Talk presented at the conference on "Causality and Physical Theories", Wayne State University, Detroit, Michigan, May 12-13, 1973.

LBL-2031 c.1

MACROCAUSALITY AND ITS ROLE IN PHYSICAL THEORIES

Henry P. Stapp

May 29, 1973

Prepared for the U. S. Atomic Energy Commission
under Contract W-7405-ENG-48

MACROCAUSALITY AND ITS ROLE IN PHYSICAL THEORIES*†

Henry P. Stapp

Lawrence Berkeley Laboratory
University of California
Berkeley, California 94720

May 29, 1973

ABSTRACT

The physical meaning of the macrocausality property of scattering transition probabilities is described, and the role of this property in S-matrix theory and other physical theories is discussed. The macroscopic causality properties of theories with shadow particles, are examined and are shown to contradict the general interpretational principles of quantum theory. Shadow particles have been introduced to remedy the unitarity difficulties of indefinite-metric field theories.

---

* This work was supported by the U. S. Atomic Energy Commission.

† Talk delivered at the conference on "Causality and Physical Theories," held at Wayne State University, Detroit, Michigan, May 12-13, 1973. To be published in the American Institute of Physics Conference Proceedings.

I. INTRODUCTION

Experience has causal properties, and these should be reflected in physical theory. However, one cannot simply deduce general theoretical causality properties directly from experiment, for experiments are neither infinitely precise nor infinitely extensive. Experiment can merely suggest possibilities, and rule out others.

The form that a theoretical causality property takes will depend on the theoretical structure in which it is imbedded. In fact, a given theoretical structure often suggests a natural causality property. For example, in quantum field theory the natural causality property is that fields at space-like-separated points commute:

$$A(x)\,A(y) = A(y)\,A(x) \qquad \text{for} \qquad (x-y)^2 < 0 . \qquad (1)$$

This commutator causality requirement appears to lead to mathematically inconsistencies, and the suggestion is often made that it may be too stringent. For it imposes precise conditions at infinitely small distances, and hence goes far beyond what experience tells us.

This lack of close connection between the commutator causality property and experiment is due in part to the lack of any close connection between the field operators of quantum field theory and experimental observables. This latter deficiency is an objectionable feature of quantum field theory. For a basic precept of quantum theory, at least at the nonrelativistic level, where the mathematical inconsistencies do not arise, is that the basic operators of the theory correspond directly to experimental observables. The logical structure of quantum theory and its connection to experience was built on this premise.

To bring relativistic quantum theory into accord with this precept Heisenberg devised S-matrix theory. This theory conforms to the basic precepts of both quantum theory and relativity theory, and it does not encounter the mathematical difficulties associated with the commutator form of the causality condition.

S-matrix theory has no observables corresponding to space-time points or to sharply defined space-time regions. Thus it might seem that S-matrix theory would have no natural causality property. This is not the case: S-matrix theory has a natural causality property, called macrocausality, which in fact plays an important role in the logical and mathematical structure of the theory.

In this talk I shall first describe the physical content of the macrocausality property. This property blends a certain intuitive idea of causality with a specific dynamical assumption. Then I shall discuss the role of macrocausality in S-matrix theory and other physical theories. Finally, I shall apply these considerations to the problem of causality in theories with shadow states.

My subject is narrower and more technical than those of most of the earlier talks. And my presentation is aimed partly at physicists who wish to understand the S-matrix causality concept. However, I shall discuss here only the physical ideas, not the mathematical details,[1] and thus hope to reach also those in the audience whose interests are mainly philosophical. Philosophers should find it useful to have a clear understanding of causality property that is more elaborate than certain traditional ones, and to see how this causality property is actually used in contemporary physical theory.

II. MACROCAUSALITY

A. General Remarks

Macrocausality deals only with those observables that occur in S-matrix theory, namely with scattering transition probabilities. These quantities can be measured to high accuracy by means of experimental arrangements of a kind that physicists actually can and do set up. This does not mean, however, that macrocausality can be derived from experiment. For macrocausality is a general property, whereas tests cover only special cases. Moreover, macrocausality refers to asymptotic distances whereas only finite distances are experimentally accessible.

Macrocausality cannot be derived from microcausality. These two causality properties are complementary. Macrocausality deals with arbitrarily large distances, whereas microcausality deals with infinitely small distances. Moreover, as will be discussed, macrocausality is equivalent to a set of analytic properties in the physical region itself, whereas microcausality implies analytic properties only outside the physical region. Thus neither one implies the other.

Macrocausality formalizes a certain physical idea, which is called the physical idea of macrocausality. This physical idea is discussed next.

B. The Physical Idea

The physical idea of macrocausality is that interactions are transmitted over <u>macroscopic</u> distances only by physical objects. This idea is a macroscopic version of the primitive idea that the world consists only of physical objects, and that these objects act on each other only by direct contact. Two examples will illustrate the main points.

Example I. A baseball is hit into a window. In this example we can identify the following features:

(a) Cause: The baseball is hit.

(b) Effect: The window breaks.

(c) Link: The baseball travels from the bat to the window. That is, a physical object travels from the space-time region of the cause to the space-time region of the effect. This is illustrated in Fig. 1.

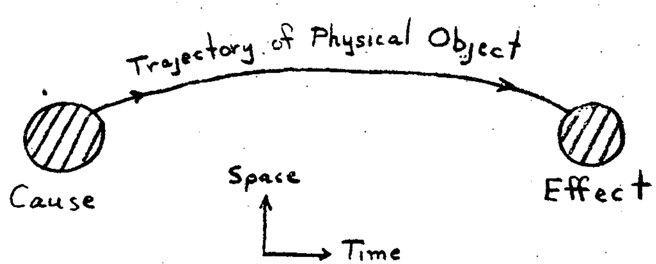

Fig. 1. A physical object travels from the space-time of the cause to the space-time region of the effect.

Example II. A set of billiards balls move about under the influence of their mutual collisions. In this case physical objects travel between the space-time collision regions. This is illustrated in Fig. 2.

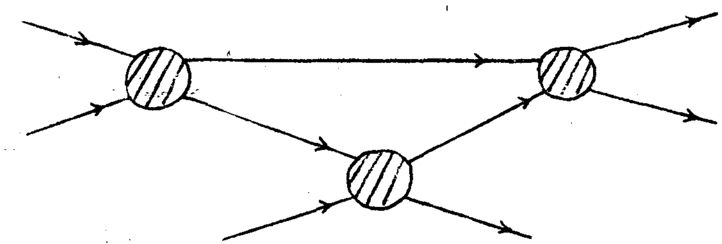

Fig. 2. Physical objects travel between the space-time collision regions. Each space-time trajectory represents the path of the center-of-mass of a physical object.

