and K approximated by a pole at $W=2m$ with residue equal to 50, which fits the one-pion-exchange amplitude approximately in the physical region. This result shows that even if the entire input amplitude were fitted by two poles at W_+ and W_-, with residues Γ_+ and Γ_-, respectively, the contribution from the anomalous threshold can at most amount to 3%. Hence, the initial assumption that its contribution is small is justified. The reason that $\bar{M}_{32}$ is small in this case is because the real parts of $W_\pm$ are exactly at threshold, and consequently $\rho_2(W_\pm)$ are very small, and so are G and K evaluated there. However, if the real parts of $W_\pm$ are large then the contribution from such complex singularities would also be large. Hence, the calculation carried out by neglecting the complex singularities are actually realistic within the framework of our program.

ACKNOWLEDGMENTS

I wish to express my gratitude to Professor P. Federbush for suggesting this problem and for his guidance throughout the course of investigation. I am also very grateful to many of my colleagues for valuable discussions and assistance; in particular, H. C. Lam, M. Forrest, M. Hoenig, D. Swanson, F. Tabakin, M. Weiss, and C. T. Wu. I am indebted to Dr. D. Sprung for a discussion of the experimental information relating to nucleon-nucleon scattering.

PHYSICAL REVIEW VOLUME 135, NUMBER 3B 10 AUGUST 1964

Exploration of S-Matrix Theory*

David I. Olive†
Department of Physics, Carnegie Institute of Technology, Pittsburgh, Pennsylvania
(Received 5 August 1963; revised manuscript received 20 April 1964)

The possibility of constructing an S-matrix theory from postulates concerning unitarity, analyticity, connectedness, the $i\epsilon$ prescription and the spin-statistics connection is explored. The existence and residues of the physical region poles are shown to follow from the connected unitarity equations. The validity of certain fundamental theorems known from field theory, Hermitian analyticity, extended unitarity, the existence of antiparticles, the substitution law for crossed processes and the TCP theorem is reduced, in simple cases, to the question of whether the S-matrix singularity structure permits specific distortions of certain paths. These distortions are shown to be possible in a "model" singularity structure consisting of the normal thresholds, and depend only upon simple properties of these singularities. It is explained that it is logically impossible to deduce the complete singularity structure without the results we are trying to prove. A suggested resolution of this difficulty is to set up a scheme of successive iterations in singularity structure to be justified by self-consistency. Then our work is the first step in such a scheme.

1. INTRODUCTION

RECENTLY some degree of understanding of the working of unitarity in S-matrix theory[1] has been developed, e.g., the way it evaluates discontinuities,[2–6] generates singularities,[6,7] and enables analytic continuations to be made onto unphysical sheets[3–5,8,9] In this sort of work a large number of properties or ingredients have been used. Apart from the quantum and Lorentz assumptions these are: (1) unitarity, (2) connectedness structure,[2,10] (3) maximal analyticity,[1] (4) the $i\epsilon$ prescription (see Sec. 3), (5) Hermitian analyticity,[2] (6) extended unitarity,[3] (7) existence of unphysical region stable poles on physical sheets, (8) the existence of antiparticles, (9) the substitution law for crossed processes, (10) the TCP theorem, (11) special physical sheet properties, (12) properties of physical region poles,[4,5] (13) connection between spin and statistics.

Several of these ideas can be grouped together. (5), (6), and (7) can be thought of as unphysical versions of the unitarity equations for T-matrix elements, valid at energies below the physical threshold of the amplitude concerned. The number of intermediate states included decreases with the energy so that (5) derives from the equation with no intermediate states and (7) from that with a single-particle intermediate state.

* This paper is a revised version of an unpublished Cambridge preprint circulated in July 1963 under the title "Towards an Axiomatisation of S-Matrix Theory." Compared with this the conclusions are restated more precisely. The work has been rearranged and more explanation given but no new results are included.

† Permanent address: Churchill College, Cambridge, England.

1 G. F. Chew, *S-Matrix Theory of Strong Interactions* (W. A. Benjamin and Company, Inc., New York, 1961).

2 D. I. Olive, Nuovo Cimento **26**, 73 (1962).

3 D. I. Olive, Nuovo Cimento **29**, 326 (1963).

4 D. I. Olive, Nuovo Cimento **28**, 1318 (1963).

5 J. Gunson (unpublished).

6 J. C. Polkinghorne, Nuovo Cimento **23**, 360 (1962); **25**, 901 (1962).

7 H. P. Stapp, Phys. Rev. **125**, 2139 (1962).

8 J. Gunson and J. G. Taylor, Phys. Rev. **119**, 112 (1960).

9 D. Zwanziger, Phys. Rev. **131**, 888 (1963).

10 H. P. Stapp, University of California Radiation Laboratory UCRL-10289, 1962 (unpublished).

(8), (9), and (10) will be referred to as the antiparticle theorems.

A satisfactory theory could only use a few members of this rather long list as axioms. It would be desirable that these axioms have a reasonably direct physical interpretation and that the theory would enable one to derive the remaining ingredients on the list and calculate the singularity structure. At present, the most attractive possibility is that of basing a theory only upon the physical unitarity relations and upon some sort of maximal analyticity postulate, in addition to the quantum and Lorentz assumptions.

The purpose of the present paper is to investigate this possibility and in particular to understand what features of such a theory could enable us to derive the unphysical unitarity relations and the antiparticle theorems. Before discussing our methods we shall briefly try to review the achievement so far in understanding the structure of S-matrix theory.

In the original investigation Stapp[7] studied the TCP theorem (10) and the connection between spin and statistics (13). He considered continuations in momentum space and used a principle of physical connection to identify certain negative energy parts of momentum space with amplitudes involving antiparticles. The TCP theorem followed from an assumption concerning single-valuedness on the physical sheet. We should like to understand the basis of these assumptions better.

At the same time he and Polkinghorne[6] showed that the unitarity equations generate a perturbation theory singularity structure[11,12] out of the normal thresholds and poles of the S matrix. To start off with however, one can deduce only the existence of physical region normal thresholds and their discontinuities, and these authors had to assume, either explicitly or implicitly, ingredients (5), (6), (7), (8), and (9), i.e., the unphysical unitarity relations and the antiparticle theorems in addition to analyticity and unitarity.

Recently, Gunson,[5] in an outstanding paper, has proposed alternative methods. He considered continuations in the space of invariants, realized the importance of connectedness (2) and physical region poles (12) and saw that the connection between particle and antiparticle poles leads to the possibility of proving crossing (9).

By a plausible procedure of subtracting out singularities from the unitarity equations he was able to enumerate and assign discontinuities to supposedly all the physical region singularities in the $3 \rightarrow 3$ amplitude lying between the 3- and 4-particle energy thresholds. These results were incorporated in his analyticity postulate, and were justified by an appeal to unspecified consistency requirements so that effectively he assumed (5) and (6), two of the unphysical unitarity relations. Also, he assumed (4), accepted Stapp's arguments concerning (8) and (10), and deduced (7) from his bound-state postulate. This played an important role in his formulation and stated that a certain two-particle normal threshold[4] was absent from a certain three-particle unphysical sheet.

All these approaches make assumptions involving either the unphysical unitarity relations or the antiparticle theorems or both and are therefore not immediately applicable in view of the fact that it is these relations which we wish to deduce. We shall still have to supplement the analyticity and unitarity assumptions (1) and (3) with assumptions concerning connectedness (2), the $i\epsilon$ prescription for physical region normal thresholds (4), and the connection between spin and statistics. The first two of these seem physically acceptable and are implicit in all the previous formulations. We have not investigated the third but understand that a new S-matrix theory derivation will shortly be available.[13]

Our methods will resemble those of Gunson[5] in exploiting the multiparticle aspects of the theory, connectedness and physical region pole structures. It is these very complications that enable us to do anything, whose sheer complexity at present prevents us from writing down our arguments in any but the simplest cases.

We find two main conclusions:

(A) The validity of the unphysical unitarity relations and the antiparticle theorems depends upon certain multiparticle features of the physical unitarity equations and upon certain plausible topological properties of the singularity structure which are satisfied in the case of a crude model of the S-matrix singularity structure consisting only of normal thresholds.

(B) Since one cannot deduce the complete singularity structure of the S matrix from the physical unitarity equations without using the theorems mentioned above, a vicious circle develops in the attempt to simultaneously deduce both the singularity structure and the fundamental theorems. We suggest that it may be possible to overcome this critical difficulty by proving that the singularities must possess some sort of hierarchical structure which will enable one to unravel the powerful consistency requirements of the theory. In this case we outline a program for generating the remaining singularities and completing our proof.

In Sec. 2 we discuss the connectedness assumption (2), write down some useful connected multiboson unitarity equations in bubble notation and develop rules for translating these into equations relating the A functions, the δ-function-free parts of the S-matrix elements.

In Sec. 3 we discuss the application of the maximal analyticity postulate to the A amplitudes as functions of invariants and formulate the $i\epsilon$ prescription assumption for the physical region normal thresholds. We ex-

[11] L. D. Landau, Nucl. Phys. 13, 181 (1959).
[12] R. E. Cutkosky, J. Math. Phys. 1, 429 (1960).
[13] H. P. Stapp (private communication).

plain why we shall henceforth ignore all singularities but normal thresholds and develop a way of understanding the overlapping normal threshold cut structure.

