International Journal of Theoretical Physics, Vol. 17, No. 10 (1978), pp. 747–852

Particle Order: A New Fundamental Concept in Hadron Physics

George Weissmann

Lawrence Berkeley Laboratory, University of California, Berkeley, Berkeley, California 94720

TABLE OF CONTENTS

0020-7748/78/1000-0747$05.00/0

To Mina, whose encouragement and support contributed so much to this work.

In Part One the concept of particle order is introduced in its simplest form, sequential order. We show how sequential order, together with the general principles of S-matrix theory, (1) reproduces the main predictions of quark models with respect to mesons, explaining their quarklike spectrum without any need to assume the existence of quarks as constituent particles; and (2) provides a theoretical foundation for duality, in particular the dual unitarization (topological expansion) approach; and hence provides a unified description for a broad range of mesonic regularities such as the origin of conserved internal quantum numbers, the nature of the mesonic spectrum, the OZI rule and its violations, Regge pole dominance, exchange and I-spin degeneracy and their breaking, the Pomeron, etc. In Part Two this scheme is extended to all hadrons by a generalization of the concept of order. Stringent self-consistency requirements are shown to determine the specific form of the theory, including such features as topological color, baryon number conservation, and a set of generalized OZI rules. The spectrum is shown to consist of nonexotic mesons and baryons, and well-defined classes of exotic particles, all obeying the zero-triality rule. An attempt is made to understand the physical interpretation of order, which may well turn out to be a most fundamental concept in hadron physics.

FOREWORD

This paper is addressed to particle physicists in general. Some familiarity with at least the basic concepts of S-matrix theory is required, but Section 2 of Part One recapitulates the main points needed to follow this paper in detail. Other than that, little specialized knowledge is needed, even about the dual topological unitarization approach out of which this work grew.

In correspondence with these goals, we have tried to avoid an excessively formal approach, and to bring out the intuitive content of our arguments as much as possible.

The material presented here constitutes original research except for Part One, Section 2, which is a review of some aspects of conventional S-matrix theory. Some of the results of Part One have already been published in summarized form as Sections 2 and 3 of a review article in *Physics Reports* (Chew and Rosenzweig, 1978). Some of the results of Part Two were developed in collaboration with G. F. Chew, J. Finkelstein, and J. P. Sursock, and have been published jointly (Chew et al., 1978), as well as reviewed in the above-mentioned review article as Chapter XIII.

A GENERAL ORIENTATION

It is the purpose of this paper to introduce a new fundamental concept into particle physics: the concept of *order*.

Order amongst what? Amongst the hadrons (incoming and outgoing) participating in a strong interaction scattering process. Thus, in a sense

that will gradually become clearer, a scattering process is an *ordered process*, with the participating particles standing in a well-defined order relationship.

How do we represent the order formally? Since the framework of *S-matrix theory* is general and elegant, and the only firmly established basis to stand on at present, we introduce order formally as an order between the particles of an *S*-matrix element (actually, *T*-matrix element, i.e., connected part). We represent the process order by a so-called *process graph* in which, roughly speaking, the vertices represent the particles of the process, and the "*edges*" (lines) connecting some pairs of vertices represent the neighbor relationships between the particles. (More precisely, a particle is represented by a connected *set* of vertices that takes the form of a tree-graph. But for baryons and mesons the above-stated version is true.) These edges will later be identified with the quark lines of quark diagrams. Each particle type has a characteristic vertex representation: e.g., mesons will be seen to correspond to two-vertices:

and baryons to three-vertices

etc. Each edge is *oriented*, and has one of three colors. Examples for a process graph are given in Figure 1:

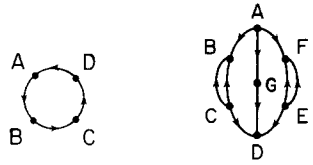
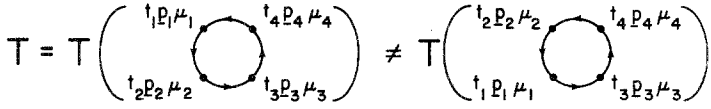


Fig. 1 (a, left) and (b, right). Here the letters *A*, *B*, . . . stand for particles of a scattering process.

To every ordered process there corresponds an ordered amplitude (by amplitude we mean specifically a connected part *T*, with the overall momentum energy δ function extracted), analogously as to every physical process there corresponds a physical amplitude representing its probability amplitude. Thus an ordered amplitude is a function of the usual particle degrees of freedom $(t_1\mathbf{p}_1\mu_1, t_2\mathbf{p}_2\mu_2, \ldots, t_n\mathbf{p}_n\mu_n)$ (where t_i stands for the particle type of the ith particle, $\mathbf{p}_i$ for its 3-momentum, μ_i for its helicity), but in addition

of the order between them, i.e., of their process-graph; thus, for example, we have



What is the logical structure of the theory we are proposing? It is in essence an ordered S-matrix theory: we postulate the existence of an ordered S matrix in the above sense, satisfying the basic postulates of Lorentz invariance, cluster decomposition, unitarity, and macrocausality (with appropriate and natural modifications for the presence of order). We also make the *bootstrap conjecture*, namely, that the ordered S matrix is uniquely defined in terms of these axioms. We do *not* need to make any specific assumptions about the spectrum of particles, or the kind of order we impose, or the presence of internal quantum numbers, etc. All this *follows* from implementing internal consistency between S-matrix properties and order. One can then prove that general S-matrix properties like discontinuity equations, normal analytic structure, dispersion relations, crossing, TCP, Hermitian analyticity and extended unitarity, Froissart bound, etc., hold also for the *ordered* S matrix. But, in addition, the theory yields a wealth of other properties, in particular duality (including exchange degeneracy, I-spin degeneracy, Regge pole dominance), and quark-model properties (zero-triality quark-model spectrum, conserved additive quantum numbers corresponding to flavors, baryon number conservation, generalized OZI rules, etc.) of the ordered S matrix.

Is the order of a process a measurable quantity? No. The particle parameters t_i, $\mathbf{p}_i$, μ_i continue to be the only measurable quantities. Hence the *physical* S matrix, which by its essence makes direct reference to observables, remains the same unordered object that we are familiar with.

Then what is the relationship between the ordered S matrix and the usual physical (unordered) S matrix? It seems that particle order is a characteristic of processes at a "deeper" level than the one we normally deal with in S-matrix theory (and in quantum mechanics in general), in which latter the degrees of freedom are all, in principle, measurable. Therefore, in order to get from ordered amplitudes (describing the statistics of scattering processes at a level where order is a meaningful concept) to the physical amplitude (describing the statistics at the observable level, where order has disappeared), we need, in some sense, to "average out the order." How is this done? A simple first approximation to the physical S matrix, the so-called *planar S-matrix approximation* (PSA), which is unordered and can thus be compared to experiment, can be constructed by adding (or, respectively, subtracting) all ordered amplitudes with a given set of particles. Its "planar" or "dual"

regularities (those regularities that go beyond those predictable from general S-matrix theory) have been approximately verified by experiment, and it seems that the PSA is quite a good approximation to the physical S matrix. However, the PSA is not exactly unitary, and thus is certainly distinct from the physical S matrix. A stepwise unitarization procedure, known as the *topological expansion*, is presumed to lead to the physical S matrix. (For the mesonic case of Part One, an ansatz for the topological expansion has been worked out in the context of dual unitarization, and applied with some success. For the general case of Part Two, an analogous procedure is being worked out by J. P. Sursock. We will not dwell on the topological expansion in this paper.) At a more practical and immediate level, order may also be seen as a device for achieving unitarity: the ordered amplitudes with their simple singularity structure are much easier to construct by bootstrap methods, and the topological expansion then yields the physical amplitudes with their full analytic complexity.

If order as such is not an observable, then what is the justification for its introduction as a fundamental concept? We feel that it is its promise to unify large parts of strong interaction phenomenology in a conceptually simple and economical scheme. The marriage of order and S-matrix theory appears capable of reproducing in a unified framework results heretofore only predicted in a piecewise fashion by a variety of models such as dual models or quark models. For mesons and their interactions most of the phenomenology has already been qualitatively understood on that basis, and for the other hadrons the outlook to do at least as well seems promising. Thus, although order as such disappears at the physical level, it leaves its ubiquitous trace in hadron physics. The quarklike structure and behavior of hadrons, in particular, may be understood as a manifestation of particle order, without the need to postulate constituent quarks as particles-that-are-not-really-particles, as in QCD.

How did the idea of particle order arise and develop? We were first led to consider it when we attempted to develop a unifying theoretical framework for the various calculational models for mesonic amplitudes known as "topological expansion" or "dual unitarization," which themselves grew out of the dual resonance model. We found that the concept of a *sequentially ordered S matrix*, corresponding to process graphs of the kind shown in Figure 1, lent itself ideally to that purpose. In Part One of this paper we describe this sequentially ordered S matrix and its properties. The close relationship between conventional (Rosner) mesonic quark diagrams and our process graphs became clear. The particles of the sequential S matrix were seen to be mesons. But all attempts to incorporate baryons into this sequential scheme failed.

Encouraged by the success of the sequentially ordered S matrix in

describing mesons, and guided by the relatedness of quark diagrams and process graphs for mesons, as well as by the failure of sequential order to accommodate baryons, we tried to construct an S matrix with a generalized order in which *any* graph could occur as a process graph. However, such a general theory was found to be incompatible with other S-matrix principles, in particular unitarity. Self-consistency forced us to restrict outselves to a certain class of graphs (so-called reducible graphs) as possible process-graphs, and also required the coloring of edges (with three colors). The outcome of this "order bootstrap" was a definite theory of a general ordered S matrix, which we present, including its derivation and properties, in Part Two. It represents the desired generalization of the sequentially ordered S matrix of Part One, in that it now also accommodates baryons and exotics. The process graphs of this general theory are still related to conventional quark diagrams, but they are more precise and richer in information in a crucial way; this helps to explain why the stubborn resistance of baryons and exotics to the conventional quark-diagram duality treatment yielded to our approach.

What is the heuristic significance, the physical interpretation, of particle order? At this point, we only speculate rather vaguely, that a "quark line" (i.e., edge) between two particles implies some form of direct causal relationship between the corresponding events. By events we mean the "actual" events such as sparks, bubbles, droplets, etc., that are *interpreted* as the manifestation of a particle, and indeed give rise to the abstraction of "particle." We will expand somewhat on this problem at the end of the paper.

PART ONE: THE SEQUENTIALLY ORDERED S MATRIX

1. INTRODUCTION

1.1. The Historical Background

1.1.1. The Dual Resonance Model. Ten years ago, a new approach developed in strong interaction physics that has continued to play an increasingly important role: duality. Its phenomenological applications have encompassed an ever-growing range of phenomena of a breadth unparalleled by any other approach. But its conceptual foundations have remained hazy, and its success somewhat mysterious. We believe that the concept of particle order that we introduce here in a systematic way is the crux of the duality approach. How did the concept of the ordered S matrix arise out of duality? To see this, we briefly review the development of duality, stressing those aspects that have a

bearing on our subject. We do not, in this Introduction, give specific references to original papers, but refer the reader to the review article (Chew and Rosenzweig, 1978) with its rich bibliography.

We begin our account in 1968, when Veneziano discovered the dual resonance model. He wrote down a specific functional form that represented a four-point function for bosons without internal degrees of freedom; this was soon generalized to an n-particle amplitude. These amplitudes displayed many of the properties required by S-matrix theory, including Lorentz invariance, analyticity (except at an infinite set of poles), pole factorization, crossing, and Regge behavior. In addition, the dual resonance model exhibited certain exact properties that corresponded to approximate regularities that were experimentally observed but never explained within the framework of S-matrix theory; we refer to regularities such as linearly rising Regge pole trajectories, absence of Regge cuts, and duality. The essential approximation that permitted one to write down the amplitudes in closed form was the narrow-resonance approximation: The only singularities of dual resonance model amplitudes are real (and hence zero-width) poles; normal thresholds and other singularities are absent. This is also the main deficiency of the model, causing it to violate unitarity in a manner that turned out to be very hard to remedy.

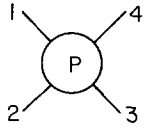
The Veneziano four-point function is the sum of three terms:

$$A(s, t) = A_{st}(s, t) + A_{su}(s, t) + A_{tu}(s, t)$$

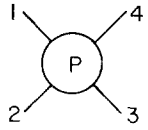
Each of these three terms can be represented as a β function, e.g.,

$$A_{st}(s, t) = -\beta \frac{\Gamma[1 - \alpha(s)] \cdot [1 - \alpha(t)]}{\Gamma[1 - \alpha(s) - \alpha(t)]}$$

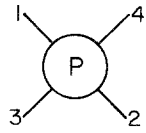
with $\alpha(s) = \alpha_0 + \alpha' s$, $\alpha(t) = \alpha_0 + \alpha' t$ and has a simple singularity structure: A_{st} has poles in the s and t channels but not in the u channel; A_{su} in s and u; and A_{tu} in t and u only. There is a formulation that represents this state of affairs nicely: the permutation ordered amplitude



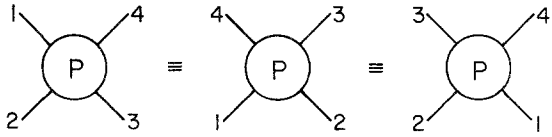
This symbol stands for that one of the three *β-fn* terms that has poles in $s_{12} \equiv (p_1 + p_2)^2$ and $s_{23} \equiv (p_2 + p_3)^2$, i.e., in those channels that are *adjacent* with respect to the cyclic ordering of the particles in



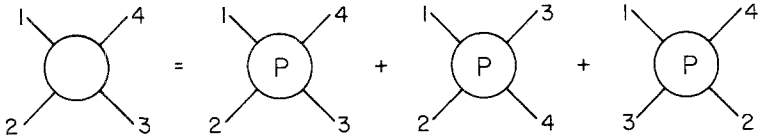
Similarly,



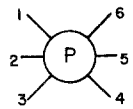
represents the term that has poles in s_{13} and s_{32} (but not in s_{12}), etc. Note that only the *cyclic* order is significant: thus



etc. In this formalism we write the Veneziano amplitude as



The n-point function can be similarly expressed as a sum of $(n-1)!/2$ cyclically distinct permutation-ordered amplitudes, each of which again has poles only in adjacent channels. For example,



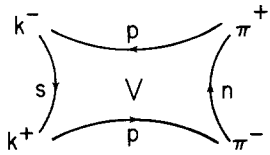
has poles in the $s_{12}, s_{23}, s_{34}, s_{45}, s_{56}, s_{61}, s_{123}, s_{234}, s_{345}$, channels, and only in these.

This representability of amplitudes as a sum of permutation-ordered amplitudes with singularities only in adjacent channels is a central feature of the dual resonance model that in and of itself guarantees its "duality properties," even when the narrow-resonance approximation is relaxed. Here, and in this whole paper, we mean by "duality properties" those properties of the dual resonance model that (a) go beyond what can be derived in general S-matrix theory and (b) do not depend on the narrow-resonance approximation. This includes properties like absence of Regge cuts, exchange degeneracy, quarklike internal quantum number structure, I-spin degeneracy, and others.

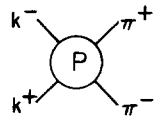
The internal (flavor) degrees of freedom, obeying $SU(n)$ symmetry, were incorporated into the model by Chan and Paton by providing each permutation-ordered amplitude with an appropriate multiplicative factor that explicitly exhibited its functional dependence on these internal degrees

of freedom. The resulting mesonic amplitudes now displayed exchange degeneracy, I-spin degeneracy (Regge pole trajectories with I spins 0 and 1 are degenerate), and an absence of exotic poles. Also, these amplitudes were interpretable in terms of the quark model. Since the Chan–Paton procedure of incorporating internal degrees of freedom with $SU(n)$ symmetry was proved to be unique, all the above-mentioned features can be considered as inherent in the dual resonance model.

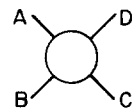
The Harari–Rosner duality diagrams or quark diagrams illustrate graphically the quark interpretation of dual amplitudes. They were originally developed to diagrammatically represent the Chan–Paton permutation-ordered amplitudes with internal degrees of freedom; but their quark interpretation is obvious, and was immediately noticed. An example of a duality diagram is



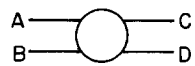
It corresponds to the permutation-ordered amplitude



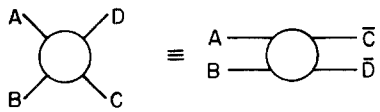
The V inside the quark diagram stands for "Veneziano" and serves to distinguish the dual resonance model amplitudes from the "planar amplitudes" that we will soon mention, and that are also represented by quark diagrams. Note that we are following the usual convention:



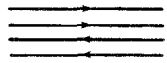
refers to an amplitude in which all particles are incoming, whereas



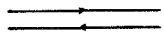
refers to A, B incoming, C, D outgoing. Thus, e.g.,



The fact that this amplitude can be written as a sum of $(p\bar{p})$ resonances in the K^+K^- channel, or of $(s\bar{n})$ resonances in the $K^-\Pi^+$) channel is nicely exhibited, as is the connection between absence of exotics and absence of singularities in the u channel: such exotic resonances would correspond to

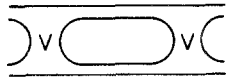


in contrast to the nonexotic mesons

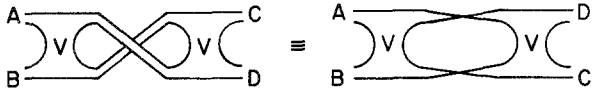


1.1.2. The Unitarization of the Dual Resonance Model: The Topological Expansion. Now while the dual resonance model amplitudes gave a nice qualitative picture of mesonic interactions, one obviously wanted to improve on it; the narrow-resonance approximation and the resulting lack of unitarity appeared to be the main source of error. In the years after 1968, attempts were made to unitarize the model while preserving its many desirable features. In analogy with Feynman perturbation theory, the dual resonance amplitudes were considered as Born terms (tree approximation) in a *perturbation expansion* whose higher terms (with loops) were to introduce resonance widths and normal thresholds. However, these higher terms turned out to be unrenormalizable, and hence the perturbation scheme appeared practically useless.

Nevertheless, Veneziano noticed that if one added together only the planar perturbation terms (such a term is called planar if the corresponding Feynman graph can be drawn on a plane without any crossing of lines, e.g.,

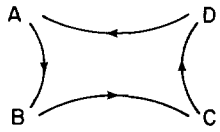


is planar, whereas

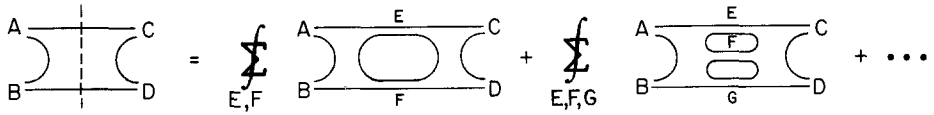


is not; obviously, this sum is to be regarded as "formal," since many of the individual terms are divergent), then the resulting "*planar amplitude*" still had the same general singularity structure as the permutation-ordered amplitudes of the dual resonance model (i.e., the Born term): they had singularities only in adjacent channels. We have already remarked that this is the essential property of the dual resonance amplitudes that we wish to preserve, since it is the basis of its "duality properties." But now these

singularities included thresholds, and resonances acquired a width. Indeed, these planar amplitudes, denoted here by



were shown by Veneziano to obey a unitaritylike discontinuity equation for adjacent channel cuts that he called "planar unitarity"; for example,



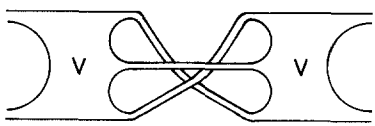
By $\sum\!\!\!\!\!\!\int_{E,F}$ we mean

$$\sum_E \sum_F \sum_{|\mu_E| \leqslant S_E} \sum_{1\mu_F| \leqslant S_F} \int \frac{d^3p_E}{(\mathbf{p}_E{}^2 + m_E{}^2)^{1/2}} \int \frac{d^3p_F}{(\mathbf{p}_F{}^2 + m_F{}^2)^{1/2}}$$

where E and F are summed over all particles, μ_E is the helicity of E, and m_E is the mass. Note that the order of the external particles (ABDC) is the same on both sides of the equation, and that internal particles do not "cross" like, e.g., this:

i.e., the terms of the right-hand side are planar. Each of the other nonplanar perturbation terms could similarly be categorized according to the topological structure of its dual perturbation graph. For example,

has a "cylinder" topology,

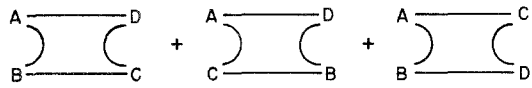


a torus topology, etc. When the total (physical) amplitude was expressed in terms of these topological terms (each of which corresponded to the infinite sum of all the perturbation terms with a particular topological characteristic), then the resulting expansion was called the *topological expansion.*

The advantage of the topological expansion over the perturbation expansion from which it was derived was that successive terms tended to be

reduced by a factor $1/N$, where N is the number of flavors, i.e., ~ 3. Chew and Rosenzweig later discovered a more important convergence mechanism related to singularity structure (Chew and Rosenzweig, 1978, Chapter 6). By contrast, analogous convergence mechanisms are absent in the original perturbation expansion owing to the largeness of the strong interaction coupling constant.

The sum of the planar amplitudes, e.g.,



(for the four-point function), represented the lowest-order term of the topological expansion; if, as expected from the presence of a convergence mechanism, this could be considered as a fairly good approximation to the physical amplitude, then the approximate "duality" of the latter finally found an explanation in the exact "duality" of the planar approximation. Furthermore, given the planar amplitudes one could in principle calculate all the other terms, and hence the physical amplitude. All this appeared quite encouraging. But how did one actually obtain the planar amplitudes? Their perturbation theory definition as an infinite series of planar unitarity products of Veneziano amplitudes was useless (owing to the divergence of the individual terms), except as a guide to their general properties.

Here was again the characteristic situation in strong interactions: owing to divergence problems, perturbation theory had to be considered as mathematical nonsense. And yet it could serve to elucidate general properties of the amplitudes, which could then be adopted, in bootstrap manner, as the starting point of a new theory. This is just what had happened in the late fifties when Feynman perturbation theory had helped to give birth to S-matrix theory. And here dual perturbation theory was to play a similar role again in the genesis of the new theoretical approach about to be described.

1.1.3. The Bootstrap Approach to Dual Topological Unitarization. An important step in this direction was taken by Chew in 1975: he suggested that the "planar unitarity" equations, together with analyticity and Lorentz invariance, and the OZI rule be regarded as *defining* the planar amplitudes, analogously to the bootstrap conjecture in S-matrix theory. The calculation of the planar amplitudes would thus be accomplished by solving, at least approximately, this "*planar bootstrap*." The planar bootstrap, particularly in its Reggeized form, promised to be simpler than the usual bootstrap for physical amplitudes; this was due in part to the simpler singularity structure in the j plane, such as the absence of Regge cuts, and also the many

degeneracies that reduced the number of Regge trajectories that had to be taken into consideration.

Assuming one could solve the planar bootstrap and calculate the planar amplitudes, then the topological expansion allowed the calculation of higher-order terms without solving any more nonlinear integral equations; this rendered the calculations practically feasible.

And even without solving the planar bootstrap, one could instead make a reasonable ansatz for it, and calculate the deviations from planar regularities such as the OZI rule, exchange and I-spin degeneracy breaking, and the Pomeron trajectory. And indeed, the phenomenological work along these lines has been eminently successful in describing qualitatively, and often quantitatively, most of the features of mesons and their interactions. What makes all this particularly impressive is the absence of any free parameters. These achievements have lent further evidence to the underlying ideas.

The reader interested in learning more about this dual topological unitarization (DTU) program, as we shall henceforth call it, is referred to the aforementioned excellent review article by Chew and Rosenzweig (1978).

The DTU program remained largely limited to mesons, attempts to include baryons having failed to a large degree. In Part Two, we will deal with this aspect of the problem in detail, and demonstrate how baryons can be included in DTU.

Having now summarized the historical background, we proceed in the next section to discuss the problems to which this paper addresses itself.

1.2. The Purpose of this Paper

At this point DTU, in spite of its phenomenological success in the meson sector, could hardly aspire to the status of a theory, but appeared more as a collection of models and calculational methods related to one another by a common theme.

The closest thing to a theoretical underpinning of DTU was represented by Chew's bootstrap approach based on the "planar unitarity" equations. And yet even this was still far from a coherent and self-contained theory. The problems began with the "planar unitarity" equations themselves. Without the background of dual perturbation theory (of which we now wanted to make ourselves independent), these seemed completely ad hoc; if they were to arise naturally out of the theory, it seemed that they should be understood as discontinuity equations derivable from some kind of unitarity relation. But which mathematical object was to be unitary? (In S-matrix theory it is the S matrix, and not the connected part that is unitary.) Were the planar amplitudes the connected parts of a sequentially ordered, unitary S matrix, and as such describable as permutation-ordered amplitudes? Or were they to be described as quark amplitudes? Or were maybe

permutation-ordered amplitudes by their very nature describable as quark amplitudes, so that these two concepts merged? Did planar amplitudes factorize? Did they obey crossing? What was their singularity structure, and the corresponding discontinuity equations? (Beyond pole and normal thresholds there was not even a hint.) Were planar amplitudes inherently mesonic, or could they be extended to baryons?

And some broader questions also posed themselves. Were planar amplitudes subsidiary quantities, derived from physical amplitudes by means of a particular approximation? Or were they fundamental quantities, independent from or even prior (in some sense) to the physical amplitudes? If so, what was their physical interpretation? What was the unifying physical principle underlying duality and quark structure?

These, and many other, unanswered questions made it painfully clear that a more fundamental approach to DTU, and duality in general, was called for.

As we worked our way through these questions, the mists of confusion gradually cleared, revealing a beautiful and simple picture: the whole structure of DTU appeared as a manifestation of a sequentially ordered, unitary S matrix. In particular, the quark structure of mesonic amplitudes could be understood on that basis.

Particle order emerged as the crucial new concept underlying the whole of duality and the quark idea, the key for a remarkable theoretical unification. In all its simplicity and power, it had to be a truly fundamental concept! And in it lay dormant the germ of a generalization that was to allow the extension of the DTU program to baryons and exotics, and possibly provide the basis for a better understanding of the nature of the quark concept, and eventually for a general theory of hadrons.

It is the purpose of Part One of this paper to introduce this new concept in the language of S-matrix theory; to construct a theory of the sequentially ordered S matrix, based on a set of basic postulates, and work out some of its properties; to show how it provides a theoretical foundation for DTU, and an alternative view of the quark concept; and to comment on some of its physical implications.