In these examples a distinction is drawn between long-range interactions and short-range interactions. The long-range interactions are those that are transmitted from one space-time region to a far-away space-time region by a physical object. These interactions fall off (in a statistical sense) at large distances only by the geometric factor associated with beam spreading. The remaining interactions are those associated with the exchanges of momentum-energy that occur when the physical objects collide. These latter interactions are associated, in various theoretical models, with potentials, or virtual-particle exchange, or unstable-particle exchange, or nonlocal interactions, or even with a breakdown of the concept of space-time at small distances.

The physical idea of macrocausality is that these remaining interactions are short range. That is, the longest-range interactions are those carried by physical objects, and hence all interactions not carried by physical objects fall off faster at large separation than those carried by physical objects.

To make this idea well defined one must identify the interactions carried by physical objects. This is done by invoking two basic ideas of relativistic mechanics.

(a) Physical Objects: Each physical object has a mass m, and the momentum energy p carried by an object equals the product of its mass with its covariant velocity v: $p = mv$.

(b) Conservation of Momentum-Energy: The momentum-energy carried into any collision equals that carried out.

These two principles, together with the requirement that the remaining interactions have short range, determine the gross features of billiard ball dynamics (see Fig. 3).

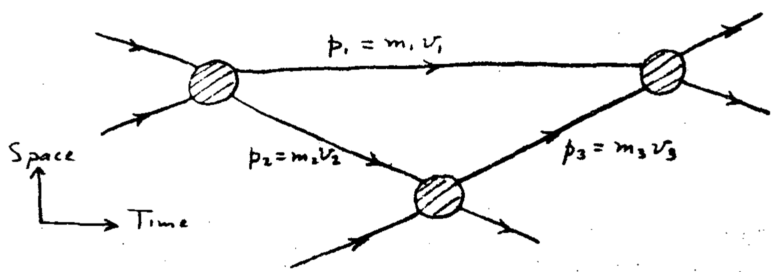

Fig. 3. A necessary condition for the reaction to occur is that the space-time collision regions can be connected by the space-time trajectories of physical objects. The momentum-energy p carried by each object must be directed along its space-time velocity v, and the proportionality factor must be the mass of that object. Momentum-energy must be conserved at each collision.

The finer details of the dynamics will depend on the precise form of the short-range interactions. However, uncertainties associated with short-range interactions can be effectively damped out by moving the physical objects farther apart.

This idea can be made precise by considering a set of scattering processes that are related to each other by space-time dilation. This dilation of the physics can be described by introducing a "scaled" coordinate system. The scaled coordinates x' are defined by $x = x'\tau$, where x represents the physical space-time coordinate, and τ is a scale parameter that tends to infinity. If one fixes the space-time trajectories in x' space then the physical objects corresponding to these trajectories are moved apart as τ tends to infinity, unless the trajectories intersect.

Any finite distance Δx shrinks to a point in x' space, as $\tau \to \infty$. Hence the x'-space image of any (finite-radius) physical object shrinks to a point. And the x'-space image of any finite-radius interaction-region shrinks to a point. Thus if all interactions not carried by physical objects had finite radius then the necessary conditions for a reaction with specified initial and final trajectories in x' space to occur for arbitrarily large τ would be this: the trajectories of the initial and final particles would have to coincide with the initial and final trajectories of a "causal network." These networks are defined in and below Fig. 4.

The quantum mechanical transition probability formula can be cast into classical form.[2] The function $w(p,x)$ is defined by a relativistic generalization of Wigner's formula:

$$w(p,x) \equiv \int \psi^*(Mv - \tfrac{1}{2} q)\, \psi(Mv + \tfrac{1}{2} q)\, e^{-iqx} (M/m)^{1/2}\, 2\pi\, \delta(q\cdot v)\, \frac{d^4q}{(2\pi)^4}$$

where

$$M = (m^2 - \tfrac{1}{4} q^2)^{1/2}$$

and

$$v = p/m \,.$$

The function $S[\{p_j, x_j\}]$ is defined in a similar way:

$$S[\{p_j,x_j\}] = \int \prod_{j=1}^{n} \left[\frac{d^4q_j}{(2\pi)^4}\, 2\pi\, \delta(q_j\cdot v_j)\, e^{-iq_j\cdot x_j}\, (M_j/m_j)^{1/2} \right]$$

$$\times\, S(\{M_j v_j \mp \tfrac{1}{2} q\})\;\; S^*(\{M_j v_j \pm \tfrac{1}{2} q\}) \,.$$

Here the upper sign is to be used for initial particle variables, and the lower sign is to be used for final particle variables, and $S(\{p_j\})$ is the usual S matrix.

C. Quantum Formulation

The physical idea of macrocausality is expressed in terms of the concepts of classical physics. From this idea one can derive some very general properties of the classical scattering transition probabilities. The quantum theoretical macrocausality property is the statement that these general properties, which follow directly from the (classical) physical idea of macrocausality, are enjoyed by the scattering transition probabilities of quantum theory.

These general properties are of the following kind: they assert that under specified conditions on the initial and final wave functions of the scattering process the scattering transition probability falls off at least exponentially as $\tau \to \infty$, due to the assumed exponential fall off of all interactions that are not carried by physical objects. For under the specified conditions the scattering process can occur only if there is at least one transfer of momentum-energy that cannot be carried by any physical object, yet must carry over a distance that increases linearly with τ. Under these conditions the exponential fall off of the scattering transition probability follows directly from the physical idea of macrocausality.

These considerations can be made quantitive by considering semi-classical models. In these models one allows momentum-energy to be transferred between particles by various possible mechanisms (see Fig. 5).

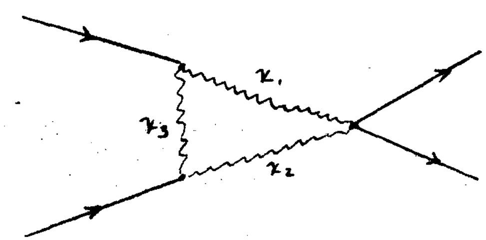

Fig. 5. Momentum-energy can be transferred between physical particles (solid lines) by various possible mechanisms (wiggly lines).

