motion of a physical system, which are in essence mappings between the continuous series of coordinates and that of time that take the form: $x = f(t)$. Different forms of the function f specify the different trajectories that correspond to the respective positions x of the system at the corresponding times t.

Rather than thinking about how we are to connect the individual spacetime points that the particle occupies, one correlates instead a *continuous* series of places that a particle occupies during its motion with a *continuous* series of corresponding times at which the particle is found at those places. In effect, we have turned our attention from the individual single stages of the motion, to a multitude of stages, the "connection" of which is described by the continuous and differentiable functions.

The central role of the notion of transmission in Salmon's theory is thus justified. Transmission, as explicated in terms of the "*at-at*" theory, equips one with these continuous, differentiable functions, which represent the equations of motion of physical systems that provide the physical connection of events situated at two locations in spacetime to amend the temporal gap. These physical connections are the causal connections we seek in Salmon's theory of physical causation.

Classical laws of motion all subscribe to this "*at-at*" idea, as a consequence of which they have become inseparable from the notion of a *continuous trajectory* in spacetime.

There is, apart from the differential representation of a particle's trajectory, also the integral approach that takes into consideration the *entire motion.* While the two approaches stand on a par when one comes to consider causation in classical physics, there exist good reasons to believe that the integral approach is the more fundamental one whereby Salmon's idea of a causal process may be extended to take account of causation in quantum physics. This task will be taken up in Chapters 5-7.

A related comment on the "*at-at*" theory of causal transmission is to do with the status of causal laws. In "*An "At-At" Theory of Causal Influence*", Salmon says the following[19],

> ***Causal Processes are of course, governed by natural laws; these laws constitute regularities whose presence can be empirically confirmed. Such regularities, presumably, represent the kinds of constant conjunctions to which Hume referred. The mark method may be said, roughly speaking, to provide a method for distinguishing causal regularities from other types of regularity in the world, including that which may be associated with pseudo-processes.***

Causal processes are governed by natural laws: a classical object moving under a conservative force obeys Newton's laws of motion for example. It is by virtue of

[19]Salmon, W.C. (1977), p.223.

these *functional laws* that a description of the transmission of marks is possible via the "*at-at*" theory. Such functional laws constitute the kind of constant conjoining regularities perceived by the Humeans. With the functional laws, the concept of the transmission of marks supplies the Humean regularities. Here, it is useful to follow Hempel[20] to make a distinction between what he calls *laws of coexistence* and *laws of succession.* Laws of coexistence expresses merely a mathematical relationship but laws of succession are concerned with temporal changes in a system. Those physical laws that fall under the heading of the laws of succession are what we regard as causal laws and a functional law like Newton's second law of motion falls neatly into such a category. Functional laws are thus elevated to the status of causal laws in this sense. It must, however, be borne in mind that these causal laws are established by induction from empirical observations and they are susceptible to be proven wrong observationally. So to crown them with the grand title of causal laws does *not* imply that they bear an element of ontological necessity. In fact, Salmon intends his causal theory to provide only a *sufficient* condition for causation to obtain in physical situations.

3.2.3 Casual Interactions

Given that marks are essential in deciding whether a process is causal or not and the production of marks is to be explicated as the concept of causal interactions between processes; this is how the notion of interaction becomes firmly anchored in Salmon's theory of causation. Interactions are important since they are responsible for the production of order and structure of causal processes. Experience tells us that after a direct physical interaction between two processes (or loosely speaking, moving objects), both will be affected in such a way that any changes so suffered in each as a result of the encounter would become correlated. Two billiard balls collide and their subsequent motion are to be constrained by a correlation governed by the law of conservation of linear momentum. In moving from the classical picture to a quantum one, the notion of interaction must now be modified to take account of the

[20]Hempel, C., (1965).

probabilistic character.

In the Müller scattering of two electrons, $e^- e^- \rightarrow e^- e^-$, the rules of quantum mechanics provide for us a recipe to obtain the so-called scattering amplitude of the interaction. The modulus square of this amplitude gives essentially the *probability for such an interaction* (the scattering) to take place. The calculation involves the considerations of both energy and momentum. The probability characterisation of interactions arises from the statistical nature of events in the world as revealed by quantum theory. However, this suggestion of characterizing interactions by statistical correlations does by no means go without reservations, for one immediately comes to notice that there are situations where although two processes are statistically correlated, the dependency arises not due to any actual physical interaction between the two, but rather to some other special background conditions. These background conditions are what generally known as *common causes.* Examples of common-cause situations are many and varied and a simple one would be that of poisonous mushrooms being sold in a supermarket and have been independently procured by two individuals X and Y who had fallen ill on account of them. There are two independent causal processes linking the consumption of the mushrooms to each individual but there is no physical interaction between X and Y whatsoever. In fact, X and Y need not be in the slightest acquaintance! The illness of X has no direct consequence upon that of Y. Besides, chance coincidences are, of course, not impossible - X could have fallen ill because of the food intake of another kind instead of the mushrooms. These considerations help to illustrate that the existence of common causes cannot be inferred with absolute certainty but only with some degree of probability. This probability is surely to increase if there are to be repeated occurrences of the coincidence. On this basis, Reichenbach[21] characterizes common causes by a species of statistical forks - the *conjunctive forks* - to capture the idea that the common cause is the connecting link that transforms an independence into a dependence,

$$P(A,B) > P(A)\times P(B) \tag{3.1}$$

[21]Reichenbach, H. (1956), p.157ff.

Formally, a conjunctive fork is best seen as an indicator of the existence of a common cause provided the following conditions are met,

$$P(A, B|C) = P(A|C) \times P(B|C) \tag{3.2}$$

$$P(A, B|\bar{C}) = P(A|\bar{C}) \times P(B|\bar{C}) \tag{3.3}$$

$$P(A|C) > P(A|\bar{C}) \tag{3.4}$$

$$P(B|C) > P(B|\bar{C}) \tag{3.5}$$

The first two conditions describe the fact that the correlation between the events A and B is dependent on the occurrence (or non-occurrence) of a third event C. The last two conditions capture the fact that the occurrences of events A and B are individually dependent on that of C.

It can be shown that the above four defining conditions for the conjunctive fork entails the "*screening-off*" condition:

$$P(B|C) = P(B|A, C) \tag{3.6}$$

To screen-off means to make statistically irrelevant. The condition stipulates that the common cause C makes each of the two effects A and B statistically irrelevant to each other. In other words, A and B are correlated only because a variation in the outcome C induces a correlated variation in the outcomes of A and B. In the absence of any variation in C, there is no correlation between A and B: there is no direct interaction between A and B even though there may be two independent processes with one leading from C to A and the other from C to B. Unfortunately, the analysis is not as straightforward as Reichenbach had anticipated. While the principle of common cause asserts that an event that qualifies as a common cause must satisfy the four defining conditions, the converse does not necessarily hold since any event C that fulfills these conditions may not qualify as a bona fide common cause of A and B. The identification of the legitimate common cause depends upon the successful identification of the physical processes leading from the common cause event to each of the two correlated subsequent events.

It is indeed Salmon's important observation that there exists a specific type of common cause that violates the conjunctive fork. This happens when two processes interact which leads to mutually correlated modifications that remain with them beyond the locus of interaction. Under such circumstances, the two processes are said to participate in a *causal interaction* and the conjunctive fork is to be replaced by what Salmon calls an *interactive fork*:

$$P(A, B|C) > P(A|C) \times P(B|C) \tag{3.7}$$

Here, the common cause C does not statistically screen-off the two effects A and B from each other. Rather, the relevant fact is that A and B are directly correlated due to the physical interaction C. The geometrical picture behind an interaction is one that involves the *intersection* (or encounter) of *two causal processes at a particular region of spacetime.*

However, one may at once raise the question if the intersection of two pseudo-processes in a region of spacetime is to be readily regarded as an interaction. Is there a difference between intersections that lead to genuine causal interactions and those that do not? The shadows of the paths of two airplanes may cross and coincide momentarily but as soon as the shadows have passed the intersection point, both move as if nothing had ever occurred. So a crucial ingredient is required to distinguish between the mere crossings of processes without any form of interaction and those crossings at which real physical interactions take place. It is often observed that if there had been an actual physical interaction, each process concerned would have suffered an enduring modification as a consequence of their encounter. Modifications can already occur whenever at least one of the processes involved is a causal process. Placing a red lens on the wall marks the light spot red. Although the spot is a pseudo-process, it nevertheless *interacts* with another causal process - the red lens - and is modified to becoming red but only at the point of intersection and not beyond. On the other hand, it is entirely possible for two causal processes to intersect without neither suffering subsequent modifications. Two light rays can pass through each other without leaving any lasting effect upon either. Both rays have the capacity to undergo physical interactions but in this instance they do not. One is thus cautioned on the fact that the requirement of both the intersecting processes be causal is a necessary though not sufficient condition of producing lasting changes in both. But of course, what we want to capture in the theory is that whenever there is a genuine physical interaction, the processes involved would suffer *lasting* modifications beyond the point of their encounter[22]. With these observations in mind, Salmon formulates the ***Principle of Causal Interaction*** (CI)[23],

> ***Let P_1 and P_2 be two processes that intersect with one another at the space-time point S, which belongs to the histories of both. Let Q be a characteristic that process P_1 would exhibit throughout an interval (which includes subintervals on both sides of S in the history of P_1) if the intersection with P_2 did not occur; let R be a characteristic that P_2 would exhibit throughout an interval (which includes sub-intervals on both sides of S in the history of P_2) if the intersection with P_1 did not occur. Then, the intersection of P_1 and P_2 at S constitutes a causal interaction if:***

(1) ***P_1 exhibits the characteristic Q before S, but it exhibits a modified characteristic Q' throughout an interval immediately following S;***

[22]This requirement implies that both processes need to be causal.
[23]Salmon, W.C. (1984), p.171.

and

(2) ***P_2 exhibits the characteristic R before S, but it exhibits a modified characteristic R' throughout an interval immediately following S.***

A few obvious comments on the above definition are in place. First, as acknowledged by Salmon himself, **CI** is clearly counterfactual[24]. The introduction of the counterfactual responds chiefly to the need to bar cases where the intersection of two pseudo-processes appears to satisfy the principle. When a red lens is installed at the rotating spotlight, the light spot on the wall would become red (and not maintaining its previous whiteness) irrespective of the fact that the spot itself has not undergone any intersection with another process. The failure of the counterfactual clause in this instance serves to dismiss a possible intersection between the spot and the red filter placed at one location of the wall as a genuine causal interaction.

Salmon resorts, as we shall see later in Section 3.3.2, to elaborated controlled experiments to assist the establishment of the truth values of these counterfactuals. His dissatisfaction with the adoption of a counterfactual condition in the definition stems from the same worry he has for **MT**, namely, that the overall objectivity of the theory may be grossly undermined.

Second, **CI** is formulated explicitly in terms of *two and two processes only* (the "X"-type interactions) and it requires the continuing existence of these two processes even though they are to suffer mutual modifications in an interaction. This is particularly problematic where there is to be creation or annihilation of processes resulting from the interaction. Examples of this kind abound in the world of elementary particle physics. A radioactive nucleus may undergo spontaneous decay into a number of particles. Salmon calls cases like these Y-type interactions. In an annihilation interaction between a positron and an electron, one finds a situation in which the e^+e^- pair "vanishes" and in their place come two different processes altogether. There are also processes such as a snake swallowing an egg, which constitutes a typical λ-type

[24]As one is informed by the phrases: "Let Q (R) be a characteristic that process P_1 (P_2) would exhibit through an interval if the intersection with P_2 (P_1) did not occur."

interaction. He admits that it would be far more desirable for a principle of causal interaction to be able to deal with these important physical scenarios as well[25],

> ***Since a large number of fundamental physical interactions are of the Y-type and of the λ-type, there would appear to be a significant advantage in defining interactive forks in terms of these configurations, instead of the x-type. Unfortunately, I have not seen how this can be accomplished, for it seems essential to have two processes going in and two processes coming out in order to exploit the idea of mutual modification.***

There is a deeper dimension to Salmon's thinking which comes to mind. He has explained that the reference to two incoming processes and two outgoing processes is required to capture the idea of mutual modification. However, the requirement is actually more stringent than what it appears to be. Each of the two incoming processes must maintain its identity over time; that is, to remain the same process in order to provide a basis for the modifications or *changes* of the respective characteristics or properties to occur. Clearly, such a requirement is quite incapable of dealing with changes that correspond to entities that pop in and out of existence. This unfortunate state of affairs is exemplified by the fundamental physical phenomena of the creation and annihilation of subatomic particles. With the very issue of identity over time at stake, are we able retain this concept of change? It is without doubt that coming into being or ceasing to exist would undeniably count as legitimate changes that may happen to an entity. We shall not dwell upon these issues here but this discussion will re-emerge again later in the chapter to follow next.

I should, however, indicate a possibly misleading point. This concerns the linguistic usage of the term "interaction". Interaction is, in its simplest form, a two-place relation between *two incoming processes.* It does seem rather odd and counterintuitive to say that a single process "interacts with itself" and in particular, in Y-type interactions where only a single incoming process is involved (for instance, in the case

[25]ibid., p.181-182.

of a spontaneous decay). However, a precise treatment according to quantum field theory tells us that it is indeed the interaction with the vacuum that induces the decay of the atom.

So, although at first sight **CI** does not seem to be able to cater for the Y-type "interactions", it is an adequate condition if these Y-type interactions are analyzed in more correct and finer[26] details.

Third, notice that neither the interactive nor the conjunctive fork enters into the definition of **CI**, though they have both played a large part in motivating the principle. And in fact Salmon thinks that statistical characterization of causal interactions is not desirable. He says, and I quote, from a footnote[27],

> ***In (Salmon, 1978), I suggested that interactive forks could be defined statistically, in analogy with conjunctive forks, but I now think that the statistical characterization is inadvisable.***

In this connection, Ben Rogers[28] has ably conducted a careful analysis on how the statistical forks, both of the conjunctive and interactive varieties, may possibly relate to the **CI**. To be sure, the conjunctive fork has been defined in terms of events and in order to make a comparison with cases of processes, one must set up a correspondence between events and processes. This is done as follows. A process is considered to be associated with a set of properties that a process of that kind may possess. The state of a process at a particular time is taken to be the set of properties it possesses at that time. In this framework, we are allowed to think about the probability of the process being in any other possible states in another time. A causal interaction would then involve a change in the probability distributions over the possible states of each process as a result of the interaction. Following Rogers, let A denotes the event where the first process has a particular property at t_2 and B be the event that the second process has a property at t_2. Also let C represent the event where the two processes have intersected at a previous time t_1 ($t_2 > t_1$) and $\bar{C}$ represent the

[26]The atom and vacuum field represent two processes.
[27]ibid., p.174.
[28]Rogers, B. (1981).

event where the two processes have not intersected at time t_1. It is obvious that the following relations hold:

$$P(A, B|C) \neq P(A|C) \times P(B|C) \tag{3.8}$$

$$P(A, B|\bar{C}) \neq P(A|\bar{C}) \times P(B|\bar{C}) \tag{3.9}$$

$$P(A|C) \neq P(A|\bar{C}) \tag{3.10}$$

$$P(B|C) \neq P(B|\bar{C}) \tag{3.11}$$

Comparing the first condition (equation (3.8)) with that which characterizes the conjunctive fork in equations (3.2-3.5), one comes to realise that while the definition of the conjunctive fork requires a cause to raise the probability of its effect, the definition of causal interaction only requires a *change* in the probability of an event with respect to those cases with interactions and those without. Another salient point should by now become transparent: the condition in equation (3.8) implied by the thesis of causal interaction is different from that respectively implied by both the conjunctive and the interactive forks. So Rogers concludes[29],

> ***As far as interactions are concerned, the important fact is that a causal interaction brings about a change (increase or decrease) in the distribution of probabilities among the possible state of affairs, measured with respect to the distribution given by the law of non-interactive evolution of the process in question.***

and[30],

[29]ibid., p.208.
[30]ibid., p.213.

> ***I have argued above that the characterization of causal interaction, (which is required to get clear about any notion of common cause) should be more general than those of either conjunctive fork or interactive fork.***

A last side-comment on **CI** concerns the temporal symmetrical character of the principle. Normally, when one speaks of two processes being "modified as a result" of their interaction, one has already assumed which are the prior states and which are the subsequent ones. Salmon considers such introduction of a prior temporal asymmetry inadvisable for it would not admit a causal theory of time. In his view, it is best to treat both the notions of causal interaction and causal transmission as temporally symmetric[31],

> ***The principle CI, as formulated previously, involves temporal commitments of just this sort. However, these can be purged easily by saying that P_1 and P_2 exhibit Q and R, respectively, on one side of the intersection at S. With this reformulation, CI becomes temporally symmetric. When one is dealing with questions of temporal anisotropy or "direction", this symmetric formulation should be adopted. Problems regarding the structure of time are not of primary concern in this book; nevertheless, I am trying to develop causal concepts that will fit harmoniously with a causal theory of time.***

3.3 The Epistemology of Causation

The problems raised by Hume's famous empiricist critique of causation pose a major challenge for later generations of scholars on the subject. Just how one comes to justify that there indeed exist causal connections in nature that are quite independent of the human intellect and at the same time conform faithfully to the Humean empirical strictures? Salmon stands up to this challenge with his 1984 theory of process

[31] Salmon, W.C. (1984), p.176, footnote.

causation.

One often infer causal claims from premises that record one's observations and causal judgements are verified by observing that certain conditions obtain. In order to adhere to *empirical* doctrines, such conditions are to be formulated in such a way that they deploy only concepts whose satisfaction is *observationally accessible.* It is vital that one does not introduce conditions that makes one unable to explain how one can know whether or not they obtain. Indeed, if causal knowledge is inferential knowledge from observations, it is all the more wanting to have an account of observational conditions on which causal judgements are based. This is precisely what Salmon sets out to achieve by trying to reduce the causal concepts of *processes, marks, transmission* and *interactions* to empirical observations. Salmon's 1984 theory is an explicit account of knowledge linking the causal relation to its conditions of observability. Disciplined by empirical observability, Salmon's theory is a promising candidate to initiate the contact between causation and a science which is founded on empirical observations like physics.

As we have already expounded in some details in the previous sections, the picture that Salmon 1984 presents to us is one such that the causal structure of the world can be seen as an attempt to *specify the relation that obtains* between two events (at two spacetime points) just in case the earlier is a causal factor in the occurrence of the latter. Here a continuous path (of some physical entity) in spacetime (a causal process in Salmon's terminology) that terminates in the two events (the causal interactions) at both ends, is a prerequisite for the two events to be connected. This mode of connection in turn warrants that the events are causally related. In essence, the causal relation is one of the connectibility between two events. This leaves us with the question of how the central notions of *process* and *interaction* may be defined in such a way to satisfy the empiricist requirement that conditions from which we are incapable of telling how we can know whether they obtain or not are excluded.

In relation to the epistemology of causal processes, we consider the following. For any two distinct points in the continuous spacetime manifold, there exists an infinity of paths connecting them. How is it possible for us to delineate those which are considered as causal? In the first instance, Salmon appeals to the theory of special relativity and relies on the possibility of physical transmission of information at velocities less than the speed of light to make the distinction between causal paths and non-causal ones. This, however, calls for an immediate refinement, for it does not take one long to realise that non-causal or pseudo-processes may travel at sub-luminal velocities, although in that case it remains incapable of transmitting information or signals. This prompts Salmon to adopt the concept of transmission of (causal) information as a criterion to decide causal processes against pseudo ones. Causal processes are distinguished by virtue of their ability to transmit *or* propagate causal information, it is to be insisted that there is *no* transmission of causal information even for $v \ll c$, as far as pseudo-processes are concerned. This, as we have discussed, is the rationale behind **MT**, which stipulates that a process, P, linking two spacetime points, c and e, is *causal* iff upon a single local interaction within P, P suffers a *lasting* modification in one of its characteristics, which would have maintained its original form had the interaction not occurred. Moreover, a "lasting" modification or change in the characteristic means that the modification is to be found at all spacetime points from c to e in the absence of further interactions. The concept of transmission, argues Salmon, has now been fully explicated in terms of the empirically observational statement: "*the modification in the characteristic of P is to be found at all spacetime points from c to e.*" - the celebrated "*at-at*" theory of causal transmission. The "*at-at*" theory, as we have already explained, forms the basis of the continuous functional laws of physics. Physical entities with equations of motion given by these laws are thus regarded as causal processes. The question of which particular subset of the spacetime manifold a causal process is comprised may now be answered by reference to the physical law that governs the particular process.

In other words, a process that qualifies as causal must be able to undergo a *a single local interaction* and an *at-at transmission* of its causally modified characteristics. To complete the task, it remains for Salmon to reduce the causal notion of interaction to observational conditions. His *Principle of Causal Interaction* (**CI**) which attempts to explicate interaction in terms of the geometrical intersection in spacetime of two processes which can be observed. However, this is complicated by the need to specify which intersections are to be considered as proper interactions. **CI** stipulates that only those intersections, as a result of which both processes suffer mutual changes in some of their own characteristics, would count as legitimate interactions. The interrelations between these concepts are depicted in the chart given in Figure 3-1 below.

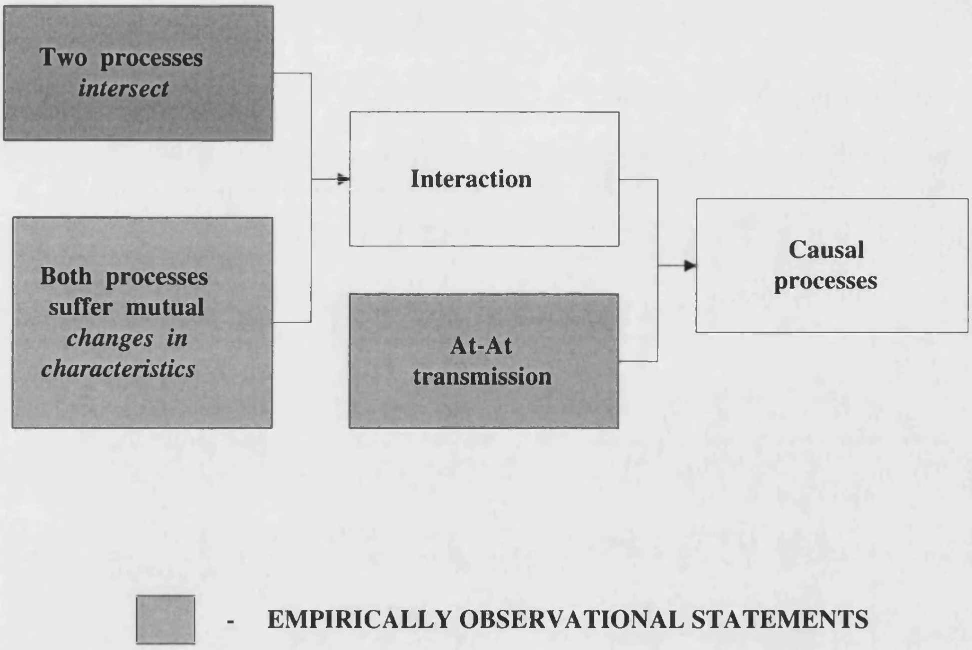

Figure 3.1: Interrelations of Causal Concepts in Salmon 1984.

At first sight, it seems that Salmon has managed to reduce the causal notions of processes and interactions to empirically verifiable observational statements, or has he not? As for causal interaction, the answer rests with whether all *"changes in the characteristics"* resulting from the interactions of processes are observable in principle or not. Cases of collision between moving billiard balls that lead to observable changes in the motions of both surely fit very neatly into the category. However, it is not hard to think of examples where there are interactions that do not lead to any observable change in the motions of the interacting objects. One can find many situations of this kind when a system of forces act in a state of equilibrium. Take for example two weights that are attached, each at one end, to a lever that keeps the latter in equilibrium (Figure 3-2). In equilibrium, forces *act* in such a way that the resultant of the observable individual effects of each weight[32] acting alone is zero and one therefore observes no movement of the lever[33].

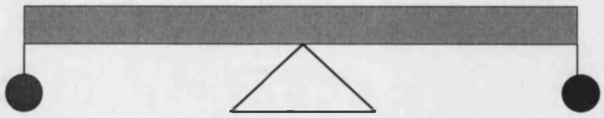

Figure 3.2: Two weights attached to a lever to maintain its equilibrium

There are several interesting aspects to this scenario. First, there is obviously an intersection of two processes at each point of contact of between a weight and the lever. However, if Salmon's **CI** is to be construed as an intersection at a "single spacetime point"[34], then for this particular system, there appears to be continuous intersections at both points of contact, at least in a temporal sense. The continuous interactions of each mass with the lever cancel the effects that correspond to each being allowed to act alone. As a result, the lever remains leveled *as if* it is free from all interactions. So it seems not always the case that interactions are to be accompanied by observable changes[35] in an entity's characteristic (the motion of the lever in this instance).

Sober [36] voices his objection to **CI** by considering a similar scenario,

> ***Suppose a billiard ball in uniform motion were bombarded by a multiplicity of influences that cancelled out each other. Many component causes impinged, but the net force was zero. The***

[32]If only one weight is attached, say, at the end A, the lever would have been tilted in an anti-clockwise direction because of the torque.

[33]Strains and stresses are regarded as immaterial.

[34]Construed as *a* particular spatial point at *a* specific time.

[35]This is reminiscent of the argument leading to the formulation of the *Principle of Virtual Work* in classical mechanics. Indeed, the great founder of analytical mechanics, Jean-Louis Lagrange (Lagrange, J-L. (1997)) opened his timeless masterpiece "*Mécanique Analytique*" by asserting that the effect of forces are only empirically known through the changes in motion: the work done they produce. So as the argument proceeds, the principle of virtual work states that the total work done by applied forces in keeping a system in equilibrium must be zero since the system exhibits no observable changes in motion.

[36]Sober, E. (1987), p.255.

> ***ball persisted with the velocity it had before. But this subsequent state was not without its causes. After all, forces impinged, and what are component forces if not component causes?***

In most circumstances, the billiard ball would have had its motion altered by the impingement of one force and the alteration in its motion is an observable fact. Upon the bombardment by a multiplicity of forces, the individual effects due to the component forces cancel each other, leaving the ball in a state that is *observationally equivalent* to where no interaction has taken place. Once again, we conclude that there are instances where genuine interactions can occur without leading to any observable changes in characteristics of an object that participates in these interactions.

But before we may decide whether there is in fact an observable change in a characteristic, we ought to be sure about which characteristics we are making reference to. This turns out to be an issue beset by ambiguities. Indeed, Kitcher[37] shows in a most strikingly extreme example that the unrestricted choice of the characteristics of intersecting processes can render **CI** empty. In his example, the *characteristic* that changes upon an intersection is taken to be the condition that for any pair of points on a continuous spacetime path, one belongs to the future lightcone of the other. So, if the respective characteristic is chosen to be *"not having intersected the other process (another continuous spacetime path)"*, then after their intersection, these would have adopted the *modified* characteristic of *"having intersected the other process (another continuous spacetime path)"*. As there has been a change in this particular *characteristic* and since it looks as if all the conditions of **CI** are satisfied, Kitcher concludes that this intersection is a bona fide causal interaction.

Take a more concrete example. In his vagueness charge against Salmon's use of certain terms such as "characteristic" and "structure" in the definitions of causal notions, Dowe provides the following[38] scenario. In the morning, the shadow of the Sydney Opera House has the characteristic of *being closer to the Harbor Bridge than*

[37]Kitcher, P. (1989), p.466.
[38]Dowe, P. (1992a), p.201.

the Opera House itself. Relatively speaking, the Opera House would have the characteristic of *being further away from the Harbor Bridge than its own shadow.* As the day progresses, the Opera House and its shadow "coincide" at noon and from that point on, the respective characteristics of the Opera House and its shadow change and reverse to those of *being closer to the Harbor Bridge than its own shadow* and *being further away from the Harbor Bridge than the Opera House itself.* So, the intersection of the Opera House and its shadow at noon constitutes a causal interaction according to **CI**.

The real trouble with this class of examples is that "relational" properties have been selected as the relevant characteristics in which changes are to take place. Arguably, such relational properties do not constitute the type of characteristics one should be interested in as far as the subject of causal interactions is concerned. But might there ever be a universal standard to rule out all those characteristics which are not desired[39]?

In particular, does one need to be in possession of the knowledge of a process being causal or not prior to deciding which characteristics do qualify as "appropriate"? This could be a legitimate problem for Salmon's 1984 program because causal processes are defined by their capabilities to propagate modifications of some of their characteristics introduced by single local interactions. The changes in the characteristics of the participating processes in a causal interaction are the means by which we are to pick out the causal processes and if these can only be identified from the knowledge that the processes which sustain such changes must themselves be causal, then the account

[39]An obvious way to do so in physics is to stipulate that only dynamical properties such as *momentum* and *energy* are to be regarded as genuine characteristics that undergo changes in a causal interaction because after all, there are exchanges of these objective quantities when physical interactions take place. Moreover, the adoption of such quantities as the relevant characteristics has the added virtue of ensuring that the changes in the processes are *correlative* since they obey conservation laws. In the collision of two billiard balls, the loss in momentum or energy in one ball is accompanied by a corresponding increase in the numerical value of the same quantity in the other; with the sole demand that the overall value of that particular quantity of the whole system remains constant. Thus, a reasonable choice for a characteristic that undergoes modifications in a physically causal interaction would seem to be a *conserved quantity* possessed by both intersecting processes. This is indeed the choice endorsed by Dowe in his 1992 formulation of the Conserved Quantity Theory of Causation.

would be blatantly circular. Both Kitcher[40] and Dowe[41] have expressed such a worry with Salmon's theory. The heart of the matter is really that since causal interaction (where there are changes in certain characteristics of the participating processes) figures prominently in the description of causal processes, must only causal processes, and not pseudo ones, undergo changes in some characteristics in a *bona fide* causal interaction?

3.3.1 The Circularity Charge

Let us put all these speculations into a more concrete form and consider Kitcher's examples of pseudo- and derivative marks. Suppose a car grazes a stone wall and gets scratched. Two processes are found to be at work - the car and the stone wall. The car acquires a scratch that represents a change in a characteristic as a result of intersecting the stone wall. This intersection thus qualifies as a causal interaction in virtue of **CI**. In addition, the mark - the scratch - is transmitted beyond the point of intersection by the moving car, the process itself. Hence, the car is a causal process in accordance with **MT**. At the same time when the car grazes the stone wall, its shadow changes its characteristics of *being the shadow of a car* to that of *being the shadow of a scratched car.* Now if one looks more carefully, also at this very moment, the shadow is intersecting with another process, namely, the patch of ground upon which it is cast. It appears that as a consequence of the intersection of the shadow and the ground, the shadow takes on a new characteristic of being the shadow of a scratched car while the patch of ground becomes darkened. There are modifications of characteristics of both of these processes - the *shadow of the scratched car* and the *patch of ground.* Besides, the "modified" characteristics of the shadow is *transmitted* beyond the locus of intersection. And thus, the shadow also qualifies as a causal process by virtue of **MT**.

At this juncture, a natural defence for Salmon is that the intersection should be "local" and "single" (i.e. without further intersections); he wants to say that the

[40]Kitcher (1989), p.463.
[41]op. cit., p.200.

genuine marking interaction is the one between the car and the stone wall and not that between the shadow of the car and the patch of ground. But this is exactly where the difficulty arises! Whereas the interaction between the car and the stone wall is a single intersection *local* to the process of the car, the single intersection between the shadow and the patch of ground is, likewise, *local* to the process of the shadow.

So the defence is, *prima facie*, of no avail here. Still, there bound to exist a subtle difference between the two respective intersections so that we are justified in claiming that one is causal while the other is not.

The interaction, the mark, in the case of the car and the stone wall is of a legitimate causal nature but the one between its shadow and the ground can only assume its status as a pseudomark. The process that carries a pseudomark owes its very existence to another process and any modification of the characteristics of the former is parasitic upon the kind of interaction that happens to the latter. Just as in the case when a red lens is placed at one location on the wall of the astrodome, the moving spot is *marked* red at that point but not beyond. As the spot moves ahead it "reverts" to a state of being white again. However, had the red lens be placed at the source, the light ray would have been marked red, as would the spot.

So a counterfactual introduced to the effect that *the spot would have not have turned red given no interaction occurs locally within itself* is false; since the spot would have turned red anyway due to the red lens installed at the source. In other words, the counterfactual serves to filter out the spot as a pseudo-process as it would have changed from a state of being white to one being red *even if* it has *not* intersected another process.

The change from the characteristic of being a white spot to that of a red one in this instance is dependent upon what has happened outside the process: namely, what has happened at the source. Hence, counterfactuals serve to eliminate processes that are dependent on an "outside source" for its survival and as so, they function as effective selectors of the sort of characteristic changes that are of relevance when it comes to the determination of a "causal" intersection of processes.

Similarly, in the case of the car and its shadow, one should expect a counterfactual statement to play also the essential role of barring a case like the shadow of a scratched car from being qualified as a genuine causal process.

Such a counterfactual would read: *the shadow would not have turned into a shadow of a scratched car given that no interaction occurs locally within the shadow itself.* However, in this instance, there *is* a *local* interaction - the intersection between the shadow and the patch of ground on which it falls - within the process of the shadow itself. Granted that an interaction and a change (into a shadow of a scratched car) have occurred, the counterfactual fails to tell that the change in characteristic has, in fact, *not* arisen from this particular interaction between the shadow and the ground. And the appeal to counterfactuals for identifying the *relevant* changes in characteristic in a *genuine* interaction breaks down.

This brings us back to address the circularity charge that Salmon's theory has been accused of. The issue of the matter is such that, in the case of the shadow, be it one of the previously un-scratched or now scratched car, it owes its very existence to the interactions between the light rays, the car and the ground; that is, to the *intersections* between *causal* processes. It thus seems obvious that a causal process is indeed needed in order to pick out those changes in the relevant characteristics arising from intersections that are deemed to be genuine causal interactions. On the other hand, in order to decide which processes are causal, a causal interaction, a change in the relevant characteristic, for distinguishing them from non-causal processes is required. Since both notions - *causal interactions* and *causal processes* - are defined in terms of each other, the account is circular. The circularity charge is shown in the diagram in Figure 3-3 below.

There are two equivalent avenues one may explore in terms of disentangling the circularity. First, instead of asserting that the only relevant changes in characteristics involved in an intersection are those sustained by causal processes, we may state in an explicit manner *just what these characteristics consist in.* Only when the ambiguities are removed and a clear definition of just what are the relevant sort of characteristics obtains, then the restriction of the requirement of just *causal* processes to intersect

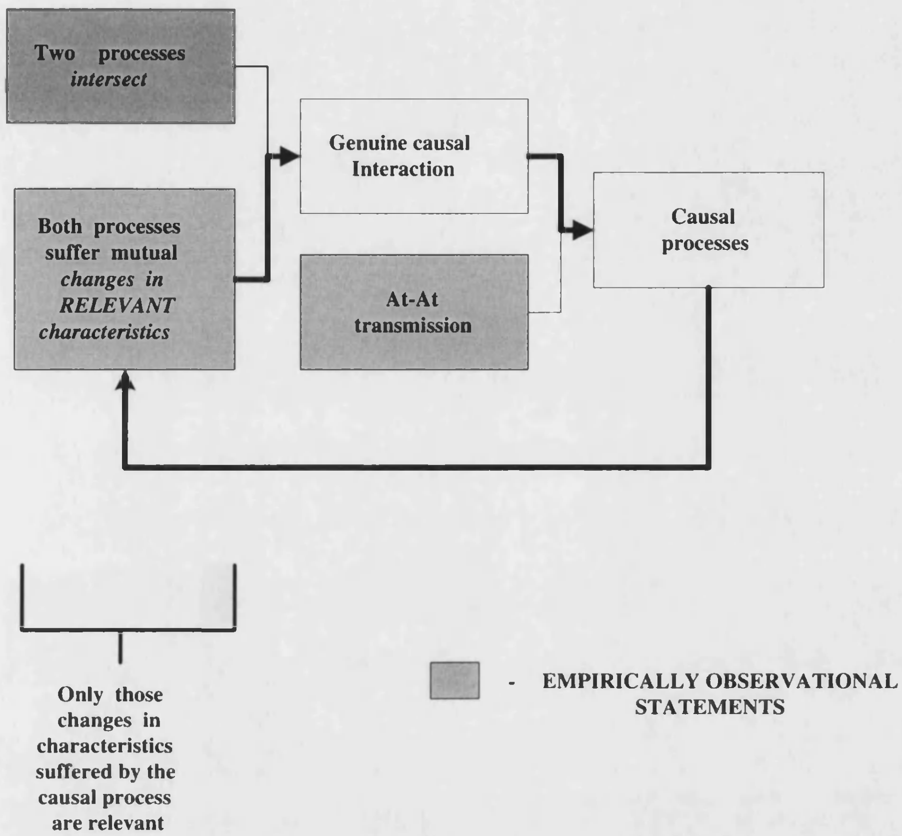

Figure 3.3: The circularity charge against Salmon 1984.

can be duly relaxed. Under such circumstances, specifications of the relevant characteristics dictate what count as genuine causal interactions. Together with the criterion of "*at-at*" transmission, they decide when a process is causal. Alternatively, one may find oneself to be more at ease with a definition of causal processes that is independent of the notion of interaction and define interaction as the intersection of causal processes. Both approaches of rescuing Salmon's theory from the charge of circularity are laid out in Figure 3-4.

One immediately comes to notice that the two schemes differ in the vital aspect of causal transmission. In fact, these represent the two approaches adopted by Salmon and Dowe in the most up-to-date versions of their respective conserved theories of causation. We shall discuss these accounts in details in Chapter 4.

In the second scheme, we see that the definition of causal processes ought to be wide enough to capture the changes of characteristics that are deemed to be relevant

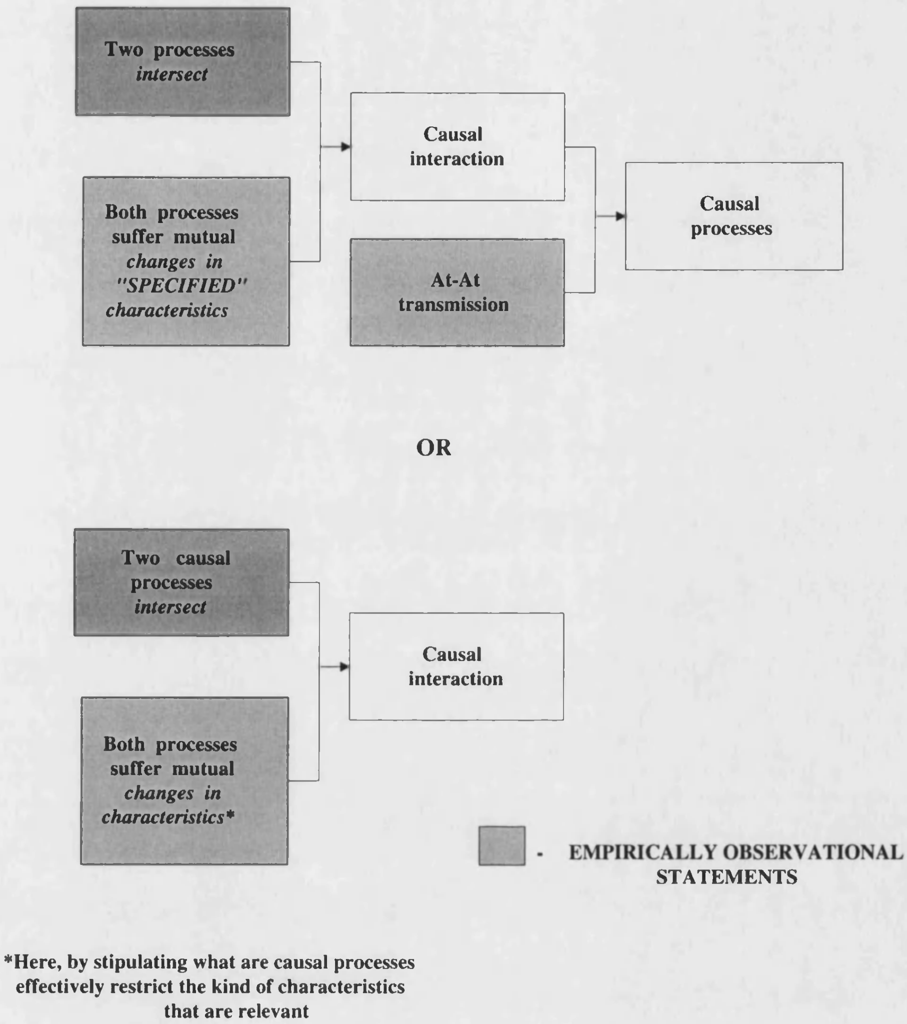

Figure 3.4: Possible approaches to the resolution of the circularity charge against Salmon 1984.

to a causal interaction. Moreover, it ought also to encapsulate the dynamics of the transmission of causal influences. We shall see in Chapter 4 that this indeed is the route that Dowe takes in his construction of the Conserved Quantity Theory.

3.3.2 The Problem of Counterfactuals

Counterfactual conditions are required by Salmon to delineate the right causal processes and causal interactions, as we have already seen in Sections 3.2.2 and 3.2.3. However, we have also learnt that counterfactuals are not always effective in telling causal processes from non-causal ones (see the example of the "shadow of a scratched car in Section 3.3.1). The problem is closely connected with the ambiguities in deciding which one should indeed be regarded as the right kind of characteristics that figure in causal interactions, which in turn find their ways into the definition of causal processes.

It is therefore appropriate to consider in some more details the use of counterfactuals in Salmon's 1984 theory since they are intimately bound up with the distinction between causal and pseudo-processes. Following Kitcher[42], we consider the example of a baseball hitting and shattering a window. The two events of *the bat hitting the ball* and *the ball shattering the window* are connected by the causal process of the baseball in flight. Suppose we focus on the window alone and call this causal process P. Also let P_t denotes the point at which the window shatters on being hit by the ball and P_t' be some earlier point at which its temperature is changed through the impact of a sudden gust of wind. Consider a segment of the causal process P that links P_t and P_t'. There are interactions at both ends of this segment: the interaction with the gust of wind P_t' and the subsequent one with the baseball at P_t. However, it is obvious that the existence of this particular structure of interactions and processes is not crucial to the causal history of the event of window shattering. As Kitcher maintains,

> ***To specify the conditions under which*** c ***and*** e ***are causally related, we need to build into the account the idea that the initial interaction produces the modification that is responsible for the characteristics of the terminal interaction. Intuitively, what is lacking is the kind of articulated structure that I envisaged in the building up of complex processes out of simple ones.***

Why do we pick the interaction of the baseball rather than the temperature change resulting from the interaction with the gust of wind as the relevant event that contributes to the window shattering? There is a good physical reason for it: the momentum of the baseball provides just the right impact force to break the window

[42]op. cit., p.470.

while the temperature differential is not substantial enough for the same purpose.

Kitcher thus argues that in order to keep a good record of the relevant causal history leading up to the event of the window shattering, counterfactuals cannot be avoided in order to stipulate and keep track of the momentum exchanges within this particular history. He claims that the causal explanation should look something like the following,

(A) If the bat had not intersected P_1 (the ball) then the momentum of P_1 would have been different.

(B) If the momentum of P_1 after its intersection with the bat had been different then the momentum of P_1 just prior to its intersection with P_2 (the window) would have been different.

(C) If the momentum of P_1 just prior to its intersection with P_2 had been different then the properties of P_2 just after the intersection would have been different.

Kitcher aims to downplay the significance of causal processes and causal interactions in favour of counterfactuals[43],

> ***The crucial point is that our claim of a causal relation between** c **and** e **depends not simply on the existence of the interactions and the processes but on our acceptance of the counterfactuals** (A)-(C) **(or of related weaker versions). We have to invoke counterfactual notions not only in characterizing the concepts of causal processes and causal interaction but also in singling out the causal processes and causal interactions that are relevant to particular events.***

But do we really need counterfactuals in so emphatical a manner as Kitcher has suggested, to specify the momentum in every stage of the causal structure? Probably not, if we hold to the principle of the conservation of momentum. The principle conveniently keeps track of the various quantities of momentum involved. Conceived as a law of nature, the principle would support counterfactual claims. So the appeal to precise statements of physical laws eliminates the need for counterfactuals.

[43]ibid.

This much being said, there is, however, a finer-grained worry when one places this rather idealized situation in a more realistic perspective. In practice, the baseball in flight intersects with numerous air molecules (on a most basic approximation) that get into its way and each of these molecules is potentially capable of imparting some of its momentum to the ball. It is conceivable that the momenta received by the baseball from all its surrounding air molecules may have a resultant effect on it so that it may be deflected from its original course of motion, or impinge on the window with a momentum which is not sufficient to break it.

There is an important moral to this scenario. The demand of no further subsequent interactions in **MT** is unrealistic. In reality, the continuous interruptions from the environment make causal transmission a too idealized concept to gain full credibility on a practical level. Even though there is a genuine causal interaction in which a causal process is marked, the mark would simply fail to be transmitted further due to other immediate interfering interactions with the environment. This in turn calls into serious doubt the deployment of the criterion of transmission as a basis of distinction between causal and pseudo-processes. Kitcher seems to think that counterfactuals are again essential to deal with continuous interactions as he writes[44],

> ***...the characteristics of a macroscopic object are (almost always) sustained by further interactions. Furthermore, the characteristics of such processes are often modified by subsequent interactions, so that, when a process is "marked" (in the ordinary sense) it is likely that the mark will be altered by a later interaction to produce a later mark.***

And again[45],

> ***We have here a sequence of causal processes and interactions, and if, we say that the final organism is marked by the initial interaction, that is because we envisage a sequence of marks***

[44]ibid., p.468.
[45]ibid., p.469.

such that each is transmitted by a causal process that interacts with another process to produce a successor mark. Our attribution is based on our acceptance of a chain of counterfactuals...

If P_n had not transmitted M_n then P_n+1 would not have acquired $M_n + 1$.

In building up a complex causal process out of elementary causal processes - that is, processes that do not interact with other processes - we need to make heavy use of counterfactuals. The sequence $P_1 ... P_n$ constitutes a complex causal process only if each P_r interacts with $P_r + 1$ so that if P_r had not been modified to bear Q_r then $P_r + 1$ would not have been modified to bear $Q_r + 1$.

However, the perennial problem haunting the enterprise of counterfactuals has been that of the determination of when a counterfactual is true, and Salmon has proposed a method for justifying one's beliefs in counterfactuals by means of controlled experiments. He suggests that science has a sound way of determining the truth and falsehood of these contrary-to-fact conditionals: by appealing to the very idea of controlled experiments so that some counterfactual statements may be determined as true and others as false.

An example of the kind of experiment offered by Salmon is as follows. Suppose in the case of the astrodome, the source is to be switched on and off one hundred times by one experimenter while a second experimenter is stationed half-way between the source and the wall whose sole task is to select at random, with the help of a random number generator he is equipped with, fifty trials in which he will mark the light ray by intercepting it with a red filter. Here, Salmon argues that if not all but only fifty trials in which the marks are made are those where the spot on the wall is red as well as the intervening stages in the beam between the point of the marking interaction and the wall, then one is justified to claim that in those fifty cases, the counterfactual "*the beam would not have turned red if the marking* *operation had not occurred*" is *true*. Moreover, Salmon has taken care in formulating the example in such a way that the introduction of the random number generator serves to maintain the objectiveness of the situation by making sure that the selection procedure is not biased towards the experimenter's subjective judgment. But are we really justified to fall back on experimental procedures to establish the truth and falsehood of counterfactual statements?

Several authors[46] have voiced their objections to this issue, which are mostly concerned with how objectivity can be obtained if the truth values of counterfactuals may only be determined in an *inductive* manner by experimental methods. It is argued that one may just be depending too much on empirical generalizations and since such generalizations remain agnostic about the existence of necessary connections as supposed by counterfactuals; experiments may therefore not be a good way for determining the truth values of counterfactual statements. In particular, Dowe has argued that the appeal to experimentation provides only epistemic but not ontic grounds for the truth values of counterfactuals.

Problems arise as to whether knowledge of the truth or falsity of the counterfactual is epistemologically accessible to us or not. Established semantical accounts of counterfactuals, such as those advanced by Stalnaker[47] and Lewis[48], make use of possible worlds and similarity relations. Such approaches are, however, of no avail to the empiricist, for the knowledge of the happenings in our world is rather useless in trying to determine what goes on in the possible worlds; unless, of course, in this instance the possible worlds serve merely as a summary of complicated facts about the real world. In practice, however, it is argued that if the various features of a possible world are *made* to be very similar to that in the actual world, one is then justified in applying our knowledge of the actual world to determine the "near-enough" truth-values of that particular possible world. This view supports the rationale behind Salmon's design of controlled experiments to test the claim that a causal interaction

[46]See for example, Kitcher, P. (1989), Dowe, P. (1992a), Fetzer, J.H. (1987) and Giere, R.N. (1988).

[47]Stalnaker, R. (1968).

[48]Lewis, D. (1973).

has obtained.

Imagine two billiard balls rolling across a table with a transparent surface and colliding. The table is illuminated from above and so that their respective shadows also intersect. Let the shadow of ball A be P_A and that of ball B be P_B. The following pair of counterfactuals seems to be true:

(i) If the worldline of ball A had not intersected that of ball B, then the direction of motion of ball B would not have altered.

(ii) If the worldline of P_A had not intersected that of P_B, then the direction of motion P_B would not have altered.

Now we want to claim that whereas the collision between balls A and B causes a change in the direction of motion of B, the intersection of the shadow of A with that of ball B does *not* cause the change in the direction of the motion of the shadow of B. That is, we want to prove that the counterfactual in (ii) is wrong and that the direction of motion of P_B would have altered even if it had not intersected P_A. Turning this around, we must set out to look for cases where P_A is found not to have intersected P_B and observe, in those instances, what happens to the direction of motion of P_B. So, we would have to keep everything as before but to invent a way to "erase" P_A. It turns out that this may readily be achieved by having part of the surface of the floor illuminated from below so that P_A (the shadow of ball A) is duly obliterated. Observation shows that the direction of motion of P_B alters even in the absence of P_A. This, argues Salmon, provides solid ground for refuting the counterfactual in (ii).

The introduction of the extra light source from below the table seems the intuitively correct experiment to perform to check the validity of (ii). Kitcher suggests that although Salmon's proposed experiment accords with our natural ideas about the causal character of this particular situation, we should be, however, looking for the *basis* of these ideas and so a separate theory is required to help us to select the *correct* kind of controlled experiments. In this vein, Kitcher asserts that the worldline P_A could alternatively have been erased by the removal of ball A out of the picture altogether. The presence of ball A is to be simulated instead by an appropriate impulse at the very moment when it would have been collided with ball B. Under this circumstance, the counterfactual in (ii) is also found to be refuted and so our aim is, again, satisfied.

In Salmon's experiment, there *is* an intersection of the paths of balls A and B, whereas in Kitcher's one there is not. Both of these "actual occurrences" have given rise to the same result so that the counterfactual in clause (ii) above is refuted. However, one immediately comes to realise that the use of an impulse in place of ball A in Kitcher's example has made the counterfactual in (i) false as well. Since now that both counterfactuals are found to be false, they are no longer capable of distinguishing between causal interactions and non-causal ones. Therefore, if our sole aim is to provide a refutation of the counterfactual in (ii) - to eliminate the intersection there described as a pseudo-interaction - then actual interactions of very different natures would do. But if we want to probe deeper, to trace the origin of *the actual* interaction which led to the situation in (ii), the counterfactuals fail to be of any assistance.

Now Kitcher's point is really that if one appeals to the similarities and differences between the experimental group and the controlled group, one must first need an account of how the trade-offs between the similarities and differences are to be made. At this juncture, we probably want to say that Salmon's experiment is still the correct one, because according to our intuition, we want to claim that the interaction between ball A and ball B is, in fact, the relevant causal interaction. But here again, from where comes our intuition? It comes from none other than the background causal knowledge we have already possessed. In Salmon's experiment, we know of the presence of both balls A and B and we are aware of the fact that their paths in spacetime have intersected. Such knowledge is, however, *not* to be obtained from the counterfactuals.

Once again, this drives us back to the very problem of how one *first* comes to *know* which one - "the intersection of balls A and B" *or* "that of ball B and the impulse" - should be deemed the correct genuine interaction. However, Kitcher is quick to note that Salmon's experiment is one which would be endorsed by practicing scientists, who design their controlled experiments by drawing heavily on their background causal knowledge. Kitcher therefore concludes [49],

> ***They endeavour to ensure that the control group and the experimental group are similar in those aspects that they take to be potentially causally relevant.***

And, continues Kitcher [50],

> ***Once we have some causal knowledge (perhaps a significant amount) then that causal knowledge can be used in the design of control experiments that will test counterfactuals in just the way that Salmon proposes. But if we are looking for a theory of how we justify counterfactuals from scratch, then the appeal to the method of controlled experiments is of no avail.***

So it would seem that Salmon's intended use of controlled experiments to study the truths of counterfactuals appears not to be entirely satisfactory and poses problems for his theory of causation, as it draws heavily on counterfactuals to tell the causal situations from non-causal ones. It has been suggested earlier that the need for counterfactuals arises from the vagueness in the reference to the *characteristics* of the physical system that are to undergo changes upon interactions, which also leads to a circularity in Salmon's theory. In order to be discharged from circularity, one may either specify the characteristics which are to undergo modifications in situations that are deemed causal *or* opt for a definition of causal processes that is quite independent of the notion of interaction. We shall see in Chapter 4 just how the issue of counterfactuals may be evaded by taking these routes.