In Sec. 4 we study physical region pole structures and deduce the equivalence of poles and particles and the form of the residues. In Sec. 5 we use the four-particle unitarity equations and a multipole structure to deduce hermitian analyticity for the $2 \to 2$ amplitude and see that similar arguments would yield extended unitarity and corresponding results for multiparticle amplitudes.

In Sec. 6 we repeat and extend Gunson's arguments to prove that antiparticles must exist, to derive the substitution law for crossed processes and to prove the *TCP* theorem. We see how to define a generalized physical sheet in terms of the newly found complete normal threshold structure.

In the conclusion we discuss the suggestion that our work can be regarded as the first step of an iteration procedure which may construct a unique, consistent S-matrix theory.

2. UNITARITY AND CONNECTEDNESS

The physical processes to be described are particle scattering processes taking place in infinite time. There are supposed to exist asymptotic states in the infinite past or future describing systems of particles, each with a definite energy and momentum. The S-matrix relates these states so that its elements are probability amplitudes for the scattering processes. In saying that the asymptotic states and the S matrix itself are well defined we are supposing that the interparticle forces have a finite range so that asymptotically the particles behave like free particles at large spatial separations.

We shall only consider spin 0 particles satisfying the Bose statistics. Then the asymptotic particle states can be set up by the usual annihilation-creation operator formalism. General considerations indicate that the normalization of a single-particle momentum eigenstate must be proportional to a three-dimensional δ function. It is convenient to choose a relativistic normalization, and in particular

$$\langle p | p' \rangle = 2p^0 (2\pi)^3 \delta(\mathbf{p} - \mathbf{p}'). \tag{2.1}$$

The creation operator for a multiparticle asymptotic state is chosen to be the product of creation operators for the individual states. The corresponding single and multiparticle phase-space integrals are, respectively,

$$S_{\mathbf{p}} = (2\pi)^{-3} \int d^4p \, \delta^+(p^2+m^2) \tag{2.2}$$

$$S_{\mathbf{p}_1,\mathbf{p}_2 \cdots \mathbf{p}_n} = (n!)^{-1} \prod_{i=1}^{n} S_{\mathbf{p}_i} \tag{2.3}$$

in the case of n identical particles.

We shall assume that the S matrix exists, and that it is unitary.

$$S^\dagger S = SS^\dagger = 1. \tag{2.4}$$

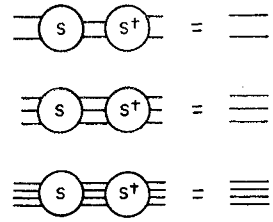

FIG. 1. The unitarity equations for certain S-matrix elements valid just above their physical thresholds. At higher energies the number of intermediate states included depends upon the energy. Equations with S and S^+ interchanged also hold.

These equations give information about the particle structure of the S matrix which can be made clearer by introducing a bubble notation. An S matrix element for the process $a \to b$ is to be represented by a bubble, with S written inside, and with the lines on the right-hand side corresponding to particles in state a and those on the left, to those in b. The S matrix element for a one-particle process, the single-particle state normalization (2.1), is represented by a single line, and the n-particle phase space integral (2.3) by n lines joining two bubbles. Then the unitarity equations for two-, three-, and four-particle scattering, valid just above the physical thresholds take the form in Fig. 1.[13a]

These equations indicate that the S-matrix elements themselves must possess a part corresponding to the right-hand side of the equations, a part which physically expresses the possibility that none of the particles may interact because of the finite range forces. In fact, there must be more structure than this because in multiparticle processes certain subsets of incoming particles may not interact with each other even though the constituent particles of each subset interact amongst themselves. The resultant connectedness structure is expressed in bubble notation in Fig. 2.

The importance of this structure in S-matrix theory has been pointed out by several authors.[2,5,10] It is of course contained within the Feynman rules for the perturbative expansion of the S matrix and it is this which is largely responsible for our physical intuition.

Having seen how it is understood physically we shall now regard the connectedness structure as a formal assumption which tells us how to subtract out the parts of the S-matrix elements containing various energy-momentum conservation δ functions in order to obtain the connected part, consisting of an over-all energy-momentum conservation δ function times a δ-function-free function which will be our candidate for analytic continuation.

Functional methods provide the most concise mathematical formulation of the connectedness structure. We define the functional $\hat{H}$ obtained from the operator

[13a] *Note added in proof.* Due to an unfortunate oversight in this diagram the 2-particle term has been omitted from the second equation and the 2- and 3-particle terms from the third. This mistake is not carried on to subsequent work.

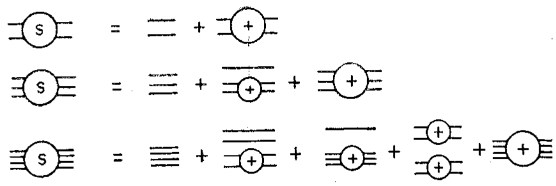

FIG. 2. Bubble notation equations representing the connectedness structure of certain S-matrix elements. The bubbles with a $+$ denote the connected part of the S-matrix bubble with corresponding external lines. A summation over all possible diagrams with similar structure is understood. Corresponding equations hold for $S^\dagger$-matrix elements.

H by

$$\hat{H}=\sum_{n=0}^{\infty}\sum_{m=0}^{\infty}(n!m!)^{-1}S_{\mathbf{p}_1\cdots\mathbf{p}_n}S_{\mathbf{q}_1\cdots\mathbf{q}_m}\bar{J}_{\mathbf{p}_1}\cdots\bar{J}_{\mathbf{p}_n}\times\langle\mathbf{p}_1\cdots\mathbf{p}_n|H|\mathbf{q}_1\cdots\mathbf{q}_m\rangle J_{\mathbf{q}_1}\cdots J_{\mathbf{q}_m}.$$

By functional differentiation with respect to the "currents" J, thereby effectively pulling out states, one can verify that Eq. (2.5) correctly expresses the structure in Fig. 2.

$$\hat{S}=\exp\hat{S}_c. \qquad (2.5)$$

Expansions of this type have been quoted in field theory[14] for the vacuum expectation values of time-ordered products. There the integration is over space-time, whereas here it is over phase space.

It is possible to obtain a closed form for the complete set of connected unitarity equations by taking the logarithm of the functional unitarity equations resulting from the insertion of (2.5) in (2.4). The functional differentiation process which yields the explicit connected equations is too clumsy to be of use. Instead it is easier to insert the connectedness expansion in Fig. 2 for the relevant S-matrix element into the equations in Fig. 1 and isolate the connected part. The remaining disconnected part automatically vanishes by lower order equations.[5,10] In Fig. 3 we use the bubble notation to write down a few unitarity equations, valid for various amplitudes just above their physical energy thresholds. For simplicity we have chosen a single identical particle theory but none of our results are restricted by the symmetrics introduced in this case.

The terms on the right-hand side (r.h.s.) are preceded by a minus if the number of bubbles labeled with a minus is even. Similar equations hold with $+$ and $-$ interchange on the rhs and with appropriate sign adjustments. The last term in Eq. (c) of Fig. 3 is interesting because it does not give rise to an ordinary normal threshold and would not occur if all the four particles involved were different and had suitable quantum numbers.

The A matrix element defined symbolically by

$$S_c=-i(2\pi)^4\delta(P_i-P_f)A \qquad (2.6)$$

is the obvious candidate for analytic continuation since it is δ-function free. Equation (2.5) demands a unique phase in S, that which will eventually lead to the Hermitian analyticity property of A. The prescription for translating bubble equations into equations involving A functions can be found by substituting (2.6) into the rules above. A common factor which is the coefficient of A in (2.6) can be cancelled and the remaining energy-momentum conservation δ functions integrated out by the introduction of loop integrations. The result is

(1) for each $+$ or $-$ bubble, $A^{\pm}$ where
$$A_{ab}^{+}=A_{ab},\ A_{ab}^{-}=A_{ba}^{*},$$
(2) for each internal line, $-2\pi i\delta^{+}(q^2+m^2)$, (2.7)
(3) for each loop $\int i(2\pi)^{-4}d^4k$,
(4) for each r particle state joining two bubbles a factor $(r!)^{-1}$.

These precisely resemble the Cutkosky rules obtained in perturbation theory,[12] but do not yet have any discontinuity content. The fact that there is no overall factor justifies our choice of normalization in (2.1) and (2.6).

3. ANALYTICITY

3.1 Decomposition of Multiboson Amplitudes and the Analyticity Postulate

The multiboson A-matrix elements are functions of real four-vectors and are invariant under proper orthochronous Lorentz transformations, and so must be functions of the Lorentz scalars which can be formed out of the four-vectors. There are two possible types, inner products of four-vectors, $z_{ij}=p_i\cdot p_j$, and determinants of four different four-vectors $\epsilon(p)$. The latter, unlike the z, change sign under spatial reversal, while, like the z, remaining invariant under complete reversal. By the rules of determinantal multiplication, any product of two such determinants is a polynomial function of the z so that all such determinants are proportional to each

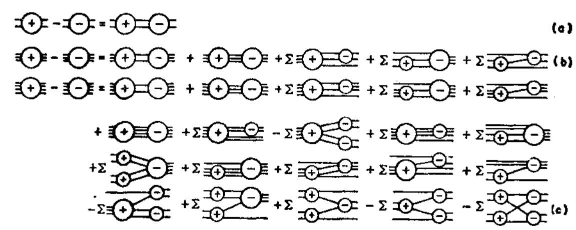

FIG. 3. Bubble notation equations (a), (b), and (c) representing the connected part of the unitarity equations valid for the $2\rightarrow 2$, $3\rightarrow 3$ and $4\rightarrow 4$ amplitudes, respectively, at energies just above their physical thresholds. The bubbles with $+$ and $-$ denote the matrix elements of S_c and $-S_c^\dagger$, respectively, and n lines joining two bubbles the n particle phase space integral (2.3). The summation indicates that over all diagrams of a similar structure. The intermediate states to be included depend upon the values of the external momenta.