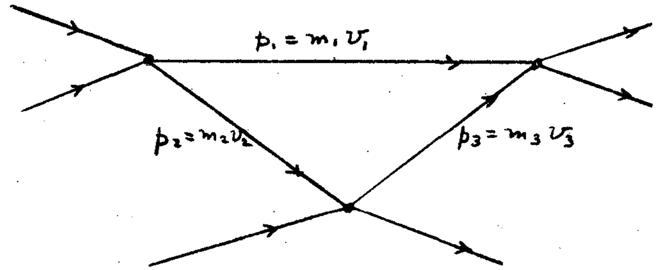

Fig. 4. A typical causal network. A causal network represents the necessary condition for a classical reaction to occur if the physical objects are point particles that interact only via point interactions.

A causal network depicts the space-time flow of conserved momentum-energy from initial particles to final particles via a network of intermediate physical particles. The momentum energy p_j carried by each particle is related to its space-time velocity v_j by $p_j = m_j v_j$. Momentum-energy is transferred between particles only at points where their trajectories intersect.

The assumption that all interactions not carried by physical objects have a finite radius is unrealistic and unnecessary. However, some assumption about the way in which those interactions fall off is needed to give precise content to the macrocausality property.

The dynamical assumption is now introduced. It is assumed that all interactions not carried by physical objects fall off at least exponentially under space-time dilation. This dynamical

assumption is analogous to the assumption that the potentials of non-relativistic theory have Yukawa-type tails.

From this exponential fall-off property one can derive analyticity properties. Weaker fall-off properties yield weaker conclusions. For example, power-law fall-off properties yield continuity properties. However, in what follows the exponential fall off is assumed.

So far the discussion has been purely classical. To pave the way to quantum theory it is useful to exhibit the classical form of the scattering transition probability formula. To do this each initial and final particle j is replaced by a statistical ensemble. This ensemble is represented by a classical probability function $w_j(p,x) \equiv w_j(\vec{p}, p^o, \vec{x}, t)$, defined by

$$\int_{\substack{\Delta \vec{p} \\ \Delta \vec{x}}} \frac{d^3\vec{p}\; d^3\vec{x}}{(2\pi)^3}\, w_j(p,x) =$$ The probability that a particle from the ensemble corresponding to particle j satisfies $(\vec{x},\vec{p}) \in (\Delta \vec{x}, \Delta \vec{p})$ at time t.

The particles in these ensembles are free. Thus the energy p^o is fixed by the mass-shell constraint. Moreover, the values of $w(p,x)$ at any one time t determines its value for all times.

The classical transition probability formula is then

$$P[\{w_j(p_j,x_j)\}] = \int \left[\prod_j \frac{d^3\vec{p}_j\; d^3\vec{x}_j}{(2\pi)^3}\, w_j(p_j,x_j) \right] S[\{p_j,x_j\}]$$

where S is the transition probability kernel. In this formula the times t_j can be chosen arbitrarily.

shall discuss the consequences of this conflict with macrocausality toward the end of my talk.

From the normal analytic structure plus unitarity one may derive all physical-region discontinuities. These discontinuities are the differences between the two different continuations of the scattering function around a physical-region singularity (see Fig. 9).

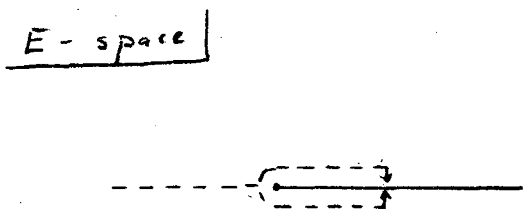

Fig. 9. The discontinuity is the difference between the functions obtained by continuing the scattering function in the two possible ways around a singularity.

Cutkosky obtained formulas for such discontinuities from perturbation theory. However, his formulas were not well defined, and his arguments were inadequate. Also, they depended on the validity of perturbation theory. Since these discontinuity formulas play a basic role in S-matrix theory--discontinuities are the S-matrix analogs of the potentials of nonrelativistic theory--it is important, from the point of view of internal cohesion, that formulas for them should be derivable from S-matrix principles.

The simplest and most important discontinuity formula is known as the pole-factorization theorem. Its simplest case is represented in Fig. 10.

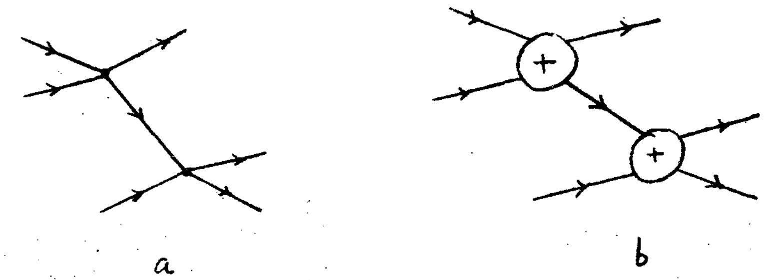

Fig. 10. The simplest case of the pole-factorization theorem. Figures a and b represent, respectively, the Landau diagram (or causal network) and the expression for the corresponding discontinuity. This discontinuity is simply the product of the two corresponding scattering amplitudes, integrated over the intermediate-particle momentum.

B. Check on Causality Properties

Macrocausality not only implies the normal analytic structure. It is also implied by it. This means that one can check the causal properties of a proposed theory by examining its physical-region analyticity properties: If the theory has the normal analytic structure then it has the macrocausality property. But if the theory has the macrocausality property then all long-range interactions are carried by physical particles. Thus the theory possesses all the general causal features that it needs to conform to ordinary macroscopic experience about causality. Any further causality requirement places conditions on the short-range structure of the theory, and hence extends causality ideas derived from macroscopic experience into realms where empirical support may be lacking.

It is interesting to compare the physical consequences of macrocausality and microcausality. This can be done by considering first the analytic properties implied by these two causality properties.

The analyticity properties implied by macrocausality are very different from those implied by microcausality. Macrocausality gives analyticity only at physical points (and hence of course in finite, but perhaps very small, neighborhoods of these real points) whereas microcausality gives analyticity only away from the physical points. By counter example it can be shown that microcausality (plus spectral conditions) can never yield analyticity in the physical region itself. Indeed, the primitive domain of analyticity in field theory includes no mass-shell points at all, either inside the physical region or outside it. However, this primitive domain can be extended by methods of analytic completion into mass-shell domains that contain physical-region points on their boundaries. The situation is schematically indicated in Fig. 11.