We will not, in this paper, dwell on the connection between the ordered and the physical S matrix (i.e., on the topological expansion), except to introduce the simple "planar S-matrix approximation" to the physical S matrix and describe some of its properties.

1.3. Outline of Part One

In Section 2 we recapitulate the main ideas of S-matrix theory as we need them.

In Section 3, we construct the theory of the sequentially ordered S

matrix in as close an analogy with usual S-matrix theory as possible. We see how the main S-matrix properties survive, sometimes in modified form, the step to the ordered theory. But in addition, the ordered S matrix exhibits the crucial features of dual models, and we are able to identify our ordered amplitudes with the "planar" amplitudes of DTU.

In Section 4, we develop a systematic understanding of the origin and role of internal quantum numbers in the framework of the usual, physical S matrix.

In Section 5, we show how the theory reproduces the standard quark-model predictions with respect to the mesonic sector, and how the formal identity of quark diagrams and process-order graphs allows us to reinterpret the quark concept in a new light. The OZI rule is seen to be a feature of ordered amplitudes.

In Section 6, we present the so-called planar S-matrix approximation, a set of unordered amplitudes obtained by summing ordered amplitudes over all orders, and find that they represent a reasonably good approximation to the physical S matrix; their main drawback is found to be their lack of exact unitarity.

2. A BRIEF REVIEW OF S-MATRIX THEORY

2.1. Introductory Remarks

S-matrix theory was developed in the late fifties and sixties, when it had become clear that quantum field theory, at least in the only form that people knew how to use it, namely, perturbation theory, was unable to deal with strong interactions. S-matrix theory was conceived as a child of quantum field theory (QFT), distilled out of a few basic properties of the QFT S matrix that were generally considered beyond dispute, without being weighed down with many assumptions of a more questionable nature that form the basis of QFT. It grew into a logically well-structured (axiomatized), beautiful, and general theory that provides essentially the only firm theoretical ground in strong interactions, besides providing an ideal conceptual framework within which to construct a self-consistent interpretation of quantum mechanics (Stapp, 1971). It is the only theory that unites quantum mechanics and special relativity in a self-consistent fashion. Moreover, it reduces to the Schrödinger theory in the nonrelativistic limit (Blancenbecler et al., 1960) and to the Feynman–Dyson theory in the weak-coupling limit (Chou and Dresden, 1967). Many subsidiary properties, and relationships between measurable quantities, were derived and experimentally confirmed. But since there was no obvious way to use the basic principles in order to numerically compute the S matrix, the question arose as to whether S-matrix theory

was actually a complete theory (which would imply that one could, at least in principle, use it to compute the S matrix numerically), or whether it was rather a conceptual framework, a language, but still lacking some dynamical principle (like the Lagrangian in field theory) to make it complete. The bootstrap conjecture takes the first stand by maintaining that the principles of S-matrix theory so restrict the possible structure of the S matrix as to admit only one solution: the actual strong-interaction S matrix. For example, unitarity (conservation of probability) ties all different regions of the S matrix inextricably together, and analyticity (macrocausality) determines a scattering function completely, given its value near any one point. Indeed, no model S matrix has been found that conforms to all of the basic requirements; if it were, and had *any* adjustable parameters, the bootstrap conjecture would be disproved. [The bootstrap version of S-matrix theory also led to a philosophy of science, new to the Western world, that stressed the role of wholeness, and self-consistency, as opposed to the traditional emphasis on analysis of a system into irreducible parts evolving in time according to fundamental differential laws (Chew, 1971).] Thus the spectrum of particles and their masses, spins, and internal quantum numbers, as well as all coupling constants, and indeed all scattering amplitudes, would follow from first principles. From the sheer complexity of the hadron spectrum it was apparent that an exact general solution of this bootstrap program surpassed human calculational ability. However, by relaxing one or the other of the basic postulates, it was possible to construct calculational models that gave good numerical results for those regions of the S matrix where that approximation was appropriate. The resulting situation is an ever-growing and often overlapping patchwork of approximate models, linked together to a coherent whole by the underlying S-matrix theory on which they are all based.

One stain in this picture developed, however, and contributed to the increasing isolation of S-matrix theory from the mainstream of particle physics: its inability to make any but a few very limited predictions about the spectrum, especially with respect to internal quantum numbers. Thus although the set of observed conserved internal quantum numbers, as well as the observed spectrum of particles, could be inserted by hand, it was done in a completely ad hoc manner. In *practice*, S-matrix theory seemed to provide few restrictions on the spectrum, far less to determine it. Baryon number, charge, strangeness, charm and their conservation, and distribution in the spectrum, and the approximate properties of $SU(a)$ multiplet structure, the OZI rule, and the duality properties could of course all be expressed in S-matrix language, but they did not seem to follow from S-matrix principles. By contrast, a field-theoretical approach that developed in the sixties, based on the idea of a small set of truly elementary particles, called quarks, of which all hadrons are considered to consist, was able to predict some of

these properties though it yielded few other dynamical predictions. However, it remains riddled with difficulties and internal paradoxes, and unless these can be resolved, it cannot aspire to the status of a self-consistent theory. One of the main paradoxes is that quarks enter the theory as particles; then a complicated set of calculations, valid only on that premise (and not actually performed till now) is to show that these quarks are "confined," i.e., *not* poles of the S matrix, and hence *not* particles, thus invalidating the original premise and calling the above calculations into question. And yet even "naive" nonrelativistic quark models, indeed especially these, yielded results that seemed beyond the calculational scope of S-matrix theory. This fact, coupled with deeply ingrained prejudice in favor of a theory based on a space-time framework and the idea of irreducible, elementary constituents, allowed the constituent quark idea to conquer the hearts and minds of most particle physicists, at the expense of the bootstrap S-matrix idea. Their general hope was that somehow a self-consistent, well-defined quantum field theory of elementary, contained quarks could be developed. The latest attempt in that direction is quantum chromodynamics (QCD); and while judgment should be withheld pending the outcome of the current effort, the internal paradox mentioned before seems built into QCD, together with the usual difficulties of any quantum field theory away from the weak-coupling limit. [The above-mentioned prejudice stems from our macroscopic experience, where the concepts of space, evolving time, and the constituent idea (A consisting of B and C) all make perfect sense; the evolution of macroscopic physics in the past 300 years, based, with so much success, on these concepts, has only led to the hardening of that prejudice. Most physicists, therefore, feel that the S matrix is not the whole story, that there must be an underlying space-time "mechanism." However, a closer examination tends to discredit this approach, and lends weight to the bootstrap approach (Iagolnitzer, 1978).

In this paper we pursue a different approach to the duality and quark properties of hadron that is more in line with the S-matrix approach. As already mentioned in the general orientation, the introduction of particle order into the S-matrix framework holds out the realistic promise of a satisfactory theory of strong interactions.

If this program succeeds (and QCD fails), has then S-matrix theory carried the day? The answer is but a qualified yes. For by introducing the mysterious and unobservable concept of order, we have subtly changed the very spirit of the theory. We have not merely added one postulate to further restrict the structure of the S matrix. The ordered S matrix is a *new object* with a richer structure, represented by a larger Hilbert space, the Hilbert space of ordered states. And very much in contrast to the physical S matrix, an ordered S-matrix element does not refer directly to any physical process,

but apparently to some underlying ordered process. [A physical process, as we use the word, is characterized by specifying fully the experimental apparatus, and the preparations and measurements made with it.] It is as if we had opened the door a crack to a hidden realm where this order is a meaningful concept. The realm itself is still shrouded in mystery, but we believe that there lie discoveries waiting to be made that may revolutionize our way of thinking about quantum mechanics and its significance.

It might well be that subsequent, more matured versions of the ordered theory will renounce the use of S-matrix terminology altogether in favor of a more basic set of concepts applicable at that level.

2.2. The General Framework

We now present a very brief review of those aspects of S-matrix theory (SMT) that the reader should be familiar with to follow the rest of the work. We sketch the axiomatic structure, introduce some important concepts and notation, and describe various properties of the S matrix (SM) that can be derived. The S matrix referred to here is of course the usual, physical S matrix. The procedures outlined here will then, in the following sections, serve as a guideline for constructing an analogous theory for the sequentially ordered S matrix. The reader interested in a more thorough presentation of the material of this chapter is referred to Iagolnitzer (1978).

In a physical scattering experiment m (usually 2) "particle preparation devices" and n "particle detecting devices," all of them macroscopically separated, interact in eventlike fashion. As the terminology suggests, one may imagine that these devices actually prepare or, respectively, detect material particles, although this is only a figure of speech; the reason it is chosen is that the ith device (more precisely, its effect on the event probability) can be mathematically described by assigning it particle degrees of freedom: t_i, $\mathbf{p}_i$, μ_i. This description is effected by representing the ith device by a "state vector" in a Hilbert space spanned by the basis vectors $|t_i, \mathbf{p}_i, \mu_i\rangle$. Each particle type t corresponds to an irreducible representation of the Poincaré group, and thus determines a mass m and a spin s. The momentum p and the helicity μ are the representation parameters. To avoid introducing density matrices, we restrict ourselves here to "pure" measurements and preparations. A device preparing (or measuring) a pure state is often assumed to produce only one type of particle; hence the general state vector corresponding to it is of the form

$$\sum_{\mu} \int \frac{d^3p}{2E} \psi_{\mu}(p_i) |tp\mu\rangle$$

note that there is no summation over t (superselection rules). In practice

this mapping of the experimental device onto its mathematical representation is determined by experimental calibration (Stapp, 1971).

The whole set of preparation devices (or, respectively, measuring devices) is represented by a vector in the direct product space of the individual particle Hilbert spaces, the "in" ("out") state vector: $|i\rangle$ ($|f\rangle$). The statistical correlations of this whole experimental system are described by means of a bounded linear operator S, defined on the Hilbert space of "in" states (and hence also on that of the "out" states). The probability of measuring $|f\rangle$ if we have prepared $|i\rangle$ is given by $|\langle f|S|i\rangle|^2$.

If the ith device is translated by the 4-vector x, then the state vector representing it, say,

$$\sum_{\mu} \int \frac{d^3p}{2E} \psi_{\mu}(\mathbf{p}) |t\mathbf{p}\mu\rangle$$

has to be replaced by

$$\sum_{\mu} \int \frac{d^3p}{2E} \exp\left(-\frac{i}{\hbar} p \cdot x\right) \psi_{\mu}(p) |t\mathbf{p}\mu\rangle$$

The above S-matrix framework is the essence of quantum mechanics; its validity has been established beyond reasonable doubt. On it one imposes four basic physical postulates to obtain the axiomatic foundations of SMT. These four postulates represent principles that are usually considered to be beyond dispute, so that the resulting theory may be considered as a firm bedrock in the otherwise rapidly shifting scenery of particle physics.

Before we introduce these postulates in Section 2.3, we comment on the question of the order in which particles occur in an SM element: this order is clearly physically irrelevant, since it merely corresponds to the arbitrary way that we have numbered the devices. For example, $|t\mathbf{p}\mu\rangle$ may be the state vector corresponding to preparation device number 1, and similarly $|t'\mathbf{p}'\mu'\rangle$ the state vector corresponding to preparation device number 2. We make the convention that $|t\mathbf{p}\mu\rangle \otimes |t'\mathbf{p}'\mu'\rangle$ denotes this state: the first vector in the product describes device 1, the second device 2. Then the vector $|t'\mathbf{p}'\mu\rangle \otimes |t\mathbf{p}\mu\rangle$ obviously describes a situation in which the device preparing $|t\mathbf{p}\mu\rangle$, namely, the one we called "number 1" before, has now been relabeled "number 2" and vice versa. Equally obvious, this relabeling does not affect the physical situation at all, so that the two state vectors $|t\mathbf{p}\mu\rangle \times |t'\mathbf{p}'\mu'\rangle$ and $|t'\mathbf{p}'\mu'\rangle \times |t\mathbf{p}\mu\rangle$ must be physically equivalent.

This in turn implies that any two SM elements differing only by such relabelings must be equal to within a phase factor, which can be shown to be ± 1. Introducing the notation

$$S(t_1\mathbf{p}_1\mu_1, t_2\mathbf{p}_2\mu_2; t_1'\mathbf{p}_1'\mu_1', t_2'\mathbf{p}_2'\mu_2') \equiv \langle t_1\mathbf{p}_1\mu_1| \times \langle t_2\mathbf{p}_2\mu_2|S|t_1'\mathbf{p}_1'\mu_1'\rangle \times |t_2'\mathbf{p}_2'\mu_2'\rangle$$

we can write

$$S(t_1\mathbf{p}_1\mu_1, t_2\mathbf{p}_2\mu_2; t_1'\mathbf{p}_1'\mu_1', t_2'\mathbf{p}_2'\mu_2') = \pm S(t_2\mathbf{p}_2\mu_2, t_1\mathbf{p}_1\mu_1; t_1'\mathbf{p}_1'\mu_1', t_2'\mathbf{p}_2'\mu_2')$$

According to the spin-statistics theorem of S-matrix theory, the minus sign applies if both t_1 and t_2 are fermions, the plus sign otherwise. For pairs of identical particles, $t_1 = t_2$, this leads to the usual Fermi and Bose symmetries.

Because of this *permutation symmetry* of the S matrix, the appropriate multiparticle Hilbert space is not the full direct product space of the one-particle Hilbert spaces, but instead an appropriately symmetrized (or, respectively, antisymmetrized) contraction thereof, the Fock space.

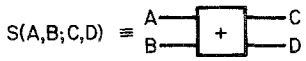
The origin of this permutation symmetry is that our language is too rich, and allows more degrees of freedom than the physical situation warrants. Specifically, the notation $S(t_1\mathbf{p}_1\mu_1, t_2\mathbf{p}_2\mu_2, \ldots)$ permits a distinction between "1" and "2" that has no counterpart in the physical system, which is characterized by the values of all the $t\mathbf{p}\mu$'s without attaching any index labels.

On the other hand, we will find the formalism of S-matrix theory (SMT) ready-made to describe the sequentially ordered SM, where the order is significant. It is this perfect fit between the formalism and what it represents that makes many of the equations wieldier in the ordered theory.

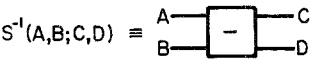
We now introduce some notation, and a few definitions. We will often let a complete particle specification $t\mathbf{p}\mu$ be denoted by letters such as $A, B, C, \ldots$; e.g., $S(A, B; C, D)$.

A *channel* τ is an (unordered), finite set of particle types: $\tau = \{t_1, t_2, \ldots, t_n\}$. It is sometimes also used to denote the set of all states constructable from the particles of the set τ.

A popular graphical notation for SM elements is the bubble notation:



and



When we do not wish to specify each particle separately, we write

About the set of all particles, the *spectrum* of the SM, we assume the following: since we are concerned with strong interactions, and hence with hadrons, there should be no zero-mass particles. The spectrum is denumerable, so that the particle-type index t may be considered an integer index.

And there is to be no accumulation point in the mass; i.e., in every finite mass-interval there are assumed to be only a finite number of particles. Initially, the S matrix is defined only in terms of stable particles; but one can then extend the definition to include unstable particles.

Having sketched the general framework, we now describe the basic postulates on which SMT is built.

2.3. The Basic Postulates

2.3.1. Poincaré Invariance. The transitions $|\langle f|S|i\rangle|^2$ of an experiment do not depend on the particular inertial frame in which it is performed: if every device is acted on by the same Poincaré transformation, the probability remains unchanged; this makes the theory conform to the special relativity principle.

One can prove that the matrix elements themselves are invariant under Poincaré transformations (without acquiring a phase factor): $\langle f|S|i\rangle = \langle f_L|S|i_L\rangle$.

Introducing a linear unitary operator L that represents the Poincaré transformation on the Hilbert space of state-vectors, we can write $\langle f|L^+SL|i\rangle = \langle f|S|i\rangle$; $LS = SL$.

Amongst the consequences of this is the conservation of 4-momentum and angular momentum. In terms of $S(t_1\mathbf{p}_1\mu_1, t_2\mathbf{p}_2\mu_2, \ldots; \ldots t_n'\mathbf{p}_n'\mu_n')$ the former implies the presence of an overall $\delta^4(p_1 + p_2 + \cdots - p_1' - p_2' - \cdots - p_n')$ factor.

Another important consequence is that an n-particle SM element can be expressed as a function of $3n$-10 relativistically invariant variables of the Mandelstam type [e.g., $S_{ij} = (p_i + p_j)^2$, $S_{ijk} = (p_i + p_j + p_k)^2, \ldots$], rather than of $3n$ 3-momentum component variables.

For strong interactions one postulates invariance under the *extended* Poincaré transformations, i.e., including space and time reflection; this implies parity conservation and the time symmetry of the SM.

2.3.2. Unitarity. The physical principle underlying this property is the "conservation of probability":

$$\sum_n \langle i|S|n\rangle|^2 = I$$

where $\sum_n$ denotes summation and integration over all possible states $|n\rangle$; more explicitly,

$$\sum\!\!\!\!\!\!\int = \sum_{N=2}^{\infty} \prod_{i=j}^{N} \sum_{t_i} \sum_{\mu_i=S_{ti}}^{S_{ti}} \int \frac{d^3p_i}{(m_{ti}^2 + p_i^2)^{1/2}}$$

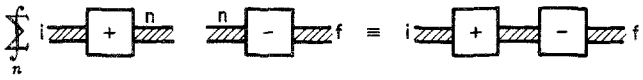
That is, if we prepare $|i\rangle$, we have to find *some* $|n\rangle$ (the basis corresponding to our detecting device) upon measurement: the probabilities have to add

up to unity. This, together with the linearity of S, is easily seen to imply the unitarity of the S operator: $S^+S = SS^+ = 1$, or in terms of SM elements

$$\sum_h \langle i|S|n\rangle\langle n|S^+|f\rangle = \sum_h \langle i|S^+|n\rangle\langle n|S|f\rangle = \langle i|f\rangle$$

The state $|n\rangle$ is called the intermediate state.

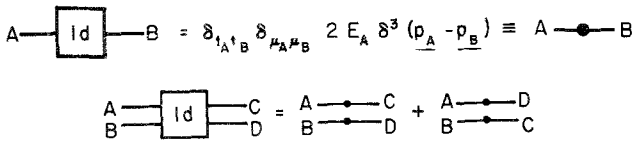
In the bubble notation the following abbreviated notation is used for the unitary product:



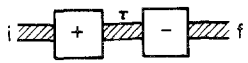
With this notation, unitarity reads:

Here

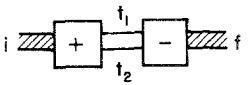
is the identity operator on the Hilbert space of states; for example



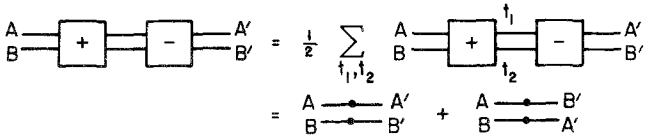
etc. Often one is interested in a partial unitarity product where the intermediate state is restricted to a particular channel τ. This is denoted by



or if we make the channel explicit, e.g.,



For example, for an elastic $2 \to 2$ process below the three-particle threshold, unitarity reads:



The symmetry factor of 1/2 has to be included to avoid overcounting. We mention this explicitly because this feature will disappear for the ordered SM.

This seemingly innocuous postulate of conservation of probability is of great power, imposing stringent conditions on the possible forms of the SM and implying many secondary properties of importance. It leads to nonlinear integral equations that provide a mathematical basis for the bootstrap.

2.3.3. Cluster Decomposition. We now introduce the *independence postulate* that two (or more) scattering experiments, set up at a great distance from one another (in space, or in time, or in both), should not influence one another appreciably. Specifically, assume a system involving m preparation and n detection devices. Now select m_1 of these preparation devices, and n_1 of the detection devices (without loss of generality, we can assume that it is the first m_1 or, respectively, n_1 of them). We say that we have divided the devices (or, respectively, the particles) of the system into two *clusters*, cluster 1 consisting of the above-mentioned m preparation and n detection devices (particles), and cluster 2 of the remaining $m_2 = m - m_1$ preparation and $n_2 = n - n_1$, detecting devices (particles). The corresponding state vectors are denoted by $|i_1\rangle$ and $|f_1\rangle$ for cluster 1, and $|i_2\rangle$ and $|f_2\rangle$ for cluster 2. In order to express the *independence postulate*, we now translate cluster 1 by the 4-vector σa, where a is an arbitrary 4-vector, and σ is a real number that we will take to infinity. Denoting the thus translated state vectors by $|i_1^{\sigma a}\rangle$ and $|f_1^{\sigma a}\rangle$, the independence postulate reads:

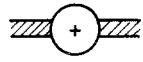
$$\lim_{\sigma\to\infty} \langle i_1^{\sigma a}| \otimes \langle i_2|S|f_1^{\sigma a}\rangle \otimes |f_2\rangle = \lim_{\sigma\to\infty} \langle i_1^{\sigma a}|S|f_1^{\sigma a}\rangle\langle i_2|S|f_2\rangle$$
$$= \langle i_1|S|f_1\rangle\langle i_2|S|f_2\rangle$$

where the last step is due to translation invariance. (Actually, one can only infer

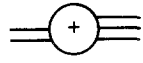
$$\lim_{\sigma\to\infty} |\langle i_1^{\sigma a}|i_2|S|f_1^{\sigma a}|f_2\rangle| = |\langle i_1|S|f_1\rangle||\langle i_2|S|f_2\rangle|$$

but it can be shown that this implies the said equation, without phase factors.)

Owing to the validity of this postulate, the introduction of the so-called *connected parts* S_c of the SM turns out to be very useful. In the bubble-diagram notation they are denoted by

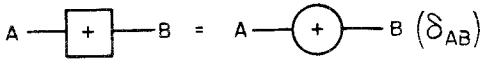


or, respectively,

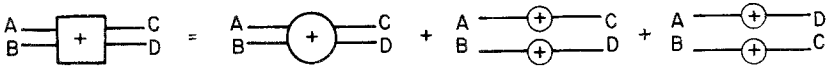


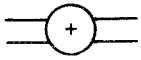
The connected parts are implicitly and recursively defined by the *cluster-decomposition equations* that break up an SM element in all possible ways into connected-part products.

For example,

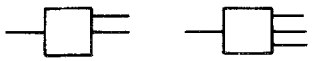


which defines 

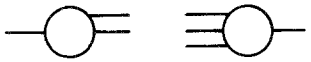


which, in conjunction with the above equation defines 

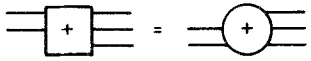
etc. Note that because of the stability of particles, objects like these:



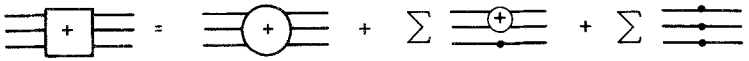
and these:



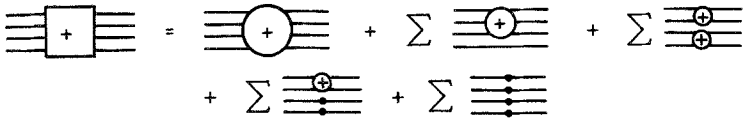
etc., do not occur. Thus, for example,



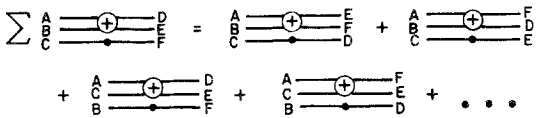
More complicated examples of cluster decomposition equations are



and

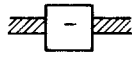


where the sums are over all topologically different ways of connecting the fixed external particles to a set of bubbles: e.g.,



For fermions one must exercise caution with the sign of a given term (for each crossing of fermion lines there is a minus sign).

For



there is an analogous cluster decomposition into connected parts:

The usual convention here is to include a factor $-I$ in the definition compared with the analogous

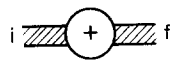
equations.

These connected parts have the important property that when a subset (cluster) of its particle states is translated to infinity as above, then the matrix element tends to zero:

$$\lim_{\Sigma \to \infty} \langle i_1^{\sigma a}| \otimes \langle i_2|S_c|f_1^{\sigma a}\rangle \otimes |f_2\rangle = 0$$

implying that except for an overall momentum-energy conserving $\delta^4(\ )$ factor, connected parts no longer contain δ functions or their derivatives, and are thus true functions, not merely distributions. These functions (i.e., the connected-part matrix elements with the d^4 factor removed) are called *T-matrix elements*, *scattering functions*, or simply *amplitudes*.

Cluster decomposition and connected parts have an interesting interpretation:

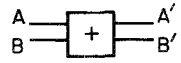


describes the probability amplitude that the physical process $|i\rangle \to |f\rangle$ takes place as a unified, causally connected process. By contrast,

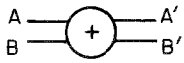
describes the probability amplitude that the physical process takes place at all, whether as a causally connected process, or in several clusters of mutually disconnected processes. The cluster decomposition equations simply express the well-known quantum mechanical principle that when a process

can be thought of as occurring in different ways ("paths"), then the amplitude of the process is the sum of the amplitudes corresponding to these different ways. Here the different "ways" are just the various causal connectedness patterns in which the scattering event can take place.

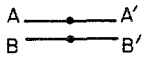
It is noteworthy that this concept of causal connectedness is not strictly speaking an observable property: when the parameters of a scattering process are such that it can actually take place in more than one connectedness pattern, then it is impossible to determine by measurement whether a particular individual event took place in one pattern or the other. For example, if in the $2 \to 2$ elastic scattering process



the wave functions $\psi^{(A)}(\mathbf{p}_A, \mu_A)$ and $\psi^{(A)}(\mathbf{p}'_A, \mu'_A)$ have a common support (i.e., there exists a p and a μ such that $\psi^{(A)}(\mathbf{p}, \mu) \neq 0$, *and* $\psi^{(A')}(\mathbf{p}, \mu) \neq 0$), and similarly for $\psi^{(B)}$ and $\psi^{(B')}$, then the process can proceed both through



and through



and there is no way to decide which one it was for any *individual* event.

But in spite of this nonobservability of connectedness, it is an extremely useful concept. We mention this point particularly because we will soon introduce another nonobservable concept about which very similar remarks can be made: particle order.

Although it is the S matrix that has a direct physical interpretation and hence is unitary, it is nevertheless the T matrix that has the simple, elegant properties (e.g., analyticity, crossing, pole factorization, etc.) and plays a central role in SMT. Indeed, once the unitarity of the S matrix has, with the help of cluster decomposition, been recast into discontinuity formulas for the T matrix, the S matrix practically vanishes from the scene.

