LECTURES IN THEORETICAL PHYSICS

Volume XIV A

Topics In Strong Interactions

Edited by

Asim O. Barut and Wesley E. Brittin
Department of Physics and Astrophysics
University of Colorado

Colorado Associated University Press
Boulder, Colorado

Colorado Associated University Press
1424 Fifteenth Street
Boulder, Colorado 80302

Library of Congress Catalog Number 59-13034
ISBN 0-87081-043-X

THE LECTURERS

Henry Abarbanel
Princeton University, Princeton, New Jersey

K. Bardakci
University of California, Berkeley, California

A. O. Barut
University of Colorado, Boulder, Colorado

Alan Chodos
University of Pennsylvania, Philadelphia, Pennsylvania

Fred Cooper
Yeshiva University, New York, New York

R. E. Cutkosky
Carnegie-Mellon University, Pittsburgh, Pennsylvania

G. Domokos
Johns Hopkins University, Baltimore, Maryland

Stanley Fenster
Argonne National Laboratory, Argonne, Illinois

M. M. Islam
University of Connecticut, Storrs, Connecticut

H. Jehle
George Washington University, Washington, D. C.

James McEnnan
University of Florida, Gainesville, Florida

H. Pfister
University of Tubingen, Tubingen, Germany

P. Ramond
National Accelerator Laboratory, Batavia, Illinois

E. Squires
University of Durham, Durham, England

Robert Sugar
University of California, Santa Barbara, California

FOREWORD

This Volume XIVA contains Lectures presented during the first part of the Boulder Summer Institute for Theoretical Physics held in June, 1971.

This part of the Institute was devoted to several active areas of the past years in Particle and High Energy Physics. Accordingly, the Volume is divided into four parts: eikonal approximations, dual resonance models, as well as other related models, and to analytic approximation methods.

We have tried to present in each area a comprehensive picture including introductory reviews as well as recent material, and different but related viewpoints.

The Institute was sponsored by the National Science Foundation.

I wish to thank the lecturers and the participants for their effort for a lively Institute and to Mrs. Charlotte Walker for her invaluable contribution to the organization of the Institute. I would like to extend my appreciation to Mrs. Walker, assisted by Mrs. Ruth Henard, for the typing of the manuscript.

Boulder, February 1972

A. O. Barut

TABLE OF CONTENTS

PART I: EIKONAL APPROXIMATIONS IN STRONG INTERACTIONS

EIKONAL APPROXIMATIONS IN COLLISION THEORY
Henry D. I. Abarbanel 3

THE RELATIVISTIC EIKONAL MODEL
Robert L. Sugar 47

EIKONAL DESCRIPTION OF HIGH-ENERGY PARTICLE SCATTERING
M. M. Islam 101

PART II: DUAL MODELS IN STRONG INTERACTIONS

AN INTRODUCTION TO DUAL RESONANCE MODELS
K. Bardakci 157

THE PRESENT STATUS OF DUAL MODELS
E. Squires 181

GROUP THEORETICAL PROPERTIES OF DUAL RESONANCE MODELS
P. Ramond 241

SELF-CONSISTENT ELECTROMAGNETIC "DUAL" AMPLITUDES AND THE ELECTRON-POSITRON SYSTEM
James McEnnan 275

DUAL LOOPS AND MULTIPLY PERIODIC FUNCTIONS
Stanley Fenster 297

PART III: NEW THEORETICAL MODELS IN STRONG INTERACTIONS

A POSSIBLE PHYSICAL BASIS FOR SOME MATHEMATICAL MODELS OF STRONG INTERACTIONS: ATOMS WITH MAGNETIC CHARGES
A. O. Barut 339

PARTICLES AS NORMAL MODES OF AN UNDERLYING GAUGE FIELD THEORY
Fred Cooper and Alan Chodos 349

INTRODUCTION TO THE CHARACTERISTIC INITIAL VALUE PROBLEM IN QUANTUM FIELD THEORY
G. Domokos 375

FLUX QUANTIZATION AND PARTICLE PHYSICS
H. Jehle 399

PART IV: ANALYTIC APPROXIMATION AND EXTRAPOLATION

PRINCIPLES OF ANALYTIC APPROXIMATION THEORY AS APPLIED TO ANALYSIS OF EXPERIMENTAL DATA
R. E. Cutkosky 423

ANALYTIC EXTRAPOLATION OF SCATTERING AMPLITUDES AND FORM FACTORS
H. Pfister 455

Part I

EIKONAL APPROXIMATIONS IN STRONG INTERACTIONS

EIKONAL APPROXIMATIONS IN COLLISION THEORY*

Henry D. I. Abarbanel**
Joseph Henry Laboratories
Princeton University
Princeton, N.J. 08540

Introduction:

In this series of lectures I would like to discuss an approximation scheme which is useful in very short wavelength or, equivalently, very high energy scattering processes or propagation of waves through a medium. We will envisage the potential causing the scattering or the index of refraction characterizing the medium to be very "smooth" so that the direction of propagation of a high frequency incident beam is essentially unaltered. Smoothness of the scattering force means, in practice, that the scale on which changes in its shape occur is much larger than a typical wavelength from the incident beam. This has the consequence that it is a good first guess to associate a modulated plane wave with the scattered or deflected beam. The modulation due to the potential can properly describe absorptive and dispersive properties of the wave propagation. One can further describe a systematic procedure for improving upon his original wave amplitude.

The usefulness of the method we will be considering is already apparent in non-relativistic potential scattering where the basic approximation provides one with a vast improvement over the Born approximation and, yet, for appropriate potentials reduces to the Born term for very large energies. The main advantage gained is a scattering amplitude which satisfies unitarity, which the Born approximation manifestly fails to do, and which is a good representation of the true amplitude over a wide range of energies.

*Work supported by the U.S. Atomic Energy Commission under Contract AT(30-1)-4159.
**Alfred P. Sloan Foundation Research Fellow.

The technique also has its place in quantum field theories where it enables one to sum up the high energy behavior of a very interesting set of Feynman graphs in a compact and useful manner. Furthermore it justifies the use of a Bessel-transform or eikonal representation of the S-matrix which is familiar from non-relativistic problems.

II. Geometrical Optics

We will first discuss the basic ideas in the context of wave propagation in a medium and then proceed to explore the properties of such an approximation scheme within potential scattering. Finally the expression of the eikonal method in quantum electrodynamics will be presented.

So let us begin by considering a wave characterized by an amplitude $A(\vec{r},t)$ propagating through a medium of spatially varying index of refraction $n(\vec{r})$. $A(\vec{r},t)$ could well represent any component of the $\vec{E}$ or $\vec{B}$ fields of an electromatnetic disturbance and satisfies the wave equation

$$\left[\nabla^2 - \frac{n(\vec{r})^2}{c^2} \frac{\partial^2}{\partial t^2} \right] A(\vec{r},t) = 0. \qquad (1)$$

We want to address the physical situation in which the variation of $n(\vec{r})$ is very slow. More precisely if a length L characterizes the spatial distance over which $n(\vec{r})$ changes significantly (say by a factor of 2) or equally well sets the scale for the gradient of $n(\vec{r})$, then we are interested in the propagation of wave whose wavelength λ is such that

$$\lambda << L \qquad (2)$$

or whose wavenumber satisfies

$$kL >> 1. \qquad (3)$$

If we are in this situation then it is physically pretty clear that the major effect of the medium will be primarily to alter the phase of the wave by k times the optical path, but that the direction of propagation will be unaltered. With this in mind let us seek a solution of the wave equation in the form

$$A(\vec{r},t) = \exp i(\vec{k}\cdot\vec{r}-\omega t)\ B(\vec{r}), \qquad (4)$$

where the free space relation $|\vec{k}|c = \omega$ is taken to be true. All of the effects of the medium or of the deviation of $n(\vec{r})$ from 1 are contained in the modulation factor $B(\vec{r})$ which we will regard as slowly varying through space.