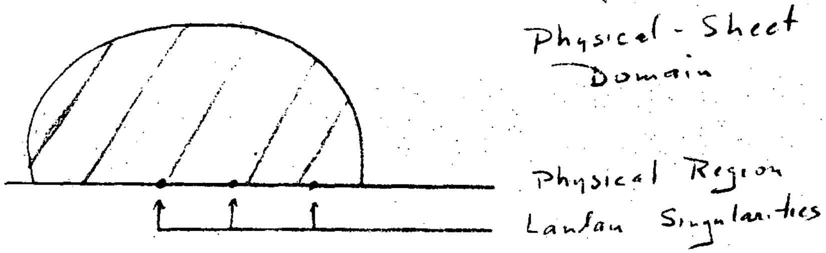

Fig. 11. Macrocausality gives analyticity in the (real) physical region, except at Landau singularities. Microcausality gives analyticity in some physical-sheet domain that contains physical-region points on its boundary.

C. Effects of Poles

To gain understanding of the physical significance of these different domains of analyticity it is useful to consider the effect on scattering transition probabilities of poles that lie in the different regions. Consider, for example, a $2 \to 2$ scattering process. Suppose, first, that the pole lies at the point $E = m - i\Gamma/2$ in the center-of-mass energy variable. And suppose this point is situated on the "unphysical sheet" reached by passing from the physical sheet through the physical region, as indicated in Fig. 12.

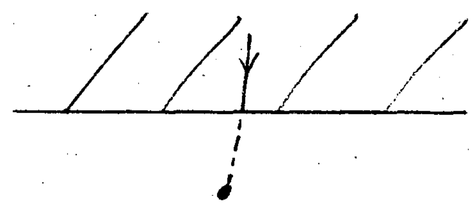

Fig. 12. A pole located at $E = m - i\Gamma/2$ on the unphysical sheet.

Suppose now that the two incoming beams intersect in a space-time region A, and that the two outgoing beams intersect in a space-time region B. (The outgoing beams are defined by the acceptance conditions of the devices that detect the outgoing particles.) Suppose A and B are both centered around the origin of space (not time) in some average center-of-mass frame, and that B is later than A by some average time t, as shown in Fig. 13.

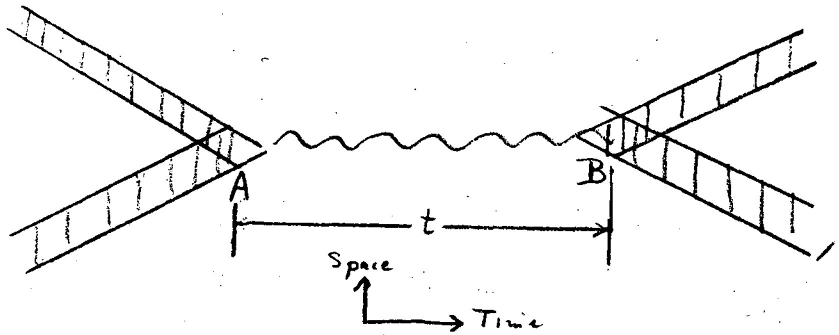

Fig. 13. The two incoming beams intersect at A, and the two outgoing beams intersect at B. The region B is later than A by the time t.

If the center-of-mass energy of the pair of incoming particles is centered around m, and the center-of-mass energy of the pair of outgoing particles is also centered around m, and if there are no other nearby singularities, then the scattering transition probability will have the behavior expected from the production and subsequent decay of an unstable particle of lifetime $1/\Gamma$. In particular, for positive t the transition probability will fall off like $\exp -\Gamma|t|$. (Omnes-type wave functions are used, with $t = \tau$.) For negative t, on the other hand, the fall off will be much faster, provided there are no other nearby singularities — on the scale of Γ. [The rate of fall off is determined by the nearness of the other singularities, and by the width of the gaussians in the Omnes-type wave functions.[3]]

Suppose, however, that the pole is situated at $E = m + i\Gamma/2$ in the physical sheet, as shown in Fig. 14. Then the situation is reversed: for large <u>negative</u> times t the scattering transition probability will have a term that falls off like $e^{-\Gamma|t|}$, whereas for

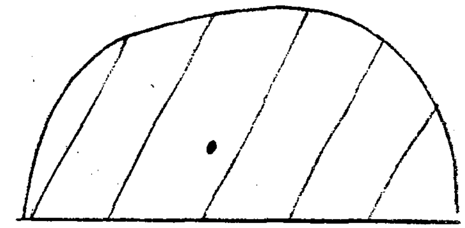

Fig. 14. Pole at $m + i\Gamma/2$ in the physical-sheet domain.

large positive times t it will fall off much faster. Thus in this case the scattering transition probability has the behavior that would correspond, not to an ordinary decaying particle, but rather to a particle that propagates <u>backward</u> in time with a decay factor $e^{-\Gamma|t|}$. Figure 15 shows the space-time configuration of the incoming and outgoing beams that would reveal this acausal effect of the pole at $m + i\Gamma/2$.

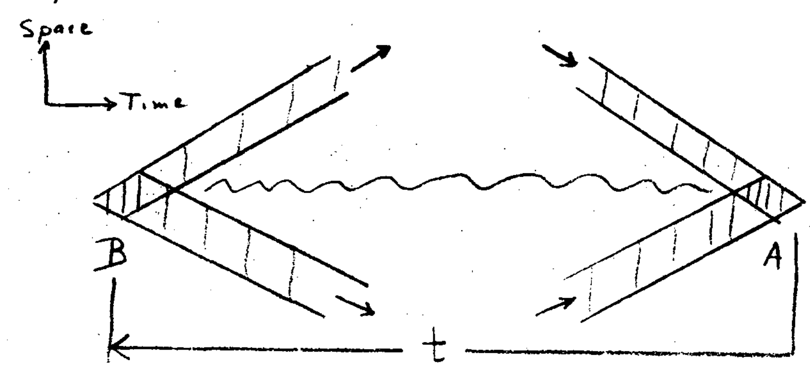

Fig. 15. The effect of the pole at $m + i\Gamma/2$. The scattering transition probability falls off like $\exp -\Gamma|t|$ for negative times, where negative times correspond to the outgoing particles being produced before the incoming particles have come together.

Microcausality allows the singularity at $m - i\Gamma/2$, which produces the causal behavior, but it forbids the singularity at $m + i\Gamma/2$.