2.3.4. Macrocausality and the Normal Analytic Structure. The principle of macrocausality states that all transfers of energy-momentum over macroscopic distances that are not ascribable to stable particles (in accordance with classical ideas) give effects that are damped exponentially with distance.

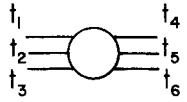
To express this more precisely, let us regard again the same situation as we did for the definition of connected parts: a scattering system divided

into two (or more) clusters, and let us translate cluster 1 by σa. Then we saw that

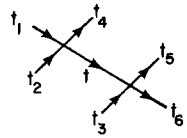
$$\lim_{\sigma \to \infty} \langle i_1^{\sigma a}| \otimes \langle i_2|T|f_1^{\sigma a}\rangle \otimes |f_1^{\sigma a}\rangle = 0$$

Now macrocausality states that this convergence (falloff) is exponential, unless the cluster parameters, in particular the momenta, are such that a particle could classically be exchanged between the clusters. A particularly simple example is a $2 \to 2$ (non-forward-elastic) scattering: if the two "in" beams are translated in such a direction that they no longer intersect (classically), then the scattering amplitude has to fall off exponentially with the translation parameter of σ. [To be precise, the width of the Gaussian wavepackets employed (in momentum space) should be shrunk by a factor of $\sigma^{1/2}$ at the same time as we translate one cluster away by σa, for reasons too involved to explain here. This amounts to taking $\hbar \to 0$, and thus to the classical limit.] This fact is closely related to the use of the Yukawa potential in the nonrelativistic approximation.

We have seen that a less than exponential decay of the amplitude (as the cluster separation is increased) is only possible if one or more particles can be exchanged between the clusters; this is also true if there are more than two clusters. Now, for a given "in" and "out" channel and given clustering of these particles, such an exchange of particles between clusters, as $\sigma \to \infty$, is possible only for those points in momentum space for which the corresponding classical multiple scattering process is possible. For example, for



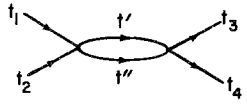
with the clustering (124) (356), the exchange between the clusters of a particle t with mass m is possible only if the momentum $k = p_1 + p_2 - p_4$ available to it is parallel (with positive coefficient) to the translation σa, and on mass shell: $k^2 = m^2$. These are exactly the conditions under which the classical double scattering



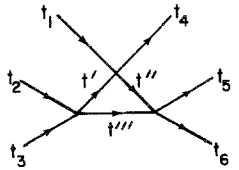
can take place, where the two interaction points are separated in space-time by the 4-vector a.

The above diagram is called a *Landau diagram*. In general, to every classical multiple scattering process there corresponds a Landau diagram.

Further examples are



and

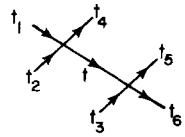


From a Landau diagram one can read off directly a set of equations called Landau equations whose so-called positive-α solutions represent the only points in momentum space where the corresponding multiple scattering process is possible. The set of all such points in general form a real analytic submanifold (of codimension 1) of the physical region called a *positive-α Landau surface.* So, to summarize: a given process $t_1 + \cdots + t_m \to t_1' + \cdots + t_n'$ can only take place as a multiple scattering process on those exceptional points of the physical region that lie on the positive-α Landau surface of that multiple scattering process.

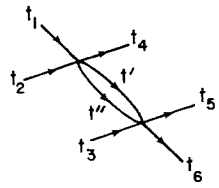
Now, for all points in the physical region of a given process that do not lie on any positive-α Landau surface, no exchange of particles between clusters is possible in the quantum mechanical case as $\sigma \to \infty$, and so macrocausality demands that the amplitude fall off exponentially with σ. But then a generalization of the Fourier theorem asserts that the scattering amplitude is analytic at that point of the physical region. Thus one can prove that the *scattering amplitude is analytic in every point of the physical region that does not lie on a positive-α Landau surface.* This does not prove that the amplitudes are singular on their positive-α Landau surfaces; no statement is made by macrocausality about the analyticity at these points. But as we will soon see, one can show from unitarity that the amplitudes do indeed become singular there. From macrocausality one can also deduce a set of so-called $+i\epsilon$ rules that specify how to analytically continue the amplitudes around Landau singularities. Thus *the analytic structure of amplitudes in the neighborhood of their physical regions is fully determined from first principles.*

Outside the physical region the singularity structure is less well known. A "*principle of maximum analyticity* (of the first kind)" is sometimes invoked there, that states that an amplitude is analytic everywhere except where singularities are demanded by unitarity. But where that is has to be deduced by a rather involved iterative process, so that the analyticity structure outside the physical region has not been worked out in all details.

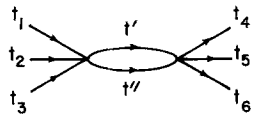
The most important and best-known Landau singularities are those corresponding to Landau diagrams describing double scattering processes: (a) *Poles*, arising from Landau diagrams like



with one internal particle. For this example, as we have seen, the Landau surface is described by $s_{124} = m_t^2$. (b) n-particle *normal threshold singularities*, arising from Landau diagram like



or



For these examples, the Landau surface is described by $s_{124} = (m_{t'} + m_{t''})^2$, or, respectively, $s_{123} = (m_{t'} + m_{t''})^2$.

2.3.5. The Bootstrap Conjecture. This is not a postulate on a par with the others, but a conjecture about the theory as defined by those postulates. The bootstrap conjecture holds that there is only one (nontrivial) SM that satisfies all the above postulates: the physical one. Hence the SM is theoretically well defined and determined by these postulates, without the need for any further dynamical principles. Everything, all the masses, coupling constants, amplitudes, and symmetries flow from self-consistency requirements; there are no free parameters.

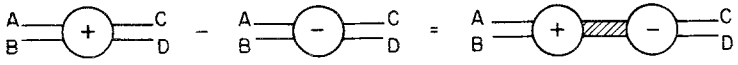
While there are some good arguments for this conjecture, and some partial bootstrap calculations have been quite successful, it is not undisputed. It serves more as a general philosophy and a methodological guideline than as an explicit logical ingredient of theoretical proofs or calculations.

Having now presented the axiomatic foundation of the SMT, we describe, in even less detail, some of the more important properties of the SM that can be derived from it.

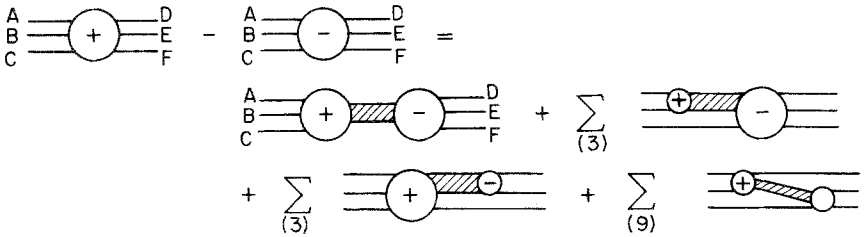
2.4. Some Important Properties of the S Matrix

If one writes down the unitarity equation

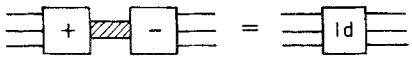
and substitutes for each of the factors its cluster decomposition into connected part products, then one obtains an equation expressing unitarity in terms of connected parts alone. Examples are



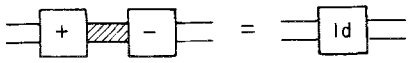
and



In the second example, the sums are again over the different ways one can attach the external particles to the bubbles, the number under the sum indicating in how many different ways that can be done. It is noteworthy that the unitarity equation

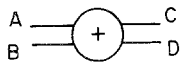


that yields the second equation also yields (from its disconnected parts) the first equation, which was originally obtained from

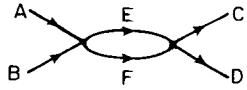


And this pattern repeats itself for higher unitarity equations. This provides a nice consistency test for the theory.

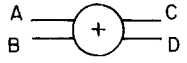
It is well known that normal threshold singularities give rise to cuts in the complex plane of the corresponding channel variable. For example, for the amplitude



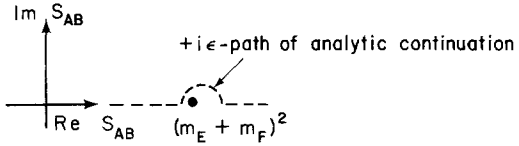
there is a two-particle normal threshold corresponding to the Landau diagram



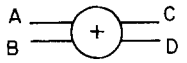
The location of this singularity is, according to the Landau equations, $s_{AB} = (m_E + m_F)^2$, and can thus be expressed in terms of one channel variable alone (this is no longer true for higher Landau singularities), making it simple to represent as a point in the complex s_{AB} plane. The $+i\epsilon$ rules derived from macrocausality then tell us that if we take



at some value $s_{AB} < (m_E + m_F)^2$, and analytically continue it to some value $s_{AB} > (m_E + m_F)^2$ by circumventing the singular point $(m_E + m_F)^2$ in a $+i\epsilon$ direction



then we obtain the physical amplitude

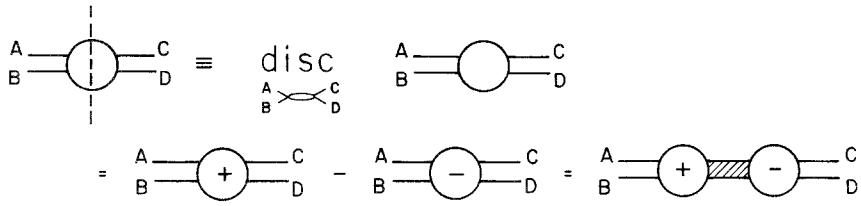


If, on the other hand, we continue on a $-i\epsilon$ path, we obtain a different function; the difference between these two analytic continuations is called the *discontinuity* around the *EF* normal threshold. One often represents this analytic structure by drawing a *cut* from the threshold along the real axis to $+\infty$, which defines a *physical* (Riemann) *sheet*. Then the discontinuity is the difference of the amplitude (on the physical sheet) above and below the cut.

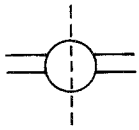
The example of the normal threshold is instructive in that all Landau singularities except for poles similarly give rise to cuts and corresponding discontinuities.

It is possible to derive a general formula that expresses the discontinuity around any physical-region Landau singularity, in the neighborhood of that singularity, in terms of physical-region amplitudes and quantities derivable from them. For the simplest case, the normal threshold discontinuity of a

$2 \to 2$ process, the answer is well known and is given directly by the connected-part unitarity equation

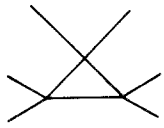


where we have introduced the customary notation

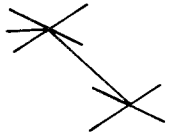


for normal-threshold discontinuities.

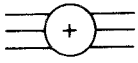
Other discontinuities are not obtained as easily directly from the unitarity equations, but there is a general procedure for deriving them, using unitarity and cluster decomposition repeatedly. For discontinuities corresponding to Landau diagrams with not more than one internal particle connecting any pair of vertices, such as



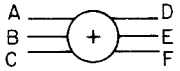
or



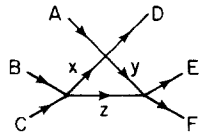
the discontinuity is simply obtained by replacing each vertex of the Landau diagram by the corresponding plus bubble



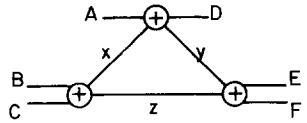
and integrating over intermediate particles lines in the usual way (as in unitarity products). For example, the discontinuity of



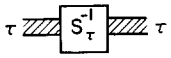
around the triangle singularity



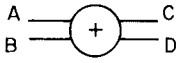
is given by



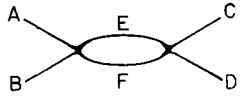
If there are two or more internal particle lines connecting a pair of vertices, then an additional factor



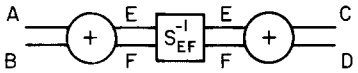
has to be inserted on that set of lines, where S_τ^{-1} is the inverse operator of the restriction of the operator S to the channel τ. For example, the discontinuity of



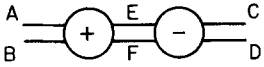
around the discontinuity



is given by



which can be shown to equal the usual result

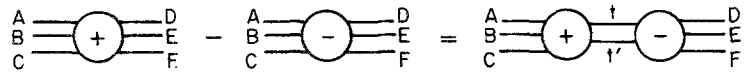


The derivation of these discontinuity formulas also demonstrates that amplitudes actually become singular on all positive-α Landau surfaces, a point that was left open as long as we only considered the effect of macrocausality.

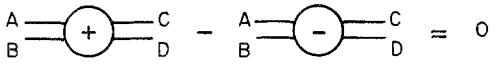
Although these discontinuity formulas are at first only derived in the vicinity of the singularity, they can be extended in most cases to the whole cut. For amplitudes with six or less particles this has been generally proved

(Coster and Stapp, 1969a, b); for amplitudes with more than six particles there are some singularities where certain difficulties appear.

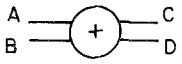
With certain weak assumptions about the analytic behavior of amplitudes away from the physical region one can prove so-called *generalized unitarity* equations that formally resemble unitarity equations, but are valid in nonphysical regions of momentum space, e.g.,



is valid in the nonphysical region $(m_t + m_{t'})^2 < S_{ABC} < (m_A + m_B + m_C)^2$. A particularly important example is the equation

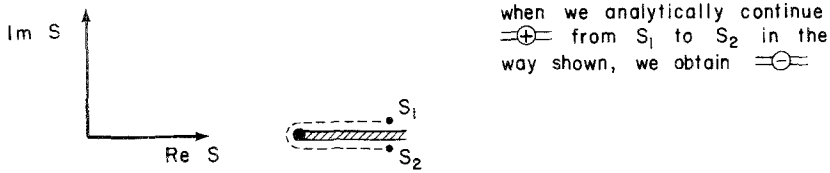


valid below the lowest two-particle threshold. It expresses the property of *Hermitian analyticity*, according to which



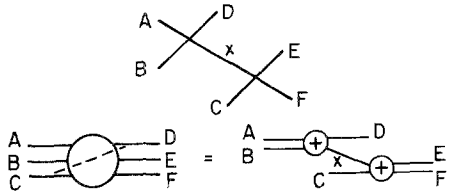
and

are analytic continuations of one another around the lowest threshold



Combined with the time invariance of amplitudes this implies their real analyticity, and the frequently encountered substitution of 2 Im T for the discontinuity of T.

A particularly important conclusion can be drawn by writing down the discontinuity equations for Landau diagrams with two vertices connected by a single internal particle, e.g.,



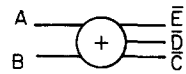
from which we can conclude that in the neighborhood of $S_{ABD} = m_x^2$,

$$T(A, B, C; D, E, F) = \frac{T(A,B; D,X)T(X,C; E,F)}{S_{ABD} - m_x^2}$$

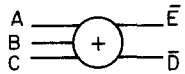
Thus $T(A, B, C; D, E, F)$ has a *pole* at $s_{ABD} = m_x^2$, and the residue of this pole factorizes into two factors, each an appropriate scattering amplitude itself. This important property of *pole factorization* establishes the identity of particles with factorizable poles of the SM. This may be seen by examining a double scattering process in space-time, with the two scattering regions macroscopically separated; only the pole contributes significantly to this process, and the pole factorization permits the interpretation of a particle X being produced in one region, propagating to the other region, and being absorbed there (Stapp, 1965). Complex poles can be interpreted as unstable particles, pole factorization now allowing the definition of amplitudes with unstable external particles, with properties similar to those of amplitudes with only stable particles.

Dispersion relations for amplitudes follow from their analyticity and the use of Cauchy's theorem. The one-variable (fixed-t) dispersion relations are well known. There is a generalization of these to n-particle amplitudes known as the *Bergman–Weil dispersion relations* (Stapp & Wright, 1968). Nothing more is needed for their proof than the analyticity structure of amplitudes on the physical sheet. They express the amplitude anywhere on the physical sheet in terms of its physical-sheet discontinuities.

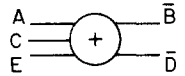
It is well known that to each particle t there is a particle $\bar{t}$, called its *antiparticle*, with the same mass and spin, and opposite internal quantum numbers; and that any two amplitudes related by *crossing*, i.e., by converting an "in"-particle t into an "out"-particle $\bar{t}$, are analytic continuations of one another. For example, the two amplitudes



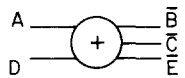
and



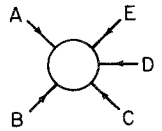
(and indeed any other amplitude related to them by a series of crossing operations, e.g.,



or



are obtained from one another by analytically continuing, say, the former function from its physical region to the physical region of the process $ABC \to \bar{D}\bar{E}$). Thus, there is one analytic function, denoted by



that describes all these processes: when all the 4-momenta are real, then it represents a physical-region scattering amplitude, namely, of the process whose "in" particles are those with positive energies and the "out" particles those with negative energies.

The above statements about the existence of antiparticles and the crossing properties of amplitudes can be proved from the principles of SMT, together with some technical assumptions about the analyticity structure of amplitude outside the physical region.

The *TCP theorem*, which states the invariance of the SM under the TCP-operation follows from crossing all "in" particles into "out" particles and vice versa, and showing that this operation takes one back to the starting point.

The *spin and statistics* theorem is also a general result of SMT: it states that integer-spin particles obey Bose statistics, half-integer spin particles, and Fermi statistics.

Amongst other properties that can be proved are various asymptotic bounds limiting the growth of amplitudes as channel invariants s tend to infinity. The most well known is the Froissart bound for the four-particle amplitude, stating that the amplitude cannot grow faster than $s^{-1} \ln s$ as $s \to \infty$ at fixed t.

We conclude this rapid review of S-matrix theory with the remark that with a further analyticity assumption (analyticity of the second kind) one obtains *Regge theory*, with its wealth of phenomenological predictions.

In the next section we construct a theory of the sequentially ordered S matrix along parallel lines. Each axiom and derived property mentioned here will have its (appropriately modified) analogue there. But additional properties (the "duality" and quark properties) will follow there that have no analog in the unordered case.

3. THE SEQUENTIALLY ORDERED S MATRIX AND ITS PROPERTIES

3.1. Ordered Processes and Ordered Channels

In Section 1 we outlined the motivation for regarding ordered processes and the corresponding ordered amplitudes. Now we proceed to

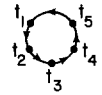
actually construct an S-matrix theory of sequentially ordered scattering processes.

3.1.1. Process Graphs. By a sequentially ordered set S we mean one on which is defined a relation between two elements called the successor relation "$\rightarrow$" with the following properties:

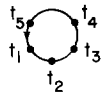
(a) For every element $t \in S$, there exists exactly one $t' \in S$ such that $t \rightarrow t'$. t' is called the *successor* of t.

(b) For every element $t \in S$, there exists exactly one $t'' \in S$ such that $t'' \rightarrow t$. t'' is called the *predecessor* of t.

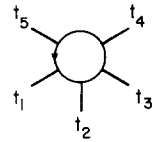
(c) For every pair of elements $t, t' \in S$ there either exists a finite sequence of elements from $S(t_1, t_2, \ldots, t_n)$ such that $t \rightarrow t_1, t_1 \rightarrow t_2, \ldots, t_n \rightarrow t'$; or there exists a finite sequence $(t_1', t_2', \ldots, t_m')$ such that $t' \rightarrow t_1', t_1' \rightarrow t_2', \ldots, t_m' \rightarrow t$; or both these sequences exist. It can easily be shown that if S is a finite sequentially ordered set it can be represented by the following graph:



We now assume that every scattering process is sequentially ordered, i.e., the set of particles participating in the process is assumed to be a sequentially ordered set. The graph of the above type that represents this order is called the process-order graph or *process graph*. For brevity we usually omit all the arrows but one, so that a typical process graph would look like this:



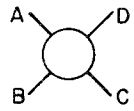
To each such ordered process corresponds an ordered amplitude



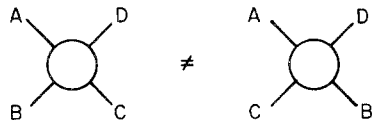
interpreted as its probability amplitude. Note that we are using the same notation for ordered amplitudes as for physical, unordered amplitudes, except that of course now the order in which the particle lines are arranged around the bubble is significant; since we will henceforth only be talking about ordered amplitudes except where explicitly stated, there should be no reason for confusion. Also, we will by convention always arrange the particles

in a counterclockwise orientation, so that we may drop the arrow altogether in the notation for amplitudes.

Note that an ordered amplitude



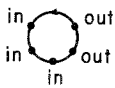
is not merely a function of the particle degrees of freedom t_A, p_A, μ_A, t_B, . . . , but also of the cyclical order in which the particles t occur:



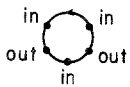
Therefore there will be $(n-1)!$ different ordered n-particle amplitudes corresponding to a given set of n external particles, compared with only one physical amplitude.

Inherent in the concept of process order is the notion that the process be *connected* (in the sense of "connected part"); the relative order of disconnected parts would not be defined. It is for this reason that we associate an ordered amplitude or *connected* part with an ordered process, rather than an ordered S-matrix element. Intuitively speaking, both the connectedness of order and the connectedness we have encountered in usual SMT (as "connected part") refer in some sense to causal connectedness. But whereas this connectedness is, in usual SMT, taken to be global and amorphous, we now introduce form into this connection by assuming each particle to be connected to specific other individual particles; in this part of the work the connectedness pattern is assumed to be sequential; in Part Two it will be generalized.

3.1.2. Ordered Channels and Channel Graphs. Some of the particles of an ordered process will be "in," the others "out." Whenever the "in" particles form a connected structure by themselves, i.e., are not interspersed with "out" particles (whence the analogous statement will be true for "out" particles), we call the ordered process an *ordered transition.* Thus

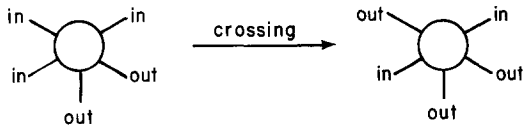


is an ordered transition, while

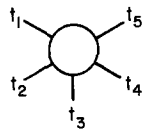


is not.

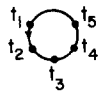
Ordered transitions play an important role because the complete set of ordered amplitudes that ultimately describe all ordered processes can be defined in terms of them alone: after we have proved crossing, we will be able to analytically continue the ordered transition amplitudes to other real momentum-energy regions where they describe general ordered processes; e.g.,



As a result there is, as we will show, *one* analytic function



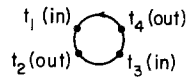
describing all processes with the process graph



regardless of which channel is involved, i.e., which of the particles are "in" and which ones "out." But we are ahead of ourselves; we merely meant to explain why it is possible to construct the theory starting with a subset of all ordered processes, namely, the ordered transitions.

Possible it may be, but why desirable? The reason is the following. In SMT, and in quantum mechanics in general, a process, and hence its amplitude, is fully determined by specifying its initial state and its final state. This is the basis for the Hilbert space formalism, with its expression for the amplitude $\langle i|T|f\rangle$. And although, e.g., crossing suggests that a scattering amplitude, in some deeper sense, describes a whole and indivisible process, nevertheless the Hilbert space formalism, dualistic as it is with its stress on initial and final states, successfully describes processes at the physical level. This dichotomy between initial and final states is of course based on the experimental dichotomy between preparation and measurement.

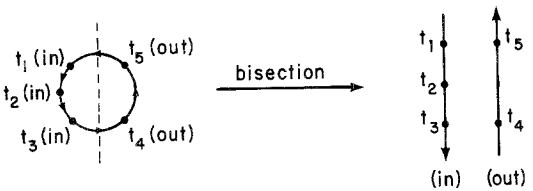
For ordered processes, such a description in terms of separate initial and final channels is in general no longer possible because the order of the process cannot in general be specified in terms of the separate orders of the initial and final particles. Already the simple example



demonstrates this clearly: there is no way that we can specify an initial ordered state and a final one *separately* that would allow us to reconstruct the process-order graph. Thus it would seem that the Hilbert space formalism, the hard core of quantum mechanics as we know it, is not applicable to ordered processes.

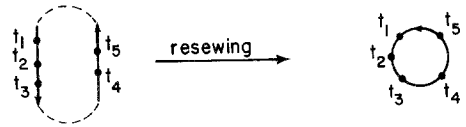
However, for the special class of processes that we called ordered transitions it is, as we shall see, possible to specify the process by specifying "in" and "out" channels separately. It is this circumstance, coupled with the fact that ordered transition amplitudes can be extended by analytic continuation to apply to all ordered processes, that allows the construction of an SMT of ordered processes.

How are ordered transitions specified by their channels? When we bisect the process graph of an ordered transition into its "in" and "out" components, we obtain two linear directed graphs representing the order of the "in" or, respectively, "out" particles separately. We call these graphs *channel graphs*:



Thus one obtains channel graphs from a process graph by *bisection.* A channel graph, together with the individual particle parameters, t, p, μ, specifies an *ordered state*; together with the individual particle-type parameters t, alone it specifies an *ordered channel.*

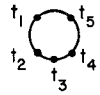
Given the ("in" and "out") ordered states of an ordered transition, we can immediately reconstruct the process graph by simply sewing the two channel graphs together in the unique way such that the two channel orientations match. This resewing of a channel graph obtained by bisection of a process graph yields back the original process graph:



Thus indeed we see that an ordered transition can be fully specified by separate specification of its channels, as claimed. Bisecting and resewing are inverse operations.

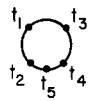
This state of affairs is by no means to be taken for granted. Had we, for

example, tried to construct an SMT of ordered processes *without orientation,* i.e., with process graphs like this



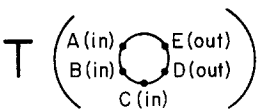
(without arrows), then we would have failed: even for ordered transitions, the initial and final channel graphs, say

(obtained as before through bisection of the process graph) do not uniquely define the process graph, since it is not clear how to sew them together; one of the two possible ways yields the original process graph, the other one

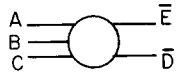


This *criterion of unique resewability* will play an important role in Part Two in determining the most general particle ordering upon which one can construct an SMT. Here, we see that the orientation of the edges of a sequentially ordered process graph is an essential, necessary feature of the theory. And since, as we will see, the existence of a nontrivial charge-conjugation operation, and of baryon number and other additive conserved quantum numbers, is closely related to the fact of edge orientation, we begin to realize the powerful results that can be generated just from self-consistency requirements. This will become much more apparent in Part Two.