[49]ibid., p.475
[50]ibid.

3.4 Discussions and Summary

In *Scientific Explanation and Causal Structure of The World*, Salmon has advanced a new model of scientific explanation that proves to be a marked improvement from his earlier program of statistical relevance (the *Statistical Relevance* or *S-R* model), which concerns the subsumption of facts under general laws for the purpose of explanation. In this new program he pays due attention to the explanation of particular facts and acknowledges the prime role that causality play in explanations. Now for Salmon, explanation is a two-tiered affair. The first level consists in the subsuming of the event-to-be-explained under a set of statistical relevance relations as in the older S-R model, but these statistical relevance relations are now to be explained in terms of the causal relations on the second level. In this chapter, we have focused on Salmon's process theory of causation which has been specially devised to provide the causal relations in support of this model of explanation.

I find Salmon's proposal attractive because it represents a serious attempt to overcome the issue of temporal contiguity between cause and effect as raised by Russell in his 1913 analysis. Salmon, with his process theory, has furnished us with spatiotemporally continuous physical processes as the causal connection between two events; and these processes serve to eliminate the temporal gap that stands opposing to temporal contiguity. It also provides a prescription for defining localized events in spacetime by reference to the interaction of causal processes. The interactions of causal processes fix the locations of causes and effects in space and time. Spatio-temporal continuous processes seem to have rescued the Humean temporal contiguity thesis from Russell's philosophical onslaught.

By appealing to the actual physical process involved in specific situations, Salmon's theory is highly local and singular. We have introduced in this chapter the various building blocks of the theory such as *process*, *the Mark Transmission* (MT) *criterion for distinguishing causal and non-causal processes* and the mechanism of *causal interaction.* In spite of the solid intuitions behind these concepts, there exist pressing philosophical issues as regard to the use of counterfactuals that beset the program.

In the Chapter 4, we shall discuss in details the subsequent developments in process causation; in particular, the Conserved Quantity Theory (**CQ**) by Phil Dowe, whose chief aim is to resolve the problem of counterfactuals.

As it stands therefore, this theory is not to be seen as complete but may only be regarded as providing a general framework for an exciting exploration into the various aspects of causality from the perspective of causal processes. The good intuitions behind Salmon's theory must remain significant for other variations of the process causal theory.

Chapter 4

Process Theories: Subsequent Developments

4.1 Introduction

Wesley Salmon's 1984 version of the process theory of causation was expounded in some detail and its relative merits and drawbacks were discussed in the previous chapter. Salmon's theory focuses on the singular aspect of physical causation by trading very heavily on the notion of spatio-temporal continuous processes. Changes in such processes occur through their physical interactions with other processes, resulting in subsequent modifications in their respective process structures. These changes, introduced through localised physical interactions, *mark* events in spacetime while causal processes provide the physical causal connection amongst these events. Recall that the major difficulty that underlies Russell's *"temporal gap"* is the superposition of events happening at discrete points in time against a background of temporal continuum. It thus seems that a promising attempt to overcome this difficulty is to superpose continuous entities on the temporal continuum. And so the appeal to the spatio-temporal continuous characteristics of causal processes might just be the solution we have been seeking to bring a closure to Russell's problematic temporal gap.

This seems plain sailing were it not for the fact that physical considerations demand Salmon to make a distinction between causal and pseudo-processes. The criterion of mark transmission is employed to distinguish causal regularities as exhibited in causal processes from the species of a sequence of "unlinked" regularities in pseudo-processes that must be sustained by an external source. A moving spot of light across the circular wall of the astrodome consists really of a sequence of spots cast by light pulses travelling from the rotating source at the center. The actual act of transmission of causal influences (or marks) by causal processes provides the "link" to enable these regularities to be deemed causal. This is the crux of Salmon's account: *not all processes of a continuous spatio-temporal character are causal; only those that are able to transmit causal influences can at the same time by this very means provide the vital links within the processes themselves.* The transmission of causal influences is of paramount importance to a process-singularist theory of causation for they ensure that all "parts" of a causal process are linked rather than it being a random conjunction of a sequence of events.

The immediate consequence of elevating the concept of transmission to the status of a fundamental causal notion is the need for defining what indeed constitutes transmission. The "*at-at*" theory, the original recipe of Russell to overcome the Zeno paradox of the flying arrow, is exploited by Salmon to provide the *mechanism* for causal transmission. The "*at-at*" notion of causal transmission, which relies entirely upon physical laws that govern the motion of a system, poses insuperable difficulties as one considers the quantum regime where the notion of spatio-temporal continuity becomes untenable. But of course, the most devastating charge to Salmon's criterion of mark transmission (MT) is the indisputable element of counterfactualness that is seen to bring damage to the objectivity intended by Salmon for this causal notion. We have argued that the counterfactual condition has been introduced into the definition of MT in order to make clear that the pseudo-processes would have been modified *even if* the interaction that is responsible for the modification resides outside the process itself. In this way, counterfactuals are essential for the distinction between causal and pseudo-processes.

However, as we have seen[1], counterfactuals are not always effective in delineating

[1]Take for instance, the counterfactual concerned is unable to dismiss the characteristic change of the shadow of a car - from one being "the shadow of an un-scratched car" to that of being "the shadow of a scratched car", due to the "local interaction" between the the shadow and the

causal processes from their non-causal counterparts, which is due to the ambiguities involved in the determination of the relevant characteristics that changes in causal interactions, and causal interactions play a major role in defining causal processes. It is thus desirable to remove counterfactuals from the theory.

It is indeed this trouble with counterfactuals that Phil Dowe sought to resolve by his *Conserved Quantity Theory* (**CQ**) in 1992[2]. **CQ** has succeeded in large measure in the simplification of Salmon's account by the definitive identification of a *causal process* with the *worldline of an object.* Because of this explicit characterization, the need to decide which are the relevant characteristics for separating the class of causal from that of non-causal processes (that has led to the use of counterfactuals in the first place) receives immediate alleviation[3]. Moreover, it is Dowe's intention for the notion of causal transmission be now incorporated in a physically natural way without the need for a separate treatment. However, in the picture that Dowe offers, changes in the causal structure of a physical process still come about as a consequence of interactions with other processes. We shall examine **CQ** in detail in Section 4.2.

Prompted by his own eagerness to eliminate the counterfactual element in the theory, Salmon welcomed the birth of **CQ** with zeal and enthusiasm. In his ensuing paper in 1994, "Causality Without Counterfactuals"[4], Salmon generously acknowledges Dowe's achievement by an *almost* complete endorsement of the worldline view of causal process and by the abandonment of the comparatively vague notion of "marks" in favour of *invariant quantities*, a somewhat modest adjustment from the use of conserved quantities in **CQ**. The adoption of invariant quantities descended from Salmon's desire to maintain causality as an objective notion. An objective notion is thought to be one that takes on the same meaning relative to all points of view. Invariant quantities are the ones picked out by the theory of special relativity to remain frame-independent, in that they bear the *same* value to every observer

patch of ground on which it falls - as an illegitimate interaction that provides the mark for alleged transmission by the shadow (Section 3.3.1).

[2]Dowe, P. (1992a) and (1992b).

[3]This is in the strict sense that we have now an empirical criterion for identifying a causal process, but we make no claim at this stage that this is a correct criterion.

[4]Salmon, W.C. (1994).

regardless of their state of motion. And so having the same value relative to all observers makes invariant quantities a suitable representation of objectivity in a physical context. This results in the ***Invariant Theory of Causation*** (**IQ**). Apart from this rather superficial modification, Salmon nevertheless sees the need to retain the notion of transmission within **IQ** despite his willingness to give up the concept of marks. The idea of causal transmission thus continues to play a central role in **IQ**.

Most recently in 1997[5], Salmon was convinced to give up his insistence on invariant quantities in favour of conserved quantities, chiefly in response to the efforts of the arguments put forth by Dowe[6] and Hitchcock[7] two years earlier in 1995. In spite of these periods of metamorphoses, Salmon has remained a dedicated disciple of the notion of causal transmission. He supports a ***Conserved Quantity theory of Causation with Transmission*** (**CQT**) in the latest episode on the philosophical scene of process causation. This latest version of Salmon's theory will form the subject of Section 4.3. There, we endeavour to understand Salmon's main argument for his seemingly obstinate hold on the notion of causal transmission. The reasonableness of the retainment of this notion within the new framework will also be assessed.

In view of the fact that the notion of a *worldline* plays such a prominent role in **CQ**, **IQ** and **CQT** - since all three take causal processes to be worldlines of objects that carry certain physical quantities - a scrutiny of the notion will serve the purpose of revealing the subtleties involved in these accounts. In Section 4.4, I provide a critique of the *Worldline View* (**WV**) of causal processes and discuss the relevance of spatiotemporal continuity to the notion of identity over time, which as I shall argue, lies at the core of both the theories of Salmon and Dowe. I shall also point out a discrepancy between the usage of the term "worldline" in both **CQ** and **CQT** and that in physics, which leads to a confusion. To remedy the situation and be in anticipation for an extension of the notion of a causal process to the quantum domain - where there is the loss of spatiotemporal continuity - a *History View* of causal processes is proposed in Section 4.5. It is indeed this ***History Conserved Quantity theory***

[5]Salmon W.C. (1997).
[6]Dowe, P. (1995).
[7]Hitchcock, C.R. (1995).

with Transmission (HCQT) that we take as a basis for extending the notion of causal processes to the quantum realm.

4.2 The Conserved Quantity Theory

Since Salmon's seminal 1984 programme on process causation, not much has taken place in this philosophical sphere until the publications of a series of papers by Phil Dowe in 1992 that aim chiefly to exorcise the specter of counterfactuals from Salmon's account. This was a crucial turning point for process causal theories and with this came the birth of the ***Conserved Quantity Theory*** (CQ).

Dowe makes it the main task of CQ to deliver Salmon's process causality from the evil of the epistemological difficulties that gave rise to the need for counterfactuals in the first instance. It is these difficulties that lay Salmon's programme open to the charge of circularity. The related arguments have been well rehearsed in Section 3.3.1 and it is not my intention to duplicate them here. However, the reader ought to be reminded that it is indicated in that earlier section that there exist two natural ways of relieving Salmon's theory from circularity and the use of counterfactuals. These are: (1) a definitive stipulation of what constitutes a causal process should be given independent of the notion of causal interaction, or, (2) the specification of which are indeed the characteristics relevant to a causal interaction (and causal processes are defined in terms of it). These two approaches are summarized diagrammatically in Figure 3-4.

Dowe opts for the strategy in (1). With this approach in mind, he introduces two definitions which comprise CQ[8],

(1) ***A causal interaction is an intersection of world lines which involves exchange of a conserved quantity.***

(2) ***A causal process is a world line of an object which manifests a conserved quantity.***

[8] Dowe, P. (1992a), p.210.

This "worldline view" (**WV**) of causal process is already embryonic[9] in Salmon's 1984 treatment where he indicates what constitutes, in his view, a process[10], *"...while processes have much greater temporal duration, and in many cases, much greater spatial extent. In space-time diagrams, events are represented by points, while processes are represented by lines"*. We can also find signs of the idea in Kitcher's 1989 critique, in the section where he argues that counterfactuals are more fundamental in the discussions of causation than causal processes and causal interactions[11]; and in there, he makes specific mention to the "worldlines" of objects. However, it is to Dowe and to Dowe alone that we owe the credit for taking the decisive step forward to make the definitive identification of a causal process with the worldline of an object.

4.2.1 Causal Processes

Let us proceed with definition (2) above. Dowe explains[12], *"A world line is the collection of points on a spacetime (Minkowski) diagram which represents the history of an object."* And a conserved quantity is[13], *"any quantity universally conserved according to current scientific theories. Some conserved quantities are mass-energy, linear momentum, angular momentum, and charge."*.

In his definition of a worldline, Dowe has tactfully avoided making reference to a "continuous sequence" of points. The deliberate expression of "the collection of points" is intended to be applicable to cases where spatiotemporal continuity does not obtain, with the discontinuous processes in the quantum domain as the candidate examples[14].

Dowe allows a wide sense of what may count as an object[15], *"An object can be anything found in the ontology of science (such as particles, waves or fields), or common sense."*, and also[16], *"The precise characterization of "object" is unimportant; what*

[9]In fact, Dowe acknowledges that he has adopted the idea from Brian Skyrms in his discussions on related matters in Skyrms, B. (1980), p.111.
[10]Salmon, W.C. (1984), p.139.
[11]Kitcher, P. (1989), p.469.
[12]ibid., p.210.
[13]ibid.
[14]Dowe, P. (1995), p.332.
[15]ibid. and Dowe, P. (2000), p.91.
[16]Dowe, P. (1992b), p.126.

matters is whether it manifests[17] *a conserved quantity.*" And again[18], "*What counts as an object is unimportant; any old gerrymandered thing qualifies. What matters is whether it may possess the right type of quantity.*" Being liberal and unrestrictive in his definition of an object, Dowe admits the worldline of something like a shadow as being, as presumably, "an object of common sense".

For him, objects[19] like shadows are allowed to have worldlines although their "worldlines" are not causal processes[20], "*A process is the world line of an object, regardless of whether or not that object possesses conserved quantities.*"

Dowe distinguishes the "worldlines" of processes like shadows from the "worldlines" of those with spatiotemporal stages that are genuinely "linked" dynamically, by stipulating that the latter - the genuine causal processes - are constituted out of events involving objects which *possess* conserved quantities[21], "*For example, a shadow is an object but it does not possess energy or momentum; it has shape but possesses no conserved quantities.*"[22].

It is perhaps not hard to see why Dowe wishes to place no restriction on the notion of an object. Empirically, given the observation that the car and its shadow both exhibit spatiotemporally continuous motion; how are we to determine empirically, as

[17]In his 1995 paper, Dowe has switched to the use of "possesses" in place of "manifests" according to a suggestion by Armstrong for fear that the latter may lead to the mistaken impression that the quantity has to be experienced by human observers (Dowe, P. (1995), p.324).

[18]Dowe, P. (1995), p.323.

[19]Contrary to the usual meaning intended for an object in the (physics) definition of a worldline that takes the object's identity over time as a primitive (as we shall argue in Section 4.4), Dowe does not make the same supposition at the outset when **CQ** was first introduced in 1992. All that is demanded of an object is its successive stages must possess a conserved quantity in order to qualify as a causal process.

[20]Dowe, P. (2000), p.90

[21]Dowe, P. (1995), p.324.

[22]From a physical perspective, a shadow is *not* in possession of physical quantities like momentum or energy. It is therefore impossible for these quantities to be transmitted by a shadow. In circumstances where it is perceived to be moving, the shadow of a moving car for example, each "stage" of the "moving" shadow is the result of an interaction between a different patch of the wall and the light that is blocked by the car. A series of static interactions provides the visual impression of a moving object - the shadow of the moving car. For this reason, something like a shadow which is not in possession of energy or momentum is not normally assigned a worldline if the meaning in physics is strictly adhered to. Incidentally, in some physics discussions authors do unfortunately refer to "spacelike worldlines", but those are only brought in to highlight the fact that they are *not possible* for there exists no genuine motion that involves transmission of energy or momentum (see for example Kopczynski, W. and Trautman, A., (1992), p.57-58.).

a matter of fact, which one is a genuine moving object (one with its successive stages linked dynamically)?

The old question of being *causally connected* versus that of *causal connectibility* re-emerges at this juncture. Causal connectibility obtains for any two events that lie within the past or future lightcones[23] of each other. Causal influences in the forms of energy, momentum or signals and so on may be transmitted from one event to another at a velocity less than or equal to the speed of light. However, being potentially connectible in this sense does not necessarily imply that there is an actual causal connection. Whether two particular events are causally connected or not depends on whether there is an actual transmission of causal influences between them.

Although the events that make up the moving shadow are spatiotemporally close to one another, as they lie in their respective past and future lightcones, there exists no actual transmission of causal influences. If the transmission of causal influences is made possible by the motion of the *one and the same* object that carries them, then in the absence of an act of transmission, there is no movement of the *one and the same* object. So there is no genuine movement of *one and the same* shadow. Why is there no genuine motion in this case? The apparent motion we perceive is only a false imagery of many individual and separate events (each corresponds to a different shadow) with no physical association at all, except for just the right spatiotemporal relations that they bear to each other that trick the naked eye[24]. Of course, the absence of transmission in this case is attributed *either* to (1) the fact that there is *no* causal influence - no conserved quantity as *possessed* by the shadow - to be transmitted, *or* (2) that the influences, the conserved quantities are not transmitted, despite the fact that they appear at different spacetime points. And Dowe's reply is that the "moving" shadow does not *possess* any conserved quantity.

Hence, the idea of ***the possession by an object of a conserved quantity*** is the **CQ** key to the identification of causal processes. The immediate philosophical

[23]Such two events are said to be "time-like" separated.

[24]One "stage" of the shadow has to cease its existence before the next one comes into being in order to give the visual impression that the one and the same object has vacated its previous location and takes up a new position at a later time.

concern is whether the stipulation of an object *possessing* a conserved quantity would serve the purpose of picking out *only* causal processes. So far, this seems to manage to distill out pseudo-processes like shadows[25]. However, we need to satisfy ourselves that the *mere possession of a conserved quantity* is indeed a sufficient condition to embody the fact that there is movement of causal influences so resulting in a causal process having one stage being the source of the next. We shall now subject this criterion to two kinds of tests.

Time-wise gerrymanders

A time-wise gerrymander is an "alleged" object that refers to different objects at different times. Consider Dowe's example[26] where the cue ball in a billiard game is struck at time t_1, which collides with the red ball at time t_2 and in turn hits the black ball at t_3. It then subsequently finds it way into the pocket at t_4. We are presented with a picture where there is transmission of energy and momentum along a spatiotemporally continuous sequence of events. Does this sequence of events constitute the successive stages of a persisting object? In particular, might one not conceive of an object x that is defined as,

x is the cue ball at $t_1 \leq t_2$,
x is the red ball at $t_2 \leq t_3$,
x is the black ball at $t_3 \leq t_4$?

One immediately notices that the condition of possession of a conserved quantity is *insufficient* to rule out this time-wise gerrymander: each of the three balls that comprised the alleged object x *does* possess energy and momentum. Troubled by this deficiency, Dowe admits that an additional constraint has to be incorporated, which amounts to a restriction on the notion of an object[27]: anything that counts as an object must satisfy the condition of being *wholly present at each time* with no

[25]As Dowe claims that shadows do not possess conserved quantities. See Dowe, P. (1992b), p.127.
[26]Dowe, P. (1995), p.323.
[27]In his more recent writings, Dowe P. (1995), p.329 and (1999), p.250.

temporal parts[28],

> ***Strict Identity (SI): an object is wholly present at each and every instant in its history.***

The "object" as referred to above obviously consists of temporal parts in the sense that a different part of the "object" exists at each different time. In other words, the "whole" object is *not* wholly present at each time (for example, at each time during the interval t_1 and t_2 only the cue ball *part* of the object is wholly present but neither is the red ball part nor the black ball part) and it is more appropriately viewed as comprising of three physical objects. The time-wise gerrymander is clearly depicted as consisting of three individual physical systems, each having its own worldline. For instance, the cue ball wholly exists at every moment between the times t_1 and t_2. And so each of these physical objects - the cue ball, the red ball or the black ball - is wholly present in each time with no temporal parts and so carry with it an unambiguous meaning of *identity over time.*

Dowe thus concludes that the condition of the *mere possession of a conserved quantity* needs to be supplemented by some thesis of identity over time, in order that it may effectively distinguish causal processes[29], "*I am not saying CQ theory needs SI. But it needs* <u>*some*</u> *thesis of identity...*".

Most recently, Dowe has taken a firm stand to impose the criterion of identity over time to restrict the scope of what is to be counted as an object in his discourse[30], "*An* <u>*object*</u> *is anything found in the ontology of science (such as particles, waves and fields), or common sense (such as chairs, buildings and people). This will include non-causal objects such spots and shadows. A process is the object's trajectory through time. That a process is the world line of an object presumes that the various time slices of the process each represent the same object, at different times, thus it is required that the object have identity over time...On the present account a*

[28]Dowe, P. (1999), p.250. **SI** carries the implication that the various time slices of the process each representing one and the same object at different times, and so the object has an identity over time.

[29]Dowe, P. (1999), p.250.

[30]Dowe, P. (2000), p.91

timewise gerrymander is not a process, for it is not the world line of an object, since objects must exhibit identity over time.". We shall examine this issue of identity over time in more details in Section 4.4, but for now we work with **SI**.

To be sure, being wholly present at each time[31] is only one out of the many characteristics of the category of objects. Being wholly present at each time, though a necessary feature for an object, is nonetheless *not* adequate to characterize a persisting object. It conveys no information on how the successive stages of a persisting object are related. Stages of objects that display qualitative similarity and spatiotemporal continuity would satisfy the condition of being wholly present at each time and can be mistaken for being the successive stages of one single persisting object. Typical examples include objects like shadows or a spot of light, the subjects of the second test to which we shall now turn.

Shadows and Spots of Light

The shadow of a moving car cast on a wall is composed of a series of interactions arising from the blocking by the car of the light that is supposed to fall onto the wall. The "moving" shadow is considered as wholly present at each time but yet every one of these "events" or "stages", as generated by a separate interaction, is an essentially different object. Entities like shadows, which satisfy the constraint of "being wholly present at each time" is to be ruled out by the criterion of the possession of a conserved quantity as an object that forms the proper subject of a causal process by virtue of **CQ**.

This is true also of the case of the moving spot of light across the wall of the astrodome. Recall that the spatiotemporal trajectory of this "moving" spot of light is everywhere continuous and hence entertains the visual impression that it is one and the same object that moves. How can one tell the motion of this object is after all an illusion? In particular, how is one to know that it is in fact not the "same" object that moves but instead an image projected from a static sequence of

[31]**SI** represents only one way of analyzing identity over time and we shall have th occasion in section 4.4 to examine the issue.

similar events? Dowe's answer would be that this object, the "moving" spot of light, is not in *possession* of a conserved quantity. Rather, it is the patch of wall which acquires the energy from the light ray but this patch of wall does not move[32], "*The object which moves, the spot, doesn't possess energy, so its worldline is not a causal process.*" Conserved quantities are not *possessed* by the "moving" spot, it is the series of segments of the wall that possess energy. Processes - the worldlines of objects - are causal by virtue of the fact that a conserved quantity is *possessed* by it. On Dowe's view, a series of world-points belonging to different physical objects are "unlinked" in a physical sense that a conserved quantity is not possessed by them.

What concerns us is whether there is much point in arguing that it is the wall and not the spot, which is in possession of energy. Whenever a patch of wall is illuminated by a light ray, there is a spot; the patch of illuminated wall *is* the spot. So there seems to be no real benefit to make a distinction between the two but as a mere linguistic maneuver. However, what is of real issue here is that even if we deem the spot of light to possess energy, there is no real action of "*carrying along*" of this energy throughout all stages of the moving spot of light, for such a "pseudo-object" is the false image of a sequence of world-points corresponding to a sequence of independent interactions involving a *different* set of objects in each: a different patch of wall and a different light ray. The objects that take part in each of these interactions would not "go on" to participate in the next interaction and hence there is simply no physical connection to provide the transmission of a conserved quantity. In other words, the "spots" (more than one object) do not move! It seems therefore, from this physical view, there is a compelling reason for driving us back to the notion of transmission, which is a fundamental causal notion in Salmon 1984. This granted, the notion of "possession" that Dowe has appealed to has proved an *insufficient* condition for identifying causal processes, for it has not succeeded in capturing the essence that one stage of the process forms the source of the next.

Let us now summarize the basic structure of Dowe's use of the criterion of "*possession of a conserved quantity*" to identify causal processes (Figure 4-1). The worldline

[32]Dowe, P. (1999), p.251.

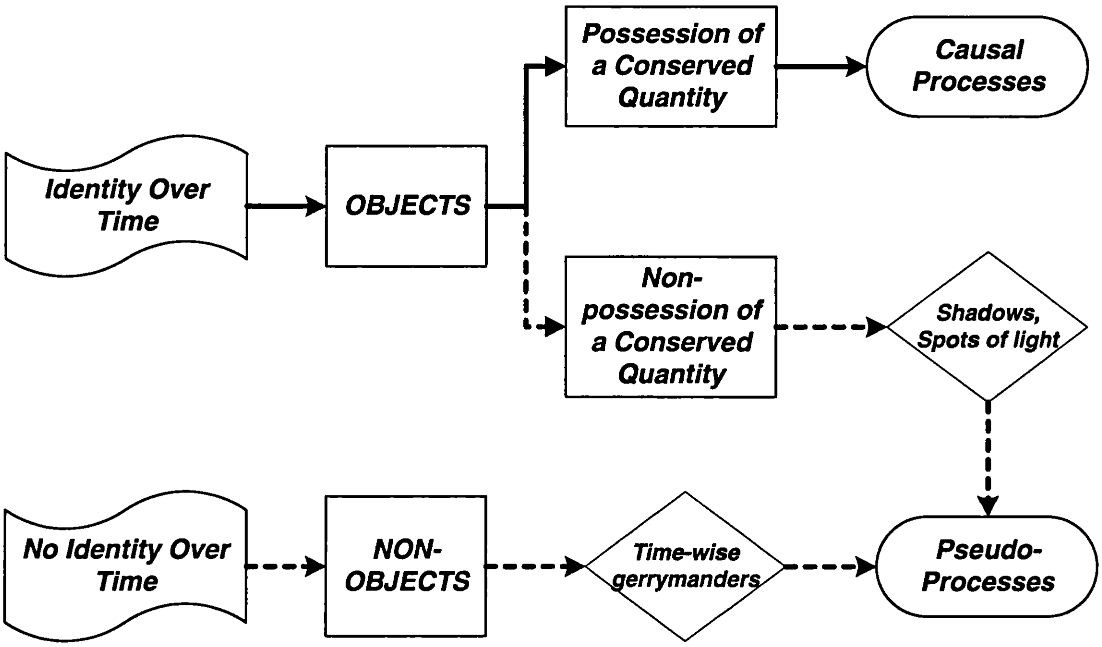

Figure 4.1: The structure of the CQ definition of a causal process.

of an object that is in possession of both a *conserved quantity* and an *identity over time* is regarded as a causal process. However, not all objects in possession of a conserved quantity qualify as causal processes, and likewise, not all entities having an identity over time qualify as causal processes. Time-wise gerrymanders that possess conserved quantities are ruled out as non-causal by the stipulation that they have temporal parts and so are not wholly present at each time.

Shadows and spots of light, on the other hand, are eliminated outright from the class of causal processes because they do not possess any conserved quantity. It is also worth pointing out that the fact that such objects do not possess any conserved quantity does not immediately condemn them to a situation of a lack of an identity over time[33]. Quite on the contrary, one needs the premise that the "object" remains the same object throughout its motion before even it is meaningful to put forward the claim that "the object" is in motion. Both the moving shadow and the moving spot of light display spatiotemporal continuity and qualitative similarity in the successive stages of their "motions". Perceptually, this dose not seem any different from the motions of a moving car or any other objects that have mass and energy. But yet,

[33]ibid.

shadows and light spots are, as we know, pseudo-processes that are incapable of propagating causal influences. Why is that so? Dowe argues that they cannot be propagating causal influences simply because they possess *none* in the first place; and so even as these objects move along their trajectories in spacetime, no causal influences are being transported. Hence on Dowe's construal, even though they have identities over time, objects that do not possess any conserved quantities are, of course, unable to propagate any of these quantities!

Dowe's strategy is clear: he intends to use "*the possession of a conserved quantity*" and "*identity over time*" as complementary criteria. On occasions when the former fails to rule out pseudo-processes, like the case with time-wise gerrymanders, he needs to fall back on the latter. However, where spatiotemporal continuity and qualitative similarity amongst the different stages of a certain process render the latter criterion useless in identifying a genuine causal process, he relies on the former to achieve the purpose.

For Dowe, an object that has an identity over time which also possesses a conserved quantity qualifies as a genuine causal process. But the vital question is whether the conjunction of both of these two criteria is indeed sufficient for us to capture the physical picture that each of the stages of this object is *actually causally connected* (with one stage being the source of the next)?

To articulate further, we consider the following counterexample. Suppose in place of a rotating spotlight shining onto the wall of the astrodome, we consider the case where the surface of the wall is consisted of a series of mini LED cells closely tiled continuously one after another in a horizontal line, with each one of these cells wired to an independent electrical circuitry (Figure4-2)[34]. Provided that these circuitries are activated in a synchronized manner so that the successive light cells are lit in the order that the previous one goes out as the next is illuminated, one thus regains

[34]It might be objected that this arrangement of LED cells forms not a continuous process, but only a discrete one. However, the objection is not as severe as it seems in anticipation of Chapters 5 to 7, where the spatiotemporal continuity in the motion of objects is relaxed in the quantum domain. This will let us focus on the real issue of how to capture our intuitions about what is wrong with such cases: either they do not constitute a worldline of *an object*, or they do not *transmit* a conserved quantity.

a situation of a moving spot of light which no doubt *possesses* a genuine conserved quantity - energy. This gives a situation where we have an object - a moving spot of light - that presents itself as a persisting object that is *in possession* of a conserved quantity.

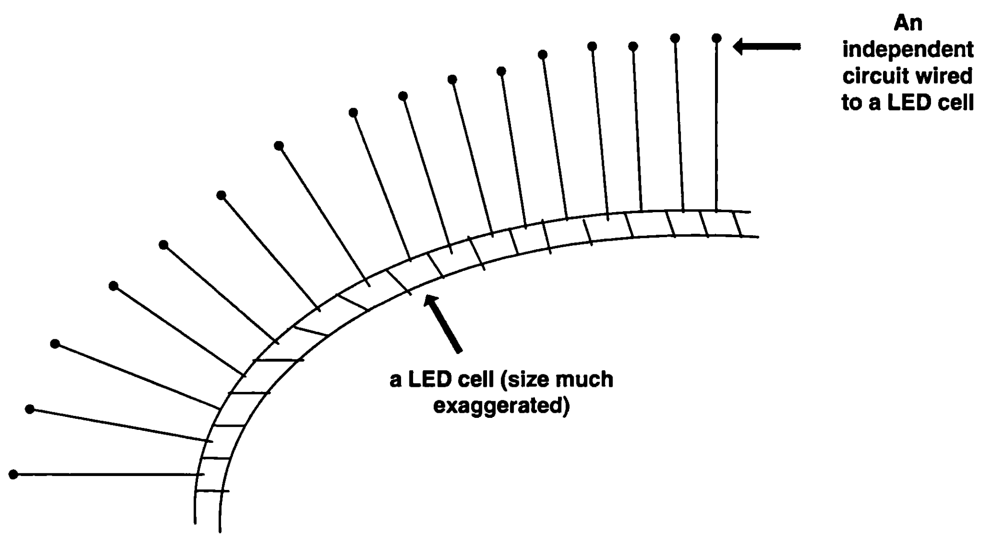

Figure 4.2: A series of LED cells each connected to an independent circuit.

However, this is a pseudo-process as we know that physically, one spot is *not* the source of the next, but each gain its energy from an independent external source; the different "stages" of the spot are not in any way causally connected. On this occasion, it would seem that Dowe's criterion cannot stand up to the test. This example is a variant of a counterexample discussed by Salmon[35], "*...we can take the worldline of the part of the wall surface that is absorbing the energy as result of being illuminated. This worldline manifests energy throughout the period during which the spot travels around the wall, but it is not the worldline of a causal process because energy is not being transmitted; it is being received from an exterior source.*" In Salmon's example, the point has been confused because he speaks of the "worldline" of the "parts" of the wall which illuminate wherever the light rays sweep across. Focusing on the *worldline* of this "object" that consists of patches of illuminated wall has left a lacuna for a defence from Dowe[36], "*The series of segments of the wall, each of which does possess*

[35]As he discusses the spot of light that moves across the wall of the astrodome, Salmon, W.C. (1994), p.308.
[36]Dowe, P. (1999), p.251.

energy, is not wholly existent at each time, so it is not the worldline of an object." I suppose a scenario along similar lines to our present example of a series of LED cells is closer to what Salmon had intended; namely, there are *independent* worldlines of objects which are wholly existent at each instant in their *own* respective history. The moving spot in our example, is wholly present at each time but at the same time it is a series of world-points, each of which belongs to the worldline of a different object.

It may be concluded from the above analysis that the possession of a conserved quantity does not warrant the definitive identification of a causal process and it needs to be supplemented by some thesis of identity over time. But Dowe's **SI** criterion has found to be too weak a condition as the last example shows. Dowe's criteria are thus incapable of eliminating all cases of pseudo-processes.

4.2.2 Causal Interaction

The notion of interaction is explicated within Salmon 1984 in terms of the geometrical idea of *intersection* - the crossings and meetings of two processes - at a localised region of spacetime. At the location of intersection, *certain* characteristics of both processes are altered. This captures vividly the intuition of the "actions-by-contact" that bring changes to the motions of the objects that come into close encounters as in classical mechanics. **CQ** stays faithful to the intuition encapsulated in Salmon's **CI** by allowing the geometrical concept of an intersection to continue to play a fundamental role in its definition of a causal interaction.

Leaving open the choice of what are to be counted as the relevant characteristics of processes that are modified in an intersection makes it necessary for **CI** to be infected with counterfactual in much the same way as **MT**. By asserting that an exchange of a conserved quantity characterizes a point of intersection of processes in a genuine interaction[37], Dowe has circumvented the need for counterfactuals. I judge this definition to have served its function well and to be physically sound.

[37]On Minkowski spacetime diagrams, physical interactions are represented by the crossings or the *intersections* of worldlines of objects that carry dynamical quantities like energy or momentum. These are the events in which the dynamical quantities of the participating processes are altered. There is an exchange - represented by a change in the numerical values - of conserved quantities at the point of intersection.

There is also this obvious virtue: unlike Salmon's **CI**, the notion of causal interaction entertained by **CQ** is not defined with respect to *two* processes. Dowe explains[38],

> ***An exchange means at least one incoming and at least one outgoing process manifest a change in the value of the conserved quantity. "Outgoing" and "incoming" are delineated on the spacetime diagram by the forward and backward light cones, but are essentially exchangeable. The exchange is governed by the conservation law.***

The upshot of relaxing the stringent demand of **CI** to requiring two incoming processes for an interaction and two "modified" processes to emerging as a result, is that one may now handle Y- and λ-type of processes. Salmon's account fails miserably to deal with these cases precisely because **CI** requires exactly *two* incoming processes and *two* outgoing ones to participate in a causal interaction. So, **CI** effectively rules out interactions that correspond to the "creation" or "annihilation" of particles like those in atomic decays.

For **CQ**, not all intersections of worldlines are causal interactions but only those that involve an exchange of a conserved quantity. When the "worldlines" of pseudo-objects (like shadows or the moving spot of light across the wall of the astrodome) intersect, the intersection would not, in principle, count as an interaction because there could not be any changes in the values of a conserved quantity, save the case that these objects do not possess any conserved quantities in the first place. The issue is, unfortunately, by no means clear-cut. There can be intersections that involve processes that are deemed as causal on Dowe's definition, are in fact, not so. Undoubtedly, any exchange of a conserved quantity associated with the intersection of these "causal" processes would be treated as a genuine interaction when they should not. The following scenario is a case in point.

We have shown how the criterion of the "*possession of a conserved quantity*",

[38]Dowe, P. (1992a), p.210.

combined with that of an identity over time, fail to eliminate all cases of pseudo-processes with the illustration by the example of the series of tightly packed LED cells that are manipulated to be illuminated one after another. For the sake of the present argument, the reader is invited to consider a peculiar arrangement of LED cells set up like the ones in Figure 4-3. As before, each cell is wired to an independent circuitry and the lighting of each is synchronized in such a manner that one comes to light up after another goes out along the series. To the light cells located along the tracks BC and BD, we have the energy of each cell reduced but demand that the total amount of energy of two light bulbs, one placed along BC and the other along BD, to light up at the same time and be equal to the value of energy of any one light bulb that is situated along AB.

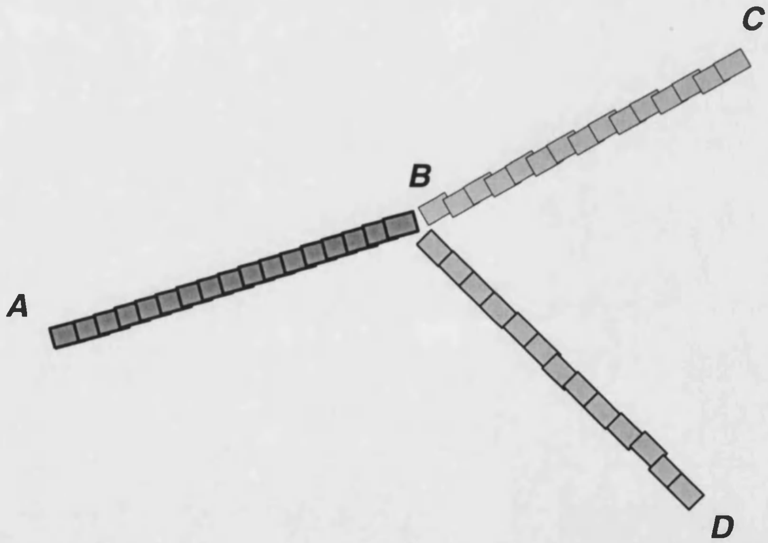

Figure 4.3: A demonstration that pseudo-processes can satisfy the **CQ** definition of a causal interaction.

We are here presented with a scenario where there is one incoming process and two outgoing processes, one which is of the Y-type interaction. Each of these processes possesses energy and thus qualifies as a causal process on Dowe's definition. Moreover, at the point intersection B, there seems to be a change in the values of energy in accordance with the conservation law.

However, the disheartening fact is that all three processes AB, BC, BD are, in fact, pseudo-processes but yet, the condition of the exchange of a conserved quantity is satisfied. Hence, the intersection at B qualifies as a *bona fide* causal interaction according to the **CQ** definition. This, of course, would not, *prima facie*, violate Dowe's definition of a causal interaction as it is not part and parcel of the definition to require the worldlines to correspond specifically to causal processes. It does, however, reveal the vexing difficulty of delineating the actual processes like the cells at B and the electric current supplied by three independent circuits with respect to which the exchange has taken place, rather than the non-causal processes like AB, BC and BD.

The failure to filter out "pseudo-intersections" can of course be traced back to the fact that the condition of mere possession of a conserved quantity has not succeeded in representing what takes to be a genuine causal connection. One natural response to this difficulty is to investigate whether the the notion of transmission, which Salmon has claimed to be the crucial element for capturing the very essence of a causal connection, would help to resolve the problem (see Section 4.3.2).

4.3 Towards a Conserved Quantity Theory with Transmission

4.3.1 A Quick Trip Through the Invariant Quantity Theory

Salmon responded positively to Dowe's critique and proposal of **CQ** in his 1994 paper titled *"Causality Without Counterfactuals"*[39]. In this work, Salmon acknowledges that the requirement of his original 1984 theory to define causal processes as ones that persist in the absence of interactions leads him to adopt the notion of a "mark" and to introduce counterfactuals to both the formulations of **MT** and **CI**. The use of counterfactuals has raised serious concerns of an epistemological nature as have been expounded masterfully by Kitcher[40]. Disturbed by these unsettling issues, Salmon welcomes **CQ** with the statement, "Dowe's proposed *conserved quantity* theory is beautiful for its simplicity"[41].

[39]Salmon, W.C. (1994).
[40]Kitcher, P. (1989).
[41]Salmon, W.C. (1994), p.303.

Inasmuch as he concurs with Dowe's results, **CQ** has not received a complete endorsement from Salmon. There are two main modifications to **CQ**, both in relation to the identification of a causal process, which in Salmon's opinion, that need to be made: (1) the replacement of conserved quantities by *invariant* quantities and, (2) the introduction of the notion of transmission.

In regard to (1), Salmon points out that a conserved quantity like energy is not invariant because its numerical value varies depending on the particular frame of reference concerned[42], "*To say that a quantity is <u>conserved</u> (within a given physical system) means that its value does not change over time; it is constant with respect to time translation. To say that a quantity is <u>invariant</u> (within a given physical system) means that it remains constant with respect to change of frame of reference.*" and[43], "*When we ask about the ontological implications of a theory, one reasonable response is to look for its invariants. Since these do not change with the selection of different frames of reference - different perspectives or points of view - they possess a kind of objective status that seems more fundamental than that of non-invariants.*"

To bring into focus the idea of causality as an invariant notion, he suggests the replacement of "conserved quantities" with "invariant quantities" in the **CQ** definition of a causal process[44], "*A causal process is a worldline of an object that manifests an invariant quantity.*".

However, Salmon confesses that this modified definition at once lends itself to grave troubles. This is because the most basic of all invariant quantities in nature is the spacetime interval between two events. This particular invariant is, unfortunately, *manifested* by processes of all sorts - causal or not - provided that a finite duration has elapsed between an initial point and a final point in that process. Salmon argues that the term "manifests" is not an adequate concept to capture the fact that something is propagated <u>within</u> a causal process[45]: "*A necessary condition for a quantity to be transmitted in a process is that it can meaningfully be said to characterize or be*

[42]ibid., p.305.
[43]ibid., p.310.
[44]ibid., p.305-306.
[45]ibid., p.306.

possessed by that process at any given moment in its history.".

Appealing to the idea that the quantity being transmitted ought to be meaningfully specified at *each moment in the history of the process* would serve to bar an invariant quantity like the spacetime interval. This is because the spacetime interval is a quantity that can only be meaningfully assigned to a process *globally* (as a property of two endpoints) rather than *locally* (as a property of every single point within the interval). While the notion of transmission encompasses exclusively the idea of possessing a property locally, that of "manifest" refers to the possession of a property either locally or globally.

Transmission therefore remains a fundamental notion and this is the important emendation of Dowe's definition[46],

> ***A causal process is a worldline of an object that transmits a non-zero amount[47] of an invariant quantity at each moment of its history (each spacetime point of its trajectory).***

In addition, Salmon offers the following revised *non-counterfactual* criterion for transmission[48],

> ***A process transmits an invariant (or conserved) quantity from A to B (A≠B) if it possesses this quantity at A and at B and at every stage of the process between A and B without any interactions in the half-open interval (A,B][49] that involve an exchange of that particular invariant (or conserved) quantity.***

[46]ibid., p.308.

[47]The phrase "manifests a non-zero amount of an invariant quantity" has been inserted to block the kind of assertion that a pseudo-process (a shadow for example) manifests an invariant quantity (like an electric charge) whose value can be treated as zero in effect. This stipulation has the repercussion that genuine casual processes which, on some occasions, do manifest a zero-value of an invariant quantity are also excluded from the definition. The photon is a case in point. It has an electric charge that equals zero. Nevertheless, it is not disqualified by the above definition from being a causal process, for, although it does not possess an invariant quantity such as an electric charge, it does so with its invariant speed. We may indeed regard the insertion of this condition as a kind of constraint to decide what are the relevant invariant quantities to be possessed by a given causal process. The vital message is that ***a pseudo-process does not possess any sort of invariant quantities whatsoever.***

[48]ibid.

[49]The half-open interval $(A, B](A \leq B)$ has been introduced to allow for the possibility that there

This two clauses above form the ***Invariant Quantity of Causation*** (**IQ**). Like **MT**, says Salmon[50], "*This definition embodies the "at-at" theory of transmission, which still seems to be fundamental to our understanding of physical causality.*"

It follows from the definitions above that a pseudo-process is one which either: (1) does not possess an invariant (or conserved) quantity or, (2) its possession of an invariant (or conserved) quantity involves "some" interactions concerning that particular quantity within the half-open interval $(A, B]$.

Shadows and "moving" light spots do not possess any invariant (or conserved) quantities and hence they both fall into the category of pseudo-processes. Others which do possess these quantities but are yet considered as non-causal because their possession of the quantity in question throughout the interval is parasitic upon the continuous replenishment of the quantity from an external source(s). The example in Section 4.2.1 of a moving light spot generated by an arrangement of mini light cells, set to illuminate one after another, illustrates. The event that gives rise to the possession of energy represents an interaction between the electric current and the filament in the light cell. As the filament heats up, it eventually reaches a point when a photon is emitted which we see as the spot of light. A similar account may be produced for the spot in the "next" position that is created by the same procedure with the details only differ in the involvement of the electric current from a different circuit and a different light cell. The resulting spot is, as a matter of fact, a series of independent interactions where there is emphatically the absence of the transmission of energy amongst them.

Likewise, the time-wise gerrymander consisted of a sequence of the motions of three billiard balls discussed in Section 4.2.1 is an object that possesses energy and momentum. This is again not a (single) causal process as there are interactions between the balls (in addition to the lack of identity over time). In this case the

is an interaction at point A that determines the amount of the quantity to be transmitted. This would also account for cases where the causal process comes into existence at A. The crucial point is that having acquired an invariant (or conserved) quantity at A, a causal process is, in principle, capable of transporting this quantity at every spacetime point between A and B. And of course, the closed interval at B gives the process the freedom to participate in interactions at B. There would be no more interactions for the rest of the spacetime points within this interval.

[50]ibid.

pseudo-process may be re-analyzed in terms of the three separate causal processes that correspond to the respective motions of the three balls.

4.3.2 The Latest Episode: A Conserved Quantity Theory with Transmission

> ***Renouncing the mark-transmission and invariant quantity criteria, I accept a conserved quantity theory similar to Dowe's - differing basically with respect to causal transmission.***

By 1997, Salmon has reverted to adopting conserved quantities in his process causal theory and put forth what I call his "***Conserved Quantity Theory with Transmission***" (**CQT**). In this section, we recount the key issues leading up to the formulation of **CQT**.

CQT represents Salmon's response to two major critiques on **IQ** from Phil Dowe[51] and Christopher Read Hitchcock[52]. Intended as a defence of **CQ**, Dowe enters into a debate over the sufficiency of the notions of "possession" and "transmission" of conserved quantities and the closely related issues of the identity over time of persisting objects. Hitchcock's comments are, however, of a broader nature and are addressed to both the process theories of Salmon and Dowe.

The basic definitions of **CQT** goes as follows[53],

> ***A causal interaction is an intersection of worldlines that involves exchange of a conserved quantity.***

> ***A causal process is the worldline of an object that transmits a non-zero amount of a conserved quantity at each moment of its history (each spacetime point of its trajectory).***

and[54],

[51] Dowe, P. (1995).
[52] Hitchcock, C. (1995).
[53] Salmon, W.C. (1997), p.468.
[54] ibid., p.462.

A process transmits a conserved quantity between A and B (A≠B) if and only if it possesses [a fixed amount of] this quantity at A and at B and at every stage of the process between A and B without any interactions in the open interval (A, B) that involve an exchange of that particular conserved quantity.

Seen as a response to Dowe's 1995 criticisms on the notions of transmission and the invariant quantities, the motivation behind Salmon's revised formulation of the process theory is clear. We shall now turn first to the definition of transmission.

First to notice is Salmon's reversion to conserved quantities in his definitions. This move has resulted from the consideration of a series of entangled issues between the notions of conserved quantities and transmission as we shall now explain. He sees the need to abandon invariant quantities in favour of conserved ones for it enables him to deal with the counterexample raised by Hitchcock[55]. The counterexample involves a shadow of an object moving across a metal plate with a uniform non-zero change density on its surface. Hitchcock's claim is that basically, the shadow may be considered to possess a constant electric charge that is both conserved and invariant, but as it glides across the plate, the shadow is not involved in any causal interaction. It thus qualifies as a causal process on both the definitions of Salmon and Dowe[56].

The portion of the surface of the plate in the shadow is in possession of a constant amount of electric charge, which is similar to the portion of wall on which a light ray falls in the case of the astrodome that possesses energy. It could therefore, at first sight, be regarded as a causal process. However, there is a subtle difference between the two cases. For the patch of wall in the astrodome, the possession of energy arises because of its interaction with the light ray, the presence of an interaction therefore dismisses the patch of wall as a candidate for a causal process. But since the metal plate possesses a constant amount of electric charge without any interactions,

[55]Hitchcock, C.R. (1995), p.314-315.

[56]As Salmon has rightly observed, and I agree, this scenario does not pose any particular problem for Dowe (Salmon, W.C. (1997), p.472.), "*Dowe will readily reject Hitchcock's example on the ground that shadows do not have electric charges; in this case, the charge belongs to the metal plate. This response is, I believe, correct.*"

it therefore appears that Salmon's criterion of transmission fails to work on this occasion.

The resolution of the difficulty lies, as Salmon maintains, with the understanding of the essence behind the notion of transmission. It makes good sense to say that something has been transmitted or has moved from one spacetime region to another, provided only that something vacates the first region and appear in the other.

The point may be illustrated with our arrangement of mini light cells, with each of these cells being illuminated by an independent circuit in order to generate a spot of light, and these spots are synchronized in such a way that one goes out as the next is lit. This synchronized sequence conveys just the right impression for one to regard the situation as a spot moving (Figure 4-4).

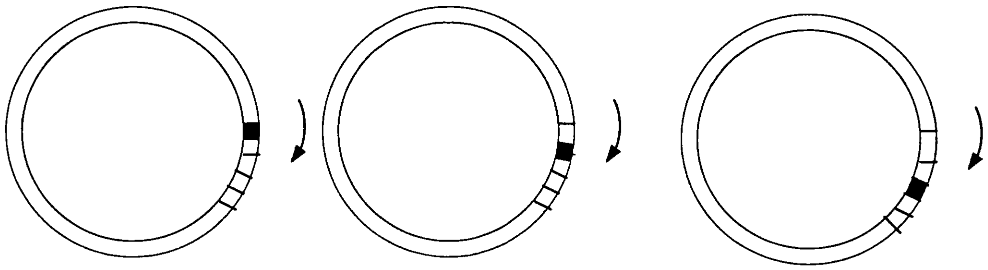

Figure 4.4: A moving spot of light vacating the last position and taking up the next.

Consider the same arrangement but this time with each light cell staying lighted once being brought into illumination as in Figure 4-5.

Two observations ensue. First, the situation depicted by Figure 4-5 would not give us the impression of a "single" moving spot but rather that something is *added* to the sequence as each light cell is being illuminated. Second, specific to this situation, an additional amount of energy in the form of light is *added* to this "spreading" process. In other words, this process does not carry a constant amount of energy. Strictly speaking, therefore, additional amounts of energy would have to be added to each stage of the process, and so there is no transference of energy from one spacetime

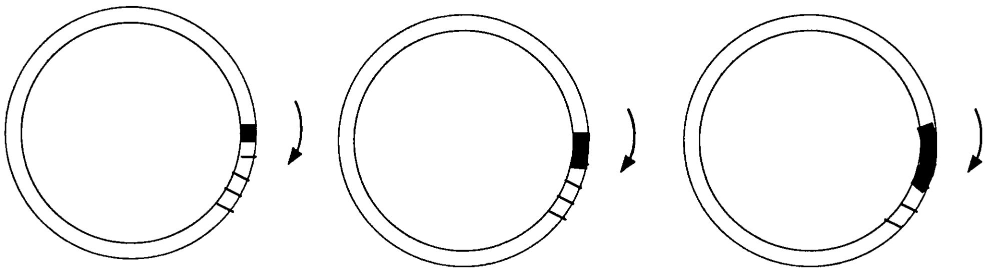

Figure 4.5: Each light cell stays lit once it is illuminated.

locale to another. Transmission then becomes an ambiguous term in this kind of situations. So although the movement of a constant amount of energy from one position to another does not guarantee[57] the distinction of causal processes, it is, nonetheless, a prerequisite for transmission to obtain.

To say that additional amounts of energy are acquired as the process develops is tantamount to the statement that energy is not conserved within the process. The non-conservation of energy may be traced to the constant acquisition of energy from some other sources which is achieved through interactions. This is what in fact happens and the process is thus readily dismissed as a non-causal one.

The same line of reasoning[58] can now be applied to discredit Hitchcock's counterexample. As the shadow moves along, the area of the plate in the shadow appears to possess a *constant* amount of charge. The main point remains that in order to regard the portions in the shadow as *possessing* a constant amount[59] of electric charge

[57]What makes us think that the light spot cast by the rotating beacon at the centre of the astrodome is "one single spot" that moves? Precisely because it appears to vacate one location as it takes up the next. We find this situation troublesome because a sequence of static interactions may, under such circumstance, bear just the correct spatiotemporal relation to satisfy an equation of motion, while having no real transfer of causal influences taking place amongst them. Both the causal programs of Salmon and Dowe aim to tell in a cogent manner, how this apparent moving object is indeed different from that of a moving physical material body. For them, the answer rests with the contention that there exist causal relations between the successive stages of the latter but not the former. Once this is recognized, the remaining tasks then amount to the expression of these causal relations in terms that abide by the empirical strictures: the relations of physics.

[58]Salmon, W.C. (1997), p.473.