If Γ is sufficiently small, and hence the lifetime $1/\Gamma$ is sufficiently long, then the acausal effects of this singularity should, in general, be observable. However, if Γ is large then these acausal would be hard to observe.

For a $2 \to 2$ reaction the pole cannot lie right in the physical region itself because of stability requirements. But if two external particles are added, in the manner shown in Fig. 16, then the

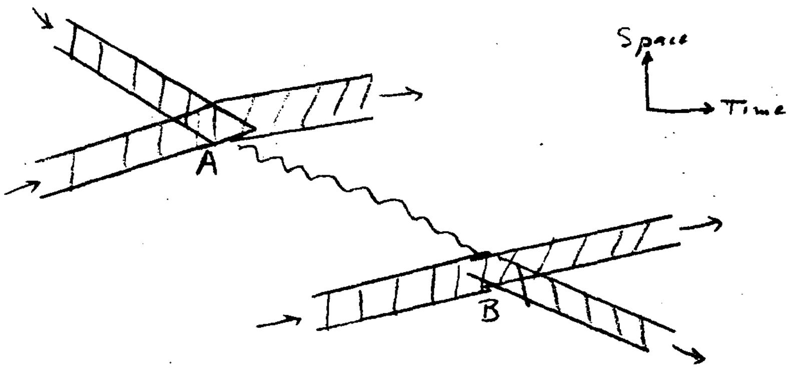

Fig. 16. Generalization of Fig. 13.

intermediate particle pole can lie in the physical region (i.e., $\Gamma \to 0$). In this case the exponential decrease factor turns into the geometrical factor corresponding to the classical spreading of the intermediate particle beams. In particular, a pole at $m - i\epsilon$ (where ϵ is infinitesimal) has an effect on the $3 \to 3$ scattering transition probability of precisely the kind that would be caused by a classical particle of mass m being produced at A and absorbed at B.

D. Theory of Measurement

These physical-region singularities, at points $m - i\epsilon$, and the pole-factorization theorem expressions for their discontinuities, play a crucial role in the theory of measurements. Bohr and Heisenberg emphasized that the consistency of quantum theory requires that the boundary between the quantum system and the (classically treated) world in which the quantum system is imbedded can in certain circumstances be shifted, so that what was originally part of the classically treated measuring device becomes part of the quantum system under consideration. This requirement was studied by von Neumann, in the framework of non-relativistic quantum theory.

The S-matrix study of this requirement is based on a generalization of the pole-factorization theorem, a special case of which is illustrated in Fig. 17.

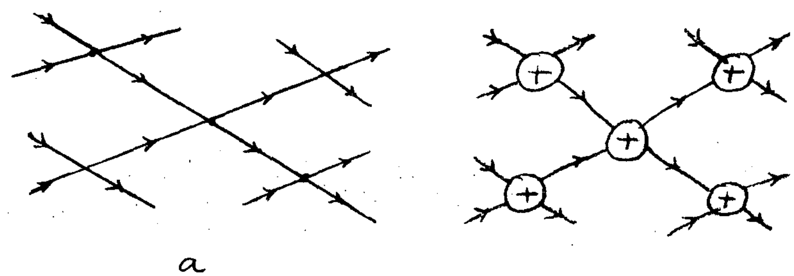

Fig. 17. A generalization of the pole-factorization theorem. Figures a and b represent, respectively, a Landau diagram (or causal network) and the corresponding discontinuity formula.

The probability that momentum-energy is transferred by a given mechanism is allowed to depend on the momenta p_j of the various particles involved, and on the various space-time intervals x_i over which the transfers carry. However, in accordance with the physical idea of macrocausality, this probability $P(p_j, x_i)$ is required to have a bound that decreases exponentially under space-time dilation:

$$P(p_j, x'_i \tau) \leqslant B(p_j, x_i)\, e^{-\gamma\tau} .$$

Here $B(p_j, x_i)$ is bounded in any bounded region in (p_j, x_i) space.

Different mechanisms can have different B and γ, but it is assumed that an upper bound on the scattering transition probability can be obtained by considering, in any finite momentum-energy range, only a finite number of different mechanisms.

Properties of scattering transition probabilities that hold in every model of the kind just described are regarded as general properties that follow directly from the physical idea of macrocausality.

It may be remarked that Planck's constant enters into S-matrix theory only as the parameter that fixes the scale of physical space-time relative to the mathematical space-time variable that occurs in the representation $\exp i\, px$ of the translation operator. Thus the space-time dilation generated by the transformation $\tau \to \infty$ is equivalent to the transformation $\hbar \to 0$. This means that the macroscopic limit $\tau \to \infty$ is equivalent to a classical limit $\hbar \to 0$. Consequently, the macrocausality property can be regarded as a form of correspondence principle: it asserts that the classical physical idea of macrocausality becomes valid in the classical limit.

III. APPLICATIONS

A. Derivation of Analyticity Properties

To derive analyticity properties from the macrocausality property one uses, for the initial and final particles, wave functions of the Omnes type:

$$\psi_j(p_j, \tau) \equiv \chi_j(p_j) \exp(i\, p_j\, a_j\, \tau) \exp -(\vec{p}_j - \vec{P}_j)^2 \gamma_i\, \tau .$$

The factor $\chi_j(p_j)$ is an infinitely differentiable function that is zero outside some finite region. The second factor generates a space-time translation by the amount $a_j\tau$. These translations move the particles apart in x space, but leave them unmoved in x' space. The third factor is a gaussian which concentrates the function near $\vec{p}_j = \vec{P}_j$ for large τ.

These Omnes functions have important properties. The width in momentum space shrinks like $\tau^{-1/2}$. Thus the width in coordinate space expands like $\tau^{1/2}$. Therefore the width in x' space shrinks like $\tau^{-1/2}$, and the x'-space trajectory region (i.e. the region where the particle is likely to be found) shrinks to a line, as indicated in Fig. 6.

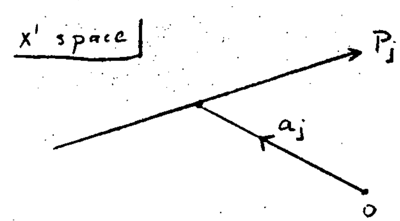

Fig. 6. The trajectory region shrinks to a classical trajectory in x' space.