Since we are using as our uncorrected wave in (4) a plane wave appropriate to free space, or n = 1 everwhere, if we are to have $B(\vec{r})$ slowly varying in space then we cannot require it to make up the difference between n = 1 and some value significantly different from that. We will require, therefore, that the deviation of $n(\vec{r})$ from 1 remain small throughout the medium. Should the medium have an index of refraction that varies slowly about some other value not equal to one, then the unmodulated plane wave should have its $\omega, |\vec{k}|$ relation chosen to reflect this, and our discussion is otherwise unaltered.

From (1) we determine an equation for $B(\vec{r})$

$$\nabla^2 B(\vec{r}) + 2i\vec{k}\cdot\nabla B(\vec{r}) + k^2[n(\vec{r})^2 - 1\,]B(\vec{r}) = 0. \qquad (5)$$

Since we will treat various spatial derivatives of $B(\vec{r})$ as small, one's inclination is to drop the $\nabla^2 B$ term with respect to the rest and evaluate the resulting $B(\vec{r})$. Indeed, that is what we shall now do in a systematic manner by rewriting the equation for B as an integral equation

$$B(\vec{r}) = B_0(\vec{r}) + \int d^3r' g_k(\vec{r},\vec{r}')\ (-\nabla_r^2, B(\vec{r}')), \qquad (6)$$

where the inhomogeneous term $B_0(\vec{r})$ satisfies

$$2ik\cdot\nabla B_0(\vec{r}) + k^2[n(\vec{r})^2 - 1]B_0(\vec{r}) = 0 \qquad (7)$$

and the Green function $g_k(\vec{r},\vec{r}')$ has the obvious property

$$\left\{2i\vec{k}\cdot\nabla_r + k^2[n(\vec{r})^2 - 1]\right\} g_k(\vec{r},\vec{r}') = \delta^3(\vec{r}-\vec{r}'). \qquad (8)$$

The approximation scheme is simply the Neumann series for the integral equation (6). Although we shall casually ignore such questions when convenient, it is perhaps useful to point out that the kernel of the integral equation contains the rather singular operator ∇^2 and that the true series expansion is unlikely to be convergent. Nevertheless a few terms in the series may provide useful, compact, and often excellent asymptotic approximation to $B(\vec{r})$.

There is no proper discussion of this matter to my knowledge so let's close our eyes and proceed.

The first approximation to $B(\vec{r})$ is a one dimensional equation expressing the variation of the modulated plane wave along its direction of propagation $(\vec{k}/|\vec{k}|) = \hat{k}$. There is no kinematical statement about the dependence on coordinates transverse to $\hat{k}$, and it is deeply in the nature of the approximation we are pursuing that these transverse degrees of freedom are decoupled from longitudinal ones. {This is the first hint that the method may not be entirely unrelated to real high energy physics.} In an heuristic manner one may simply attribute this to inertia; namely, fast moving things primarily keep moving forward. This persistance of the initial directions and consequent straight like paths is labeled the eikonal approximation.

We will solve for $B_o(\vec{r})$ by using these remarks to motivate the separation of space into a two vector $\vec{b}$ orthogonal to $\hat{k}$ and a piece along $\hat{k}$. Thus one writes

$$\vec{r} = \vec{b} + \lambda\hat{k}, \tag{9}$$

and the equation for B_o becomes

$$2ik \frac{\partial}{\partial\lambda} B_o(\vec{b} + \lambda\hat{k}) + k^2 [n(\vec{b} + \lambda\hat{k})^2 - 1]B_o(\vec{b} + \lambda\hat{k}) = 0. \tag{10}$$

This can be immediately integrated to yield

$$B_o(\vec{b} + \lambda\hat{k}) = \exp \frac{ik}{2} \int_o^\lambda d\tau[n(\vec{b} + \tau\hat{k})^2 - 1], \tag{11}$$

with the boundary condition that $B_o(\vec{b}) = 1$. The full wave amplitude becomes

$$A(\vec{r},t) = \exp i\left\{\vec{k}\cdot\vec{r} + \frac{k}{2} \int_o^\lambda d\tau[n(\vec{b} + \tau\hat{k})^2 - 1] - \omega t\right\}, \tag{12}$$

which for real $n(\vec{r})$ is the alteration of an ordinary plane wave by a phase reflecting the properties of the medium. For complex n we can get some insight into the properties of (12) by writing

$$n \approx 1 + i \text{ Im } n, \tag{13}$$

so that A is

$$A(\vec{r},t) \approx \exp i\{\vec{k}\cdot\vec{r} - \frac{k}{2}\int_o^\lambda d\tau\ (\text{Im } n\ (\vec{b}+\tau\hat{k}))^2 - \omega t\}x$$

$$\exp - k\int_o^\lambda d\tau\ \text{Im } n(\vec{b}+\tau\hat{k}). \qquad (14)$$

The last term represents the exponential damping of the incident wave that one ordinarily associates with an absorptive medium.

When the index of refraction varies slowly in space, as we have assumed all along, then the damping term in $A(\vec{r},t)$ is approximately

$$\exp - k\int_o^\lambda d\tau\ \text{Im } n(\vec{b}+\tau\hat{k}) \approx \exp - k\lambda\ \text{Im } n(o), \qquad (15)$$

whose form is that of a standard

$$\exp - \tfrac{1}{2}[\text{Path length/mean free path}]. \qquad (16)$$

Since the mean free path is inversely proporional to the density of scattering centers, N, times the cross section for interaction, σ, we find

$$\sigma = \frac{2k\ \text{Im } n(o)}{N}. \qquad (17)$$

To evaluate the next correction to the modulation function $B(\vec{r})$ it is necessary to exhibit the Green function $g_k(\vec{r},\vec{r}')$ corresponding to the one dimensional propogation of waves. It is convenient to seek this function in the "mixed representation" where we Fourier transform away the $\vec{r}'$ dependence. So let us define

$$\tilde{g}_k(\vec{r},\vec{q}) = \int d^3r'\ e^{i\vec{q}\cdot\vec{r}'}\ g_k(\vec{r},\vec{r}'), \qquad (18)$$

which satisfies the differential equation

$$\left[2i\vec{k}\cdot\nabla_r + k^2[n(\vec{r})^2 - 1]\right]\tilde{g}_k(\vec{r},\vec{q}) = e^{i\vec{q}\cdot\vec{r}}. \qquad (19)$$

We will seek a solution to this in the parametric form

$$\tilde{g}_k(\vec{r},\vec{q}) = e^{i\vec{q}\cdot\vec{r}}\ i\int_o^\infty d\sigma\ e^{2i\sigma\vec{k}\cdot\vec{q}} + F_k(\sigma,\vec{r}). \qquad (20)$$

The equation (19) then requires of us that

$$i\int_0^\infty d\sigma\ [k^2(n(\vec{r})^2 - 1) + 2i\ \vec{k}\cdot\nabla_r F_k(\sigma,\vec{r}) - 2\vec{k}\cdot\vec{q}]e^{2i\sigma\vec{k}\cdot\vec{q}+F_k(\sigma,\vec{r})} = 1. \tag{21}$$