---

[14] K. Symanzik, *Hercegnovi Lectures on High Energy Physics* (Yugoslav Atomic Energy Commission, 1961).

other with a rational function of z as coefficient. Because of these properties, the multiboson amplitude can be decomposed linearly into scalar and pseudoscalar parts:

$$A(p)=A_s(z)+\epsilon(p)A_p(z), \tag{3.1}$$

where A_s and A_p are functions of inner product invariants z only.

Given the variables appearing in (3.1) and the sign of the time-like component of one of the particle momenta, which is, of course, fixed for the physical process in question, one can easily construct a set of corresponding momenta and see that this set is unique to within a proper orthochronous Lorentz transformation. Thus, Eq. (3.1) completely exhibits the dependence of the physical A amplitude. In view of some deductions we make in Sec. 6 it is important to note that it is not in order to include a term involving the sign of the time-like component of a vector which is the difference of sums of incoming and outgoing momenta because this quantity is determined by the variables already included in (3.1)

Our preliminary analyticity assumption is that the A_s and A_p functions can be analytically continued from the physical region meeting only those singularities required by the unitarity equations, and, in the case of A_p, additional possible poles at the zeros of $\epsilon(p)$, i.e., part of the boundary of the physical region, behaving not worse than $\epsilon(p)^{-2}$. We are forced to include the possibility of such kinematic poles because, if we did not, we would find that the pseudoscalar function A_p corresponding to a different choice of momenta in $\epsilon(p)$ would have to possess such poles, and then our analyticity assumption would be inconsistent in that it depended upon a special choice of $\epsilon(p)$.

We choose to consider functions of scalar inner product invariants rather than four-momenta because we wish to incorporate Lorentz invariance in a simple way, and not have to bother with Hall-Wightman theorems, and because the fundamental singularities, the normal thresholds, will appear in a simpler form. Also, awkward mass-shell conditions are avoided, but all this is at the expense of having nonlinear Gram determinant conditions[15] expressing the dimensionality of momentum space, and unitarity integration hypercontours with complicated boundaries.

Later we shall discuss how various singularities, normal thresholds and others, may enter the physical region. This eventuality makes it immediately necessary to add an extra clause to our analyticity postulate to the general effect that the physical amplitude is the boundary value of an analytic function. Because of various complications introduced by the multiparticle features we shall have to state this in a different way. We imagine the real sections of the singularity hypersurfaces dividing the physical region into segments and express the assumption in terms of the properties of the physical region paths of analytic continuation which relate the physical amplitudes in different segments.

The imaginary distortions necessary to enable a path of continuation to link the physical amplitude in two adjacent segments of the physical region need only be infinitesimal. The assignment of imaginary distortions must be such that any closed physical region path can be contracted to zero.

The second requirement ensures that two different paths lying within a physical region and leading from the physical amplitude at a point A to the physical amplitude at a point B do indeed encounter the same single valued function. The requirement that a path could be distorted through an intersection of singularity surfaces with linearly dependent normals (a pinch configuration)[6] restricts the possible relative assignment of imaginary distortions to the singularities. It is even possible to imagine singularity configurations in which it cannot be satisfied. It is apparently a consistency requirement upon the theory, whose understanding would clarify the status of the analyticity assumption, that the unitarity equations cannot generate such configurations.

Later when we discuss the normal thresholds we shall make more specific assumptions about the imaginary distortions.

3.2 Nomenclature

We shall call a combination of external lines of the amplitude for a given physical process a channel, and two channels whose lines are disjoint and exhaustive a reaction. This usage is wider than that conventional in nuclear theory since at present we can only associate a well-defined physical state with a channel composed of lines either all incoming or all outgoing, and since the only "reaction" describing a physical process is that composed of the incoming and outgoing channels. Later we shall see that all these "reactions" can be related to physical processes but at present we shall use the words in the rather loose sense that has become standard in the literature on S-matrix theory.

An amplitude with n external lines has $2^{n-1}-n-1$ associated different reactions providing we exclude reactions with single-particle channels and do not distinguish the direction of the reaction. The channel invariant variable is the square of the energy in the given channel C

$$s_C=-(\sum_{i\epsilon C}\pm p_i)^2, \tag{3.2}$$

where $\pm p_i$ are the momenta of incoming and outgoing lines, respectively. By energy-momentum conservation this quantity is identical for each channel in a reaction.

We shall consider the space of these $2^{n-1}-n-1$ different channel invariants in preference to the Z's above.

[15] V. E. Asribekov, Zh. Eksperim. i Teor. Fiz. 42, 565 (1962) [English transl.: Soviet Phys.—JETP 15, 394 (1962)].

Although the channel invariants are linearly related, the advantage is that it will be easier to discuss a specific normal threshold singularity in terms of the channel invariant in which it appears fixed.

The channel invariants for a given process can be classified into four types which we shall refer to as

(1) total energy,
(2) subenergy (composed of lines either all incoming or all outgoing),
(3) momentum transfer (composed of one incoming line and one outgoing line),
(4) crossenergy (composed of both incoming and outgoing lines but not 3).

In the physical region the momentum transfers cannot support real intermediate states whereas the cross energies can run from momentum-transfer-like values to energy-like values and are obviously the predominant type of variable in processes involving very many particles.

We shall say that two reactions overlap when the channels from the different reactions always have external lines in common. Thus, the cross-energy reaction overlaps the energy reaction whereas the subenergy reaction does not.

3.3 Normal Thresholds and Their $i\varepsilon$ Prescriptions

Without knowing anything about Hermitian analyticity, the statement that A^-, defined in (2.7), is the opposite boundary of A^+, one can deduce the existence of normal threshold singularities in any channel invariant that can support a real intermediate state whose threshold is within the physical region of the amplitude considered. As the channel invariant increases past the threshold, a new term enters the unitarity relation corresponding to the newly possible intermediate state. Examples of such terms are the terms with single-particle intermediate states in cross energies Eqs. (b) and (c) of Fig. 3 and the thirteenth term drawn on the right-hand side of Eq. (c) of Fig. 3. More examples occur in the unitarity equations valid at higher energies.

The unitarity equations also suggest the existence of further normal thresholds in the energy and subenergy variable occurring at the physical thresholds of these channels and at lower values. We cannot properly deduce the properties of these singularities until we have derived the unphysical unitarity equations operating in their vicinitiy.

When the energy in, say, the $3\rightarrow 3$ amplitude approaches its physical threshold, the three-particle threshold, the inequalities imposed by the requirement that we consider points within the physical region cause the subenergies to simultaneously approach their physical thresholds, the two-particle thresholds. This sort of behavior does not apply to the normal thresholds within the physical region, and it is characteristic of these that it is possible to move across them one at a time while remaining in the physical region.

The physical-region normal thresholds divide up the physical region into segments, and the question arises as to how the physical amplitudes in these segments are related analytically. Consider a special case. According to the idea of maximal analyticity the path in the appropriate channel invariant plane connecting the two physical amplitudes on each side of the singularity ought to be distorted into either the upper or the lower half-planes, rather than wind round the singularity. We supplement our analyticity assumption by assuming that the $+i\epsilon$ prescription in the correct invariant is appropriate for all the physical region normal thresholds in all amplitudes. Perturbation theory is the guide here but we have strong suspicions that the assumption can be removed.

This formulation of the $i\epsilon$ prescription shows that it has nothing to do with arrangements of cuts. Also, by associating an $i\epsilon$ with a channel invariant only when it nears a normal threshold, we avoid the difficulties and ambiguities which arise in trying to reconcile the constraint equations, linear and Gram determinant with an $i\epsilon$ associated with each energy variable. This is providing we allow the possibility that some variable may be temporarily depressed below the real axis when remote from a singularity.

Since A_{ab} and A_{ba} contain the same singularities, $A_{ab}^-=A_{ba}^*$ contains them also, but, because of the complex conjugation, follows the $-i\epsilon$ prescription.

Since different forms of physical unitarity equation hold on either side of a physical region normal threshold it is possible to use methods established elsewhere[3,5] to evaluate the discontinuity across the corresponding cut as a unitarity type of integral involving the expected single intermediate state. We emphasize that it is not necessary to know about Hermitian analyticity in order to do this but only that the $-$ amplitudes possess opposite $i\epsilon$ prescriptions to the $+$ amplitudes with respect to the same singularities. Thus, we know almost immediately that unitarity equations are associated with discontinuities. A difficulty in the argument[3] is that one must continue the unitarity equations and that this process may generate singularities of a new type.[16,17] We shall not derive or use such results except in the special case of single-particle thresholds when we shall obtain the pole residue.

3.4 Normal Threshold Model

Although, in principle, a systematic analysis of these other physical region singularities is possible at this stage, little is known about them yet, and we shall ignore them. This is not too serious in view of the fact that the sort of properties we would require of them

[16] R. E. Cutkosky, Rev. Mod. Phys. **33**, 448 (1961).
[17] P. V. Landshoff, Phys. Letters **3**, 116 (1962).

seem to be assured by our analyticity postulate, and since we are forced to ignore another class of singularities for more fundamental reasons.