Following a notation analogous to that of usual SMT, we denote the amplitude

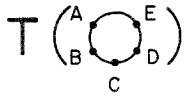


corresponding to an ordered transition by

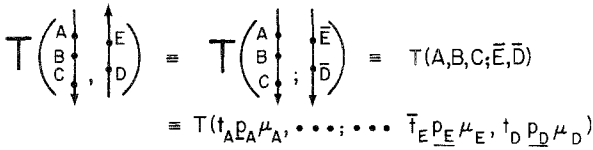


Here the bar above the particle symbols signifies the transformation $\overline{(t, \mathbf{p}, \mu)}\,(\bar{t}, \mathbf{p}, -\mu)$; at this point this is not more than a convention, which will, however, be justified when we prove crossing.

A nondiagrammatic notation follows from the fact that



can be written as



This notation is outwardly identical with the one employed in usual SMT for amplitudes. But here the order of the particles in the argument of the function T is allowed to express the channel order. Thus this argument order, redundant in usual SMT, and eliminated there by imposition of permutation symmetry, comes to its right in ordered SMT.

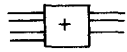
A result of this fortunate circumstance (that we can employ the same formalism for ordered as for physical amplitudes) is that many of the equations and proofs of usual SMT continue to hold for ordered SMT, as will be seen.

In terms of Hilbert space language the difference between physical and ordered SM is that the former is defined on a Hilbert space of states in which the n-particle direct-product spaces have been contracted to their symmetric or, respectively, antisymmetric subspaces, whereas the latter is defined on the full (and therefore ordered) direct-product space.

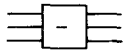
3.1.3. Ordered S Matrix. So far we have introduced only the ordered connected parts or amplitudes (these are, as we saw, the basis objects of our theory and describe the probability amplitude of an ordered process); but if we want to introduce the postulate of (ordered) unitarity, we will need the ordered analog of the S matrix, defined by its cluster decomposition into ordered connected parts.

Since unitarity yields a host of important properties for the ordered amplitudes (like pole factorization, discontinuity formulas, crossing, etc.) we are loath to discard it as a tool. But in order to include it in our arsenal we have to pay the price of introducing an artificial concept, namely, the ordered S matrix, which neither, as the physical S matrix, describes the probability amplitude of observable processes, nor, as the ordered connected part, or for that matter the physical connected part, describes the probability amplitude of hypothetical elementary processes. Indeed, the only reason we introduce it is to be able to postulate its unitarity, which can then, via cluster decomposition, be converted into the desired properties of the connected parts.

We denote the ordered S-matrix elements by



and the matrix elements of the inverse S'' by



Again, there can be no confusion with the physical S matrix, since all S matrix elements will be assumed to be ordered unless specifically mentioned.

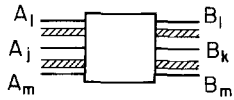
3.2. The Basic Postulates of Ordered S-Matrix Theory

Upon the general ordered S-matrix (OSM) framework described in the previous section we now impose four postulates analogous to those of usual SMT, as described in Section 2, which are assumed to fully define the theory in bootstrap fashion.

3.2.1. Poincaré Invariance. It is introduced in formally the same way as for the physical SM, and implies the same consequences, since a Poincaré transformation does not affect the order of the particles of a process, but only their individual momenta and helicities.

3.2.2. Cluster Decomposition. We postulate the independence property as for the physical SM, but with the modification, natural if we consider the interpretation of order as causal order, that the clusters consist of *adjacent* particles.

Specifically, given an ordered SM element



choose an arbitrary subset (cluster) of adjacent particles; if it is to correspond to a nonzero amplitude, it must contain both "in" and "out" particles, so its general form is $\{A_1, \ldots, A_j; B_1, \ldots, B_k\}$, with $1 < j < m$, and $1 < k < n$. Call this cluster 1, and its complement $\{A_{j+1}, \ldots, A_m; B_{k+1}, \ldots, B_n\}$ cluster 2. Then if we translate the particles of cluster 1 by σa, where a is an arbitrary 4-vector, and σ a real number, then as $\sigma \to \infty$, according to the independence postulate, the ordered SM element factors into the product of the two separate cluster SM elements:

$$\lim_{\sigma \to \infty} \langle A_1^{\sigma a} \cdots A_j^{\sigma a}| \otimes \langle A_{j+1} \cdots A_m|S|B^{\sigma a} B_2^{\sigma a} \cdots B_k^{\sigma a}\rangle \otimes |B_{k+1} \cdots B_n\rangle$$
$$= \langle A_1 \cdots A_j|S|B_1 \cdots B_n\rangle \cdot \langle A_{j+n} \cdots A_m|S|B_{k+1} \cdots B_n\rangle \quad (3.1)$$

When the two clusters do not consist of adjacent particles, then each subset (of a cluster) that consists of adjacent particles may be considered as an independent cluster as far as the formulation of the independence postulate is concerned. Thus this case is reduced to the case of p ($p < 2$) clusters of adjacent particles, which we treat later.

Next, we define a cluster decomposition of SM elements into sums of products of connected parts. Our guiding principle in constructing the cluster decomposition will be that the left-hand side of the independence postulate should tend to zero when a cluster of its particles is translated to infinity:

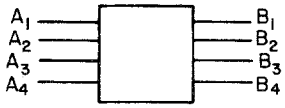
$$\lim_{\sigma \to \beta} \langle A_1^{\sigma a} \cdots A_j^{\sigma a}| \otimes \langle A_{j+1} \cdots A_m|S_c|B_1^{\sigma a} \cdots B_k^{\sigma a}\rangle \otimes |B_{k+1} \cdots B_n\rangle = 0 \quad (3.2)$$

The cluster decomposition will certainly express the ordered SM element as a sum of products of connected-part matrix elements. Any particular term of the sum, being a product of l connected parts, corresponds to a particular of the external particles into l ordered clusters. The question is only which of all such possible partitions are "legal," i.e., actually occur in the cluster decomposition: the specification of the set of legal partitions defines the cluster decomposition.

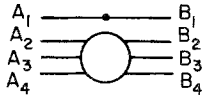
The correct rule for cluster decomposition, i.e., the one which, together with the independence postulate (3.1), guarantees the above equation (3.2), turns out to be the one that one might well guess intuitively.

Let us consider partitions into two clusters first. Which ones are legal? The answer is: those where the two clusters each consist of adjacent particles, i.e., the so-called adjacent partitions, obtained by bisection of the process graph. And since any cluster consisting of either only "in" or only "out" particles yields a zero connected part, we consider only those adjacent partitions where each cluster contains both "in" and "out" particles.

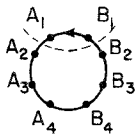
To illustrate this with an example, regard the cluster decomposition of



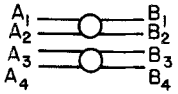
into $l = 2$ terms:



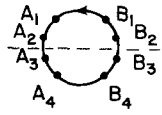
corresponding to



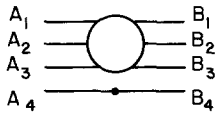
, or,



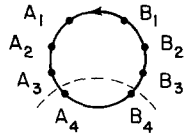
corresponding to



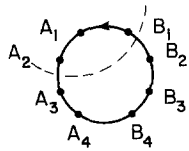
and



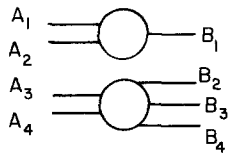
corresponding to



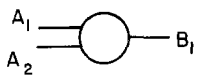
This exhausts the possibilities, since, e.g.,



would lead to

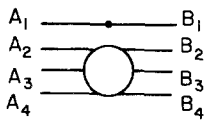


which is zero because

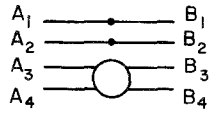


is zero (all particles are assumed stable).

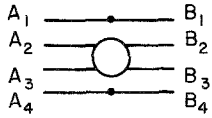
For $l = 3$ the legal partitions are those obtained from the $l = 2$ partitions by further adjacent partitioning of either of the two clusters. For example, in the above example the partition



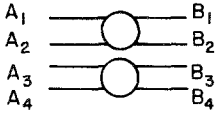
leads to



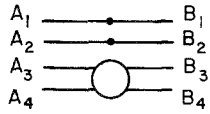
and



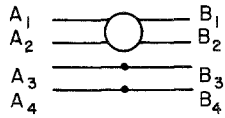
while



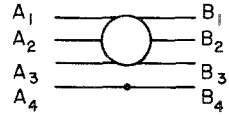
leads to



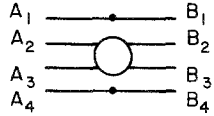
and



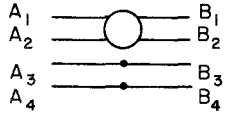
while



leads to

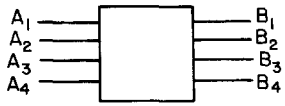


and

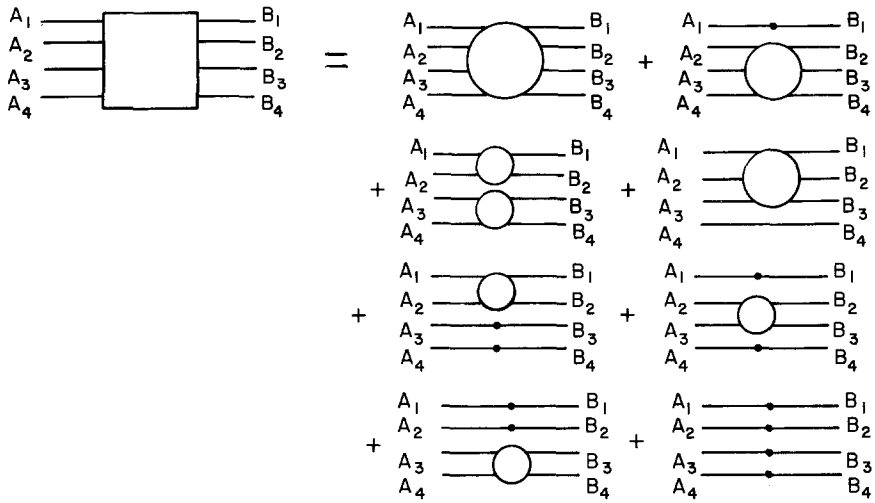


Of course each term of a kind only has to be counted once. The $l = 4$ terms are obtained from the $l = 3$ terms in a similar way, etc.

The complete cluster decomposition of

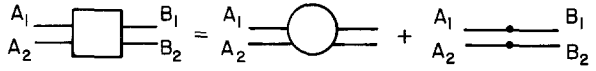


now reads

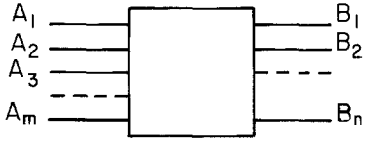


Note that there are far fewer terms than in the corresponding cluster decomposition of the physical S matrix, for which there are $4! = 24$ terms corresponding to just the one term

alone. Another simple example of ordered cluster decomposition is



The general result can easily be formulated in terms of "in" and "out" channels: to cluster-decompose a given ordered SM element



choose an integer l, such that $0 < l < \min(m, n)$. Partition the "in"-channel graph into l connected, nonempty subgraphs in an arbitrary way; do likewise with the "out"-channel graph. Then combine the top "in"

channel with the top "out" channel to form a connected part, do likewise with the second "in" and "out" channels, etc., to form the product of l connected parts. Do the partitioning in all different ways, and for all integers l with $0 < l < \min(m, n)$, and add all the terms together. In equation form this instruction reads, with $j_0 \equiv 1$, $k_0 \equiv 1$, $J_l \equiv m$, $k_l \equiv n$,

$$\langle A_1, A_2, \ldots, A_m|S|B_1, B_3, \ldots, B_n\rangle = \sum_{l=0}^{\min(m,n)} \sum_{j_1=1}^{m-l} \sum_{k_1=1}^{n-l} \sum_{j_2=j_1+1}^{m-l+1} \sum_{k_2=k_1+1}^{n-l+1} \sum_{j_3=j_2+1}^{m-l+2} \sum_{k_3=k_2+1}^{n-l+2} \cdots \sum_{j_l=j_{l-1}+1}^{m-1} \sum_{k_l=k_{l-1}+1}^{n-1} \prod_{i=0}^{l-1} \times \langle A_{j_i+1}A_{j_i+2}\cdots A_{j_{i+1}}|S_c|B_{k_i+1}B_{k_i+2}\cdots B_{k_{i+1}}\rangle \quad (3.3)$$

This is the general equation of ordered cluster decomposition. It may be regarded as defining the ordered S matrix. [By contrast, in physical SMT it is the S matrix that is regarded as the primary object, since it is directly connected with observation, and the cluster decomposition is regarded as recursively and iteratively defining the connected parts.]

We now prove the claim that the independence postulate (3.1), in conjunction with the equation (3.3) that we have chosen as our cluster decomposition, implies the crucial property (3.2). The proof, similar in its basic idea to the proof of the analogous statement in usual SMT, proceeds by induction. First, we assume that each cluster consists of one "in" and one "out" particle each. Employing the notation $\langle A_1^{\sigma a}| \otimes \langle A_2|S|B_1^{\sigma a}\rangle \otimes |B_2\rangle \equiv \langle A_1^{\sigma a}A_2|S|B_1^{\sigma a}B_2\rangle$ for the SM element whose first cluster has been translated by the 4-vector σa, equation (3.3) then reads

$$\begin{aligned}\langle A_1^{\sigma a}A_2|S|B_1^{\sigma a}B_2\rangle &= \langle A_1^{\sigma a}A_2|S_c|B_1^{\sigma a}B_2\rangle + \langle A_1^{\sigma a}|S_c|B_1^{\sigma a}\rangle\langle A_2|S_c|B_2\rangle \\ &= \langle A_1^{\sigma a}A_2|S_c|A_1^{\sigma a}B_2\rangle + \langle A_1|S|B_1\rangle\langle A_2|S|B_2\rangle\end{aligned}$$

hence

$$\lim_{\sigma\to\infty}\langle A_1^{\sigma a}A_2|S_c|B_1^{\sigma a}B_2\rangle = \lim_{\sigma\to\infty}\langle A_1^{\sigma a}A_2|S|B_1^{\sigma a}B_2\rangle - \langle A_1|S|B_1\rangle\langle A_2|S|B_2\rangle = 0$$

by virtue of equation (3.1). Thus (3.2) is true for this case.

Now assume that (3.2) is true for a $(r-1)$-particle SM element. Now regard $\langle A_1\cdots A_m|S|B_1\cdots B_n\rangle$, with $m+n=r$. By (3.3) we have

$$\begin{aligned}&\langle A_1^{\sigma a}A_2^{\sigma a}\cdots A_j^{\sigma a}A_{j+n}\cdots A_m|S_c|B_1^{\sigma a}B_2^{\sigma a}\cdots B_k^{\sigma a}B_{k+1}\cdots B_n\rangle \\ &\quad = \langle A_1^{\sigma a}\cdots A_j^{\sigma a}A_{j+1}\cdots A_m|S|B_1^{\sigma a}\cdots B_k^{\sigma a}B_{k+1}\cdots B_n\rangle \\ &\qquad - \langle A_1^{\sigma a}\cdots A_j^{\sigma s}|S_c|B_1^{\sigma a}\cdots B_k^{\sigma a}\rangle\langle A_{j+n}\cdots A_m|S_c|B_{k+1}\cdots B_n\rangle - R_1 - R_2\end{aligned}$$

where R_1 consists of all terms of the cluster decomposition of

$$\langle A_1\cdots A_m|S|B_1\cdots B_n\rangle$$

whose partitions into clusters correspond to subpartitions of the partition $(A_1 \cdots A_j; B_1 \cdots B_k)(A_{j+n} \cdots A_m; B_{k+n} \cdots B_n)$, and R_2 consists of all other terms.

Now, as $\sigma \to \infty$, R_2 tends to zero, since each of its terms contains as a factor at least one S_c-matrix element with some particles translated to infinity and others not, so that this matrix element, containing less than r particles, becomes zero by induction hypothesis.

On the other hand, $\langle A_1^{\sigma a} \cdots A_j^{\sigma a} | S | B_1^{\sigma a} \cdots B_k^{\sigma a} \rangle + R_1$ is, by virtue of (3.3), equal to

$$\langle A_1^{\sigma a} \cdots A_j^{\sigma a} | S | B_1^{\sigma a} \cdots B_k^{\sigma a} \rangle \langle A_j + 1 \cdots A_m | S | B_{k+1} \cdots B_n \rangle$$
$$= \langle A_1 \cdots A_j | S | B_1 \cdots B_k \rangle \langle A_{j+1} \cdots A_m | S | B_{k+1} \cdots B_n \rangle$$

so that

$$\lim_{\sigma \to \infty} \langle A_1^{\sigma a} \cdots A_j^{\sigma a} A_{j+1} \cdots A_m | S_c | B_1^{\sigma a} \cdots B_k^{\sigma a} B_{k+1} \cdots B_n \rangle$$
$$= \lim_{\sigma \to \infty} \langle A_1^{\sigma a} \cdots A_j^{\sigma a} A_{j+n} \cdots A_m | S | B_1^{\sigma a} \cdots B_k^{\sigma a} B_{k+1} \cdots B_n \rangle$$
$$- \langle A_1 \cdots A_j | S | B_1 \cdots B_k \rangle \langle A_{j+1} \cdots A_m | S | B_{k+1} \cdots B_n \rangle = 0$$

by (3.1). This proves (3.2).

A generalization of (3.1) from two to p clusters follows from stepwise application of (i). Let

$$\langle i | S | f \rangle = \langle i_1 | \otimes \langle i_2 | \otimes \cdots \otimes \langle i_p | S | f_1 \rangle \otimes | f_2 \rangle \otimes \cdots f_p \rangle$$

and let $a_1, a_2, \ldots, a_p$ be a set of p distinct 4-vectors. Then the generalized independence postulate reads

$$\lim_{\sigma \to \infty} \langle i_1^{\sigma a_1} | \langle 1_2^{\sigma a_2} | \cdots \otimes \langle i_p^{\sigma a_p} | S | f_1^{\sigma a_1} \rangle \otimes \cdots \otimes | f_p^{\sigma a_p} \rangle$$
$$= \langle i_1 | S | f_1 \rangle \langle i_2 | S | f_2 \rangle \cdots \langle i_p | S | f_p \rangle \qquad (3.4)$$

This property of S implies, in exactly the same way as above, the generalization

$$\lim_{\sigma \to \infty} \langle i_1^{\sigma a_i} | \otimes \cdots \otimes \langle i_p^{\sigma a_p} | S_c | f_1^{\sigma a_1} \rangle \otimes \cdots \otimes | f_p^{\sigma a_p} \rangle = 0 \qquad (3.5)$$

An equation similar to (3.5) is valid, too, when the clusters are not composed of adjacent particles.

Consequently, one can in the usual way deduce that the ordered amplitudes, i.e., connected-part matrix elements without the overall $\delta^4(\ )$ factor, do not contain any δ functions or their derivatives, and are thus not merely distributions, but true functions.

The fact that we were able to deduce the desired equation (3.5) from a reasonable independence postulate (3.1) for the ordered SM indicates

that we have defined the latter in terms of ordered amplitudes by reasonable cluster-decomposition equations (3.3).

3.2.3. Unitarity. We postulate the unitarity of the ordered SM; $S^+S = SS^+ = 1$, or in terms of matrix elements:

$$\sum_n \langle i|S|n\rangle\langle n|S^+|f\rangle = \sum_n \langle i|S^+|n\rangle\langle n|S|f\rangle = \langle i|f\rangle$$

which is formally identical with the unitarity of the physical SM.

The difference between physical and ordered unitarity lies hidden in the difference between the physical (symmetrized) and the ordered Hilbert space. Thus, e.g., for the ordered Hilbert space, we have

$$\langle A_1 \cdots A_m|B_1 \cdots B_n\rangle = \delta_{mn}\langle A_1|B_1\rangle\langle A_2|B_2\rangle \cdots \langle A_m|B_m\rangle$$

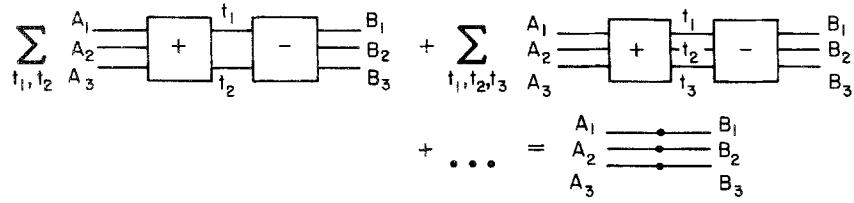
(rather than a sum over all $m!$ permutations terms) and

$$1 = \sum_n \sum_{A_1} \sum_{A_2} \sum_{A_n} |A_1 \cdots A_n\rangle\langle A_1 \cdots A_n|$$

That is, the sum-integral over all intermediate states implies a sum over all different channel orders.

In the diagrammatical notation, unitarity reads

or, e.g.,

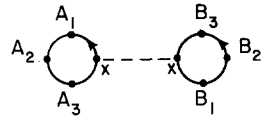


Note that the intermediate lines never cross, since the intermediate ordered states on both bubbles are the same, implying the same channel order.

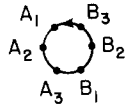
3.2.4. Macrocausality. In usual SMT macrocausality was seen to imply that an amplitude is analytic in the neighborhood of the physical region except possibly on its so-called positive-α Landau surfaces, where it could become singular. Each Landau surface corresponds to a distinct multiple scattering process, symbolized by a Landau diagram in which each vertex represents one of the simple scattering processes of which the multiple scattering process is composed.

Here, both the simple and the multiple scattering processes are ordered. Therefore, we first have to examine how simple ordered processes can combine to form a composite (multiple scattering) process, and how the order of this composite process is determined by the orders of the constituent processes.

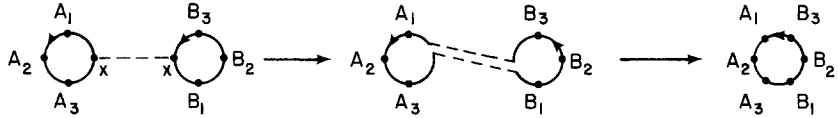
We proceed recursively: first we consider how two simple processes join to form a double scattering process. The simplest case is when the double scattering consists of two processes linked by exchange of a single particle, which we describe by the diagram



This corresponds to the presence of a pole. In this case, the joint process is again ordered and has the process graph

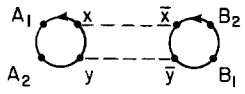


obtained by erasing the internal particle X and joining the two process graphs by their free edges in such a way that the orientations match:

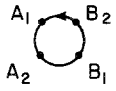


Let us call this procedure the *composition* of the two process graphs to the composite process graph. This process graph is said to represent the *global order* of the composite scattering process.

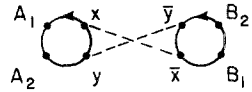
If, instead of one particle, several particles are exchanged between the two processes, then the resulting double scattering process is again ordered if it corresponds to an ordered channel exchange, i.e., if the exchanged internal particles are adjacent and occur in the same order in both processes; the order of the joint process is again obtained by erasing the internal (exchanged) channel and joining the two process graphs by their free edges such that the orientations match. For example, the global order of



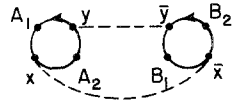
is



whereas



or

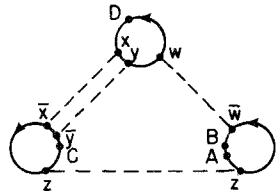


do not have a global order. The first of these examples corresponds to the presence of a normal-threshold discontinuity.

When there are three or more ordered subprocesses linked together to multiple scattering process by exchange of particles then the resulting joint process is ordered if and only if we can legally compose (according to the rules described above) pairs of process graphs step by step until we have obtained one single process graph; when this is possible, this process graph represents the order of the joint process. This case corresponds to one of the higher Landau singularities (e.g., triangle or box singularities). It is easy to see that if this stepwise composition can be carried out in any order of succession, then it can be carried out in every order, and yields the same result.

The criteria for the stepwise fusion to be possible, and thus for the joint process to be ordered, are as follows.

(a) All the external particles of a subprocess must be adjacent and "outside," i.e., not contained in any polygon formed by internal lines (for example,

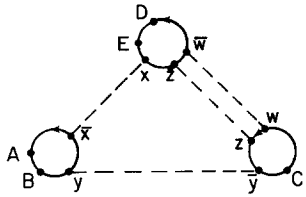


is excluded, because A, B, C are "inside").

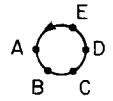
(b) The set of particles exchanged between any two subprocesses must be adjacent and in the same order in both process graphs.

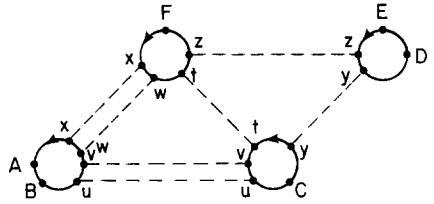
(c) The whole diagram must be drawable on a plane without any crossing of intermediate particle lines.

We illustrate these rules by a few examples:

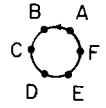


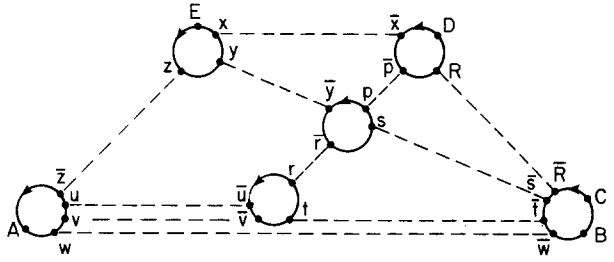
has the global order



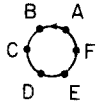


has the global order

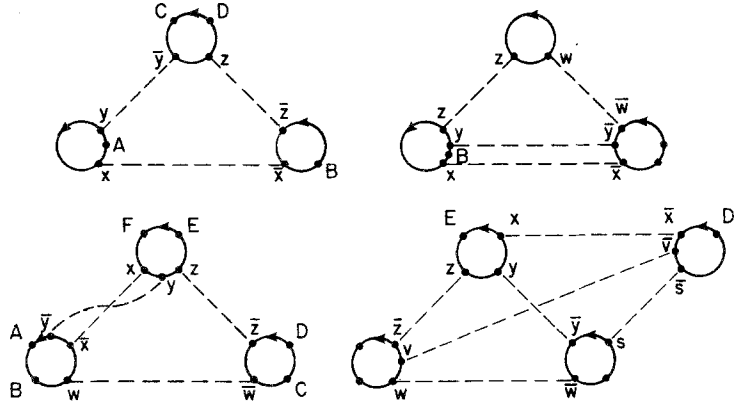




has the global order



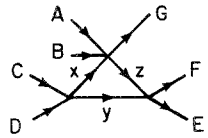
On the other hand, the following multiple scattering processes have no global order:



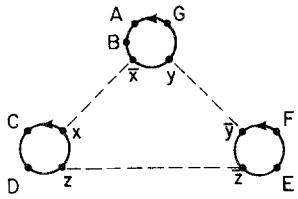
So far we have merely stated these composition rules without proof. In fact, they are generated by unitarity, which demands the presence of singularities wherever multiple scattering processes with a global order, as just defined, can take place; this will be seen in the next section.