[59]It has been pointed out to me that Hitchcock's example may be modified so that a source of a positive charge (say) induces a negative charge on a localized area on the metal plate as it moves along from underneath the plate. This scenario would then produce the observation that the quantity of (negative) charge vacating one place and moving into another; as well as having the advantage that it is a continuous process. I owe this example to Dr. Carl Hoefer.

without interaction, namely, that a *constant amount* of electric charge is *transmitted*, this quantity of charge has to vacate the previous portion and subsequently appears at the next portion where the shadow falls. As Salmon argues[60],

> ***Informally we want to say that electric charges are carried by particles like electrons and protons; they are transmitted between different spatiotemporal regions by the movement of such particles. This involves the passage of electric charges from one locale to another, thereby augmenting the electric charges already there. The same consideration applies to the intermediate spacetime locations - that is, the electric charge in question must vacate its location at one stage of the process and appear at the other stages of the process at the appropriate times. Otherwise, the electric charge would not be a conserved quantity.***

The contribution of conserved quantities to the thrust of his argument prompts Salmon to regard conserved quantities as the appropriate dynamical quantities that are possessed by causal processes.

In practice, however, argues Dowe, a physical object is generally not confined within a closed system that is shut off from all interactions. Especially, it undergoes continuous interactions with its environment. The constant bombardment by air molecules in its surroundings and the continuous gravitational interactions are not the only but two of the most encountered instances. These constant interactions render the concept of the transmission of a constant dynamical quantity[61] (and in turn the notion of a causal process) vacuous. And Dowe concludes[62],

> ***For this reason the CQ theory does not require that a causal***

[60]ibid., p.472-473.

[61]It would only be consistent, in the absence of interactions, to speak about a fixed (or constant) amount of a quantity to be transmitted. This is because the value of a quantity to be transmitted changes whenever there is an interaction. Hence, this reading of the meaning of transmission restricts one to a constant amount of a transmitted quantity.

[62]Dowe, P. (1995), p.331.

possess a constant amount of the relevant quantity over the entire history of the process.

As part of his defence of the notion of "possession" against that of Salmon's "transmission", Dowe also takes issue with the direction of causation in the definition of transmission.

On the concept of transmission, Salmon remarks[63], "*Dowe says that it is unnecessary; I claim that it is indispensable.*" And so much for the dichotomy in their respective positions.

Clearly, **CQ** takes a neutral position[64] with respect to temporal direction and, can therefore be said to be essentially a temporally-symmetric theory[65]. The expression "from A to B" within Salmon's new definition of transmission[66] does, however, strongly suggests a temporal directionality in which transmission is to take place. This immediately finds itself at odds with the "*at-at*" condition which is satisfied, as long as the quantity concerned (be it invariant or conserved) is *possessed* at every stage between A and B. Thus, Dowe claims[67], "*Indeed, the notion of transmission turns out to involve no more than what is contained in the notion of "possession". I conclude that there is no advantage in appealing to "transmission" (as in the **IQ** theory) rather than "possession" (as in the **CQ** theory).*"

Responding to Dowe's challenge to the notion of transmission, Salmon effected a change in the revised definition of transmission with the replacement of the directional expression "from A to B" by "between A and B"[68]. The aim is to restore temporal

[63]ibid., p.466.

[64]Both Salmon and Dowe have at times expressed wishes to pursue a process theory that is non-committed to a preferred temporal direction on which a causal theory of time may be grounded. Any *a priori* commitment to a temporal direction is considered to place such a programme under threat.

[65]This is because "possession" is a temporally symmetrical concept.

[66]Salmon, W.C. (1994), p.308.

[67]Dowe, P. (1995), p.326.

[68]Another minor point of emendation to note is the alteration from a half-open interval $(A,B]$ $(A < x \leq B)$ to an open interval (A,B) $(A < x < B)$. This serves to make precise the idea that a causal process is capable of transmitting causal influence without resorting to external sources in-between interactions. The half-open interval $(A < x \leq B)$ may give the incorrect impression of the process entering upon an interaction at B. Given that it has not taken part in any interaction, the process ought still to be possessing the conserved quantity in question at point B. However, a genuine interaction at B can run the risk of being ambiguously counted against the causal process,

symmetry in order to facilitate the construction of a temporally-neutral causal theory on Dowe's suggestion. This granted, however, Salmon still remains loyal to the notion of transmission[69],

> ***The issue concerns the concept of transmission, and it is centered on the final clause of the definition, "without any interactions...that involve an exchange of that particular conserved quantity" ...From my point of view the crucial question regarding causal processes is what they do on their own without outside intervention. My answer is that they transmit something - e.g., conserved quantities, information, or causal influence - and it is by virtue of such transmission that events at A and B are causally related.***

We shall now scrutinize this dispute of transmission over possession. In the example of the three alignments of LEDs (Figure 4-3) given at the end of Section 4.2.2, Dowe's criterion of *mere possession* has been shown to fail to disqualify AB, BC and BD as pseudo-processes. And also, the intersection at B is identified as a causal interaction by **CQ**; although, in fact, this causal interaction at B is due *not* to the intersection of the processes AB, BC and BD, but rather to that between the three LED cells (each positioned along one process around the vicinity of point B) and the independent circuit that each is connected to.

Can **CQT** do better[70]? Firstly, there are interactions that involve exchange of a conserved quantity within the processes AB, BC and BD and so there is *no* transmission of a conserved quantity along each of these processes. Hence, all three are pseudo-processes. However, the intersection of these three "worldlines", right at the localized point B, does involve an exchange.

---

which has been able to sustain its own existence and to transport causal influences of its own accord at every spacetime point from A up until the encounter at B. In order to remove this ambiguity and at the same time to allow for the fact that a proper interaction at B should not frustrate our determination of whether a given process is genuinely causal, there is the need to exclude interactions at B.

[69]ibid.

[70]We ended the discussion in Section 4.2.2 by suggesting to investigate whether the notion of transmission would help to resolve the problem.

So, we have arrived at the following situation. Dowe's **CQ** criterion of mere possession cannot pick out AB, BC and BD as pseudo-processes and it also deems the interaction at B as a genuine causal interaction. It is thus concluded that the intersection of the "causal process" (by virtue of **CQ**) AB, BC and BD has given rise to this interaction. On the other hand, Salmon's **CQT** can correctly identify AB, BC and BD as pseudo-processes, but it appears that we now have a peculiar scenario whereby the intersection of these pseudo-processes has given rise to a genuine causal interaction!

Therefore, even though it would seem that **CQT** does offer more than **CQ**, since it has successfully identify AB, BC and BD as pseudo-processes with its criterion of transmission, it too, however, is incapable of expressing the fact that the intersection at B is not an interaction *with respect to* these three specific processes in question.

How may one remedy the situation? What we really need to capture is the fact that the causal interaction - the intersection with an exchange of a conserved quantity - at B has arisen from the intersections of the three LEDs and the current in the wires connecting each to their respective circuits; that is, from the intersections of causal processes!

A natural move is to state in explicit terms that only *intersections of causal processes* involving the exchange of a conserved quantity are deemed as causal interactions.

Intuitively, when pseudo-processes interact, one would not expect any one of them to suffer any changes as a consequence of their intersections. Two shadows may momentarily overlap but as they "move" away from the point of intersection, they recover their respective shapes. However, the situation becomes more delicate with pseudo-processes such as two spots of light moving across a wall intersecting for instance. At the point of intersection, there *is* exchange of energy, a conserved quantity; although the exchange takes place between each of the two light rays and the area of the wall on which they both fall.

A crucial observation ensues. In each of the above scenarios where there is an actual exchange of a conserved quantity when two pseudo-processes intersect, the exchange is found to occur between the causal processes that the pseudo-ones originate. The two moving spots of light are the product of two light rays sweeping across the wall, and so are the processes AB, BC and CD being the product of the interactions amongst the different light cells and circuitries.

Given that there is a change in the value of a conserved quantity in an interaction and the change has to obey the constraints imposed by the conservation laws; only causal processes *with the capacities to transmit conserved quantities* may participate in interactions. It is impossible for a pseudo-process that lacks the capability to transmit a conserved quantity[71] to participate in an interaction - an intersection where there is an exchange of a conserved quantity.

To labor the point, imagine the following situation. A red-filter is placed a shade above the area of the wall of the astrodome where a beam of light falls and the spot of light becomes red as a result. There seems to be an exchange of a conserved quantity that involves a pseudo-process like the light spot. On careful reflection, at the very point of intersection, there are several processes at work. There are the red-filter, the light beam and also the patch of wall on which the beam falls. The red-filter absorbs all other wavelengths in the beam of white light and allows only red light through. As a "by-product", the spot on the wall becomes red. The exchange of energy, a conserved quantity, has taken place between the red-filter and the beam of white light, with both of these being causal processes. But at this very point of intersection, is there much meaning in speaking about the spot as *independent* from the patch of wall that is illuminated by the red-filtered light? Especially as we are just concerned with the "light spot" at the point of intersection and that point only, and not the *moving* spot. The crux of the matter, as I see it, is this: insofar as the "static" spot is concerned, there is no physical difference between the illuminated patch of wall and the light spot itself, it is purely a linguistic maneuver to *call* the the illuminated

[71]It may be objected that an object like the gerrymander given in Section 4.2.1, although a pseudo-process, does, however, possess conserved quantities. But since a pseudo-process of such kind may be easily re-analyzed as a composition of a number of causal processes, the term "pseudo-process" would be reserved for processes which involve objects that do not in fact possess any conserved quantity.

patch of wall a "spot" of light. But it would matter as soon as the "*movement*" of the spot comes into the picture, the moving spot thus becomes a process, albeit a pseudo-one, with its successive stages bearing no cause-effect relation to each other. The all important message is that at the static level, should there be an exchange of a conserved quantity on the point of intersection, the exchange is to be accounted for in terms of causal processes that also intersect at that very point. But how do we know which processes are causal? This is only possible through the consideration on the "*dynamical*" level of these processes, which concerns the relation and non-relation between the successive stages of the processes.

At the locus of the intersection, one may, of course, consider the part of the "worldline" of a pseudo-process, namely, the light spot. However, this pseudo-process plays a redundant role in the actual exchange for it neither receives nor dispatches any conserved quantity. I would therefore suggest a slight modification to Dowe's definition of a causal interaction,

> ***A causal interaction is an intersection of causal processes that involves exchange of a conserved quantity.***

The stipulation of causal processes as the proper participants in intersections that give rise to genuine interaction alleviates our difficulty and with this modification, **CQT** is now capable of identifying all three processes AB, BC and CD as pseudo-processes, as well as dismissing intersection at B as a genuine interaction *with respect to* these pseudo-processes.

But now, a deeper worry ensues. The incorporation of causal processes into the definition of causal interaction drives **CQT** to a circularity, because causal processes are themselves defined in terms of transmission, which is in turn defined with reference to interaction[72]. So, it does seem that an independent definition of causal processes

[72]The circularity runs as follows. The decision of the intersection at B being a non-causal interaction with respect to the processes AB, BC and CD hinges upon the premise that these three are pseudo-processes. Why are these non-causal processes? It is because they do not transmit conserved quantities. Why do they not transmit conserved quantities? It is because their possession of conserved quantities obtain by the intersections of causal processes that involve an exchange of a conserved quantity.

is required for circumventing the circularity[73].

Recall that the one essential feature of causal processes that the notion of transmission has all the while been trying to capture is that, causal processes have the ability to sustain themselves without any intervention from outside the process. And so the locution "without any interactions that involve an exchange of that particular conserved quantity" deserves a further scrutiny.

The understanding that a causal interaction is not just any odd intersection of processes[74] but only those at which the exchange of a conserved quantity occurs, carries the implication that there can be intersections of processes which do not count as interactions. Hence, the following two[75] propositions are true:

(i) *there is no intersection of worldlines and* ***hence*** *no exchange of any conserved quantity.*

(ii) *there is an intersection of worldlines* ***but*** *without an actual exchange of a conserved quantity.*

Both propositions represent cases of *no* interaction and are therefore relevant to the definition of causal processes.

It turns out that a slight modification by incorporating (ii) into the definition would serve our purpose. This is attained simply by the replacement of the term "interaction" (that is defined by causal processes) by "intersection", which results in the notion of transmission, and hence that of causal processes, being independent from the notion of interaction.

Let us now state the three definition of the revised form of **CQT**,

> ***A causal process is the worldline of an object that transmits a non-zero amount of a conserved quantity at each moment of its history (each spacetime point of its trajectory).***

[73]Incidentally, this is one of the two ways out of the circularity charge as indicated earlier in Section 3.3.2 (Figure 3-4).

[74]We do not specify the exact nature of the processes at this point and the reason will be apparent from the discussion to follow.

[75]The possibility of an exchange of an invariant quantity without an intersection is discarded.

A process transmits a conserved quantity between** A **and** B **(**$A \neq B$**) if and only if it possesses [a fixed amount of] this quantity at** A **and at** B **and at every stage of the process between** A **and** B **without any <u>intersections</u> in the open interval (**A**,** B**) that involve an exchange of that particular conserved quantity.

A causal interaction is an intersection of <u>causal processes</u> that involves exchange of a conserved quantity.

It remains to be shown how Salmon's scheme works to identify casual processes in concrete settings. In particular, we are interested in whether Salmon's criterion is capable of telling causal processes from non-causal ones.

Firstly, the worldlines of objects like shadows and a moving spot of light originating from a rotating beam some distance away are at once ruled out (as pseudo-processes) by the definition since they are not in possession of any conserved quantity.

The worldline of an object that possesses a conserved quantity would be regarded as a causal process even in the presence of an intersection with the worldline of another object, provided the other object is *not* in possession of any such conserved quantity so that no exchange can take place on intersection. An example is a ball placed in the shadow of a car. The ball is an object possessing energy while the shadow is not. There is no exchange of energy between the ball and the shadow. So the worldline of the ball is a causal process.

Yet another example is the meeting of two light rays which passes each other without any exchange of conserved quantity occurring. Both of these objects possess energy and momentum and since the intersection does not alter their dynamical makeups, these are not interactions which would count against their worldlines being causal processes.

The ultimate test, perhaps, rests with the kind of example such as a process that appears as a moving spot of light created by the successive illumination of a series of mini LED light cells aligned in a continuous array. Each stage of all three processes AB, BC and BD *is* an intersection involving the exchange of energy and it follows therefore that there is *no* transmission taking place along each of these processes, because possession of energy is by virtue of these intersections. And since there is no transmission of a conserved quantity, these three are considered *not* as causal processes. Hence, by our revised definition of causal interaction, which specifies causal processes as the proper subjects that undergo intersection in a legitimate interaction, the intersection of the processes AB, BC and BD is *not* a causal interaction; since the three are not proper worldlines of objects (which are now understood to be causal processes). That is, the exchange at B should not count as a causal interaction *with respect to* the processes AB, BC and BD.

However, *relative to* the particular light cells situated around point B and the circuitries that they are connected to, the exchanges are indeed *genuine* interactions. It is because the processes of the light cells and the current contained in the wires can sustain themselves without any interaction, and as such are regarded as causal processes.

Given that an intersection of processes that involves an exchange of a conserved quantity has taken place in a localized region of spacetime, one would be able to declare it a genuine interaction with respect to *only* the causal processes that intersect.

In this section, we have shown that with a slight modification, the notion of transmission is more effective in delineating causal processes from non-causal ones, than the notion of possession.

4.4 Continuity, Identity Over Time and the Worldline View of Causal Processes

In this section, I shall consider a cluster of closely related issues that are of direct importance to our present investigation. In particular, I shall argue that **CQT**[76] and **CQ** both trade heavily on the notion of spatiotemporal continuity. This is because both adopt the view of a causal process as the worldline of an object, which is essentially a continuous entity in spacetime. However, as we have seen, not the continuous motions in spacetime of all "*objects*" can be counted as causal processes, with the

[76]Now in its revised form given in Section 4.3.2.

revolving spot of light as the candidate example of a pseudo-process. This is what has driven both Salmon and Dowe to resort to their respective criteria of "transmission" and "possession" to distinguish between causal and non-causal processes. Both criteria, must nevertheless, be complemented by a thesis of identity over time of an object - the stipulation that the object is *one and the same.*

With the foregoing in mind, this section is organized in two main parts. It is shown in Section 4.4.1 how the notions of *causal processes*, *spatiotemporal continuity* and *identity over time* are related. I shall argue that the kind of identity over time needed for causal processes must enter as a fundamental supposition. In Section 4.4.2 that follows, a discrepancy between the usage of the term worldline in both **CQT** and **CQ** and that in physics will be pointed out. The clarification of this discrepancy helps to pave the way for extending the notion of causal processes to quantum physics.

4.4.1 Spatiotemporal Continuity, Identity Over Time and Causal Processes

In explaining how the notion of spatiotemporal continuity got bound up with that of causal processes, Salmon says[77],

> ***...I shall assume without further ado that we can observe many macroscopic physical objects, processes and events, and that we can legitimately infer the existence of such entities when they are not actually being observed. We infer the continuous existence of the kitchen clock while we are not at home, and explain the positions of its hands in terms of continuous processes that have transpired in our absence. We maintain that the planet Mars exists and moves in a continuous path during the day and at times when the sky is obscured by clouds... If any serious question arises, it can in principle always be settled by making an appropriate kind of observation. This approach,***

[77]Salmon, W.C. (1984), p.206.

which conforms to common sense, enables us to endow our world with a great deal of spatiotemporal continuity.

To explicate the inference to spatiotemporal continuity, Salmon considers C.J. Ducasse's example of the mouse in the basement of his home. Upon the discovery of a little mouse in his basement, Professor Ducasse, a great animal lover, was said to have carefully constructed a special trap which enabled him to capture the mouse without bringing any harm to it. The creature was then duly released in a vacant field by the good professor while on his way to work. After this pattern of events had been repeated on a number of successive days with the mouse caught on each occasion looking very similar, the professor decided to "mark" the mouse he caught by a small blob of white paint before having it released in the usual manner. That same very night, the trap was set once again and much to his expectation, the mouse caught was one with a white mark on his head. In this example the physical process, the mouse, was observed by Ducasse only at disconnected times; but says Salmon[78], "*we have no doubt that the process itself possessed spatiotemporal continuity. Ducasse employed the mark method to ascertain whether he was dealing with a single causal process (rather than many different mice), and there is no doubt that the process itself possessed spatiotemporal continuity of the transmission of the mark.*" And he seems to think that[79], "*This example is unproblematic, for the process is one which is in principle observable throughout its duration.*"

So, **MT**, the criterion for mark transmission that embodies the "*at-at*" theory of transmission is regarded as the first entry-point where the notion of spatiotemporal continuity comes into Salmon's 1984 causal theory. According to the "*at-at*" theory, the transmission of a mark or more generally of causal influences, is made possible by the correlation of the continuous one-dimensional *time* sequence with the term(s) of the three-dimensional continuous *spatial* sequence. In other words, there ought to be a continuous spatiotemporal trajectory of the mark in question and any suggestion of *action-at-a-distance* is unacceptable.

[78]ibid., p.208.
[79]ibid.

As we have pointed out, transmission is a form of motion whereby causal influences move from one spacetime locale to another. This point merits further elucidation. To proceed, Russell's "*at-at*" theory of motion, on which the "*at-at*" theory of transmission is based, deserves a closer examination.

The concept of motion, Russell remarks, is logically subsequent and founded on the idea of *occupying a place at a time* and also that of *change.* One sees, one senses or even feels something changes, from the changing moods of deciduous trees to the tragic feeling of love slipping away. In each and every instance when a change is perceived or sensed, a difference in the very object or state of affairs *prior* and *subsequent* to the change is what leads one to the conclusion that a change has taken place. This essential difference *in time* is what has been captured most vividly by Russell's notion of change. Change, as defined by Russell[80], is the *difference, in respect of truth and falsehood, of the same proposition concerning an entity occurring at two different times.* Thus, mere existence on this definition constitutes a basis for change since if the proposition "A exists" is *true* at time T but *false* at T_1 when A ceases to exist, then this amounts to a difference in regard to the respective truth and falsity of the same proposition at two moments of time T and T_1.

Notice that the definition does not immediately lend itself to the conclusion that change ought to assume a continuous nature with respect to, at the very least, the continuity of time as taken in the Cantorian sense. For the difference is one that refers to two specific moments of time (with a certain order) and this is meaningful whether the temporal sequence is *continuous* or *discrete.* However, change can be of a continuous kind if the characteristic to undergo change forms a continuous sequence that is correlated with a corresponding continuous temporal sequence. Indeed, for a change to be considered as continuous, it is *neither* sufficient to have, (a) only a continuous sequence of the subject of the propositions alone[81] *nor* (b) a continuous temporal sequence on its own, for if one of the two correlating sequences is discrete in nature, this would render the change to be discontinuous. However, it should be

[80]Russell B., (1903), p.469.

[81]An example is the continuous sequence of spatial coordinates that is represented by the real number series.

pointed out that the converse is *not* true since the correlation between two continuous sequences does not imply that a change need necessary be continuous. This point will receive a detailed analysis below.

The essence of motion rests upon the fact that *different places are occupied at different times* by *an object.* From a dynamical viewpoint, the relevant purpose of the object is to establish a correlation between all moments of time and some points of space in order to generate functional dynamical laws.

It must be once again emphasized that up until now there exists no necessary demand for the notion of continuity in the definition of motion based on Russell's criterion for change. In actual fact, this observation has been confirmed by Russell in the *Principle of Mathematics*[82],

> ***The motion is continuous** if the correlating relation R defines a continuous function. It is to be taken as part of the definition of motion that it is continuous... **This is an entirely new assumption**, **having no kind of necessity**, but serving merely the purpose of giving a subject akin to rational dynamics.*

As the notion of continuity is neither a formal requirement for *change* nor *motion* but merely serves as an additional assumption, it is important to assess the impact of such a supposition. In particular, we ask what indeed is the role played by the notion of spatiotemporal continuity? We *infer* from the spatiotemporally continuous motion of an object that it is the *one and the same* object that moves through spacetime. In other words, spatiotemporal continuity is brought in to ensure that the object in motion has an *identity over time.* But is this inference a sound one?

Consider a moving billiard ball that collides head-on with one that stays at rest and promptly sets the latter into motion upon impact. There, what is displayed before our very eyes are continuous spatiotemporal tracks. Indeed, it is this perceptual impression of the continuous movements of the balls that lends us the confidence and belief in the causal inference from the moving ball to the collision and from

[82]Russell, B. (1903), 473.

the collision to the subsequent setting into motion of the stationary ball. To quote Elizabeth Anscombe, "*a causal explanation would be it (an object) went along some path from A to B*".[83] We are confident that the moving ball is the legitimate cause of the subsequent motion of the stationary one. Especially, if one thinks along the line of Anscombe, then the causal influence is made possible by a physical connection provided by the "*same*" object - the moving ball.

But is spatiotemporal continuity alone a sufficient condition for causal inference via a physical connection? Imagine, however, the somewhat miraculous scenario where right after the moment of impact, the previously stationary ball vanishes and in its place (at the same spatial point) pops into existence another ball which resembles the former in so remarkably every detail that *no observable difference* is detected. Are we to conclude that this is the same ball as the original granted that this "other" ball is to move off in the very same direction as the original one would have done, and with just the correct amount of momentum and energy? Given the knowledge that what appears to be the same ball is in actual fact not, we now ask whether we still find ourselves confident to make the causal inference of the "moving ball setting the one initially at rest into motion"? I think not. We have here an instance of spatiotemporal continuity but *not* physical connection. It seems plausible that the answer centers on whether additional information concerning the rather strange acts of disappearance and reappearance of the ball is accessible to us or not[84]. If we are not furnished with background knowledge of such kind, as it seems we never are, the spatiotemporal continuous paths of the balls would deem to be good reason for us to infer that it is the one and the same ball.

Or we may also imagine a world where there exist those rather strange fantasy machines like "table canceller" and "table creator" as thought up by Sidney Shoemaker[85]. The "table canceller" is supposed to make tables vanish instantly at desired

[83] Anscombe, G.E.M. (1974), p.150.

[84] For the observation of the spatiotemporal continuous motion of the ball may arise from either the scenario of *being one and the same ball* or *being two indiscernable balls.* Since both of these scenarios cannot be distinguished by spatiotemporal continuity, one therefore is not allowed to assert that the ball is, in fact, one and the same.

[85] Shoemaker, S. (1979). Reprinted in French, P.A., Uehling, T.E. Jr. and Wettstein, H.K. (eds.) (1984), p.326.

locations and the "table creator" is to perform the reverse task of instantly creating tables at specified locations. Now consider a table standing in the middle of a dinning hall and the "table canceller" is set to cancel this very table whereas the "table creator" is set to create a table, that replicates the original in every aspect, to appear at the same spot. When the buttons on both machines are pressed simultaneously, an observer would not notice that the "original" table has been *replaced* by the "new" table. The stages of both the original and the new tables are in spatiotemporally continuous succession and display remarkable qualitative similarity and resemblance.

However, one would definitely not regard the stages of the "original" table and those of the "new" one as the successive stages of *one* persisting table *if* granted the knowledge of the presence of both the "table canceller" and the "table creator" machines and the fact that both machines were set to work at the same time. Without this knowledge, a spatiotemporally continuous sequence of events, which possesses so striking a qualitative similarity seems sufficient for one to conclude that there is indeed *one* persisting object. Here again, although spatiotemporal continuity does not warrant a physical connection, it seems a good enough reason for inferring that it is the one and the same table all along.

The proper relation between the notions of spatiotemporal continuity and causal transmission can be established as follows:

A: Being the same object

B: The assumption of spatiotemporal continuity

C: Causal transmission (as a kind of motion)

So the argument goes as follows:

(P_1)	**C→A**	Being the same object is necessary for causal transmission.
(P_2)	**A→B**	Spatiotemporal continuity is necessary (on an empirical level) to determine whether it is indeed the same object.

Therefore, **C→B** Spatiotemporal continuity is determined to be necessary for causal transmission .

On the physical front, as one speaks of the motion of an object, one does, albeit in an implicit manner, make the assumption of the object being one and the same. And it seems that the observation of spatiotemporal continuity in the motion of an object grants us the confidence and justification for such doing. Even though, of course, on the empirical level, one still cannot be certain whether the scenario of being "one and the same object" obtains or not.

Therefore, on the basis of physics, the spatiotemporal continuous motion is the connection between events in spacetime[86]. And this connection is what we deem as the *causal* connection on a singular level. This is how the notion of spatiotemporal continuity is being tied to that of causation in the context of classical physics.

It should therefore come as no great surprise that the issue of spatiotemporal continuity has much to be blamed for the causality crisis brought on by the quantum revolution. In non-relativistic quantum mechanics, one finds discontinuous leaps in the positions of a microscopic object while the characteristic of space and time as the respective three-dimensional and linear continua enters the theory at a most fundamental level. I have especially in mind the quantum mechanical phenomena of "tunneling" where a particle is confined within a given potential well that is classically forbidden for it to overcome since its total energy is lower than that of the well. Quantum mechanically, the particle may nevertheless be found to be able to "tunnel" through the well and appear on the other side of it. There is the pressing issue of how the particle gets from a location within the potential well to one outside *without traversing the intermediate positions*! These discontinuous leaps give a glimpse of the difficulty quantum theory poses for our much cherished intuitive correspondence between continuity and causation. This is due essentially to the absence of spatiotemporal continuity, an absence of a well-defined trajectory. Our intuitions are not tutored to be at ease with the notion of discreteness.

[86]As this physical connection needs only be satisfied on one occasion for singular causal inference, it is deemed a *sufficient* condition for causation on a singular level to obtain.

Now how can one even be certain that the particle found outside the barrier is *one and the same* that had been imprisoned inside it some moments ago? How then are we to suppose the continual existence of the particle?

The continual existence of the particle requires it to be existing in space as well as in time. However, it is sufficient for the claim of an object being one and the same if it enjoys only continual existence in time. Consider a simple case where an object, situated at a location A, vanishes at time t_A and an "indiscernable" object, resembling the first in every imaginable aspect, appears at a different location B at a later time t_B (Figure 4-6).

Figure 4.6:

As no object is to be found anywhere in space between these two times, one is unable to lodge the claim of *the object being one and the same as substantiated by the continual existence in space*, for it is obviously not the case. One could, at best, resort to the continual existence in time of the object in order to carry through the claim of the object being one and the same. This is possible because we may entertain the thought that the object may still exist during the times between t_A and t_B, not in physical space[87] as we perceive it, but rather some other manifestations of it, for instance.

[87]Such a maneuver is bound to attract controversy, for although we have posited the continual existence *or* the persistence through time of the object, what meaning may be attached to it being "existing in time *but not* in physical space"? Is it tenable to segregate existence in time from that in physical space? Events occupy specific spacetime locations as and when they occur, but some events have the tendency to remain in one's memory over a period of time, usually ones that had brought great delight or tragedy, long after they had ceased existing in space. However, I should refrain from such discourse which concerns the mind and soul but rather restrict my reference to the phrase "existence in space" in a much narrower sense to ontological entities in the physical world. Physical events, objects and processes all exist in both space and time. Therefore, the possession of a location in *both* space and time is a vital precondition of physical existence. This is the picture presented to us by non-relativistic classical mechanics.

The absence of a complete spatiotemporal record of a tunneling particle in a potential well would seem to run contrary to our supposition of continuance in existence. If such a continuance is found to stand on shaky grounds, I ask, with what right are we to speak of the particle being "*one and the same*" object?

One notes that, as I have argued, the continuous path of a physical object in spacetime is vital to the claim that it is in fact the *same* object which takes up the different positions at different times. It is therefore essential and inherent in the classical idea of matter that the object persists through time and retains its identity when it is in motion. This is why the locution "the same object" loses its entire meaning when an "object" *hops* or *leaps* through space without visiting the intermediate positions. This is why we ought not to speak legitimately of (classical) "motion" when the object takes a leap through spacetime, recalling that our object, the material point, has been brought in to impose important kinematical conditions for capturing the essence of motion. This is how the evaporation of spatiotemporal continuity has rendered the concept of motion empty.

The difficulty, I maintain, is due to the fact that at least on a classical level, the idea of being the same object, which is crucial to the related ideas of motion and causal inference, is itself found to be parasitic upon the notion of continuity. But must it be so?

As we have seen, the notion of continuity enters into the "*at-at*" theory of motion as an *extra assumption.* What is, however, of vital importance is that motion presupposes the presence of ***a*** material thing or matter. By virtue of its occupation of different points in space at different times the material thing provides the correlations between space and time that represent states of rest and motion. In order to fulfill its role as a correlator, *this material thing is to possess the feature of persisting through time and maintaining its identity during motion.*

This is indeed the kind of identity over time we seek for causal processes: the stipulation that it is the one and the same object that is in motion. This is required before one may meaningfully speak of the notions of an object in motion and that of the transmission of causal influences. This supposition releases one's thought from spatiotemporal continuity which is incidental to classical physics, one can begin to speculate on the kind of discrete motions of objects as portrayed in quantum physics.

I would like to end this section on a different note. Salmon's process theory has also been criticized for its incapability to deal with cases of action-at-a-distance like Newtonian gravity; for there is seemingly no spatiotemporally continuous connection existing between the two objects involved in such an interaction. There appears to be an exchange of momentum without any actual spacetime encounter of the two objects. One finds an analogous situation in classical electromagnetism that had proved to be equally perplexing before Maxwell, who duly resolved the conundrum by introducing the model of an electromagnetic field, which acts as the mediator of the forces between two charged bodies. Electromagnetic fields prevail through spacetime and thus form a good example of causal processes[88]. Geodesics of objects, from the standpoint of Einstein's Theory of General Relativity provide further support, as Salmon argues[89], "*...any such falling object is a single causal process that is free from interactions.*"

To be sure, an object is in free fall unless some other force *other than gravitation* is exerted on it. Consider the simple case of a ball resting on a table top. The reaction exerted by the table prevents the ball from falling. The point where the ball touches the table top is an intersection between the two processes of the ball and the table top, which involves an exchange of momentum in a continuous manner. We are thus faced with the difficulty that the ball (or the table) cannot be regarded as a causal process on the **CQT** definition.

But as with the continuous interactions with the air-molecules in its surrounding, the energy exchanged in these continuous interactions with gravity is considered as peripheral to the major "chunck" of energy that constitute the ball (our causal process), which is transmitted by the ball by virtue of its identity over time.

[88]It may be argued that the object whose worldline is causal process has finite spatial extent, and as such, electromagnetic fields would be disqualify. However, as we shall see, by the adoption of the "history" view in place of the worldline view would enable us to count electromagnetic fields as causal processes.

[89]ibid., p.465.

4.4.2 A Critique on the "Worldline View" of Causal Processes

In the present section we endeavour to acquire a better understanding of how causal processes come to get bound up with the worldlines of objects. The concept of a *worldline* defined as the collection of events happened at an object (its history) was introduced into physics through Einstein's special theory of relativity (**SRel**). **SRel** rejects the view of absolute space and time in favour of the four-dimensional *spacetime* continuum. In this treatment, different observers assign different sets of spatial and temporal coordinates to a region of spacetime and each one of these coordinate systems corresponds to a different frame of reference. There are different descriptions of the *same* event for observers from different perspectives because of their different choices of the set of spacetime coordinates - the different frames of reference. In the theory, events are the ones that assume a concrete reality, for although described by different spacetime coordinates, they are the representations of the actual occurrences in the world and these are indisputable, empirical facts, irrespective of which frame of reference an observer happens to be in.

It was realized by Hermann Minkowski, a German mathematician, in the early 1900's[90] that all the results of **SRel** can be represented geometrically in the so-called Minkowski spacetime diagram. In this diagram, the vertical-axis represents the time coordinates while the horizontal-axis represents the spatial coordinates[91], and the spacetime locations of the occurrences of events are unambiguously denoted by *points*[92] on the diagram. All the events happening to a physical system - its *history* - form a continuous sequence of points in this diagram. This continuous sequence of points taken collectively is the *"worldline"* of the system. It is important to recognize the fact that the concept of the worldline is defined *with respect to a single physical system or object.*

Most textbook definitions of a worldline focus on this point and the aspect of

[90]Minkowski, H. (1908). Reprinted in Lorentz, H.A. et al. (1923), p.84.

[91]This is usually understood to describe the three-dimensional location but since in practice, the diagram is presented on a two-dimensional plane, so the three spatial coordinates are "collapsed" into a one-dimensional coordinate without any loss of generality.

[92]Minkowski calls these "world-points", ibid., p.76.

spatiotemporal continuity that is associated with it. As an illustration, we need only examine a few of the definitions below,

> ***The history of a particle consists of a continuous sequence of events.*** (Synge[93])

> ***Lines describing the history of point objects are called worldlines, or spacetime trajectories.*** (Mermin[94])

> ***We can construct a simple x-t coordinate system, on which we draw "worldlines" showing the development of the system in space and time. The worldline of any given particle is just a graph of its position as a function of time, it provides a complete picture of the history of the particle as observed within a given frame of reference.*** (French[95])

> ***The track of a particle through four-dimensional spacetime is called its worldline.*** (Eddington[96])

> ***We describe the world by listing events and showing how they relate to one another...Now we turn to a whole chain of events, events that track the passage of a particle through spacetime. Think of a speeding sparkplug that emits a spark every meter of time read on its own wristwatch. Each spark is an event; the collection of spark events forms a chain that threads through spacetime, like pearls. String the pearls together. The thread connecting the pearl events, tracing out the path of a particle through spacetime, has a wonderfully evocative name: worldline. The sparkplug travels through spacetime trailing its worldline behind it.*** (Wheeler and Taylor[97])

[93]Synge, J.L. (1972), p.9.
[94]Mermin, N.D. (1968), p.157.
[95]French, A.P. (1968), p.74.
[96]Eddington, A. (1978), p.78.
[97]Wheeler, J.A. and Taylor, E.F. (1992), p.143-144.

and last but not least, from the grand master himself,

> ***To avoid saying "matter" or "electricity" I will use for this something the word "substance". We fix our attention on the substantial point which is at the worldpoint x, y, z, t and imagine that we are able to recognize this substantial point at any time. Let the variations dx, dy, dz of the space coordinates of this substantial point correspond to a time element dt. Then we obtain, as an image, so to speak, of the everlasting career of the substantial point, a curve in the world, a worldline, the points of which can be referred unequivocally to the parameter t from $-\infty$ to $+\infty$.*** (Minkowski[98])

The above cited passages call our immediate attention to two important aspects of a worldline. First, a worldline is a *continuous* series of events occurred at a physical object or system. Second, the spatiotemporally continuous characteristics of a worldline is described by its dynamical trajectory in spacetime and this trajectory is governed by the laws of classical mechanics. This second point is evident from the expressions "...a graph of its position as a function of time" (French, op. cit.) and "the track of a particle through four-dimensional spacetime" (Eddingtion, op. cit.).

One simple but essential point to notice is that the notion of a worldline is defined for a *single object* (e.g. a particle, a sparkplug or a substantial point). We shall see in a little while that this seemingly trivial observation is in fact of paramount importance, as it is through this simple fact that the notion of identity over time becomes associated with what I shall call the "worldline view" (**WV**) of causal process.

A physical object or system has yet to receive a precise specification. From the above cited passages, a physical object is one that is often construed as a particle or in the usual abstraction of classical mechanics, a material point - one that possesses energy or momentum. The object acts as the *carrier* of this energy and momentum either when it is at rest or in motion. Minkowski speaks of an object as a "substantial"

[98]Minkowski, H. (1908). Reprinted in Lorentz, H.A. et al. (1923), p.76.

point, one that is understood to be in possession of matter and energy. The sparkplug of Taylor and Wheeler is also one which possesses energy and momentum, and these quantities are transmitted along the spacetime trajectories of the sparkplug by virtue of its motion.

In summary, a worldline in the terms of physics is defined with respect to a material point particle[99]; one which is in possession of energy and momentum. As the particle moves along in spacetime, this energy and momentum is transported from one region of spacetime to another.

It ought to be put into strong emphasis that it is by means of the occupation of successive spacetime points by *the one and the same object* that energy and momentum of the object can be meaningfully said to be transferred from one spacetime locale to another. This is because we may only speak of motion in a legitimate manner if we refer to the same object. Put simply, it is part and parcel of the definition of motion for the object to remain the *same* object throughout its motion[100]. Philosophically speaking, this concerns the *identity over time* of an object. The question arises as to how might one be sure that the sequence of successive events - the occupation of a different spatial point at each time by an entity - corresponds to the successive stages of *one* persisting object? Not any odd sequence will do, for that may result in one that consists of the stages of different objects. Intuitively, the successive stages of a persisting object must somehow hang together in a distinctive way. We conceive of some sort of special connection amongst the stages or events that belong to a persisting object.

In a similar vein, a worldline is not just any odd continuous sequence of events in spacetime but one in which the constituent events are linked together in a way that they can be rightfully regarded as the successive stages of a single persisting object. This special relation of connection, in physical terms, as I would argue, is the transmission of energy or momentum from one event to the next. Transmission of energy and momentum thus provides the connection that obtains between any two

[99]Without the loss of generality, I shall use the terms "particle" and "object" interchangeably on the understanding that they both refer to a point-particle in the usual sense of abstraction.
[100]Refer to the argument in Section 4.2.2.

successive stages of a persisting physical object. It aims to capture the fact that one event is the source of the next within this special sequence.

For a sequence of events to correspond to the successive stages of a single object, each of these events would represent *the* object as occupying a certain spatial location at a certain time. This is, however, not too helpful since one is indeed supposing it is "the" (or *one single*) object that takes up those locations in spacetime. We shall therefore retreat to a weaker claim: *given a spatiotemporal sequence of events, each of these events represents an object as occupying a certain spatial location at a certain time.* This statement encompasses the possibility that each of these events may correspond to *an* object (or *any* object for that matter) - similar to the original one in every visual detail and possessing dynamically the right amount of energy and momentum, to be at the appropriate location at the appropriate time - in order to make it looks as if it is the one single object that is in motion. Of course, to opt for the weaker claim obliges one to subscribe to the notion of "possession" as the correct criterion, rather than the more elaborate concept of transmission. But is the notion of possession a sufficient criterion for our purposes?

At first sight, to investigate the issue of the identity over time of an object seems entirely a philosophically motivated pursuit that is of no physical relevance; since physics proceeds with the supposition of the one and the same object as the prerequisite condition for motion. It is, however, far from being so. Recall our familiar example of the light ray and the spot of light it cast on a far wall in the astrodome. "*The* spot of light" which evolves around the wall is an illusion: it is not the same spot that moves but rather, a sequence of all similar spots (in the sense that they are produced by the same physical mechanisms), each appearing at the right place at the appropriate time to create the image of *one* moving spot. In this instance, this continuous sequence of events corresponds not to the history of a single object, but rather, to a "pseudo-sequence" constituted out of different objects. Despite the fact that energy is found to appear (be possessed) at every stage of the pseudo-sequence, there is no actual movement of the *same* energy[101] from one place to another. In

[101]Here, we take the view that the rest mass of an object, for example, is a form of energy by virtue

contrast, the *same* energy of a single object - that of the photon - is transmitted from one place to another within the light ray; it captures the intuition that one event is the source of the next.

It is therefore reasonable, for a theory of physical causation, to associate *causal processes* with *physical objects.* The physical relation that connects the stages of a physical object provides the causal connection. Pseudo-processes are non-physical because of the lack of a physical connection amongst the stages of the process. So worldlines - which are essentially the histories of physical objects - constitute causal processes.

Both the theories of Salmon and Dowe place no restriction on the notion of an object at the outset. When it comes to be confronted with the decision as to what kinds of objects are physical (or causal) and what are non-physical (or non-causal), Dowe falls back on the criterion of the "*possession of a conserved quantity*" (as supplemented by identity over time) by the object in question while Salmon opts for the "*transmission of a conserved quantity*". In other words, both find it necessary to re-capture the physical relation by their respective criteria of "possession" and "transmission" as a trade-off to the "unqualified" notion of an object.

Being liberal in employing the notion of an object, Salmon and Dowe allow the assignments of "worldlines" to non-physical objects. This is a philosophical move prompted by the desire to distinguish between processes with none of its successive stages bearing any physical relation to each other, and those whose stages are genuinely linked by physical relations; given that both types of processes exhibit a kind of constancy in their dynamical motions. The physical relations they aim to capture have to do with, of course, the transport of physical quantities like energy or momentum along the spacetime trajectories of physical objects.

The term "worldline" has, as we have seen, a more specific meaning in physics than what Salmon and Dowe have both intended. Let us recapitulate briefly the four main characteristics of a worldline of an object as conceived from the viewpoint of physics,

of $E = mc^2$, which then sanctions us to speak of an identity over time for this chunk of energy.

(1) the worldline is a sequence of events that corresponds to the history of a *physical* object - an entity that is in possession of energy and/ or mass (which include both massive and massless entities).

(2) it is a *continuous* sequence of events happening at "*the*" object and it supposes an identity over time of a persisting object[102].

(3) the worldline of an object need not be confined to an object being in uniform motion. It simply describes the motion of the object[103].

(4) the different "worldpoints" that constitute the worldline are physically connected in the sense that one serves as the source of the next as the object, the carrier of dynamical quantities like energy and momentum, moves from one spacetime point to the next. The physical connection is provided by the *movement* of "one and the same" object. It is indeed this physical connection that makes possible the correspondence between worldlines and causal processes.

The term "worldline" as employed by Salmon and Dowe in their causal theories have been stripped-off all these fine physical properties but left only with the feature of it being a sequence of points in spacetime. Starting with this bare minimum, their strategy then consists in the recovery of all the other remaining physical details above by the respective criteria of "transmission" and "possession" of conserved quantities. An immediate question that is of interest to us is: are these criteria capable of recovering the physical essence of a worldline? We now consider this question in relation to each criterion in turn.

Dowe: "the possession of a conserved quantity"

On an empirical level, the movement of a quantity of energy and momentum by means of the motion of a physical object from one spacetime locale to another seems

[102]This notion of identity over time as encapsulated in the concept of an object with respect to the worldline is quite different from Dowe's SI. While SI requires the object to be wholly present at each time and can hence, in principle, admit indiscernable objects to be the stages of a single "moving" object (i.e. one that has spatiotemporal stages). Here, our requirement that being the one and the same object is a prerequisite for motion immediately rules out indiscernable objects.

[103]For example, it may describe the motion of an object under a constant conservative force.

to, *prima facie*, consist in no more than the *possession* of these quantities successively at the appropriate spacetime points along the trajectory of the object. And so the condition of "the possession of a conserved quantity" conforms to feature (1) above of the physical worldline.

Counterexamples such as the series of successively illuminated light cells and Hitchcock's shadow that moves over a uniformly charged metal plate, seem to pose too tremendous a difficulty for the criterion of possession to overcome. Both examples show that there can be processes that are in possession of conserved quantities and thereby satisfying Dowe's condition of being causal, but yet whose stages are not physically connected. Hence, the condition cannot apply universally to distinguish causal processes.

Despite the failure to capture the important feature of a physical connection (clause (4) above), it is still useful to see how successful has Dowe been with **CQ** in recovering the other physical characteristics.

As regards the identity over time of a persisting object (clause (2) above), Dowe needs, as he admits, an additional premise of identity (e.g. **SI**)[104] to dismiss cases like the above counterexamples (time-wise gerrymanders for instance). But I maintain that in order for the statement "a light spot (or a shadow) is in motion" to be meaningful, there is already the supposition of an identity over time of a persisting object.

Dowe has spoken specifically about his theory not requiring the object to possess a constant amount of a conserved quantity. Practical cases of continuous interactions of the object with its environment support this flexibility in his **CQ** theory. Physical objects with respect to which worldlines are defined are not restricted to the possession of a constant amount of a conserved quantity; that is, they are not constrained to uniform motion. Therefore, Dowe's condition agrees with clause (3) above.

[104]However, in his most recent writing, Dowe no longer wish to commit himself to **SI** but maintains that he prefers to leave the notion of identity of an object as primitive in his theory (Dowe, P. (2000), p.101-107).

The term "worldline" as understood in physics, refers to a spatiotemporally continuous sequence of events, as contrary to Dowe's argument otherwise[105]. Granted what Dowe really has in mind is a sequence of events corresponding to the successive stages of a moving object[106], it would be far more cogent to abandon the use of the term "worldline" and adopt a comparatively neutral term, which will serve to convey the idea of a sequence of spacetime points but without committing to all the fine details of a worldline (see Section 4.5).

Salmon: "the transmission of a conserved quantity"

First of all, Salmon's theory lacks but also stands in need of an identity over time of an object, for the same reason that it is not possible to speak about the motion of any object - both physical and non-physical - if that object is not considered as one persisting object.

He makes no scruples about the aspect of spatiotemporal continuity of his causal processes. In fact, as one may see, he has indicated at several places that he deploys the concept of spatiotemporal continuity to establish the identity over time of a persisting object (Section 4.4.1), in accordance with the usual empirical analysis of the concept of identity.

Although spatiotemporal continuity stands as a fundamental feature of Salmon's causal processes, not all spatially and temporally continuous processes are causal: one needs only consider spatiotemporally continuous processes like a moving shadow or a moving light spot across the wall. Salmon realizes the important distinction that divides causal and non-causal processes and has brought his analysis closer to that in physics by appealing to the notion of transmission (clause (4) above).

In Salmon's theory, transmission is explicated as the possession of a conserved quantity (therefore, clause (1) is satisfied) *without any further interaction*, except for the one which facilitates the acquisition of the said quantity in the first place. This is done in the hope of capturing the fact that genuine physical processes sustain their own existence and dynamical makeups so that an earlier stage of it provides the source

[105] Dowe, P. (1995), p.332.
[106] On which an identity over time is presumed.

of a conserved quantity to be possessed by the next to come.

Transmission has also a spatiotemporal continuous nature, as Salmon has intended through the "*at-at*" theory. The notions of spatiotemporal continuity and transmission coalesce to form the essence of a genuine persisting physical object (clause (2) above).

In spite of all these niceties, there is, however, one caveat. The notion of transmission encompasses only a fixed amount of conserved quantity is to be transmitted. An object vacates one spacetime point as it takes up occupation of the next and at the same time, the conserved quantities it possesses leave one spacetime region and get delivered to the next. This is what is usually conceived as an essential part of the meaning of motion. As we have explained, it is not entirely feasible to restrict oneself to the stringent requirement of the transmission of a constant amount of a conserved quantity because of the non-practicality of completely closed systems in nature. The insistence of the transmission of a constant amount of a conserved quantity may thus be criticized as being too idealized.

For Salmon though, it remains essential to make such an abstraction in order to highlight the fact that any changes in the value of the conserved quantity in question is to be identified as the consequence of a causal interaction.

Notwithstanding this basic intuitive insight, the notion of a worldline as given in physics has not been restricted to the motion of a free particle - one that is free from the influence of interactions. A worldline may well represent the motion of a massive body under the influence of an applied force, most likely in which case, a changing amount of momentum (or energy) is transmitted along (clause (3) above).

Once again, as Salmon wishes to "build on" the concept of the "worldline" (a comparatively minimal notion that as it is defined by physics), it would also, as we shall immediately come to in the next section, be appropriate to trade the term "worldline" for a more neutral term.

4.5 The "Histories View" of Causal Processes: A Modest Emendation

In the last section we bring to the reader's attention that the definition of a worldline in physics refers not indiscriminatingly to any sequence of events with respect to an object in spacetime, but rather with quite a specific meaning attached to it.

Both the approaches of Salmon and Dowe take the intuition of a "sequence of events" of any object (either physical or non-physical) and then impose on their respective conditions of the "*transmission of a conserved quantity*" and the "*possession of a conserved quantity*" to determine whether such a sequence does indeed represent a causal sequence or not.

To eliminate the confusion arising from the disparity on the notion of a worldline as understood from the physical viewpoint and the meaning as intended by Salmon and Dowe, I propose a slight adjustment to the terminology: the adoption of the comparatively less restrictive term "***history***" in place of "*worldline*". I say less restrictive because a history is here understood to be a sequence of events that *may or may not* bear causal relations to each other. And the stages of a history need only be related in a temporal manner. In particular, one may now conceive of *continuous* as well as *discrete* histories.

Hence, I would suggest a slight emendation to the definition of a causal process as follows,

> ***A causal process is the history of an object that transmits a conserved quantity.***

With a "*history*" understood to be "*a sequence of events in spacetime*". On this definition, the worldline of a material particle qualifies as a causal process. Shadows and moving light spots have histories, their stages form spacetime sequences but there is no transmission of conserved quantities along them. These histories are *not* worldlines. There are both causal and non-causal histories in the world.

An object can be anything found in the ontology of science or common sense, but it must have an identity over time in the primitive sense that one may only meaningfully speak of the motion of an object if one refers to "*the one and the same*" entity[107]. On this definition, moving spots and shadows are objects, although their tracks through spacetime are not causal processes as ruled out by the condition of transmission.

As in Salmon's **CQT**, the stipulation that transmission only obtains if the conserved quantity in question is to be found at every spacetime point *without further intersections that involve exchange of a particular conserved quantity* at any of these points remains valid. In cases of pseudo-processes, the presence of further interactions is a prime requirement for the process to survive; and there are many objects, each coming into existence at the spacetime points where those interactions occur.

With a causal process now being construed as the history of a sequence of events with respect to an object that involves the transmission of a conserved quantity, our previous definition of a causal interaction still applies,

> ***A causal interaction is an intersection of causal processes that involves exchange of a conserved quantity.***

The clauses form the ***History Conserved Quantity theory with Transmission*** (HCQT).

The move from "worldlines" to "histories" brings the analysis closer to the intended strategies of both Salmon and Dowe. With the neutral notion of history as merely a sequence of points in spacetime, the analysis is effectively free from any presumed meaning attached to the term worldline and one may now impose an *extra* criterion of transmission for a history to enable it to qualify as a genuine causal process. This has also acquired the added benefit of being in line with the physical point of view.

Another important reason for favoring histories over worldlines in the definition for causal processes arises, of course, from the issue of spatiotemporal continuity. It has been pointed out that worldlines as conceived from the standpoint of physics refers to the spatiotemporally continuous trajectories of material objects. Histories,

[107]This supposition allows one to speak about *an* object in discrete motion.

on the other hand, denotes sequences of events in spacetime that may or may not be continuous[108].

Spatiotemporal continuity has been the main weapon engaged by critics of process causal theories to claim that these theories not capable of dealing with physical processes in the quantum domain where discreteness in the dynamics of a physical system is a distinctive feature. We shall show in Chapters 6 and 7 how one may still be able to speak meaningfully about the history of a physical system in the context of quantum theory.

There are more subtle considerations for the preference of the term "*history*". To illustrate the extreme, one approach to quantum gravity is to take the view that spacetime itself is a discrete entity that is endowed with a temporal structure[109]. From such a perspective, it remains sensible to speak about a history as a sequence of events, while on the other hand, calling such a sequence a worldline would be blatantly inappropriate. The change of terminology thus allows one to investigate the question of causality in these wider contexts of modern physics.

4.6 Interlude

There is one connecting theme for the previous chapters: the idea of a spatiotemporally continuous path. This idea forges an intimate link between *physical* and *causal* connections. In the language of process theories, causal processes, which are essentially the continuous motions of material objects, transmit causal influences across spacetime regions. This transmission of causal influences is regarded as the causal connection among physical events. The spatiotemporally continuous character of a causal process requires that it be present *at* every point in space and *at* every instant of time in the intermediate spacetime region between the two events it is supposed to connect.

So any program of causation that is founded on the notion of continuous paths in

[108]It is the very task of our analysis to lay down sufficient conditions to help identifying which of these sequences are indeed causal processes.
[109]See for example, Bombelli, L. et al. (1987).

spacetime is doomed to the most miserable fate in situations where spatiotemporal continuity seems untenable. The quantum revolution has given us ample reasons to believe that the classical notion of a definite spatiotemporal trajectory can no longer be preserved. With this untimely passing of spatiotemporal continuity, we ask: what is left of causality?