Now integrate by parts on the $2\vec{k}\cdot\vec{q}$ term

$$-i\int_0^\infty d\sigma\ e^{2i\sigma\vec{k}\cdot\vec{q}\ +\ F_k(\sigma,\vec{r})}_{d\sigma} = -\int_0^\infty d\sigma\Big(\frac{d}{d\sigma}\ e^{2i\sigma\vec{k}\cdot\vec{q}}\Big)e^{F_k(\sigma,\vec{r})}$$

$$= e^{F_k(o,\vec{r})} + \int_0^\infty d\sigma\ e^{2i\sigma\vec{k}\cdot\vec{q}\ +\ F_k(\sigma,\vec{r})}\ \frac{\partial F_k(\sigma,\vec{r})}{\partial\sigma}\ , \tag{22}$$

dropping the "surface term" at $\sigma = +\infty$. (One may append a $+i\epsilon$ to the $\vec{k}\cdot\vec{q}$, if you wish.) Now let us choose $F_k(o,\vec{r})=0$, then there results an equation for $F_k(\sigma,\vec{r})$

$$i\ k^2(n(\vec{r})^2 - 1) - 2\vec{k}\cdot\nabla_r F_k(\sigma,\vec{r}) + \frac{\partial F_k(\sigma,\vec{r})}{\partial\sigma} = 0. \tag{23}$$

This has the solution which vanishes at $\sigma = 0$:

$$F_k(\sigma,\vec{r}) = -i\int_0^\sigma d\tau\ k^2[n(\vec{r} + 2\tau\vec{k})^2 - 1]. \tag{24}$$

The straight line or eikonal Green function is now

$$\tilde{g}_k(\vec{r},\vec{q}) = e^{i\vec{q}\cdot\vec{r}}\ i\int_0^\infty d\sigma\ \exp\Big\{2i\sigma\ \vec{k}\cdot\vec{q} - i\int_0^\sigma k^2[n(\vec{r}+2\tau\vec{k})^2-1]d\tau\Big\} \tag{25}$$

This may be understood if we regard σ as a kind of path length parameter so that g_k is the resultant wave gotten by integrating over all paths of the "free wave" exp $2i\sigma\vec{k}\cdot\vec{q}$ modified by the effects of the medium contained in the integral over $n^2 - 1$. The form (25) may be cast into the form of an integral over our first modulation function, $B_o(\vec{r})$, if you desire.

The next approximation to $B(\vec{r})$ is now given by combining (25) with the integral equation (6)

$$B(\vec{r}) = B_o(\vec{r}) + \frac{\int d^3q}{(2\pi)^3} \int d^3r' \; \tilde{g}_k(\vec{r},\vec{q}) e^{-i\vec{q}\cdot\vec{r}'} \left(-\nabla_{r'}^2 B_o(r')\right) \tag{26}$$

$$= B_o(\vec{r}) + i\int_o^\infty d\sigma \; \exp\left\{-i\int_o^\sigma k^2 [n(\vec{r} + 2\sigma\vec{k})^2 - 1]\right\}(-\nabla^2 B_o(\vec{r}+2\sigma\vec{k})). \tag{27}$$

This is not an unexpected result being a one dimensional path integral along the direction $\hat{k}$ of unperturbed functions $B_o(\vec{r})$.

Before proceeding onto potential scattering let's briefly note another variation on the theme we have been pursuing. The manner in which we found $g_k(\vec{r},\vec{q})$ is suggestive of seeking the entire modulating function $B(\vec{r})$ as an exponential integral. So we write

$$B(\vec{r}) = e^{\tilde{F}(\vec{r})} \tag{28}$$

Using the equation for $B(\vec{r})$ we arrive directly at a differential equation for $\tilde{F}(\vec{r})$

$$\nabla^2\tilde{F}(\vec{r}) + \nabla\tilde{F}(\vec{r})\cdot\nabla\tilde{F}(\vec{r}) + 2i\vec{k}\cdot\nabla\tilde{F}(\vec{r}) + k^2(n(\vec{r})^2 - 1) = 0. \tag{29}$$

We will again treat the derivatives as small and formally write

$$2i\vec{k}\cdot\nabla\,\tilde{F}(\vec{r},\eta) + k^2(n^2(\vec{r}) - 1)$$

$$= -\eta\left[\nabla^2\tilde{F}(\vec{r},\eta) + \nabla\tilde{F}(\vec{r},\eta)\cdot\nabla\tilde{F}(\vec{r},\eta)\right], \tag{30}$$

in which η is a derivative labeling parameter which we will set equal to one at the end of time. Then we seek a power series solution for $\tilde{F}(\vec{r},\eta)$

$$\tilde{F}(\vec{r},\eta) = \sum_{m=o}^{\infty} \eta^m \tilde{F}_m(\vec{r}). \tag{31}$$

The equation for $\tilde{F}_o(\vec{r})$ is familiar and has the solution

$$\tilde{F}_o(\vec{b} + \lambda\hat{k}) = \frac{ik}{2} \int_o^\lambda d\tau \left[n(\vec{b} + \tau\hat{k})^2 - 1\right] \tag{32}$$

which reproduces the result (11) for $B_o(\vec{r})$. The advantage of the present method is that when we seek the next correction which satisfies

$$2i\vec{k}\cdot\nabla\tilde{F}_1(\vec{r}) = -\left[\nabla^2\tilde{F}_o(\vec{r}) + \nabla\tilde{F}_o(\vec{r})\cdot\nabla\tilde{F}_o(\vec{r})\right], \tag{33}$$

the solution to which is elementary, we arrive at a corrected $B(\vec{r}) = \exp(\tilde{F}_o(\vec{r}) + \eta\tilde{F}_1(\vec{r}))$ which is in the form of a product rather than a sum. This small improvement is useful for numerical work with modulated waves. Further, every order in η leads to the same equation (33) which makes higher order iterating straightforward.

III. Potential Scattering

We now turn our attention to the use of eikonal or straight line approximations in the scattering of point particles from fixed potentials in non-relativistic quantum theory. Some years ago at this very institute R. Glauber[2] delivered a series of quite lucid lectures on this subject and we shall not overlap with his presentation except in the results. He emphasizes the modulated planewave approach of the previous section. We shall have our eyes slightly more focused on the quantum field theory to come in the next section.

The problem is to find an expression for the scattering amplitude $f(\vec{k}_f,\vec{k}_i)$ for a particle of mass m to go from initial momentum $\vec{k}_i$ to final momentum $\vec{k}_f$ in the presence of a potential $V(\vec{r})$. The differential cross section for this process is simply

$$\frac{d\sigma}{d\Omega_{\hat{k}_f}} = |f(\vec{k}_f,\vec{k}_i)|^2. \tag{33}$$

We will find it convenient to discuss instead of f the T matrix

$$T(\vec{k}_f,\vec{k}_i) = -\frac{2\pi\hbar^2}{m} f(\vec{k}_f,\vec{k}_i), \tag{34}$$

which satisfies the Schrodinger integral equation

$$T(\vec{k}_f,\vec{k}_i) = \tilde{V}(\vec{k}_f - \vec{k}_i) + \int \frac{d^3q}{(2\pi)^3} \tilde{V}(\vec{k}_f - \vec{q}) \frac{1}{\frac{\hbar^2k^2-\hbar^2q^2}{2m} + i\varepsilon} \times$$

$$\times \; T(\vec{q},\vec{k}_i) \tag{35}$$

where $\tilde{V}(\vec{\Delta})$ is the Fourier transform of the potential at momentum transfer $\vec{\Delta} = \vec{k}_f - \vec{k}_i$ and $\varepsilon > 0$ guarantees outgoing waves. We will consider (35) as the matrix element of operators T,V, and G_o and write

$$T = V + V\, G_o\, T = V + T\, G_o\, V, \tag{36}$$

where $G_o^{-1}(E) = E - \vec{p}^2/2m + i\varepsilon$ taking $E = \hbar^2k^2/2m$ which is the energy of the beam and $\vec{p}$ the familiar momentum operator.

The formal solution of (36) is easily constructed

$$T = V + VGV = VGG_o^{-1} = G_o^{-1}GV, \tag{37}$$

where the full Green function

$$G^{-1} = E - \vec{p}^2/2m - V + i\varepsilon \tag{38}$$

appears.