Whenever a multiple scattering represented by a graph L has a global order, represented by a process graph G, then at appropriate momenta the ordered process with process graph G can proceed as a multiple scattering of the type L. Therefore, according to the ideas of macrocausality, L may be interpreted as the (ordered) Landau diagram of that process. And then the same procedure used in usual SMT, and outlined in Section 2.3, leads to the conclusion that an *ordered amplitude A is analytic everywhere in the neighborhood of the physical region except on the positive α Landau surfaces* (whose locations are determined in the usual way, by the Landau equations) corresponding to Landau diagrams with a global order identical to that of the amplitude A.

An alternative notation for ordered Landau diagrams, which corresponds to the notation for usual (unordered) Landau diagrams, is possible here (but not for the generalization of Part Two):

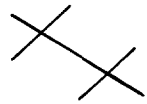


stands for

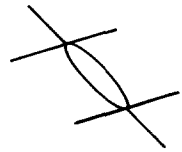


Note that here the order of the particle lines around a vertex is significant, whereas in the usual Landau diagrams it is not.

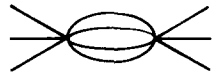
An important difference between physical and ordered amplitudes is that the latter have far fewer singularities because only Landau diagrams with the appropriate global order can contribute. Let us consider the most familiar and important Landau singularities, namely, poles such as



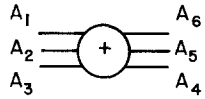
and normal threshold singularities such as



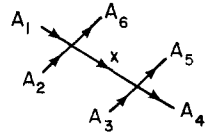
or, respectively,



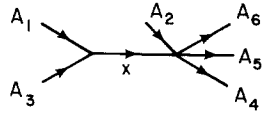
Given an ordered amplitude, such as



then the general results presented above imply that this amplitude has poles and normal threshold cuts only in *adjacent* channel invariants. For example, it will have a pole in s_{612}, corresponding to the ordered Landau diagram



but never in s_{13}: the corresponding Landau diagrams, such as



cannot have the global order

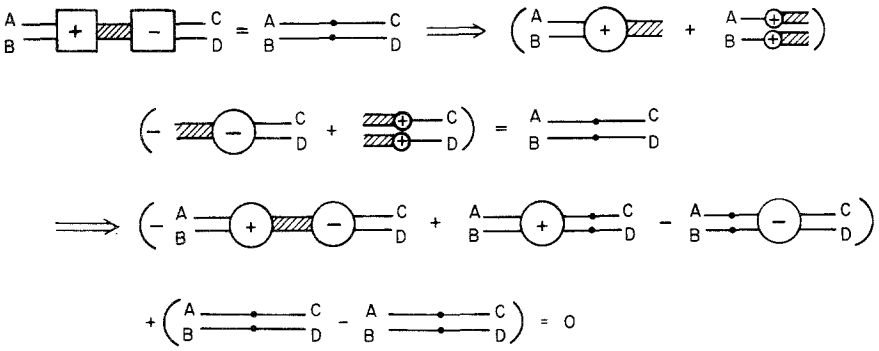
The $+i\epsilon$-rules for ordered amplitudes are identical to those for physical amplitudes, as are the Landau equations that determine the locations of the Landau singularities.

3.3. Properties of the Ordered S Matrix

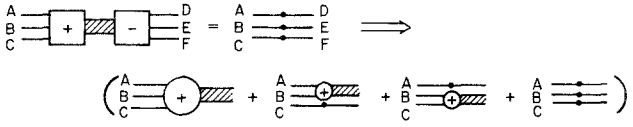
In this section we show how various properties of the physical S matrix carry over into the domain of the ordered S matrix. In many cases, the detailed proofs need not be presented here since they proceed formally analogously to those of physical S-matrix theory; in these cases we comment mainly on the differences and modifications.

3.3.1. Unitarity Relations for Connected Parts. In order to obtain these we write down the unitarity equation

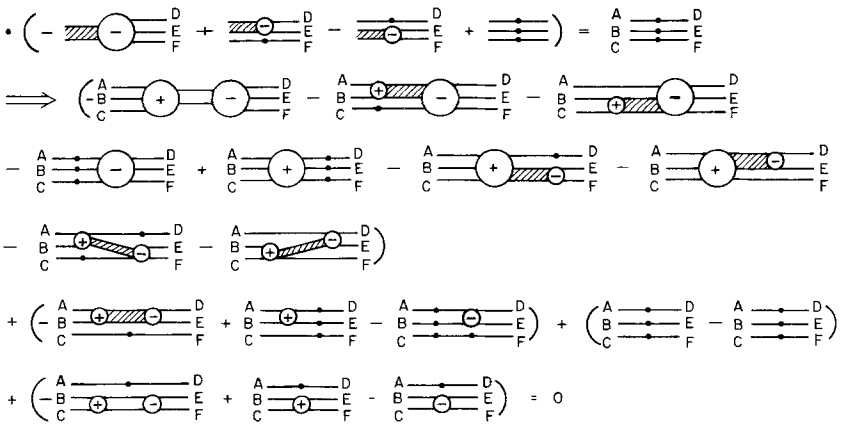
and substitute both SM elements by their cluster decomposition into products of connected parts. Then, in the resulting equation, we regard separately all the terms with the same δ-function structure, i.e., the same connectedness structure: these constitute the desired unitarity equations:



Then the first term in parentheses represents one of the unitarity equations, while the second is in this case trivial. Another example is the following:

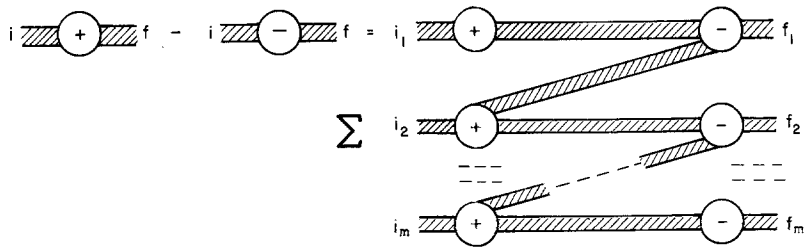


continued overleaf

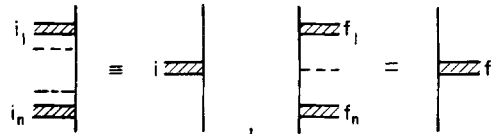


Since each of the terms in parentheses has a different δ-function structure, they must vanish separately. Note that the second term in parentheses reproduces the unitarity equation of our previous example. This constitutes a consistency test that the ordered SMT passes in this case as well as in all others examined.

The general form of such a unitarity equation is



where the sum is over all the decompositions of the initial and final channels into clusters preserving the order and yielding a connected product. Here, as in the above examples, ▨ stands for a set of *noncrossing* particle lines, i.e., for an ordered channel. And



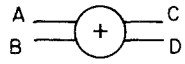
i.e., the external particles occur in the same order on both sides of the equation.

3.3.2. Identification of Ordered Amplitudes with the "Planar Amplitudes" of DTU. Since the connected-part unitarity equations we derived above coincide formally with the "planar unitarity" equations obtained in dual

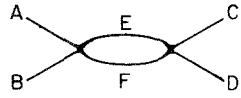
perturbation theory for the planar amplitudes, we may identify our ordered amplitudes with the planar amplitudes of DTU, and consider the theory of the sequentially ordered S matrix as providing a theoretical foundation of DTU. In Section 5 we will show that ordered amplitudes can also be represented as quark amplitudes, which will complete the identification. (DTU is often expressed in terms of quark amplitudes.)

3.3.3. Discontinuity Equations. Just as for physical amplitudes, the repeated appropriate use of the above unitarity equations yields, for any given ordered amplitude and one of its Landau singularities (corresponding to one of the ordered Landau diagrams with the appropriate global order), an expression for the discontinuity of the amplitude around the singularity that is valid at least in the neighborhood of that singularity, and often outside it, too. The procedure is formally identical with that used for the physical amplitudes, and so we will not comment on it further; it also yields the same result, namely, the discontinuity formulas described in Section 2.3. Of course, the operators occurring in these equations now have to be interpreted as operating on the ordered Hilbert space.

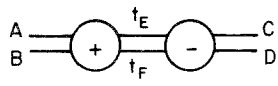
For example, for the two-particle normal threshold discontinuity of the amplitude



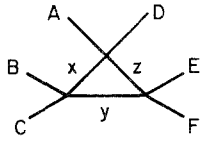
corresponding to the Landau diagram



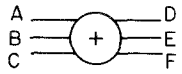
this formula yields the well-known expression



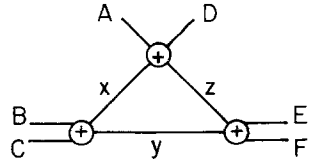
For the



singularity of

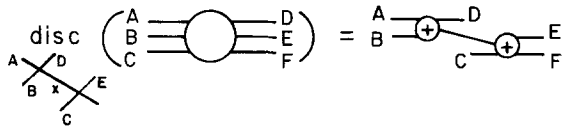


the discontinuity is given by



The extension of these formulas outside the immediate neighborhood of the appropriate singularity *might* possibly be valid for all discontinuities, because for the physical amplitudes the intractable Landau singularities examined till now all corresponded to nonplanar Landau diagrams, and these do not occur in the ordered case. This simplification of the analytic structure of ordered amplitudes might prove to be very important both theoretically and for practical purposes.

3.3.4. Pole Factorization and the Particle Pole Identity. Exactly as for physical amplitudes, the discontinuity formulas for pole singularities, e.g.,



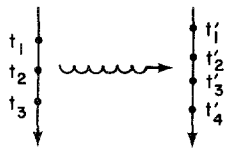
i.e., disc. $T(A, B, C; D, E, F) = T(A, B; D, X)T(X, C; E, F)\delta(S_{ABD} - m_X{}^2)$, from which it follows that $T(A, B, C; D, E, F)$ has a pole at $s_{ABD} = m_X{}^2$ whose residue is factorizable in the usual way. Thus the *particle-pole* identity continues to hold for the OSM, and scattering events dominated by this pole can be interpreted as a double scattering. This is a particularly crucial result since it permits us to interpret the OSM as a bona fide S matrix.

3.3.5. The Physical-Region Analytic Structure of Ordered Amplitudes. While ordered macrocausality states that physical-region singularities can occur *nowhere but* on the positive-α Landau surfaces it does not demand that amplitudes actually become singular there. That they do become singular there is implied by the discontinuity equations (i.e., by unitarity and cluster decomposition), as the example of the pole, in the last paragraph, illustrates. Thus the physical-region analytic structure of the ordered amplitudes (location and type of singularities of the otherwise analytic amplitudes, discontinuity formulas, and $+i\epsilon$ rules) is fully determined from first principles, as was the case for physical amplitudes.

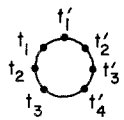
As far as poles and normal thresholds are concerned, the main difference between physical and ordered amplitudes lies in the fact, mentioned earlier, that ordered amplitudes have these singularities only in adjacent channels.

3.3.6. Dispersion Relations. Due to this analytic structure, the same kinds of dispersion relations are valid as for physical amplitudes. But due to far smaller number of cuts, they are considerably simpler to handle.

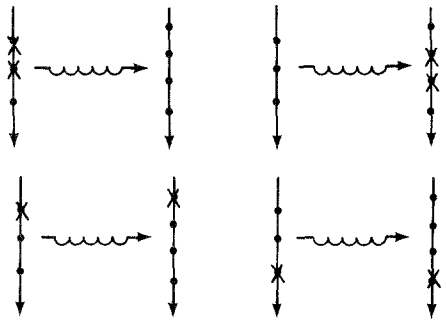
3.3.7. Antiparticles and Crossing. We introduced the concept of process order (and corresponding process graphs) without, up to this point, really making full use of it: we have described ordered transition amplitudes as transitions between ordered channels whose order was described by channel graphs; a complete description consisted of specification of the individual particle parameters t_i, p_i, μ_i and the initial and final channel graphs



The additional information contained in the process graph



by placing all the particles, "in" and "out," on a *common* line and thus defining a *joint* order of "in" and "out" particles has so far remained redundant. However, there have been strong hints that there is a justification for the concept of process graph: e.g., the concept of particle adjacency, which has played such an important role, is natural and easy to formulate in the framework of process graphs, but not so in the channel-graph formalism, where such seemingly different configurations as the following four:



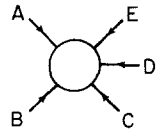
are lumped together under the label "adjacent."

But now that we will present the property of ordered crossing, the process graph will finally appear in its full significance. Ordered crossing

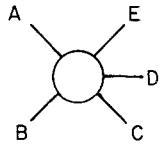
states that all ordered transition amplitudes *with the same process graph* are analytic continuations of one another (in the usual sense of crossing, where incoming particles with positive energy become outgoing antiparticles with negative energy).

Thus to every process graph, e.g.,

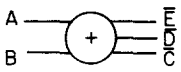
there corresponds an analytic function, denoted by

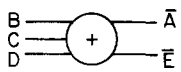


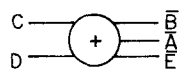
or just



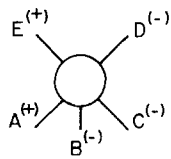
or by $T(A, B, C, D, E)$, which describes all ordered transition amplitudes with that process graph, e.g.,



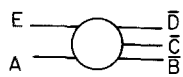




(these three are all related by ordered crossing), in the following sense: when all momenta are real, and the energies of an adjacent set of particles, e.g., E and A, are positive, and the others negative, then

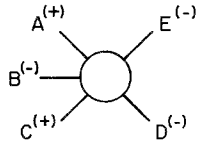


in that region equals

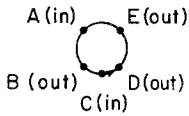


The path of analytical continuation is given by the $+i\epsilon$ rules.

In real momentum regions where the set of particles of positive energy (and hence the set of particles with negative energy) is not adjacent, e.g.,

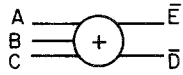


we can regard this function as *defining* the probability amplitude of the ordered process

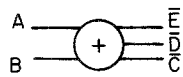


which is not an ordered transition. Thus, as promised at the beginning of this chapter, ordered transition amplitudes generate, by analytic continuation, amplitudes for all ordered processes. These newly defined amplitudes play just as important a role in the calculation of physical amplitudes as the ordered transition amplitudes themselves, as will be shown in Section 6.

The *proof* that (a) to every particle t there exists an antiparticle $\bar{t}$ with the same mass and spin, but with opposite internal additive quantum numbers; and (b) when one takes the appropriate analytic continuation of, say



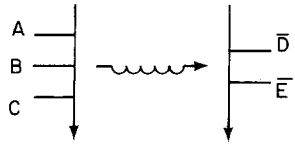
to a real region where $E_\epsilon < -m_c$, one obtains



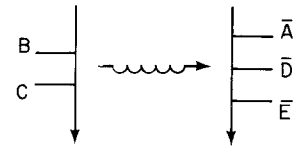
proceeds in a manner closely analogous to that of usual SMT, so we need not present it here in detail. We remark that owing to the present lack of precise knowledge about the singularity structure of amplitudes outside the physical region, one needs to make a certain rather innocuous assumption about that structure in the proof, which is therefore based not solely on the four basic postulates; this is as in usual SMT.

However, since order is a novel dimension here, we need to comment on how its behavior under crossing, namely, the *invariance of the process graph under crossing*, arises from the reasoning of the proof. In other words: why can, e.g., the first (or, respectively, last) particle of an "in" channel of an ordered transition amplitude only be crossed into the first (last) position of the "out" channel, and an intermediate particle not be crossed at all

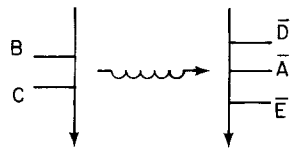
within the framework of ordered transition amplitudes. For example, the ordered transition amplitude



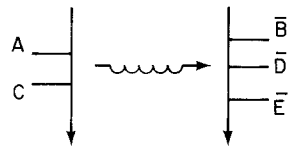
goes by crossing into



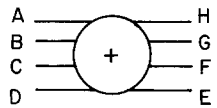
but not into



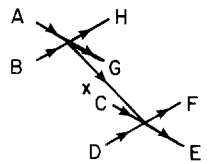
and B cannot be crossed at all into any ordered transition amplitude such as into, e.g.,



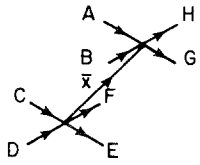
To understand the reason for this, we regard the proof in broadest outline. For example, crossing for the five-particle ordered amplitude is proved by regarding the ordered amplitude



at a pole in S_{ABHG}, corresponding to the ordered Landau diagram



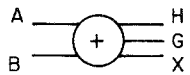
From the analyticity properties of ordered amplitudes one can then easily show that the amplitude must have another ordered Landau diagram



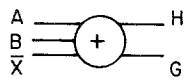
corresponding to a pole $\bar{X}$ in s_{ABHG} of the same mass and spin but opposite internal additive quantum numbers: this pole is called the antiparticle $\bar{X}$ of X. It can then be shown that the residues of these two poles

or, respectively,

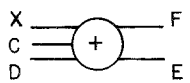
are analytic continuations of one another; with the help of Lorentz invariance one can then deduce that the separate factors



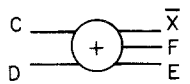
and



or, respectively,



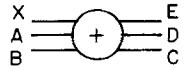
and



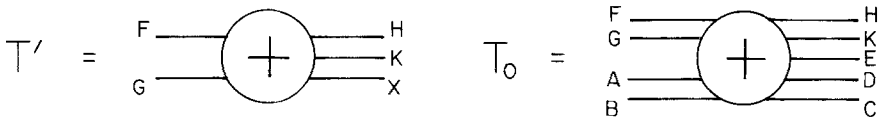
are analytic continuations of one another, which concludes the proof of ordered crossing.

Now we can answer the question posed above as to why the process graph of an ordered transition amplitude remains invariant under crossing to another ordered transition amplitude. In order to be crossed at all, a particle X of an ordered transition amplitude T has to occur as a pole in some larger ordered *transition amplitude* T_0, whose factorized residue is a product of T and some other ordered transition amplitude T'. This can only

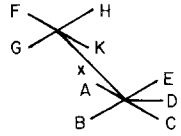
occur if and only if X is the first or the last particle of its channel in T: if it is the first, e.g., in



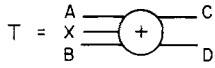
then, e.g.,



and the pole

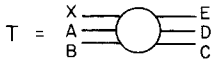


will do the job; and analogously, if X is the last particle, if it is intermediate, e.g.,

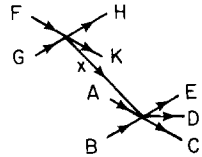


then no T_0 with pole X and T as a residue factor could possibly be an ordered transition amplitude.

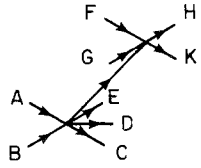
Now regard the case where X is the first particle of its channel in the ordered amplitude T,



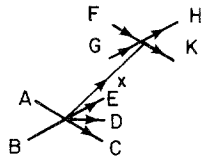
and T', and hence T_0, chosen as in our previous example, with X as a pole in T_0 corresponding to the Landau diagram



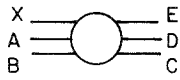
Then the only other Landau diagrams corresponding to poles of T_0 in the variable s_{FGHK} are of the kind



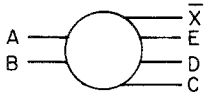
and the proof shows specifically that the antiparticle pole whose existence can be deduced from the analyticity properties corresponds to the Landau diagram



so that crossing relates



with



and with no other ordered amplitude. All this goes to show that the process graph of an ordered amplitude remains invariant under crossing.

3.3.8. The TCP Theorem. The famous TCP theorem that states the invariance of the SM under the TCP transformation follows from repeated crossing, until all "in" particles have been converted into "out" antiparticles with PT-transformed space-time variables p, μ and vice versa and then showing that this series of analytic continuations has taken one back to the starting point in p space, so that the amplitude has the original value again:

$$T(t_1 p_1 \mu_1, t_1 p_2 \mu_2, \ldots, t_m p_m \mu_m; t_i p_i \mu_i, \ldots, t'_n p'_n \mu'_n) = T(\bar{t}'_n p'_n - \mu'_n, \ldots, \bar{t}'_n p'_n - \mu'_n; t_m p_m - \mu_m, \ldots, \bar{t}_1 p_1 - \mu_1)$$

or in terms of the invariant functions:

$$T(t_1 p_1 \mu_1, t_2 p_2 \mu_2 \cdots t_n p_n \mu_n) = T(t_n p_n - \mu_n, \ldots, t_1 p_1 - \mu_1)$$

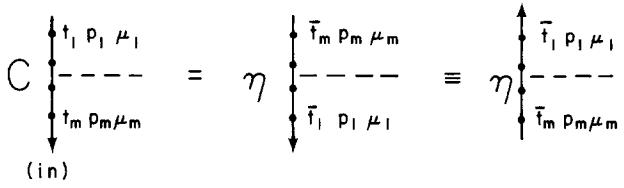
3.3.9. Charge Conjugation. Since we see from the above that

$$TCP|t_1 p_1 \mu_1, t_2 p_2 \mu_2, \ldots, t_m p_m \mu_m\rangle = \langle \bar{t}_m p_m - \mu_m, \ldots, \bar{t}_1 p_1 - \mu_1|$$

and the action of TP is known to be $TP|t_1 p_1 \mu_1, \ldots, t_m p_m \mu_m\rangle = \langle t_1 p_1 - \mu_1, \ldots, t_m p_m - \mu_m|$, one can deduce the action of the charge-conjugation operation C to be

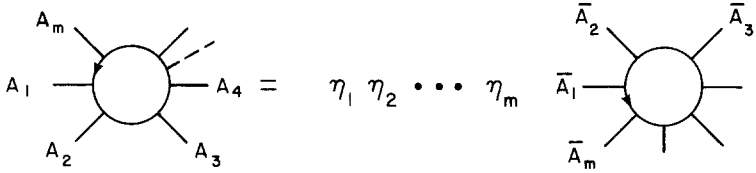
$$C|t_1 p_1 \mu_1 \cdots t_m p_m \mu_m\rangle = \eta|\bar{t}_m p_m \mu_m, \ldots, \bar{t}_1 p_1 \mu_1\rangle$$

or pictorially

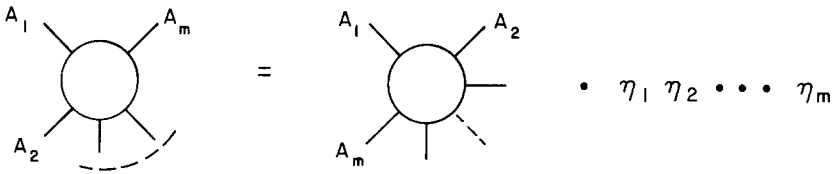


where η is ± 1. Comparing the first and the last expression, we see that C can be regarded as converting t into $\bar{t}$, and reversing the direction of the arrow, i.e., *reversing the direction of the edge orientation.* Note that this is not merely an arbitrary definition of C; C is canonically defined as $T \cdot P \cdot$ (TCP), where (TCP) is the operation determined in the above proof. The phase factor η [which is again the product of the C values (± 1) of the individual particles: $C|t\rangle = \eta_i|t\rangle$], and $t \to \bar{t}$ are familiar from ordinary SMT, but the reversal of edge direction is a specific feature of the ordered theory, which entails some novel consequences.

If we postulate the C invariance of the ordered SM, then



and if all the particles A_i are self-conjugate, then we have



Since there is no reason why the two amplitudes with reversed orientations should have identical values, we *cannot conclude that the product of charge-conjugations η_i has to be 1*; it could just as well be -1. This was first observed by Y. Eylon (1976, 1978). This is in contrast to the situation in usual SMT, where, since there is no reversed order to differentiate them, the two amplitudes are identical, and hence the product of η_i's is unity, which yields the well-known selection rule.

However, we need not be alarmed: as we will see in Section 6, this

selection rule is restored when we pass to the unordered planar approximation of the physical SM.

3.3.10. Other Properties. It is easy to check that the various other properties of the SM mentioned in Section 2, such as Hermitian analyticity, extended unitarity, and the Froissart bound, continue to hold in the ordered framework, since the proofs still go through. The one exception is the spin-and-statistics theorem, for which there can obviously be no counterpart here, since there is no permutation symmetry. With the passage to the planar and higher SM approximations, however, the question will become relevant again since they are unordered, and hence require permutation symmetry. Another property that, we believe, holds for ordered amplitudes, but has not been proved yet, is the *pole conjecture*: any nonzero ordered amplitude has poles in all adjacent channels. (So far, we have only established that it *may* have poles in those channels.)

3.3.11. Duality Properties. Duality is a somewhat ambiguous word that has been used in different contexts with different meanings. One frequent use is to describe the property of, say, a four-particle amplitude to be expressible either in terms of fixed-t dispersion relations or *alternatively* in terms of fixed-s dispersion relations, just in terms of the appropriate discontinuities and without any arbitrary subtraction constants. But this is just the statement of the Regge postulate, since the subtraction constants correspond to the Kronecker deltas in the j plane which are ruled out by analyticity of the second kind. Thus this kind of duality is a direct consequence of SMT. In phenomenology "duality" is used in a less precise but more practical way: there the statement is that the amplitudes can be approximated either by a *few* leading resonances in the s channel, or a *few* leading Regge poles in the t channel, but not by a superposition of both.

The sense in which the word "duality" is used in this work refers to the property of the dual amplitudes to be the sum of permutation-ordered amplitudes with pole and normal-threshold singularities only in adjacent channels; this property brings with it the (approximate) exchange degeneracy, I-spin degeneracy, and Regge pole dominance of physical amplitudes.