Later on we shall see that we have to make arguments concerning certain unphysical regions lying on the real axis below the physical threshold in the various energy-like variables. Amongst the singularities expected to lie in this region are ones corresponding to Landau diagrams with three-line vertices. At the present stage in the argument we cannot possibly deduce the existence of such singularities, far less their properties. As an example we mention the triangle diagram which sometimes causes an anomalous threshold in two-particle scattering amplitudes. According to our present point of view this is a very sophisticated singularity because, as was discussed by Polkinghorne,[6] it is generated by the momentum transfer in the unitarity integral hypercontour striking a crossed reaction stable pole. But, as yet, we can neither deduce the existence of crossed reactions or of the stable poles in unphysical regions. Similarly we cannot deduce the existence of any singularities generated by normal thresholds in unphysical regions whether below physical thresholds in energy-like variables or in momentum transfer channels.

Thus we cannot deduce the singularity structure without the fundamental theorems and we cannot deduce the fundamental theorems without the singularity structure. This vicious circle is a fundamental obstacle to the deductive approach to S-matrix theory. Gunson's answer[5] to this problem was to postulate a discontinuity scheme involving all the unphysical unitarities and to appeal to unspecified consistency requirements. This is unsatisfactory to our point of view because we wish at least to understand the nature of the consistency requirements and would prefer to see whether they can be unravelled so that the whole theory can be put on a deductive basis.

The method we shall suggest, and whose details and logical status we shall discuss at greater length in the conclusion, is to set up a scheme of successive iterations in terms of the singularity structure of the S matrix. The singularity structure that can be deduced is used to derive the fundamental theorems within this "approximation" and then these theorems are used, within their "approximation" to generate further singularity structure and so on.

In the following sections on the unphysical unitarity relations and the antiparticle theorems, but not in that concerning physical-region poles, we shall understand that we ignore all singularities but normal thresholds. At first we shall only be able to include the physical region ones but will find that eventually all of them can be taken into account. The motives for this rather crude simplification can be variously ascribed, according to the taste of the reader, to either our inability to do better because of our present lack of understanding of the S-matrix singularity structure, our desire to present our methods in a simple context, or our belief that we are carrying out the first step in an iteration scheme which will eventually produce a consistent S-matrix theory.

3.5 Overlapping Cuts and Independence

First we must clarify the normal threshold structure that we do have.

A complication arising in multiparticle amplitudes is the occurrence of overlapping normal cuts when several different channels can simultaneously support real intermediate states while being related by the constraint equations.[18] According to the linear and Gram determinant relations there are only $3n-10$ independent variables. Suppose we select an independent set and fix all but one w. As it varies along its real axis and encounters its own normal thresholds it causes the redundant variables to vary similarly. The singularities arising when they strike their normal thresholds is reflected in the w plane as the overlapping cuts. The appropriate distortions for these depends upon the relevant constraint equations. Two apparently coincident singularities can be separated by adding a suitable imaginary increment to one of the independent variables which is not itself at a normal threshold.

Since we can pass over one normal threshold at a time we can encircle one at a time and we shall want to understand the relative structure of the normal thresholds in different channels, whether for example, paths encircling a w normal threshold and an overlapping normal threshold commute. Looking at just one plane we cannot answer such questions but, if we choose a set of independent invariants including the two whose normal thresholds we are interested in, and choose values of invariants such that we can encircle either normal threshold without interference from other normal threshold in redundant channels,[19] we see that paths encircling the two thresholds do commute since they are fixed singularities in independent variables.

Translated back into the w plane the result is a nontrivial relation between the overlapping singularities and the w singularities which we shall call "independence." It will be mentioned later that this property can be shown to be consistent with the unitarity equations in a nontrivial way. It means that we can correctly represent the multiparticle normal threshold structure by thinking of a direct product of planes, one for each channel invariant, in which our point of observation moves in each plane subject to the linear and Gram determinant relations. As yet, all we know about the structure of the individual planes is the existence of the physical region normal threshold singularities. We cannot yet deduce the existence of normal

[18] P. V. Landshoff and S. B. Trieman, Nuovo Cimento 19, 1249 (1961).

[19] Such interference does sometimes occur when we include further singularities and then our result is modified.

thresholds in the unphysical part of the energy and subenergy channels or in any part of the momentum transfer channels. Although we cannot even deduce the existence of fixed normal thresholds in the unphysical regions reached by encircling the physical region normal threshold it follows from independence that we would find exactly the same overlapping cut structure.

Later in the paper we shall be able to fill in our independence picture by deducing these extra normal thresholds. The independence structure will enable us, in a convenient and unambiguous way, to denote the generalized unphysical sheets reached by encircling sequences of normal thresholds by these sequences themselves. This is the generalized path notation introduced in a previous paper.[4]

Because of the constraint conditions it may sometimes be impossible to add positive imaginary increments to a set of energy variables simultaneously and this means that it will not always be possible to move directly into the physical sheet, as defined later, from the physical region.

3.6 A Theorem

We conclude this section by mentioning an important result concerning the singularities of unitarity-like integrals. If we consider a reaction R of a unitarity-like integral consisting of two bubbles joined by an intermediate state in a reaction R', then, if, and only if, R and R' overlap, in the sense above, the function as a whole possesses no normal threshold singularities in the reaction R. When the bubbles are given a perturbation theory structure, the result follows from the Cutkosky rule[12] that, when looking for the Landau singularities of a perturbation theory discontinuity integral, one cannot contract the lines bearing δ functions. An argument due to Polkinghorne[6] indicates that such a result is the consequence of the unitarity-like structure of the integral rather than the detailed properties of the bubbles and we take this to be the case. The result is applicable to unitarity-like integrals and to discontinuity integrals, which differ from the former in the assignment of boundary values. The argument could obviously be developed to obtain results concerning integrals with more structure.

4. PHYSICAL REGION POLES

4.1 Consistency

According to the rules (2.7), multiparticle unitarity equations, e.g., Eqs. (b) and (c) of Fig. 3, possess terms

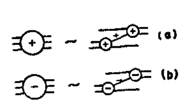

FIG. 4. Equation (a) indicates the existence and residue of a physical region pole occurring in the $3\rightarrow 3$ + amplitude. Equation (b) indicates the corresponding pole and residue in the − amplitude.

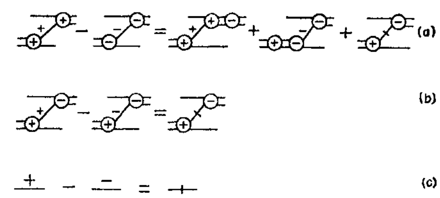

FIG. 5. These equations indicate the consequences of inserting the pole structures of Fig. 4 into the 3-particle unitarity equation, Eq. (b) of Fig. 3, and isolating the dominant pole terms.

proportional to a single-particle mass-shell δ function. These immediately suggests the existence of a single-particle pole within the physical region at the point in question. This pole could be understood physically, in the case of three-particle scattering, for example, as due to the dominance of processes in which two of the particles scatter on each other, and then one of these particles scatters again on the third incident particle. Diagrammatically such a process could be represented by Eq. (a) in Fig. 4.

The internal line represents a propagator function D^+ containing the pole. This picture suggests that the residue to the pole is proportional to the product of two mass-shell $2\rightarrow 2$ amplitudes.

We shall adopt this viewpoint, and show that it is consistent with the unitarity equation [Eq. (b) of Fig. 3] and identifies the function D. This argument and the result [Eq. (a) in Fig. 4] were referred to in Ref. 4. We shall then argue that the pole is the only permissible singularity at the point in question, and that the residue is uniquely determined. In addition, all such poles must correspond to particles.

It follows that the complex conjugate amplitude for the inverse process possesses the pole given by Eq. (b) in Fig. 4 where the internal line labeled with a minus represents the complex conjugate of D^+.

We pick out and equate the dominant parts of terms in Eq. (b) of Fig. 3 possessing the one-particle pole structure in question. Only the first two terms on the rhs and the terms included in the sums but not explicitly drawn, lack this. This rather obvious conclusion is a simple case of the theorem mentioned at the end of the previous section. We obtain, in bubble notation, Eq. (a) of Fig. 5.

According to the rules (2.7) the crossed line bears $-2\pi i$ times a mass-shell δ function. On using two-particle unitarity Eq. (a) of Fig. 3 to simplify the first two terms on the right we obtain, on cancellation, Eq. (b) of Fig. 5. Cancelling the scattering amplitude factors Eq. (c) (Fig. 5) results, or, in terms of the D propagator function,

$$D^+(u)-D^-(u)=-2\pi i\delta(m^2-u)\,.$$

Assuming D is analytic near the point concerned, and remembering that $D^-=D^{+*}$ by definition the dominant part of the solution is

$$D^+(u)=\pm(u-m^2\pm i\epsilon)^{-1} \tag{4.1}$$

or any appropriate linear combination of the two.

According to our $i\epsilon$ assumption we shall adopt the $+i\epsilon$ prescription for all singularities. We have shown that it is consistent to assume the behavior, in terms of A amplitudes

$$A_{33}\sim A_{22}A_{22}/u-m^2+i\epsilon\,. \tag{4.2}$$

The particularly simple form of the denominator is a consequence of our normalization conventions (2.1) and (2.6).

It would seem plausible that this sort of argument works in general and we shall assume this. Indeed, Stapp[13] has succeeded in developing a generalized bubble notation which demonstrates this.