We want to examine an approximation scheme to T which is based on high energy or short wave length of the initial beam. Let us imagine that the potential has a "range" characterized by a length a, and that it is for all purposes zero outside that distance. It will be useful to keep the old favorite

$$V(\vec{r}) = \frac{V_o a}{|\vec{r}|} \exp - (|\vec{r}|/a), \tag{39}$$

in mind. We want to consider the limit of geometrical optics wherein the wave length associated with the incident beam is much smaller than a; that is

$$ka >> 1. \tag{40}$$

Further we wish to assume that the potential is both "smooth" enough and small enough so that the wave is not strongly distorted as it passes through. Smoothness can

be stated by asking that the fourier components contained in the potential be small for large values in Fourier space. To make this quantitative consider $|V(\vec{\Delta})|^2$ to be a measure of the probability of finding momentum $|\vec{\Delta}|$ in this distribution be small with respect to k or

$$\langle|\vec{\Delta}|\rangle = \int_0^\infty d\Delta\ \Delta\ |V(\vec{\Delta})|^2 \Big/ \int_0^\infty d\Delta\ |V(\Delta)|^2 \ll k. \qquad (41)$$

Physically this means that the momentum which can be transferred to the beam by the potential is small relative to the beam momentum, and, thus, most scattering will be in a cone about the forward direction. It is clear for dimensional reasons that (41) is equivalent to (40) for most potentials.

We can guarantee that the potential is "small enough" by asking that its "depth" be small compared to E, or for the Yukawa potential in (39) simply

$$|V_o| \ll E. \qquad (42)$$

We now can motivate the idea that during the scattering the expectation value of $\vec{p}$ will never appreciably differ from the incoming momentum $\vec{k}_i$ and that an expansion of G^{-1} about $\vec{p} = \vec{k}_i$ might be warranted. To that end we write

$$G^{-1} = E - \frac{\vec{p}^2}{2m} - V + i\varepsilon + \frac{(\vec{p} - \vec{k}_i)^2}{2m} - \frac{(p - k_i)^2}{2m}, \qquad (43)$$

$$= \left(\frac{k^2 + k_i^2 - 2\vec{p}\cdot\vec{k}_i}{2m} - V + i\varepsilon\right) - \frac{(p - k_i)^2}{2m}. \qquad (44)$$

The idea is to treat the second term as small in the evaluation of the T matrix via (37) and make a perturbation expansion in that small quantity. This was suggested by Glauber[2] and carried out elegantly by Blankenbecler and Sugar[3].

Let us call the unperturbed Green function G_i and the perturbation F_i, so

$$G_i^{-1} = \frac{k^2 - 2\vec{p}\cdot\vec{k}_i + k_i^2}{2m} - V + i \tag{45}$$

and $$F_i = (\vec{p} - \vec{k}_i)^2/2m. \tag{46}$$

In the spirit of our argument we could equally well have made an expansion about the final momentum $\vec{k}_f$ since this is supposed to differ only slightly from $\vec{k}_i$. We would be led then to

$$G^{-1} = G_f^{-1} - F_f \tag{47}$$

with

$$G_f^{-1} = \frac{k^2 - 2\vec{p}\cdot\vec{k}_f + k_f^2}{2m} - V + i\epsilon \tag{48}$$

and

$$F_f = (\vec{p} - \vec{k}_f)^2/2m. \tag{49}$$

A symmetric form of approximation to G is then suggested. We note that

$$G = G_i + G_i F_i G = G_i + G F_i G_i, \tag{50}$$

and $$G = G_f + G_f F_f F = G_f + G F_f G_f. \tag{51}$$

Using these together yields

$$G = \tfrac{1}{2}(G_i + G_f) + \tfrac{1}{2}G_f (F_i + F_f) G_f + G_f F_f G F_i G_i. \tag{52}$$

The contribution to the T matrix from the zeroth order term in F's is equal to

$$T_E = V + \tfrac{1}{2}V(G_i + G_f)V, \tag{53}$$

$$= T_{Ei} + T_{Ef} \tag{54}$$

with $$T_{Ei} = \tfrac{1}{2}(V + V G_i V), \tag{55}$$

and a similar expression for T_{Ef}. We may formally recast T_{Ei} into the form

$$T_{Ei} = \frac{V}{2}\left(1 + \frac{}{\left(\frac{k_i^2 + k^2 - 2\vec{p}\cdot\vec{k}_i}{2m}\right) - V + i\epsilon}\, V\right) \tag{56}$$

$$= \tfrac{1}{2}\, V\, G_i\left(\frac{k_i^2 + k^2 - 2\vec{p}\cdot\vec{k}_i}{2m}\right). \tag{57}$$

So that the matrix element $\langle \vec{k}_f \,|T_{Ei}|\, \vec{k}_i \rangle$ is

$$\langle \vec{k}_f \,|T_{Ei}|\, \vec{k}_i \rangle = \lim_{k^2 \to k_i^2} \langle \vec{k}_f \,|V\, G_i\, |\vec{k}_i \rangle \frac{(k^2 - k_i^2)}{4m}. \tag{58}$$

This is most easily evaluated in a mixed representation to yield

$$\langle \vec{k}_f \,|T_{Ei}|\vec{k}_i \rangle = \lim_{k^2 \to k_i^2} \int d^3r\, e^{-i\vec{k}_f\cdot\vec{r}} V(\vec{r}) \frac{(k^2 - k_i^2)}{4m}\, G_i(\vec{r},\vec{k}_i), \tag{59}$$

where the Green function $G_i(\vec{r},\vec{k}_i)$ satisfies

$$\left[\frac{k^2 + k_i^2 + 2i\,\vec{k}_i\cdot\nabla_r}{2m} - V(\vec{r})\right]\, G_i(\vec{r},\vec{k}_i) = e^{i\,\vec{k}_i\cdot\,\vec{r}}, \tag{60}$$

whose resemblance to (19) is not entirely coincidental. Having solved this equation before we will not delve into fancier techniques to do it again[4], but remind ourselves of the answer

$$G_i(\vec{r},\vec{k}_i) = i\int_0^\infty dt\, \exp\left\{i\frac{(k^2 - k_i^2)}{2m}t - i\int_0^t d\tau\, V\left(\vec{r} - \frac{\vec{k}_i\tau}{m}\right)\right\}. \tag{61}$$

When we take the limit to go onto the energy shell, we find

$$\langle \vec{k}_f \,|T_{Ei}|\, \vec{k}_i \rangle = \tfrac{1}{2}\int d^3r\, V(\vec{r})\, e^{-i\vec{\Delta}\cdot\vec{r}}\, \exp -i\int_0^\infty d\tau\, V\left(\frac{\vec{r}-\vec{k}_i\tau}{m}\right), \tag{62}$$

and for the term in which we expand about $\vec{k}_f$ the same operations yield

$$\langle \vec{k}_f \,|T_{Ef}|\, \vec{k}_i \rangle = \tfrac{1}{2}\int d^3r\; V(\vec{r})\; e^{-i\vec{\Delta}\cdot\vec{r}}\; \exp - i \int_0^\infty d\tau\; V(\vec{r}+\frac{\vec{k}_f \tau}{m}) . \tag{63}$$

We may understand these formulae as saying that, for (62) say, the particle begins the scattering process by entering as a plane wave exp i $\vec{k}_i \cdot \vec{r}$ and picks up a phase due to movement through the potential along the direction $\hat{k}_i$ until it reaches the potential at $\vec{r}$. At this point it scatters once, as in a Born approximation, and leaves as a plane wave along $\hat{k}_f$.