We recognize that the planar SM approximation to be introduced in Section 6 is going to be exactly dual in this last sense, too, because of the property of ordered amplitudes to have poles and normal thresholds in adjacent channels only.

3.3.12. Some Regge Properties of Ordered Amplitudes. With the further assumption of analyticity of the second kind, one can construct a Regge theory of ordered amplitudes quite analogous to the usual one.

One difference arises because, say, the four-particle ordered amplitudes

have discontinuities in only two of the three channel invariants, say s and t. Therefore, the two signatured Froissart–Gribov amplitudes

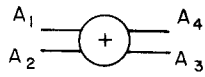
$$A_{j_s}(t) = \frac{1}{2\pi i}\int_{-\infty}^{\infty} da Q_j(z)[D_s(t, z) \pm D_u(t, z)]$$

are equal since $(D_u = 0)$; signature need not be introduced; the Regge trajectories are *strongly exchange degenerate*:

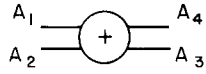
$$\alpha^+(t) = \alpha^-(t), \qquad \beta^+(t) = \beta^-(t)$$

Another difference arising from the absence of the u-channel discontinuities is presumably the absence of any j-plane singularities other than moving Regge poles (in particular there are to be no Regge cuts and fixed poles). The reasons for this conjecture, based on the duality of the amplitudes, are outlined in Chew and Rosenzweig (1978).

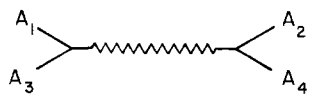
Another interesting point is that, given an ordered amplitude



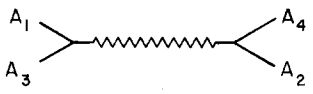
then its Froissart–Gribov projection $A_{j_{13}}(s_{12})$ cannot have a Regge pole in j_{13}, since this would give rise to a pole of



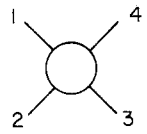
in the variable s_{13}; but we know that such a pole does not exist. Thus there are no Reggeon contributions such as



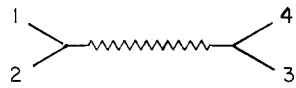
or

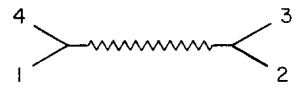


Indeed,

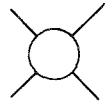


can only have Reggeon contributions of the form

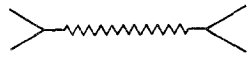




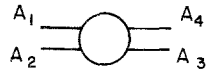
i.e., the order of the particles in



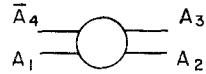
and



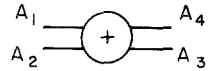
has to coincide. This, especially when coupled with the absence of other Regge singularities, strongly suggests that



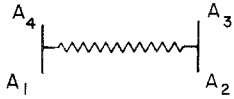
does not have a backward peak (where forward means that p_4 is parallel to p_1), nor does



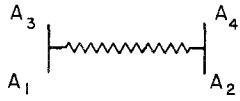
3.3.13. The Connection between Channel Order and Rapidity Order. Regard an ordered transition amplitude, e.g.,



At high energies ($s_{12} \gg 1$), it will be dominated by the leading Regge pole contribution



(the other possible contribution

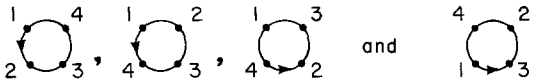


is absent, as we saw).

But at high energies this term has a sharp forward peak within which most events will fall, and thus the rapidity order of the "out" particles coincides with the channel order.

For multiparticle "out" channels the same is true; rapidity order is strongly correlated with channel order.

So can we use this effect to measure the order of individual processes, at least at high energies? This is not possible, since, e.g., the orders



all contribute to that forward peak.

4. INTERNAL QUANTUM NUMBERS IN *S*-MATRIX THEORY

4.1. Introductory Remarks

We have already mentioned several times that the ordered SM has a quarklike spectrum of particles; a statement like that implies certain statements about the possible values of a set of internal quantum numbers.

In our theory particles are labeled by an integer index t only; there is no a priori assignment of internal quantum numbers. If we therefore wish to make predictions about the internal quantum numbers (IQN) of the particles of the spectrum, we have to *assign* them in a way determined by the theory. This assignment necessitates a systematic understanding of the nature of internal quantum numbers in SMT.

This section is dedicated to the development of such an understanding within the framework of conventional, unordered SMT. The results obtained will be useful in the following section in working out the quarklike quantum number predictions of the ordered theory.

Traditionally, internal quantum numbers have been introduced into SMT in a rather unreflected, ad hoc way, so to say "by hand." Specifically, it has been customary to make the following implicit assumptions:

(a) that there be a set of n additive, conserved, integer-valued channel functions $Q^{(1)}, Q^{(2)}, \ldots, Q^{(n)}$, called internal quantum numbers (IQN). By channel function we mean that $Q^{(i)}(t_1 p_1 \mu_1, t_2 p_2 \mu_2, \ldots, t_m p_m \mu_m) \equiv Q^{(i)}(t_1, t_2, \ldots, t_m)$. By additive we mean that

$$Q^{(i)}(t_1, t_2, \ldots, t_m) = Q^{(i)}(t_1) + Q^{(i)}(t_2) + \cdots + Q^{(i)}(t_m)$$

(b) that the complete set of internal and space-time quantum numbers of a particle t: $\{Q^{(1)}(t), \ldots, Q^{(n)}(t), m(t), s(t), p(t)\}$ completely specify the particle t; this set of values may thus be used as a canonical labeling system for the particles, instead of the arbitrary integer index t, or an equally arbitrary name like "proton."

Within this framework, the question as to the spectrum of particles can be formulated in the following way: For which set of integer values of the

$Q^{(i)}$'s, positive value of m, nonnegative integer or half-integer value of s, and ± 1 value of P, does a particle t exist with these values? Or in other words: if an initial state (and therefore, by conservation, the final state) of an amplitude has a set of IQN values $Q^{(1)}, Q^{(2)}, \ldots, Q^{(n)}$, an energy m, an angular momentum s, and a parity P, then does that amplitude have a pole there? To know the answer to this question is to know the spectrum of particles.

We already mentioned in Section 2.1 the unsatisfactory situation that SMT has so far not come up with a calculational method of answering this question, in spite of the bootstrap "profession of faith" that the answer is uniquely determined by the theory and lies latent within it. Therefore, in most calculational applications of SMT one introduces the particles (and their parameters) relevant to the calculation by hand, usually as determined from experiment.

In Section 5 we will see how the IQN values of the spectrum of mesons are *predicted* by the theory of the sequentially ordered SM; as is the OZI rule; in Part Two this is extended to all hadrons.

In this section we examine the status of the above assumptions (a) and (b), i.e., the origin of the IQN formalism in which the specific spectral predictions of the next section are expressed.

4.2. Defining the Problem

What is the origin of the internal quantum numbers and the properties postulated for them in (a) and (b)? The naive notion that certain "charges" simply exist in nature, are indestructible (and hence conserved), occur in multiples of a unit charge, and that it is an intrinsic property of particles to "carry" characteristic amounts of these charges, is rooted in a mechanistic way of thinking totally inappropriate in relativistic quantum mechanics; besides, it is hardly illuminating, since there is no inner logic connecting these assumptions with those of SMT; nor even with one another; they are just grafted onto SMT for convenience.

From the SMT point of view, any particle property should be reducible to observable concepts, such as in particular its interactions with other particles. Since the SM is defined in terms of the particle parameters t, p, μ, with t an integer index, any *other* particle property such as its IQN values has to be derivable from the SM; i.e., given the SM (numerically), we have to be able to assign each particle and channel its set of IQN values.

In the case of the IQN, it is well known how these assignments are performed: it is an observed fact that the SM obeys certain *selection rules* that can be simply expressed by assigning each particle an appropriate set of additive IQNs and demanding their conservation. These selection rules

belong to a class of selection rules that we will call *sector selection rules*; they make no reference to the space-time parameters p, μ of the particles, but, by definition, rule out transitions between certain pairs of channels: $\tau \equiv (t_1, \ldots, t_m) \not\to \tau' \equiv (t_1', \ldots, t_n')$. It is the selection rules that are the most basic reality; IQNs are simply an elegant way of expressing them.

Now, from the SM point of view, it is very natural to ask the following questions: Is the structure of SMT such as to demand the presence of certain sector selection rules? Is every possible channel selection rule expressible as the conservation law for an appropriately defined additive quantum number? If not, then how? Must any particle always be assigned in *integer* multiple of some unit? Is a particle fully specified by its IQN and space-time quantum number values? It is these questions that we will examine now.

We will see that all sector selection rules can be expressed as conservation laws for appropriately defined additive or multiplicative quantum numbers [a multiplicative internal quantum number $\hat{Q}$ of order n is a phase factor $e^{2\pi i m/n}$ (with m an integer, $0 \leqslant m < n$), such that $\hat{Q}(t_1, t_2, \ldots, t_k) = \prod_{i=1}^{k} \hat{Q}(t_i)$] that every additive IQN may be scaled so as to adopt integer values for all particles; and that no two particles may have the same values of all IQNs and m, s, P simultaneously (whence these values determine the particle type t).

4.3. Assumptions, Definitions, and Some Simple Facts

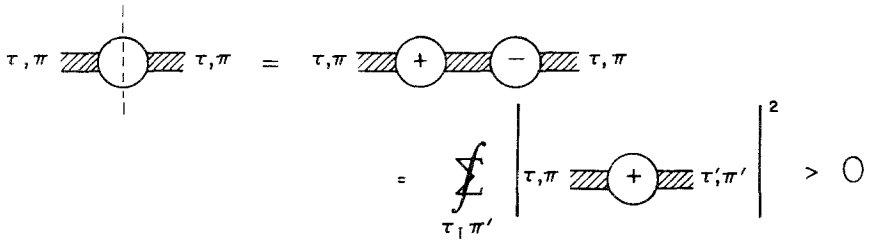
In this section we work within the framework of physical unordered SMT and its basic assumptions (see Section 2, particularly the end of Section 2.2, where the assumptions about the spectrum of particles are specified).

A finite set of particles $\tau = \{t_1, t_2, \ldots, t_n\}$ is called a *channel*. The term "set" here is actually a misnomer, since the same particle may occur several times in one and the same channel. A *state* of this channel is defined by specifying $\Pi \equiv \{\mathbf{p}_1, \mu_1; \mathbf{p}_2, \mu_2, \ldots, \mathbf{p}_n, \mu_n\}$. This state may be denoted by $|\tau, \Pi\rangle$. We call the Hilbert space of all states $|\tau, \Pi\rangle$ with fixed τ the channel space of τ.

$S(\tau; \tau')$ is that function of (Π, Π'), such that $[S(\tau; \tau')](\Pi, \Pi') \equiv S(\tau, \Pi; \tau'\Pi')$. Thus $S(\tau; \tau') = 0$ means $S(\tau, \Pi; \tau'\Pi') = 0$ for all Π, Π'. And similarly for $T(\tau; \tau')$. Two channels τ, τ' are said to *directly communicate* if and only if $T(\tau; \tau') \neq 0$; this is denoted by $\tau \approx \tau'$. Two channels τ, τ' are said to *communicate* if and only if there is a finite sequence of channels $\{\tau_1, \tau_2, \ldots, \tau_m\}$ such that $\tau \approx \tau_1, \tau_1 \approx \tau_2, \tau_2 \approx \tau_{31} \cdots \tau_m \approx \tau'$ (included is the case where $m = 0$, i.e., $\tau \approx \tau'$).

A channel τ is called noninteracting if $T(\tau, \tau') = 0$ for all channels τ'. Otherwise it is called *interacting*. Note that if and only if a channel τ

is interacting, then $T(\tau;\tau) \neq 0$. To see that this is true, regard the discontinuity of the $\tau \to \tau$ forward elastic amplitude



(by hypothesis, since τ was assumed to be interacting). Thus the forward elastic amplitudes of interacting channels are always nonzero.

We will soon see that for the physical SM, all channels are interacting; for the ordered SM that will not be the case.

The communication relation "$\sim$" is an equivalence relation on the set of all interacting channels, since for these it is reflexive ($\tau \sim \tau$), and symmetric ($\tau \sim \tau' \Rightarrow \tau' \sim \tau$) and transitive ($\tau \sim \tau'$, $\tau' \sim \tau'' \Rightarrow \tau \sim \tau''$) by definition. Therefore, the set of all interacting channels decomposes into disjoint equivalence classes S_i of interacting channels in such a way that any two channels from the same S_i communicate, but no two channels from different S_i's communicate. These equivalence classes are called *sectors*. Note that two channels from the same sector need not communicate *directly*. (We will see that all channels from the same sector are assigned identical IQN values, while conversely channels from different sectors differ in at least the value of one IQN.)

A *selection rule* demands that, for a given state $|i\rangle$, there exist a subspace F_i (of the Hilbert space of final states), of nonzero measure, such that for every $|f\rangle \in F_i \langle i|S|f\rangle = 0$. We know of three kinds of selection rules:

(a) Space-time selection rules, involving only p's and μ's; they are generated by Poincaré invariance (4-momentum, angular momentum, and parity conservation). Occasionally they can block the *direct* communication between two channels completely.

(b) Sector selection rules dividing the set of all channels into mutually noncommunicating sectors. It is these that will give rise to additive and multiplicative conserved IQNs.

(c) Selection rules generated by internal symmetries of the $SU(n)$ kind, in particular $SU(2)$. These occasionally block the direct communication of two channels; however, in general they cannot be expressed in terms of channels, but only in terms of linear combinations of states from different channels.

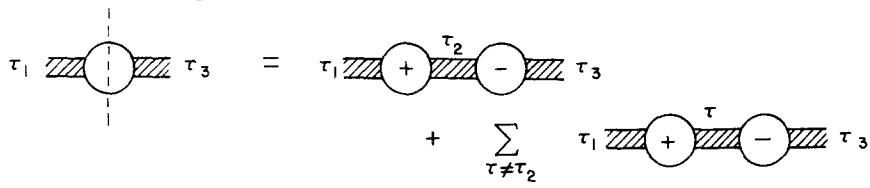
An important remark is that if two one-particle channels $\{t\}$ and $\{t'\}$ both communicate directly with a channel τ, then the particles t and t' cannot share the same values of mass, spin, and parity simultaneously: If $\{t\} \approx \tau$, $\{t'\} \approx \tau \Rightarrow (m(t), s(t), p(t)) \neq (m(t'), s(t'), p(t'))$. For if this were not the case, then since $T(\tau; \tau)$ has both t and t' as factorizable poles, the net effect would be a nonfactorizable pole in $T(\tau; \tau)$ of definite spin and parity, and of course mass. But pole factorization is a direct consequence of unitarity, and has to be demanded for a particle pole. Thus we would have a pole not corresponding to a particle, this pole would mediate long-range effects like an ordinary particle, e.g., a double scattering

it would correspond to an irreducible representation of the Poincaré group; and yet it would not be a particle, since the information the second scattering system gets about the first one cannot be condensed into particle degrees of freedom. One might even argue that t and t' should not have the same mass (even if they differ in s and P), if they both communicate directly with a common channel τ. One might hope to extend the argument to the case where there is no such common channel τ with which both t and t' communicate directly, but instead a chain of channels τ_i such that $\{t\} \approx \tau_1, \tau_1 \approx \tau_2, \ldots, \tau_n \approx \{t'\}$; (this is admittedly somewhat academic since we know of no case where two channels from the same sector have no common channel τ that they both communicate directly with); with this generalization, we have the theorem: No particles from the same sector can be degenerate in mass, spin, and parity simultaneously (or, respectively, in mass alone).

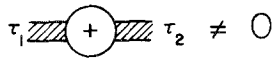
4.4. Direct Communication of Channels within a Sector

Regard three channels τ_1, τ_2, τ_3 such that $\tau_1 \approx \tau_2, \tau_2 \approx \tau_3$. We ask the question whether this implies that $\tau_1 \approx \tau_3$. If it does, then "$\approx$" is transitive, and then all channels within a sector would communicate directly, and the distinction between communication and direct communication would be eliminated.

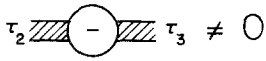
Unitarity implies



and since



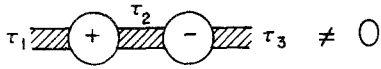
and



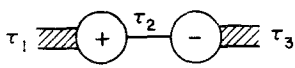
one might be tempted to conclude that at least the first term on the right-hand side is nonzero, and hence in all probability the whole expression.

But this argument is fallacious; for one thing, owing to the presence of the discrete parameters s, P, the product

might well be identically zero even if the factors are not. For example, if channel τ_1 consisted of two scalars, τ_3 of a scalar and an isoscalar, and τ_2 was an arbitrary three-particle channel all in the same sector, then angular momentum and parity would prohibit direct communication between τ_1 and τ_3, even though $\tau_1 \approx \tau_2$, $\tau_2 \approx \tau_3$. Furthermore, even if the product is nonzero, the integral over the τ_2 phase space need not be; and even if



the other terms on the right-hand side might still cancel it; all that is required for this to be possible is that there be other channels τ with the same threshold energy and number of particles as τ_2, since only such terms would have the same analytic form as the τ_2 contribution so that a finite number of them could cancel it. For example, τ_1 and τ_3 could have a distinct and definite value of C or G parity, whereas τ_2 has no such definite value. Or τ_1 could consist of two $I = 0$ particles, τ_2 of two $I = 1$ particles, and τ_3 of an $I = 0$ and an $I = 2$ particle. And indeed, as the above examples show, selection rules that block the direct communication of certain pairs of channels within a sector do exist, even if they tend to be exceptional. In the special case where τ_2 is a one-particle channel, however, it is true that $\tau_1 \approx \tau_2$, $\tau_2 \approx \tau_3$ implies that $\tau_1 \approx \tau_3$. This is so because now



is not an integral, but merely a product of two nonzero factors; and since there can be no other particle of the same mass in this sector, as we saw before, the finite sum of remaining terms on the right-hand side cannot reproduce a τ_2 pole and hence cannot cancel the τ_2 term. Owing to the

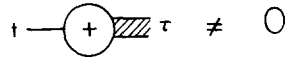
assumption that there are no zero-mass particles and no accumulation points in mass, the number of channels that can contribute to a unitarity sum at a given energy is finite. And a finite number of pole terms with masses $\neq m$, together with a finite number of terms corresponding to other kinds of singularities cannot add up to a pole of mass m. Therefore the proof goes through and shows that $\tau_1 \approx \tau_3$, as claimed.

As a result, all the channels of a sector that communicate directly with a given particle t communicate directly with one another. Thus to every particle t in a sector S there exists a set $S^{(t)}$ of channels communicating directly with t and with one another; the union of all these $S^{(t)}$ is equal to S itself, owing to the pole conjecture. But the $S^{(t)}$ are not disjoint; indeed most of the channels will be in all $S^{(t)}$ simultaneously.

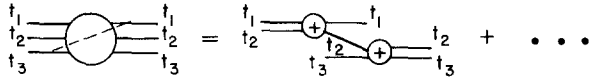
4.5. All Channels of the Physical SM Are Interacting

We now show that in a sense the concept of "interacting channel" is trivial for the physical SM, since *all* channels are interacting. This constitutes a major difference between the physical and the ordered SM: for the latter it is precisely the fact that not all channels are interacting that gives rise to flavor.

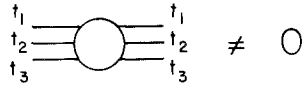
First we note that every one-particle channel $\{t\}$ is interacting (since as t is a particle of the SM spectrum, there must be at least one nonzero amplitude with t as one of the external particles:



Next we observe that *if* all *two*-particle channels are interacting, then all channels are interacting. To see this, choose an arbitrary three-particle channel $\tau = \{t_1, t_2, t_3\}$. Then



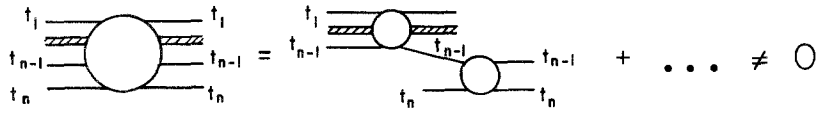
Since the first term on the right-hand side is nonzero by hypothesis, and the remaining terms cannot cancel it (as we saw),



and so τ is interacting; since τ was arbitrary, all three-particle channels are interacting.

The same kind of reasoning shows that if all $(n-1)$-particle channels

are interacting, then all n-particle channels are, too: choose an arbitrary n-particle channel $\tau = \{t_1, \ldots, t_n\}$. Then



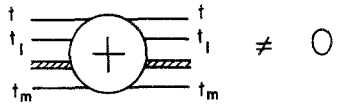
(by hypothesis), and so τ is interacting. Thus, by induction, we have shown that if all two-particle channels are interacting, then all channels are interacting. So if there are to be any noninteracting channels at all, then there have to be noninteracting two-particle channels.

Now it is clear that any two particles t_1, t_2 occur together as external particles of a nonzero amplitude if and only if $\tau = \{t_1, t_2\}$ is interacting. This is so because, if t_1, t_2 occur together on a nonzero amplitude, then by successive crossing one can bring t_1 and t_2 to one side (say, as initial particles), and all the other particles to the other side:

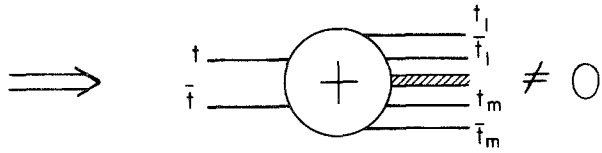
Therefore, $\tau = \{t_1, t_2\}$ is interacting. The converse is trivial.

Next, we show that the relation "$\backsim$" between two particles t_1 and t_2 defined by "$t_1 \backsim t_2 \Leftrightarrow \tau \equiv \{t_1, t_2\}$ interacting," is an equivalence relation.

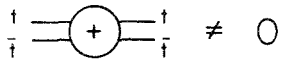
The relation is *reflexive*: $t \backsim t$ for all t. This is because since t is a particle, it occurs in a nonzero amplitude, from which we see that there is an interacting multiparticle channel that contains t; for this channel,



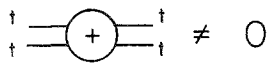
and so by crossing



therefore



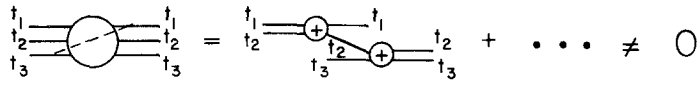
and so by crossing



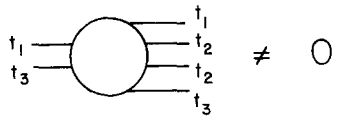
thus $\{t, t\}$ is interacting, for all t, and hence by definition $t \backsim t$ for all t.

The relation is *symmetric*: $t_1 \leftharpoonup t_2 \Leftrightarrow t_2 \leftharpoonup t_1$. This is clear since $\{t_1, t_2\} \equiv \{t_2, t_1\}$; and so if one of these channels is interacting, so is the other.

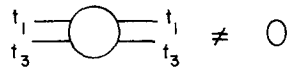
The relation is *transitive*: $t_1 \leftharpoonup t_2$, $t_2 \leftharpoonup t_3 \Rightarrow t_1 \leftharpoonup t_3$: If $\{t_1, t_2\}$ and $\{t_2, t_3\}$ are interacting, then



and so by crossing



and hence (discontinuity relation)



$\{t_1, t_3\}$ is interacting.

Since $\leftharpoonup$ is an equivalence relation, the set of all particles decomposes into equivalence classes such that any two particles from the same equivalence class form an interacting channel, whereas any two particles from different equivalence classes form a noninteracting channel. As we saw before, this implies that no two particles from different equivalence classes can occur in a nonzero amplitude together. Taking note of this fact, we see from the cluster decomposition that the SM factorizes into a product of S matrices, each defined on the Hilbert space generated by the particles from one of the equivalence classes alone. Each of these SMs is unitary and has all the other properties of a physical SM.

The physical interpretation of this situation, if there is more than one equivalence class, is that of a set of mutually noninteracting worlds, one corresponding to each equivalence class of particles. And since we are living in one of these worlds, and are not concerned (here) with metaphysics, we may define the sub-S matrix pertaining to our world as *the* physical SM.

In this sense, all channels of the physical SM are interacting.

4.6. The Sector Group

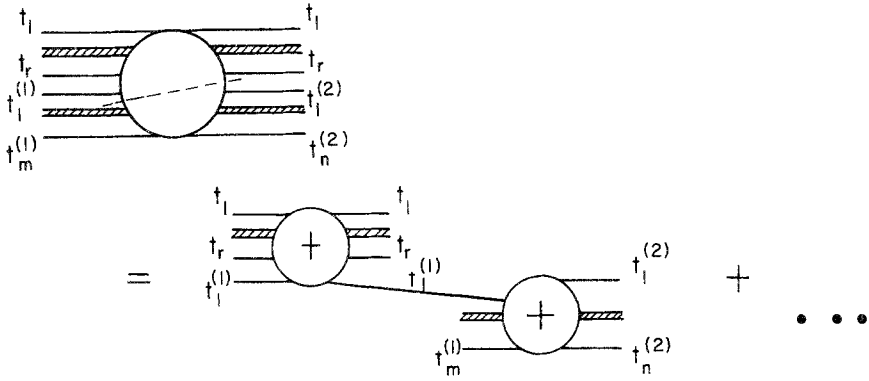
We are now ready to introduce the concept of the *sector group*. The elements of this group are the sectors of the physical SM. We may denote a sector S by $[\tau]$, where τ is any one of the channels contained in S. Given any two channels, τ and τ', we define $\tau \circ \tau'$ to be the channel that contains

all the particles of τ and those of τ'; if $\tau = \{t_1, \ldots, t_m\}$, and $\tau' = \{t'_1, \ldots, t'_n\}$, then $\tau \circ \tau' \equiv \{t_1, t_2, \ldots, t_m, t'_1, \ldots, t'_n\}$.

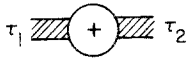
The group multiplication "$\circ$" on the set of sectors is defined by $[\tau] \circ [\tau'] \equiv [\tau \circ \tau']$, where the operation of the right-hand side was just defined above.

We first have to show that this operation is well defined, i.e., that the product $S_1 \circ S_2$ of two sectors S_1 and S_2 does not depend on which particular channels are chosen to represent the sectors in the defining equation: if $S_1 = [\tau_1] = [\tau'_1]$, and $S_2 = [\tau_2] = [\tau'_2]$, then $[\tau_1 \circ \tau_2] = [\tau'_1 \circ \tau'_2]$ is to be proved; to do this we proceed in three steps.