In their struggles to recover old, cherished classical intuitions from the new quantum theory, the founding fathers of quantum mechanics addressed the same issue, as can be seen from the following excerpts[110],

> ***The very nature of the quantum theory thus forces us to regard the spacetime coordination and the claim of causality, the union of which characterizes the classical theories, as complementary but elusive features of the description, symbolizing the idealization of observation and definition respectively.***

and commenting on the discrepancy between Einstein's position and that of his own[111],

> ***Yet, a certain difference in attitude and outlook remained, since, with his mastery for coordinating apparently contrasting experience without abandoning continuity and causality, Einstein was perhaps more reluctant to renounce such ideals than someone for whom renunciation in this respect appeared to be the only way open to proceed with the intermediate task of coordinating the multifarious evidence regarding atomic phenomena, which accumulated from day to day in the exploration of this new field of knowledge.***

Compared to Bohr, Heisenberg took an even bolder attitude towards this issue of the continuous versus the discrete[112],

[110]Bohr, N. (1928). Reprinted in Wheeler, J.A. and Zurek, W.H. (eds.) (1983), p.89.

[111]Bohr, N. (1949). Reprinted in Wheeler, J.A. and Zurek, W.H. (eds.) (1983), p. 14.

[112]Heisenberg, W. (1927). Translated into English as "The Physical Content of Quantum Kinematics and Mechanics" by J.A.W. and W.H.Z. (1981), reprinted in Wheeler, J.A. and Zurek, W.H. (1983), p.62.

> ***The physical interpretation of quantum mechanics is still full of internal discrepancies, which show themselves in arguments about continuity versus discontinuity and particle versus wave.***

and he continues[113],

> ***If one considers, for example, the notion of a particle in one dimension, then in continuum theory one will be able to draw (Figure 4-7a) a worldline x(t) for the track of the particle (more precisely, its center of gravity), the tangent of which gives the velocity at every instant. In contrast, in a theory based on discontinuity there might be in place of this curve a series of points at finite separation (Figure 4-7b). In this case it is clearly meaningless to speak about one velocity at one position (1) because one velocity can only be defined by two positions and (2), conversely, because any one point is associated with two velocities.***

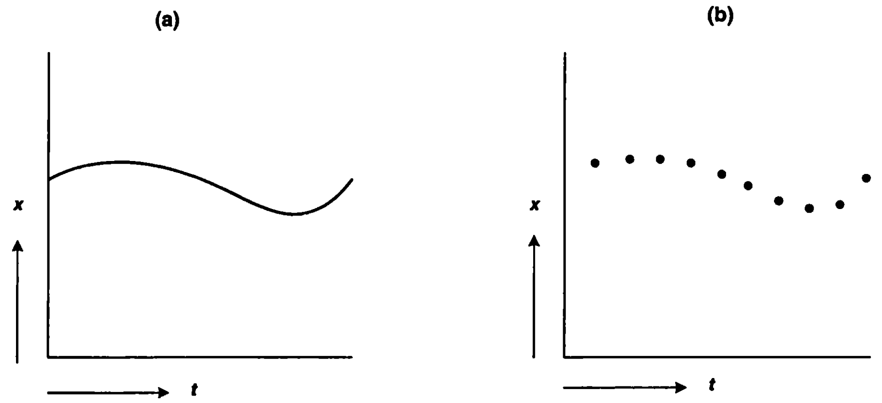

Figure 4.7: Continuous versus discrete (diagrams reproduced from Heisenberg's paper, in Wheeler-Zurek (1983), p.63.).

Causal influences, in the form of dynamical conserved quantities such as energy and momentum, get transported from one spacetime locale to another by the physical

[113]ibid., p.63.

objects that possess them. The lack of definite object paths now makes it unclear how these quantities may be transmitted.

The difficulty has to do mainly with the fact that our "classical" notion of causality - which has been founded on the idea of the spatiotemporally continuous motions of physical objects - is employed in the attempt to understand quantum phenomena. In the effort to preserve this notion of causality, it is natural to recover this essence of spatiotemporal continuity within the quantum formalism.

Despite the absence of continuous spacetime paths of the physical objects, it is often thought that all is not lost for the notion of spatiotemporal continuous evolution in the quantum world. What evolve continuously in spacetime are no more the concrete physical objects we perceive, but in their places come instead some abstract mathematical construct - the wavefunction ψ - the modulus square of which provides the *probability of locating an object at a certain point in spacetime.* The notion of an object being at a certain location in spacetime is now only meaningful insofar as an actual measurement is concerned. It is only through a measurement on the physical system that we "actualize" a value for a dynamical observable. The series of points depicted in Figure 4-7(b) above corresponds to these instances of measurements when the acts of measurement themselves contribute to "fixing" these values. Between successive measurements, one is not permitted to speak of the system as having definite values for a dynamical variable. As so, measurements thus furnish us with a sequence of discrete points with no possibility of interpolating between any two with a continuous spectrum of definite values.

Without continuous spectra of definite values for dynamical variables, we are left with no clear notion of a path of an object in physical spacetime. However, at every spacetime point between two measurements, we have a probability of locating the particle upon being measured. Moreover, this probability can be shown to evolve in spacetime as governed by the Schrödinger equation.

Granted that the notion of causal transmission is dependent upon the existence of a spatiotemporally continuous path[114] and that now there is the continuous spatiotemporal evolution of the probability function, might we not entertain the possibility of recovering a notion of transmission of a physical quantity via this evolution of the probability function? This suggestion, of course, immediately commits one to the conviction that the probability function is a legitimate property of the physical system under investigation. This would undoubtedly involve the interpretation of probability in quantum mechanics, a subject which is by no means uncontroversial.

In particular, one ought first to be clear about which specific view of ψ one is to adopt. For, an assessment of ψ as a legitimate property of a physical system is only meaningful with reference to an interpretation of it.

The view taken in this work is based on the simple observation that there are *two* levels of superposition at work in nature, in both classical and quantum physics: (1) the superposition of potentials and, (2) the superposition of the effects of these potentials.

Classically, these two levels coincide but the distinction becomes crucial in the quantum mechanical context. I shall argue that the superposition of potentials is fundamental to the understanding of the behaviour of ψ.

With this understanding of ψ, we are then able to proceed, in Chapter 6, to investigate if ψ may indeed be regarded as a legitimate property of a physical system, and the extent to which its continuous evolution in spacetime satisfies the "*at-at*" criterion of causal transmission. Then in Chapter 7, we shall discuss how our interpretation of ψ as the superposition of potentials fits naturally to the Feynman Path Integral formulation of quantum mechanics, which represents an appropriate description of causal processes in quantum physics.

To set the scene for these forthcoming discussions, Chapter 5 is devoted to the principle of superposition in quantum physics and how the principle makes contact with probabilities.

[114]This is because causal transmission in the manner of Salmon's "*at-at*" theory consists in a dynamical variable (like position or momentum for example) of the physical system assuming some definite value at every spacetime point between two events. In other words, a spatiotemporally continuous path of values for the dynamical variable.

Chapter 5

On the Way to Quantum Paths: The Superposition of Potentials

5.1 Introduction

This chapter is about the principle of superposition. It defends two central distinctions, each of which is often ignored or overlooked; and in both cases, conflating the distinctions can generate considerable confusion. The first begins from noting the simple distinction between effects and what brings them about, which I shall call "potentials"[1]. Though a simple distinction, it is one that is not always kept clearly in mind when it comes to discussions of the principle of superposition. In line with the distinction between effects and the "potentials" that produce them, I shall point out (Section 5.4.2) that there are two different principles of superposition - *superposition of potentials* and *superposition of effects.* In classical mechanics, with vector addition of forces (the "potentials") and vector addition of accelerations (the effects), we see both principles at once. In classical waves, we see only the superposition of potentials but not that of effects. In quantum mechanics, it is again the principle of superposition of potentials that matters.

The second distinction is that between the sum-rule of probability and the superposition principle. In many discussions of the two-slit experiment, it seems as if the sum-rule of probability is taken to express a special instance of superposition for the

[1] I use the word "potential" rather than "cause" because the former will also apply to probabilistic situations where one speaks of a tendency to generate a particular result. But I do not mean to refer to energy potentials.

setup in which photons are known to pass through one slit or another. I defend this point by looking at how probabilistic potentials should be conceived in the case of a classical die, and I conclude the chapter by contrasting a classical die, for which the sum-rule of probability holds, and the principle of superposition fails, and a quantum die, for which the opposite is the case.

In order to set the stage for the arguments in Section 5.2, I begin by introducing the feature of superposition in quantum mechanics via the Schrödinger wavefunction. This quantum nature of superposition is very different from its classical counterpart, which is discussed in Section 5.3. In Section 5.4, I illustrate the quantum nature of the principle of superposition by the two-slit experiment and we can see how the superposition of potentials is preserved in quantum phenomena even that on the level of effects fails.

5.2 Superposition and the Schrödinger Wavefunction

The beginning of the quantum era represents a curious conceptual struggle in the history of physics. In the transition from the classical to the quantum world, the issue of the continuous trajectories were fiercely debated.

Niels Bohr also found it necessary to inject the idea of discontinuous jumps in positions into the explanation of the stability of the atom. According to classical electrodynamics, the orbiting electron in an atom is destined to continuously radiate energy that would lead to its eventual spiraling into the nucleus. Bohr's quantum hypothesis consists in two parts:

(1) the electrons move around the nucleus only in a certain number of allowed orbits (states), each with a well-defined energy. That is to say that the energy of the atom is quantized and,

(2) the electrons only radiate when they jump from one orbit to another (not passing through any point in between). The radiation associated with such a transition is thus expected to come in quanta of light (photons). An electron may emit a

photon of energy $h\nu$ when making a transition from a higher energy level to a lower one; and conversely, it would jump from a lower energy level to a higher one upon absorbing a photon,

$$E_i - E_f = h\nu$$

with E_i and E_f being the energies of the initial and final stationary states respectively. The stability of an atom may now be explained by the assertion that if the atom finds itself in its lowest energy state (ground state), it cannot radiate and hence remains stable.

Bohr's model of the atom also provides a fitting explanation for the appearance of the line structures in the spectrum of the hydrogen atom (the Balmer series) that is in contradiction with the classical theory.

The picture of electrons in captivity within the atom only being permitted to reside in specific energy states (stationary states) has a strong analogy in classical wave phenomena. Stationary waves are a property of waves in confinement. As a guitar string is plucked, the stationary waves form a discrete pattern of harmonics. This led de Broglie to the inspiration that atomic electrons are, after all, confined waves that produce a spectra of discrete stationary states. This marks the inception of the idea of *wave-particle duality* and the famous de Broglie wavelength for "matter waves" λ:

$$\lambda = h/p$$

with h and p being Planck's constant and the particle's momentum respectively.

The route to the Shrödinger equation is now clear. If matter is capable of behaving as a wave, there should indeed be an accompanying *wave equation* to describe the evolution of such a matter wave. This is what Shrödinger proceeded to search for in 1926[2].

[2]Schrödinger, E. (1926). English translation in Schrödinger E. (1928) and Ludwig, G. (ed.) (1968), p.94-105.

Recall that the classical form of the equations describing wave phenomena is of a second-order differential[3] character:

$$\frac{\partial^2 y}{\partial x^2} = \frac{1}{\omega^2} \cdot \frac{\partial^2 y}{\partial t^2} \tag{5.1}$$

with ω being the speed of propagation of the wave under investigation.

The symmetry between the propagation of the wave in opposite directions (i.e. in both the $+\mathbf{r}$ and $-\mathbf{r}$ directions) implies that the solutions to these equations should take the general form of a plane wave propagating in the $\mathbf{r}$-direction with wavevector $\mathbf{k}$ and amplitude A:

$$y = A\, exp\left[i(\mathbf{k} \cdot \mathbf{r} - \omega t)\right] \tag{5.2}$$

Classically, we only admit the real part of the solution and discard the imaginary part. As a first step in deriving the wave equation for matter waves, we must incorporate the two quantum conditions for its momentum and energy. Given the de Broglie wavelength $\lambda(= h/p)$ and that the wavevector $\mathbf{k}$ is related to λ through the relation $k = 2\pi/\lambda$, we obtain,

$$p = \hbar k \qquad (\hbar = h/2\pi) \tag{5.3}$$

As for the energy of this matter wave, we start from the formula for a quantum of energy, $E = h\nu$. With $\omega = 2\pi\nu$, this can be re-written as $E = \hbar\omega$. Substituting for k and ω into equation (5.2) gives,

$$\psi = A\, exp\left[i(\mathbf{p} \cdot \mathbf{r} - Et)/\hbar\right] \tag{5.4}$$

[3]The second-order differential nature stems from the consideration of the dynamics of the system by applying Newton's second law ($F = ma$) to a small region of the system.

We may now proceed in two steps:

Step 1: Differentiate (5.4) with respect to t^4,

$$\frac{\partial\psi}{\partial t} = (\frac{-iE}{\hbar})\psi \tag{5.5}$$

$$\Rightarrow \quad i\hbar \cdot \frac{\partial\psi}{\partial t} = E\psi \tag{5.6}$$

Step 2: Differentiate equation (5.4) twice with respect to $\mathbf{r}$,

$$\nabla^2\psi = -\frac{p^2}{\hbar^2} \cdot \psi \tag{5.7}$$

$$\Rightarrow \quad p^2 = -\hbar^2 \frac{\nabla^2\psi}{\psi} \tag{5.8}$$

(with the Laplacian $\nabla^2 \equiv \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$)

The energy term in equation (5.6) may also be written as the sum of a kinetic and a potential component (V),

$$E = \frac{p^2}{2m} + V \tag{5.9}$$

Substitute equation (5.8) into equation (5.9) yields,

$$E = -(\frac{\hbar^2}{2m\psi}) \cdot \nabla^2\psi + V \tag{5.10}$$

Finally, replacing the E term in equation (5.6) above by the expression in (5.10) leads to the celebrated Schrödinger (time-dependent) equation that governs the dynamics of a matter wave,

[4] $A^2 \equiv 1$, as A is treated as a normalization constant.

$$i\hbar\frac{\partial\psi}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2\psi + V\psi \tag{5.11}$$

One immediately observes that this equation is linear - the variable ψ and its derivatives appear only in the first power and in separate terms. Linearity gives rise to two important consequences.

First, being first-order in time, the equation uniquely defines the entire evolution of ψ, provided the latter is specified at a given initial instant. The dynamical state of the physical system is determined once ψ is known. For this reason, the *wavefunction* ψ is commonly called the state function.

Second, the linearity of the Schrödinger equation carries the implication that any linear combination of the functions displaying the form of a plane wave as in equation (6.2) is also a solution for the same equation. This is the important property of *superposition* which, as we shall explain, plays a crucial role in determining quantum behaviours.

To be sure, one may speak sensibly about the temporal evolution of ψ only provided one understands the meaning of ψ and what it represents. A good grasp of its meaning is all the more essential if the evolution of ψ is to be justifiably treated as the paradigm of a continuous spatiotemporal physical process in the quantum domain. Any attempt with a view to shedding light on the wavefunction ψ invariably leads to the consideration of interpretational matters concerning its meaning.

One must, however, realise that we are, after all, supplying a description of matter (with a particle in the simplest form), and a natural question arises as to whether one can somehow reconcile this "spreading" wave form with the image of a particle being an entity that is relatively localized in space. In other words, one must find a way of making sense of how ψ describes, at least to a good approximation, the classical motion of a particle having both reasonably definite values of momentum and energy. It turns out that a certain degree of localization may be achieved by the *superposition*[5] of several of these different plane wave solutions to form what is known as a *wavepacket*,

$$G(x,t) = \int_{-\infty}^{+\infty} A(k)\; exp\,[i(\mathbf{k}\cdot\mathbf{x} - \omega t)]dk \tag{5.12}$$

where $\omega = \omega(k)$. The linearity of the Schrödinger equation guarantees that the resulting wavepacket is also an acceptable solution.

The classical relation between the velocity and momentum of the particle should hold for the wavepacket for it to be a close analogue to a classical particle. This would mean for the wave group described by equation (5.12) to be traveling with one characteristic group velocity. Moreover, the condition of localization makes it a necessity to narrow the propagation vectors $\mathbf{k}$ included in the superposition to a fairly small range. That is, it is supposed that the function $A(k)$ is nonzero only for a small range of values centered around a particular value of k,

$$A(k) \neq 0, \quad k_o - \varepsilon < k < k_o + \varepsilon, \quad \varepsilon << k_o$$

For a small range of values in the vicinity of k_o, $\omega(k)$ can be expanded as,

$$\omega = \omega_o + (k - k_o)\frac{d\omega}{dk}_{k_o} + \ldots \tag{5.13}$$

With the approximation of ω given in equation (5.13) and ignoring higher-order terms, equation (5.12) may be written as,

$$G(x,t) \approx \int_{-\infty}^{+\infty} A(k) \exp[i(kx - (\omega_o + (k - k_o)\frac{d\omega}{dk}))t]dk \tag{5.14}$$

$$= \int_{-\infty}^{+\infty} A(k) \exp[ikx - i\omega_o t - ik\frac{d\omega}{dk}t + ik_o\frac{d\omega}{dk}t]dk \tag{5.15}$$

[5]For the sake of simplicity, we consider only plane waves in the x-direction, although the result may be readily generalized to cover three-dimensional waves.

$$= \exp[i(k_o x - \omega_o t)] \int_{-\infty}^{+\infty} A(k) \exp[ikx - ik_o x - i(k - k_o)\frac{d\omega}{dk}t]dk \qquad (5.16)$$

$$G(x,t) \approx \exp[i(k_o x - \omega_o t)] \int_{-\infty}^{+\infty} A(k) \exp[i(k - k_o)(x - \frac{d\omega}{dk}t)]dk \qquad (5.17)$$

The integral in equation (5.17) when evaluated gives the form,

$$B\,[x - \tfrac{d\omega}{dk}t]$$

and equation (5.12) thus becomes,

$$G(x,t) \;=\; B\,[x - \frac{d\omega}{dk}t] \exp[i(k_o x - \omega_o t)] \qquad (5.18)$$

Equation (5.18) represents a product of an envelope function B and a plane wave as shown in Figure 5-1. It describes the propagation of a group of waves for which the envelope or group velocity is given by,

$$v_g \;=\; d\omega/dk$$

and this velocity v_g is to be identified as the velocity of the associated particle.

The visualization of a material particle as a localized wavepacket also paves the way for understanding the Heisenberg uncertainty principle. Unlike its classical counterpart, a particle so conceived is forbidden by its "wave-like" nature from possessing simultaneous precise values of position and momentum, and the construction of a wavepacket by means of superposition has unquestionably provided a powerful illustration of wave-particle duality in quantum mechanics.

One should, however, be cautioned on the potential misrepresentation of this picture. This picture draws too strongly on the superposition principle as one that has

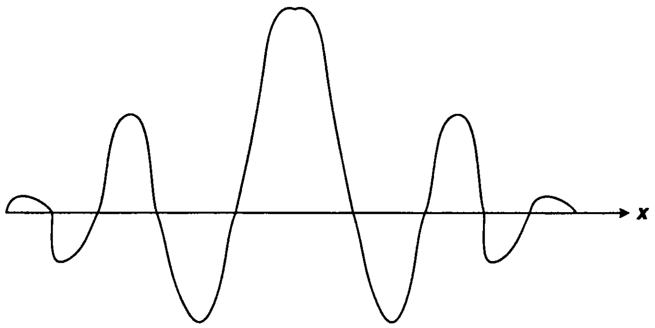

Figure 5.1: Schematic representation of a complex wavepacket moving in the x-direction.

become familiar to us through classical wave phenomena. Indeed, it has been stressed many times in various expositions of the subject[6] that the *quantum* treatment of the *principle of superposition* is of an entirely different character from that of classical physics and has in turn far reaching implications for the interpretation of ψ. We must now elucidate the peculiar nature of this principle in relation to quantum mechanics. In order to be able to compare and contrast with its application to classical physics, we first introduce the principle of superposition in the classical context.

[6]Most notable and illuminating of which can be found in Dirac, P.A.M. (1958), Chapter 1, p.4-18.

5.3 The Superposition of Classical Waves

The *principle of superposition* as a characteristic of physical systems described by linear equations, has long been recognized in classical wave phenomena. In that context, the primary question of interest is the resultant produced by either two or more harmonic[7] vibrations or non-periodic disturbances (single pulses for instance) that act together simultaneously in the same region of space and on the same medium[8]. A simple scenario is that of two pulses traveling down a piece of string for instance. The principle of superposition stipulates that the resultant pulse is simply taken to be the sum of the individual ones. In other words, the total displacement produced as a resultant of several disturbances is the *vectorial sum* of each of the individual constituent disturbances.

Superposition holds because the component forces acting are *independent*: they all contribute towards the overall effect on the medium on which they act but each remains unaffected by the presence and actions of the others. The addition of forces in classical mechanics is also based on this supposition, although it is often considered most remarkably demonstrated in wave phenomena. Consider two opposite pulses that are equal in magnitude traveling towards each other (Figure 5-2). When these pulses meet, the resulting displacement is given as the sum of the separate displacements of the pulses as in Figures 5-2(b) and (c). But once they move through each other, both of them would in no way suffer any permanent alteration and behave as though the other is not present (Figure 5-2(d)).

It is illustrated in Figure 5-2(b) that at some points, the two pulses add to produce a maximum displacement of the resultant (x_1 and x_3) while at some others (x_2), they do so to result in a minimum overall displacement. These phenomena are known as *constructive* superposition and *destructive* superposition respectively.

The phenomena of constructive and destructive superpositions in Figure 5-2 are of a transient nature - they appear only as and when the two pulses meet. However,

[7]We often associate periodic motions with sinusodal waves but it is important to note that the principle of superposition applies regardless of the shapes of the disturbances.

[8]Light waves, however, can act in the vacuum.

(a)

(b)

(c)

(d)

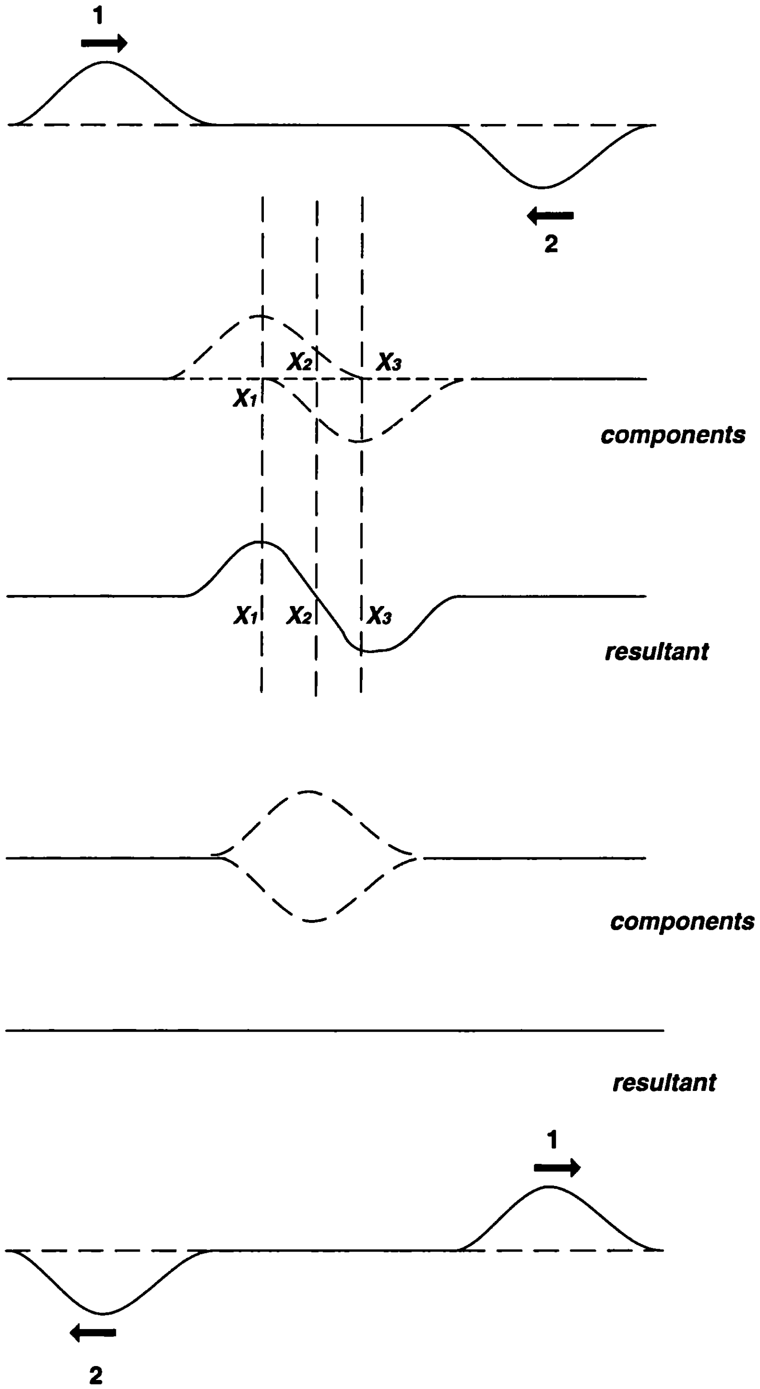

Figure 5.2: The superposition of two equal but opposite pulses traveling towards each other.

if the two pulses are replaced by two wave trains that display periodicity, a steady pattern of constructive and destructive superpositions may then be maintained. This, as we often call it, is the phenomenon of *interference.* It applies commonly to water waves as well as light waves. Interference is a direct physical consequence of the superposition principle. In optical interference, this gives rise to the observation of bright and dark fringes when a number of light waves coexist in a region of space.

Interference can sometimes be a misleading word for it may suggest that the component light waves really "interact" to produce the resulting fringes. A genuine interaction, as we have emphasized, brings about permanent changes to the interacting systems beyond the locus of the interaction. However, interference is fundamentally a phenomenon of superposition and it is indeed a vital feature of the principle that the components act independently and do not suffer any alterations in their motions due to their encounters with each other. In this regard, A.A. Michelson issues a word of caution[9],

> ***When two similar wave-trains traveling in approximately the same direction are superposed, the resulting motion may be greater or less than that of the components, according to the difference of phase of the components. Thus if the two wave-trains are simple harmonic and meet in the same phase, the amplitude will be doubled and the intensity quadrupled. If, however, the phases be opposite, the resulting amplitude (and intensity) will be zero. In this case, the two wave-trains are said to "interfere", and the resulting phenomenon is known as "interference". The term is not very well chosen, for in fact each train produces its own effect quite independently of the other, but it has been in use so long that it would not seem wise to alter it.***

Michelson's caution is echoed by R.E.I. Newton[10],

[9] Michelson, A.A. (1927), p.10.
[10] Newton, R.E.I. (1990), p.78.

Sometimes these phenomena are called constructive and destructive interference respectively... The term can be confusing, as we have seen, the waves do not interfere[11] but move through each other unaltered; at the place of meeting, they act as though the other wave were not present and the total displacement is obtained by adding together two separate displacements.

As a first impression, the adjectives "constructive" and "destructive" could be misleading since they may convey the impressions that the creation and destruction of energy have occurred. But this is certainly not the case because if less light is reaching a given point (of low intensity), more light will be reaching some other point correspondingly. Interference merely effects a re-distribution of light energy and the conservation of energy is therefore upheld, with the total (numerical sum) energy of the two waves to remain constant.

The bright and dark fringes in an optical interference pattern represent locations of reinforced and weakened intensities. Intensity is usually taken as the measure of the rate of energy flow per unit area perpendicular to the direction of propagation of a wave and it is proportional to the square of the amplitude of the wave,

$$Intensity \propto (amplitude)^2$$

To show how interference is described mathematically, we consider the superposition (R) of two plane waves of the same frequency ω, but allow for the different phases ϕ_1 and ϕ_2[12],

$$R = A_1\cos(\omega t + \phi_1) + A_2\cos(\omega t + \phi_2) \tag{5.19}$$

This is equivalent to[13],

[11]This, I take it as the common usage of the word "interfere" as a somewhat genuine interaction in which interacting agents suffer permanent changes beyond the focus of encounter.

[12]ϕ_1 and ϕ_2 are normally called phase constants. Changing the values of these quantities merely makes all the events in the cycle happen earlier or later by the same amount without altering the physical relation or the sequential order of events within the cycle.

[13]For classical waves, only the real part of the linear combination of the two wave solutions is admitted.

$$R = Re[A_1 \exp i(\omega t + \phi_1) + A_2 \exp i(\omega t + \phi_2)]$$

$$R = [A_1 \exp(i\phi_1) + A_2 \exp(i\phi_2)] \cdot \exp(i\omega t) = A_R \exp(i\phi_R) \cdot \exp(i\omega t) \qquad (5.20)$$

with,

$$A_R \exp(i\phi_R) = A_1 \exp(i\phi_1) + A_2 \exp(i\phi_2) \qquad (5.21)$$

Since the intensity is proportional to the square of the amplitude, it follows therefore that the resultant intensity due to the superposition of the two waves is given by,

$$[A_R \exp(i\phi_R) \cdot A_R \exp(-i\phi_R)] = [A_1 \exp(i\phi_1) + A_2 \exp(i\phi_2)][A_1 \exp(-i\phi_1) + A_2 \exp(-i\phi_2)]$$

$$A_R^2 = A_1^2 + A_2^2 + A_1 A_2 [\exp(i(\phi_2 - \phi_1))] + A_1 A_2 [\exp(-i(\phi_2 - \phi_1))]$$

$$= A_1^2 + A_2^2 + A_1 A_2 \{[\exp(i(\phi_2 - \phi_1))] + [\exp(-i(\phi_2 - \phi_1))]$$

Given that: $\exp(i\theta) + \exp(-i\theta) = \cos\theta + i\sin\theta + \cos\theta - i\sin\theta = 2\cos\theta$, therefore,

$$A_R^2 = A_1^2 + A_2^2 + 2A_1 A_2 \cos(\phi_2 - \phi_1) \qquad (5.22)$$

Let I_R, I_1 and I_2 denote A_R^2, A_1^2 and A_2^2 respectively, we have,

$$I_R = I_1 + I_2 + 2\sqrt{I_1 I_2} \cos(\phi_2 - \phi_1) \qquad (5.23)$$

In the special case where the amplitudes of the two waves are equal with $I = I_1 = I_2$,

$$I_R = 4I \cos^2[\frac{1}{2}(\phi_2 - \phi_1)] \qquad (5.24)$$

The influence of the phase difference $(\phi_2 - \phi_1)$ is obvious from equation (5.24). Depending on the value of $\cos^2[\frac{1}{2}(\phi_2 - \phi_1)]$, which varies between -1 and +1, I_R can take on any value between 0 and $4I$. These represent cases of alternating dark and bright fringes observed at the locations of minimum and maximum intensities within the interference pattern. One sees that when $(\phi_2 - \phi_1)$ equals zero (when there is no phase difference between the two waves and they are said to be *in phase*[14]), a reinforcement of intensities that amounts to four times the original intensities due to the aggregation of the two waves obtains.

The crossed-term $2A_1A_2\cos(\phi_2-\phi_1)$ in equation (5.22) and the corresponding term $2\sqrt{I_1I_2}\cos(\phi_2 - \phi_1)$ in equation (5.23) both describe the appearance of interference effects. They represent a correction to the discrepancy between the resultant of the two waves as obtained by the simple addition of the two intensities and that which is actually observed - the alternate bright and dark fringes. This correction explains how the difference in the phases of the two waves accounts for the variation in the resulting intensity that has not arisen from the direct sum of the separate intensities. This is what interference consists in and to quote Feynman[15],

> ***This correction we call the interference effect. It is really only the difference between what we get simply by adding the intensities, and what actually happens.***

It ought to be pointed out that although the principle of superposition is widely applicable to any linear system, only those that follow periodic motions may exhibit phase differences (which are essentially the *time differences* in the phases of the wave motions), and so only such systems are capable of displaying interference effects. This is why interference is considered a characteristic feature of periodic or wave motions.

Optical interference is most readily observed in the famous Young's double-slit experiment (Figure 5-3). In its paradigmatic form, the apparatus consists of a source of monochromatic light (to ensure coherence) and light emitted from this source is to meet a screen on which there are two narrow slits. These slits, S_1 and S_2, divide the original wavefront into two separate ones and light subsequently emanating from them are exactly in phase and constitute two coherent secondary sources. These

[14]Two waves that are in phase are called *coherent.*

[15]Feynman, R.P. (1963a), p.29-7.

waves then superpose to give an interference pattern in the form of bright and dark fringes on a second (detection) screen placed some distance away.

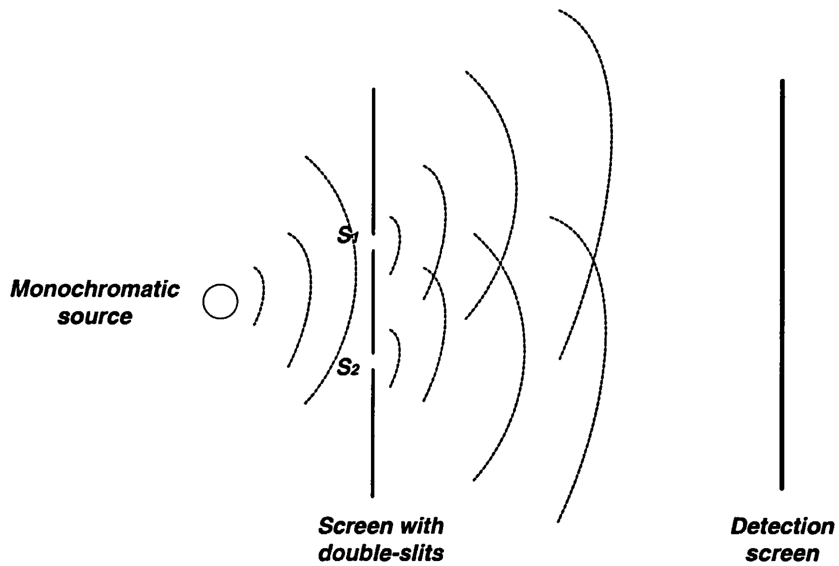

Figure 5.3: Schematic of the Young's double-slit experiment.

The presence of interference in this setup is attributed to the mis-matches of the phases of the two light waves as introduced by a path difference[16] relative to a specific point in space, as in Figure 5-4 below.

In order to superpose constructively at P, the two waves must be found to be in phase at that point. For this to happen, the path difference d must equal to an integral number of whole wavelengths. Conversely, if d is equal to an odd number of half-wavelengths, the two waves arriving at P would be found to be *out of phase* and a case of destructive interference obtains. These events of constructive and destructive superpositions produce an alternate pattern of bright and dark fringes on the detection screen which are equally spaced.

It can be shown that for this experimental arrangement, the resultant intensity varies between maxima and minima according to the square of the cosine of the phase difference between the two components, which is similar to the more general form of equation (5.24) given earlier,

[16]The slight difference between the corresponding successive parts of both waves to reach a certain point in space.

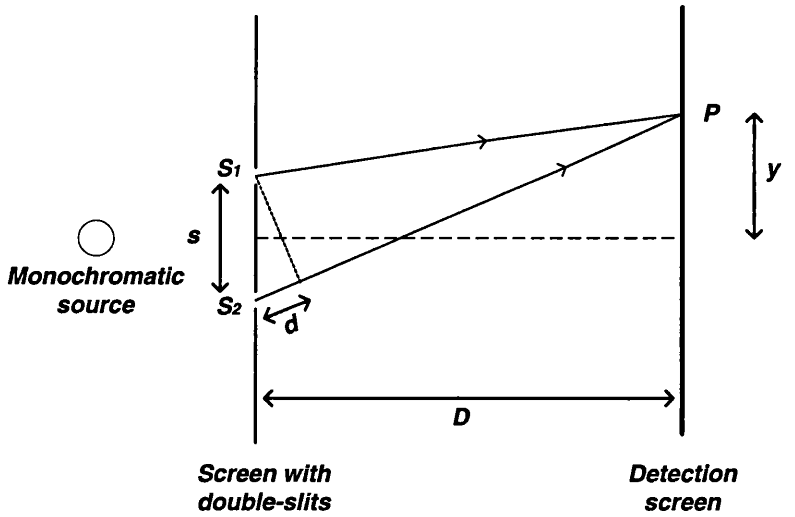

Figure 5.4: Light rays emanating from slits S_1 and S_2 are in phase but in order to reach a certain P on the screen, light from S_2 is required to traverse an extra distance d, the path difference.

$$I = 4I_o\cos^2(ys\pi/D\lambda) \tag{5.25}$$

where y is the distance of point P from the horizontal central axis, s is the separation between the two slits and D is the distance between the two-slits and the detection screen.

This predicts, in theory, a constant alternating picture of bright and dark fringes (non-localized fringes) as shown in Figure 5-5.

In practice, the fringes are only significant where the intensity from each slit is large and where the intensities are approximately equal. And so the actual observed pattern often shows that the brightness of these fringes being modulated by the diffraction pattern[17] (Figure 5-6) dominated by a central maximum.

We have gone into quite some length to explain the principle of superposition

[17]Diffraction occurs whenever a portion of the wavefront is obstructed in some way. The various segments of the wavefront that propagate beyond the obstacle interfere, causing the particular energy-density distribution referred to as the diffraction pattern. There is, however, no significant physical distinction between interference and diffraction.

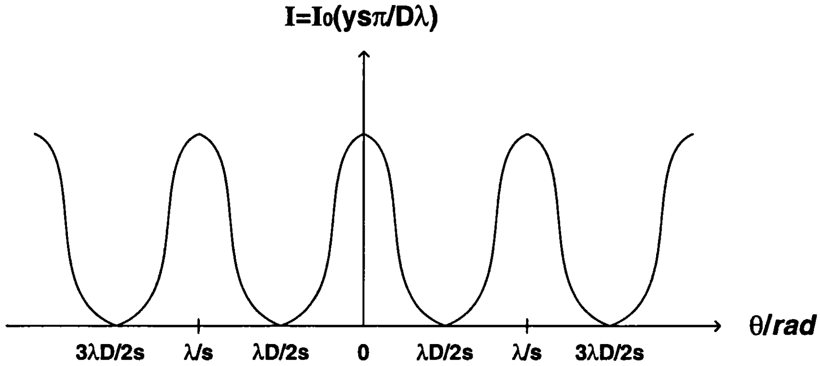

Figure 5.5: Idealized plot of intensity-versus-distance from the central axis in a double-slit interference pattern.

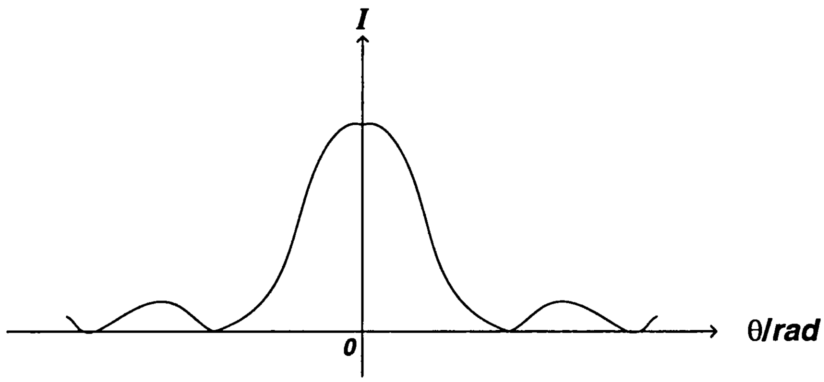

Figure 5.6: Actual interference observed in double-slit interferometers due to effects of diffraction.

as it is understood in classical physics. This sets the stage for a fuller appreciation of the very different quantum treatment of the superposition principle. Let us now summarize the essential features of the *classical* principle of superposition.

The principle applies to physical systems that obey linear equations. The essence consists in the fact that the component forces acting in a superposition stay unaltered and *act in an independent manner* so that their resultant effect on the system can be represented as *a direct sum of the separate individual effects*, as produced by each of the constituent forces.

In statics, forces acting concurrently on a body in equilibrium give a resultant effect of zero as the individual force vectors cancel each other. Likewise in dynamics, the resultant force on a body has an effect that is equivalent to the combined effects of the constituent forces. All these work because Newton's second law is a linear equation.

Thus viewed, the principle of superposition can be applied very generally. However, when applied to systems that undergo periodic motions like waves, the mechanism of superposition leads to the consequence of interference that is unique to these modes of motion. Interference phenomena are possible only for periodic motions because these are the ones which have the privilege of enjoying the prospects of sharing possible differences in phases with other similar motions of its kind. Non-periodic motions do not entertain any possibility of incurring phase differences and hence while the principle of superposition works for both kinds of motions, only periodic ones exhibit interference effects.

Wave-like periodic motions spread out in space. Being a phenomenon characteristic of and unique to wave-like motion, interference is a non-localized affair that cannot happen at only one single point in space. In fact, its non-local character ensures that where energy seems to vanish at points of minimal intensity, it is compensated correspondingly at points of maximal intensity. Because of its prevalence through space, conservation of energy within a wave is upheld in an instance of interference.

5.4 The "Quantum" Principle of Superposition and the probabilistic interpretation of ψ

5.4.1 The Strange Case of Quantum Interference

In the preceding discussion, the principle of superposition has been introduced in the context of wave motions. For waves produced by coherent sources, the principle leads to the observation of the classic phenomenon of interference, which has been well-demonstrated in the Young's double-slit experiment.

It is, indeed, a central claim of quantum mechanics that under appropriate conditions, particles (or waves) can exhibit wave-like (or particle-like) behaviours. As we have already shown, superposition enables us to conceive of a particle as being manifested as a localized wavepacket. The introduction of the wavepacket concept is an attempt to bring into reconciliation two seemingly contradictory characters of a particle: its *localization in space* and its *exhibition of wave-like behaviours* under certain appropriate conditions. For an electron, one such condition that brings out its wave-like manifestation would be its confinement within an atom, as described above. Of course, one may raise the immediate objection that the existence of stationary states of an atomic electron merely supplies de Broglie with the inspiration to draw the bold analogy with confined classical waves. Nevertheless, the analogy seems a good one. The subsequent formulation of the wave equation by Schrödinger and its impressive agreements with the bulk of experimental evidence confirm the validity of this inference. However, this wave-like behaviour is not to be confused with a physical wave as conceived in the classical sense. As it happens, the "wave" nature of an electron is far removed from the usual conception. In this section, we endeavour to probe further into the realm of the quantum "wave".

Light waves superpose to give interference patterns. No one finds it necessary to dispute this wavy behaviour of light. The quantum revolution has, however, brought to our knowledge another mode of manifestation of light: as photons - the discrete quanta of light energy having the hallmark particle characteristic of localized positions[18] in space. It is a legitimate question to ask what sanctions us to infer the existence of this corpuscular manifestation of light?

Modern versions of the Young's double-slit experiments employ photodetectors that ride on motor-driven slides that scan the interference patterns. These photodetectors detect "*whole*" photons: the detector issues a "*click*" every time a (whole) photon with a specified amount of energy (and not a fraction of it) is received. It is an "all-or-nothing" event: either the detector receives a photon of a specific energy or it doesn't. Because photodetectors have dimensions that take up relatively localized positions in space, the registration of each photon establishes the corresponding particle-like localization of these photons. The detection of photons by photodetectors therefore confirms for us their particle nature. As Dirac has put it most elegantly[19], "*the individuality of the photons are preserved.*"

Taking into account the reality of photons, a beam of light may now be thought of as made up of a large number of photons aggregating to give the beam its appearance of a continuous character.

Light, both as waves or as photons, undoubtedly produces optical interference in the Young's double-slit apparatus. It turns out that re-analyzing the interference results *on a photon basis* reveals disturbing quantum features of these "particles". In particular, it leads us to the appreciation of the nature of quantum superposition and the probabilistic interpretation of the wavefunction ψ.

One important assumption one needs to bear in mind as one switches from a wave-ontology to a particle-ontology is the notion of localization in space. Entities having localized positions in space are expected to trace out more or less *definite paths* during their motions, as represented by a sequence of localized positions in spacetime. In the Young's double-slit setup, this corresponds to the association of a definite "*continuous*" trajectory with each of the photons leaving the source, traversing "one"

[18]To be precise, these localized positions are established only when the photons are detected. Although energy quantization does not suggest localization, observation has indicated that photons can follow narrow paths and that they are not spread out everywhere.

[19]Dirac, P.A.M. (1958), p.6.

of the slits and eventually reaching the detector.

Our immediate challenge is this: how might one utilize the idea of a discrete unit (a photon) having localized spatial extent that follows a definite trajectory in spacetime, to explain the appearance of the alternating bright and dark fringes as observed in the double-slits interference pattern?

A basic observation ensues. Since only spatially localized photons are detected, the "spread-out" interference pattern is generated and gradually built up from the arrival of photons at the detector. In other words, the "spread-out" pattern suggests itself as some kind of a *measure* relating the photons and their locations; the distribution thus lends itself naturally to statistical and probabilistic considerations.

Interference as a Measure of the Probable Number of Photons Present at a Location

There are two sorts of measures one may establish between the photons and their locations. First, we note that the "*intensity of light*" is defined as the *time average of the amount of energy crossing in unit time a unit area perpendicular to the direction of energy flow*[20]. Expressed in terms of photons,

$$\text{Intensity} = \text{Number of photons arriving} \times \text{Energy of each photon} \quad (E = \hbar\omega)$$

This leads to the association of the measure with the population density of photons at a certain location y along the vertical-axis on which the detector is scanning. This rather straightforward interpretation of the interference pattern gives us information of the *likely* or *probable number of photons present at a certain point.* Hidden in this apparent intuitive simplicity is a subtle point the recognition of which would no doubt save us from committing a significant error.

By the "probable number of photons present at a point", we may seem to refer specifically to photons that are simultaneously[21] present at that point.

[20] Born, M. and Wolf, E. (1980).

[21] When speaking about the probability of a number of photons at a certain location, we have as yet made no scruples of the relative timing of arrivals of these photons at that location. With respect

To proceed, we make the identification that a maximum (a bright fringe) within the interference pattern corresponds to the place frequented by the highest number of photons (either simultaneously or successively) and a minimum (a dark fringe) represents a position where there should hardly be any photons in presence (Figure 5-7).

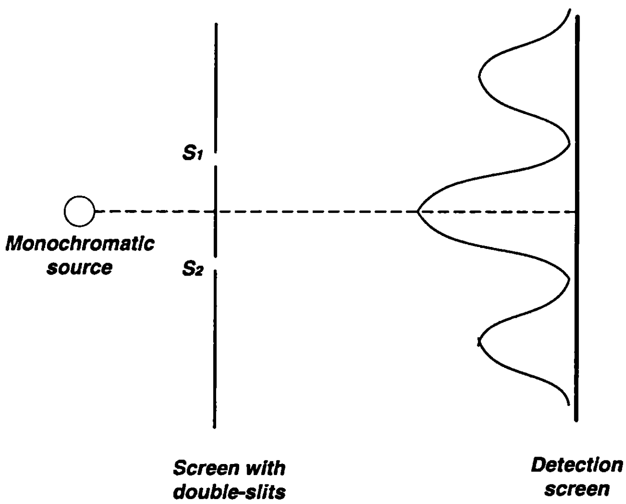

Figure 5.7: Young's double-slit experiment - can the interference pattern observed be explained by the photon basis?

In particular, the locations of minima may correspond either to points at which *not a single photon has called* or *photons that are simultaneously present somehow interact* to result in the consequence of a nil number of photons.

We first examine this possibility of "interactions" amongst photons on their simultaneous[22] arrivals at the detectors.

to timing, it is reasonable to make the assertion of two photons emerging from a different slit may arrive at the same location at the same time, although we are most inclined to think that photons from the same slit would have to arrive at a specific spot one after another in a successive sequence. Still, due to the finite spatial dimension of the detector and the slit-width, we also allow for the possibility of two (or more) photons reaching the detector at the same time given that they have both passed through the same slit. There are photodetection devices that are capable of detecting two photons arriving (at the same spatial point) at the same time. These are normally characterized by the appearance of "two photon" peaks on their responsivity curves. I thank Professor Yanhua Shih for this information.

[22]Since a photon is annihilated upon its detection by the photodetection device, it follows therefore that any possible interactions between two photons arriving successively can be safely ruled out. And strictly speaking, the interaction should happen just before the photons are annihilated by the detector.

To be sure, the photons are, of course, entitled to "interact" well before reaching the detector, but here, we focus on the possibility of their interactions at the detector as we endeavour to establish the the interference pattern as a measure of the probable number of photons present at a point; with such a point being a position of the detector and the collection of detection events at these points make up the interference pattern.

An explanation is called for as to how two photons[23] are supposed to "*interact*" in such a way that no photon is found at the minima as a result. An obvious suggestion is the two photons would have to *annihilate*[24] each other so that no photon is left at these spots. Similarly, the two photons would have to interact to produce a total of four photons at the central maxima (equation (5.25)[25]). But where does the energy go in the former case and from where comes the extra energy in the latter? Bearing in mind the fact that unlike a physical wave, these photons are particle-like entities with no "continuous" spatial extension and hence offer no opportunity for "squeezing" or "re-distribution" of energy from one spot to another. The respective annihilation and creation of photons at the various minima and maxima would then have to mean that energy conservation is blatantly violated at these locations. One therefore concludes that no interaction of the right sort is possible for the interference pattern to stand for the measure of the probable number of photons simultaneously present at one place.

One is perhaps quick to respond by insisting that, because of possible scattering actions at the slits, no photon needs arrive at points of minima and several photons may get scattered to be collected at places of maxima, which accumulate to produce the intensity at the maxima.

Given this line of reasoning, most photons of a definite energy are expected to pass through the the slits in a straight line unhindered, while only a relatively small

[23]As a minimum (at least in a classical sense), one requires two entities for an interaction to take place.

[24]Here, we are *not* using the term annihilation as in the formal sense of quantum field theory.

[25]It can be seen from the equation that at a maxima, the cosine function takes on its maximum value 1 where the overall intensity equals four times the initial intensity. Since the intensity serves as an indication of the number of photons present (with the initial intensity corresponding to the presence of one photon) it follows that there is a total of four present at a maxima.

proportion of them would hit the edge of the slit at such angles and get deflected off their prescribed courses. In other words, the two "centered" regions that are in line with the center of both slits is predicted to receive the highest concentration of photons (Figure 5-8).

Now it turns out that this kind of scattering hypothesis may be readily tested by examining what happens at each slit and the resulting scattering pattern. This procedure is easily carried out by having the other slit blocked off while observing one of them.

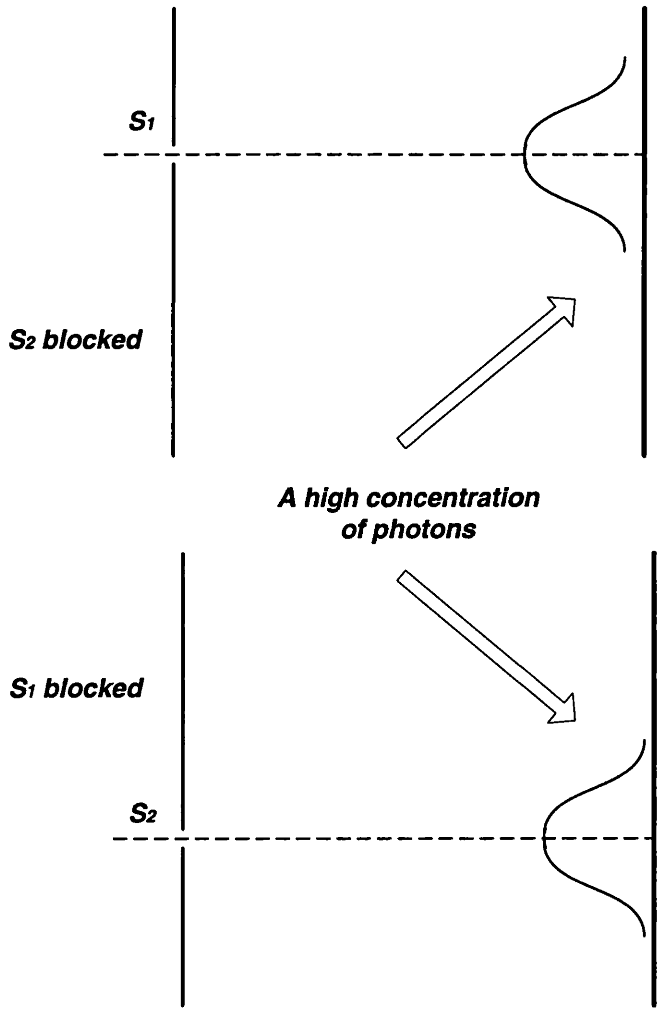

Figure 5.8: Results of scattering of photons at the slits.

Two sharp peaks indicating the concentration of photons around the central-axes through the two slits support our prediction which, I hasten to emphasize, is based on the *classical* scattering of a particle-like entity having quite definite momentum and location in space. With no hesitation, one would confidently apply the *principle of superposition* to both of these distributions to obtain the resultant scattering pattern *when both slits are open*, which is shown in Figure 5-9 below.