If we decompose three space as we did previously into a piece along $\hat{k}_i$ and coordinates orthogonal to $\hat{k}_i$ we may cast $T_{Ei}(\vec{k}_f, \vec{k}_i)$ into

$$T_{Ei}(\vec{k}_f,\vec{k}_i) = \tfrac{1}{2}\int d^3r\; e^{-i\vec{k}_f\cdot\vec{r}}\; -i\,\frac{m}{k}\int_{-\infty}^{\lambda} d\tau\; V(\vec{b} + \hat{k}_i\,\tau) , \tag{64}$$

in which our modulated plane wave of before is now exhitited. {The boundary condition on it is now taken at $-\infty$.} Indeed, recall the general form for the full T-matrix

$$T(\vec{k}_f,\vec{k}_i) = \int d^3r\; e^{-i\vec{k}_f\cdot\vec{r}}\; V(\vec{r})\; \psi_{k_i}^{(+)}(\vec{r}), \tag{65}$$

where $\psi_{k_i}^{(+)}(\vec{r})$ is the solution to the Schrodinger equation having as boundary conditions incoming plane waves $e^{i\vec{k}_i\cdot\vec{r}}$ and outgoing spherical waves. Our method here has been precisely to make a modulated wave approximation for $\psi_{k_i}^{(+)}(\vec{r})$.

If we introduce the shorthand for the phases associated with the eikonal approximation

$$\chi_i(\vec{r}) = \int_0^\infty d\tau\; V(\vec{r} - \vec{k}_i\tau/m), \tag{66}$$

and

$$\chi_f(\vec{r}) = \int_0^\infty d\tau\; V(\vec{r} + \vec{k}_f\tau/m), \tag{67}$$

the full first eikonal approximation reads

$$T_E(\vec{k}_f,\vec{k}_i) = T_{Ei}(\vec{k}_f,\vec{k}_i) + T_{Ef}(\vec{k}_f,\vec{k}_i)$$
$$= \tfrac{1}{2}\int d^3r\, V(\vec{r})\, e^{-i\vec{\Delta}\cdot\vec{r}} \left[e^{-i\chi_i(\vec{r})} + e^{-i\chi_f(\vec{r})}\right] \tag{68}$$

Consider now the limit in which the energy is very large and neglect the difference between k_i and k_f in the phases χ. When we write

$$\vec{r} = \vec{b} + \lambda\hat{k}$$

and note that $\vec{r}\cdot\vec{\Delta} \approx \vec{b}\cdot\vec{\Delta}$, T_E may be expressed as

$$T_E(k^2,\Delta) = \int d^2b \int_{-\infty}^{+\infty} d\lambda\, V(\vec{b} + \lambda\vec{k})\, e^{-i\vec{\Delta}\cdot\vec{b}}\, e^{-i\frac{m}{k}\int_{-\infty}^{\lambda} d\tau V(\vec{b} + \tau\hat{k})} \tag{69}$$

$$= \int d^2b\, e^{-i\vec{\Delta}\cdot\vec{b}}\, \frac{ik}{m} \int_{-\infty}^{+\infty} d\lambda\, \frac{d}{d\lambda}\, e^{-i\frac{m}{k}\int_{-\infty}^{\lambda} d\tau\, V(\vec{b} + \tau\hat{k})} \tag{70}$$

$$= (ik/m) \int d^2b \left(e^{-i\chi(\vec{b},k)} - 1\right) e^{-i\vec{\Delta}\cdot\vec{b}}, \tag{71}$$

where

$$\chi(\vec{b},k) = \frac{m}{k} \int_{-\infty}^{+\infty} d\tau\, V(\vec{b} + \tau\hat{k}). \tag{72}$$

In these expressions $\vec{\Delta}$ is to be interpreted as a two vector whose magnitude is the momentum transferred by the potential

$$\Delta^2 = 2k^2\,(1 - \cos\theta), \tag{73}$$

$\cos\theta = \hat{k}_f\cdot\hat{k}_i$. The form (71) of the eikonal approximation as a two dimensional integral over the space impact parameters $\vec{b}$ is probably most familiar.

We can cast the whole eikonal expression into a two dimensional form by introducing the Fourier transform of the potential $\tilde{V}(\vec{q})$ into (72)

$$\chi(\vec{b},k) = \frac{m}{k}\int_{-\infty}^{+\infty} d\tau \int \frac{d^2q}{(2\pi)^2}\,\frac{d(\vec{q}\cdot\hat{k})}{2\pi}\, e^{i\mathbf{q}\cdot\vec{b} + i\vec{q}\cdot\hat{k}\tau}\,\tilde{V}(q,\vec{q}\cdot\hat{k})$$

$$= \frac{m}{k}\int \frac{d^2q}{(2\pi)^2}\, e^{i\,\mathbf{q}\cdot\vec{b}}\,\tilde{V}(q), \qquad (74)$$

where $\mathbf{q}$ is the two vector $\vec{q} - \hat{k}(\vec{q}\cdot\hat{k})$. The eikonal T-matrix T_E is thus revealed to be dependent in its dynamical content, namely its dependence on V, only through those coordinates transverse to the high speed initial beam. The longitudinal dependences have become essentially a kinematic structure appearing through the magnitude of k_i or k_f only. This is a real virtue of eikonal approximations since in actual high energy collisions there appears to be a decoupling between longitudinal and transverse degrees of freedom. More precisely stated we would note that the eikonal form for T provides a natural framework in which to describe the transverse dynamics with the more trivial longitudinal properties already accounted for. What one must do still is to demonstrate the relevance of the form of (72) and (74) for relativistic problems, but we shall soon come to that.

At this point, however, it will be useful to discuss the particular potential

$$V(\vec{r}) = -\,Z\,e^2\,\frac{e^{-\mu|\vec{r}|}}{|\vec{r}|} \qquad (75)$$

both for its obvious physical interest and as a particular exercise in our formalism. First we evaluate $\chi(\vec{b},k)$

$$\chi(\vec{b},k) = \frac{m}{k}\int_{-\infty}^{+\infty} d\tau\,(-Ze^2)\,\frac{e^{-\mu\sqrt{b^2+\tau^2}}}{\sqrt{b^2+\tau^2}}, \qquad (76)$$

$$\text{or } \chi(\vec{b},k) = -2\,\frac{mZe^2}{k}\int_0^{\infty} d\theta\, e^{-\mu\, b\,\cosh\,\theta}$$

$$= \frac{-2mZe^2}{k}\,K_0(\mu b), \qquad (77)$$

where $K_o(\mu b)$ is a modified bessel function. The eikonal T-matrix is then

$$T_E(\Delta,k) = \frac{ik}{m} \int d^2b\, e^{-i\vec{\Delta}\cdot\vec{b}} \left[e^{\frac{2im\mathcal{Z}e^2}{k} K_o(\mu b)} - 1 \right]. \quad (78)$$

For large k and fixed Δ, the only term which survives is the k independent term which a moment's contemplation shows to be the Born term:

$$T_E(\Delta,k) \underset{\substack{k\to\infty \\ \Delta \text{ fixed}}}{\sim} \tilde{V}(\Delta) + 0(1/k). \quad (79)$$

This is known to be the correct limit for the potential scattering T-matrix. Clearly, T_E has this limit for a wide class of potentials. Further I will let the reader verify that our T_E satisfies the optical theorem when we approximate the angular integral involved by its value near $\Delta \approx 0$. That is, we take T_E seriously only within the forward scattering cone but argue that it constitutes so much of the actual amplitude that we may use it over all angular intervals.

Suppose we now examine the eikonal T_E for very small values of μ; namely, as we go over to the Coulomb potential. For small argument $K_o(\chi)$ behaves as

$$K_o(\chi) \approx -(\gamma + \log\frac{\chi}{2})\quad (1 + \frac{\chi^2}{4} + \ldots) + \chi^2/4 + \ldots \quad , \quad (80)$$

where $\gamma = 0.577\ldots$.