(1) First we show that if two channels τ_1, τ_2, communicate directly, $\tau_1 \approx \tau_2$, then for any channel τ one has $\tau \circ \tau_1 \approx \tau \circ \tau_2$. So let us assume $\tau_1 = \{t_1^{(1)}, t_2^{(1)}, \ldots, t_m^{(1)}\}$, $t_2 = \{t_1^{(2)}, t_2^{(2)}, \ldots, t_n^{(2)}\}$, $\tau = \{t_1, t_2, \ldots, t_n\}$, and $\tau_1 \approx \tau_2$. Then



The first term of the right-hand side is nonzero, since it is a product of a (nonzero) elastic amplitude, and of



which is nonzero by hypothesis. And for the usual reasons this term cannot be canceled by the remaining terms. Therefore, $\tau_1 \circ \tau_2 \approx \tau_2 \circ \tau$.

(2) If any two channels τ_1, τ_2 communicate, $\tau_1 \sim \tau_2$, then for every channel τ, $\tau \circ \tau_1 \sim \tau \circ \tau_2$ holds.

For if $\tau_1 \sim \tau_2$, then by definition there exists a finite sequence of channels $\{\tau^{(1)}, \tau^{(2)}, \ldots, \tau^{(n)}\}$, such that $\tau_1 \approx \tau^{(1)}, \tau^{(1)} \approx \tau^{(2)}, \ldots, \tau^{(n)} \approx \tau_2$. Applying the lemma of (1) repeatedly, this implies that $\tau \circ \tau_1 \approx \tau \circ \tau^{(1)}$, $\tau \circ \tau^{(1)} \approx \tau \circ \tau^{(2)}, \ldots, \tau \circ \tau^{(n)} \approx \tau \circ \tau_2$ and hence $\tau \circ \tau_1 \sim \tau \circ \tau_2$, as claimed.

(3) Now we prove the main claim: consider any four channels τ_1, τ'_1, τ_2, τ'_2 such that $\tau_1 \sim \tau'_1, \tau_2 \sim \tau'_2$. Then by (2) $\tau_1 \circ \tau_2 \sim \tau'_1 \circ \tau_2$, and also

$\tau_1' \circ \tau_2 \sim \tau_1' \circ \tau_2'$; hence by the transitivity of "$\sim$", $\tau_1 \circ \tau_2 \sim \tau_1' \circ \tau_2'$, as claimed. In other words, if $[\tau_1] = [\tau_1']$, and $[\tau_2] = [\tau_2']$, then $[\tau_1 \circ \tau_2] = [\tau_1' \circ \tau_2']$, hence the group multiplication "$\circ$" on the set of sectors is well defined.

This multiplication is associative and commutative because the operation "$\circ$" of channel composition is, by definition; so

$$[\tau_1] \circ ([\tau_2] \circ [\tau_3]) = [\tau_1] \circ [\tau_2 \circ \tau_3] = [\tau_1 \circ (\tau_2 \circ \tau_3)] = [(\tau_1 \circ \tau_2) \circ \tau_3]$$
$$= [\tau_1 \circ \tau_2] \circ [\tau_3] = ([\tau_1] \circ [\tau_2]) \circ [\tau_3],$$

and

$$[\tau_1] \circ [\tau_2] = [\tau_1 \circ \tau_2] = [\tau_2 \circ \tau_1] = [\tau_2] \circ [\tau_1],$$

The unit element of the group is given by $[\tau \circ \bar{\tau}]$, where τ is an arbitrary channel; it is well defined since for any two channels τ, τ', we have $\tau \circ \bar{\tau} \approx \tau' \circ \bar{\tau}'$ (obtained by crossing from the statement $\tau \circ \tau' \approx \tau \circ \tau'$), and thus $[\tau \circ \bar{\tau}] = [\tau' \circ \bar{\tau}']$.

The inverse element $[\tau]^{-1}$ of any sector $[\tau]$ is the sector containing the charge-conjugate channel: $[\tau]^{-1} \equiv [\bar{\tau}]$, since $[\tau] \circ [\bar{\tau}] = [\tau \circ \bar{\tau}] =$ unit element.

Thus the sectors indeed form an *Abelian group* under the operation "$\circ$": the sector group.

4.7. The Origin of Internal Quantum Numbers

Any set of elements of a group such that no element of the set can be expressed as an appropriate product of the others is called independent. For any denumerable Abelian group we can always choose a set of independent generators $\{S_1, S_2, \ldots\}$ (depending on the group, it could be a finite or an infinite set), such that every group element S can be expressed as $S = S_1^{n_1} \circ S_2^{n_2} \circ \cdots \prod_i S_i^{n_i}$, where the n_i's are integers only a finite number of which are nonzero. If the order of a generator S_i (i.e., the smallest positive integer r_i such that $S_i^{r_i} =$ unit element) is finite, then we make the additional convention that $0 \leqslant n_i < r_i$ (since n_i is only defined modulo r_i). Then the representation $S = \prod_i S_i^{n_i}$ is unique, and the set of integers $\{n_1, n_2, \ldots\}$ fully characterizes a group element.

We now apply all this to the Abelian sector group: then we choose a set of sectors $\{S_1, S_2, \ldots\}$ that form an independent generator set. And every sector S is then characterized by its set of integers $\{n_1, n_2, \ldots\}$, defined as above by $S = \prod_i S_i^{n_i}$.

We now define a set of sector functions $Q_i(\tau)$, one for each generator S_i, in the following way: for a sector $[\tau] = \prod_i S_i^{n_i}$, $Q_i([\tau]) \equiv n_i$. The full

set $\{Q_1(\tau), Q_2(\tau), \ldots\}$ thus characterizes the sector to which that channel belongs: $[\tau] = \prod_i S_i^{Q_i([\tau])}$ in terms of the fixed set of generators S_i.

If S_i is of infinite order, $r_i = \infty$, then $Q_i(\tau)$ is an *additive*, conserved quantum number; additive, because if $\tau = \prod_j S_j^{n_j}$, $\tau' = \prod_i S_j^{n'_j}$, then

$$Q_i(\tau \circ \tau') = Q_i([\tau \circ \tau']) = Q_i([\tau] \circ [\tau]) = Q_i\left(\prod_j S_j^{n_j} \cdot \prod_i S_j^{n'_j}\right)$$

$$\times\, Q_i\left(\prod_j S_j^{(n_j + n_j')}\right) = n_i + n_i' = Q_i(\tau) + Q_i(\tau')$$

and conserved because "in" and "out" channels are in the same sector, and thus have the same Q_j's.

If S_i is of finite order r_i, then we may still define a corresponding conserved quantum number $Q_i(\tau)$ as before, but it is not additive, since $Q_i(\tau)$ is only determined modulo r_i, and we have chosen the convention $0 \leqslant Q_i(\tau) < r_i$; thus $Q_i(\tau \circ \tau') = Q_i(\tau) + Q_i(\tau') - mr_i$, where m is a non-negative integer chosen so that $0 \leqslant Q(\tau \circ \tau') < r_i$.

Therefore, when r_i is finite, $Q_i(\tau)$ is not that useful. Instead, we define

$$\hat{Q}_i(\tau) \equiv \exp\left[\frac{2\pi i Q_i(\tau)}{r_i}\right]$$

This conserved quantum number has the nice property that it is multiplicative:

$$\hat{Q}_i(\tau \circ \tau') = \exp\left[\frac{2\pi i Q_i(\tau \circ \tau')}{r_i}\right] = \exp\left\{\frac{2\pi i (Q_i(\tau) + Q_i(\tau') - mr_i)}{r_i}\right\}$$

$$= \exp\left[\frac{2\pi i Q_i(\tau)}{r_i}\right] \cdot \exp\left[\frac{2\pi i Q_i(\tau')}{r_i}\right] = \hat{Q}_i(\tau) \cdot \hat{Q}_i(\tau')$$

Thus each generator of the sector group corresponds to an additive conserved internal quantum number Q_i or a multiplicative one $\hat{Q}_i$, depending on whether it is of infinite or finite order.

Our choice of the set of generators was to a large extent arbitrary. If we choose another set $\{S_1', S_2', \ldots\}$, with $S_i = \prod_k S_k'^{l_{ik}}$, then the corresponding Q_i''s are related to the Q_k's by a linear combination $Q_k = \sum_i l_{ik} Q_i'$, in the case when all generators are of infinite order.

We have now seen that sector selection rules are always expressible as conservation laws of appropriately defined additive and/or multiplicative internal quantum numbers; that the additive ones have integer values; and from the nondegeneracy of particles within a sector we know that specifying the IQNs (which specifies the sector) and (m, s, P) specifies the particle.

5. QUARK-MODEL PROPERTIES OF THE ORDERED S MATRIX

5.1. Introductory Remarks

Why can the quantum number and multiplet structure of all known hadrons be predicted by a simple nonrelativistic model according to which each hadron consists of a characteristic combination of fundamental constituents called quarks? These quarks q_k, of which there are only a few different kinds (although the number of them required to explain experimental data is increasing by the year) differing in "color" and "flavor," are allowed to combine according to the zero-triality rule, e.g., $(q_i\bar{q}_j)$ (nonexotic mesons), $(q_kq_jq_k)$ or $(\bar{q}_i\bar{q}_j\bar{q}_k)$ (baryons or antibaryons, respectively), $(q_iq_j\bar{q}_k\bar{q}_1)$ (baryonium), etc. Until recently all hadrons found in nature were nonexotic mesons and baryons (or, respectively, antibaryons).

Some readers may be inclined to answer the above question in a very straightforward way: Because hadrons *do* consist of quarks, plus maybe some other elementary constituents, like gluons. That these "particles" do not occur in nature, but merely manifest themselves indirectly in the hidden inner recesses of hadrons, is attributed to their "confinement." This view, that some gauge theory of local elementary fields like QCD will emerge as the ultimate particle theory, is popular today. We have already voiced our reservations against this approach in Section 2.1 and will not pursue the matter any further here. Since these more sophisticated versions of quark theories have not till now produced any observable predictions in strong interactions that go beyond those of the naive nonrelativistic quark models (and indeed strictly speaking do not even reproduce these, since the confinement problem is not solved at this point), we will in the following be referring only to the latter.

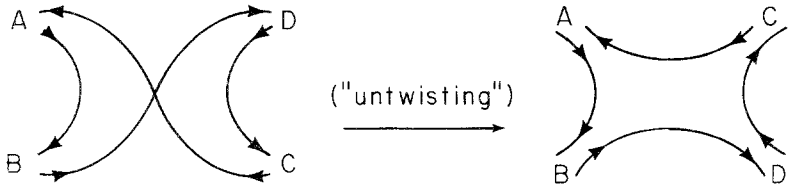
Here we wish to show that the ordered SM approach is able to reproduce, without having been constructed specifically for that purpose, most of the predictions of the quark model in the field of strong interactions. And in so doing we will gain a completely new perspective on some of the basic concepts of the quark model. In particular, quark diagrams will be seen to be equivalent to process-order graphs, and quarks as order relationships between particles rather than as particles themselves. In this chapter we will deal only with nonexotic mesons; the generalization to all hadrons is presented in Part Two.

We first show that so-called order selection rules (which prohibit, say, a particle t' from being the successor of another particle t:

in any nonzero amplitude) allow us to categorize all particles and ordered channels into ordered sectors denoted by (i, j) in such a way that a particle t' from (k, l) can be the successor of a particle t from (i, j) in a nonzero ordered amplitude if and only if $j = k$; here the integer indices $i, j, k, \ldots$ are called flavor indices. As a result, the edges of the process graphs can be labeled with the flavor indices and are seen to be formally identical to conventional mesonic quark diagrams. Thus the ordered amplitudes are seen to be, by their very nature, quark diagram amplitudes obeying the OZI rule, and the particles are identified as nonexotic mesons. This identification is confirmed when, with an eye on the transition to the unordered, physical SM, and based on the ideas of Section 5, we assign additive conserved internal quantum numbers to these particles, and find that they coincide with the mesonic quantum numbers predicted by the quark model and experimentally observed. Finally, after some remarks about internal symmetry and multiplet structure, we compare the quark and ordered SM points of view and their successes for strong interactions.

5.2. Process-Order Graphs and Quark Diagrams

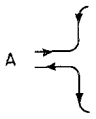
A mesonic quark diagram can always be drawn in exactly one way as a planar diagram, i.e., without any quark lines crossing one another on the plane of the paper. For example,



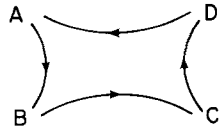
This defines for each quark diagram a unique sequential directed order identical to that of a process graph.

Then what is the *formal* difference between a quark diagram and a process-order graph? First of all, there is a difference in the way that particles are represented: in process graphs as a 2-vertex

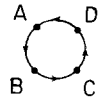
in quark diagrams as



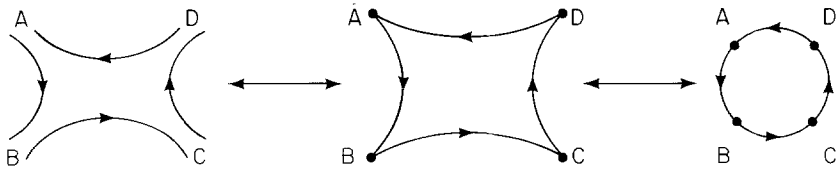
This difference, however, is superficial: The graph



contains no more and no less information than the graph



They can be converted into one another by bringing the two quark lines of a particle to a point and then deforming to a circle



But there is a second, more important difference between quark diagrams and process graphs: in the former, the quark lines are differentiated by a flavor label (and in some models by a color label) whereas an edge of a process graph is simply an edge, without further qualification.

However, in the following we will show that flavor labels are a natural feature of the edges of a process graph, if there are order selection rules at work in the OSM; and with that the formal equivalence of quark diagrams and process graphs will have been established.

5.3. Order Selection Rules

We discover the concept of order selection rule immediately when we ask ourselves the question whether all ordered channels are interacting (in the sense of Section 4): Is $T(\tau, \tau) \neq 0$ for every ordered channel τ? In the case of the physical, unordered SM we were able to prove it, but here that proof does not go through; the reason it does not go through is that now an ordered channel is not just a set of particles of an ordered amplitude, but a set of *adjacent* particles of an ordered amplitude.

Thus while we prove in a formally identical way (see the proof in Section 4) that "if all ordered two-particle channels are interacting, then all ordered channels are interacting," we can no longer argue that the two properties of two particles t, t' of (a) (t, t') being an interacting ordered channel, (b) t and t' occurring in the same nonzero ordered amplitude are

equivalent, since even if t and t' occur in the same nonzero amplitude, they cannot be made to become adjacent by crossing (crossing preserves the order); and hence the rest of the argument that all channels must be interacting breaks down.

Thus there could well be certain noninteracting ordered channels. And if there are, then the proof shows that there must be noninteracting ordered *two*-particle channels. In a sense these are the basic source of noninteractingness: if an ordered channel is noninteracting it is always because some ordered two-particle subchannel is noninteracting.

Let us regard such a noninteracting ordered two-particle channel, say (t, t'). Then the channel graph

cannot appear in any nonzero ordered amplitude. We call this an *order selection rule* since it rules out certain order patterns in the ordered SM, i.e., excludes certain ordered channels as noninteracting.

We will now examine what inferences we can draw if we assume that there are order selection rules at work in the OSM. In order to study them in isolation we make the assumption that there are no *other* channel selection rules at work in the OSM, in particular not the kind studied in Section 4 that give rise to additive conserved quantum numbers. To be more precise, we will assume that *if every bisection of a process graph yields two interacting channels, then the corresponding ordered amplitude is not identically zero*. We will see later, when this hypothesis is translated into the language of the physical, unordered SM, that what remains of the order selection rules is a set of so-called *canonical conserved additive internal quantum numbers* that are precisely the set of additive IQNs found in nature; that no additional additive or multiplicative IQNs of the type permitted by the structure of SMT (see Section 4) occur in nature: the order selection rules we are about to examine are thus the origin of all IQNs known. We find this intuitively appealing, since order selection rules are "local," referring as they do only to the ability of a particle to link up with a neighbor.

5.4. Successor and Predecessor Classes

Choose any particle t. We call any particle t' such that (t, t') is an interacting ordered channel a *successor* of t, and any particle t'' such that (t'',t) is interacting a *predecessor* of t. The set of all successors of t is called the *successor class* of t, denoted by S_t; the set of all predecessors of t is called the predecessor class of t, denoted by P_t. If there are no order selection rules,

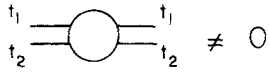
then for every t, S_t and P_t contain all particles. Otherwise some particles will be in S_t (or, respectively, P_t), while other particles will not. The various sets S_t will form a covering of the set of all particles, since every particle is the successor of at least one other particle (and similarly for P_t). Thus, by

will occur in a nonzero ordered amplitude if and only if $t' \in S_t$, or what is equivalent, $t \in P_{t'}$.

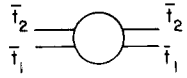
We now prove the crucial property that the covering S_t is *disjoint*, i.e., two successor classes S_{t_1}, S_{t_2} are either equal or they have no particle in common: either $S_{t_1} = S_{t_2}$ or $S_{t_1} \cap S_{t_2}$ (the empty set).

To prove this, we first make two observations:

(a) If (t_1, t_2) is interacting, then so is $(\bar{t}_2, \bar{t}_1)$. This is because if

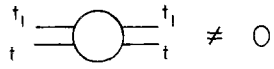


then by C invariance

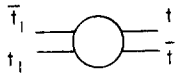


is nonzero, too.

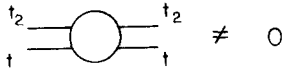
(b) If t_1 and t_2 have a common successor t, then $\bar{t}_2$ is a successor of t_1, and $\bar{t}_1$ a successor of t_2: if $t \in S_{t_1}$, $t \in S_{t_2} \Rightarrow \bar{t}_1 \in S_{t_1}$, $\bar{t}_2 \in S_{t_2}$. This is so because by hypothesis



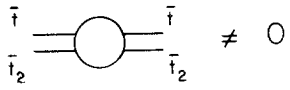
and therefore by crossing



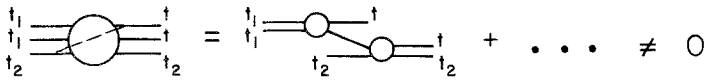
is too; also



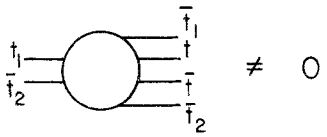
and hence by C invariance



So, by the ordered discontinuity relation,



since the first term on the right-hand side is the product of two nonzero factors and cannot be canceled by any further terms for the usual reason (see Section 4). Hence by crossing



and thus $(t_1, \bar{t}_2)$ interacting; and by (a) $(t_2, \bar{t}_1)$ is interacting too. And quite analogously, if two particles t_1 and t_2 have a common predecessor t, then $\bar{t}_1 \in P_{t_2}$, $\bar{t}_2 \in P_{t_1}$.

We are now ready to prove the disjointness of successor classes. Let us assume that two successor classes S_{t_1} and S_{t_2} are not disjoint: then they contain a common particle, say t: $t \in S_{t_1}$, $t \in S_{t_2}$. But then, as we just saw, $\bar{t}_2 \in S_{t_1}$ and $\bar{t}_1 \in S_{t_2}$. Now choose an arbitrary particle $t' \in S_{t_1}$. Because $\bar{t}_2$ and t' have a common predecessor, namely, t_1, we conclude that $t_2 \in P_{t'}$, i.e., $t' \in S_{t_2}$. Thus an arbitrary particle $t' \in S_{t_1}$ was found to be $t' \in S_{t_2}$; hence $S_{t_1} \subset S_{t_2}$. Analogously one proves $S_{t_2} \subset S_{t_1}$. Hence $S_{t_1} = S_{t_2}$. Thus we have shown that two successor classes are either disjoint or equal.

In an analogous way one shows that predecessor classes are either disjoint or equal, and that the predecessor classes thus form a disjoint covering of the set of all particles.

We now note that there is a one-to-one relation between predecessor and successor classes: choose any particle t, and any successor of t, say t'. Then $t \in P_{t'}$, $t' \in S_t$. Then, as we saw, any $t_1 \in P_{t'}$ is a predecessor of any $t_1' \in S_t$; and any $t_1' \in S_t$ is the successor of any $t_1 \in P_{t'}$. This defines a one-to-one relationship between predecessor and successor classes: for a predecessor class P and a successor class S that correspond to one another, every particle from S is a successor of every particle from P, and vice versa.

Till now we have labeled successor and predecessor classes by one of the particular particles whose successor or predecessors they are: S_t, P_t. Now we label the distinct successor and predecessor classes with integers i, j: S^i, P^j in such a way that a successor class S and a predecessor class P receive the same label i if and only if they correspond to one another in the above sense. The labels i, j are called *flavor labels*. Owing to the one-to-one correspondence between successor and predecessor classes the indices i and j range over the same index set.

Given any successor class S_i, and an arbitrary particle $t \in S_i$; then since, as we saw, $(\bar{t}, t)$ is always an interacting channel, we have $\bar{t} \in P_i$. And the converse statement is also true. Thus we see that $\bar{P}_i = S_i$, $\bar{S}_i = P_i$.

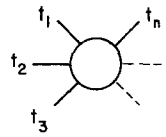
5.5. The Flavor Classes (i, j)

Regard any particle t. It is in a definite predecessor class P_j, and in a definite successor class S_i. Then we say that t is in the flavor class (i, j): $t \in (i, j)$. Obviously, $\bigcup_j (i, j) = S_i$, and $\bigcup_i (i, j) = P_j$, and $(i, j) = S_i \cap P_j$.

The way we have labeled the S's and P's correspondingly, we see easily that if $t \in (i, j)$, and $t' \in (k, l)$, then (t, t') is an interacting channel if and only if $j = k$. And similarly for an ordered channel $(t_1, t_2, \ldots, t_m)$, with $t_1 \in (i_1, j_1)$, $t_2 \in (i_2, j_2), \ldots, t_m \in (i_m, j_m)$, the conditions $i_2 = j_1, i_3 = j_2, \ldots, i_n = j_{n+1}$ are sufficient and necessary for it to be interacting. That they are necessary is obvious; that they are sufficient follows from the fact that every two-particle subchannel is interacting, which, as we saw, guarantees that the whole channel is.

Since, as we saw, $\bar{S}_i = P_i$, $\bar{P}_i = S_i$, it is clear that if $t(i, j)$, then $\bar{t} \in (j, i)$.

If one assumes, as we have done, that there are no channel selection rules except for order selection rules, then *the amplitude*



is nonzero (as a function) *if and only if* $j_1 = i_2, j_2 = i_3, \ldots, j_n = i_1$.

We now define a flavor classification (i, j) for interacting channels, too: an (interacting) ordered channel $\tau = (t_1, \ldots, t_m)$ is said to be in the (i, j) *sector* if and only if $i_1 = i, j_m = j$. This set of ordered (i, j) channels is indeed an (ordered) sector in the sense of Section 4, since every (i, j) channel was just seen to communicate with every other (i, j) channel, but with no other (i', j') channel.

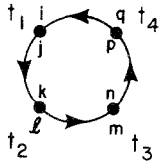
It is easy to see that, owing to the pole conjecture, particles of every class (i, j) actually exist.

5.6. Flavor-Labeled Process Graphs and Quark Diagrams; the OZI Rule

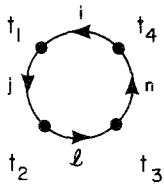
Graphically, a particle is represented as

If t is in (i, j), then as we saw, i labels the predecessor class and j the successor class of t; graphically we represent this as

Since a process graph like

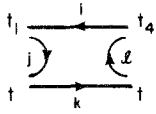


only corresponds to a nonzero amplitude if $j = k$, $l = m$, $n = p$, and $q = i$, we see that in any nonzero ordered amplitude we can consistently flavor-label the edges of the process graph:



With the alternative representation of particles

this process graph looks like



which we recognize as a conventional mesonic quark diagram. And we see that an ordered amplitude is only nonzero if it can be written in this way as a connected flavor-labeled quark diagram; this is the *OZI rule*, which is thus automatically obeyed by the OSM.

So we have demonstrated that sequentially ordered amplitudes can, in and of themselves, be considered as mesonic quark-diagram amplitudes, and the external particles be identified as nonexotic mesons. This identification will be confirmed when we assign internal quantum numbers

to the particles, and find them to obey the usual mesonic pattern. The edges of process graphs are found to correspond to the quark lines of quark diagrams.

So far we have not made any statement about the number of different flavors; it could be one (no order selection rules), several, or infinite. Experimentally, we need at least five flavors to understand the present data, but the number is growing, and there seems no a priori reason to exclude even the possibility of infinitely many flavors, "unfrozen" at ever higher energies. Using just the simple considerations that we have, there is no way to theoretically determine the number of flavors. To achieve this, a detailed dynamical bootstrap analysis would be necessary.

We would like to stress that although flavor labeling is useful, it is not an indispensable part of the formalism of ordered SMT; we were able to construct the whole theory in Section 3 without ever mentioning flavor. The flavor information is implicitly contained in the particles. In this respect the situation resembles that of the usual SMT, where IQNs are similarly useful but formally dispensable.

5.7. The Canonical Internal Quantum Numbers

In a nonzero ordered amplitude, every edge that starts out from a particle with a successor index i has to end up in a particle with the predecessor index i. Hence for every flavor i, the number of times it occurs as a successor index minus the number of times it occurs as a predecessor index must be zero for every nonzero ordered amplitude.

For any particle of set of particles we denote the number of times that a flavor i occurs as a successor index minus the number of times it occurs as a predecessor index (i.e., the number of i edges leaving the set minus the number of i edges entering it) the *canonical additive internal quantum number* $Q^{(i)}$, or simply the canonical IQN. Thus for every flavor index i there is a corresponding $Q^{(i)}$. As we just saw, the $Q^{(i)}$'s are *conserved* in ordered amplitudes.

By definition $Q^{(i)}$ has the same value for all particles in the same flavor class (j, k); it is a class property. Specifically, $Q^i(j, k) = \delta_{ik} - \delta_{ij}$. The same formula also holds true for ordered *channels* in the class (j, k). Thus for particles and ordered channels $Q^{(i)}$ can only take on the values $-1, 0, +1$. For nonadjacent sets of particles it can take on any integer value.