4.2 Uniqueness

We shall now show that this dominant behavior is unique. Suppose it were possible for the amplitude to have some other dominant behavior nearby, due, for example, to some other singularity superimposed upon the pole, or a pole with a different, possibly nonfactorizing residue. This must, according to our assumptions, possess a $+i\epsilon$ prescription in the amplitude, and in the inverse amplitude (trivially so for identical particle processes) an so a $-i\epsilon$ prescription in the complex conjugate inverse amplitude. Equation (a) (Fig. 2) is linear in the $3\rightarrow 3$ amplitude apart from the "inhomogeneous" last term. The difference amplitudes $d^\pm$ between the two kinds of behavior must satisfy the "homogeneous" equation (a) of Fig. 6.

Two functions equal along part of the real axis, possessing singularities just off this region, but with opposite $i\epsilon$ prescriptions can only both be regular. This conclusion is expressed in Eq. (b) (Fig. 6).

This integral equation is invertible because the scattering amplitude possesses no poles in its physical region, and in fact, satisfies the Fredholm condition,[3] and we see d^+ itself must be regular.

Thus, a pole with the residue [Eq. (a) (Fig. 4)] is the unique possibility. Also the existence of a pole with nonzero residue is only possible if there is a δ function then present, as in Eq. (a) (Fig. 5), so that given a pole, there must be a corresponding particle in the completeness relation.

4.3 Further Comments

Later on we shall often want to continue pole terms like (4.2). Knowledge of the singularity structure of the complete amplitude enables us to continue the residue, providing we remain on the mass-shell section prescribed by putting the denominator of the pole term equal to zero, because singularities in the residue must appear in the complete amplitude (but not necessarily

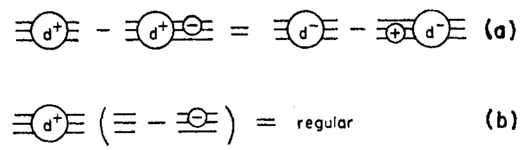

FIG. 6. These are the equations satisfied by the difference d between two different possible dominant behaviors.

vice versa). A sufficient condition for the residues of a given pole at two points to be continuations of each other is that there exists a path of analytic continuation in the complete amplitude joining the two points and lying within the mass shell section.

If we had investigated the factorization Eq. (a) (Fig. 4) at a higher energy, e.g., above the four-particle threshold, there would have been more terms in Eq. (b) (Fig. 3) and so Eq. (a) (Fig. 5). Instead of Eq. (a) (Fig. 3) we would have to use three-particle unitarity for the two-particle amplitude, and would then arrive at the same factorization Eq. (a) (Fig. 4). This is expected, because we could have continued Eq. (a) (Fig. 4) in the energy and subenergies to the higher energy region. The reason for the agreement is that we had already assumed that the $i\epsilon$ prescription in the two-particle amplitude energy and the three-particle amplitude subenergy for the three-particle threshold are identical, whereas we could have deduced this fact. This illustrates how the $i\epsilon$ prescription must satisfy certain consistency requirements.

In multiparticle amplitudes there also exist more complicated pole structures with several physical region poles and we must check that the residue factorizes as expected. Assuming that we have justified all single-pole factorizations by the methods just given, we can factorize the amplitude by considering one pole at a time, and so arrive at the expected answer. The fact that the result must be independent of the order chosen tells us that the $i\epsilon$ prescription for a given pole must be the same in all amplitudes. Alternatively the whole factorization structure could be justified in one step by insertion into the appropriate unitarity equation and this is illustrated in Sec. 5.1.

If one were to assume that all A matrix elements for processes involving more than a certain number of particles vanish, then all the lower ones must vanish also, because the lower ones must be factors in residues of the vanishing amplitudes. We mention this because an analogous conclusion has been reached in field theory.[20]

5. THE UNPHYSICAL UNITARITY RELATIONS

5.1 Preliminary

We mentioned that the Hermitian analyticity and extended unitarity equations can, for two-particle

[20] O. W. Greenberg and A. L. Licht, J. Math. Phys. 4, 613 (1963).

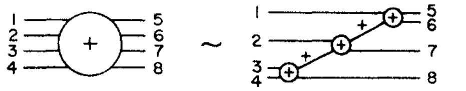

FIG. 7. This equation denotes the particular pole structure in the $4 \rightarrow 4$ amplitude which is under consideration. It follows from this that a similar equation holds with each $+$ replaced by a $-$.

scattering amplitudes at least, be thought of as unphysical versions of the unitarity equations for the A amplitudes in the sense that they operate at energies below the physical thresholds of the amplitudes concerned and that the number of intermediate states to be included is determined by this energy, just as in the physical unitarity equations. In particular, when the energy is below the lowest threshold, the unphysical unitarity relation contains no intermediate states and simply equates the functions obtained by continuing the $+$ and $-$ amplitudes down the energy axis with $+$ and $-i\epsilon$ prescriptions, respectively. Thus, the original $+$ and $-$ amplitudes are continuations of each other. In fact, opposite boundary values of the same function onto the energy axis and the result is called Hermitian analyticity. It has been shown that this fundamental property of the A amplitudes, which means that unitarity indeed evaluates the discontinuity, holds in field theory,[2,21] and in multichannel potential theory, irrespective of whether any special invariance principles, such as time-reversal invariance, operate. "Real analyticity" is then a special case of Hermitian analyticity which is valid only when the amplitude happens to be symmetric.

Hermitian analyticity is not necessarily a relativistic phenomenon (and so cannot depend upon the TCP theorem as was erroneously stated in[2]), but depends upon the Hermiticity of the Lagrangian or potential or else the reality of the energy spectrum. In S-matrix theory it will appear to depend upon the corresponding properties, the unitarity of the S matrix and the reality of the stable particle masses and also upon the $i\epsilon$ prescription for the physical region normal thresholds.

We now propose to use pole factorizations to show that the unphysical unitarity equations are contained within the multiparticle unitarity equations. The only type of channel invariant which can take physical values above or below the lowest continuum threshold is the cross energy. Apparently the simplest case in which we can use physical region pole factorizations to isolate this structure is the four-particle unitarity equation [Eq. (c) of Fig. 3], and the interesting term is that with a possible two-particle intermediate state in the cross energy. In the part of the physical region, just above the four-particle threshold, where this term does contribute, there is present the physical region pole structure shown in Fig. 7. To check this assertion we pick out from Eq. (c) (Fig. 3) those terms which, according to the overlapping channel theorem mentioned at the end of Sec. 3, possess this structure (Fig. 7) and write the result as the equation in Fig. 8. Since the poles are in the physical region we are entitled to equate the pole terms found by inserting the structure in Fig. 1, and the corresponding factorization for the minus amplitude, and also Eqs. (a) and (b) of Fig. 4 to obtain, in the notation of Sec. 4, Eq. (a) (Fig. 9). We have numbered the corresponding terms in Fig. 8 and Eq. (a) of Fig. 9.

Using two-particle unitarity [Eq. (a) of Fig. 3] to simplify terms [1] minus [3] to [1′] and terms [2] minus [4] to [2′], we obtain Eq. (b) (Fig. 9). We now use Eq. (c) of Fig. 5 to simplify terms [1′] minus [5] to [1″] and terms [2′] minus [6] to [2″] and obtain Eq. (c) of Fig. 9. As a result of these manipulations only three terms remain, and these now have the outside bubbles and the internal lines as common factors which we can cancel off to obtain Eq. (d) shown in Fig. 9. This is precisely Eq. (a) of Fig. 3 and the result confirms the structure in Fig. 7. We shall omit the uniqueness argument corresponding to that in Sec. 4.

5.2 Hermitian Analyticity

The above analysis suggests that in order to obtain the Hermitian analyticity relation for the middle bubble in Fig. 7, we must consider the part of the four-particle amplitude physical region above the four-particle threshold when the term [7] in Fig. 8 does not contribute. This happens when the external momenta are so prescribed that, in terms of the labeling of Fig. 7, particle 8 comes off with so much energy that the energy of the 234 system cannot produce another two particles. At the same time, the terms [5] and [6] in Fig. 8 cannot contribute, and this means that the pole structure (Fig. 7) is not present in this new part of the physical region.

Let us call these two parts of the $4 \rightarrow 4$ $+$ and $-$ amplitude physical regions $R_1^{\pm}$ and $R_2^{\pm}$ (in an obvious notation), respectively. They differ in that the values taken by the crossenergy variables of interest S_{2348}, S_{156}, and S_{348} are lower in R_2 than in R_1 and that the regions are separated by the S_{2348} two-particle normal threshold. We shall not assume that there is any path of continuation relating R_2^+ and R_2^-, i.e., that the $4 \rightarrow 4$ amplitude itself obeys any sort of Hermitian analyticity relation.

According to our analyticity postulate the $-$ amplitude can be continued from R_2^- to R_1^- by a path within the physical region. Suppose we continued the whole unitarity equation valid in R_2 so that the $-$ amplitude

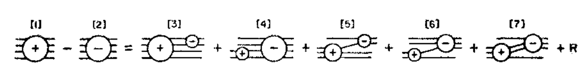

FIG. 8. This equation consists of those terms in the four-particle unitarity equation [(c) of Fig. 3] possessing the pole structure shown in Fig. 7. The terms in the equation are numbered for future reference.