So $T_E(\Delta,k)$ is approximately

$$T_E(\Delta,k) \underset{\mu\to 0}{\sim} \frac{i\,k}{m} e^{\frac{-2im\gamma\mathcal{Z}e^2}{k}} \int d^2b\, e^{-i\vec{\Delta}\cdot\vec{b}} \left[\left(\frac{\mu b}{2}\right)^{\frac{-2im\mathcal{Z}e^2}{k}} - 1 \right] \quad (81)$$

The - 1 term doesn't contrkbute for $\Delta \neq 0$ and we will now

agree to stay away from that point. The remaining integral is

$$T_E(\Delta,k) \underset{\mu\to 0}{\sim} \frac{2\pi i k}{m} \frac{e^{\frac{-2im\gamma Ze^2}{k}}}{\Delta^2} \left(\frac{\mu\Delta}{2}\right)^{\frac{-2imZe^2}{k}} \times$$

$$\times \int_0^\infty dx\, J_0(x)\, x^{1-\frac{2imZe^2}{k}}, \qquad (82)$$

or upon dropping an irrelevant phase and a tricky phase associated with the $\mu\Delta$ term, we may write

$$T_{E\ \text{Coulomb}}(\Delta,k) = -\frac{4\pi Ze^2}{\Delta^2} \frac{\Gamma(1 - iZe^2m/k)}{\Gamma(1 + iZe^2m/k)}. \qquad (83)$$

This function has a set of poles at the points

$$1 - \frac{iZe^2m}{k} = -n, \quad n=0,1,2, \qquad (84)$$

or $$E_n = \frac{\hbar^2k^2}{2m} = \frac{-Z^2Ry}{(n+1)^2}, \qquad (85)$$

where $$Ry = \frac{me^4}{2\hbar^2} = 13.5 \text{ e V}. \qquad (86)$$

There are familiar poles, and it is rather a pleasant surprise to find them in a T-matrix whose realm of validity we suspected to be for $E \to \infty$, not for E below the threshold in the bound state region. What explanation can we find for this? Perhaps the following heuristic argument will appeal to the reader (perhaps not): Binding by a potential is a non-perturbative process involving an infinite number of interactions with that potential. Consider a particle which has just interacted with a potential . If it received a large momentum kick from that interaction it will have a very small overlap with the bound state wave function it is trying to generate. One might well expect that, therefore, the primary contribution to binding comes from the low momentum components of the potential, i.e. those governing the long wavelength

behavior of V. Since these are what we have summed up in the eikonal approximation we are not so surprised the binding energies are reasonable. What one should, I think, be surprised at is the fact that the exact answer emerges. This is not the only surprise the Coulomb potential has ever produced. A more significant test of T_E would come in examining the binding energies for the Yukawa potential with $\mu \neq 0$; this has not yet been done.

Now instead of evaluating the corrections to the first eikonal approximation which are contained in our expression (52) for G, let us mention a technique of approximation which is similar to that we employed for the modulating function in the previous section. As we now know the problem we are interested in is essentially that of constructing G. When we have G then we can find $\langle \vec{k}_f | T | \vec{k}_i \rangle$ by

$$\langle k_f | T | k_i \rangle = \lim_{k^2 \to k_i^2} \left[\frac{k^2 - k_i^2}{2m} \right] \int d^3 r e^{-i\vec{k}_f \cdot \vec{r}} V(\vec{r}) \; G(\vec{r}, \vec{k}_i) \tag{87}$$

and we know $G(\vec{r}, \vec{k}_i)$ satisfies the differential equation

$$\left[E + \frac{\nabla_r^2}{2m} - V(\vec{r}) \right] G(\vec{r}, \vec{k}_i) = e^{i\vec{k}_i \cdot \vec{r}} , \tag{88}$$

so that $H(\vec{r}, \vec{k}_i) = e^{-i\vec{k}_i \cdot \vec{r}} \; G(\vec{r}, \vec{k}_i)$ will obey

$$\left[E + \frac{(\nabla_r + i\vec{k}_i)^2}{2m} - V(\vec{r}) \right] H(\vec{r}, \vec{k}_i) = 1. \tag{89}$$

Now we seek a solution of this in the form[5].

$$H(\vec{r}, \vec{k}_i) = i \int_0^\infty d\lambda \; e^{-i\left(\frac{k^2 - k_i^2}{2m} - i\varepsilon\right)\lambda + F(\vec{r}, \vec{k}_i, \lambda)} , \tag{90}$$

which has the virtue that when V = 0, F = 0 and

$H(\vec{r}, \vec{k}_i) = \left(\frac{k^2 - k_i^2}{2m} - i\varepsilon \right)^{-1}$ which is expected. In the now familiar manner we will find the differential equation

for F to be

$$\left(\frac{i\vec{k}_i\cdot\nabla_r}{m} + \nabla_r{}^2/2m+i\frac{\partial}{\partial\lambda}\right)F(\vec{r},\vec{k}_i\lambda) - V(\vec{r})+\nabla_r F(\vec{r},\vec{k}_i,\lambda)\cdot\nabla_r F(\vec{r},\vec{k}_i\lambda) = 0. \quad (91)$$

There are two routes to take now: (1) treat the $\nabla^2 F$ and $(\nabla F)\cdot(\nabla F)$ terms perturbatively as we have done above or (2) expand F in powers of the potential by giving it a scale factor g and writing [5].

$$F(\vec{r},\vec{k}_i,\lambda) = \sum_{n=1}^{\infty} g^n F_n(\vec{r},\vec{k}_i,\lambda). \quad (92)$$

For variety we will pursue the latter path. The lowest order term $F_1(\vec{r},\vec{k}_i,\lambda)$ satisfies

$$\left(\frac{i\vec{k}_i\cdot\nabla_r}{m} + \frac{\nabla r^2}{2m} + i\frac{\partial}{\partial\lambda}\right)F_1(\vec{r},\vec{k}_i,\lambda) = V(\vec{r}), \quad (93)$$

and has the solution which vanishes at $\lambda = 0$:

$$F_1(\vec{r},\vec{k}_i,\lambda) = -i\int_0^{\lambda} d\tau \int\frac{d^3q}{(2\pi)^3}\,\tilde{V}(\vec{q})\; e^{i\vec{q}\cdot\vec{r}-i\frac{\vec{k}_i\cdot\vec{q}}{m}\tau - i\tau q^2/2m}, \quad (94)$$

which differs from our previous form by the $q^2/2m$ term. This can be regarded as a form of curvature correction to the straight line approximation of our other examples. The T-matrix which results from this method is from (87)

$$T(\vec{k}_f,\vec{k}_i) = \int d^3r\; e^{-i\vec{\Delta}\cdot\vec{r}}\; V(\vec{r})\; e^{F(\vec{r},\vec{k}_i,\infty)} \quad ; \quad (95)$$

this is a general result. Inserting (94) gives this version of the eikonal approximation. Again the major virtue of this technique is to provide a final expression for the T-matrix which is a single integral rather than a sum of terms.

The potentials we have treated to this point are $\vec{r}$ dependent potentials only. It is of some interest to explore the eikonal T-matrices for certain energy dependent

potentials also. The first one we shall explore is an old friend, the spin orbit potential,

$$V = V_1(r) + \vec{\sigma}\cdot\vec{L}\, V_2(r) \ , \tag{96}$$

where we take V_1, and V_2 to be central and the $\vec{\sigma}$ operator could refer to any spin, but we'll be happy with spin $\frac{1}{2}$. The operator $\vec{L}$ is, of course, $\vec{r} \times \vec{p}$.