We may, if we wish, attribute a "charge" $Q^{(i)}(j) = \delta_{ij}$ to every edge ("the j edge carries the charge $Q^{(j)}$"), although nothing is gained by that terminology. Then $Q^{(k)}(i, j) = Q^{(k)}(i) - Q^{(k)}(j)$. These charges $Q^{(i)}(j)$ can

be changed by an arbitrary additive constant $Q_0^{(i)}$ without modifying the $Q^{(i)}$ assignments of particles.

Since every particle or set of particles has an equal number of edges entering it and leaving it, we have $\sum_i Q^{(i)}(kl) = 0$ for all k, l. [In Part Two, for generalized order, this will no longer be true: there $\sum_i Q^{(i)} \equiv 3B$, where B is the baryon number.] For example, we have one less nontrivial conservation law than flavors (e.g., for a one-flavor theory, there is no nontrivial conservation law). We may therefore eliminate one of the $Q^{(i)}$'s as superfluous. If we introduce the conventional names for the five flavors observed till now, namely, u, d, s, c, b, then the reduction from five to four additive IQNs may be carried out by eliminating $Q^{(u)}$ and $Q^{(d)}$ and replacing them by the quantity $Q^{(e)} \equiv -\frac{1}{3}Q^{(d)} + \frac{2}{3}Q^{(u)} - \frac{1}{3}Q^{(s)} + \frac{2}{3}Q^{(c)} + \frac{1}{3}Q^{(b)} \cdots$; the quantity $Q^{(el)}$ is called the electric charge; it is defined in this particular way for reasons that are outside the scope of a pure strong-interaction approach. Thus we remain with a set of conserved quantum numbers $Q^{(el)}, Q^{(s)}, Q^{(c)}, Q^{(b)}$, that coincides exactly with the usual electric charge, strangeness, charm and beauty, etc.

5.8. The Significance of the Canonical IQNs for the Physical SM

We saw in Section 4 that the simplest way to formulate the sector selection rules of the physical, unordered SM is to define a set of additive and/or multiplicative quantum numbers for each particle; the conservation of these quantum numbers then expresses the sector selection rules.

Similarly, for the order selection rules of the ordered SM, the canonical formalism was found to be one in which we defined flavor labels for the two edges attached to every particle; the order selection rules were then expressed by demanding that the two flavor labels thus attached to the two ends of an edge are equal. Since we assumed that order selection rules were the only channel selection rules of the ordered SM, the flavor labels are sufficient to express all channel selection rules.

We also found that the order selection rules implied the existence of a set of canonical additive conserved quantum numbers $Q^{(i)}$, one for each flavor. But the converse is not true: merely postulating $Q^{(i)}$ conservation does not imply the order selection rules, but merely a fraction of their statement. For one thing, additive conservation rules are oblivious to particle order: they distinguish only between different sets of external particles, but not between different orders of one and the same set.

So the best one could hope for would be that $Q^{(i)}$ conservation would allow us to determine for which sets of particles there is *any* order for which the ordered amplitude is nonzero.

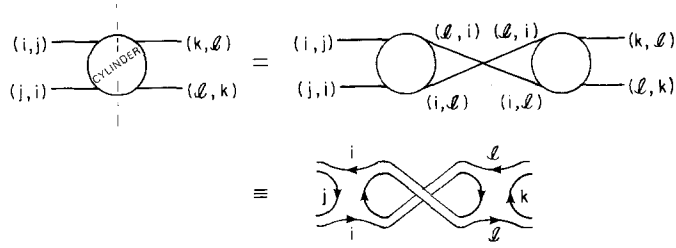
But the fact is that $Q^{(i)}$ conservation does not even accomplish that completely. For particle sets like $A \in (i, j)$, $B \in (j, i)$, $c \in (k, l)$, $D \in (l, k)$, or, e.g., $A \in (i, j)$, $B \in (j, k)$, $C \in (k, i)$, $D \in (l, l)$, $E \in (m, n)$, $F \in (n, p)$, $G \in (p, m)$ that can be drawn as several disconnected process graphs but not as one single one, we know that all ordered amplitudes are zero; and yet they are allowed by $Q^{(i)}$ conservation.

But if $Q^{(i)}$ conservation is so ineffective in describing the selection rules of ordered amplitudes, then why do we bother to introduce the $Q^{(i)}$'s at all? The answer is that as long as we remain in the domain of ordered amplitudes, there is no advantage to be gained by their introduction.

It is only when we begin to construct the physical SM from ordered amplitudes by means of the topological expansion that the real significance of the $Q^{(i)}$'s emerges.

The first step of this construction yields the so-called planar amplitudes (an unordered approximation to the physical amplitude obtained by adding all ordered amplitudes with a given set of external particles, as described in Section 6). Here the reason for the introduction of the $Q^{(i)}$'s does not become clear yet, since, by their definition these planar amplitudes are nonzero if and only if the particles can be arranged in any order such that the corresponding (connected) process graph can be consistently flavor-labeled (OZI rule); thus the less restrictive $Q^{(i)}$ conservation rules are redundant (since they are necessary but not sufficient).

But already the next higher term of the topological expansion, the cylinder amplitude, breaks the OZI rule, as the example



demonstrates, but the $Q^{(i)}$'s are still conserved in this term, and indeed in all terms of the topological expansion, because a quark line of a given flavor starting out on an external particle has to end up on an external particle, still with the same flavor. Thus, for the physical amplitudes, all that remains of the order selection rules are the $Q^{(i)}$ conservation laws. This, then, is the true significance of the $Q^{(i)}$'s: the sector selection rules to which their conservation corresponds are the vestige, at the physical level, of the order selection rules at the ordered level. And, on the other hand, owing to our

postulate that there be no selection rules other than order selection rules for the ordered S matrix, the $Q^{(i)}$'s are the *only* conserved internal quantum numbers.

5.9. And If There Were Additional Selection Rules?

Although we have found our postulate—that there be no selection rules besides order selection rules for the ordered SM—completely satisfactory, we were nevertheless interested in the consequences of relaxing that postulate. Since the question appears largely academic, however, we just present a summary of results.

The methods of Section 4 can be brought to bear on the problem in somewhat modified form. If $\tau = (t_1, t_2, \ldots, t_m)$, $\tau' = (t_1', t_2', \ldots, t_n')$ are two ordered channels, then $\tau \circ \tau'$ is defined to be the ordered channel $(t_1, t_2, \ldots, t_m, t_1', \ldots, t_n')$. This operation is now obviously not commutative. Since $\tau \circ \tau'$ need not be interacting even if τ and τ' are both interacting, we cannot in general define the sector group by $[\tau] \circ [\tau'] = [\tau \circ \tau']$, since if $\tau \circ \tau'$ is not interacting, then $[\tau \circ \tau']$ is not defined. But if we restrict ourselves to particles (and hence also ordered channels) from the flavor class (i, i), for any flavor i, then the group multiplication is well defined, and the sector group [of (i, i) channels] can be introduced as in Section 4. However, it need not be Abelian. One can show that the sector group of (i, i) channels is isomorphic to the sector group of (j, j) channels for any two flavors i, j. And furthermore, knowledge of that group structure implies a complete knowledge of the communication structure of the ordered SM.

If the (i, i) sector group is trivial, then we are back to our postulate: there are no selection rules except for order selection rules. If it is not trivial, but at least Abelian, then that implies the existence of further conserved internal quantum numbers (additive or multiplicative) besides the canonical ones. These additional additive IQNs have the property that they adopt every integer value from $-\infty$ to $+\infty$ for every class of particles (i, j) separately. To prove this, we require the pole conjecture. Any additive conserved IQN that is a function of the (i, j) labels of the particle alone, is a linear combination of canonical $Q^{(i)}$'s. If we make the assumption that all (i, i) particles are self-conjugate, then besides the canonical IQNs only multiplicative IQNs of order 2 are permitted.

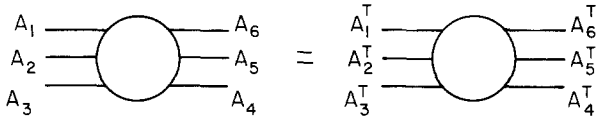
If the (i, i) sector group is non-Abelian, then we are faced with a new kind of selection rule that cannot be expressed as additive or multiplicative IQN conservation (and also not as an order selection rule). For example, the ordered channels (t, t') and (t', t) might both be (i, i) channels and yet not communicate.

5.10. Internal Symmetries and Multiplet Structure; I-Spin Degeneracy

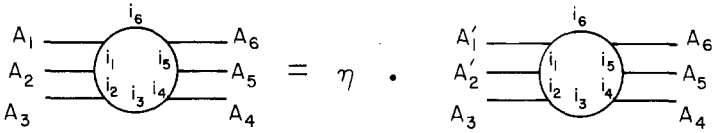
So far, we have not regarded internal symmetries, and there seemed to be no intrinsic reason to introduce them; possibly, a more in-depth study of ordered SMT will turn up such reasons. But meanwhile we know empirically that the flavors are related to one another by a "broken" $SU(n)$ symmetry group, the $SU(2)$ subgroup of which, relating u and d, is exact up to within the order of electromagnetic corrections. Therefore, we introduce these symmetries by postulate in a manner formally identical to that of quark models. The fact that these symmetries have to be introduced by postulate rather than follow from the inner logic of the theory is common to the quark model and the ordered SM approach; but in the former this ad hoc approach seems unavoidable, whereas a bootstrap treatment of the ordered SM might conceivably yield a derivation of internal symmetries together with the number of flavors based on consistency requirements.

The postulate of exact $SU(2)$, or, respectively, broken $SU(n)$, yields exactly the same predictions about the multiplet structure of the spectrum as it does in the usual framework; we need not elaborate on it here.

$SU(n)$ symmetry as postulated above is "global" in the sense that it demands the equality of two ordered amplitudes related by an $SU(n)$ transformation T that acts on *every* particle of the amplitude: e.g.,



where A_i^T is the state obtained from A_i by application of T. But we now impose a more stringent symmetry condition which is "local" in the sense that it demands the equality of two ordered amplitudes related by a transformation that affects the flavor of one single edge. For example, we have



where η is a phase-factor, and A_1' and A_2' are the well-defined particles that A_1 and A_2 are transformed into when the flavor i_1 is transformed into the flavor i_1'. Note that A_3, A_4, A_5, A_6 are not affected by this transformation, which can thus not be considered as a usual symmetry transformation acting on initial and final states as a whole. We see that this "local" symmetry implies the usual "global" $SU(n)$ symmetry but not vice versa.

We now make a few remarks about I-spin degeneracy. $SU(2)$ symmetry

results in degenerate I-spin multiplets of particles, corresponding to irreducible representations of $SU(2)$, and characterized by a value of I. $SU(2)$ does not demand the degeneracy of different multiplets. For the physical SM general principles indeed prohibit exact I-spin degeneracy; since it would entail that two particles from the same sector (i.e., with all additive IQNs equal), but differing in the value of I, would be degenerate. But in Section 4 we observed that two particles from the same sector cannot be degenerate.

For the ordered SM, the situation is different. Here, the "local" symmetry demands that, e.g., for every particle t_{pp} from the flavor class (p, p) there exist a set of degenerate particles t_{pn} in (p, n), t_{np} in (n, p), and t_{nn} in (n, n). Now, in the ordered case, t_{pp} and t_{nn} are in different sectors (OZI rule) and so there is no obstacle to their being degenerate.

We need not regard t_{pn} and t_{np} further since, having distinct values of canonical IQNs, in particular electric charge, they will continue to belong to different sectors even when we pass by the topological expansion to the corresponding physical particles. But t_{pp} and t_{nn}, having the same set of canonical IQNs, begin to communicate already at the cylinder level of the topological expansion, and so we know that the degeneracy must be broken there. As always in the case of a degeneracy, we can choose the basis in an arbitrary way. At the ordered level, the basis $\{t_{pp}, t_{nn}\}$ is appropriate for most purposes, in particular to express the order selection rules. But if we are interested in the basis that corresponds to the physical particles, then the basis

$$\left\{\frac{t_{pp} + t_{nn}}{2^{1/2}}, \frac{t_{pp} - t_{nn}}{2^{1/2}}\right\}$$

is appropriate, because it is these linear combinations that are unmixed by $SU(2)$ and therefore correspond to the physical particles, with I spin 1 and 0, respectively. In terms of this basis the degeneracy is expressed as I-spin degeneracy.

6. THE PLANAR S-MATRIX APPROXIMATION

6.1. The Elimination of Process Order: The PSA and the Topological Expansion

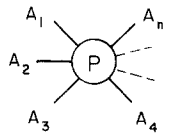
We have already mentioned that the ordered amplitudes introduced in this work do not represent directly measurable quantities, since they refer, besides to the individual particle parameters t, p, μ, which are measurable, also to the order between them, which is not.

In order to pass from ordered to physical amplitudes, we have to somehow eliminate, "average out," the order.

This corresponds to passing from the statistical treatment of a level of description characterized by order (namely, ordered SMT) to the statistical treatment of the usual, "physical" level of description (namely, the usual SMT), which does not contain the concept of order.

How to make this transition is not yet known in all generality; it constitutes a novel problem. The topological expansion of DTU, which, as mentioned in Section 1, grew out of dual perturbation theory, is an attempt to solve just this problem. But it suffers from basic theoretical defects, and cannot be regarded as the final solution. Nevertheless, it appears to contain correct elements, and has achieved remarkable phenomenological successes. We may therefore hope to draw our inspiration at least in part from the topological expansion when we attack the problem systematically in a future work.

In this section, we merely regard the lowest-order term of the topological expansion, which we call the PSA (planar S-matrix approximation). It consists of approximating a physical amplitude $T(A_1, A_2, \ldots, A_n)$



by the sum of all $(n-1)$ different ordered amplitudes with the "same" set of ordered particles $\{A_1, A_2, \ldots, A_n\}$ in all different cyclic orderings. Thus the PSA is an unordered amplitude defined in terms of the particles of the ordered SM spectrum.

Why did we put "same" in quotation marks? The reason is that we know from DTU that higher terms of the topological expansion shift the Regge trajectories and couplings. If the total shift resulting from the topological expansion is not too large, and, above all, if no poles disappear and no new poles are created at any (finite) step of the topological expansion, so that we can keep track of the individual poles in the iterative progression from the ordered to the physical SM, then there is a clear correspondence between ordered and physical particles. It is this presumed correspondence we refer to when we talk about the "same" particles; without it we could not even compare the PSA and the physical SM.

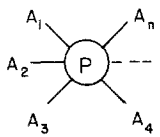
The prescription for obtaining the PSA by adding all the ordered amplitudes with the given set of external particles is the most straightforward way of "averaging out" the particle order. It corresponds to the quantum mechanical rule according to which the amplitude of a process is

equal to the sum of amplitudes corresponding to the various paths the process can take. In this case process order plays the role of path.

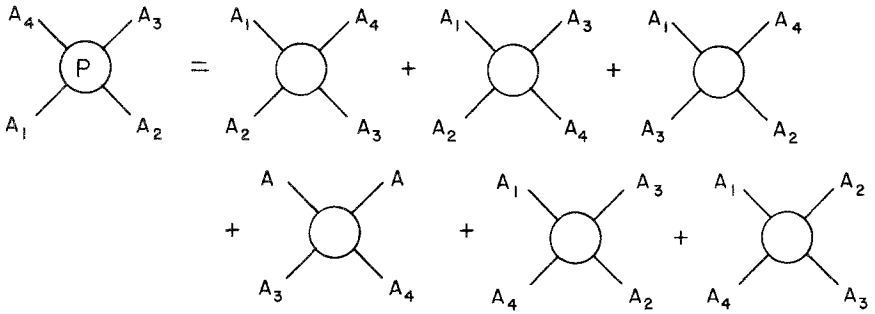
As plausible as this prescription for eliminating process order may seem, it is only approximately correct in the numerical sense, and the path analogy is probably conceptually inappropriate. The main defect is the demonstrable lack of unitarity of the PSA. On the other hand, the unitarity-violating terms can be shown to be relatively small, and the PSA obeys many of the general SM properties such as Poincaré invariance, macrocausality, and normal analytic structure, the independence property and cluster decomposition, pole factorization, asymptotic bounds, Hermitian analyticity, Bose statistics, etc. In addition, it has the specific duality and quark properties that it inherits from the ordered amplitudes by virtue of its linear definition in terms of them: weak exchange degeneracy, I-spin degeneracy, absence of Regge cuts and long-range correlations in rapidity, OZI rule, all of which are found to be approximately valid experimentally; and of course the quark-model mesonic spectrum derived in Section 5. These features let us expect that the PSA represents a reasonably good approximation to the physical SM, and hence a good starting point for the topological expansion.

6.2. Some Properties of the PSA

Definition and Notation. We denote the planar amplitudes by

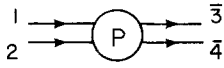


So according to our definition



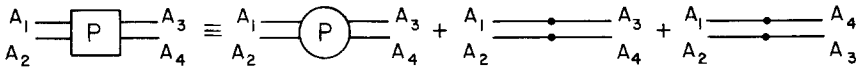
It is obvious from this definition that planar amplitudes are unordered, since they are symmetrized with respect to particle order.

When referring to a particular channel, say $1, 2 \to \bar{3}, \bar{4}$, we write the above amplitude as

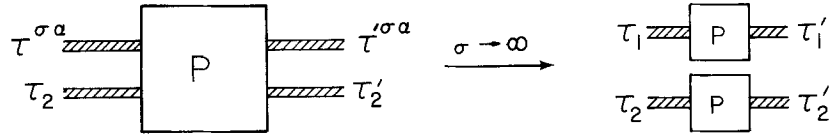


with the usual conventions about positive energy, etc.

Cluster Decomposition and the Planar SM. The planar SM is defined in terms of planar amplitudes by the usual (unordered) cluster decomposition equations, e.g.,



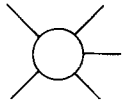
With this definition, the independence property for planar amplitudes, namely, that they tend to zero when a subset of particles is translated to infinity, is converted into the independence property of the planar SM:



Lorentz Invariance. That planar amplitudes are Poincaré invariant follows trivially from the corresponding property of the ordered amplitudes.

Analyticity. Planar amplitudes are analytic everywhere in the physical region except on positive-α Landau surfaces; this is because they have all the singularities that any of the ordered amplitudes (whose sum they are) have. Its discontinuities in "exotic" channels will be zero.

Crossing. Crossing for planar amplitudes follows from their very definition. But the basis for that definition, the existence of the (ordered) functions



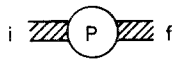
depends on the ordered crossing property of ordered amplitudes.

Spectrum. By its definition, the PSA has the same particle spectrum as the ordered SM, namely nonexotic mesons. We saw in Section 5 that the internal quantum numbers assigned to these mesons are just those predicted by the quark model.

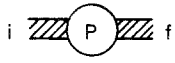
Pole Factorization. This is usually a direct consequence of unitarity; indeed it could be called one-particle unitarity. But although unitarity does not hold for the PSA, pole factorization does. This result is of crucial importance since it allows a consistent SM interpretation of the PSA. And

to the extent that planar amplitudes are resonance dominated, pole factorization alone, without any other considerations, implies approximate unitarity.

To show pole factorization, we regard a planar amplitude

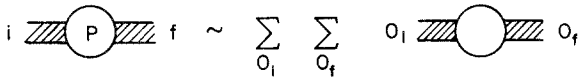


at one of its poles X. Now

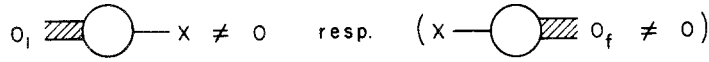


is the sum of all ordered amplitudes with the set of external particles occurring in the channels i and f. Amongst those will be ordered amplitudes where all initial particles are adjacent (ordered transition amplitudes) and others, where initial and final particles are interspersed. The latter do not contribute to the pole in question, and can therefore be ignored in this context.

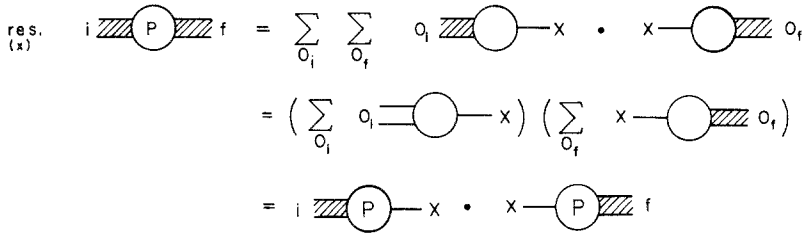
Hence near the pole



where O_i stands for any ordered initial channel consisting of the same particles as i (but in different order); and O_f analogously. Only those channels O (O_f) for which



contribute to the above sum, so let us restrict the sum to these. Owing to the pole factorization or ordered amplitudes, the pole residue is then equal to



which is nothing but pole factorization.

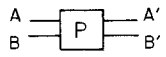
Other SM Properties. Hermitian analyticity and the Froissart bound of the PSA follow immediately from their validity for ordered amplitudes. Bose statistics arises from the full symmetrization of the particle order due to summation over all such orders; this leads to a permutation symmetry of the PSA analogous to that of the physical SM; and as we saw in Section 2, it is this permutation symmetry that gives rise to Bose statistics.

Duality Properties. The I-spin degeneracy and absence of Regge cuts and fixed j poles of the ordered amplitudes trivially entails that of planar amplitudes. And since the position of a Regge trajectory is certainly not changed by superposition, the positive and negative signature poles still coincide: $\alpha_j^{(+)}(t) = \alpha_j^{(-)}(t)$. However, the residues $\beta_j^{(+)}(t)$ and $\beta_j^{(-)}(t)$ are now in general distinct. This is called the *weak exchange degeneracy* of the PSA. What could conceivably happen (but does not appear to happen for mesons) is that one signatured trajectory could be wiped out because the corresponding residues all vanish. This may be what is happening for some baryon trajectories.

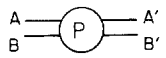
OZI Rule. Obviously, a planar amplitude is nonzero only if at least one ordered amplitude in the sum is nonzero, i.e., may be drawn as a legal quark diagram. This is the usual OZI rule; it is exact for the PSA.

C Symmetry. We noted in Section 2.4 that even if all external particles of an ordered amplitude are self-conjugate, this does not lead to a selection rule because the product of all charged conjugations need not be $+1$. For the PSA, this is no longer true: if the product of charge conjugations is $=1$, then every ordered amplitude in the sum cancels against the amplitude with opposite orientation, and so the PSA is zero. Thus the PSA obeys the same C-selection rule as physical amplitudes.

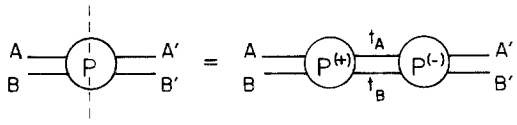
Lack of Unitarity. We demonstrate with the simplest example that the PSA is not unitary. If



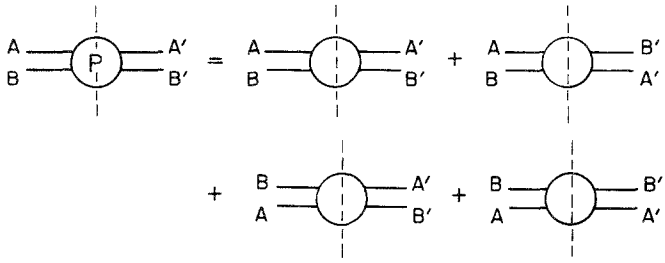
were unitary, then



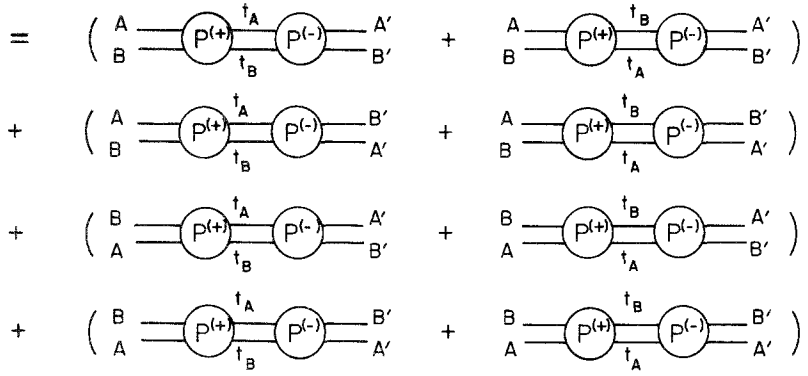
would, below the inelastic threshold, obey the discontinuity equation



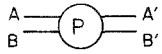
On the other hand, we *know* from the discontinuity equations for ordered amplitudes that



(the other two ordered amplitudes do not contribute to the discontinuity)

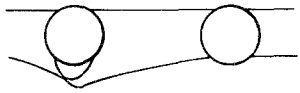


When we compare these two expressions for



we see that the eight terms of the latter expression cancel against eight of the $6 \times 6 = 36$ expressions of the former, leaving us with the demand that the sum over the 28 other unitarity products be zero if the PSA were indeed to be unitary. These 28 "nonplanar" products are of the type

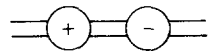
(products of ordered transition amplitudes with crossing intermediate lines), and



or

etc. (at least one of the factors is not an ordered transition amplitude).

Although the sum of these terms is in general nonzero, and therefore the PSA not unitary, it can be shown that these nonplanar terms are suppressed with respect to the planar terms



and therefore the unitarity violation of the PSA is not as serious as it might appear. This smallness of unitarity violation was also implied by pole factorization to the extent that the PSA is resonance dominated.

Experimental Verification of Ordered SMT. We have already seen that a variety of duality and quark-model properties are predicted, and experimentally confirmed.

Since ordered SMT provides a theoretical foundation for DTU, the phenomenological success of that approach also constitutes a powerful confirmation of the ideas presented here. Readers interested in that aspect are referred to Chew and Rosenzweig (1978) and the bibliography to be found there.

We now turn our attention to a generalization of the concept of order that will allow us to broaden the scope of the theory to include all hadrons.

6.3. Outline of Part Two

In Section 7, we recognize the need for a generalization of the ordered SM scheme to include baryons and exotics, and review some of the difficulties involved. We see why sequential order cannot accommodate baryons, and develop the framework of a more general particle order to achieve that.

In Section 8, we present those concepts and facts from graph theory that will be required to develop the theory.

In Section 9, consistency requirements between general principles of SMT and particle order are exploited to determine the specific form of the theory.

In Section 10, we review various properties of the general ordered SM, and the axiomatic structure of the theory. Also, a planar approximation to the physical SM is proposed.

Section 11 brings the conclusion, and an outlook on possible further developments.

In an epilogue ("Philosophical Postscript") we speculate about a possible interpretation of particle order and some of the broader consequences that might follow from this concept.