More quantitatively, one finds that the probability that the particle lies in any closed bounded region in x' space that does not intersect the classical trajectory drops exponentially to zero. And, similarly, the probability that the particle has $\vec{p}$ in any closed bounded interval that does not intersect $\vec{P}$ goes exponentially to zero. Thus when viewed in x' space the particle goes over, in effect, to a free particle, modulo effects that fall off exponentially as $\tau \to \infty$.

Using these properties of the Omnes wave functions one may show that the macrocausality property implies the normal analytic structure.[1] This normal analytic structure consists of two properties. The first is that the physical-region singularities of scattering functions are confined to Landau surfaces. These surfaces, discovered by Landau, contain all perturbation theory physical-region singularities. That is, the functions represented by Feynman diagrams have physical-region singularities only on these surfaces.

Landau derived equations that defined these surfaces. Later Coleman and Norton pointed out that Landau's equations are just the condition that the Feynman diagram be interpretable as a causal network. This connection between causal networks and Landau surfaces is the root of the connection between macrocausality and the normal analytic structure.

The second part of the normal analytic structure consists of the $i\epsilon$ rules. The rules assert that the physical scattering functions on different sides of the Landau singularity surfaces are all parts of one single analytic function. And these rules specify precisely how this function should be continued around each Landau surface to reach the physical scattering function on the other side of that surface (see Figs. 7 and 8).

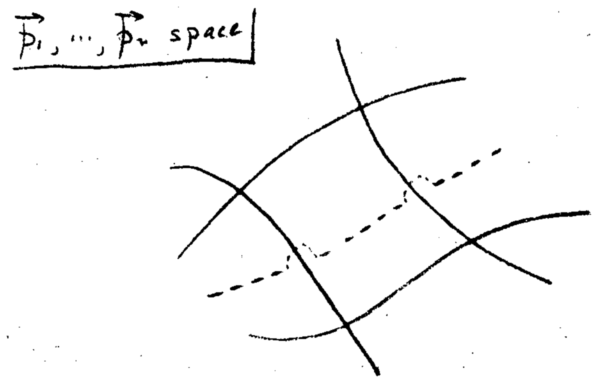

Fig. 7. The $i\epsilon$ rules specify the path of continuation that connects the physical scattering functions on different sides of Landau surfaces.

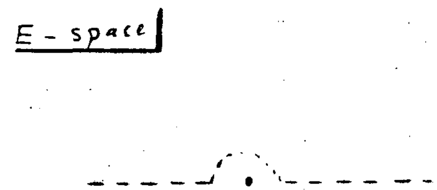

Fig. 8. In an appropriate energy variable the Landau surface is (locally) a point, and the physical continuation passes into the upper-half plane.

The fact that the scattering function is one single analytic function is neither trivial nor obvious. In fact, in theories with shadow particles of the kind discussed in the preceding talk by Professor Sudarshan the scattering function is not a single analytic function. Thus these theories lack the macrocausality property. I

The important point is that asymptotically only the singular part of the scattering amplitude contributes. Thus, if the space-time separations between the five collision regions in Fig. 18 are all large (see Fig. 18) then the scattering amplitude for the overall $6 \to 6$

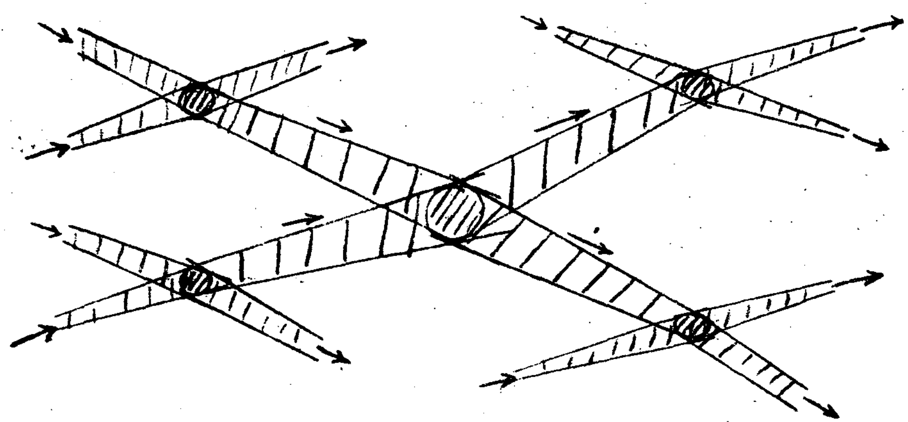

Fig. 18. A space-time process corresponding to Fig. 17.

process can be replaced by its singular part, which is exhibited in Fig. 18b, and the transition amplitude takes the factorized form

$$\int \prod_{i=1}^{4} \left[\psi_i^{(*)}(p_i) \frac{d^3 p_i}{2p_i^{\,o}(2\pi)^3} \right] \; S(p_i)$$

where

$$\psi_i^{(*)}(p_i) \equiv \int S_i(p_i, \, p_{ij}) \left[\prod_{j=1}^{3} \psi_{ij}^{(*)}(p_j) \frac{d^3 p_j}{2p_j^{\,o}(2\pi)^3} \right] .$$

Here $S(p_i)$ is the S matrix for the central $2 \to 2$ process in Figs. 17 and 18, where the index i runs over the four outer processes. The four functions $S_i(p_i, \, p_{ij})$ are the S matrices for these four outer $2 \to 2$ processes. The various $\psi_i^{(*)}$ are ψ_i or ψ_i^*

according to whether particle i is an incoming or outgoing particle for the central reaction, and the $\psi_{ij}^{(*)}$ are ψ_{ij} or ψ_{ij}^* according to whether ij labels an incoming or outgoing particle of the <u>ith</u> outer reaction.

The first two outer reactions (reading from left to right in Fig. 18) can be regarded as the reactions in which the two incoming particles of the central reaction are prepared. And the final two outer reactions can be regarded as the reactions that detect the two outgoing particles of the central reaction. Thus the factorized formula for the transition probability shows the consistency between the interpretations in which the outer reactions are considered, alternatively, as integral parts of the overall $6 \to 6$ process, or as the reactions that prepare and detect the incoming and outgoing particles of the central $2 \to 2$ reaction.

IV. THEORIES WITH SHADOW PARTICLES

A. The Measurement Problem

Theories with shadow particles encounter problems concerning measurement, which will now be discussed. The discussion is based on the foregoing discussion of the theory of measurements.