The Green function we wish is

$$G^{-1}(E) = E - \frac{\vec{p}^2}{2m} - V_1(r) - (\vec{\sigma} \times \vec{r})\cdot\vec{p}V_2(r) + i\epsilon \tag{97}$$

We will approximate this by expanding about $\vec{k}_i$ as usual and treating the remainder as a perturbation. The Green function for the basic eikonal path will be

$$G_i^{-1}(E) = \frac{k^2+k_i^2-2\vec{p}\cdot\vec{k}_i}{2m} - V_1 - (\vec{\sigma}\times\vec{r})\cdot\vec{k}_i V_2 + i\epsilon. \tag{98}$$

The T-matrix which results is easily read off from (71) and (72). We have

$$\chi(\vec{b},k) = \frac{m}{k}\int_{-\infty}^{+\infty} V_1(\vec{b}+\tau\hat{k})d\tau + m\sigma\cdot(\vec{b}\times\hat{k})\int_{-\infty}^{+\infty} V_2(\vec{b}+\tau\hat{k})d\tau, \tag{99}$$

$$= \frac{m}{k}\chi_1(\vec{b}) + m\,\vec{\sigma}\cdot(\vec{b}\times\hat{k})\chi_2(\vec{b}), \tag{100}$$

and

$$T_E^{\text{Spin orbit}}(k,\Delta) = \frac{ik}{m}\int d^2b\, e^{-i\vec{\Delta}\cdot\vec{b}}\, e^{-i\frac{m}{k}\chi(\vec{b},k)}\left\{\cos mb\chi_2(\vec{b}) - \right.$$

$$\left. -i\,\vec{\sigma}\cdot(\vec{b}/b \times \hat{k})\sin mb\,\chi_2(\vec{b})\right\}. \tag{101}$$

There are two basic differences between this result and (71). First, of course, there is the additional spin dependent term. More significant, however, is the limit as $k\to\infty$ for fixed momentum transfer Δ. Whereas the ordinary central potential yielded a constant in k which we recognized to be the Born approximation, for this potential we see that in this limit the χ_1 term is irrelevant and we are

left with

$$T_E^{\text{Spin orbit}}(k,\Delta) \underset{\substack{k\to\infty \\ \Delta\text{fixed}}}{\sim} i\,k\,F(\Delta). \tag{102}$$

Now this is an intriguing result for we know that in actual high energy scattering experiments the elastic scattering amplitude does not retreat to the Born approximation but becomes very close to being purely imaginary and grows with the center of mass energy W as W^2. Here we have a clue as to how this might transpire. We have yet to discuss relativistic scattering, but we are encouraged to seek the equivalent of a spin-orbit potential and must discover how k goes over into W^2.

Next consider the potential

$$V = \vec{L}^2\, V_o\, e^{-\mu r}/r \tag{103}$$

where V_o is a constant. This is not a terribly realistic form for V, but it will illustrate an interesting point. As usual we expand the full Green function about $\vec{p} = \vec{k}_i$ or $\vec{k}_f$ and via (71) and (72) construct T_E. The eikonal phase is

$$\chi(\vec{b},k) = 2\frac{m}{k}\,(k^2b^2)\,V_o\,K_o(\mu b), \tag{104}$$

and

$$T_E(k,\Delta) = \frac{i\,k}{m}\int d^2b\; e^{-i\vec{\Delta}\cdot\vec{b}}\left[e^{-2imkb^2V_oK_o(\mu b)} - 1\right]. \tag{105}$$

Now for fixed Δ and large k we encounter a rather different situation from our previous examples. In this case the phase in the exponential oscillates wildly and we must seek the point of stationary phase; that is, we must search for a minimum of

$$V_o\; 2m\; k\; b^2\; K_o(\mu b). \tag{106}$$

The small b region does not contribute much to the limit we are taking. However, if we note the large b behavior

of $K_o(\mu b)$, namely

$$K_o(\mu b) \underset{b\to\infty}{\sim} \left(\frac{\pi}{2\mu b}\right)^{\frac{1}{2}} e^{-\mu b} \left[1 - \frac{1}{8\mu b} + \ldots\right], \qquad (107)$$

we see that we may achieve a stationary phase in the region of b space where

$$b \approx \log \frac{2m\, k\, V_o}{\mu^2}, \qquad (108)$$

and that the leading behavior of the eikonal amplitude is

$$T_E(k,\Delta) \underset{\substack{k\to\infty \\ \Delta \text{ fixed}}}{\sim} i\, k(\log k)^2, \qquad (109)$$

dropping some uninteresting constants.

Two comments are in order about this result. First, the same leading behavior would occur if higher powers of energy dependence were allowed to enter the model potential (103), so that (109) is in some sense a maximum high energy behavior. Second, were we to replace k by W^2 to pretend we were relativistic, we would recognize this behavior as exactly the Froissart bound. Since that bound is deeply ingrained in our beliefs of what is consistent with unitarity, it is pleasant indeed to have an approximation scheme which respects it.

This ends what I have to say about the use of eikonal methods in potential scattering. There are a number of real uses of these methods in nuclear physics where potential theory appears to be important. I refer you again to Glauber's lectures and subsequent work by many physicists for discussion of these interesting topics. As for us, we will proceed on to field theory.

IV. Field Theory

Since relativistic field theory is much richer in its physical content than potential scattering, it will take a bit more cleverness to extract from it the equivalent of our eikonal approximations. The approach we will take is

to effectively reduce field theory to a problem in potential scattering and then begin approximating. The basic observation necessary for this task was made by Feynman and Schwinger so long ago that we have ceased teaching it in our schools. Namely, particle emission, absorption, and exchange which we think are the basic processes in the description of forces among relativistic quantum objects are all contained in the knowledge of the response of a particle to external potentials (or as some would put it more forcefully these days, external c-number sources.) More precisely, if we know the amplitude T(A) for motion in an external potential A(x) we can extract from it the amplitude for emission and absorption of the quanta associated with the quantum numbers of A(x) by varying T(A) with respect to A(x). Since A(x) is a function, we will need to take functional derivatives, but we'll proceed as if they were ordinary derivatives and leave any problems to the squeamish.

We will deal in these lectures exclusively with quantum electrodynamics of (usually) massive photons. So the object of interest to us is the amplitude for some process occurring in an external c-number potential $A_\alpha(x)$. This potential satisfies

$$(\partial^2 + \mu^2)\, A_\alpha(x) = J_\alpha(x), \tag{110}$$

where $J_\alpha(x)$ is the conserved external current giving rise to $A_\alpha(x)$ and μ is the mass associated with the quantized A_α field. We establish our notation a bit more by reminding ourselves that the solution to (110) involving causal propagation is

$$A_\alpha(x) = \int d^4 y\; D_+(x - y)\; J_\alpha(y), \tag{111}$$

with

$$D_+(x) = \int \frac{d^4 q}{(2\pi)^4}\; \frac{e^{-iq\cdot x}}{\mu^2 - q^2 - i\epsilon}\,. \tag{112}$$

If we are given T(A), the transition matrix element for the processes in $A_\alpha(x)$ and wish instead the amplitude for the same process to occur accompanied with the emission of a photon (of mass μ) with wave function $G_\beta(x)$, then we ask how T(A) varies as we alter $A_\alpha(x)$, then replace at x $A_\alpha(x)$

by $\mathcal{A}_\alpha(x)$ and integrate over all x. Thus we arrive at

$$\int d^4x \; \frac{\delta T(A)}{\delta A_\alpha(x)} \; \mathcal{A}^*_\alpha(x) \Big| \tag{113}$$

as the amplitude for T plus the emission of a photon $\mathcal{A}_\alpha$. If we are interested in the final process in the absence of the external field, we simply set A - 0 at the end of all operations.