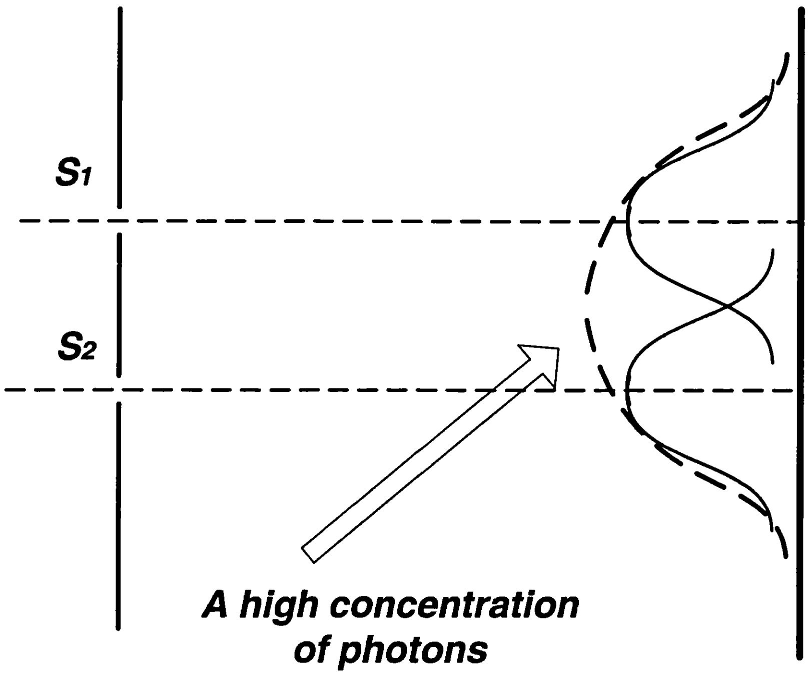

Figure 5.9: The overall distribution of photons (as indicated by the dotted curve) through both slits according to the prediction of classical scattering.

The resultant distribution shows a "normal" curve peaking at the central axis mid-way between the two slits. It does not result in the interference pattern as it is observed experimentally! From this observation we conclude that both the classical model of scattering, and the hypothesis of "no interaction amongst photons arriving at the same location", fail on this occasion to support the proposition that *the interference pattern represents a measure of the probable number of photons present simultaneously at a location.*

There remains a potential loophole that must be amended. In the course of exploring the possibility of interactions amongst the simultaneously arriving photons, we have supposed that those photons that arrive *in succession* do not interact. As a confirmation of this supposition, the intensity of the source is greatly reduced to such an extent that, at any one time, only one photon is allowed to "pass through the screen with the slits" and reach the detector[26].

However, one is surprised to find the most baffling consequence! An interference pattern, much like the one obtained in the case of high photon intensity, is gradually built up even in this low photon intensity case. Reducing the intensity serves merely to reduce the rate at which the pattern builds up. Interestingly, the first experiment of this kind was performed in 1909 by G.I.Taylor (later Sir Geoffrey) at Trinity College, Cambridge[27] from which he concluded, "*In no case was there any diminution in the sharpness of the pattern...*"

With photons reaching the detector one at a time, the suspicion of any interaction amongst photons simultaneously arriving at the same spot can be dispelled. The only other possibility one may reasonably entertain in this situation is that the photons suffer scattering as they pass through their respective slits *in such a way* that each of them gets delivered to take up its rightful position within the interference pattern. To examine this hypothesis, we resort to the procedures of observing each slit (with the other one closed) in turn.

The observation of each slit on its own provides convincing evidence of the classical scattering of photons of a certain momentum and energy, as has been already demonstrated above in the case of high photon intensity. With both slits kept open, one should again expect the resultant effect of scattering to obtain as a result of the superposition of the two individual results (as in Figure 5-9 above). But once again disappointingly, the application of the principle of superposition on this occasion is *unable* to reproduce the interference pattern observed experimentally.

It may now be supposed that *the act of closing one of the slits* might have had an influence on the final observation so that the interference pattern becomes effectively "washed out". This contention is countered by the results of an experimental arrangement with three-level atoms[28] used as detectors at the slits, that are employed

[26]In this circumstance, one should expect the detector to register distinct "click"s upon the individual arrival of each photon.

[27]Taylor, G.I., (1909), p.115. For an up-to-date account of Taylor's vast contributions to different areas of physics, see Brenner, M.P. and Stone, H.A. (2000).

[28]Scully, M.O. and Drühl, K. (1982).

to track the photons *while having both slits kept open at the same time.* This arrangement also fails to reproduce the interference pattern. It seems that whenever we care to look to obtain some information about the photon's trajectory (by locating it at the slit for example), the interference pattern disappears. It is only when we are complacent enough to remain ignorant about the photon's whereabouts between the two events of its emission by the source and its arrival at the photodetector, that we are able to retain the interference pattern.

This state of affairs is all very puzzling, for our notion of a particle having a definite trajectory now appears to depend on our observation of it. Classically, a particle does have a definite trajectory. Our measurement, for instance, of its position x in space, is regarded as a faithful reproduction of its already "prescribed value" of position. We can be certain about this because given a value x at a time t, we may predict another value x_1 at any later times t_1 by the equation that governs its motion. If at time t_1, a measurement of its position is carried out, one would then find the value to agree with our predicted value x_1. In other words, the particle takes up all the predicted positions with *probability one.* This is possible because the particle *possesses* these attributes. And so, irrespective of whether any actual position measurement is performed, the particle follows an already prescribed trajectory.

In the case of a photon traversing the two-slit apparatus, our measurement appears to be capable of changing the final positions of the photon from a "smooth" scattering distribution to one showing interference.

What conclusion might one draw from this case of low photon intensity?

Interference as a measure of the probability of locating a particle

Recall that earlier in the discussion, we have indicated there are *two* sorts of measures associated with the photons and their locations within the interference pattern. Because of the violation of energy conservation, we have argued that the interference pattern is ***not*** a measure of the total number of photons that are simultaneously present at a point. This is an appropriate point to explore the possibility of the second kind of measure. The *distinct and successive arrivals of the individual photons*, in the case of low photon intensity, have quite unambiguously suggested the association of the distribution of photons within the interference pattern with a measure of the *probable number of photons arriving at a certain point*[29].

Put slightly differently, the measure tells us how likely a certain place is for a photon to land. The interference pattern thus measures the *probability of locating the photon at a certain point.* We have seen that the interference pattern is described by the intensity (equation (5.25)) that results from the overlapping (superposition) of two plane wave solutions ψ_1 and ψ_2. Granted the intensity is proportional to the *modulus square* of the amplitudes, it follows that the probability of finding a photon at a certain point is to be associated with $|\psi_1+\psi_2|^2$. In other words, it is the modulus square of a wavefunction that is given a probabilistic interpretation as *the measure of the probability of locating a particle (in our particular case, a photon) at a certain point upon our measurement.* This, incidentally, is the usual statistical interpretation of the wavefunction ψ via $|\psi|^2$.

To understand fully how the idea of a wave is integrated into the notion of probability, we need to probe deeper into their intricate relation as presented in quantum mechanics.

It is observed that although on the level of the wavefunction ψ, the principle of superposition remains valid in the sense that the two wavefunctions ψ_1 and ψ_2 add to produce a resultant whose modulus square describes the interference pattern, it is important to stress that ontologically speaking, ψ does in no way stand for a physical wave, as in the sense of classical physics. This is because the interference pattern observed consists, after all, of individual photons whose "discrete" reality we have so convincingly established. The particle nature of photons - its distinctive aspect of localization in space - prevents any possibility of a re-distribution of energy and thus, *in theory*, inhibits the attainment of interference patterns. In contrast, a physical wave provides a continuous medium so that re-distribution of energy can be duly effected. The wavefunction is therefore best conceived only as an instrument by which probabilities are introduced into the quantum formalism.

[29]Either simultaneously or successively.

This somewhat "instrumentalist" construal receives further support as one considers carefully the two "waves", one "emanating" from each of the slits supposingly, to be simultaneously present and superpose to result in the observation of the interference pattern. One may feel, at this juncture, persuaded by the concept of *wavepackets* to associate each of the "waves" with one photon. This appears a sound possibility at first sight when one refers to cases with *two different* photons - each traverses a different slit *at the same time* - where one may unambiguously identify each of the two waves with each of the two photons.

However, in the case of low photon intensity where individual photons arrive at the detector successively one after another, this rather simple-minded one-to-one association of "*a*" *photon* with "*a*" *wave* through the wavepacket concept is no longer feasible. The reason is this: there is no doubt that the interference pattern only obtains when two "waves" are introduced and superposed. For low photon density, it now appears that we are left with no other option but to relate *both* of these "waves" at the slit to *one* photon. Speaking in a figurative manner, this would have to mean something like associating the *two wavepackets* at the slits with the *same* photon. One perhaps would not find this entirely unacceptable, for might it not be the case that a photon - the wavepacket - leaves the source, *gets divided into two* when reaching the slits and later re-combine to one wavepacket upon detection? This kind of reasoning would be deemed sound in cases where a particle is composed of "physical waves" and the localized feature of the particle arises from the superposition of these "extended" physical waves[30].

If there should be a one-to-one correspondence between a wavepacket and a particle, then in the situation just pictured, the two wavepackets resulting from the division of the original one at the slits are to be associated with a *part* of the photon that the original wavepacket represents. In other words, should this state of affairs really

[30]A good example is that of *solitons.* These are relatively stable nonlinear waves with their envelopes (which itself is of periodic wave form) formed of several waves of different amplitudes. The term was first used by Zabusky and Kurskal in a paper published in 1964 and the authors discovered that these solitary waves can pass through one another without deformations due to collisions and hence the name "solitons" to signify their behaviours as particles. See, Hasegawa, A. (1992), especially chapters 1-4.

obtain, one would then be able to observe, say, a "half-photon" - one that is of half the amount of energy of the original photon for instance - at each of the slits. But as far as the evidence goes, when a detector is placed near each slit, one receives distinct "click"s that record photons coming from either slit but never both. Moreover, each click indicates that there is no variation in the energies of the photons detected. Both of these observations confirm that the photons do *not* get split at the slits. What is more, is the interference pattern vanishes once we start trying to track down which path - *through either slit 1 or slit 2* - a photon follows!

The conclusion is somewhat mind-boggling. Two "waves" must be simultaneously present in order to superpose to achieve the observed interference pattern. But on the other hand, we are compelled to accept the association of both waves simultaneously with the same photon. Here goes Dirac's famous dictum[31], "*each photon then interfere only with itself*". The one-to-one correspondence between a wavepacket and a particle is indeed too simplified for the understanding of the present situation. Yet, we must strive to devise an "intelligible" means to associate two simultaneous waves with one and the same photon.

The interpretation of the interference pattern as a measure of the probability of locating a photon at a certain point provides an important clue. Analogously, it seems plausible to associate the "wave" at each of the two slits with a measure of the probability of locating the photon at that particular slit.

If this procedure is followed, one is landed with the advantage of relating *two potential alternatives* rather than two hard-fact eventualities[32] to the same photon. The two subsets of events: "*the photon passes through slit 1*" and "*the photon passes through slit 2*" are treated as mutually exclusive[33] and together they form a sample space for the probability measures in accordance with the classical (Kolmogorovian)

[31]Dirac, P.A.M. (1958), p.9.

[32]Here, we refer to two "physical waves" that carry energy, and hence the conservation of energy needs to be taken into account when they are both associated with the same photon that is observed to traverse only one slit.

[33]Two sets are said to be *mutually exclusive* or *disjoint* if they share *no* common element. The fact that each photon (as a particle) is found to traverse either slit 1 or slit 2 but not both at the same time sanctions the partition into two mutually exclusive subsets.

probability theory[34].

If indeed the two alternative subsets of events corresponding to "*the photon passing through slit 1*" and "*the photon passing through slit 2*" are mutually exclusive, so that the photon may either come through slit 1 or slit 2 but *not* both *at the same time*, then the classical theory of probability stipulates that:

$$P(1) + P(2) \;=\; P(1\,or\,2) \tag{5.26}$$

where $P(1)$, $P(2)$ and $P(1\,or\,2)$ denote the probabilities for the photon traversing *slit 1*, *slit 2* and *either slit 1 or slit 2* respectively.

Equation (5.26) represents the important property of finite additivity that obtains as a consequence of the disjointness of the subsets (the two alternatives) under consideration.

Now if the probabilities are represented by the physical distribution of photons (due to scattering), then it would seem, as similar to our previous argument, that the probability of the photon going through either slit 1 or slit 2 ($P(1\ or\ 2)$)is equal to the sum of the two physical distributions of scattering ($P(1) + P(2)$). As it should be expected, photons coming through either slit 1 or slit 2 would be scattered into positions *independent of the presence of the other slit.*

This is often seen to be the "rendezvous" for the sum-rule of probability (equation (5.26)) and the principle of superposition. In such a regard, one thinks of the physical scattering distributions as representing the "individual effects" of a photon coming through slit 1 or slit 2 and the direct sum of the two distributions as representing the "resultant effect" of both slits being open so that both possibilities are made available. Thus viewed, the sum-rule would seem to be expressing the same content as the physical principle of superposition. As attractive as it may seem, this simple correspondence is not exact though it does bring to our attention one vital element of our analysis; namely, probabilistic considerations ought somehow be connected to the physical situations through the principle of superposition. The question that concerns

[34]This property of disjointness of the subsets of events makes possible the finite additive property of probability measures. See Kolmogorov, A.N. (1950), Chapter 1.

us immediately is how should such a connection be made?

5.4.2 When Physics meets Probability: the Superposition of Potentials

Before attempting an answer to the last question posed in the foregoing section, we need first understand some relevant aspects about both the principle of superposition and the simple concept of probabilities. As a useful and instructive illustration, we compare and contrast the cases of *applied forces on a mass* and the classic example of the *casting of a die.*

Forces acting on a mass m

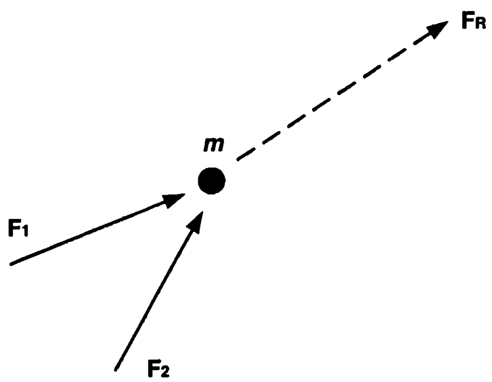

Figure 5.10: Two forces acting simultaneously on the mass m.

Consider a case where two external forces $\mathbf{F}_1$ and $\mathbf{F}_2$ are applied simultaneously to a mass m. The laws of mechanics tell us that the resultant effect generated may be attributed to the effect of a single force $\mathbf{F}_R$ (which is given by the vectorial sum of $\mathbf{F}_1$ and $\mathbf{F}_2$) as in the diagram (Figure 5-10) above. This is because the concerted effort of both of these forces equals the sum of the effect of each component force.

The principle of superposition is at work: the sum of the individual effects produced by the separate forces equals the one resultant effect generated by the forces acting together. This principle holds because the resultant force ($\mathbf{F}_R$) is the sum of the individual constituent forces ($\mathbf{F}_1 + \mathbf{F}_2$) and this fact has manifested itself in the effects that the forces produce collectively on the mass. Now we ask the deceptively trivial question of how does one ascertain the principle quantitatively?

Strictly speaking, the measurements we make of the so-called "forces" are only in a sense indirect. Following Newton's second law of motion, we measure "forces" by their effects - the accelerations they produce on physical systems that reduce finally to spatial and temporal measurements.

We may at this juncture alter our vocabulary slightly and in place of "forces", we say that the measurements of acceleration produced on the system represent the effects of the "*potentials*" or "*capacities*" to produce these accelerations. One may be suspicious whether there is any real advantage attached to what appears to be a mere change of terminology. The reader is, however, assured that the motive for this change will become apparent as the discussion develops. In the meantime, I would remain unspecific about the exact meaning and nature of these potentials, which we shall have occasion to elucidate shortly.

In order to establish the validity of the principle of superposition, the following steps are followed. First, we measure the acceleration produced by the potential $\mathbf{F}_1$ by subjecting the mass m to the influence of $\mathbf{F}_1$ alone. Next, the same procedure is repeated but this time with m being placed under the influence of $\mathbf{F}_2$ alone. Now we may combine (mathematically) the two potentials $\mathbf{F}_1$ and $\mathbf{F}_2$ to arrive at the resultant potential $\mathbf{F}_R$. This is achieved by taking the vectorial sum of $\mathbf{F}_1$ and $\mathbf{F}_2$ to obtain $\mathbf{F}_R$. The crucial point underlying this step is that $\mathbf{F}_R$ is *computed* on the tacit assumption that the potentials $\mathbf{F}_1$ and $\mathbf{F}_2$ *co-exist* and *act simultaneously* on m. Moreover, $\mathbf{F}_1$ and $\mathbf{F}_2$ both act on m *independently* in the sense that one potential *does not come under the influence of the other* or *its action is not affected by the other.*

The final step consists in subjecting m to both the potentials $\mathbf{F}_1$ and $\mathbf{F}_2$ (*actually*) acting simultaneously and measuring the overall acceleration of m. One should then find,

$$\underbrace{\frac{\mathbf{F}_1+\mathbf{F}_2}{m}}_{\text{measured}} = \underbrace{\frac{\mathbf{F}_R}{m}}_{\text{computed}} \tag{5.27}$$

The combined effect of $\mathbf{F}_1$ and $\mathbf{F}_2$ acting together is no different from the resultant effect of the separate effects due to $\mathbf{F}_1$ and $\mathbf{F}_2$ individually. Thus, we have demonstrated, empirically, the validity of the principle of superposition. The principle applies because the two potentials giving rise to the resultant effect act *simultaneously* ***and*** independently.

The principle of superposition can be quite unambiguously *demonstrated* to hold in the case of (non-periodic) potentials as in the above because the effects we measure as produced by these potentials also obey the simple additive rule that exists between the potentials. For two potentials that both exhibit periodicity, the relation between the effects are, however, not as straightforward. In such circumstances, the principle of superposition fails to apply on the level of the observed effects, despite the fact that the underlying potentials still follows a simple additive sum as in the previous example of non-periodic mechanical potentials. There is not the slightest doubt that each of these periodic potentials is capable of acting on the medium independently as one may readily verify by the closure of one of the slits, for instance, in the double-slit experiment for light waves. However, observable interference effects are present when both potentials[35] are found to act together simultaneously in the same region of spacetime (or on the same medium). This resultant effect of the two potentials acting concurrently is *not* simply the direct sum of the two individual effects of each potential acting independently. Does it mean that, after all, the two potentials affect each other when being put to work at the same?

The reply to this question hinges on one important fact: the relation between the *potential* and the *measure of its effect.* In the case of classical mechanics, the potentials, (the forces) bear a linear relation to the effect (the accelerations we measure) they each produce. On the other hand, in an optical interference experiment, the effect we measure is the intensity I of the electric-field $\mathbf{E}$ that bears a quadratic relation (as $I \propto |\mathbf{E}|^2$) to the potential $\mathbf{E}$ that gives rise to it. In spite of this quadratic dependence, useful information of the underlying relation between the potentials may, however, be extracted as follows. The intensities $I(1)$ and $I(2)$ are measured by the

[35]With the proviso that these are in coherent motions.

alternative closure of one slit while observing the other. From the values of these measurements, one computes $\mathbf{E}_1$ and $\mathbf{E}_2$ respectively. The next step consists in taking the vectorial sum of $\mathbf{E}_1$ and $\mathbf{E}_2$ to give $\mathbf{E}_R$, the resultant electric-field. The quantity I_R, the intensity of the resultant can now be calculated theoretically and its value is compared against the actual observed intensity $I(1+2)$ when both slits are open and the two wavy potentials through the slits are allowed to act together. One finds,

$$I(1+2) \;=\; I_R \tag{5.28}$$

The absence of discrepancy in the observed and theoretical values of the resultant intensity of the two light waves brings out a crucial point. In this case, even though the observed effects $I(1)$, $I(2)$ and $I(1+2)$ being measured do not in fact stand in a relation of superposition (i.e. $I(1+2) \neq I(1)+I(2)$), the underlying potentials $\mathbf{E}_1$, $\mathbf{E}_2$ and $\mathbf{E}_R$ that give rise to them do satisfy such a relation ($\mathbf{E}_1 + \mathbf{E}_2 = \mathbf{E}_R$), for if they do not, taking the modulus square of $\mathbf{E}_R$ (obtained by simply adding $\mathbf{E}_1$ and $\mathbf{E}_2$) then provide us with an intensity, say, I'_R, that would *not* have been the same as the observed intensity $I(1+2)$. In that circumstance, I_R would then be seen as an indication of some kind of interaction that has taken place between $\mathbf{E}_1$ and $\mathbf{E}_2$ when they are put together to work (see Figure 5-11). Hence, the very fact that equation (5.28) above holds is good evidence for the assertion that the direct sum $\mathbf{E}_R$ of the two independent potentials[36] $\mathbf{E}_1$ and $\mathbf{E}_2$ is indeed the *correct* potential that is responsible for the observed interference effect.

The conclusion one may draw from the above argument is that the two periodic potentials remain physically independent of each other as they operate at the same time, in pretty much the same way that mechanical forces do. What is remarkable is that since the potentials are of a periodic nature, the difference in their phases, which is a relation only existing when *both are present*, manifests itself in the effect[37] (the

[36] These act independently as may be verified by the fact that they can be separately subjected to individual measurements as in the above.

[37] The crossed-terms responsible for the interference effects appear as we take the modulus squares of $\mathbf{E}_1$ and $\mathbf{E}_2$ to obtain the intensity that agrees with the observed quantity.

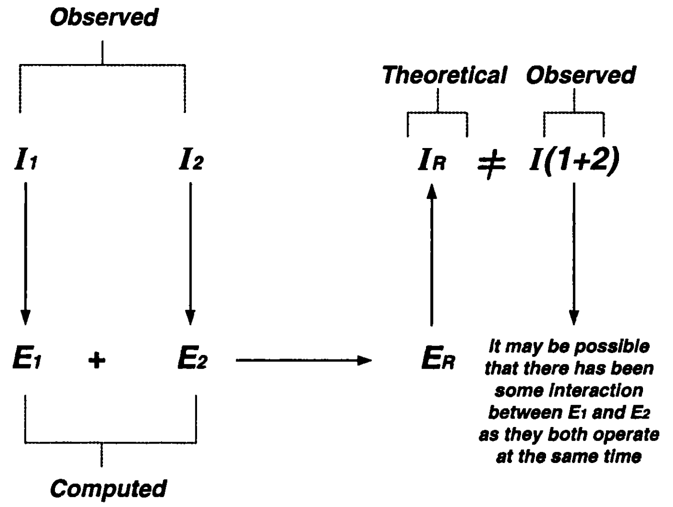

Figure 5.11: Possible interaction between two electric-fields

interference pattern observed) when they are brought together to act at the same time.

Painstakingly, we have gone into great length to discuss the principle of superposition in both the contexts of *non-periodic* and *periodic* potentials. The point I am trying to put across is this: the principle of superposition is often stated as applied to the effects with respect to the potentials from which they originate. However, it is amply clear that there are *two levels of superposition at work.* One is the superposition of the effects and the other is what I would call the superposition of the potentials that produce these effects. In the previous example of non-periodic mechanical forces, the principle of superposition applies on both the level of effects ($\mathbf{a}_1 + \mathbf{a}_2 = \mathbf{a}_R$) as well as that of the potentials ($\mathbf{F}_1 + \mathbf{F}_2 = \mathbf{F}_R$) because both the potentials and their effects are related in a *linear* manner. But in the case of periodic potentials, the principle fails on the level of effects ($I(1) + I(2) \neq I(1+2)$) while it is still found to hold on the level of potentials ($\mathbf{E}_1 + \mathbf{E}_2 = \mathbf{E}_R$).

We can therefore have situations where potentials acting independently at the same time to produce an overall effect that differs from the addition of the individual effects produced separately by each of the potentials acting alone. The "difference" between the observed resultant effect and the simple theoretical sum of the individual effects consists in the important fact that a special kind of relation (it is the phase difference in the case of waves) manifests itself only when *all the potentials are in action simultaneously*, as a consequence of which this "*aggregate property*" gets translated into the observable effect, while the respective actions of the potentials remain unaffected by those of the others[38].

Of these two levels of superposition, it seems to me that the superposition of potentials is more fundamental because it provides the assurance that the potentials maintain their physical independence when they operate simultaneously in the same region of space (and on the same medium); and much of our physics is founded on this important supposition. One may still suspect that the distinction of these two levels of superposition is merely a verbal exercise. This distinction, however, is of a significant impact when one comes to consider the concept of probability, the subject to which we would promptly turn.

The cast of a die

Consider the simple case of throwing a fair die. Given the perfect symmetry of the die, we assert that there are six equally likely possible outcomes but *only one* would materialize in the next throw of the die. By the six possible outcomes, we mean the six possibilities that correspond to each of the six faces of the die to show in the action of one throw of the die. These are six "*potential*" eventualities associated with this physical situation.

The six possibilities are said to be of equal chance to occur given the symmetry of the die. The phrase "of equal chance" may be operationalized by conducting a long sequence of throws of the same die[39] and recording the number of throws in which each face obtains. If the total number of throws is sufficiently large, the proportion of the number of throws in which a particular face shows to the total number of throws would become very close to one-sixth in the long run[40]. The fraction "$\frac{1}{6}$" is thus taken

[38]Otherwise, a constant phase difference may not be maintained and the interference effects become unattainable.

[39]Theoretically speaking, one may also adopt the equivalent procedure of throwing a large number of unbiased dice of the same kind simultaneously.

[40]Strictly speaking, this is only achieved in the limit where the number of throws approaches infinity but in practice, a large number of throws is usually considered acceptable. See the discussion of the Laws of Large Number by Renyi (1970). Renyi, A. (1970).

as a *numerical measure* of the potential of, say, the face 4 of the die showing.

This immediately leads us to the characteristic feature of indeterministic, chancy events: one may speak of the potential of a given situation being present or manifesting but there is *no certainty* of a potential to actualize a particular result, in contrast to the cases of both the non-periodic mechanical and periodic wavy potentials. The "*probabilistic*" potential, which nevertheless represents a *tendency* to realise a specific result, only reveals itself retrospectively in a long sequence of repetitive measurements.

One often speaks about an entity *tending* towards a particular position when all that is meant is the entity needs not necessarily take up that particular position but in *a large proportion of times* it does. Thus, the "particular position" provides a reference point for the entity to tend towards in the long run. The concept of *tendency* (a *potential* in the probabilistic context) presupposes the idea of repeated measurements; one is only able to measure such a tendency through a sequence of repeated measurements.

In the limit, a potential like a mechanical force is found to *always* produce the *same* result - the *same* acceleration on a physical system. In this *deterministic* setting, we see an extreme degree of tendency of the system to move at a certain acceleration when it comes under the influence of a particular force. Here, the system would be left with no alternative but to accelerate by the *same* amount when the *same* force is applied.

One usually finds this not too illuminating as one can quite easily supply the argument that, after all, we know that there is one potential - the one mechanical force acting - and this always provides the same acceleration. To raise the question of how one knows that this potential "always produces" the same acceleration would no doubt provoke an extensive excursion into the terrain of the problem of induction. We shall not detain ourselves with such a consideration.

But what I would focus on is a subtle difference between a deterministic potential in classical mechanics and a probabilistic one such as, for instance, the tendency of the face "4" of a die showing in one throw. In the former, we are confident that the acceleration of the mass is produced by the applied force because we are certain about the action of this one potential, and also because that we are at will to subject the mass to the influence of this one potential; to manipulate the potential so to speak. In contrast, in the latter case of the cast of an unbiased die, we are faced with six *potentials* ***or*** *tendencies* and there is only one possible outcome out of the six that would be realised upon the throw[41]. There is therefore, a *one-to-one correspondence* between the acting potential and its effect in the former but this correspondence is found to be absent in the latter: before the throw of the unbiased die, the six potentials $(\phi_a, \phi_b, \phi_c, \phi_d, \phi_e, \phi_f)$ are all present but only one outcome materializes once it is cast (Figure 5-12).

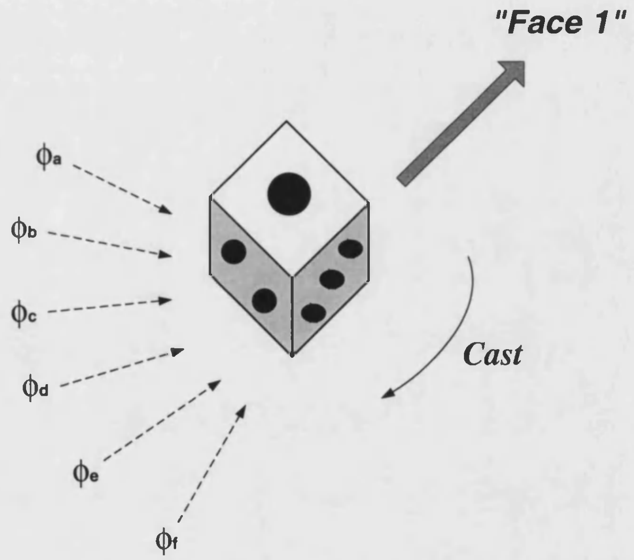

Figure 5.12: Six probabilistic potentials present before the casting of an unbiased die.

The claim that the six potentials co-exist before the throw is based on the physical symmetry of the system in question and hence, there does not seem to be any contrary reason for us to think that there should be a smaller or greater number of potentials in operation. The co-presence of the six potentials does, however, in no way commit

[41]Moreover, these probabilistic potentials may not be manipulated in the sense that one cannot be sure which one of the six possible outcomes would realise in the next throw of the die. One may perhaps suggest at this point that the potentials may be "manipulated" by physically altering the die to a biased one in which all other potentials but one are effectively eliminated. However, since the biased die represents a very different physical situation to the unbiased one, this sort of manipulation can then be ruled out.

them to be in action at the same time *or* for each of them to be in a one-to-one correspondence with the six possible outcomes[42]. Several tentative scenarios thus arise. (1) One may take the view that *all* the ϕ's operate in such a manner to result in just one outcome materializing. This would mean, for example in Figure 5-12 that, the potentials $\phi_a, \phi_b, \phi_c, \phi_d, \phi_e$ and ϕ_f all operate to generate the outcome that "face 1" results upon a cast of the die. This is in analogy with the case of mechanical forces where the component potentials act together to produce a single resultant effect. Here, the outcome that has materialized is thought of as the resultant effect of the actions of all the potentials. And in order to account for the reason why this particular outcome has come to dominate rather than the other five[43], some intricate mechanisms of the potentials in actions would have to be supposed. (2) One may also take the view that only some of the ϕ's are in action. An example would be that, say, only the potentials $\phi_a, \phi_c, \phi_d, \phi_e$ have operated (but not ϕ_b, ϕ_f) to generate the outcome of "face 1" of the die resulting. As similar to the reasoning in (1), how this outcome has come to be realized as a consequence of the aggregate effort of these ϕ's must again be explained. (3) Given the fact that, say, "face 1" has actualized in one throw of the die, we infer retrospectively that *only* the potential ϕ_a, the tendency exclusive to the production of the outcome "face 1", and no other, has acted on this particular throw.

Our usual method of measurement of these probabilistic tendencies (or potentials) - the assignment of probability measures - by reference to relative frequencies has a great deal to do with why it seems more reasonable for us to favour the view given in (3) over those of (1) or (2). We shall now try and provide a justification of this reasoning.

To dispose of the views in (1) or (2), it is useful to consider an analogous situation in the case of physical potentials. Let us suppose there are two forces $\mathbf{F}_1$ and $\mathbf{F}_2$ present at the same time but unlike $\mathbf{F}_1$, $\mathbf{F}_2$ is not acting on the mass m (Figure 5-13). As a illustration, we may think of $\mathbf{F}_1$ as an ordinary mechanical force and $\mathbf{F}_2$ as,

[42]Hence, lower-case letters a, b, c, d, e, f have been used to label the six potentials in order to avoid possible confusion if these are labeled by the numbers 1 to 6 instead.

[43]That is, it is "face 1" that turns up rather than the five remaining faces of the die.

perhaps a magnetic force generated by a small magnet. Provided the mass m is not composed of magnetic material (but a small plastic sphere, say), the motion of m would not be affected by the presence of $\mathbf{F}_2$ on a macroscopic level.

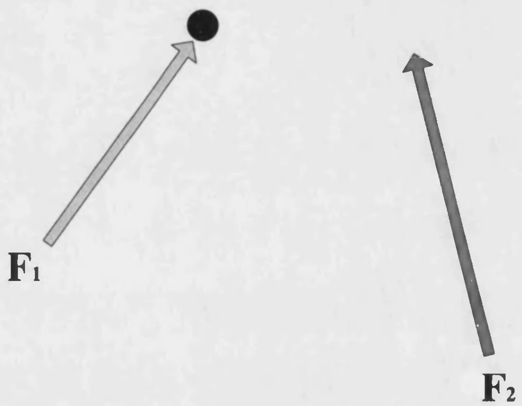

Figure 5.13: Two physical forces enjoy a simultaneous presence but not in operation at the same time.

In this circumstance, the acceleration of mass m in the presence of *both* $\mathbf{F}_1$ *and* $\mathbf{F}_2$ is measured to be of the same value as the one due to $\mathbf{F}_1$ acting alone (with the magnet removed for example). That is, only the effect $\mathbf{a}_1$ (the acceleration of mass m produced by $\mathbf{F}_1$) of $\mathbf{F}_1$ manifests, given that $\mathbf{F}_1$ and $\mathbf{F}_2$ are both present. One may then infer that $\mathbf{F}_1$ alone acts on the occasion even in the presence of $\mathbf{F}_2$.

This kind of reasoning motivates the analogy with probabilistic potentials as sketched under view (3) above: although six probabilistic potentials $\phi_a, \phi_b, \phi_c, \phi_d, \phi_e$ and ϕ_f may all be present, it is, however, due solely to the action of ϕ_a that the outcome of "face 1" shows. Similar to the case of mechanical potentials, such a view is, of course, based on the assumption that there exists a one-to-one correspondence between the potential and its effect. This one-to-one correspondence is what underlies the relative frequency method as a means of measuring probabilities. For the case of the casting of a die, we take the fraction,

$$\frac{Number\ of\ throws\ with\ \text{“}face\ 1\text{”}}{Total\ number\ of\ throws} \tag{5.29}$$

as meaning the tendency for the die, out of a long sequence of throws, to produce a certain number of outcomes with "face 1" showing.

In order words, we choose this fraction, this number, as measuring the *one* potential that gives rise to *this* particular outcome and hence a *one-to-one correspondence* is established. Likewise, we collect all the results with "face 2", "face 3" and so on, to obtain similar fractions as the measures of the respective tendencies of "face 2 showing", "face 3 showing" etc. Such one-to-one correspondences between the tendencies towards particular results and the results themselves are implied by the method we adopt for the measurements of these tendencies. So on each occasion where a certain "tendency" has been seen to have operated, the other five potentials, though with their presence prior to the cast as real as $\mathbf{F}_2$ as in the above example, have somehow become "barred" from their actions.

The difficulties that plague the views expressed in (1) and (2) (and thus induce us to pursue view (3)) can be traced to the property of mutual exclusiveness of the outcomes of an experiment of probabilistic character. It is, indeed, a hallmark feature of probabilistic considerations that the causes for mutually exclusive possibilities (the tendencies or potentials) can co-exist before an event but the same cannot be said of the respective eventualities that each of them leads to. The problem hinges upon the usual, and perhaps misguided, impression that being co-existing, all tendencies would have to exert influence or to act on the die at the same time. Following this direction of thought, we find ourselves in deep puzzlement with regard to the transition that would have to take place in order to go from a situation where *all* potentials act prior to the cast, to one where only a single potential has arguably been in actual action as known from after the cast. However, by carefully distinguishing between a situation where all potentials may be *present but not acting* and one that where they are *all present and acting*, one is able to relieve oneself from such a confusion.

In a metaphysical vein, one may proceed further and argue that the very fact that "six" (and *no more* or *no less*) distinct potentials have been associated with the die is due entirely to the six distinct outcomes used as the measures according to the relative frequencies method. Hence, this one-to-one correspondence is implicit when "six" potentials are referred to. It follows that the question of whether all these six potentials or any combination of them act concurrently on the die to produce a certain result would become automatically dissolved.

To make our argument more transparent, we consider a number of probabilistic potentials $\phi_I, \phi_{II}, ..., \phi_n$ where n can be any natural number $(n \in N)$. We do not require any one of these ϕ's to be put into a one-to-one correspondence with the eventual outcomes. Under such a circumstance, the ϕs in action must "interact" in some ways in order to produce the six outcomes. Given this flexibility, the ideas expressed in views (1) and (2) thus become plausible. There are, however, two difficulties. First, an explanation, most probably a mechanism in terms of some intricate "interactions" amongst the ϕ's must be provided. Second, there exists, in theory, no upper limit to the number of potentials that can be brought in, and hence, there is every possibility of an infinite number of the various combinations of the deifferent ϕ's that may lead to the same outcome. I am inclined to reject both of these hypotheses on the grounds of their inherent non-testability.

As discussed, our method of measuring these probabilistic potentials and their observed outcomes affords us with the view of a one-to-one correspondence between a probabilistic potential and "its" effect. Of course, the fact that prompts us to feel even the slightest discomfort in imposing this correspondence is due entirely to our inability to "isolate" one single potential and place the die under the influence of this particular potential. Unlike mechanical forces, these probabilistic potentials are not open to being isolated. It is worth emphasizing that it is characteristic of probabilistic potentials to be known to have operated *only after* their outcomes are actualized.

A further worry ensues. The discussion so far on probabilistic potentials has undoubtedly conveyed to the reader an image of a reality of these potentials that seems to be on a par with that of a "real" physical potential like a mechanical force. The worry concerns the justification of conferring on these probabilistic potentials such an ontological status. As a mark of prudence, we ought to provide some rationale for why one is permitted to think in this manner.

There is at least one well-known view on probability theory that takes an affirmative attitude on this issue. In his development of an objective concept of probability in physics and the sciences, Sir Karl Popper adopts this view[44],

> ***Propensities, it is assumed, are not mere possibilities but are physical realities. They are as real as forces, or fields of forces. And vice versa: forces are propensities. They are propensities for setting bodies in motion. Forces are propensities to accelerate, and fields of forces are propensities distributed over some region of space and perhaps changing continuously over this region. Fields of forces are fields of propensities. They are real, they exist.***

If the one-to-one correspondence between a probabilistic potential and the effect it generates is assumed, we are then led to conclude that in each throw of the die, all six potentials ϕ's are co-present but only one out of the six actually operates to bring about the outcome. I should, however, point out that the supposition of such a one-to-one correspondence does not come into any conflict with the doctrine of indeterminism. The crucial point is, after all, the knowledge of a particular ϕ has acted on a particular occasion is obtained *only after* the corresponding outcome has materialized[45]. Since the supposition that *a certain potential ϕ has acted* may only be inferred retrospectively after the throw and not before, our knowledge of a particular ϕ to operate in the next trial remains entirely dictated by chance.

In the language of sets, the *mutually exclusive* or *disjoint* sets of outcomes lead directly to the property of finite additivity of probability measures - the *sum-rule* of probabilities. The *sum-rule* is an expression of the fact that the members of the union of a collection of disjoint sets is equal to the sum of the individual members belonging to each of the constituent sets. In the example of the cast of a perfect die, the sum-rule is stated as follows,

---

[44]Popper, K.R. (1990), p.12.

[45]With mechanical potentials, we may at will subject the system to one single potential and measure its result. With probabilistic potentials, we cannot "isolate" one potential ϕ from the others and "subject" the die to this particular potential.

$$P(1\ or\ 2\ or\ 3\ or\ 4\ or\ 5\ or\ 6) \;=\; P(1) + P(2) + P(3) + P(4) + P(5) + P(6) \quad (5.30)$$

If each of the P(1), P(2), ..., P(6) represents a fraction of the total number of throws, this would simply mean that the sum of the number of throws in which "face 1" *or* "face 2" *or* so on obtains is equal to the total number of throws carried out. This is a trivially true statement as we have measured each of the individual P(1), P(2),...etc. by partitioning the total number of throws that have been performed.

However, notice the use of the disjunctive connective "*or*" in the above. This is essential because the possible outcomes are mutually exclusive with respect to *one throw.* Two or more outcomes cannot be occurring *at the same time. This means that the sum-rule of probability is <u>not</u> an expression of the principle of superposition in probabilistic examples - neither superposition of potentials nor superposition of effects.* This is a vital point to bear in mind when one comes to unravel the intricate relationship between *the principle of superposition* and *the sum-rule of probability* in the double-slit experiment.

The Double-Slit Experiment Revisited: the Sum-Rule - versus - the Principle of Superposition

In the double-slit experiment, the randomness lies with the emission of photons from the source. The source emits each photon in a random direction but once it is emitted in a particular direction, the photon would follow a *definite trajectory* was it to behave as a corpuscular entity on a particle picture. The two slits, thus, merely serve as filters to select out those photons that travel in these two specific directions out of the total number of photons, which have been emitted in all the different directions. Their function resembles our selection by inspection of the number of throws of a die in which "face 1" or "face 2" shows in a sequence of throws for example (the sequence, of course, includes throws with faces other than these two faces showing). As selectors of two particular subsets of photons, the presence of the slits does not influence the photons in any other significant way apart from the scattering occurring at the edges of the slits. The two distributions of photons coming through slit 1, $D(1)$, and through slit 2, $D(2)$, show two distinctly separate subsets out of the set of all the photons generated at the source. One should then expect to find the total number of photons passing through slit 1 *or* slit 2 to be equal to $D(1) + D(2)$ given the two assumptions that:

(i) each photon has a *definite trajectory* in a certain direction once emitted in that direction and,

(ii) the presence of the slits serves *only* the function of selecting out the photons which travel in two specific directions (out of the total number emitted) and *they do not in any way alter the motions of the photons significantly.*

So, one expects the simple relation $D(1\,or\,2) = D(1)+D(2)$ to hold with $D(1\,or\,2)$, $D(1)$ and D(2) denoting *the total number of photons passing through either slit 1 or slit 2*, *the number of photons passing through slit 1* and *the number of photons passing through slit 2* respectively.

Translating these values into probability measures that are taken as relative frequencies, we obtain,

$$\frac{D(1\ or\ 2)}{N} = \frac{D(1)}{N} + \frac{D(2)}{N} \tag{5.31}$$

where N is the total number of photons emitted by the source in any direction.

Now equation (5.31) may readily be rewritten as,

$$P(1\ or\ 2) = P(1) + P(2) \tag{5.32}$$

with $P(1\,or\,2)$, P(1) and P(2) denoting the corresponding fractions in equation (5.31) respectively.

We have arrived at the "*sum-rule of probability*" (equation (5.32)) via the physical distributions of photons *and* the two assumptions (i) and (ii) stated above. Hence, the validity of the sum-rule for the probability of photons at the slits is dependent on these two assumptions. In particular, the assumption in (i) that each photon follows a definite trajectory has served the important purpose of "partitioning" the photons into disjoint subsets; that is, once it is sent travelling in a specific direction, a photon would continue in that direction if its motion is left unhindered (albeit the possible scattering as it traverses a slit). We shall, however, see below that this assumption, and in fact, both of the assumptions (i) and (ii) are to be abandoned in the reconciliation of our thinking with quantum phenomena.

Because of the fact that the D's are representations of physical effects, it would perhaps seem natural for one to associate the relation $D(1 \, or \, 2) = D(1) + D(2)$ with the superposition principle. Above all, is it not so that the relation expresses the fact that *the combined effect equals the sum of the individual constituent effects*? To entertain such a correspondence, as we shall see, would be a mistake, for, as we argued in the last sections, the crux of the principle of superposition (whether of potentials or of effects) is that the resultant is supposed to be the aggregate result of the individual effects *happening at the same time* due to the *simultaneous actions of the respective potentials* that give rise to them. One sees at once that this character of the principle of superposition is at odds with the idea of disjointness, which implies that the potentials cannot be simultaneously in operation. But there are more subtleties to consider before a conclusion may be reached.

It is customary to think that whenever the principle of superposition obtains, one "observes" several effects present at the same time, each due to a different potential. In reality, however, one never observes these several effects simultaneously in a strict sense. One perceives only the resultant, the *one* effect that can *be proved to be equal to* the sum of the individual effects indirectly, as has been already discussed above in both the situations of non-periodic mechanical forces and periodic wavy potentials. Despite this reality (which incidentally forms a crucial point of our argument to come), it seems harmless to continue thinking that "several effects occur at the same time whenever the resultant obtains" provided the principle of superposition holds. But this thought is immediately met with a difficulty when one considers probabilistic potentials.

In the classic probabilistic situation of die casting and also in the double-slit scenario (particularly apparent with the low photon density case with one-photon being emitted at a time), it would be very odd to think of two simultaneous effects of the same kind[46] existing in the same system because of the mutually exclusive nature of these effects.

In the case of die casting, I have argued that it is incorrect to say that each of the results, which corresponds to the showing of each face of the die, is a "combined" or "resultant" effect of the potentials. This is in part because it is not possible to subject the die exclusively to any individual one of these potentials and then ascertain the particular effect due to it[47]. Our knowledge of these potentials is only retrospective. The best one can achieve is to infer from the observation of an outcome that the *one* potential, which has given rise to the outcome[48], has operated in that particular instance. Our method of assignment of the probability measures has induced this kind of inference. The difficulty lies in the fact that we are not free to manipulate the die (before the cast) so that we may know that just one or several of these potentials are acting.

Furthermore, because of this one-to-one correspondence we have assumed for a particular outcome i and the potential ϕ_i that is supposed to bring it about, the disjointness of the set of all outcomes $\{i\}$ implies also the disjointness of the operating ϕ_i's. In physical terms, this is tantamount to the statement that the ϕ_i's are prohibited from acting together at the same time. But the ability of the potentials to be in action at the same time is necessary if the corresponding effects are to be produced together, which is vital to the idea behind the principle of superposition, we therefore conclude

[46]By simultaneous effects, we require that more than one potential are acting on the system at the same time. Since we have assumed that there are one-to-one correspondences between the potentials and their outcomes, it follows therefore that the disjoint outcomes make it impossible for two potentials to act simultaneously on the die in one throw.

[47]In the case of non-periodic mechanical forces, the resultant we observe (albeit it appears to do not harm to think of it as several effects occurring at the same time) is one single effect when the system is placed under the influence of more than one potential at the same time. In principle, the resultant is one single individual effect in its own right, it is equal to the sum of the effects only in the sense that the former has been found to be empirically equivalent to the latter. We observe *just* the overall effect and not the several individual effects at the same time.

[48]Assuming, of course, a one-to-one correspondence between the potential and the effect it generates.

that the relation $D(1 or 2) = D(1)+D(2)$[49] and its derivative, the sum-rule in equation (5.31) that arises from the disjointness property of the potentials, are not expressions of the principle of superposition (neither of potentials nor of effects).

Now comes the new twist in the tale. Experimentally, we find when both slits are open, we do not obtain the resultant distribution $D(1\ or\ 2)$ as given by the simple sum, $D(1) + D(2)$, when no attempt is made to observe which slit each photon has come through. Rather, we observe a "resultant" distribution $D(1\ or\ 2)$ that exhibits an interference pattern characteristic of the superposition of two periodic potentials or forces.

Recall that earlier in our analysis, we argued that in the two-slit experiment, superposition does not apply on the level of effects (interference), it nonetheless applies on the level of the potentials that generate these effects. In addition, the phase difference, a kind of "relational" property brought about by the simultaneous presence of both potentials that act together on the same system, manifests itself in the resultant effect. We also note that in the case of optical interference, the underlying potentials obey the superposition relation: $\mathbf{E}_R = \mathbf{E}_1 + \mathbf{E}_2$, which states explicitly the fact that the potential resulting from the two potentials acting together is simply the (vectorial) sum of the two individual potentials.

Applying the same reasoning to the double-slit experiment, we now demand both of the potentials $\phi(1)$ and $\phi(2)$, which correspond to the two tendencies for the photon to pass through slit 1 and slit 2 respectively, to *act together on the same photon at the same time* in order to produce the interference pattern, that is,

$$\phi(1\ or\ 2) \;=\; \phi(1) + \phi(2) \tag{5.33}$$

As similar to the foregoing discussion on real physical waves, it may be readily verified[50] that,

[49]Which strictly, is a relation amongst the effects but is deemed to hold also for the potentials because of the one-to-one correspondence between the potential and the effect.

[50]$\phi(1)$ and $\phi(2)$ are obtained separately by observing the two distributions $D(1)$ and $D(2)$. Since $D(1)$ and $D(2)$ represent the physical intensities we measure, they are thus proportional to the potentials $\phi(1)$ and $\phi(2)$ that give rise to them. Having evaluated $\phi(1)$ and $\phi(2)$, the sum of these two quantities are taken to give $\phi(1)+\phi(2)$ and the modulus square $|\phi(1)+\phi(2)|^2$ of which is compared

$$|\phi(1+2)|^2 = |\phi(1)+\phi(2)|^2 = |\phi(1)|^2 + |\phi(2)|^2 + crossed\,terms \qquad (5.34)$$

At this juncture we find ourselves landed on unfamiliar landscape. The probabilistic potentialities whose actions are supposedly mutually exclusive can nevertheless be in operation on the same physical system (the photon) at the same time[51]!

The Quantum Priniciple of Superposition

Now that quantum mechanics has revealed to us such strange modes of behaviours of the probabilistic potentials, it is appropriate to assess its impact on our conception of physics. We are concerned with, in particular, the status of the assumptions in (i) and (ii) in light of these new results. In this regard, one must look into the two contexts of "*observation*" and "*non-observation*" of photons at the two slits; as these are the conditions under which the two modes of action of the potentials are brought to light and would have important repercussions for both assumptions. Let us first state the relevant facts with respect to each of these scenarios.

(a) *Observation of photons at the slits*

Each photon is observed to traverse one of the two slits when appropriate devices are situated at each of the two slits. The consequence of the assumption in (i) that photons follow definite trajectories agrees with the observations. Moreover, the overall distribution of photons with respect to both slits staying open is found equal to the sum of the two individual distributions obtained by observing each individual slit in turn. In this case, the sum $D(1)+D(2) = D(1\,or\,2)$ holds and it may be translated trivially, as we have seen in equation 5.31, into the sum-rule of probability. The very fact that this simple sum holds provides convincing evidence for the slits behaving as mere "selectors" of two subsets of

to $|\phi(1+2)|^2$ as observed.

[51]When no effort is spent to find out which slit the photon has come through.

photons that are emitted in these particular directions; for they merely help us to pay attention to photons travelling in these two directions. The presence of the slits is supposed not to affect the prescribed trajectories of the photons in this case.

We have argued previously that the sum $D(1) + D(2) = D(1 \textit{ or } 2)$ is *not* a statement of the principle of superposition (neither of potentials nor of effects). As far as actual observations at the slits are concerned, the probability of finding a photon at a location x (on the detection screen) obeys the simple sum-rule of the classical probability calculus. The respective observed photon distributions at the two slits conform to the assumption that each photon *has followed a definite trajectory all along.* The definiteness of the photon trajectories make us certain about the mutual exclusiveness of the probabilistic potentials. No two potentials are in operation on the same system at the same time[52]. In the presence of observational procedures carried out in the vicinity of the slits, the probabilistic potentials seem to behave *as if* the photons have had well-defined definite trajectories once they are emitted from the source.

(b) *Non-observation of photons at the slits*

In the event where no effort is made to observe which of the two slits a photon has traversed, one is left with a very different overall distribution of photons: one that shows the characteristic pattern of optical interference. The appearance of this "interference modulation" leads us to reject the assumption that the photons have, after all, followed definite trajectories all along[53].

Since a distribution of photons without interference modulation is a consequence of the hypothesis that each photon travels a definite path, the appearance of interference would then mean that this hypothesis is to be given up in the first instance. Hence, we conclude that if no attempt is made to locate the photons

[52]As evidenced by the distinct clicks at the detectors placed at the slits.

[53]For if they had, the overall distribution with both slits open would equal to the sum of the individual distributions that correspond to photons observed to have traversed each separate slit (such as in case (a) above).

at the slits, it cannot be said from which respective slit each photon has come through, and therefore, *we are not allowed to claim that the photons have been following definite trajectories.*

The presence of interference modulation of the overall photon distribution is attributed to the superposition of the probabilistic potentials $\phi(1)$ and $\phi(2)$. As long as no observation is made to discover through which slit a photon has passed, these potentials, far from enjoying a mutually exclusive relation in their operations, now act in a simultaneous[54] manner on the photon. The superposition of probabilistic potentials is preserved insofar as one shows no attempt to track down the photon.

However, when no observation is made to locate the photons at the slits, both slits seem to play an active role in attaining the interference effects[55]. The observed distribution is one that is characteristic of double-slit interference. One could have also worked with a larger number of slits, three or four for example, and been able to obtain patterns that show the characteristics of three and four-slit interference respectively.

In the light of the evidence, it seems inevitable that we would have to, at the very least, re-adjust our conviction on the much cherished assumptions of (i) and (ii) that are attached to a classical particle. Classically, our investigation into the nature of a physical system through the various performances of measurement on it is deemed to be an exact representation of its underlying reality. Our observations thus faithfully reveal this underlying reality. The reality is thought to be objective and this amounts to a statement of the kind that the moon *is* there with a definite position (relative to some vantage point) even if not a single soul cares to gaze at it. In the quantum realm, it is no longer possible for us to say that the quantity we measure is what it is *as if the measurement had not happened.* To put the point bluntly, our observation of an

[54]As characteristic of being in a superposition.