[21] M. A. Rashid and A. Syed, Nuovo Cimento 28, 107 (1963).

is continued along this path while the $+$ amplitude follows a related path. For singularities not generated by the equation being continued, the $i\epsilon$ distortions in this related path are those appropriate to the $-$ rather than the $+$ amplitudes and so this path leads from R_2^+ to a region R_1^i on an unphysical sheet.

We know that the $-$ amplitude possesses the physical region pole structure of the type in Fig. 7 in R_1^-, but only know that the $+$ amplitude possesses this structure in R_1^+ but not necessarily in R_1^i. The sufficient condition that it does, and that the residues in R_1^i and R_1^+ are continuations of each other, is that there exists a path connecting R_1^+ to R_1^i lying in the section of the $4\to 4$ amplitude invariant space prescribed by the mass shell constraints $S_{156}=S_{348}=m^2$. We know that there is a path from R_1^+ to R_1^i, namely the physical region path from R_1^+ to R_2^+ followed by the path above from R_2^+ to R_1^i, and the question is whether this path can be distorted to lie within the mass-shell section.

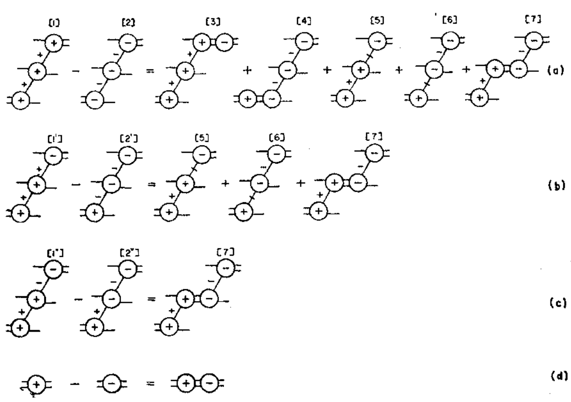

FIG. 9. These equations indicate the consequences of inserting the pole structure in Fig. 7, the corresponding one with $-$'s and those in Fig. 4 into the relevant terms of the four-particle unitarity equation valid above the s_{2348} two-particle threshold, namely, the terms appearing in Fig. 8.

We cannot answer this question in general since the $4\to 4$ amplitude singularity structure is not yet understood. In our "approximation" explained in Sec. 3, we retain only normal thresholds and the question is easily settled. The situation is represented diagrammatically in Fig. 10. Since S_{156} and S_{348} enter the problem symmetrically, we have simplified the picture by assigning a single axis to these variables. The known path from R_1^+ to R_2^+ to R_1^i runs over the S_{2348} two-particle normal threshold with a $+i\epsilon$ distortion and returns with the opposite distortion, and can easily be distorted to lie in the mass shell section when it just loops the normal threshold.

According to this distorted path the residue of the pole structure in R_1^i differs from that in R_1^+ in that the middle bubble is evaluated in a region reached when its energy variable (S_{2348}) encircles its two-particle threshold in an anticlockwise direction. We shall denote this

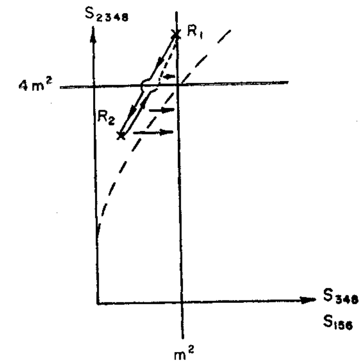

FIG. 10. Diagram representing the S_{2348}-joint S_{348}, S_{156} plane of the $+$ and $-$ $4\to 4$ amplitudes. The vertical and horizontal lines are the single-particle pole and two-particle normal thresholds, respectively. The broken parabolic arc represents the boundary of the physical region. The path with arrows is associated with the $+$ amplitude and runs from R_1^+ to R_2^+ to R_1^i. The large arrows indicate the deformation of this path to the mass shell section.

by placing an i in the middle bubble instead of a $+$. The unitarity equation valid in R_2 has not changed its form in the continuation to R_1. We now know which terms of the equation must possess the pole structure in R_1, namely terms [1], [2], [3], and [4] in Fig. 8. Also we have found the residues corresponding to each term and can equate the pole terms, obtaining Eq. (a) shown in Fig. 11. This differs from Eq. (a) of Fig. 9 in that i replaces $+$ on the middle bubble of the pole structure whenever it appears, the $i\epsilon$ prescription on the poles lines is absent and terms [5], [6], and [7] are absent. By the same manipulation which lead to Eq. (b) of Fig. 9 we obtain Eq. (b) of Fig. 11 and hence Eq. (c) of Fig. 11 which states that the i and $-$ amplitudes are equal. This is our desired Hermitian analyticity relation and states that if the $+$ amplitude is continued along the "path of Hermitian analyticity" round the two-particle threshold in a counterclockwise direction then the $-$ amplitude is obtained.

The result implies that the two particle threshold is a singularity and with Eq. (a) of Fig. 3 gives its discontinuity.

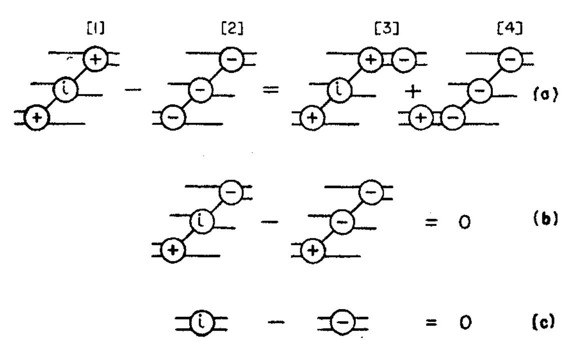

FIG. 11. These indicate the consequences of inserting the pole structures discussed in the text into a continuation of the four-particle unitarity equation valid below the s_{2348} two-particle threshold.

5.3 Generalizations

By studying more complicated versions of Eqs. (a) to (c) of Fig. 3 we could extend the analysis to inelastic two-particle scattering amplitudes. Then we could obtain the extended unitarity equations, infer the existence of normal threshold singularities below the physical threshold, and see that the path of Hermitian analyticity connecting the physical A_{ab} boundary value to the physical $A_{ba}{}^*$ boundary value runs down the energy axis with a $+i\epsilon$ prescription, encircles the lowest threshold, and returns with a $-i\epsilon$ prescription. In such cases it is known in perturbation theory that the path of Hermitian analyticity may be modified by the existence of anomalous thresholds. It is important that our method should be able to explain this possibility and, in fact, it can, but we shall not include this argument in the present paper.[22]

Sheer algebraic complication precludes an analysis of higher unitarity equations to obtain corresponding results for multiparticle amplitudes. In principle it is possible to derive generalized unitarity formulas for multiparticle amplitudes. From these there would follow, by methods of an established type,[3] equations for the individual single variable discontinuities across the normal threshold singularities lying at or below the physical thresholds in the energy and subenergy variables. We shall not discuss the form or properties of these important formulas here.[22] Nevertheless, we shall anticipate that in terms of the independence picture we can insert the newly found normal threshold singularities in the energy and subenergy planes and that the generalized path of Hermitian analyticity consists of a product of paths of the type described above in the planes corresponding to energy-like variables. We also expect the scalar and pseudoscalar amplitudes appearing in Eq. (3.1) to be separately Hermitian analytic.

A third sort of unphysical unitarity relation is that valid at a single particle threshold on the physical sheet of the $2\rightarrow 2$ scattering amplitude, and it tells us that this amplitude possesses a pole on the physical sheet with residue the product of coupling constants corresponding to the three-line vertices. The term in Eq. (c) of Fig. 3 which could produce such a pole in the middle bubble of the pole structure in Fig. 7 is the fifth drawn on the right-hand side, but we cannot apply our method unless we know that the $2\rightarrow 3$ production amplitude possesses a similar pole in its momentum transfer variable. This is tantamount to what we are trying to prove. We mention an alternative approach to this unphysical unitarity relation in the next section.

6. THE ANTIPARTICLE THEOREMS

We shall argue that the necessity for the existence of antiparticles, the validity of the substitution law for crossed processes and the TCP theorem follow from our postulates within our normal threshold approximation. We shall use an argument due to Gunson[5] who first suggested the derivation of the substitution law, and repeat it here because we wish to amplify it and because we feel he failed to realize its full significance when applied to the other points. He had accepted the arguments of Stapp[7] concerning the TCP theorem and this involved several assumptions which were unsatisfactory from our point of view.

[22] We hope to discuss this elsewhere.

6.1 The Existence of Antiparticles

According to relativistic kinematics, the energy of a particle is given in terms of its momentum by

$$p^0=(\mathbf{p}^2+m^2)^{1/2}. \qquad (6.1)$$

Consideration of the negative root leads, via the hole theory,[23] to the concept of antiparticle, but does not tell us that the antiparticle must, of necessity, exist in relativistic physics. It is the second quantization that forces a Lorentz invariant field operator to possess a negative energy part which creates states interpretable only as those of an antiparticle which carries physical energy momentum but quantum numbers opposite to those of the original particle. In the axiomatic formulation, weak local commutativity does this,[24] and moreover, leads to the TCP theorem. Local commutativity goes further, giving us forward tube analyticity in momentum space and the substitution law for crossed processes. In S-matrix theory we cannot just assume that antiparticles exist by analogy, or because asymptotic fields are causal, but must find the corresponding principle that forces them upon us. Since, in field theory, the theorems on antiparticles follow from the "causal" assumptions, we would expect analyticity to be the corresponding operative assumption in S-matrix theory. We shall find that, in as far as we can yet understand the singularity structure of the S matrix, this is indeed so, and conclude that the experimental detection of antiparticles is, in a sense, a partial verification of the analyticity postulate.