Consider a theory with shadow particles, of the kind discussed by Professor Sudarshan in the preceding talk.[4] Suppose there is a shadow particle of mass m . Consider a $3 \to 3$ scattering process in which the three incoming particles and the three outgoing particles are all ordinary (i.e., non-shadow) particles. And suppose the incoming and outgoing beams are arranged as shown in Fig. 19.

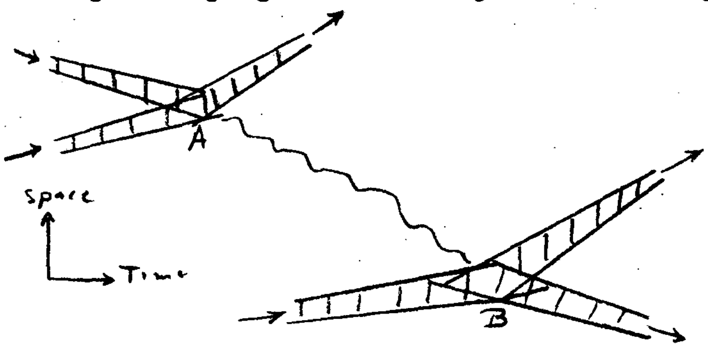

Fig. 19. The incoming and outgoing beams of the $3 \to 3$ scattering process are arranged so that two of the incoming beams and one of the outgoing beams intersect in a space-time region A, and so that the other two outgoing beams and the other incoming beam intersect in a space-time region B. The outgoing beams are defined by the acceptance conditions of the measuring devices that detect the outgoing particles.

Suppose the momentum-energies of the three external particles that intersect at A are such that a particle of mass m and momentum-energy k could be produced in this subreaction. And suppose the momentum-energies of the three external particles that intersect at B are such that their momentum-energy imbalance would be corrected by an extra incoming particle of mass m and momentum energy k. And suppose the locations of A and B are such that a space-time trajectory with direction $v = k/m$ connects A to B, as shown in Fig. 20.

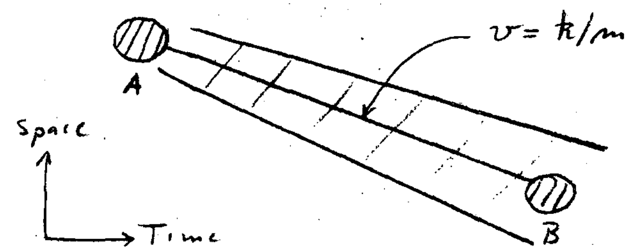

Fig. 20. The space-time region B lies in the region where a particle of mass m and momentum energy k could go if it were produced in A. The region of space-time corresponding to the various values of k that are compatible with the momentum-energy ranges in the incoming and outgoing wave functions is also shown.

If the particle of mass m were an ordinary (i.e. non-shadow) particle, then there would be pole in the scattering amplitude at

$m - i\epsilon$. The effect of such a pole is to give a contribution to this scattering process of exactly the kind that would be expected if a particle of mass m were produced at A and absorbed at B. In particular, the pole-factorization property ensures that the dominant contribution to the $3 \to 3$ scattering transition amplitude, for large separation between A and B, would have the form[5]

$$\int \frac{d^3p}{2p^o(2\pi)^3} \psi^*(p)\, \phi(p) \equiv \langle \psi \mid \phi \rangle$$

where

$$\phi(p) = \int S_A(p, p_j) \prod_{j=1}^{3} \left[\psi_j^{(*)} \frac{d^3p_j}{2p_j{}^o(2\pi)^3} \right]$$

and

$$\psi^*(p) = \int S_B(p, p_j) \prod_{j=4}^{6} \left[\psi_j^{(*)} \frac{d^3p_j}{2p_j{}^o(2\pi)^3} \right],$$

and all momentum-energy vectors are on-mass-shell. If the formula for $\psi(p)$ is substituted into the expression for the transition amplitude $\langle \psi \mid \phi \rangle$, then the result can be interpreted by saying that a particle of mass m and wave function $\phi(p)$ is produced in the reaction at A and detected in the reaction at B .[5]

If the particle of mass m is a shadow particle then the rules set forth by Sudarshan and co-workers[4] say that the S matrix for this $3 \to 3$ process should be calculated by using the principal-value resolution of the pole singularity at $E = m$. That is, one should use

$$\frac{1}{2} \left[\frac{1}{m - i\epsilon} + \frac{1}{m + i\epsilon} \right]$$

instead of the usual retarded propagator resolution $[1/(m - i\epsilon)]$.

The effect of this change on the transition probability rates predicted under the conditions represented in Figs. 19 and 20 is to decrease them by a factor of four. For in these situations only the retarded part of the propagator contributes significantly, and hence the factor of one-half occurring in front of the retarded part of the principal-value propagator produces a factor of one-quarter in the scattering transition probabilities. This means that the shadow particle can be detected by its interaction at B with ordinary particles, but that the probability of its being found at B is decreased by a factor of four.

The fact that the shadow particle can be detected in this way far away from the region in which it was formed conflicts with the ideas of shadow theory. For shadow particles are supposed to contribute to the dynamics, yet not appear as physically observed particles.

The problem, however, is that dynamics cannot be separated from observation. For what is observed is dynamical effects. If the long-range dynamical effects corresponding to a particle are present, then this particle is present. For in quantum theory a physical particle is nothing more than the physical effects that we associate with a particle.

The point, then, is that the effect of the retarded part of the principal-value propagator is to ensure that the shadow particle will propagate through the space-time region indicated in Fig. 20, in the physical sense that it can be detected in this region by probes consisting of ordinary particles.

Since the long-range dynamical effects corresponding to the reaction at B are present it is hard to understand how a charged shadow particle could fail to produce also tracks in a cloud chamber. For the two effects do not seem qualitatively different.

The obvious way out of these difficulties is to make the masses of all shadow particles complex. Then these particles would be unstable, and hence would not contribute to the asymptotic states. This is the strategy of Lee and Wick.[6] But Sudarshan and co-workers do not require their shadow-particle masses to be complex, and in fact usually deal with cases in which the shadow-particle masses are real.

B. Causality Problem

The difficulties just discussed arise from the retarded part of the shadow-particle propagator. The advanced part leads to other difficulties.

The advanced part of the shadow-particle propagator produces acausal precursor effects. In particular, it generates contributions to reactions of the kind shown in Fig. 21.