[55]Contrast this with the occasion where a photon has been observed to traverse a slit and followed a definite trajectory, the slit serves mainly the function of selecting and helping us to concentrate on that particular subset of photons (out of the total number emitted by the source). The slit does not in any significant way influence the emitted photons except for minor scattering effects at its edges.

electron seems to force it into a behaviour that is not what it would have manifested if left on its own.

I shall now sum up the main findings in this section. Potentials are dispositions or capacities arising from a given physical situation that induce a physical system to generate certain observable effects. There are probabilistic as well as non-probabilistic potentials. A mechanical force acting on a mass produces an acceleration of the latter is a good example of a non-probabilistic potential; while the disposition to produce "face 1" (say) in a long sequence of die rolling counts as a probabilistic potential.

Probabilistic potentials differ mainly from their non-probabilistic counterparts in that their actions may only come to be known retrospectively *after* a long sequence of measurements. Our usual method of assigning probabilities through the relative frequencies of a set of mutually exclusive outcomes leads us to the conclusion that *one and only one probabilistic potential may operate in a particular trial.* In other words, a *one-to-one correspondence* is established between each probabilistic potential and the outcome that it gives rise to, which thus means that these potentials themselves form a mutually exclusive set. This feature of mutual exclusivity applies not to the possibility of co-existence of the potentials but rather to their respective actions on the *same* physical system at *a* time: the potentials may co-exist but yet may not necessarily be in concurrent operations. In fact, implicit in the frequency interpretation is the message that the potentials *cannot* be in concurrent operations, which is in accordance with classical probabilistic reasoning. As such, the sum-rule is not a statement of the principle of superposition, since potentials in a superposition must act concurrently.

Quantum mechanics reveals to us a very different facet of these probabilistic potentials, but certainly not one that should have taken us by great surprise as generally is popularized[56]. When speaking about the dynamics of a system in the probabilistic

[56]Since our knowledge with regard to the actions of probabilistic potentials can only be gained retrospectively, we are thus incapable of isolating one particular potential and controlling its operation on a system in order to bring about a particular outcome. Because of this, we cannot determine, before the event (or experiment), which of the two possibilities - of *the potentials being mutually exclusive in their actions in one trial* **or** whether *they all (or any combination of them) act simultaneously to produce the desired result* - really obtains. Our assignments of probability measures in terms of relative frequencies seem to have eliminated the latter possibility in favour of the former: the probabilistic potentials do not act together according to our usual interpretation of probabilities

context, as in the Young's double-slit illustration, the origin of the mutual exclusiveness of the probabilistic potentials is attributed to the physical fact that classical particles follow definite trajectories in space and time. In the absence of any effort to find out about the information regarding the spacetime trajectory of a quantum particle (without observations at the slits for instance), we uncover that the probabilistic potentials can indeed *all* act concurrently to produce the interference distribution that is observed. The conception of a particle having a definite spacetime path must be given up.

The foregoing observations can be visualized pictorially in the contexts of the "classical die" (Figure 5-14) and the "quantum die" (Figure 5-15).

For the classical die, the probabilistic potentials are measured by the frequencies of their outcomes relative to the total number of throws. In figure 5-14, only after, and not before the outcome "face 1" has shown are we able to infer that ϕ_1 has acted. We are unable to exert any control over which particular ϕ's is to operate on the die in the next trial. This feature of retrospective knowledge contributes largely to the shaping of probabilistic reasoning.

Especially, because of our lack of control of these potentials, it is meaningless to speak about which outcome is to actualize in the *next* throw of the die. The "strength" of these potentials are measured by the frequencies of occurrence relative to each other in a long sequence of trials. Chance lies in the fact we cannot know for sure which outcome is to materialize in the next trial to come.

The quantum mechanical results bestow on us an entirely novel feature of these probabilistic potentials. On every occasion, *all* six potentials operate concurrently on the "*quantum die*" (Figures 5-15 and 5-16). Furthermore, these potentials must exhibit periodicity by necessity, for it is the overall difference in their relative phases when the potentials are operating simultaneously, which accounts for the modulation of the distribution of the outcomes by interference.

as relative frequencies.

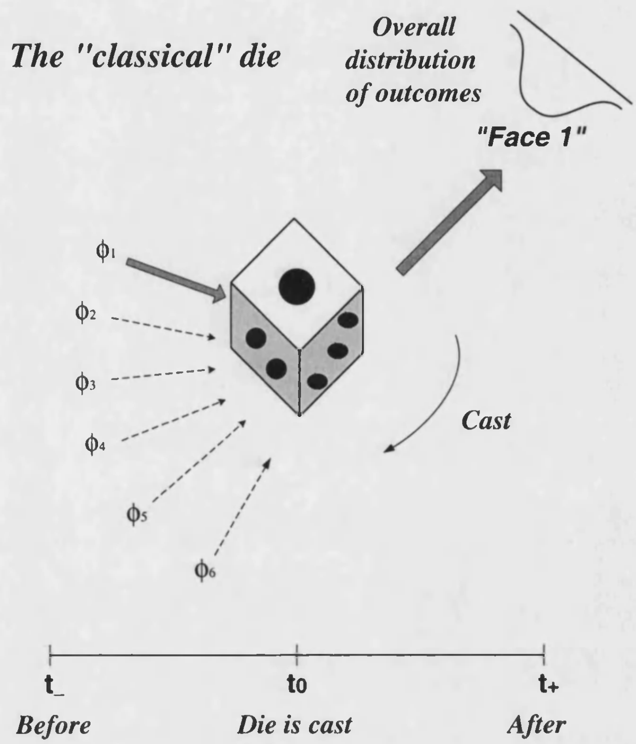

Figure 5.14: The classical die.

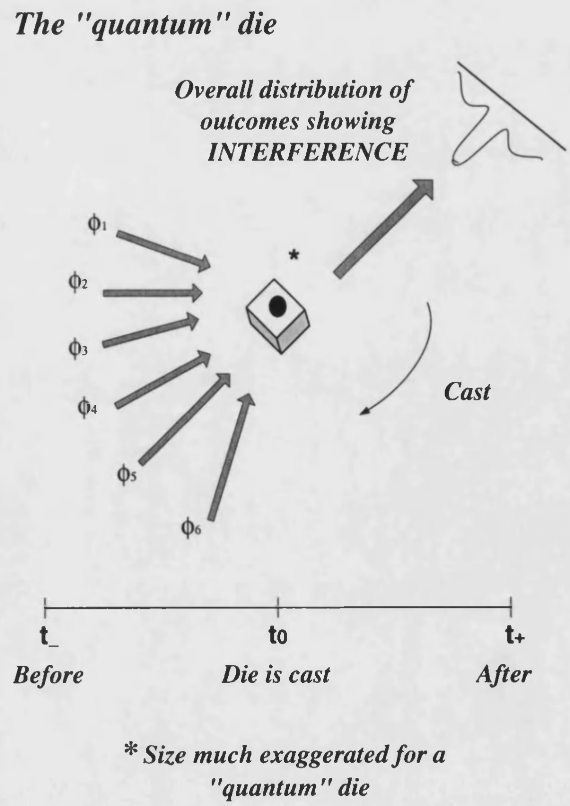

Figure 5.15: The quantum die.

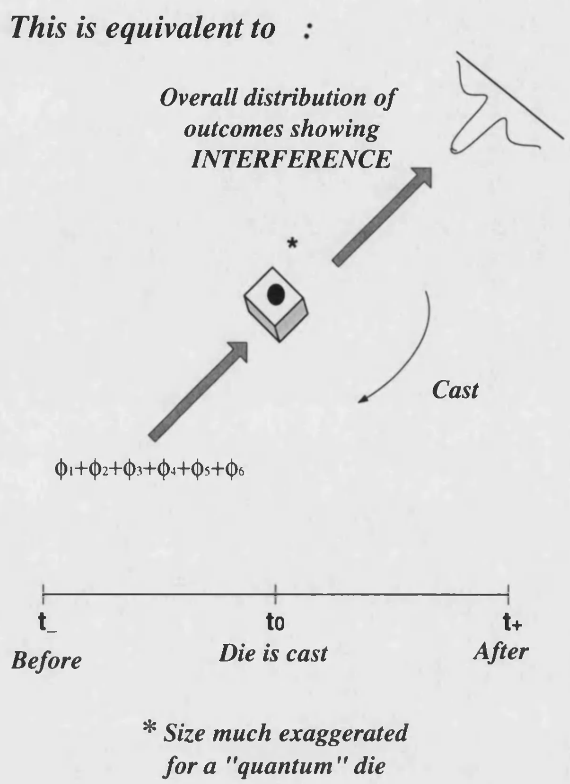

Figure 5.16: The quantum die.

Granted that all potentials must act simultaneously to result in a statistical distribution with interference modulation, would this signal the return of a sort of determinism (with respect to the behaviours of the potentials) in the naive sense that it can be determined that all potentials are to operate conjointly on every occasion? The reply to this question is an emphatic "NO". This is because we know not which phase difference dominates each time the die is rolled (or, as each photon is emitted in the more realistic double-slit experiment). Similar to the classical case, the nature of such knowledge is only retrospective, acquired only when the outcomes have been known. It is in this sense that quantum probabilistic potentials are chancy.

Chapter 6

The Wavefunction and Causal Processes

6.1 Introduction

In this chapter, I first argue (Sections 6.2-6.4) that the Schrödinger evolution of the wavefunction ψ (as a representation of mutually acting potentials) of a physical system may be regarded as a causal process (in the manner of **HCQT**), which transmits a conserved quantity $|\psi|^2$ according to the "*at-at*" criterion.

I then proceed (Section 6.5) to point out that this view is not entirely satisfactory because of the conceptual problems associated with both ψ and $|\psi|^2$. First, the need to satisfy the "*at-at*" criterion for causal transmission means that $|\psi|^2$ must be locally conserved, and this in turn implies that ψ is a complex quantity[1]. It is argued that by treating seriously the distinction between potentials and their effects, one is then able to dissolve this worry. Second, and more importantly, $|\psi|^2$ as a probability measure represents strictly an aggregate property relative to an ensemble of systems of the same kind. I maintain that one can take a single-case propensity interpretation of the potential ψ such that its effect $|\psi|^2$ can be viewed as ascribing a property to a single system, as in line with the singular aspect of process causation. Once again, this drives home the relevance of the distinction between potentials and their effects.

[1]We shall see in Section 6.3 that being a locally conserved quantity requires $|\psi|^2$ to satisfy a continuity equation to the effect that the sum of the two terms corresponding to the time rate of change of $|\psi|^2$ and a probability current density defined in terms of ψ is equal to zero. In order to ensure that the probability current density is not always equal to zero, ψ must be a complex mathematical quantity.

Third, as illustrated in the two-slit experiment, the way the probabilistic potentials act at the source (or at least before the slits) depends on whether there "*will be*" measurements at the slit. This, I contest, is more reasonably expressed by maintaining that the behaviours of the potentials depend on the entire situation from start to finish. This situational dependence of the behaviours of probabilistic potentials undermines the Schrödinger wavefunction as a causal process because of the wavefunction's differential nature that focuses on one event at any one time. A more suitable candidate for a causal process in quantum mechanics, as I argue (Section 6.6), is via an integral approach to quantum mechanics that takes into account this essential situational aspect of probabilistic potentials, as we shall see in Chapter 7.

6.2 Process causation and Quantum Physics

As I remarked in the interlude (Section 4.6), the harmonious link between *causation* and *continuity* has suffered a brutal severance in the face of quantum mechanics when the notion of a continuous spacetime path becomes untenable. This is nowhere expressed as vehemently as by Ernst Cassirer in a discussion on "Causality and Continuity"[2],

> ***If one reflects once more on the evolution of philosophical and physical thought, in which the concepts of causality and continuity are more and more welded together, it becomes understandable how difficult it must have been to rend asunder the bonds holding them together. For this nothing less than the mighty explosive of the quantum theory was necessary. And in the meantime the belief that causality and continuity are indissolubly bound together and interdependently had taken such deep roots that an abolition of this union was considered by the representation of the new view as an abolition of causality itself. The scission which separated continuity from causality***

[2]Cassirer, E. (1956), p.163.

was considered fatal.

Fatality, I suppose, applies to both continuity and causality. Quantum considerations point to every evidence that the dynamics of objects in the world is by its very nature a discrete one and discontinuity, having dethroned continuity, has assumed its rightful place as the ruling monarch of microphysical processes.

Continuity, inasmuch as it has been a beautiful concept adhered to, is now, in this cruel sense, dead and gone. Causality, which had all those times enjoyed a faithful partnership with it, has been naturally regarded by many to suffer the same terrible doom and has been driven to its demise.

The motivation of this chapter and the next stems from the somewhat grandiose hope of making sense of the concept of causality with respect to the quantum domain. While the attitude I take in the present work is optimistic, it is, however, not one without caution.

To assess the degree of devastation the teachings of quantum mechanics has brought to the well-entrenched traditional thoughts on physics and philosophy, we need examine, so to speak, the manner of the alleged death of continuity. We need to understand how exactly the formalism of quantum theory has introduced and encouraged the vision of discontinuity. We need to gain this understanding in order to be able to identify within the quantum formalism the degree to which the notion of continuity has failed. Only then are we entitled to make an informed judgement on whether causality may still be awarded a sensible meaning in the realm of quantum physics.

The process causal theories of Salmon and Dowe have contributed much towards formalising a precise relation between causation and physical processes within the classical setting. This is especially true of Salmon's latest ***Conserved Quantity Theory with Transmission*** (CQT) and my proposed ***History Conserved Quantity Theory with Transmission*** (HCQT), where one sees in a clear manner the crystallization of the idea of transmission of causal influences via the continuous paths of objects in spacetime. These theories thus provide a solid platform for investigating the departure from the traditional views of causality under the regime of quantum mechanics.

As preluded in the opening section of the last chapter, a notion of continuity in terms of a *spatiotemporally continuous course of some entity* does exist in quantum mechanics. And indeed, it plays a crucial role in the usual formulations of the theory.

In particular, I have in mind the formulation of quantum mechanics in terms of the evolution in spacetime of the complex mathematical entity, the so-called wavefunction, according to the Schrödinger equation, as we have discussed in Section 5.2.

We note that the appeal to the Schrödinger dynamics of the wavefunction of a physical system has already been suggested in the writings of several critics of Salmon's original 1984 process theory as the issue of spatiotemporal continuity was addressed. Amongst these, the critiques of Woodward[3] and Forge[4] are most telling. Woodward, especially, proposes the treatment of a determinate probability distribution[5] evolving continuously in spacetime as a causal process in line with Salmon's theory[6].

This idea of a wavefunction evolving in spacetime also finds its echoes in a recent collection of papers by Salmon on the subject[7],

(a) ***I do not believe that quantum indeterminacy poses any particular problems for a probabilistic theory of causality, or for the motion of continuous casual processes. This quantum indeterminacy is, in fact, the most compelling reason for insisting on the need for probabilistic causation.***

In discussing the example of atomic transitions, Salmon says[8],

(b) ***But in the absence of outside influence, of that sort, it simply "transmits" the probability distribution through a span of time.***

[3] Woodward, J. (1989).
[4] Forge, J.C. (1982).
[5] This is usually interpreted as the modulus square of the wavefunction.
[6] Woodward, J. (1989), p.376.
[7] Salmon, W.C. (1980b). Reprinted as a newly extended version with additional footnotes in Salmon (1998a). See (1998a), p.231, footnote n.19.
[8] ibid., p.227.

And on speaking about the system of the decaying atom[9],

(c) ***...the case is analyzed as a series of causal processes, succeeding one another in time, each of which transmits a definite probability distribution - a distribution that turns out to give the probabilities of certain types of interactions. Transmission of a determinate probability distribution is, I believe, the essential function of causal processes with respect to the theory of probabilistic causality.***

As he further speculates[10],

(d) ***Each transition event can be considered as an intersection of causal processes, and it is the set of probabilities of various outcomes of such interactions that constitute the transmitted distribution... While it is true that the photon, whose emission marks the transition from one process to another, does not exist as a separate entity prior to its emission, it does constitute a causal process from that time on. The intersection is like a fork where a road divides into two distinct branches - indeed, it qualifies, I believe, as a bona fide interactive fork. The atom with its emitted photons (and absorbed photons as well) exemplifies the interplay of causal processes and causal interactions upon which, it seems to me, a viable theory of probabilistic causality must be built.***

He is even more precise in his recent essay "*Indeterminacy, Indeterminism and Quantum Mechanics*"[11],

Schrödinger introduced a wave equation to characterize the evolution of a quantum mechanical system. With this equation it is possible to calculate the state of the system at a later

[9] ibid., p.227-228.
[10] ibid.
[11] Salmon, W.C. (1998b), p.270-271.

time, given its state at an earlier time. Schrödinger's equation establishes a deterministic relationship between the state of the system at one time and its state at a later time. Of course, this later state does not embody precise values for both position and momentum.

I think everyone would agree that the Schrödinger evolution of the wavefunction of a physical system, or the probability distribution it generates (as suggested by Salmon and Woodward) are natural choices to consider in the attempt to extend the notion of a spatiotemporally continuous process to the quantum realm.

6.3 The Wavefunction and Conserved Quantity

We note from several of Salmon's remarks (Section 6.2) his suggestion of a possible generalization of the process causal theory to the domain of quantum physics via the Schrödinger dynamics of the wavefunction ψ. In particular, he maintains that the *transmission of a definite probability distribution* is the essential function of quantum causal processes. By a "definite probability distribution", Salmon refers to, of course, the modulus square[12] of the wavefunction ψ, with its usual interpretation as the probability of finding a particle at a position x upon a measurement carried out at a time t.

A number of comments are in order. The entity that undergoes transmission - that evolves in spacetime (according to the Schrödinger equation) - is *not* the probability distribution $|\psi|^2$ but ψ[13] itself. And so we have this picture: a quantum process ψ transmitting a quantity $|\psi|^2$.

In order to qualify as a causal process à la **HCQT** (and **CQT**), the condition of the "*transmission of a conserved quantity*" must be fulfilled. So if indeed transmission

[12]Mathematically, it is the *modulus square* of the wavefunction $|\psi|^2$ and not the wavefunction ψ itself that represents the measure of probability.

[13]Following our arguments in the previous chapter, one conceives of some kind of *probabilistic potentials* (or *propensities*) that provide the physical system with the tendencies to manifest different values of a dynamical variable, whose relative frequencies in the distribution are the measures of the probabilities of the individual values being attained. ψ may now be associated with such probabilistic potentials.

is achieved by following a Schrödinger evolution and ψ is the quantity that evolves as so, then we are faced with the immediate task of showing: (1) $|\psi|^2$ (as a property of the system) is a conserved quantity and, (2) its dynamical evolution in spacetime does indeed satisfy the "*at-at*" criterion of causal transmission. To my knowledge, this treatment is lacking in Salmon's recent arguments and it is my hope to fill in the lacunae in this section.

We deal first with the issue of whether the quantity $|\psi|^2$ may indeed be considered a conserved quantity.

Given that the wavefunction ψ is to be interpreted as a probability amplitude[14] in relation to the location of a particle, the "motion" of ψ is closely related to the motion of the particle in spacetime. It is useful, therefore, to introduce the concept of a *probability density current* to describe the flow of probability - to describe how probability changes in space and time.

Consider an infinitesimal region of space with volume $dxdydz \equiv d\tau$ and we denote the probability density - *the probability per unit volume evaluated at a point* $\mathbf{r}$ *at time* t - to be $\rho(\mathbf{r}, t)$. The probability that a particle is located within this tiny region at time t is thus given by,

$$\rho(\mathbf{r}, t)d\tau \tag{6.1}$$

We must now impose the important condition that there is definitely one particle somewhere at time t,

$$\iiint \rho(\mathbf{r}, t)d\tau = 1 \tag{6.2}$$

The integral is then to be taken over all regions where the particle might be found.

[14]The term "probability amplitude" is used in quantum mechanics to denote a quantity whose modulus square gives the probability measure.

Since the probability density $\rho(\mathbf{r}, t)$ for finding the particle is proportional[15] to $|\psi(\mathbf{r}, t)|^2$, it follows therefore,

$$\iiint |\psi(\mathbf{r}, t)|^2 d\tau = 1 \tag{6.3}$$

provided that particles are neither added nor removed. In other words, equation (6.3) must hold true at all times and not just at $t = t_0$. Furthermore, the probability of finding the particle over all spacetime regions must be conserved. In particular, if the probability of finding the particle in the same bounded region of space *decreases* as time goes on, then the probability of finding it outside of this region must *increase* correspondingly by the same amount. This implies that we may indeed define a probability density current $\mathbf{j}$ and show that the quantities $\mathbf{j}$ and $\rho(\mathbf{j}, t)$ together satisfy a *continuity equation.* This can be shown by examining the time rate of change of $\rho(\mathbf{r}, t)(\equiv |\psi(\mathbf{r}, t)|^2 = \psi^\star \psi)$,

$$\frac{\partial \rho}{\partial t} = (\frac{\partial \psi^\star}{\partial t})\psi + \psi^\star(\frac{\partial \psi}{\partial t}) \tag{6.4}$$

Using the Schrödinger equation (5.11) in the form of,

$$\frac{\partial \psi}{\partial t} = \frac{1}{i\hbar}(-\frac{\hbar^2}{2m}\nabla^2\psi + V\psi) \tag{6.5}$$

Equation (6.5) is substituted into equation (6.4) to obtain,

$$\frac{\partial \rho}{\partial t} = -\frac{\hbar}{2im}[\psi^\star(\nabla^2\psi) - (\nabla^2\psi^\star)\psi] + \frac{\psi^\star(V\psi)}{i\hbar} - \frac{(V^\star\psi^\star)\psi}{i\hbar} \tag{6.6}$$

[15]We have taken $\rho(\mathbf{r}, t) = |\psi(\mathbf{r}, t)|^2$. Strictly speaking, however, $|\psi(\mathbf{r}, t)|^2$ is only proportional to $\rho(\mathbf{r}, t)$ if not for the condition we impose on $\psi(\mathbf{r}, t)$ that if there is one particle at an initial time t_0, then $\iiint |\psi(\mathbf{r}, t_0)|^2 d\tau = 1$. For those ψ's that do not satisfy this condition, they can nevertheless be normalized (achieved by the multiplication by an appropriate constant) in order to be kept consistent with this condition.

The last two terms cancel because the usual potential energy V is a real multiplying function. The remaining two terms can be rearranged to give,

$$\frac{\partial \rho}{\partial t} = -\nabla \cdot \frac{\hbar}{2im}[\psi^{\star}(\nabla\psi) - (\nabla\psi^{\star})\psi] \tag{6.7}$$

Thus, if the probability current density is *defined* as,

$$\mathbf{j} \equiv \frac{\hbar}{2im}[\psi^{\star}(\nabla\psi) - (\nabla\psi^{\star})\psi] \tag{6.8}$$

This would then result in the continuity equation,

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{j} = 0 \tag{6.9}$$

This equation represents the general form whenever something is described by a density ρ ("so much" per unit volume), a current density $\mathbf{j}$ ("so much" per unit time per unit area perpendicular to the flow) and conservation, in the sense that our something is neither created nor destroyed. Generally speaking, equation (6.9) applies not only to quantum probabilities, but also to a diverse variety of situations like the flows of electric charge, gas or even traffic.

Examing carefully the expression for $\mathbf{j}$ above (see equation (6.8)), we note that ψ is, by necessity, a *complex* quantity. For if it is a real function instead, then $\mathbf{j}$ would always be zero.

In the above, the presence of a continuity equation is utilized as an indication of the existence of a conservation law. There are, however, two kinds of conservation laws: *global* and *local* ones, the distinction of which is of relevance to our discussion. To illustrate the difference between these two kinds of conservation laws, we take for instance the simple case of electric charge. The statement that *the total charge in the universe is a constant (independent of time)* is an example of a global conservation law. It implies that any decrease of charge in one region of the universe is compensated by an increase of the same amount of charge in another region at the same time. In particular, if these two are vastly separated spacelike regions, then this would mean action-at-a-distance. Non-local processes are thus sanctioned by global conservation laws *per se.*

Contrary to common thought, it is worth noting that aside from our demand that the probability of locating the particle in a region of space decreases as that in an adjacent region increases, there is no *a priori* logical requirement for the probability of locating a particle in a region of spacetime to obey both global and local conservation laws. In particular, the probabilistic interpretation does not prohibit cases of global-and-not-local conservation. It makes perfect sense to ask if the probability of locating the particle in one region of spacetime is reduced, then does the probability of finding it in a region spacelike-separated from the first increase correspondingly? The complex character of ψ appears to have been forced upon us by the demand to preserve the local conservation of its effect, the probability density $|\psi|^2$ (see equation (6.8) above). A net transport is also essential for the condition of transmission to be fulfilled, granted that the notion of transmission requires its subject to vacate a previous occupied location in spacetime to take up a different one.

To eliminate non-local effects, one needs also *local* conservation laws and these are expressible in terms of *continuity equations* that bear the form of equation (6.9). This equation is often used in hydrodynamics for a fluid of density P and current $\mathbf{j}$ in a medium without *source* or *sink*[16]. To satisfy the condition of conservation locally, we take a small region of the fluid with a closed surface. The *locality* condition is then brought in by the consideration of the flux out of one volume (say 1) over this surface moving into an *adjacent and neighboring* volume (2) *over the same surface* (Figure 6-1).

The continuity equation is intended to capture this aspect of local conservation of a physical quantity: the flux coming out of one side of the surface immediately

[16]It is useful to introduce the concept "flux lines" or "streamlines" in such a context; these are directed lines (or curves) that indicate at each point the direction of a vector field. The flux of a vector field is analogous to the flow of an incompressible fluid such as water. Now for a volume with an enclosed surface found within the fluid, there will be an excess of outward or inward flow through the surface only when the volume contains, respectively, a *source* or a *sink.* That is, a net positive divergence indicates the presence of a source (a region where the flux lines originate) of fluid inside the volume, whereas a net negative divergence indicates the presence of a sink (where the flux lines terminate).

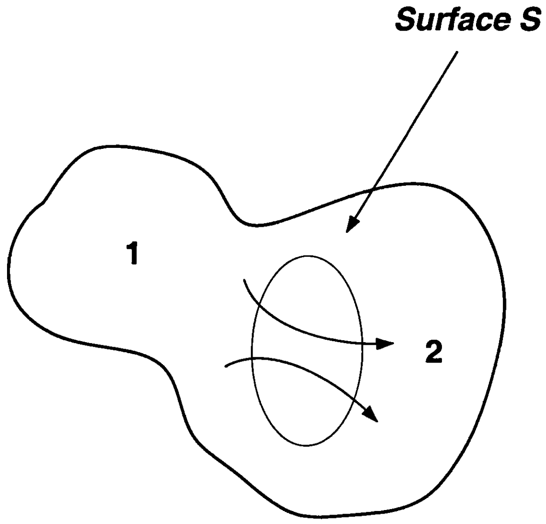

Figure 6.1: Flux lines "flowing" out of one volume (1) into a neighboring volume (2) over the same surface S.

passes through to the other side of the same surface. It is in this way that the spatiotemporal continuity in the motion of the flux lines is secured. Our interpretation of the probability of finding a particle somewhere in a region of spacetime as $|\psi|^2$ $(\psi^{\star}\psi)$ leads to the continuity equation (6.7) with the V-terms being conveniently cancelled out. This continuity equation describes the local conservation of probability density.

Incidentally, the total probability for finding the particle anywhere in the universe is also a globally conserved quantity. This can be easily shown as follows,

$$\langle\psi(t)|\psi(t)\rangle = \langle\psi(0)|U\dagger(t)U(t)|\psi(0)\rangle = constant \tag{6.10}$$

$$constant = \int\int\int \langle\psi(t)|x,y,z\rangle\langle x,y,z|\psi(t)\rangle dxdydz \tag{6.11}$$

$$= \int\int\int \langle\psi(t)|\mathbf{r}\rangle\langle\mathbf{r}|\psi(t)\rangle d^3\mathbf{r} \tag{6.12}$$

$$= \int\int\int \psi\star(\mathbf{r},t)\psi(\mathbf{r},t)d^3\mathbf{r} = \int\int\int P(\mathbf{r},t)d^3\mathbf{r} \tag{6.13}$$

We have thus arrived at the following situation. The quantity ψ evolves in space-time according to the Schrödinger equation but, as it does so, it transmits a conserved quantity $|\psi|^2$ that satisfies the condition of global as well as local conservation.

We need then to examine more closely the extent to which $|\psi|^2$ may be considered a *dynamical property* (like energy or momentum) possessed by the physical system, in the same manner as the acceleration (as the result of the action of an external force $\mathbf{F}$) on a corpuscular mass is considered a proper dynamical property of it. We say that a particle in motion *has* or *possesses* an acceleration that is considered as a dynamical property[17]. So if in place of an external force, we have the probabilistic potential ψ operating on the particle, are we then allowed to attribute its "effect" - the probability density $|\psi|^2$ of locating the particle upon a position measurement at a certain place - as a legitimate property possessed by the particle as similar to the case of acceleration?

Forces in classical physics are *defined* in terms of the effects (the accelerations) they produce on a point mass m through Newton's Second Law of motion, namely: $\mathbf{F} = m\mathbf{a}$. Notice that here the potential ($\mathbf{F}$) and its effect ($\mathbf{a}$) on a physical object are treated in such a way that the potential is described by its effect, which represents a dynamical property that the mass m has acquired due to the influence of (or via its interaction with) the potential. Moreover, a one-to-one correspondence exists between the potential and its effect so that the same potential always gives a unique value of acceleration *ceteris paribus*; even though the actual magnitude of the acceleration depends also of the value of the mass. In this circumstance, the potential is quantified in terms of its effect and no distinction is made between the two, at least on the level of mathematical treatment. The potential is what is postulated to be a reason for the particle to acquire the particular dynamical property in question. The crucial idea is that, in the presence of the manifestation of a potential, the *possession* of the dynamical property is characterized by the assignment of a unique numerical value

[17]The term property in physics refers to the values of some dynamical quantities lying in certain specified ranges. Classical physics makes the distinction between *internal* properties like mass or charge and *external* or *dynamical* properties such as position, velocity and acceleration whose numerical values specify how the object appears in relation to the framework of space and time. For a nice discussion, see Isham, C.J. (1995), p.56ff.

of this property to the physical system concerned. What is of importance is that the effect forms the contact point between the potential and the property acquired by the system due to the influence of this potential.

The requirement of possessing one unique numerical value for any dynamical property poses potential difficulties when one comes to think about probabilistic systems. Given similar (and what, for all intent and purposes, may often be viewed as identically repeated) dynamical conditions such as the application of the same force in casting the same (unbiased) die for a great number of times, only one out of all six possible outcomes obtains in any one cast. It would seem, *prima facie*, in such circumstance that the *same* potential is capable of leading to six different values of the same dynamical property that corresponds to the orientation in space of the die. The one-to-one correspondence between the applied force and the orientation of the die seems to be absent in such a case. It is, however, possible to restore this one-to-one correspondence by resorting to the consideration of probabilistic potentials as we have done in our previous chapter[18].

As we saw in the last chapter, working with six probabilistic potentials and six outcomes, the one-to-one correspondence between each potential and its associated outcome may now be restored, albeit in a retrospective manner. There is one and only one outcome for each trial and hence only one result is assigned to it. It is in this way that the probabilistic potentials (ϕ_1,...,ϕ_6) which is dependent on both the external applied force and the mass (i.e. an intrinsic property) of the die finds its manifestation in the dynamical property of the system, which in this instance is the spatial orientation of the die.

This granted, it would seem to follow that by the consideration of probabilistic potentials as the tendencies to produce mutually exclusive outcomes in a given situation, we can attribute genuine properties to the system. Every spatial orientation

[18]Instead of a material potential, our applied force, we speak about the six probabilistic potentials with each corresponding to the tendency of one face of the die showing in a trial. It is understandable that one might feel an unease towards this rather untangible and elusive concept of probabilistic potentials. One way of bringing one's thought more at ease with the concept is to regard each one of these probabilistic potentials as taking account of the slightly varied initial conditions of the trials whose difference is practically immaterial. This way of viewing the probabilistic potentials gives more of a physical content to these potentials.

that represents one face of the die is a genuine dynamical property (similar to that of position) of the die in much the same way as the different accelerations of the same mass produced by the different magnitudes of the applied force. Here, we have the different probabilistic potentials accounting for the differences in the spatial orientations of the die.

In the case of mechanical potentials, the *effects* get translated into the property on which it is exerting influence; these are the properties of the system that gets conserved. For instance, when acted upon by an applied force $\mathbf{F}$, a mass m accelerates, this acceleration manifests as an increase in the conserved quantities $(E, \mathbf{p})$ m possesses. It is already clear from this simple example that the potential is in a sense, something *external* to the system. It is tied closely to the variables $\mathbf{x}$, $\mathbf{p}$ that describe the state of the system but nonetheless, it is *not* the state of the system.

In quantum mechanics, this level of clarity seems to have been obscured by probabilistic considerations. I would, however, argue that the basic structure is very similar and this hinges upon our interpretation of ψ as a probability-generating potential.

Being granted the same level of ontological reality as mechanical potentials, ψ, as a probability-generating potential is bound up intimately with the "particle" it is supposed to influence. Again, analogously to the mechanical sense, it is something that is, in a very loose manner of speaking, "external" to the system. Since $|\psi|^2$ makes the eventual contact with the dynamical properties of the system, it appears therefore, by analogy with the case of forces and accelerations, reasonable to regard $|\psi|^2$ as a property of particles like photons[19].

6.4 The Wavefunction as a Causal Process

I shall now argue that ψ qualifies as a causal process à la **HCQT** on the basis that, (i) the quantum process ψ constitutes a history and, (ii) it transmits a conserved quantity according to the "*at-at*" criterion.

[19]Here, the probability distribution is revealed by the repeated measurements of an ensemble of photons rather than only one of them; and so strictly speaking, the distribution is a property of the *whole* ensemble rather than of a single system.

ψ evolves continuously in spacetime according to the Schrödinger equation. One may consider this evolution as describing a continuous history in the sense that at each spacetime point, the modulus square of ψ gives the probability of locating the particle at that point upon measurement.

Having decided that $|\psi|^2$ may be regarded as a property of a physical system and also that it is a conserved quantity, we now come to the second question of whether the "*at-at*" criterion of causal transmission is satisfied. The essence of the "*at-at*" criterion consists in the fact that the subject of transmission, in this case the conserved quantity $|\psi|^2$, is to be found at every spacetime point between two specified destinations (without intersections that involve exchange of a conserved quantity).

Indeed, the continuity equation[20] (6.7) has furnished us with the information of $|\psi|^2$ at every spacetime point. Thus, the spatiotemporal evolution of $|\psi|^2$ fulfills the "*at-at*" criterion of causal transmission.

However, transmission now takes the form of the possession of a probability distribution for a particle to be found at a location, defined at every spacetime point along each possible path between two specified destinations (the source and the detection screen in the two-slit experiment for example). Formally, the possession of this conserved quantity should be achieved in the absence of intersections with other processes, which involve exchange of the same quantity. In regard to probability distributions, what might one refer to as an intersection and what would indeed constitute an exchange?

Here, one is unable to speak about the trajectories of objects intersecting geometrically, as there is no longer any meaning of a definite trajectory for the object. Instead, one now thinks of a certain probability associated with two causal processes to intersect. This probability manifests itself in changes in the probability distributions (e.g the weights of the different basis states) that may naturally be regarded as an indication that an exchange has taken place.

This granted, one must, however, be specific about the sort of intersections that

[20]This being a first-order differential equation serves the same purpose for $|\psi|^2$ as the Schrödinger equation does for ψ.

induce changes in probability distributions. This is a relevant consideration because there exists in quantum physics a very special class of intersections that changes the status of a physical system so drastically from being in a state of having no definite dynamical property (superposition) to one that does (eigenstate). These are intersections which constitute measurements on the system.

Because of its significant dominance in quantum phenomena, there is every reason to treat intersections that induce probability changes as falling into two categories: those that preserve the basic feature of superposition of probabilistic potentials and those that changes these superposing behaviours of the probabilistic potentials (measurements). It is therefore appropriate to stipulate that the latter kind of intersections giving rise to changes in probability must be absent in a causal process in quantum physics.

6.5 The Problematic Wavefunction

The spatiotemporally continuous path followed by a particle is the paradigm of a causal process in classical physics. It is in virtue of the motion of the particle that conserved quantities like energy and momentum are carried along from one spacetime locale to another. These quantities are the causal influences that produce dynamical changes when the particle comes into interaction with another. Central to this picture of causation is the idea that there is the instantiation of a unique definite value of a conserved quantity at every spacetime point traversed by the particle during its motion. The definite instantiations of these conserved quantities are regarded as properties of the particle and these properties form part of the description of its state.

In the transition to quantum mechanics, the particle "loses its path": one is no longer sanctioned to speak of the particle as traversing each and every point in-between two specified destinations in spacetime. The dire consequence that follows concerns, of course, the manner one is to think about the transmission of these conserved quantities in the absence of a carrier travelling continuously in spacetime. This is where the notion of transmission, as essentially the causal connection in physical processes, reaches a blind alley. Without this important causal ingredient, one would feel compelled to expel the concept of causation from the quantum language.

Almost every effort to rescue what remains of causation from the quantum onslaught focuses on physical quantities that preserve the feature of spatiotemporal continuity. But this by no means a straightforward task. For the start, because of the need to satisfy local conservation as demanded by the condition of transmission, ψ is, by necessity, a complex quantity. But this should not raise such an alarm as it often has done. For, strictly speaking, there is no restriction on a potential to be a mathematically real quantity, provided the effect it produces can be measured as a real quantity[21]. However, even accepting the fact that ψ is a complex quantity does not cause us immediate concern if the distinction between potentials and their effects is taken seriously as we have done, there are problems associated with $|\psi|^2$ (although as a real quantity) as a property of the system.

We have already discussed (Section 6.3) the sense in which $|\psi|^2$ can be considered as a conserved quantity transmitted by the wavefunction ψ as it evolves continuously in spacetime. In place of the instantiation of a unique definite value of a dynamical quantity, one speaks of the instantiation of a definite probability distribution of dynamical values (upon a measurement) at every spacetime point. If there exists not a definite value of a previously recognized dynamical value at a particular spacetime point, in what right then are these distributions to be considered as a property of the system? To answer this question, as it turns out, we need to think more carefully about ψ itself. Strictly speaking, $|\psi|^2$ gives only the *probability* of locating a particle at a position x upon a *measurement* performed there. The reference to "probability" and "measurement" poses several difficulties. I will first address those in relation to probability.

There is one chief problem, much discussed in the voluminous literature on the subject, of $|\psi|^2$ being interpreted as a probability distribution and its being regarded as a property of a system. This is to do with the fact that probability measures are

[21] One measures a classical mechanical force by the acceleration it produces on a physical system. The acceleration takes on a value given by a real number, but no corresponding restriction applies to the force itself.

defined in terms of a long sequence of repeated trials or equivalently, of an ensemble of identical trials. A probability distribution then, at best, represents some kind of *"aggregate" property* that manifests and reveals itself only in the the presence of a long sequence of repeated trials of an experiment or over identical experiments conducted on an ensemble of the same kinds of systems simultaneously. It is indeed with the very nature of such "aggregate" behaviours that statistical techniques deal. This is so whether one is dealing with probabilities in classical or quantum physics. In the classical experiment of the tossing of an unbiased coin, a distribution of "heads" and "tails" obtains only after a large number of repeated throws of the same coin (or throwing a large number of identical coins at the same time). In the two-slit experiment, the statistical distribution of the positions of the photons (the interference pattern) takes shape only after *many* photons have been deposited on the detection screen. In other words, the distribution is obtained not by the observation of one single photon, but a whole ensemble of them, as a result of the repeated firing of photons at the slits.

Generally speaking, then, it seems inappropriate to regard $|\psi|^2$ as a property of any single individual member of the ensemble. It is apt to trace the roots of this difficulty to the classical frequency view of probability (as advocated by Von Mises[22] and codified in an axiomatic form by Kolmogorov[23]) to which the interpretation of $|\psi|^2$ adheres. According to these theories, probabilities are deemed meaningful insofar as long sequences of either repeated trials or repeatable conditions are concerned. And so the frequency interpretation of probability, which hinges upon the notion of repeatability, is often thought to render it meaningless to speak about the probability of a single trial.

For the frequentist there is no meaning in the pronouncement of a statement like "the probability of a single throw of an unbiased coin showing a head", because it is strictly, with respect to the condition of *repeatability*, incorrect to give probability assignment to a single trial. Rather, we are primarily concerned with the situation

[22]Von Mises, R. (1963).
[23]Kolmogorov, A.N. (1950).

where say, m heads obtain over a long sequence of n trials of the throwing of the same coin. The probability of obtaining a head *with respect to this sequence of experiments* is $\frac{m}{n}$.

The stringent requirement of repeatability is then taken to forbid the assignment of probabilities to single trials. From the perspective of the two-slit experiment, $|\psi|^2$ signifies the probability of finding a particle at a certain location with respect to a long sequence of repeated deliveries of photons on the screen through the two slits. It does not, strictly, provide the probability of locating the particle in any one particular trial.

It has been a long-standing challenge for workers in the foundations of probability to try and make sense of singular probability statements. It remains a task no less significant for physics to tackle the analogous issue of the probability of a single event. This is of particular relevance in relation to the process causal theories of Salmon and Dowe that aim to provide a framework for singular causation; namely, to be able to speak about causation with respect to the ascription of properties to a single physical system such as an electron or a photon.

Of the various treatments that attempt to provide a meaning for singular probabilities (popular mostly amongst the subjectivist types of interpretations of probability), I find Popper's *propensity theory* attractive since it was put forward with a view of probabilities as an objective feature of the world by providing an underlying substratum for relative frequencies. All the more importantly, it is consistent with my interpretation of ψ as a probabilistic potential. Let us now very briefly review the propensity theory.

Propensity, for Popper, represents a tendency to realize *an event* and such a tendency is, according to him[24], "*inherent in every single throw*[25]." And the connection of propensities and relative frequencies consists in the fact that[26], "*we estimate the measure of this tendency or propensity by appealing to the relative frequency of the*

[24]Popper, K.R. (1990a), p.11.

[25]In his discussion of the example of the probability for the event "two turning up" in the cast of a fair die.

[26]ibid.

actual realization in a large number of throws; in other words by finding out how often the event in question actually occurs."

And again[27], *"So, instead of speaking of the possibility of an event occurring, we might speak more precisely, of an inherent propensity to produce, upon repetition, a certain statistical average."*

With the notion of propensity, Popper introduced an objective theory of probability based on two-tiers: the *potentials for the system* to realize certain results *and* the *relative frequencies* of the occurrences of the results that measure these potentials. In contrast, the traditional frequentist takes the "single-tiered" view: a sufficiently long sequence of trials giving rise to a distribution of a set of mutually exclusive outcomes and probabilities are equated to the limiting relative frequency of occurrences of these outcomes. They make no distinction between the series of outcomes and the manners in which they have been generated. The lack of consideration of this important connection is the main inhibition of the frequentist view to the provision of a meaning to singular probability.

By appealing to propensities, Popper[28] puts probability-generating potentials on a par with the more familiar types of potentials like physical forces. And even more compellingly[29], *"The introduction of propensities amounts to generalizing and extending the idea of forces again."*

After all, one deals with physical situations where a die is thrown or a coin is tossed and inevitably, the outcomes are dependent on the particular physical situation under examination. A biased coin produces more instances of "heads", for example, than a normal unbiased one. A die landing on a slotted surface would have a chance of having one of its edges rather than a face of it pointing upwards. In each of these examples, the effects - the "probabilities" - are parasitic upon the physical situation as a whole: the system and its environment. This is indeed an important observation by Popper[30], *"I had stressed that propensities should not be regarded as* <u>*properties*</u>

[27]ibid.
[28]ibid., p.12.
[29]ibid., p.14.
[30]ibid.

inherent in an object, such as a die or a penny, but that they should be regarded as inherent in a situation (of which, of course, the object was a part). I asserted that the situational aspect of the propensity theory was important, and decisively important for a realistic interpretation of quantum theory."

Bearing a marked resonance with our already much discussed probabilistic potentials, propensities give physical reality to each single trial: *a potential or tendency to produce a certain result.* But of course, it is worth emphasizing that for Popper's propensities as well as for our probability-generating potentials, the measures of their "strengths" are by means of relative frequencies of the outcomes of an experiment in a *retrospective* manner.

As Popper has observed, propensities (as well as our probability-generating potentials) are situation dependent and therefore should not be regarded as *just* the properties inherent in the objects themselves. We have already argued, in addition, it is the effects of the probability-generating potentials (or propensities) that find their manifestations in the dynamical properties of the physical objects that come under their influence. When an external force is applied to an object, the object acquires an acceleration (depending also, of course, on the mass of the object) that is to be treated as a legitimate property of the object. There exists a one-to-one correspondence between the potential and the manifestation of its effect. Such a correspondence is, as we have maintained, crucial to the assignments of properties to a system since the manifestations of the effects of the potentials usually take the form of one single definite value of a dynamical variable. While the *effects* of the probability-generating potentials (or propensities) find their manifestations as properties in the object upon which they act, the probability-generating potentials (or propensities) themselves do not.

Likewise, in the scenario of die-rolling, the introduction of probability-generating potentials (or propensities) re-establishes this one-to-one correspondence, facilitating the assignment of a property to the system in terms of the manifestation of the effect of a particular probabilistic potential; even though the knowledge of this can only be inferred in a retrospective fashion.

The outcome of each trial depends not only on the inherent symmetry of the die, for example, its being six-sided, but also on the conditions under which the throw is performed, such as in the presence of a strong gust of wind and a slotted surface for landing; these conditions invariably find their way into being the dynamical properties of the die, which in turn determine its twists and turns in space that lead to its eventual orientation.

So, after a long sequence of trials, one would then be able to make a claim that the potential ϕ_1 has *actually* operated on the die, on every occasion that "face 1" of the die shows. On each of these occasions, the die acquires this property - its orientation with "face 1" showing - through the action of ϕ_1 alone and the one-to-one correspondence between the potential and its effect thus ensures the assignment of this property to the die.

Hence, the consideration of propensities or probability-generating potentials renders the assignment of a property (in the sense of one unique value of a dynamical variable) possible in every single trial (in each single throw of the die for example). In every trial, one and only one out of the six potentials has operated and this *is* the objective fact the knowledge of which is, however, not available to us *a priori*. It is also in this sense that the measures of these potentials - the probabilities in terms of relative frequencies - are considered objective probabilities. There is a physical process behind every trial that consists of a probability-generating potential and its respective result. The propensity theory therefore provides a basis for and serves well the purpose of considering singular probabilities.

This works in part because of the fact that in a classical probabilistic situation, once one of the "competing" potentials has become actualized (or crystallized), the others would be regarded as not having acted. Put differently, the propensities facing a physical system forms a mutually exclusive set.

Given the intimate relationship between the propensities and the dynamical properties of a physical system, one is rightly to regard the physical "state" of a system as encoding not only information about the internal properties such as its mass but also that of the effect (like acceleration in the case of mechanical potentials) of the potentials acting upon it.

In this sense, one regards the state of a system as a catalogue of its dynamical properties, both internal and external. Moreover, it only appears to be appropriate to see the state as a property of the system - a summary of all its internal and external attributes. But, of course, the assignments of the external attributes are possible given the existence of a one-to-one correspondence between the acting potential and its effect that manifests indeed as a dynamical property of the system.

In a similar vein, the wavefunction ψ of a physical system encodes information about both the internal and external properties of the system. Take for instance, the distribution of photons on the detection screen changes as a large multiple of slits are introduced. The change in the number of slits alters the possible potential routes of each photon, the consequence of which has found its manifestation in a different statistical distribution.

The decision of whether one may also, as in the classical case, regard ψ as appropriately a property of the system (in the loose sense of being the tendency to produce a probability distribution $|\psi|^2$) depends largely on the existence of a one-to-one correspondence between the propensity and its effect. It transpires that if the photons are measured and tracked down in order to find out from which slit they have each emerged, this one-to-one correspondence exists. However, when such observations are performed, we recover the "classical case" of a "definite trajectory" for each photon[31], but at the expense of the vanishing of the interference pattern obtained when no such observation is carried out.

The dependence of the one-to-one correspondence on *acts of observation* of the photons has cast serious doubts on whether $|\psi|^2$ ought indeed be regarded as objectively a real property of the system. We shall now address the issue of *measurement* in relation to the interpretation of $|\psi|^2$.

Granted that our classical intuition tells us for any quantity associated with a system to be considered a legitimate property of it, the quantity must be possessed by

[31]Each has given rise by a separate probabilistic potential that forms the member of a mutually exclusive set.

it. Possession is to mean that a definite value of that attribute can be assigned at each spacetime point (although the actual value may vary from one point to another) along the trajectory of the system *regardless of whether a measurement is to be performed on it or not.* The possession of a property in the classical sense forms an objective element of reality. In quantum mechanics, the possession of a classical dynamical property by a physical system is complicated by the fact that it can only be said to possess such a property, in the sense that there is one definite value of that quantity at every point of its spacetime path, *only when it is measured.* The possession of a dynamical property no longer has an objective reality that is divorced from the very acts of measurement.

The best way to understand how our acts of observing a physical system change the "state" of it (speaking in a figurative manner) is to turn again to the two-slit experiment, Figure 6-2.

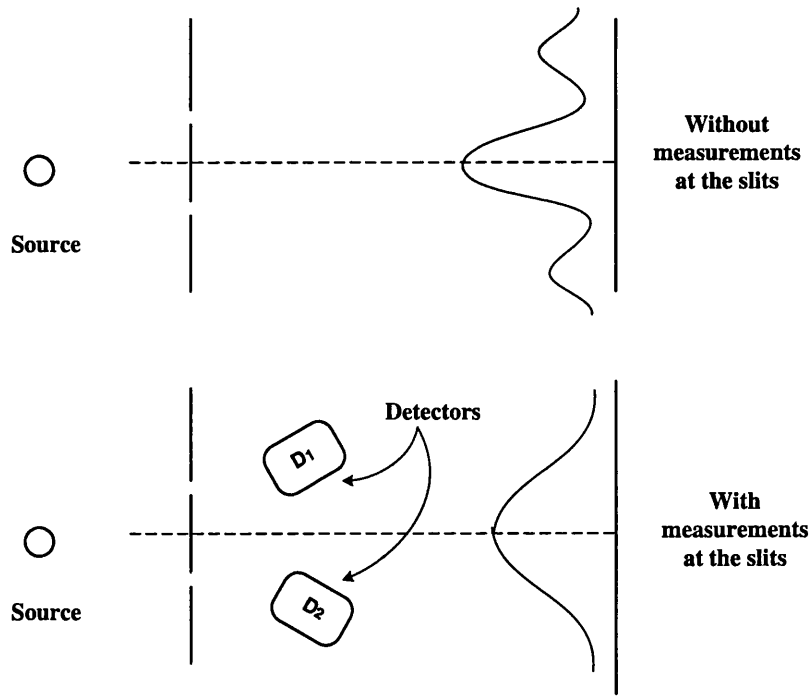

Figure 6.2: The changes in the final observed photon distribution as measurements of the positions of the photons are carried out at the slits.

We have seen how strikingly different the behaviours of photons can be under the two scenarios of "*with measurement*" and "*without measurement*" at the slits and the

With Measurement	*Without Measurement*
Distribution of photons shows *no* interference pattern.	*Distribution of photons shows an interference pattern.*
Indicates that the *sum-rule* of probability is obeyed and the outcomes form a mutually exclusive set.	*Indicates that the sum-rule of probability is violated and the outcomes no longer form a mutually exclusive set but one that is exhaustive.*
A mutually exclusive set implies that there exists a one-one correspondence between each operating potential and its effect, which in turn means that the potentials also form a mutually exclusive set.	*An exhaustive set makes possible the simultaneous actions of all the potentials.*
Definite trajectories for the photons result.	*No definite trajectory for the photons.*

Table 6.1: Differences in the two scenarios of "with measurements" and "without measurements" at the slits in the double-slit experiment.

results are summarized and laid out in Table 6-1.

Without any measurement of the positions of photons at the slits, one obtains a distribution of photons on the detection screen showing the distinctive characteristic of interference and under such circumstance, it is concluded that *the probability-generating potentials* form not a mutually exclusive but a set whose members are in operation simultaneously. In this case, there cannot be an assignment of a definite trajectory for a photon.

On the contrary, if there is an attempt of observation of the position of the photons by devices situated close to each of the two slits, that would result in a distribution of photons that obtains, which agrees with the supposition that each photon has followed a definite trajectory all along right from the very moment it was emitted at the source. In addition, this would mean that the probability-generating potentials do not operate together.

The logic central to this argument is as follows: the two modes of behaviours of the probabilistic potentials - whether they act *independently but simultaneously* ***or*** *act in* *a mutually exclusive fashion* - depends on two *different* situations. These correspond respectively to the presence and absence of detection devices set at the slits to observe the passage of photons. It may now be argued that a "future" event (with respect to the time of emission of a photon by the source) can take an active part in shaping the actions of the probabilistic potentials that are to govern the "fate" of the photon. In particular, with observations conducted at the slits, the distribution of photons that has the characteristic of a set of mutually exclusive potentials describes a reality *as if* the photons had all along followed a definite prescribed trajectory once it is emitted from the source.