Assuming that our results on physical region poles can be reworked in more general theories, we see that there could occur the pole due to a certain particle in the $4\rightarrow 4$ amplitude with the residue indicated in Fig. 12(a). This pole occurs in what we shall refer to as the particle pole part of the physical region and is fixed in the appropriate cross energy variable. If the antiparticle existed, it would give rise to essentially the same pole in another part of the same physical region, the antiparticle part, with the residue indicated by Fig. 12(b). As we see from the diagrams the flux of

[23] P. A. M. Dirac, *Quantum Mechanics* (Oxford University Press, New York, 1958).

[24] S. Weinberg (private communication) has pointed out that the requirement that the field theory give rise to a Lorentz invariant S matrix implies the existence of antiparticles. It also implies local commutativity and hence analyticity.

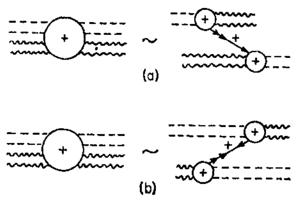

FIG. 12. The particle and antiparticle poles in a certain $4 \rightarrow 4$ amplitude are represented in (a) and (b), respectively. The straight, wavy, and dashed lines indicate particles with different quantum numbers. The single arrow represents the flow of quantum number and the double arrow the flow of positive energy.

positive energy is reversed with respect to the flux of quantum number in 12(b) as compared with 12(a), so that the antiparticle is a manifestation of the negative root in (6.1).

The question is whether the existence of the pole in the antiparticle region can be deduced from the existence of the pole in the particle region. There is no continuous transition from particle to antiparticle, that is, the path lying along the pole line passes through an unphysical region of the complete amplitude. Since poles persist in functions of many complex variables, the pole must persist in the antiparticle region, but not necessarily on the correct sheet, that unique sheet reached by a path within the physical region, which is distorted according to the physical region $i\epsilon$ prescription.

If it is possible to distort this physical region path so that it lies on the mass shell section then the pole certainly must exist in the antiparticle region and its residue must be a continuation of the residue of the particle pole. In general we have no means of arguing that such a distortion through an unphysical region is possible, but within our normal threshold approximation the answer is simply found.

The situation is represented diagrammatically in Fig. 13.

The normal thresholds in reactions other than that with the pole cannot impede the distortion since the path just slides along them. The only conceivable difficulty is that the path could encircle a normal threshold in the reaction with the pole. The $i\epsilon$ prescription assumption forbids this and since the path need only pass through values below the mass shell value in the pole invariant it need not even cross a normal threshold in this variable.

We have shown in Sec. 4 that physical region poles must represent particles to be included in the completeness relation, and this one in the antiparticle region must now represent the antiparticle of the original since it bears physical energy momentum but opposite quantum numbers. This interpretation is confirmed by the possibility of forming multi-antiparticle states by extension of the argument. The antiparticles, therefore, must be included in the completeness relation, and antiparticle amplitudes must be defined satisfying systems of unitarity equations so that the analysis of the previous section is applicable.

An interesting difference between this argument and the field theoretic one is that this necessarily depends upon a theory with interaction. Physically this is reasonable because in a universe consisting only of noninteracting particles, antiparticles could not be created.

6.2 The Substitution Law for Crossed Processes

Gunson has pointed out that if the residues of the poles in the particle and antiparticle regions are continuations of each other, then one might expect the individual factors to be related similarly so that the substitution law for crossed processes would result.

But this argument is not quite correct, and what does follow is that, if we consider all factorizations due to these poles we can relate the continued amplitude $A_{ab}{}^{c}$ and the crossed amplitude $\tilde{A}_{ab}$ in the following way:

$$\begin{aligned} A_{ab}{}^{c} &= \alpha \tilde{A}_{ab} \text{ if particle crossed is initially incoming} \\ &= \alpha^{-1} \tilde{A}_{ab} \text{ if particle crossed is initially outgoing.} \end{aligned} \tag{6.2}$$

α is a constant number depending only upon the particle crossed. We shall prove that it is a phase factor which can be taken as unity so that the crossed and continued amplitudes are indeed identical.

We shall understand that a superscript placed on an amplitude denotes that that amplitude has been continued along a path designated by that symbol. Thus, c refers to the path to the crossed physical region and is obtained from the path in the larger amplitude connecting the particle and antiparticle regions by taking suitable projection onto the relevant subspace of variables. $\bar{c}$ denotes the complex conjugate path to c.

In order to prove that α is a phase factor we first note that, if $A_{ab}{}^{c}$ and $\tilde{A}_{ab}$ have identical paths of Hermitian analyticity, then these amplitudes are related to the corresponding amplitudes for the inverse processes and

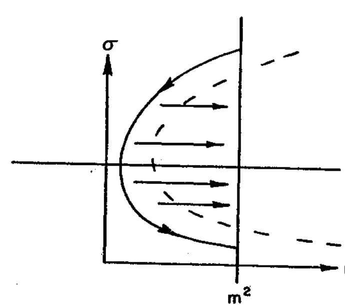

FIG. 13. u, the variable with the pole at m^2 is plotted against a typical variable σ. The broken arc represents the boundary of the physical region and the curved path with the arrows the physical region path running from the particle region to the antiparticle region. The large arrows indicate the deformation of this path to the mass-shell section.

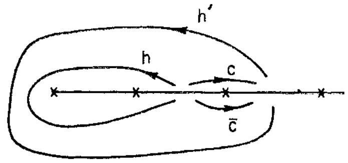

FIG. 14. The plane represents one of those in our independence picture corresponding to a variable taking energy-like values which crosses into one also taking energy-like values. h represents the path of hermitian analyticity for the original amplitude, c the path of continuation to the crossed physical region, and $\bar{c}$ the complex conjugate of this. h' is the path of Hermitian analyticity for the crossed amplitude.

a restriction on α results. The paths of Hermitian analyticity of $\tilde{A}_{ab}$ and $A_{ab}{}^{c}$ are h' and h_c and are defined by $\tilde{A}_{ab}{}^{h'}=\tilde{A}_{ba}{}^{*}$ and $h_c=c^{-1}h\bar{c}$ so that $A_{ab}{}^{ch_c}=(A_{ba}{}^{c})^{*}$. Supposing that the particle crossed is in state a, and comparing these equations with the relations (6.2) we find

$$|\alpha|^2\tilde{A}_{ab}{}^{h_c}=\tilde{A}_{ab}{}^{h'}. \tag{6.3}$$

According to our simplified picture of the singularity structure and the idea that generalized paths of Hermitian analyticity are products of single-plane paths of Hermitian analyticity, we see that $h_c=h'$, and so $|\alpha|=1$ by Eq. (6.3) (see Fig. 14).

This result can also be derived without such specific assumptions. According to (6.3) there are two different sheets of the crossed amplitude on which this function differs only by a positive real factor. This sort of structure seems unreasonable and would have to be the consequence of an un-unitarity-like discontinuity formula. Assuming that such behavior is not permissible we deduce $|\alpha|=1$, $h_c=h'$. This last relation with the definition $h_c=c^{-1}h\bar{c}$ implies that $h'=c^{-1}h\bar{c}$, an interesting general relation between paths relating different physical boundary values.

We now argue that α is an arbitrary phase factor which we are at liberty to choose as unity, and so finally obtain the substitution law for crossed processes. If we select a particular type of particle in the theory, and modify the theory by multiplying each S-matrix element by $e^{i\theta}$ each time it appears in an initial state and by $e^{-i\theta}$ each time it appears in a final state, then the equations of the modified theory, unitarity, and connectedness are formally identical with the original ones.

Such a transformation, really just a phase transformation of the particle creation operators, can be carried out independently of the antiparticle and be used to alter the phase of α, and, in particular, make it unity. The reason that this complication of having a particle-antiparticle relative phase does not usually occur in field theory is that there a natural phase for particle momentum states exists, that in which the wave function is simply e^{ipx} with no extra factor. We shall take it that this phase is fixed for all particles so that the substitution law for crossed processes always holds.

An extra step is necessary to obtain crossing for two particle amplitudes. Then two lines must be crossed simultaneously by considering a twin pole structure.

6.3 The *TCP* Theorem

The *TCP* conjugate of a given multiboson amplitude is constructed by "crossing" all incoming particles (or antiparticles) into outgoing antiparticles (or particles) bearing the same physical energy momentum, and treating all outgoing particles and antiparticles similarly. In bubble notation this means that along a given external line, the outflow of energy momentum is reversed while the outflow of quantum number is preserved. The values of the corresponding channel invariants and determinants of four vectors are the same since both are unchanged by complete reversal.

The *TCP* theorem states that an amplitude and its *TCP* conjugate are equal. Although this is contained within crossing, we shall obtain a picturesque direct proof by considering a four-pole structure in an $8\rightarrow 8$ amplitude, as shown in Fig. 15(a).

The complete $8\rightarrow 8$ amplitude is continued within its physical region to the part where the internal pole lines now correspond to antiparticles and where the invariants formed out of the momenta born by the internal lines assume their original values. This structure is illustrated in Fig. 15(b).