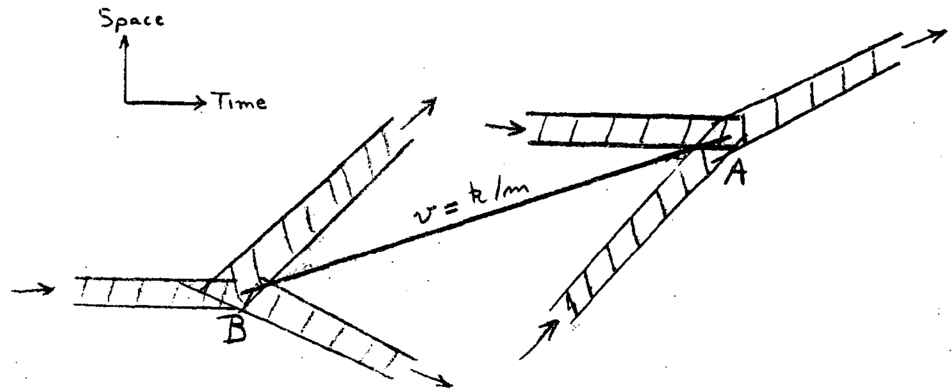

Fig. 21. A scattering process exhibiting the acausal effect. In this $3 \to 3$ process two of the incoming particles collide at A, and one outgoing particle emerges. The missing energy-momentum k is compatible with that of a shadow particle. The other two outgoing particles are observed to emerge from a region B, which lies in the intersection of the third incoming beam with a space-time trajectory that starts at A and moves backward in time along a space-time line that is parallel to the momentum-energy vector k .

The problem, now, is that the outgoing particles from B can, in principle, be detected before the incoming beams aimed at A are turned on. And the experiment can be set up so that these incoming beams are turned on if and only if the particles from B are not detected. On the other hand, by making the incoming beams sufficiently intense one can arrange that quantum theory will predict this: if the incoming beams are turned on then particles from B will almost surely be detected. And the set-up can be such that quantum theory will also predict this: if the incoming beams are not turned on then particles from B will almost surely not be detected.

This gives a "causal loop" similar to those discussed in earlier talks: if particles from B are detected then the beams will not be turned on, and quantum theory will predict that particles from B will almost surely not be detected. Conversely, if particles from B are not detected then the beams will be turned on, and quantum theory will predict that particles from B will almost surely be detected.

It is logically impossible for these statistical predictions of quantum theory to be borne out in a sequence of repetitions of this experiment. Thus quantum theory must, by logical necessity, fail to correspond to experience in the way that quantum principles demand. Thus the introduction of the principal-value propagator in the manner prescribed by shadow theory is incompatible with the basic interpretational principles of quantum theory.

The above argument is based on the Copenhagen interpretation of quantum theory. That is, quantum theory is viewed as fundamentally a procedure by which scientists make predictions about what they will observe under specified conditions. And the wave function is viewed as the quantum theorist's representation of an idealization of the finite system that he is examining, rather than some absolute representation of the world itself.

This Copenhagen view places the scientist and his macroscopic measuring devices outside the quantum system. Thus the quantum system is "open", in the sense used in earlier talks. The scientists sets up

the experimental conditions and is, as far as quantum theory is concerned, a free external agent.[7]

The causality problem just discussed, unlike the measurement problem discussed earlier, is not resolved by simply making the shadow-particle masses complex. For if the unstable shadow particles have sufficiently long lifetimes then by making the incoming beam sufficiently intense one could, in principle, still construct experimental arrangements that would lead to the contradictions with quantum theoretical principles. Moreover, even for shadow particles with small lifetimes there are two-particle branch-points at $m + m^* = 2 \text{ Re } m$ that lie in the physical region itself, and which would give acausal effects that have a power-law fall off, rather than an exponential fall off.[6] Though these acausal effects would in practice be small, they would generally lead in principle to causality problems of the kind just discussed.

A central question to which this conference has addressed itself is whether causality requirements have the force of logical necessity, or are mere expressions of convention or prejudice. Logical necessity can, of course, operate only within a logical or theoretical framework. However, within a given general theoretical framework causality requirements can be a logical necessity. The example discussed in this section illustrates this point.

FOOTNOTES AND REFERENCES

1. The mathematical details are given in D. Iagolnitzer and H. P. Stapp, Commun. Math. Phys. 14, 15 (1969).

2. H. P. Stapp, Foundations of S-Matrix Theory. I. Theory and Measurement, Lawrence Berkeley Laboratory preprint LBL-759, June 13, 1972.

3. These results follow directly from the procedures of Ref. 1.

4. Shadow particles are discussed in literature in the following papers: E. C. G. Sudarshan, Fields and Quanta 2, 175 (1972); C. A. Nelson and E. C. G. Sudarshan, Phys. Rev. D6, 3658 (1972); E. C. G. Sudarshan and C. A. Nelson, Phys. Rev. D6, 3678 (1972); A. M. Gleeson, R. J. Moore, H. Rechenberg, and E. C. G. Sudarshan, Phys. Rev. D4, 2242 (1971).

5. These questions are considered in more detail in H. P. Stapp, Phys. Rev. 139, 257 (1965); and in D. Iagolnitzer's article in Lectures in Theoretical Physics, 1968, XID (Gordon and Breach, 1968), ed. K. T. Mahanthappa and W. E. Brittin.

6. T. D. Lee in Proceedings of the International School of Physics "Ettore Majorana", 1970, ed. A. Zichichi (Academic Press, New York, 1971). Cf. also T. D. Lee and G. C. Wick, Phys. Rev. D3, 1046 (1971).

7. There is an opposing naive view of quantum theory that holds that the entire world is represented by a wave function. This view entails, however, either that the superposition principle fails to hold universally, in which case the theory is not quantum theory, or that the world we know is one of a continuously infinite collection of similar worlds, all but one of which must remain

forever unobservable.[8] The need to accept such a metaphysical assumption is a big price to pay for shadow particles. Moreover, a technical problem arises. For this interpretation is based on the idea that there is a Schroedinger equation that governs the temporal evolution of the world's wave function. Shadow theory, on the other hand, has been formulated in the S-matrix framework. Thus additional work would be needed to show that shadow theory can be generally formulated in terms of an evolving wave function of the world.

8. These opposing interpretations of quantum theory are discussed in H. P. Stapp, Am. J. Phys. 40, 1098 (1972). References are given there.

TECHNICAL INFORMATION DIVISION
LAWRENCE BERKELEY LABORATORY
UNIVERSITY OF CALIFORNIA
BERKELEY, CALIFORNIA 94720