In contrast, when no attempt whatsoever is made to observe the passage of photons, the "interfering" distribution of photons, which reveals the characteristics of an exhaustive set of probabilistic potentials acting in concert, leads to the description of an alternative reality *as if* there were no definite trajectory, for the photon right from their emission from the source.

6.6 Towards Feynman's Paths

An important observation emerges. Because of the fact that this seeming dependence on the entire physical situation inevitably encompasses events that occur at *different times* (for instance, the observation of the passage of photons as a "later" event in relation to the "earlier" emission event at the source) for quantum phenomena, our formulation must take into account the overall influence of events at *more than one time* on the behaviours of the probabilistic potentials; rather than one that pays attention to the separate influence of an event at each time on those potentials.

In the language of mathematics, this is tantamount to saying that an integral formulation of quantum mechanics is more revealing than a differential one. This is by no means trivial for it gives us an ontology in which it is the entire physical situation consisting of a system and its environment that shapes the behaviours of the probabilistic potentials (with the probabilistic potentials ψ representing not only the system but also its environment). One thinks no longer of ψ as something that attaches solely to the photon and that its mere function is to represent the state and evolution of the photon, quite independent of everything else.

This view of ψ releases one's thought from the conceptual difficulties associated with the conventional view that ψ is describing only the system, because the wavefunction ψ is now conceived as something that does not merely describe the system alone but rather, is dependent on the system and its environment - the situation as a whole. ψ does, however, find its manifestations in the dynamical properties of the system through its effect $|\psi|^2$.

We may now extract the following essence concerning the nature of ψ. (1) ψ *represents the set of probability-generating potentials that acts upon a physical system like an electron or a photon*, (2) ψ *is essentially periodic* and, (3) ψ *depends on the entire situation* (both on the system and its environment).

Treating ψ as a propensity has resolved for us the problem associated with the ensemble interpretation. One may now speak about ψ, the superposition of probability-generating potentials, as operating on *each individual* system. It has already been pointed out that when the situational aspect is put into focus, ψ is most perspicuously represented in an integral rather than a differential formulation. Moreover, also because of the dependence on the system's environment, ψ has to be relieved from being conceived as a wavepacket that is solely tied to the system itself.

To accommodate this all-encompassing nature of ψ, we need a formulation which "tunes-down" the emphasis on "*wavepackets* as *particles*" - the direct identification of the wavepacket (as mathematically represented by the wavefunction ψ) as the particle itself.

Tying too tightly the motion of a particle with the continuous advancement in spacetime of the wavefunction has also the unfortunate consequence of leading to picturing a shift in the behaviours of the potentials (which is my version of what is generally called "wavefunction collapse"). Imagine the scenario of low photon density in the two-slit experiment (i.e. one photon passes through the apparatus at one time). One first observes the distribution of photons on the detection screen that shows interference characteristics. By placing devices at the slits to track the photons, the distribution *changes* to one in which the interference disappears. If the wavefunction describes the motion of a photon as some form of a localized manifestation of a "physical wave" (in the usual sense of waves), one is compelled to describe the "change" in the behaviours of the probabilistic potentials as corresponding to the change as experienced in the physical situation (the introduction of the devices).

The distribution with interference indicates that the photon is subjected to a superposition of probabilistic potentials and hence nothing may be said of it having a well-defined trajectory *before* it reaches the detectors located at the slits. As one of these two detectors is triggered (indicating the registration of a photon), one of the two superposing potentials would have then become realized while the other would have failed to act, to account for an observed distribution that shows no interference! And it is therefore, to be concluded that once the position measurement is carried out to locate the photon at the slit, a *change* from the scenario where those propensities superpose to another *after observation at the slits* where one and only one of the propensities operates. Schematically, all these may be neatly summarized by the "time-line" in Figure 6-3.

One is then led naturally to puzzle over how the act of observation *selects* or "*single-out*" one particular potential and "dis-associates" its action from that of the rest of the superposition. This is the familiar territory of the topic of wavefunction collapse. Measurement is thus granted a very special status - this is the very event that is held primarily responsible for the change, which takes place between the two modes of behaviours of the probabilistic potentials.

The problem is encountered when the wavefunction is viewed as a representation of the state of the photon and no more than that. This view of a wavefunction immediately points to the usual differential approach, which describes the continuous spatiotemporal motions of particles in the classical context. However, a better perspective can be gained from an integral approach that incorporates a sequence of events together. One may wonder at this point that given the integral approach serves, in most classical cases, as an alternative view of describing the dynamics of the system to the differential view and that both views are deemed equivalent, does it not make our argument for the integral view mere verbal gibberish?

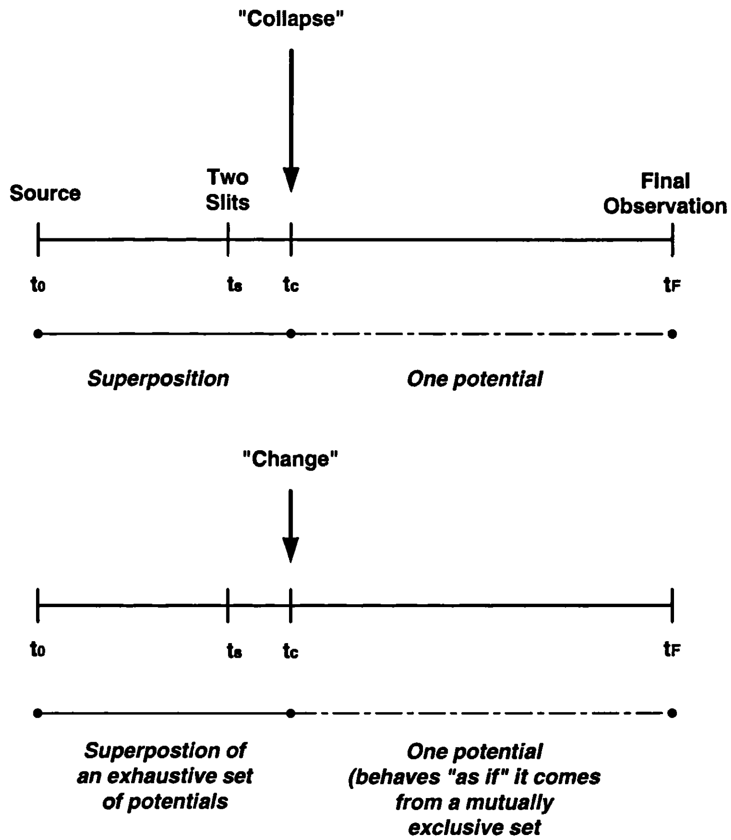

Figure 6.3: The chronology of "wavefunction collapse" in the two-slit experiment.

I would argue that it does not. In the context of quantum mechanics and in light of the view that the wavefunction is situation dependent, the integral approach provides far more insights into what is happening. In particular, there is this fallacy: the hypothesis that the probabilistic potentials are in a state of superposition *before* it reaches the detector is a testable and hence, in Popper's words, a falsifiable one. Why is that so? One usually supposes that without the detectors at the slits, the photons are propelled by potentials that organize them into an interference pattern and *as soon as* detectors are activated at the slits, the interference effect is no longer present. It is then deemed reasonable to conclude, as has often been argued, that before the photons reach the detectors, they *would have been* propelled by potentials in a superposition. But the crux of the matter remains: only those photons travelling through the apparatus *before* the activation of devices at the slits are found to exhibit interference, which evidences the fact that those photons are subjected to the action of a set of superposing probabilistic potentials. However, once the devices are activated, they influence only those later photons that reaches the detectors from that very moment on and these photons behave "*as if*" they are emitted into definite trajectories leading from the source to one of the slits.

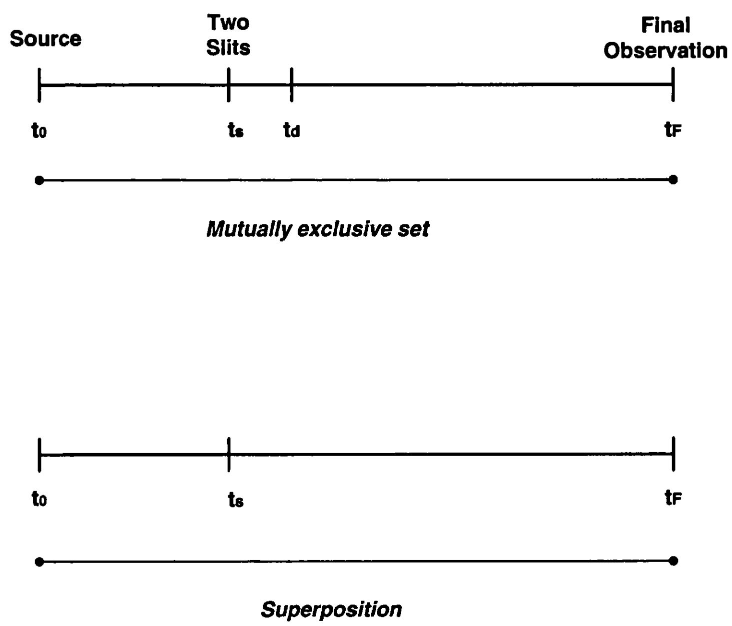

Figure 6.4: No "wavefunction collapse"!

It is therefore to be concluded that the influence of the observations of photons at the slits not only affects affairs at that spacetime point but also has far-reaching influence on the state of play of an extended spacetime region stretching from where the source is situated to the location of the slits. It changes not only the behaviours of the propensities at one point but also that in an extended spacetime region. And this *is* the crucial point.

This again suggests that integral approaches, based on the focus of a sequence of events as a whole rather than on attention to each event within the sequence at a time, seem to be the most appropriate tool for studying quantum phenomena.

To be sure, it is important to note it is not my claim that the appeal to integral approaches would help to explain away or even resolve the problems of measurement and wavefunction collapse. Needless to say, a proper treatment of these topics is far beyond the scope of the present work. Rather, I wish to point out that integral approaches do help to focus on the situational aspects of quantum phenomena; it highlights the amazing effects the mere observations of position (in terms of localization measurements) have on the behaviours of the probabilistic potentials in an extended spacetime region and not just that at a specific spacetime point.

Our strategy is now clear. We need to free ourselves from the burden of a too intimate bond between the wavefunction ψ and the particle that ψ *is supposed to describe*. A more general view of the quantity ψ should be adopted to take full recognition of the situational aspect in the description of quantum phenomena. Indeed, it is even advisable to abandon altogether the use of the term "wavefunction" that is understood as governing the evolution of a particle in spacetime in a differential manner. We should, however, retain and work with its most vital essence: the *probability-generating potentials.*

In place of the wavefunction ψ, an alternative to a causal process is sought with an emphasis on the following features: (1) the periodic nature of the probability-generating potentials, (2) the accommodation of a set of exhaustive alternatives in addition to a mutually exclusive set and, (3) the situation dependent aspect in quantum phenomena and, (4) the ability to speak about the probability of a single system.

Fortunately, such an approach does exist and is widely used today as an important tool in theoretical physics. It is the *Path Integral Formulation of Quantum Mechanics*, as put forward by Richard Feynman in his 1942 doctoral dissertation. In the next chapter, we shall first give an exposition of the formulation and then conclude with a discussion of how this approach may bear on the notion of a causal process in the quantum-mechanical setting.

6.7 Postscript: Bohmian Quantum Mechanics

However, I should conclude this chapter with a brief postscript on how, as some may think, that David Bohm's version of quantum mechanics can fit in with the process causality view. I also explain why I have chosen not to enter into an extensive discussion of the Bohmian formulation.

Bohm recasts the Schrödinger formulation of quantum mechanics in a form[32] that permits one to consider the motion of a particle being subjected to two kinds of potentials[33]: the usual classical potential energy V and the so-called quantum potential U. The quantum potential produces highly non-classical effects and is thus responsible for the quantum behaviour of the particle in question (interference effects in the double-slit arrangement for example).

We are therefore presented with a picture in which a "guiding" quantum potential U that acts on the particle and propels its motion through spacetime. In particular, Bohm's formulation provides the virtue that once the initial position x_0 has been specified, the trajectory $x(t)$ is uniquely determined. In regard to the ontological picture Bohm offers, Cushing comments[34],

All of the mathematical details aside, what Bohm did was to

[32]Bohm begins with non-relativistic Schrödinger equation, which is taken as given and not derived, and he defines the wavefunction ψ in terms of two *real* functions R and S such that $\psi = R\exp(\frac{iS}{\hbar})$. Taking $|\psi|^2$ as P, he was able to arrive at a continuity equation in the form of equation 6.9, which then allows the interpretation of P $(= |\psi|^2)$ as the probability density for the distribution of particles. See Cushing, J.T. (1994), Appendix 1.1, p.61.

[33]The term potential used here in the context of Bohmian quantum mechanics should not be confused with the probability-generating potentials that have been introduced earlier.

[34]Cushing, J.T. (1994), p.43.

take the Schrödinger equation, which has the form of a wave equation and hence naturally invites a wave (or perhaps, a wave-particle) interpretation, and re-expressed it in a form similar to Newton's second law of motion, which naturally invites a particle interpretation in terms of trajectories.

It seems, *prima facie*, that particles having definite trajectories within the Bohmian framework provide just the right paradigm for causal processes in quantum mechanics. For, one expects a well-defined particle trajectory would somehow render the notion of *at-at* transmission of causal influences meaningful on the quantum level. Bohmian trajectories thus look the right kind of analogue we seek for a causal process (à la **HCQT**) in the quantum realm.

Despite the nice and simple ontology that Bohmian quantum mechanics offers, there are, however, features of the formulation that discredit it from being the most appropriate candidate for a quantum causal process. We shall here consider two that have caused worries for the Schrödinger wavefunction ψ to become quantum causal processes.

First, I address the issue of probability. The interpretation of $|\psi|^2$ intended by Bohm as the probability density of particles is not meant to be inherent in the conceptual structure, but merely as a result of our ignorance of the precise initial conditions of the particles. This is similar to the case of random walks followed by particles in Brownian motion. However, when it comes to the consideration of probabilities, one still deals with an ensemble of particles. For Bohm, the quantum potential is not a superposition of probability-generating potentials, and the probabilistic interpretation of $|\psi|^2$ remains a classical one in that the probability-generating potentials forms a mutually exclusive set (as particles are presumed to have well-defined trajectories). It fails to capture that quantum mechanics is an *inherent* probabilistic theory and the important essence of the quantum behaviours of probability-generating potentials, namely, their simultaneous actions on the physical system, which I maintain is of fundamental significance.

Second, with its differential nature, the formulation is not well placed for the consideration of events that occur in extended spacetime regions.

As a consequence of these issues, even though the Bohmian formulation provides a coherent transition from the classical to the quantum motion of a particle, it lacks the ability to incorporate the simultaneous behaviours of probability-generating potentials at the fundamental level. Moreover, the emphasis of the situational aspect of quantum phenomena, which involves the consideration of events at more than one time, is much hindered by its differential nature. I therefore conclude that Bohmian quantum mechanics remains subordinate to the Feynman Path-Integral approach, which as we shall see in the next chapter, allows ones to incorporate the simultaneous actions of the probability-generating potentials at a fundamental level; and its integral structure accommodates most naturally the consideration of events at more than one time.

Chapter 7

Feynman's Path Integrals (Sum-Over-Histories) Formulation of Quantum Mechanics and Causal Processes

7.1 Introduction

In view of the conceptual problems associated with the interpretation of the wavefunction ψ, and also because of the fact that the descriptions of quantum phenomena require the consideration of events at different times, Chapter 6 concluded that an integral approach to the formulation of quantum mechanics (which takes into account the situational aspects of probability-generating potentials) might provide a more appropriate setting than the usual Schrödinger differential approach for the extension of the notion of a causal process to the quantum realm. Such a scheme - the ***Feynman's Path Integral Formulation of Quantum Mechanics*** - is discussed in the present chapter.

To appreciate the full impact of Feynman's approach, we must first come to gain some familiarity, in Section 7.2, with the role of the *Principle of Least Action* in classical mechanics. There, we learn how a classical system picks its one trajectory between two locations in spacetime out of the many that are presented to it with the help of the calculus of variations.

Feynman's Path Integral approach takes full advantage of the semantics of "*all* *possible paths*" that are presented to a physical system. These "possible paths" are construed as the different alternative configurations for the dynamics of the system, and this is where the formulation makes contact with probabilistic considerations. The rationale behind this approach is discussed in Section 7.3.

The Path Integral approach is then introduced in Section 7.4, where we shall show in details the procedure of " *summing-over-all-paths*" to obtain the overall probabilistic potential that operates on the system. I then conclude in Section 7.5 that the probability amplitude generated by the "sum-over-all-paths" in Feynman's approach may be regarded as a quantum causal process in the spirit of **HCQT**.

Finally, I summarize the main theses of Chapters 5 to 7 in Section 7.6.

7.2 The Principle of Least Action

Before embarking on an exposition of Feynman's formulation, it is instructive to consider first the *Principle of Least Action* in relation to classical mechanics. This enables us to compare and contrast in a fruitful manner its use in classical mechanics with the principle as applied to quantum mechanics.

The Principle of Least Action first came to acquire a mathematical definite form in the hands of Euler and Lagrange. We shall now introduce the principle by considering the simple case of a classical particle of mass m moving along one-dimension under a potential $V(x)$. According to Newton's Second Law,

$$m\frac{d^2x}{dt^2} = -\frac{dV}{dx} \tag{7.1}$$

Given the initial values of the position and velocity of the particle, we are able to calculate similar quantities for a later time $t + \Delta t$ and the process may be repeated to "inch" forward to a still later time $t + 2\Delta t$ and so forth. The rationale behind this *differential* approach is that being in a definite position with a definite velocity constrains the particle to one particular motion along one trajectory in spacetime.

An equally valid way to formulate the problem is to ask: given that the particle is at positions x_i and x_f at times t_i and t_f respectively, what is there to distinguish the actual trajectory $x_{cl}(t)$ from *all other trajectories* or *paths* that connect these points? Under the terrain of determinism - in the crudest form, the thesis that *dynamical data given at an earlier time completely determine those at a later time* - this seems a question that is easily tackled. It is because the actual trajectory is constrained by the initial data given. This constraint serves to select the actual path from all other possible ones that the particle *would have followed* if given different sets of initial data. This is the usual differential approach to the problem. On the other hand, the integral approach is global in nature in the sense that it attempts to determine in one stroke the entire trajectory $x_{cl}(t)$, in contrast to the local approach that concerns itself with the particle's plan of action in the next infinitesimal stretch of time.

At this point, it seems a sheer matter of preference for one to adopt either the differential or integral approach to the problem. But a significant advantage of integral approaches over that of differential ones has already emerged even in the domain of classical physics. It is a necessity for the differential approach to introduce special point coordinates to maintain the correct bookkeeping of the motion of the particle in the next infinitesimal spacetime point. Its exact formulation depends heavily on the choice of coordinates and can usually get, in a technical sense, rather complicated and cumbersome. Whereas in the integral approach, one deals with the "comparisons" of the magnitudes of actions (as measures of motion) of the possible paths and differences in the magnitudes are *invariant* with respect to any system of coordinates.

In the integral approach, the motion of a particle is seen to be constrained by events at more than one time. Given the events that the particle is at position x_i at t_i and x_f at t_f, it must have had certain definite (although we are not interested in their individual respective values) velocities to enable it to get from x_i to x_f in the time interval $t_f - t_i$. In other words, it must have taken up specific positions in-between those two end-points and with specific velocities during the interval in order to move from one point to another in the prescribed time. The emphasis here is placed not on these individual values but rather on *the path as an entirety.* Take for instance the "crossed-over" point M in Figure 7-1 that belongs to both the actual trajectory and another possible candidate (a). This piece of information alone is not useful at all in telling the actual path from the rest, for it can either be part of the actual motion or that of the virtual motion (a) . Likewise, it is equally non-illuminating and un-helpful if even a greater number of individual "crossed-over" points such as X, Y and Z are specified. In the integral approach where one focus on the path as a whole, any local information in regard to an individual spacetime point becomes redundant as any of these points may appear on the course of more than one path.

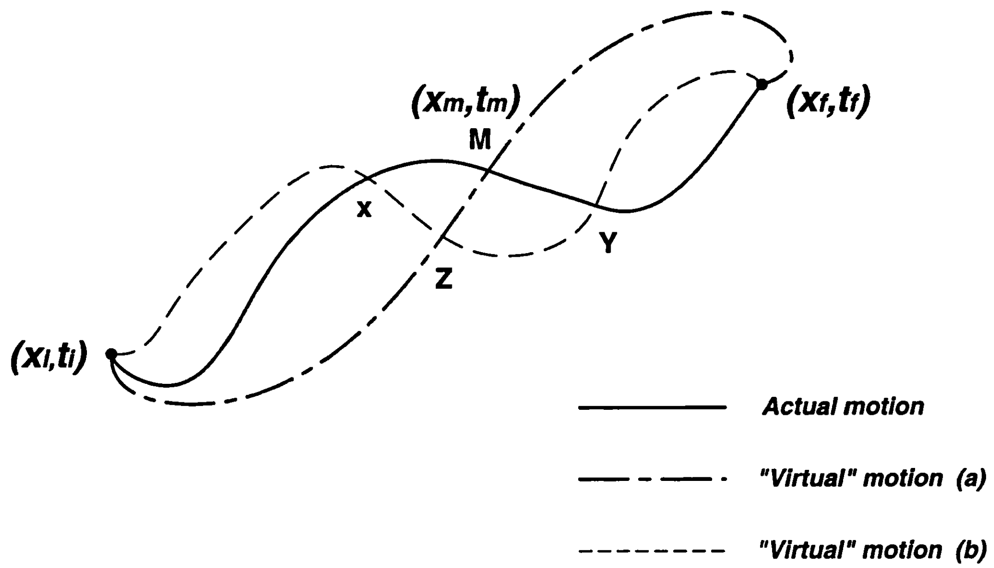

Figure 7.1: "Possible motions of a particle between two specific spacetime points.

Disregarding the individual intermediate spacetime points, what other means may one fall back on in order to distinguish the actual path of the particle between two specific points?

We find a useful hint in this original idea of minimum principles: Nature favours the one motion with the minimum *expenditure of effort.* It transpires that the phrase "expenditure of effort" has already appeared in the context of the familiar notions of "work" and "energy". When one speaks of an expenditure of some kind, one refers to a specific amount of a certain quantity. When an external force $\mathbf{F}$ is applied to a mass m and the latter is moved through a distance ds from point A to point B, the "effort" of **F** on m is then *quantified* by the notion of "*work done*", dW[1], by **F**. This is represented mathematically by the expression: $W = \int_A^B \mathbf{F}\cdot d\mathbf{s}$. W is a scalar quantity (it does not depend on direction) that describes the entire motion: the traversal of the distance $d\mathbf{s}$ by m under the influence of **F**. In this sense, W is a "global" quantity that refers to the path as a collective whole. In particular, it is futile to speak about the work done at an isolated spacetime point.

The concept of work thus forms the link between the abstract phrase "*expenditure of effort*" and the physical notion of *energy* in the following manner. Given $\mathbf{F} = m\ddot{\mathbf{r}}$, by Newton's Second Law,

$$\int_A^B \mathbf{F}\cdot d\mathbf{r} = m\int_A^B \ddot{\mathbf{r}}\cdot d\mathbf{r} = \frac{m}{2}\int_A^B \frac{d}{dt}(\dot{\mathbf{r}}\cdot\dot{\mathbf{r}})dt \tag{7.2}$$

Since $\dot{\mathbf{r}}\cdot\dot{\mathbf{r}} = v^2$, with v being the scalar magnitude of the velocity of the particle. It follows, therefore,

$$W = \int_A^B \mathbf{F}\cdot d\mathbf{r} = \frac{1}{2}m\int_{v_1}^{v_2} d(v^2) = \frac{1}{2}mv_2^2 - \frac{1}{2}mv_1^2. \tag{7.3}$$

The above equation shows that the work W done in moving the mass m from location A to B is equal to the *kinetic energy* ($\frac{1}{2}mv^2$) acquired by the system. So the "work done" by the force is "turned" into a kind of mobile (kinetic) energy of the object that is being moved. To make the correspondence more precise, it is reasonable to conceive of the work W as the dissipation of some form of energy and the loss of this energy is exactly compensated by the gain in kinetic energy by the mass m. Denoting this diminution of energy by the quantity $-dV$, we have,

$$-dV = dW = \mathbf{F}\cdot\mathbf{s}$$

$$dV = -\mathbf{F}\cdot\mathbf{s}$$

[1]Strictly speaking, the work done is not just equal to "force times distance" as often stated. Rather, it is the "component of force along the path times the path length", since force is a vector quantity with its effects depending on the direction of its application.

$$\Rightarrow \quad \mathbf{F} = -\frac{\partial V}{\partial s} \tag{7.4}$$

Provided that the quantity $\mathbf{F}\cdot\mathbf{s}$ is a perfect differential, there exists a function V that depends only upon the initial and final positions of the system, which is independent of the actual path taken between these positions[2]. V is called the "potential energy" and is defined for every point in space.

To be precise, it is the difference $\triangle V$ in the potential energy between two spacetime points that is independent of path and this difference finds its way into an increase (which is also in the form of a difference) in the *total* kinetic energy of the particle in its traversal from one point to another. It is indeed the total differences of both kinds of energies at each intermediate spacetime point of each path that are path independent (the velocities of the particle may differ), and this accounts for the fact that some paths require the particle to take a longer time to traverse given the *same* total amount of energy.

The brief digression has led us to this important clue: the concept of energy may be employed as our means to distinguish the actual path of a particle's motion from all other possible ones. The quantity we shall work with is, however, not the total energy but rather a quantity that corresponds to the *difference* between the kinetic and potential energies at each point along a path.

We shall now explain how the actual path followed by the particle is to be distinguished mathematically by the integral method.

One starts by first defining some quantity for each *tentative* path that is used for comparison. This quantity is the function $\mathcal{L}$, called the *Lagrangian* and given by $\mathcal{L} = T - V$, where T and V are the kinetic and potential energies of the particle. Since the kinetic and potential energies depend on $\dot{x}$ and x respectively, it follows that $\mathcal{L}$ is a function of both of these quantities.

Having defined the quantity $\mathcal{L}$ for comparison, one must, for each tentative path

[2]This is only true if in the ideal case where there is no other means of dissipation of potential energy (like "heat" to the environment), except for its being converted solely into the kinetic energy of the mass.

$x(t)$ connecting the endpoints (x_i, t_i) and (x_f, t_f), compute the *action* $\mathcal{S}[x(t)]$ of the path according to the following expression,

$$\mathcal{S}[x(t)] = \int_{t_i}^{t_f} \mathcal{L}(x, \dot{x}) dt \tag{7.5}$$

$\mathcal{S}$ is a function of the entire path and so depends on the function $x(t)$. This dependence of $\mathcal{S}$ on another function ($\mathcal{L}$) makes $\mathcal{S}$ *a function of a function* and the special name *functional* is introduced to denote this fact. It ought to be stressed that $\mathcal{S}$ depends not just on the value of x at some time t. To give the favour of a functional, we consider the following.

In Figure 7-2, we consider three different paths A, B and C, connecting the two endpoints (x_i, t_i) and (x_f, t_f), that the particle could have traveled along. For each of the three curves, there is a position x as well as a velocity $\dot{x}$ for the particle at each point along the curve. These data generate the three "base" curves C_A, C_B and C_C in the three plots (a), (b) and (c) respectively (these curves[3] are not replicas of the original curves A, B and C given that the latter all lie in three-dimensional space and time). There is to every point on the curves C_A, C_B and C_C a value corresponding to the Lagrangian $\mathcal{L}$. The collection of values forms the curves $\mathcal{L}_A$, $\mathcal{L}_B$ and $\mathcal{L}_C$. In other words, at every point $(x, \dot{x})$, there corresponds a number that gives the Lagrangian at that point since the $\mathcal{L}$'s themselves are functions of x and $\dot{x}$.

The functionals $\mathcal{S}_A$, $\mathcal{S}_B$ and $\mathcal{S}_C$ may now be calculated for each of the paths (Figures 7-2 and 7-3), bearing in mind that each point on the curves $\mathcal{L}_A$, $\mathcal{L}_B$ and $\mathcal{L}_C$ is itself a function of x, $\dot{x}$ and t. The values of $\mathcal{S}_A$, $\mathcal{S}_B$ and $\mathcal{S}_C$ are then obtained by the familiar summation procedure of finding the area under a curve - the Riemann integral method. The value of each of these areas corresponds to the functionals $\mathcal{S}_A$, $\mathcal{S}_B$ and $\mathcal{S}_C$ respectively. It is obvious that the value of $\mathcal{S}$ is determined by the overall shape of the *entire* curve $\mathcal{L}$ and not just by selected points on $\mathcal{L}$.

With values of the action $\mathcal{S}$ computed for each path, we can now compare these

[3]The specification of a particle's motion by its position and velocity (rather than position and time) generates a curve in phase-space instead of the three-dimensional real space.

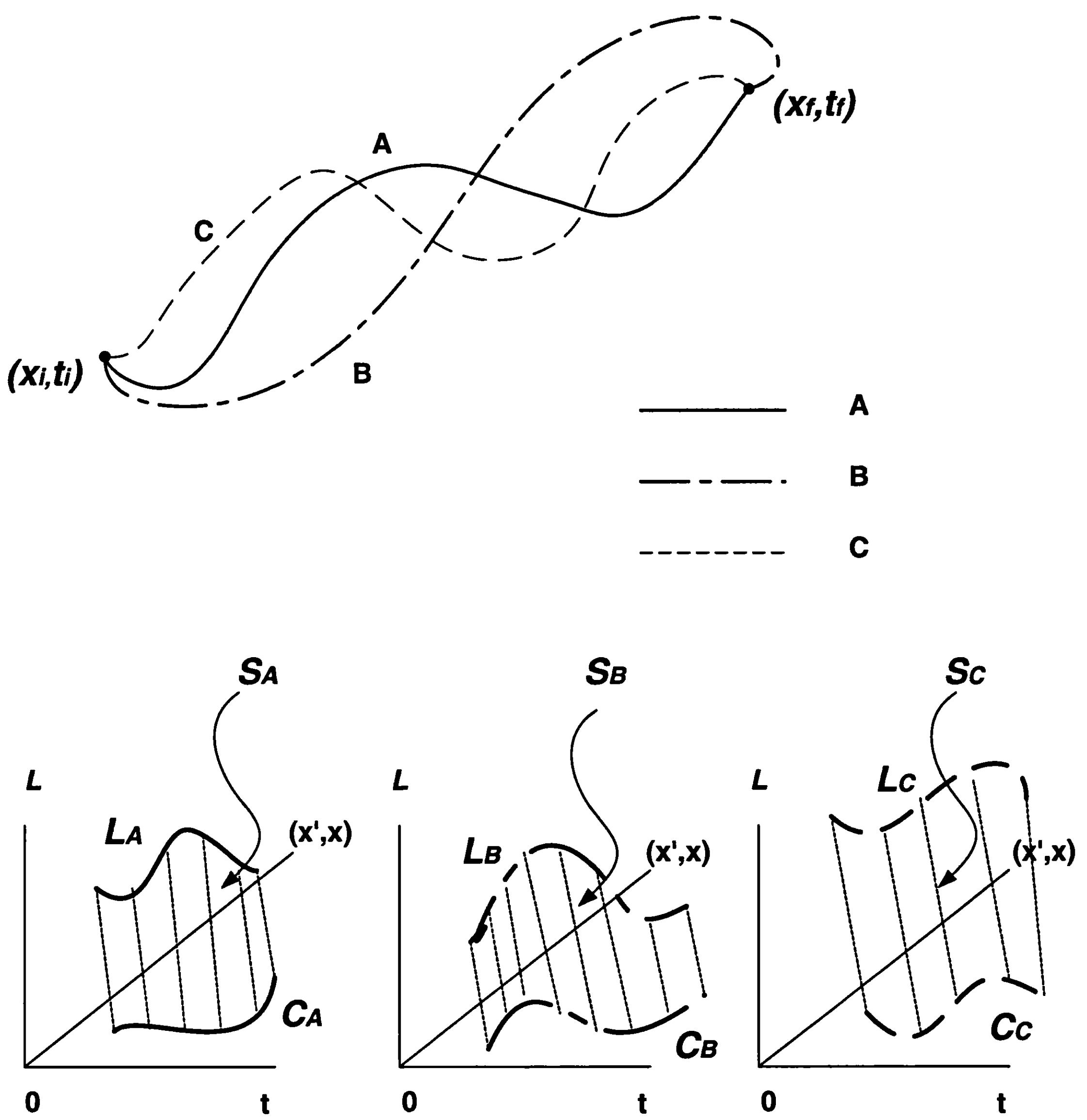

Figure 7.2: Graphical illustration of the notion of a functional.

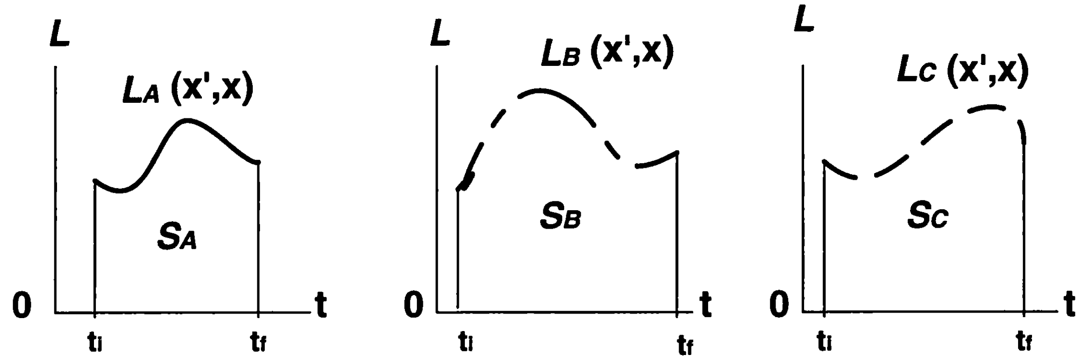

Figure 7.3: Graphical illustration of the notion of a functional.

values by imposing the criterion that *the classical path is the one whose action has the minimum value.* We are to find out for which curve the value of $\mathcal{S}$ is the least.

The calculus of variations exempts us from the seemingly impossible chores of having to compute the action for every single path connecting the two spacetime points[4]. This is rendered possible by the crucial observation that for the *true* path - the one with the minimum action - no change in the value of the action would be experienced by another path that differs from this path only by a tiny amount, so that in the *first order*, there is *no* difference in the values of the actions.

One thus starts with the postulation of a true classical path $x_T(t)$ (although unknown) describing the motion of the particle. And to $x_T(t)$, a "small" fluctuation or variation $\eta(t)$ would take us to a very closely neighboring path: $x(t) = x_T(t) + \eta(t)$. The action of this neighboring path can be obtained by the substitution of $x(t)$ into the expression for $\mathcal{S}$,

$$\mathcal{S} = \int_{t_1}^{t_2} [\frac{m}{2}(\frac{dx_T}{dt} + \frac{d\eta}{dt})^2 - V(x_T + \eta)]dt \tag{7.6}$$

Multiplying out the quadratic term and expanding the V-term as a Taylor series gives,

$$\mathcal{S} = \int_{t_1}^{t_2} \frac{m}{2}[(\frac{dx_T}{dt})^2 + 2\frac{dx_T}{dt}\cdot\frac{d\eta}{dt} + (\frac{d\eta}{dt})^2 - (V(x_T) + \eta V'(x_T) + \frac{\eta}{2}V''(x_T) + ...)]dt \tag{7.7}$$

Re-arranging and leaving only terms up to the first order in equation (7.7) then yields,

$$\mathcal{S} = \int_{t_1}^{t_2} [\frac{m}{2}(\frac{dx_T}{dt})^2 - V(x_T) + m\frac{dx_T}{dt}\cdot\frac{d\eta}{dt} - \eta V'(x_T)]dt \tag{7.8}$$

The first two terms in equation (7.8) correspond to the action $\mathcal{S}_T$ for the true path x_T and so the remaining two terms may be regarded as the difference $\delta\mathcal{S}$ in the actions of the true and varied paths respectively,

$$\delta\mathcal{S} = \int_{t_1}^{t_2} [m\frac{dx_T}{dt}\cdot\frac{d\eta}{dt} - \eta V'(x_T)]dt \tag{7.9}$$

[4]In principle, there exists no limit for the number of paths that may connect two points in spacetime and that would mean one has to engage oneself in an arduous task of computing and comparing a potentially infinite number of actions.

The ingenuity of the method consists in the very fact that even though the form of x_{τ} still remains unknown at this stage, but since by definition, its action $\mathcal{S}_{\tau}$ is a minimum; any variation in $\mathcal{S}_{\tau}$ would then be in the first order. This means that the expression $\delta\mathcal{S}$ above must be zero, regardless of whatever value η assumes. The question is therefore reduced to how one may make the integral in equation (7.9) above zero.

The integration of the first term in the above expression may be carried out by the procedure of "integration by parts". Recall that,

$$\begin{aligned} \tfrac{d}{dt}(\eta f) &= \eta\tfrac{df}{dt} + f\tfrac{d\eta}{dt} \\ \textstyle\int f\tfrac{d\eta}{dt} &= \eta f - \textstyle\int \eta\tfrac{df}{dt}dt \end{aligned}$$

With $f = m\frac{dx_{\tau}}{dt}$, equation (7.9) then becomes,

$$\delta\mathcal{S} = \int_{t_1}^{t_2}[m\frac{dx_{\tau}}{dt}[\eta(t)]_{t_1}^{t_2} - \int_{t_1}^{t_2}\frac{d}{dt}(m\frac{dx_{\tau}}{dt})\eta(t)dt - \int_{t_1}^{t_2}\eta(t)V'(x_{\tau})dt \tag{7.10}$$

Conveniently, the (first) integrated term in equation (7.10) vanishes because all paths - the actual and the varied ones - share the same two end-points so that $\eta(t_1) = \eta(t_2) = 0$.

The remaining terms may now be gathered together to arrive at,

$$\delta\mathcal{S}' = \int_{t_1}^{t_2}[-m\frac{d^2x_{\tau}}{dt^2} - V'(x_{\tau})]\eta(t)dt \tag{7.11}$$

Although small in magnitude, the variation $\eta(t)$ is capable of taking *any* value. It may be as small as one would like so that the "varied" path stays very close to the actual one. It follows therefore that the only way to make $\delta\mathcal{S}'$ zero is by demanding the expression within the square brackets to be equal to zero,

$$-m\frac{d^2x_{\tau}}{dt^2} - V'(x_{\tau}) = 0 \tag{7.12}$$

$$\Rightarrow -V'(x_\tau) = m\frac{d^2x_\tau}{dt^2} \tag{7.13}$$

And Equation (7.13) is none other than Newton's Second Law $F = ma$.

We began with the observation that for $\mathcal{S}$ to be a minimum, there ought to be no change in the action in the first-order for a nearby path that results when the actual path is varied by a minute amount $\eta(t)$ (with $\delta\mathcal{S} = 0$). $\delta\mathcal{S} = 0$ is satisfied when the actual path x_τ (with the minimum value of $\mathcal{S}$) obeys Newton's second law. Hence, we have demonstrated that the principle of least action implies that the path having the minimum action is the classical one satisfying Newton's law. This simple example illustrates for us the integral approach to a classical mechanical problem and demonstrates its equivalence to the usual differential approach by way of the familiar Newton's second law.

An important remark is in place. The above derivation of $\mathcal{S}$ has not shown that $\mathcal{S}$ is a *minimum*; indeed it may well also have been a *maximum*. This is what we have loosely meant by "least" when referring to the first-order change in the value of $\mathcal{S}$ as being zero when the path is changed.

7.3 The Principle of "Least" Action in Quantum Mechanics

Nearly six decades ago in the year 1942, Richard Feynman published a doctoral thesis with the title "*On the Principle of Least Action in Quantum Mechanics*"[5]. The material contained therein later appeared in a well-known paper in the Physical Review as "*Spacetime Approach to Non-Relativistic Quantum Mechanics*"[6]. In this work, Feynman draws on a proposal by Paul Dirac[7] to develop an integral approach to quantum mechanics that is shown to be equivalent to both the Schrödinger differential and the Heisenberg matrix mechanics formulations.

[5] Feynman, R.P. (1942). (See also Feynman, R.P. and Hibbs, A.R. (1965)).
[6] Feynman, R.P. (1948). Reprinted in Schwinger, J. (ed.) (1958), p.321-341.
[7] Dirac, P.A.M. (1933). Reprinted in Schwinger, J. (1958), p.312-320.

Consider the following comment from Feynman[8], as he tries to give an "animated" account of the action of a photon - a quantum of light energy - as described by the integral approach,

> *But all your instincts on cause and effect go haywire when you say that the particle decides to take the path that is going to give the minimum[9] action. Does it "smell" the neighboring paths to find out whether or not they have more action? In the case of light, when we put blocks in the way so that the photons could not test all the paths, we found that they couldn't figure out which way to go, and we had the phenomenon of diffraction.*
>
> *Is the same thing true in mechanics? Is it true that the particle doesn't just "take the right path" but it looks at all other possible trajectories? And if having things in the way, we don't let it look, that we will get an analog of diffraction? The miracle of it all is, of course, that it does just that! That's what the laws of quantum mechanics say. So our principle of least action is incompletely stated. It isn't that a particle takes the path of least action but that it smells all the paths in the neighborhood and chooses the one that has the least action by a method analogous to the one by which light chose the shortest time. You remember that the way light chose the shortest time was this: If it went on a path that took a different amount of time, it would arrive at a different phase. And the total amplitude at some point is the sum of contributions of amplitude for all the different ways the light can arrive. All the paths that give wildly different phases don't add up to anything. But if you can find a whole sequence of paths which have phases almost*

[8]Feynman R.P. (1963b), p.19-9.
[9]Strictly speaking, an extremum.

all the same, then the little contributions will add up and you get a reasonable total amplitude to arrive. The important path becomes the one for which there are many nearby paths which give the same phase.

It is just exactly the same thing for quantum mechanics. The complete quantum mechanics (for the non-relativistic case and neglecting electron spin) works as follows: The probability that a particle staring at point 1 at the time t_1 will arrive at point 2 at time t_2 is the square of a probability amplitude. The total amplitude can be written as the sum of the amplitudes for each possible path - for each way of arrival. For every $x(t)$ that we could have - for every possible imaginary trajectory - we have to calculate an amplitude. Then we add them all together. What do we take for the amplitude for each path? Our action integral tells us what the amplitude for a single path ought to be. The amplitude is proportional to some constant times $\exp^{iS/\hbar}$, where S is the action for that path. That is, if we represent the phase of the amplitude by a complex number, the phase angle is $S/\hbar$. The action S has dimensions of energy times time, and the Planck's constant $\hbar$ has the same dimensions. It is the constant that determines when quantum mechanics is important.

The consideration of the "*different possible paths*" that the motion of a particle may take makes the integral approach especially suitable to formulate a theory of probabilistic nature such as quantum mechanics. One may now speak comfortably and legitimately of a *single* particle being confronted with a number of *possibilities* corresponding to these different paths.

In the usual (Schrödinger) differential approach to quantum mechanics, we study the evolution of a wavefunction. But we also know that quantum systems reveal corpuscular properties. Following Born, we reconcile the ideas of a "definite" trajectory with probability considerations by the use of the concept of the wavefunction. As we have discussed, this is achieved by *interpreting* the wavefunction ψ as an entity whose modulus square gives the probability distribution when an observation is made. It then appears that the "particle" does *exist as* a woolly ball of possibilities when it is not localized by an act of observation. There seems to be a discontinuity in the ontology between the wavefunction as representing mere possibilities and the particle. Moreover, statistical behaviours only reveal themselves *en mass* insofar as an ensemble is concerned. This brings further complication as to whether ψ may indeed be only meaningfully associated with the ensemble rather than the individual members of the ensemble, which are deprived of their supposed mathematical representation.

At this point comes Feynman's enchanting magic. By drawing a close analogy with classical wave phenomena, he was able to recognize the essence of quantum mechanics: *the simultaneous actions of a multitude of co-present possibilities for a single physical system.* The classical integral approach to mechanics, founded on the least action principle, leads to the selection of the actual path in spacetime. But the integral formulation of quantum mechanics appeals to the intricate "mingling" of the "*various possible routes*" that are open to the system - the particle - in question.

With Feynman's integral approach, we have relieved ourselves from the shaky pursuit of injecting the probabilistic feature in the the quantum mechanical formulation via the interpretation of the wavefunction and all the conceptual difficulties to go with it. Rather, we may now rightfully entertain a picture with *a single system being subjected to different possibilities.* In this sense, the probabilistic feature is introduced into the theory from the outset. However, as we shall see, quantum mechanics is no ordinary stochastic theory. It is dissimilar from any stochastic theory that is disciplined by the axioms of classical Kolmogorovian probability measures. The considerations of probabilities have to be drastically revised to accommodate the fact that the alternative possible routes are *not* necessarily mutually exclusive in their operations on the physical system. In fact, this new mode of behaviour (of the probabilistic potentials) is central to the explanation of the existence of interference effects. With these remarks, we now turn proper to an exposition of Feynman's ***Sum-Over-Histories***, or the more commonly-known ***Path Integral Formulation of Quantum Mechanics***.

7.4 Feynman's Path Integral (Sum-Over-Histories) Formulation of Quantum Mechanics

> ***To summarize, when I was done with this, as a physicist I had gained two things. One, I knew many different ways of formulating classical electrodynamics, with many different mathematical forms. I got to know how to express the subject every which way. Second, I had a point of view - the overall spacetime point of view - and a disrespect for the Hamiltonian method of describing physics.***
>
> ***The character of quantum mechanics of the day was to write things in the famous Hamiltonian way - in the form of a differential equation, which described how the wavefunction changes from instant to instant, and in terms of an operator $\mathcal{H}$. If the classical physics could be reduced to a Hamiltonian form, everything was all right. Now, least action does not imply a Hamiltonian form if the action is a function of anything more than positions and velocities at the same moment. If the action is of the form of the integral of a function (usually called the Lagrangian) of the velocities and positions at the same time, $S = \int \mathcal{L}(\dot{x}, x)dt$, then you start with the Lagrangian and then create a Hamiltonian and work out the quantum mechanics, more or less uniquely. But this thing[10] involves the key variables, positions, at two different times, and therefore it was not obvious***

[10] By which he refers to an action for the motions of charges.

what to do to make the quantum-mechanical analogue.

These two excerpts have been taken from Feynman's Nobel Lecture in 1966, "*The Development of the Space-Time View of Quantum Electrodynamics*"[11]. Feynman's original motivation for developing an integral approach to quantum mechanics stems from the desire to generalize the Feynman-Wheeler absorber theory of classical electrodynamics to the quantum regime. The absorber theory is centered upon the ideas of "advanced" and "retarded" waves to deal with the feature of back-reaction amongst charges and this necessitates the consideration of the phenomena at more than one time. Thus, the integral approach that encompasses events at more than one time is a natural approach. At that time, the only known method for making the transition from classical to the quantum regime was via the Hamiltonian formulation that describes events at each particular time. This set Feynman to make a desperate search for an alternative method of quantization via the integral route.

One indisputable milestone in Feynman's quest is a paper by Dirac in 1933[12] in which some preliminary explorations concerning a Lagrangian approach to quantum mechanics is discussed. The important idea taken over by Feynman from Dirac's paper is that of a quantum mechanical *transformation function* that connects the values of a set of variables at one time t to that at another time T. Such a function has been shown by Dirac to be "*analogous to*" the quantity $\exp^{i\int_T^t \mathcal{L}dt/\hbar}$. It was Feynman who took the bold step forward and asserted that the vague phrase "*analogous to*", as hesitatingly uttered by Dirac, should indeed be construed as "*proportional to*". From this he convincingly demonstrated the recovery of the Schrödinger equation by the identification of the transformation function in quantum mechanics as exactly equal to $\exp^{i\int_T^t \mathcal{L}dt/\hbar}$. This forms the starting point of our introduction to Feynman's Path Integral method.

We first return to this "classical" question: in order to get from a certain point A in spacetime to another point B, which route should a particle take? This question is meaningful only insofar as there exists more than one route, in principle, that the

[11] Feynman, R.P. (1966), p.36.
[12] Dirac, P.A.M. (1933).

particle may follow. By "in principle", I mean in the absence of any considerations of whether it is in fact dynamically viable for the particle to do so. In a *prima facie* sense, there exist in principle, infinitely many paths that correspond to any combination of spacetime points along which the particle travels. Each of these paths represents a distinct possible route from A to B.

Classically, as we have seen, out of the many possibilities, the particle follows the *one* that comes under the dynamical constraint of the value of its action being an extremum. Given infinitely many possibilities, there is one that is unique and actually realized.

The ontology of paths via the integral approach as opposed to the usual differential framework sets the scene for probabilistic considerations already in a classical setting[13]. It would therefore seem most natural to adopt such an ontology as we make the transition to quantum mechanics.

The real value of Feynman's integral approach to quantum mechanics consists in the realization of the full potential of the concept "*all possibilities*"! In quantum mechanics, all the "possible paths" act in concurrence to result in one single or resultant probabilistic potential that governs the motion of the particle.

The distinction between "*one out of many possibilities actualizing*" (as in the classical domain) and "*all possibilities actualizing*" is reminiscent of our previous discussion of the difference in the behaviours of classical and quantum probabilistic potentials. Classical probabilistic potentials form a mutually exclusive set while their quantum counterparts do not form an exclusive one. The latter, as I have argued, is essential for the explanation of interference phenomena. It must be borne in mind that the objects (the electrons or photons) do not interfere but rather, the probabilistic potentials confronting them do!

This is what Feynman's formulation does. Bypassing all the messy ontology that burdens the "mythical" wavefunction, the theory goes straight to *periodic possibilities*

[13]In fact, the concept of "paths" has been widely used in physical processes. An example of that is Brownian motion in which a particle undergoes a large number of "random" interactions and hence may be conceived of as being pursuing a kind of random-walk motion. Such classical processes are dealt with probabilistically.

and their *concurrent actions.* As such, the theory deals with probabilistic considerations directly. This is introduced into two fundamental postulates[14],

Postulate I

If an ideal measurement is performed to determine whether a particle has a path lying in a region of spacetime, then the probability that the result will be affirmative is the absolute square of a sum of complex contributions, one from each path in the region.

Postulate II

The paths contribute equally in magnitude, but the phase of their contribution is the classical action (in units of $\hbar$); i.e. the time integral of the Lagrangian taken along the path.

Some of the terms that appear in the above postulates call for further clarification. First, by an "ideal measurement", Feynman refers to[15], "*...a measurement is made which is capable only of determining that the path lies somewhere within R (R is the spacetime region concerned)...The measurement is to be what we might call an "ideal measurement". We suppose that no further details could be obtained from the same measurement without further disturbance to the system. I have not been able to find a precise definition*".

The main obstacle in framing a precise definition of an "*ideal measurement*" stems from the fact that the very act of locating the particle by a position measurement alters the dynamics of it. An "ideal measurement" is therefore best to be thought of as a kind of "quasi-measurement" that serves really as a means for us to know that the particle's motion is contained *somewhere in a region of spacetime* but with no exactitude to pin it down to precise spatial points. This is best illustrated yet

[14]Feynman, R.P. (1948), p.371 and also in Schwinger, J. (ed.) (1958), p.325.
[15]ibid., p.370.

again by the two-slit experiment (see Figure 5-5). The recording of whether a photon has fallen somewhere or another on the detection screen constitutes a measurement to evidence the fact that after its emission from the source, (the dynamics of) the photon has been confined to the region of spacetime between the source and the detection screen (via, of course, the two slits). This is the kind of measurement taken to be an ideal measurement and can be contrasted with a measurement that is performed, for instance, at the slit in order to find out whether it has traversed the slit or not. With the latter kind of measurement, one does not obtain a distribution of photons showing the distinctive interference pattern on the detection screen. Such kind of measurements bring inevitable disturbances to the system as more information is extracted from it. So why must one be so elaborate on the notion of an ideal measurement in postulate I?

The distinction of both types of measurements was first made by Wolfgang Pauli[16]. Pauli called an ideal measurement a *measurement of the first kind*[17], "*...if the result of using the measuring apparatus is not known, but only the fact of its use is known (the measured quantity is unknown after the measurement, but is determinate),* ***the probability that the quantity measured has a certain value is the same both before and after the measurement.***"

One can see from Pauli's definition the crucial feature for an ideal measurement as one which brings no change to the probability that the quantity "measured" has a certain value. In the double-slit experiment, the presence of the slits has restricted the number of paths - the number of possible routes of the photon - to two, and these are the two contributions that find their ways into the calculation of the probability. One may think that this manner of restricting the number of possible routes of the photons by the two slits could constitute a kind of measurement. Although this is a reasonable way of thinking, "measurements" of such kinds would not generally be considered as actual measurements of the precise position of each photon. This is because, as far as each photon is concerned, it is still presented with two possibilities

[16]Pauli, W. (1980), p.75.
[17]ibid.

as they are emitted from the source and these possibilities interfere! The entire setup, from the source to the detection screen via the two-slits together constitute an ideal measurement. We *know* that photons are emitted from the source and they eventually reach the detection screen. There is also the obvious information that they somehow traverse the two-slits (notice that I have taken care not to say "*one*" of the slits) and this accounts for the appearance of the interference pattern. In the case of no observation of photons at the slits , we know *before* the "measurement" (between the photon and the two slits) that there is the probability for the photon to traverse the two slits. However, once the photon has "passed" the slits (that is, *after* the "measurement") and reaches the detection screen, all we know is *still* that the photon had a certain probability to traverse each of the slits. It is in this sense that our knowledge has not altered.