According to the results on physical region poles in Sec. 4, the residues are just the products of amplitudes indicated by the diagrams. In particular, the factor corresponding to the middle bubble in Fig. 15(b) is the *TCP* conjugate of that occurring in Fig. 15(a). We also know, by an extension of previous arguments, that the total residues are continuations of each other, and that, by our choice of particle-antiparticle relative phase, the crossed production amplitudes are continuations of the uncrossed ones. It follows that the *TCP* conjugate $2\rightarrow 2$ amplitude is a continuation of the original $2\rightarrow 2$ amplitude, the path of continuation C_{TCP} being determined by the path joining the pole

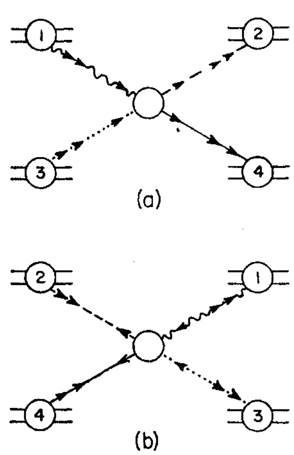

FIG. 15. Four-pole structures in an $8\rightarrow 8$ amplitude. In (a) the poles correspond to particles and in (b) to antiparticles. The single arrows denote flow of quantum number and the double arrows the flow of positive energy. Pluses have been omitted on the bubbles and pole lines.

structures in Fig. 15 in the $8 \to 8$ amplitude physical region. In order to return to its starting value in the space of channel invariants, as it does, the path C_{TCP} must pass over any normal thresholds twice, and must be distorted the same way each time because of the physical region $i\epsilon$ prescription in the $8 \to 8$ amplitude. Thus, C_{TCP} returns to the same point on the same sheet, and the TCP theorem follows, given our preliminary understanding of the singularity structure of the S matrix.

Since the determinants of four vectors are invariant under complete reversal they return to their original value in the continuation between particle and antiparticle pole structures, and the argument is generalizable to multiboson amplitudes. Similarly, if spin-$\frac{1}{2}$ particles were included and a γ-matrix formalism used, the same arguments would apply to the functions of invariants, and the spin part of the TCP theorem would be accounted for by the properties of the γ matrices.[2]

6.4 The Physical Sheet

As we prove crossing we can deduce the existence of new singularities. In particular, the existence of normal thresholds in those momentum transfer channels of the uncrossed amplitude which cross into subenergies follows from the unitarity and extended unitarity equations valid in the crossed channel. Thus, with the observations of Sec. 5, we have deduced the existence of all the normal thresholds singularities for continuum states and can insert them into our independence picture. By Hermitian analyticity and independence, they appear in both the upper and lower limits onto the positive real axis in each of the $2^{n-1}-n-1$ channel invariant planes.

Collecting our results on Hermitian analyticity, crossing, and TCP, we see that $2(2^{n-1}-n-1)$ different boundary values of a single analytic function onto this singularity structure, describe $4(2^{n-1}-n-1)$ different n-particle physical processes. In the independence picture the continuation between crossed physical regions is an appropriate movement along the real axis with a $+i\epsilon$ prescription.

The region which connects all these different physical boundary values is the product of cut planes and will be called the physical sheet. Its structure is a consequence of the original $i\epsilon$ prescription for physical region normal thresholds. When we have found the further singularities, we expect this region to have analytic properties simpler than those of any comparable regions.

By crossing the appropriate two lines in Fig. 4a we can deduce that A_{33} must possess a single-particle pole in the energy channel on the physical sheet as defined above. By considering the system of unitarity, or rather normal threshold discontinuity equations which link A_{22} to A_{33} it is possible to deduce that A_{22} possesses a stable pole on its physical sheet also. Because we must allow for the possible occurrence of multiple zeros and poles in the various amplitudes appearing the argument becomes too complicated for inclusion in the present paper.[22]

It seems surprising that such an apparently simple result is the consequence of such a roundabout argument, but in view of the comments at the end of Sec. 5 we have seen no alternative approach. Gunson[5] has demonstrated that it is also an extraordinarily powerful result.

All the derivations given in this paper can be repeated in the presence of all the normal threshold singularities. Thus, we have obtained a scheme in which all the normal threshold singularities are included but everything else ignored, and in which all the fundamental theorems can be derived. For this scheme to mean anything it must, at least, be self-consistent. Examples of consistency requirements concern the behavior of the individual normal threshold discontinuity formulas mentioned in Sec. 5, with respect to independence and crossing, but we shall not discuss here how they are indeed satisfied.[22]

7. DISCUSSION

Multiparticle features of the physical unitarity equations have enabled us to develop methods which relate the validity of the unphysical unitarity relations and the antiparticle theorems to properties of the S-matrix singularity structure. In each case the condition reduced to a question of distorting a certain well defined path in a certain way. We were unable to verify this in general because we do not yet know enough about the singularities involved. We considered a "model" or "approximation" taking only normal thresholds into account and found that the distortion was permitted and that this depended only on rather weak properties of the singularities such as independence and $i\epsilon$ prescriptions and not upon any detailed topological properties of the normal thresholds (which are only known for two-particle thresholds[3]).

It is possible to imagine unfavorable configurations, but we are encouraged by the fact that these are forbidden in physical regions, at least, by the part of the analyticity postulate which permits arbitrary distortion of a path therein.

It does not appear possible to deduce much of the S-matrix singularity structure from the physical unitarity equations without introducing the unphysical unitarity relations and antiparticle theorems. Since the specification of these results must involve paths of continuation passing through regions where there may lie singularities we cannot yet deduce, a severe difficulty arises which is characteristic of S-matrix theory rather than our particular method.

In the circumstances, the obvious and apparently only thing to do is to try to set up an iteration procedure whereby one takes into account only those singularities whose discontinuities we can deduce and, having derived the fundamental theorems in this approximation, use the theorems to generate more singularity structure, and repeat the process. The justification is to be self-consistency within the "approximation" at each step because, if we do eventually obtain a self-consistent solution it must be a correct solution.

Our results could be regarded as a first step in this direction. We have derived the fundamental theorems and in principle, all the normal threshold discontinuities in the step which takes only normal thresholds into account and are ready to use known methods[6,7] to deduce more singularity structure and repeat the process. So far we are encouraged by the fact that the method does seem to lead to self-consistent results within the normal threshold approximation and that it seems possible to proceed further in an unambiguous way.

In general, a procedure whereby we ignore and later include singularities does not make mathematical sense because of ambiguities in sheet structure. That it should work in this particular case would require that the S matrix singularity structure should have some special properties which we shall refer to as the hierarchical structure.[25] For example, it would seem that a path starting from a normal threshold singularity and looping some singularity other than a normal threshold must, on returning to its starting point find a new normal threshold on the new sheet. Such conclusions could be modified if we allowed a preferred sheet, the physical sheet. It is interesting that these requirements appear to be related to those mentioned above.

It is unsatisfactory to make the hierarchical structure an extra assumption since it is probably not independent of the unitarity and analyticity assumptions. Admittedly it is not clear either whether analyticity is independent of unitarity. A satisfactory possibility is that the hierarchical structure can be proved at the start.

Although one knows no details one knows that any given singularities of the S matrix must be generated by the unitarity equations in one of the three ways known: the change in form of the equations (which generates normal thresholds), the Stapp-Polkinghorne endpoint mechanism[6,7] (other Landau singularities, pseudothresholds, second type), or by integral equation methods[4,5,8,9] (unstable poles and cuts). This knowledge may be enough to deduce the desired information. If the hierarchical structure can be proved we obtain the important bonus that, since our method of construction seems unique, the solution so constructed is unique.

Let us suppose these surmises are correct and look at later steps in the iteration scheme. Since we know all the individual normal threshold discontinuities we can generate further singularities[6,7] and evaluate their discontinuities[6] by known methods. Various clues[26] have been found as to how the consistency requirements determine on which sheet the new singularities must lie. We must show that the solution is consistent at each stage, e.g., different unitarity equations can generate a given singularity and its discontinuity and ought, for consistency, to yield the same result. We mentioned that this was so for normal thresholds and Mandelstam[27] has verified this in the case of the square diagram. Since we appear to be reproducing a Landau singularity structure[11] with Cutkosky rules[12] for the discontinuities (but with the refinement that summation effects are automatically accounted for[4]) we can expect consistency in general because of the resemblance to perturbation theory which is consistent. Apart from these consistency requirements, questions of dynamical consistency may also enter which would determine the initial masses.

It is the opinion of the author that the results obtained here are favorable to the possibility of building up a complete, unique, and consistent theory from unitarity and analyticity postulates, but that the aim of further work must be to gain a deeper understanding of the analyticity assumption and of the hierarchical structure which seems to play a key role in unravelling the powerful consistency requirements of the S matrix.

ACKNOWLEDGMENTS

I would like to thank Professor G. F. Chew, Professor R. E. Cutkosky, and Dr. J. C. Taylor, my supervisor in Cambridge, for encouragement and help. I would like to thank Dr. I. T. Drummond, Dr. P. V. Landshoff, and especially Dr. J. Goldstone for interesting discussions and Professor H. P. Stapp for helpful criticisms and suggestions. I would also like to thank the DSIR for a maintenance grant during my stay in Cambridge when most of this work was carried out.

[25] We do not wish to imply the meaning Polkinghorne (Ref. 6) gave to this phrase although there is probably a relation between the two usages.

[26] J. C. Polkinghorne, Phys. Rev. **128**, 2898 (1962).

[27] S. Mandelstam, Phys. Rev. **112**, 1349 (1958).