Moreover, upon the closure of one of the slits, the photons are given one distinct possible route and the interference pattern vanishes. Therefore, the manipulation of the apparatus within the spacetime region alters the behaviours of the photons and this is possible only if the photons have their existence within that very region. An ideal measurement therefore preserves the essential concurrent behaviours of probability-generating potentials in the quantum domain, which give rise to interference.

It must, however, be noted that Postulate I makes only a somewhat vague mention of "complex contributions" and the exact nature of these are clarified in Postulate II. Postulate I states that these complex contributions, one from each path in the region, are the main ingredients for the computation of the probability that a particle would make a transition from one spacetime point to another (both being those defining the region).

There is one more important point that emerges from Postulate I. Feynman gives the definition of a path as follows[18],

> ***A path is first defined only by the positions x_i through which it goes at a sequence of equally spaced time, $t_i = t_{i-1} + \epsilon$.***

[18]Feynman, R.P. (1948), p.371.

The notion of a path is intended by Feynman to encompass a wider meaning than just the spatiotemporal path of a corpuscular entity that is relatively localized in space. As he remarks[19],

> ***When the system has several degrees of freedom the coordinate space x has several dimensions so that the symbol x will represent a set of coordinates $(x^{(1)}, x^{(2)}, ..., x^{(k)})$ for a system with k degrees of freedom. A path is a sequence of configurations for successive times and is described by giving the configuration x_i or $(x^{(1)}, x^{(2)}, ..., x^{(k)})$, i.e. the value of each of the k coordinates for each time t_i.***

One should not be seized by surprise that the concept of a path is to be generalized this way, especially when we recall that the term *path* is intended to mean a *possibility* or an *alternative.* This generalization enables us to deal with fields, a good example of a physical system with several degrees of freedom. A field is defined by a collection of spatial variables at each time and the collection as a whole changes from one time to another. The configuration at each of the successive times may therefore be thought of as collectively forming a "path". We see once again that the more appropriate terminology of a "*history*" may be adopted to accommodate a temporal sequence of extended spatial configurations.

Postulate II serves two important functions,

(1) it provides the recipe for calculating the sum of the complex contributions and,

(2) it injects dynamical content into the probabilistic consideration by defining the phase of a contribution with the classical action along a path.

Discussions of the second postulate invariably center upon one worrying issue: the assumption that *all* paths contribute *equally in magnitude.* Given that every path, including the classical one $x_{cl}(t)$, carry the same weight, how can it be possible to regain classical mechanics if the classical path is not in some sense favored?

[19]ibid.

Three observations help to dispel this worry. First, it has been stated clearly in Postulate I that it is not the sum of the complex contributions but rather, the *modulus square* of this sum that gives the probability measure. Thus, the usual consideration of (equal) weights in the context of classical probabilities is not necessarily expected to strictly apply. The second observation has to do with the way we approach the problem. Be it classical or quantum mechanics, we start with the *same* problem: given that a particle travels from point A to point B, which route should it take?

In the absence of information otherwise (for example, that there are obstacles placed between those two points), we must give all conceivable routes from A to B a democratic treatment and assign equal weight to each of them. The classical path is picked out by the requirement that it is one with its action being an extremum when compared to those actions along all the other paths between the two points. This requirement is an extra assumption notwithstanding that it is effective in recovering the classical path as described also by Newton's second law. Different weights should therefore not be attached to any path in a *prima facie* fashion.

The third observation concerns the fact that if the complex contributions are to be identified with the probability-generating potentials that we have introduced, then we may consider the superposition of these contributions in an analogy to the superposition of forces in classical dynamics. It is true that when two forces, one that has twice the magnitude of the other, act simultaneously on a particle, the motion of the particle would naturally be more biased towards the influence of the larger force. In the case of the probability-generating potentials, the differences in magnitude of their respective influences have found their manifestation in the phases of the paths via the action functions $\mathcal{S}$'s.

Furthermore, the superposition of these complex contributions is a one-to-one affair - each potential contributes only once, as similar to the case of the forces so that the larger and the smaller forces both contribute once. In the case where they are two identical forces, we see them as a "single" resultant force with double the magnitude but this "single" force still contributes once. In a way the "weights" have been *smuggled in* through the magnitudes of the forces in the classical scenario; here the weights find their ways in through the phases. We ought also to emphasize that the phases are periodic in character so that the overall phase difference is responsible for the appearance of the interference pattern.

Feynman's prescription is most simply illustrated by way of an elementary example: the *free particle.* See Figure 7-4.

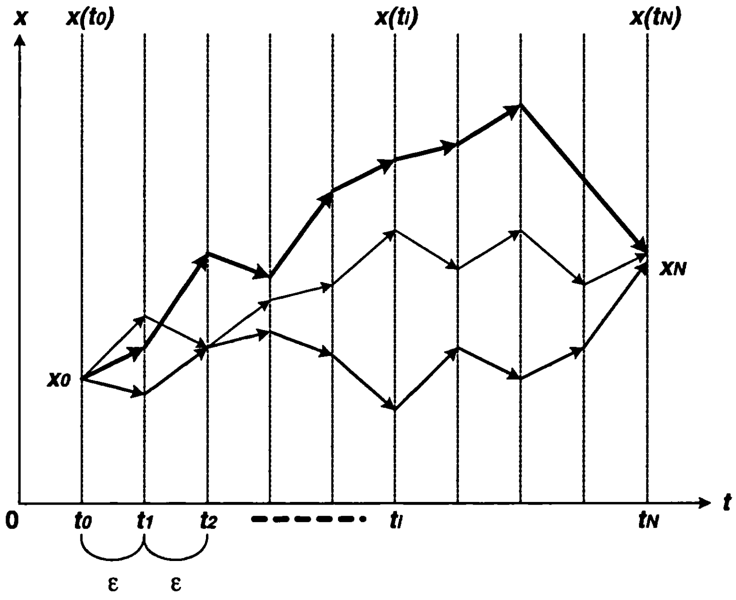

Figure 7.4: The discrete approximation to a path $x(t)$.

Figure 7-4 shows three examples of a path, labeled by $x(t)$, that could be followed by a particle in reaching the spacetime point (x_N, t_N) from (x_0, t_0). Following Feynman, the sequence of times $t_0 < t_1 < t_2 ... < t_N$ is taken to be of equal spacing ϵ. At each of these time points, the particle can take up a position anywhere between $-\infty$ and $+\infty$. The procedure of "*summing over all paths (or histories)*" amounts to integrating over each $x(t_i)$ (that is, all values of x at each time point) from $-\infty$ and $+\infty$ to take account of all the possibilities. Once all the individual integrations at each of the x_i's are carried out, we then let N go to the continuum limit ($N \rightarrow \infty$) to signify that the integration is performed over every point t in the interval t_0 to t_N.

Denoting the quantity whose modulus square is the probability of finding the path of a particle to lie in a specific region of spacetime R by $\phi(R)$, the two postulates may now be combined to yield $\phi(R)$ as follows,

$$\phi(R) = \lim_{N\to\infty,\epsilon\to 0} A\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}...\int_{-\infty}^{\infty}\exp[i\mathcal{S}[x(t)]/\hbar] \times dx_1 dx_2...dx_{N-1} \tag{7.14}$$

Notice that the integration over all paths is carried out for all values of $x(t_i)$ that lie between x_0 and x_N. The two endpoints x_0 and x_N are left out of this procedure as they are both fixed and shared by all paths. The factor A in front is required to reach the correct scale for $\phi(R)$ when the limit $N\to\infty$ is taken.

To simplify the form of equation (7.14), we introduce the "shorthand" $\mathcal{D}[x(t)]$ to symbolize the fact that an operation of integrating over all paths is performed and the equation may be rewritten as[20],

$$\phi(R) = \int_{x_0}^{x_N}\exp[i\mathcal{S}[x(t)]/\hbar]\mathcal{D}[x(t)] \tag{7.15}$$

In order to proceed, we must obtain an expression for the action. As one needs only to deal with the presence of kinetic energy of the free particle, its action takes the most simple form,

$$\mathcal{S} = \int_{t_0}^{t_N}\mathcal{L}(t)dt = \int_{t_0}^{t_N}\frac{1}{2}m\dot{x}^2 dt \tag{7.16}$$

For our "discretized" paths, $\dot{x}$ is approximated by $\frac{x_{i+1}-x_i}{\epsilon}$ and equation (7.16) is thus modified as,

$$\mathcal{S} = \sum_{i=0}^{N-1}\frac{m}{2}(\frac{x_{i+1}-x_i}{\epsilon})^2\epsilon \tag{7.17}$$

where $\epsilon = (\frac{t_N - t_0}{N})$. And so equation (7.14) becomes,

[20]The factor A has been absorbed into $\mathcal{D}[x(t)]$.

$$\phi(R) = \lim_{N\to\infty,\epsilon\to 0} A\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}...\int_{-\infty}^{\infty}\exp[\frac{i}{\hbar}\frac{m}{2}\sum_{i=0}^{N-1}\frac{(x_{i+1}-x_i)^2}{\epsilon}] \times dx_1dx_2...dx_{N-1} \tag{7.18}$$

To facilitate further computation, let us now switch variables from x_i to y_i by the transformation,

$$y_i = (\frac{m}{2\hbar\epsilon})^{\frac{1}{2}}x_i \tag{7.19}$$

We then evaluate,

$$\phi(R) = \lim_{N\to\infty} A'\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}...\int_{-\infty}^{\infty}\exp[-\sum_{i=0}^{N-1}\frac{(y_{i+1}-y_i)^2}{i}] \times dy_1dy_2...dy_{N-1} \tag{7.20}$$

where $A' = A(\frac{2\hbar\epsilon}{m})^{\frac{(N-1)}{2}}$.

Let us now begin by performing the y_1 integration and focus on the part of the integrand just involving y_1,

$$\int_{-\infty}^{\infty}\exp\{-\tfrac{1}{i}[(y_2-y_1)^2+(y_1-y_0)^2]\}dy_1$$

$$= \int_{-\infty}^{\infty}\exp\{i({y_2}^2-2y_1y_2+{y_1}^2+{y_1}^2-2y_1y_0+{y_0}^2)\}dy_1$$

$$= \int_{-\infty}^{\infty}\exp\{i[2{y_1}^2-2(y_2+y_0)y_1+({y_2}^2+{y_0}^2)]\}dy_1$$

$$= \exp[i({y_2}^2+{y_0}^2)]\int_{-\infty}^{\infty}\exp[2i{y_1}^2-2i(y_2+y_0)]y_1dy_1$$

$$= \exp[i({y_2}^2+{y_0}^2)]\int_{-\infty}^{\infty}\exp\{-(-2i){y_1}^2+[-2i(y_2+y_0y_1)]\}dy_1 \tag{7.21}$$

Using the following result for Guassian integrals,

$$I_0(\alpha,\beta) \;=\; \int_{-\infty}^{\infty} \exp^{-\alpha x^2+\beta x}\,dx \;=\; \exp^{\frac{\beta^2}{4\alpha}}(\frac{\pi}{\alpha})^{\frac{1}{2}} \tag{7.22}$$

with $\alpha = -2i$ and $\beta = -2i(y_2 + y_0)$, the integral in equation (7.21) thus becomes,

$$\exp[i({y_2}^2+{y_0}^2)]\exp[\frac{(-2i)^2(y_2+y_0)^2}{4(-2i)}](\frac{\pi}{-2i})^{\frac{1}{2}}] \tag{7.23}$$

$$= \exp[i({y_2}^2+{y_0}^2)]\exp[\frac{(y_2+y_0)^2}{2i}](\frac{i\pi}{2})^{\frac{1}{2}} \tag{7.24}$$

$$= (\frac{i\pi}{2})^{\frac{1}{2}}\exp[\frac{-2{y_2}^2-2{y_0}^2+{y_2}^2+2y_2y_0+{y_0}^2}{2i}] \tag{7.25}$$

$$= (\frac{i\pi}{2})^{\frac{1}{2}}\exp[\frac{-(y_2-y_0)^2}{2i}] \tag{7.26}$$

In a similar manner, it may also be shown that the part of the integrand just involving y_1 and y_2 is evaluated as,

$$[\frac{(i\pi)^2}{3}]^{\frac{1}{2}}\exp[\frac{-(y_3-y_0)^2}{3i}] \tag{7.27}$$

and so forth.

And so the pattern emerges that if the process of integration is carried out $(N-1)$ times, it will become,

$$[\frac{(i\pi)^{N-1}}{N}]^{\frac{1}{2}}\exp[\frac{-(y_N-y_0)^2}{Ni}] \tag{7.28}$$

In terms of the original coordinates x_i,

$$[\frac{(i\pi)^{N-1}}{N}]^{\frac{1}{2}}\exp[\frac{-m(x_N-x_0)^2}{i2\hbar N\epsilon}] \tag{7.29}$$

At last, we are in a position to evaluate $\phi(R)$. Equation (7.18) becomes,

$$\phi(R) = A(\frac{2\hbar\epsilon}{m})^{\frac{N-1}{2}} \cdot \frac{(i\pi)^{\frac{N-1}{2}}}{N^{\frac{1}{2}}} \exp[\frac{im(x_N - x_0)^2}{2\hbar N\epsilon}] \tag{7.30}$$

Thus,

$$\phi(R) = A(\frac{2\pi\hbar\epsilon i}{m})^{\frac{N}{2}} \cdot (\frac{m}{2\pi\hbar N\epsilon i})^{\frac{1}{2}} \exp[\frac{im(x_N - x_0)^2}{2\hbar N\epsilon}] \tag{7.31}$$

We may now complete the integration procedure by letting $N \rightarrow \infty$, $\epsilon \rightarrow 0$ with $N\epsilon \rightarrow (t_N - t_0)$. This is meaningful provided,

$$A = [\frac{2\pi\hbar\epsilon i}{m}]^{\frac{-N}{2}} \equiv B^{-N} \tag{7.32}$$

Following Feynman, it is conventional to associate a factor $\frac{1}{B}$ with each of the $N-1$ integrations and the remaining factor with the overall process. Therefore, the precise meaning of the antecedent "sum-over-all-paths" consists in,

$$\int \mathcal{D}[x(t)] = \lim_{N\rightarrow\infty, \epsilon\rightarrow 0} \frac{1}{B} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} ... \int_{-\infty}^{\infty} \frac{dx_1}{B} \frac{dx_2}{B} ... \frac{dx_{N-1}}{B} \tag{7.33}$$

where $B = (\frac{2\pi\hbar\epsilon i}{m})^{\frac{1}{2}}$.

The probability of finding the particle somewhere in the region R is obtained by taking the modulus square of $\phi(R)$. In order to distinguish it from its modulus square, $|\phi(R)|^2$, we call the quantity $\phi(R)$ the "*probability amplitude*". It remains to be shown that $\phi(R)$ is equivalent to the propagator in the more familiar Schrödinger dynamics.

To give a favour before attempting a more vigorous derivation, we may consider a region of spacetime bounded by the temporal coordinates t and t' and call it R'.

Analogously, another similar region of spacetime may be defined by t' and t'' and we call it R''. Defining R' and R'' in this manner with respect to temporal coordinates enables one to make the assertion that since R' lies entirely previous to R'', a probability can be associated with the path of the particle that *had been* in region R' and *will be* in region R'', given that t' is the present time. The amplitude we compute for region R' only depends on those times all previous to t' and emphatically so, it does not depend on the happenings to the system in those times that comes after t'.

We must first, however, remind the reader briefly about the propagator in the Schrödinger dynamics. One starts with the Schrödinger equation in the form of,

$$i\hbar|\dot{\psi}\rangle = H|\psi\rangle \tag{7.34}$$

and the propagator is a unitary matrix which evolves the state vector $|\psi(0)\rangle$ at $t = 0$ to $|\psi(t)\rangle$ at $t = t$,

$$|\psi(t)\rangle = U(t)|\psi(0)\rangle. \tag{7.35}$$

It can be shown that in the basis of energy eigenvectors (those of the Hamiltonian $\mathcal{H}$), $U(t)$ takes the form,

$$U(t) = \sum_{\alpha}\sum_{E}|E,\alpha\rangle\langle E,\alpha|\exp\frac{-iEt}{\hbar} \tag{7.36}$$

where α is a degeneracy label.

For a free particle, the time dependent Schrödinger equation is reduced to,

$$i\hbar\frac{\partial|\psi\rangle}{\partial t} = H|\psi\rangle = \frac{p^2}{2m}|\psi\rangle. \tag{7.37}$$

The energy eigenfunctions of the equation are solutions of the plane-wave form,

$$|\psi\rangle = |E\rangle \exp\frac{-iEt}{\hbar}. \tag{7.38}$$

Feeding this into equation (7.37) above, one then obtains the time independent Schrödinger equation for $|E\rangle$,

$$H|E\rangle = \frac{p^2}{2m}|E\rangle = E|E\rangle \tag{7.39}$$

with $p = \pm(2mE)^{\frac{1}{2}}$.

Hence, there are two eigenstates associated with each eigenvalue E,

$$|E_+\rangle = |p = (+2mE)^{\frac{1}{2}}\rangle$$

$$|E_-\rangle = |p = (-2mE)^{\frac{1}{2}}\rangle \tag{7.40}$$

.

One may use instead the compatible observable p that is non-degenerate and in the basis of the eigenkets $|p\rangle$ and according to equation (7.35), $U(t)$ becomes,

$$U(t) = \int_{-\infty}^{\infty} |p\rangle\langle p| \exp(\frac{-ip^2t}{2m\hbar})dp. \tag{7.41}$$

To show its equivalence to $\phi(R)$, $U(t)$ must be expressed in the x-basis and in terms of an orthogonal set of eigenvectors $|x\rangle$,

$$\langle x_N|U(t)|x_0\rangle = \int_{-\infty}^{\infty} \langle x_N|p\rangle\langle p|x_0\rangle \exp(\frac{-ip^2t}{2m\hbar})dp. \tag{7.42}$$

With $\langle x_N|p\rangle = \frac{1}{(2\pi\hbar)^{\frac{1}{2}}}\exp(\frac{ipx_N}{\hbar})$ and $\langle x_0|p\rangle = \frac{1}{(2\pi\hbar)^{\frac{1}{2}}}\exp(\frac{ipx_0}{\hbar})$, equation (7.42) becomes,

$$\langle x_N|U(t)|x_0\rangle = (\frac{1}{2\pi\hbar})\int_{-\infty}^{\infty}\exp(\frac{ip(x_N-x_0)}{\hbar})\cdot\exp(\frac{-ip^2t}{2m\hbar})dp. \tag{7.43}$$

The integral is evaluated by reference to the standard Gaussian integral,

$$\int_{-\infty}^{\infty}\exp(-ay^2+by)dy = (\frac{\pi}{a})^{\frac{1}{2}}\exp(\frac{b^2}{4a}) \tag{7.44}$$

with $a = \frac{-it}{2m\hbar}$ and $b = \frac{i(x_N-x_0)}{\hbar}$. Therefore,

$$\langle x_N|U(t)|x_0\rangle = (\frac{1}{2\pi\hbar})\cdot(\frac{2m\hbar\pi}{it})^{\frac{1}{2}}\exp\{[\frac{i(x_N-x_0)}{\hbar}]^2\times\frac{2m\hbar}{4it}\} \tag{7.45}$$

$$= (\frac{m}{2\pi\hbar it})^{\frac{1}{2}}\exp[\frac{im(x_N-x_0)^2}{2\hbar t}] \tag{7.46}$$

With $t = t_N - t_0$,

$$\langle x_N|U(t)|x_0\rangle = (\frac{m}{2\pi\hbar i(t_N-t_0)})^{\frac{1}{2}}\exp[\frac{im(x_N-x_0)^2}{2\hbar(t_N-t_0)}] \tag{7.47}$$

which is the quantity $\phi(R)$ obtained by the path integral method earlier.

We would also remark that one may make the transition to classical mechanics by appealing to the path-integral idea that the propagator is a sum over all paths of possibilities. In the classical regime where $\hbar \to 0$, the exponent in the expression for $\phi(R)$ (equation(7.31)) becomes a rapidly oscillating function over each of the variables x_i. As the x_i's vary, the positive and negative contributions very nearly cancel each other. The region at which x_i contributes most dominatingly is when the phase of the exponent varies *least* rapidly with x_i. In other words, the overall contribution that obtains is when $\frac{\partial S}{\partial x_i} = 0$ for all values of x_i with S being the action of a path.

The classical path is then the one with its action suffering no change in the first-order upon varying the path. What the path-integral has succeeded in telling us is that nature does *not* choose to pick the path with the action at an extremum as followed by a classical (macroscopic) system. Rather, because of the vast cancellations of the contributions from most paths in the classical regime as the quantum of action $\hbar$ approaches zero, only the one that is least affected by the fluctuations contributes to the overall dynamics and this is the classical path. On the classical level, the multitude of possibilities go through a procedure of self-adjustment (via phase cancellations) to leave us with one unique path.

This concludes our brief exposition of the basic ideas behind the *Feynman's Path-Integral Formulation of Quantum Mechanics.*

7.5 Feynman Paths and Causal Processes

In the final analysis, whether the "Feynman paths" qualify as causal processes is parasitic upon the fact of how well they fulfill, with suitable adjustments, the causal notions of *histories, possession and transmission* in accordance with **HCQT**.

HCQT defines a causal process as the "history of an object that transmits a conserved quantity " with a "history" understood to be a sequence of events - either continuous or discrete - in spacetime. An object must have an identity over time (in the sense already discussed in Section 4.5), and we admit also entities with infinite degrees of freedom like fields in addition to those which are spatiotemporally localized.

Since it is in the nature of quantum phenomena that events at different times are to be considered together in a collective manner, so the idea of a history as referring to an entire sequence of events is best suited for thinking about quantum physics. Feynman's histories - the possible paths - therefore provide just the paradigm for such a kind of history. Moreover, each of these histories corresponds to a possible dynamical evolution of the system in spacetime, giving a natural setting for incorporating probabilistic considerations at the fundamental level. In particular, the *sum-over-all-paths* procedure (that results in an overall probability amplitude $\phi(R)$) reveals the peculiarity in the behaviours of probability-generating potentials - their simultaneous actions - which are in sharp contrast to their behaviours in the classical domain. As such, a causal process in the Feynman Path-Integral formulation is the overall probability amplitude - which arises from the sum-over-all-paths procedure - for the object to make a transition from one spacetime region to another.

Next, we must examine the sense in which such Feynman histories can be said to possess a conserved quantity and to transmit this quantity. One crucial element must not be overlooked. For causal transmission to make sense at all, some physical quantity or more precisely, some conserved quantity - the subject of transmission - must be present.

It is at once noticed that it is rather futile to speak about the possession of a conserved quantity in quantum mechanics, with the notion of possession so bound up with the definite values of the dynamical attributes of a physical system. Quantum mechanics tells us that it is no longer possible for us to entertain the idea of physical attributes having definite values before a measurement[21]. This being said, however, quantum systems do experience exchanges of energy and momentum as they interact. This is possible if the system is, in a sense, in possession of such quantities. But the notion of possession must now be widened to express the fact that even though no definite amount of a conserved quantity can be associated with the system, the system is to be thought of as having these quantities in the following sense[22].

"Classical" conserved quantities such as energy and momentum find their ways into the Path-Integral formalism through the phase of each possible path or alternative. This is because the phase of each path is proportional to $\exp(\frac{i}{\hbar})\mathcal{S}[x(t)]$ and $\mathcal{S}[x(t)]$, the classical action is dependent on $\mathcal{L}(\dot{x}(t), x(t))$, which is an explicit energy expression for *every spacetime point* along the path. Also because of the $\dot{x}(t)$ dependence of $\mathcal{L}$, momentum is considered as obtaining at every spacetime point along the path.

We now turn to the notion of transmission. Central to the criterion of causal transmission is the "*at-at*" condition, which stipulates that a "quantity" is said to be transmitted from a spacetime point A to another point B, if (and only if) the said quantity is *possessed at* A *and at* B *and at every stage of the process between* A *and* B, without any intersections that involve exchange of a conserved quantity.

Under the usual differential approach, given the quantity is to appear at A and then at B, it is therefore quite natural to ask how the quantity gets from point A to point B. Salmon's reply, which has been rehearsed so many times, is that the quantity occupies all the intermediate points in turn between A and B. It is not necessary, says Salmon, to ask how the quantity get from one point to the next. However, it can be easily argued, as have been quite ably so by various critics, that the differential

[21]A measurement would put the system into the eigenstate of the physical attribute in question. System in an eigenstate of a dynamical observable has a definite value of that attribute.

[22]It may be argued that since the overall Feynman probability amplitude is equivalent to the wavefunction, might we not consider the modulus square of this quantity as a conserved quantity as previously in Chapter 6? However, it must be borne in mind that we do not, strictly speaking, have a notion of wavefunction here and so it is only consistent if we refrain from speaking about conserved quantities via the continuity equation as derived from the Schrödinger evolution.

approach emphasizes on the transition from one intermediate point to another.

In Feynman's approach, a phase is associated with each possible path between two points. Within the *sum-over-paths* procedure, one finally takes the limit $\epsilon \longrightarrow 0$ so that one essentially considers the entire set of the spacetime points between the two specific points A and B; and thus skillfully avoids the question of how to get from one (intermediate) point to the next. In this manner, the Path-Integral approach considers all the spacetime points at once in-between two specific points. There is a single probability amplitude $\phi(R)$ (whose modulus square gives the probability) of finding the particle to make a transition from its position at point A to that at point B. In effect, a transition probability $|\phi(R)|^2$ between two spacetime point is what is rightly referred to when one speaks of transmission in quantum mechanics. The probability for a particle to make a transition from one spacetime point to another is thus dependent on the energy and momentum being defined at every spacetime point in-between the two endpoints. As so, the condition of "possessing a conserved quantity at every spacetime point" in the "*at-at*" criterion is satisfied.

Hence, the Path Integral paradigm stays far more akin and faithful to Salmon's original intention for the "*at-at*" criterion, but with also the benefit of circumventing the need to address the question of "how exactly the intermediate points are traversed?"

However, the notion of transmission is defined in the absence of intersections that involve exchange of a conserved quantity. To be sure, any exchange of conserved quantities like energy or momentum would find its way in the adjustment of the phase. As I have argued in Section 6.4, one ought to specify that intersections, which destroy the quantum essence of superposition are the ones that should be absent in causal processes. These are the non-ideal measurements as opposed to the ideal ones discussed in Section 7.4 above.

Recall that the intended function of the stipulation of "the absence of any intersections that involve exchange of a conserved quantity" within the definition of transmission is to capture the fact that processes that are genuinely causal are able to sustain themselves without the assistance of outside sources. So is there any sense this important feature of transmission has been captured?

One thing we can be sure is that if the view of the superposition behaviours of probabilistic potentials in the quantum regime is taken seriously, then the description of a physical process like a photon is very much intertwined with its environment. In the two-slit experiment, the probability amplitude for the location of the photon takes account of the presence of the two slits. Such an "intersection" between the photon and the slits, however, does not sustain the photon in the sense that the environment may be altered (with the two slits removed for example) but we still obtain a probability distribution of the photon, albeit one that is different from the distribution when the slits are present. And this gives an indication that the photon is not sustained by the interactions with the slits.

Conveniently, one finds also, in the Path-Integral paradigm, the reply to the objections raised by the modalist. The chief criticism from the modal camp on the process causation view hinges upon the concern that the view does not explain why the particle takes a particular course but not others. In the context of classical physics, it is reasonable to put forth such an argument. The way that classical mechanics is formulated makes it remain agnostic as to this line of attack. Classically, the least action principle would not improve the situation: the appeal to the path being one with its action at the extremum would only delay the issue, for the modalist can effortlessly come up with a similar objection that it has not been explained as to why the path with an extremum action ought to be the one that the particle chooses to follow.

The Path-Integral paradigm, however, provides a twist in the debate. On this view, one considers *all* possibilities - all potential paths - the superposition of which results in a single "effective" probabilistic potential to "guide" the particle along in spacetime. This gives a stern reply to the modalist - it provides just a mechanism to pick out the "favored" path. This is demonstrated most convincingly in the transition from the quantum to the classical regime in the limit where $\hbar \rightarrow 0$. It is emphatically *not* the case that the particle chooses a particular path. Rather, all possibilities are utilized with each contributing equally. However, as $\hbar \rightarrow 0$, the phases associated with each path superpose either constructively or destructively, leaving only the one, which corresponds to the constructive superposition of phases, to survive and dominate and this *is* the classical trajectory.

7.6 Postlude

It is the aim of the last three chapters of this thesis to generalize the notion of causal processes in the theories of Salmon and Dowe to the realm of the microscopic quantum world. Chapter 5 explained the basic principles and peculiar features in the physical descriptions of quantum phenomena. In particular, the idea of *probability-generating potentials* was put forward to highlight the important differences between classical and quantum phenomena. It was argued that the difference essentially consists in the behaviours of the quantum probability-generating potentials from their classical counterparts. Quantum probability-generating potentials superpose and act in a simultaneous manner on a physical system, while classical ones operate in a mutually exclusive manner so that if one has been found to have acted on a certain occasion, the others would have ceased from their operations accordingly[23]. The superposition of probability-generating potentials is used to explain the appearance of interference characteristics exhibited in the probability distribution of photons in the paradigmatic example of two-slit experiment. Similarly, mutual exclusivity experienced in the activities of the probabilistic potentials is responsible, as has been argued, for the observation of normal statistical behaviours as expected from classical probability theory.

The double-slit experiment also brings to light the intimate relationship between the notion of a spacetime path of a particle - which forms the single most important element in the notion of a causal process - with probabilistic considerations. For the classical probabilist, randomness lies with the emission of photons by the source but once emitted into a particular direction, a photon is expected to follow a "prescribed" definite track. If its track is so directed to enable it to pass without hindrance through

[23]In other words, classical probability generating potentials cannot be operating in a simultaneous manner.

"*one of the two slits*", then it would certainly register itself on the detection screen at the far end. Under this circumstance, the two slits serve the purpose of selecting two *mutually exclusive* subsets of particles (in the sense that having a definite trajectory means the particle cannot be at the same time be travelling along another distinct trajectory) out of all subsets that correspond to the possible emission of photons in a 4π solid-angle.

For the quantum probabilist, things take on a very different outlook altogether. The interference pattern displayed in the distribution of recorded photons is not one that is expected from classical reasoning. Classically, the possessions of definite trajectories imply that the trajectories of the emitted photons, with especially the two subsets selected by the slits, form a mutually exclusive set; as a result of which a distribution that obeys the sum-rule of probability (one that shows no interference) obtains.

However, the actual distribution with the characteristics of interference is not one that arises should the sum-rule be conformed to. From this we conclude that the "trajectories" do not form a mutually exclusive set. In other words, the photon is *in some sense* "associated with" both trajectories. I have taken extra care not to use the expression "*travelling in*" (both trajectories), for it would prove most disconcerting as to the meaning of an spatiotemporally localized corpuscular entity to be "spreaded-out" between two trajectories - two distinctively localized lines of spacetime points - that only extended entities in spacetime are capable of.

The appearance of the interference characteristics further supports a view of an extended entity. Thus the "wavefunction" whose mathematical form is inspired by the extended physical wave was invented. The very fact that the interference pattern emerges even with the individual photons reaching the screen *one at a time* implies that such a "waveform" cannot be identified with that of a classical wave, and we have seen that it had to acquire its meaning through a probabilistic interpretation (its modulus square gives the probability distribution of photons).

The whole procedure seems artificial: having to associate a wave with the particle[24] and then to impose an (probabilistic) interpretation on it. In addition, this differential approach that deals with the wavefunction evolving forward at each time presents tremendous difficulties when an event at a different spacetime point appears to play a crucial part in determining the behaviour of the wavefunction at a specific spacetime point. A concrete example, which I discussed in the second half of Chapter 6, is when detectors are placed at the slits to track which of the two slits a photon has come through, and the presence of these detectors alters dramatically the behaviour of the wavefunction at a spatial point some distance away.

In view of these difficulties, it is appropriate to adopt a formulation of quantum mechanics that takes seriously the nature of the interdependence amongst events at different times. We have seen that such an approach exists and was introduced by Feynman as the Path-Integral approach to quantum mechanics. In this approach, the notion of a "path" enters at the most fundamental level. This is not a path in spacetime in the ordinary construal of the word, but is closely related to the latter as it corresponds to the possibility of a particle taking on specific values for certain dynamical variables at specific spacetime points. No special favoritism is shown at the outset as to which one of all the possible paths is to dominate. Rather, each of these possibilities contributes on an equal basis to the final eventuality. However, in order to take into account the full consequences of quantum phenomena, one must extract and take over from the wave description the periodic nature of these potentials. *Periodicity* is then introduced as the *phase* of each "path". As these paths or potentials operate, their relative phases may either reinforce or cancel each other to various degrees, in much the same way that real physical waves interfere. The final "dominating" potential is the product of the aggregation of the phases of each contributing potential or possibility. It is also through the notion of "paths" as alternatives that probabilistic considerations enter at the most fundamental level.

But the question of interest is whether there is a difference between the Feynman

[24]This alone appears to be a natural move as a "tightly-packed" wavepacket formed by the superposition of a large number of Fourier modes would give the localized property of a corpuscular entity.

approach and that of the wavefunction. I would say "yes from standpoint of process causation". The argument goes as follows. In wave mechanics, one speaks of the wavepacket as composed of a large number of Fourier modes. Because of the principle of superposition, a linear combination of *any* number of these Fourier modes would satisfy the Schrödinger equation. The larger the number of modes superposed, the more localized is the wavepacket. This localization of the wavepacket is then identified with the precision in locating the particle. However, despite its relatively localized position, a wavepacket is a "spread-out" entity and its "fuzziness" depends largely upon the number of Fourier modes that are in superposition. Because of the uncertainty, one may not speak about the trajectory of a particle in a straightforward sense. As a result, this poses problems on a fundamental level for engaging in causal talk in the quantum domain .

Feynman's Path-Integral approach, on the other hand, has a distinct advantage over the Schrödinger wavefunction approach. In this picture, the classical character of the localization of a particle is very much retained. And this particle would then be subjected to the action of probabilistic potentials as analogous to the case of forces acting on particles that we find ourselves so familiar with in classical physics. This enables us to adopt a kind of propensity view like the one expounded by Popper and to speak of probability on a singular level. The concerted effort of the probabilistic potentials acts on every single particle. Classical determinism consists in the fact that our acting potential, the mechanical force, produces the "same" single effect on a particle. In contrast, the aggregate action of the quantum potentials produces not one but a multitude of effects on an ensemble of identically prepared particles. In spite of the probabilistic complications, a photon emitted from a source is *sent through* the two-slits apparatus (and gets registered on the screen) by these quantum probabilistic potentials. Arguably, a formulation with these potentials placed at its core forms a natural foundation for causation on the quantum level.

There are causal processes à la **HCQT** in quantum mechanics. These are the probability-generating potentials - as probability amplitudes obtained by the "sum-over-all-paths" procedures in the Feynman Path Integral approach. They provide an underlying mechanism for the production of quantum phenomena; they propel the system through a multitude of possibilities. Because of the all encompassing aspect of the system and its environment, these causal processes in quantum mechanics amount to more than just the spatiotemporal trajectory of the system itself.

The entanglement of probabilistic potentials imposes upon us the view of these histories as "holistic" entities, rather than paying attention to the individual members of the sequence. In this way, Salmon's intention for causal processes as "ropes of causation" is fulfilled more profoundly on the quantum level.

However, there remain unfinished businesses. It is natural for one to proceed to elucidate the notion of interaction in the quantum context via the Path-Integral approach to the subject of quantum field theory.

There exist recent approaches[25] in theoretical physics that are founded on the notion of "histories" and it would be of great interest to analyze how the ideas behind these approaches bear on process causation.

Finally, the very structure of spacetime with respect to its continuous nature has deep implications for the notion of causal process and causation and it would indeed be a holy grail of both philosophy and physics to understand the nature of spacetime and its relation to causation.

[25] I have in mind, in particular, the so-called "Consistent Histories" or "Decoherent Histories" approaches of Omnes, R. (1994), Griffiths, R.B. (1984) and Gell-Mann, M. and Hartle, J.B. (1990). For a shorter and user-friendly introduction, see Gell-Mann, M. (1994). A detailed survey can be found in Omnes, R. (1992).

Bibliography

[1] Anscombe, G.E.M. (1974), "Times, Beginnings and Causes" in the *Proceedings of the British Academy*. Reprinted in G.E.M. Anscombe, *Metaphysics and the Philosophy of Mind*, Collected Philosophical Papers Volume II, Basil Blackwell, Oxford, p.148-162.

[2] Aspect, A., Grangier, P. and Roger, G. (1982), "Experimental Realization of Einstein-Podolsky-Rosen-Bohm *Gedankenexperiment*: A New Violation of Bell's Inequalities", *Physical Review Letters*, 49(2), p.91-94.

[3] Asquith, P.D. and Giere, R.N. (eds.) (1980), *Proceedings of the 1980 Biennial Meeting of the Philosophy of Science Association*, Volume 2, Philosophy of Science Association.

[4] Baldwin, J.M.(ed.), (1902), *Dictionary of Philosophy and Psychology*, Volumes I and II, Macmillan and Co. Limited.

[5] Bastin, T. (ed.) (1971), *Quantum Theory and Beyond: Essays and Discussions from a Colloquium*, Cambridge University Press.

[6] Bell, J.S. (1964), "On the Einstein Podolsky Rosen Paradox", *Physics*, 1, p.195-200. Reprinted in Wheeler and Zurek (eds.) (1983), p.403-408.

[7] Bohr, N. (1928), "The Quantum Postulate and the Recent Development of Atomic Theory", *Nature*, 121, p.580-590. Reprinted in Wheeler and Zurek (eds.) (1983), p.87-126.

[8] Bohr, N. (1949), "Discussion with Einstein on Epistemological Problems in Atomic Physics" in Schilipp (Ed.) (1949), p.200-241. Reprinted in Wheeler and Zurek (eds.) (1983), p.9-49.

[9] Bombelli, L., Lee, J., Meyer, D. and Sorkin, R.D. (1987), "Spacetime as a Causal Set", *Physical Review Letters*, 59, p.521.

[10] Born, M. and Wolf, E. (1980), *Principles of Optics*, sixth edition, Pergammon Press.

[11] Brenner, M.P. and Stone, H.A. (2000), "Modern Classical Physics Through the Work of G.I. Taylor", *Physics Today*, 35(5), p.30-35.

[12] Cassirer, E. (1956), *Determinism and Indetreminism in Modern Physics - Historical and Systematic Studies of the Problem of Causality*, Translated by Benfry, O.T., Yale University Press.

[13] Cushing, J.T. (1994), *Quantum Mechanics - Historical Contingency and the Copenhagen Hegemony*, The University of Chicago Press.

[14] Dieks, D. (1981), "A Note on Causation and the Flow of Energy", Erkenntnis, 16, p.103-108.

[15] Dirac, P.A.M. (1933), "The Lagrangian in Quantum Mechanics", *Physikalische Zeitschrift der Sowjetunion*, Band 3, Heft 1, p.64-72. Reprinted in Schwinger (ed.), (1958), p.312-320.

[16] Dirac, P.A.M. (1958), *The Principle of Quantum Mechanics*, Cambridge University Press.

[17] Dowe, P. (1992a), "Wesley Salmon's Process Theory of Causality and the Conserved Quantity Theory", *Philosophy of Science*, 59, p.195-216.

[18] Dowe, P. (1992b), "An Empiricist Defence of the Causal Account of Explanation", *International Studies in the Philosophy of Science*, 6(2), p.123-128.

[19] Dowe, P. (1995), "Causality and Conserved Quantities: A Reply to Salmon", *Philosophy of Science*, 62, p.321-333.

[20] Dowe, P. (1999), "Good Connections: Causation and Causal Processes", in *Causation and Laws of Nature*, Sankey, H. (ed.) (1999), p.247-263).

[21] Dowe, P. (2000), *Physical Causation*, Cambridge Studies in Probability, Induction and Decision Theory, Cambridge University Press.

[22] Ducasse, C.J. (1926), "On the Nature of the Observability of the Causal Relation", *Journal of Philosophy*, 23, p.57-68. Reprinted in Ernest and Sosa (eds.) (1993), p.125-136.

[23] Eddington, A. (1920), *Space, Time and Gravitation - An Outline of the General Relativity Theory*, Cambridge University Press.

[24] Ehring, D. (1986), "The Transference Theory of Causation", *Synthese*, 67, p.249-258.

[25] Fair, D. (1979), "Causation and the Flow of Energy", *Erkenntnis*, 14, p.219-250.

[26] Fetzer, J.H. (1987), "Critical Notice: Wesley Salmon's *Scientific Explanation and the Causal Structure of the World*", *Philosophy of Science*, 54, p.597-610.

[27] Feynman, R.P. (1942), *On the Principle of Least Action in Quantum Mechanics*, Ph.D. Thesis, Princeton University, Princeton, NJ, Ann Arbor, University Microfilms Publication no. 2948.

[28] Feynman, R.P. (1948), "Spacetime Approach to Non-Relativistic Quantum Mechanics", *Physical Review*, 76, p.367-387. Reprinted in Schwinger (ed.) (1958), p.321-341.

[29] Feynman, R.P. (1963a), *Feynman's Lectures on Physics*, Volume I.

[30] Feynman, R.P. (1963b), *Feynman's Lectures on Physics*, Volume II.

[31] Feynman, R.P. (1966), "The Development of the Space-Time View of Quantum Electrodynamics", Nobel Foundation. Reprinted in *Physics Today*, 19, August 1966, p.31-44.

[32] Feynman, R.P. and Hibbs, A.R. (1965), *Quantum Mechanics and Path Integrals*, McGraw-Hill, Inc.

[33] Forge, J. (1982), "Physical Explanation: With Reference to the Theories of Scientific Explanation of Hempel and Salmon", in McLaughlin, R. (ed.) (1982), p.211-229.

[34] French, A.P. (1968), *Special Relativity*, MIT Introductory Physics Series, MIT Press.

[35] French, P.A., Uehling, T.E. and Wettstein, H.K. (eds.) (1984), *Causation and Causal Theories*, Midewest Studies in Philosophy, Volume 9, Minneapolis: Univeristy of Minnesota Press.

[36] Garrison, J.C., Mitchell, M.W., Chiao, R.Y. and Bolda, E.L. (1998), "Superluminal Signals: Causal Loop Paradoxes Revisited", quant-ph/9810031, e-print archive at http://xxx.lanl.gov.

[37] Gell-Mann, M. and Hartle, J.B. (1990), "Quantum Mechanics in the Light of Quantum Cosmology", in Zurek, W.H. (ed.), (1990).

[38] Gell-Mann, M. (1994), *The Quark and the Jaguar: Adventures in the Simple and the Complex*, Little Brown, London.

[39] Giere, R.N. (1988), "Scientific Explanation and teh Causal Structure of the World by Wesley Salmon", *Philosophical Review*, 97, p.444-446.

[40] Griffiths, R.B. (1984), "Consistent Histories and the Interpretation of Quantum Mechanics", *Journal of Statistical Physics*, 36, p.219-272.

[41] Hasegawa, A. (1992), *Optical Solitons in Fibres*, Springer-Verlag, Berlin.

[42] Heisenberg, W. (1927), "über denschaulichen Inhalt der quantentheoretischen Kinematik and Mechanik", Seitschrift für Physik, 43, p.172-198. Translated into English as "The Physical Content of Quantum Kinematics and Mechanics" by Wheeler, J.A. and Zurek, H.W. (1981) and reprinted in Wheeler and Zurek (eds.) (1983), p.62-84.

[43] Hempel, C. (1965), *Aspects of Scientific Explanation*, New York, Free Press.

[44] Hirsch, E. (1982), *The Concept of Identity*, Oxford University Press.

[45] Hitchcock, C.R. (1995), "Discussion: Salmon on Explanatory Relevance", *Philosophy of Science*, 62, p.304-320.

[46] Hume, D. (1888), *A Treatise of Human Nature*, second edition (1978), with revised text and notes by P.H. Nidditch, Clarendon Press, Oxford.

[47] Isham, C.J. (1995), *Lectures on Quantum Theory: Mathematical and Structural Foundations*, Imperial College Press.

[48] Jackson, F. (ed.) (1991), *Conditionals*, Oxford University Press.

[49] Kitcher, P. (1989), "Explanatory Unification and the Causal Structure of the World", in Salmon, W.C. and Kitcher, P. (eds.) (1989), p.410-505.

[50] Kline, A.D. (1985), "Humean Causation and the Necessity of Temporal Discontinuity", *Mind*, 94(376), p.550-556.

[51] Kolmogorov, A.N. (1950), *Foundations of the Theory of Probability*, English translation by Morrison, N., Chelsea Publishing Company, New York.

[52] Kopczynski, W. and Trautman, A. (1992), *Spacetime and Gravitation*, John Wiley and Sons.

[53] Lagrange, J-L. (1997), *Analytical Mechanics*. Translated from *Mecanique Analytique*, nouvelle édition of 1811 and edited by Boissonnade, A. and Vagliente, V.N., *Boston Studies in the Philosophy of Science*, Volume 191, Dordrecht London, Kluwer Academic.

[54] Laudauer, R. (1993), "Light Faster than Light?", *Nature*, 365, p.692.

[55] Lewis, D. (1973), *Counterfactuals*, Blackwell, Oxford.

[56] Lorentz, H.A., Einstein A., Minkowski, H. and Weyl H. (1923), *The Principle of Relativity: A Collection of Original Memoirs on the Special and General Theory of Relativity* (with notes by A. Sommerfeld). Translated by Perrett, W. and Jeffrey, G.B., Methuen and Co. Ltd., London.

[57] Ludwig, G. (ed.) (1968), *Wave Mechanics*, Selected Readings in Physics, Pergamon Press Ltd.

[58] Ma, C.K.W. (2000), "Causation and Physics", *Fortschritte Der Physik*, 48(5-7), p.631-641.

[59] McLaughlin, R. (ed.) (1982), *What? Where? When? Why? - Essays on Induction, Space and Time, Explanation*, D. Reidel Publishing Company.

[60] Mermin, N.D. (1968), *Space and Time in Special Relativity*, McGraw-Hill, Inc.

[61] Michelson, A.A. (1927), *Studies in Optics*, Pheonix Science Series, University of Chicago Press, (1968 reprint).

[62] Minkowski, H. (1908), "Space and Time" (A translation of an address delivered at the 80th Assembly of German Natural Scientists and Physicians, at Cologne, 21 September 1908). Reprinted in Lorentz, H.A. et al. (1923), p.75-91.

[63] Newton, R.E.I. (1990), *Waves Physics*, Edward Arnold.

[64] Omnès, R. (1992), "Consistent Interpretation of Quantum Mechanics", *Reviews of Modern Physics*, 64(2), p.339-382.

[65] Omnès, R. (1994), *The Interpretation of Quantum Mechanics*, Princeton Series in Physics, Princeton University Press.

[66] Pauli, W. (1980), *General Principles of Quantum Mechanics*, translated by Achuthan, P. and Venkateson, K. from the article "Die allgemeinen Prinzipien der Wellenmechanik", which appeared in Volume V, Part I of Handbuch der Physik (1958), Flügge S. (ed.).

[67] Penrose, R. (1971), "Angular Momentum: an Approach to Combinatorial Spacetime", in Bastin, T. (ed.) (1971), p.151-180.

[68] Popper, K.R. (1990a), "A World of Propensities: Two Views of Causality" in Popper (1990b), p.1-25.

[69] Popper, K.R. (1990b), *A World of Propensities*, Thoemmes Antiquarian Books Ltd.

[70] Reichenbach, H. (1956), *The Direction of Time*, University of California Press.

[71] Renyi, A. (1970), *Foundations of Probability*, Holden-Day Inc.

[72] Rescher, N. (1968), *Studies in Logical Theory*, Oxford, Blackwell.

[73] Robb, A. (1936), *A Theory of Time and Space*, second edition, Cambridge University Press.

[74] Rogers, B. (1981), "Probabilistic Causality, Explanation, and Detection", *Synthese*, 48, p.201-223.

[75] Rovelli, C. and Smolin, L. (1995), "Spin Networks and Quantum Gravity", *Physical Review D*, 52, p.5743-5759.

[76] Russell, B. (1903), *The Principles of Mathematics*, Cambridge University Press.

[77] Russell, B. (1914), "On the Notion of Cause, with Applications to the Free-Will Problem" in *Our Knowledge of the External World*, Open Court Publishing Company. 1993 edition by Routledge Inc., p.214-246.

[78] Russell, B. (1917), "On the Notion of Cause" in *Mysticism and Logic*, George Allen and Urwin, p.180-208.

[79] Russell, B. (1948), *Human Knowledge: Its Scope and Limits*, New York, Simon and Schuster. 1992 edition by Routledge Inc.

[80] Salmon, W.C. (1977), "An "At-At" Theory of Causal Influence", *Philosophy of Science*, 44, p.215-224.

[81] Salmon, W.C. (1980a), "Causality: Production and Propagation", in Asquith, P.D. and Giere, R.N. (eds.) (1980), p.49-69. Reprinted in Sosa, E. and Tooley, M. (eds.) (1993), p.154-171.

[82] Salmon, W.C. (1980b), "Probabilistic Causality", *Pacific Philosophical Quarterly*, 61, p.50-74. Reprinted in Sosa, E. and Tooley, M. (eds.) (1993), p.137-153 and as a newly extended version with additional footnotes in Salmon (1998), p.208-232.

[83] Salmon, W.C. (1984), *Scientific Explanation and Causal Structure of the World*, Princeton Univeristy Press.

[84] Salmon, W.C. and Kitcher, P. (eds.) (1989), *Minnesota Studies in Philosophy of Science*, Volume 13. Minneapolis: University of Minnesota Press.

[85] Salmon, W.C. (1994), "Causality Without Counterfactuals", *Philosophy of Science*, 61, p.297-312.

[86] Salmon, W.C. (1997), "Causality and Explanation: A Reply to Two Critics", *Philsophy of Science*, 64, p.461-477.

[87] Salmon, W.C. (1998a), *Causality and Explanation*, Oxford University Press.

[88] Salmon, W.C. (1998b), "Indterminacy, Indeterminism and Quantum Mechanics", in Salmon, W.C. (1998a), p.261-281.

[89] Sankey, H. (ed.) (1999), *Causation and Laws of Nature*, Australasian Studies in History and Philosophy of Science Volume 14, Kluwer Academic Publishers.

[90] Schilipp, P.A. (1949), *Albert Einstein: Philosopher-Scientist*, Open Court Publishing Inc.

[91] Schrödinger, E. (1926), "Quantisierung als Eigenwertproblem", *Annalen der Physik*, 79, p.361-376.

[92] Schrödinger, E. (1928), "Quantization as a Problem of Proper Values (First Communication)", *Collected Papers on Wave Mechanics*, Blackie and Son, London.

[93] Schwinger, J. (ed.) (1958), *Selected Papers on Quantum Electrodynamics*, Dover Publications Inc.

[94] Scully, M.O. and Drühl, K. (1982), "Quantum Eraser: a Proposed Photon Correlation Experiment on Observation and "Delayed Choice" in Quantum Mechanics", *Physical Review A*, 25, p.2208.

[95] Shih, Y.H. and Alley, C.O. (1988), "New Type of Einstein-Podolsky-Rosen-Bohm Experiment Using Pairs of Light Quanta Produced by Optical parametric Down Conversion", *Physical Review Letters*, 61(26), p.2921-2924.

[96] Shoemaker, S. (1979), "Identity, Properties and Causality", in French P.A. et al. (eds.) (1984), p.326.

[97] Skyrms, B. (1981), *Causal Necessity: A Pragmatic Investigation of the Necessity of Laws*, New Haven, yale University Press.

[98] Sober, E. (1987), "Review Article: Explanation and Causation", *British Journal for the Philosophy of Science*, 38, p.243-257.

[99] Sosa, E. and Tooley, M. (eds.) (1993), *Causation*, Oxford Readings in Philosophy, Oxford University Press.

[100] Stalnaker, R.C. (1968), "A Theory of Conditionals", in Rescher, N. (ed.) (1968). Reprinted in jackson, F. (ed.) (1991) , p.28-45.

[101] Synge, J.L. (1972), *Relativity: The Special Theory*, North-Holland.

[102] Taylor, E.F. and Wheeler, J.A. (1966), *Spacetime Physics*, W.H. Freeman and Company.

[103] Taylor, E.F. and Wheeler, J.A. (1992), *Spacetime Physics*, second edition, W.H. Freeman and Company.

[104] Taylor, G.I., (1909), "Interference Fringes with Feeble Light", *Proceedings of the Cambridge Philosophical Society*, Vol.XX, p.114-115.

[105] Von Mises, R. (1963), *Mathematical Theory of Probability and Statistic*, Academic Press.

[106] Wheeler, J.A. and Zurek, W.H. (eds.) (1983), *Quantum Theory and Measurement*, Princeton University Press.

[107] Woodward, J. (1989), "The Causal Mechanical Model of Explanation", in Salmon, W.C. and Kitcher, P. (eds.) (1989), p.357-381.

[108] Zurek, W.H. (ed.) (1990), *Complexity, Entropy and the Physics of Information*, Santa Fe Institute Studies in teh Science of Complexity No. 8, Addison-Wesley, Redwood City, California.