An example of this may be useful. Suppose we have an electron described by the Dirac equation

$$(\not{p} - e\not{A} - m)\,\psi = 0 \tag{114}$$

in an external potential $A_\alpha(x)$.[7] The amplitude for the electron to go from momentum p, and spin λ_1 to p_2 and λ_2 to second order in the action of $A_\alpha(x)$ is as usual

$$T(p_1 \to p_2; A) = \int d^4z d^4y \;\; \bar{u}(p_2,\lambda_2)\; e^{-ip_2 z}\; e\gamma\cdot A(z) \;\times$$

$$\times\; S(z-y)\; e\gamma\cdot A(y)\; e^{ip_1\cdot y}\; u(p_1,\lambda_1), \tag{115}$$

with S(z-y) the electron causal propagator. From this we may find the amplitude for the electron to scatter to first order in the potential $A_\alpha(x)$ and also emit a photon of wave function $\mathcal{A}_\alpha(x) = \epsilon'_\alpha\, e^{iq_2\cdot x}$. Using (113) there are two terms

$$e^2 \epsilon'^{\alpha *}\, \bar{u}(p_2,\lambda_2) \int d^4z d^4y \Big\{ e^{-i(q_2+p_2)\cdot z}\; \gamma_\alpha\; S(z-y)\,\gamma\cdot A(y) e^{ip_1\cdot y}$$

$$+ e^{-ip_2\cdot z}\; \gamma\cdot A(z) S(z-y)\; \gamma_\alpha e^{i(p_1-q_2)\cdot y} \Big\}\, u(p_1,\lambda_1). \tag{116}$$

These are just the well known Feynman graphs which include emission of the photon of momentum q_2 before and after the action of the potential A_α.

Also contained in (115) is the amplitude for the absorption of a photon of momentum q_1 and then the emission of q_2 while the electron goes from p_1 to p_2. We arrive at this piece of the Compton effect by applying (113) to our last result

$$T(p_1,\lambda_1 + q_1,\epsilon \to p_2,\lambda_2 + q_2,\epsilon') =$$

$$e^2\ \epsilon'^{\alpha *}\ \epsilon^{\beta}\ \bar{u}(p_2,\lambda_2) \int d^4z d^4y \left\{ \right.$$

$$e^{-i(p_1+q_2)\cdot z}\ \gamma_\alpha S(z-y)\gamma_\beta e^{i(p_1+q_1)\cdot y} + e^{-i(p_2-q_1)\cdot z}\gamma_\beta S(z-y)\gamma_\alpha \times$$

$$\left. \times e^{i(p_1-q_2)\cdot y} \right\}\ u(p_1,\lambda_1), \qquad (117)$$

again a well known result for the Born approximation to the Compton effect.

Finally there is another piece of information in (115), namely the amplitude for p_1 to go to p_2 including radiation corrections to order e^2. That is, the first order in e^2 corrections to the electron propagator $S(x-y)$. To exhibit this we need consider the possibility that the electron emits a photon of momentum q and then reabsorbs it, the initial emission acting as a source for the subsequent absorption. To emit the photon we take one derivative of $T(A)$ and to absorb it another. Since it is the same photon which is emitted and absorbed we must supply a $D_+(x-y)$ to take it from one act to the other. The resultant amplitude for $p_1 \to p_2$ to our order in e^2 is

$$\tfrac{1}{2}\int d^4x d^4y\ \frac{\delta}{\delta A_\alpha(x)}\ \ g_{\alpha\beta}D_+(x-y)\ \frac{\delta}{\delta A_\beta(y)}\ T(p_1 \to p_2;A), \qquad (118)$$

the factor of $\frac{1}{2}$ arising because we integrate over all possible positions of photon creation and annihilation.

We are prepared now to construct a framework in which to describe interaction. Essentially we shall ask that we know how a particle, call it number 1, moves in an external potential $A_{1\alpha}$, while, independently, number 2 moves in $A_{2\alpha}$. Next we ask of 1 that it emit a photon via $\delta T_1(A_1)/\delta A_{1\alpha}(x)$, then allow the photon to go to y by $D_+(y-x)$ and be absorbed by 2 there via $\delta T_2(A_2)/\delta A_{2\alpha}(y)$. The propagation function $D_+(z)$ acts as the potential or mediator of the reaction. If our experimence with potential theory is relevant (and it is) we shall expect to find that the two particle scattering amplitude is at very large energies of an eikonal form with the eikonal phase given by an

appropriate integral over $D_+(x)$.

Before we launch into more formalism it will be useful to see what this "appropriate integral" must be. Recall that in potential theory one spatial direction $\hat{k}$ was singled out and that the eikonal phase was a path integral over the potential along this direction. In relativistic scattering two "directions" are singled out by the particles coming along $\pm \hat{k}$. Take the spatial direction to be the z-axis, then for very large momenta the trajectories of colliding particles are near the light cones $z = \pm t$. We may anticipate that the eikonal phase will be proportional to

$$\int_{-\infty}^{+\infty} dz dt \; D_+(x,y,z,t), \tag{119}$$

which is $$\int \frac{d^2q}{(2\pi)} \frac{e^{+iq\cdot\vec{b}}}{\mu^2 + |q|^2} = K_0(\mu b) \tag{120}$$

with $\vec{b} = (x,y)$. Furthermore, since in electrodynamics we are dealing with the relativistic analog of the spin orbit force, (96), we can expect that no energy dependence will appear in the eikonal phase.

To demonstrate that these expectations are fulfilled we need to construct an equation for the Green function of an electron moving in an external field $A_\alpha(x)$. Were there no self-interaction, we know from the Dirac equation that the Green function $\mathcal{G}_A(x,y)$ would satisfy

$$[i \not\partial_x - m - e \not{A}(x)] \mathcal{G}_A(x,y) = \delta^4(x-y), \tag{121}$$

with appropriate boundary conditions. With self action the Green function must be altered to reflect the possibility of the electron providing its own potential $A^\alpha{}_{self}(x)$. This self potential can be found by asking how the electron responds to an external source J_α and (121) is altered to read

$$\left[i \not\partial_x - m - e\left(\not{A}(x) - i\gamma_\beta \frac{\delta}{\delta J_\beta(x)}\right)\right] G_A(x,y) = \delta^4(x-y), \tag{122}$$

and at the same time the potential $A_\alpha(x)$ must satisfy

$$(\mu^2 + \partial^2)\, A_\alpha(x) = J_\alpha(x) + \frac{i\,e}{2}\, \mathrm{tr}\, \gamma_\alpha\, G_A(x,x). \qquad (123)$$

The derivation of these equations is given by Schwinger[6] and by Fradkin[5]. We pass over that but pause to interpret them. The appearance of $\delta/\delta\, J_\beta$ in (122) is exactly what one would expect since it is the response to an external source that provides a potential. The self potential of an electron is thus provided by its own response to an outside source. In (123) the "extra" term on the right is just the source term of photons due to the electron current $\bar{\psi}(x)\, \gamma_\alpha \psi(x)$. It gives rise to vacuum polarization effects in photon propagation and is the origin of most of the interesting non-linear aspects of field theory.

Our first approximation is to either ignore this polarization term or to replace it by some appropriately path averaged equivalent. That is we begin our discussion of motion in an external potential by ignoring the ability of photons to create pairs or by replacing the photon propogator $D_+(x-y)$ as given above by an effective $\tilde{D}$ reflecting some knowledge of the off shell behavior of D. This is tantamount to discarding in graphical language all those diagrams in which an electron pair lying on a photon line connects to anything else. With this approximation we may eliminate $\delta/\delta J$ in favor of $\delta/\delta A$ in the equation for $\mathcal{G}_A$ yielding

$$\left(i\not{\partial}_x - m - e\left[\not{A} - i\gamma_\beta \int d^4z\, D_+(x-z)\, \frac{\delta}{\delta A_\beta(z)}\right]\right) G_A(x,y) = \delta^4(x-y) \qquad (124)$$

The advantage of this result is immediate since a formal solution to (124) is

$$G_A(x,y) = \exp\left(-\tfrac{1}{2}\int d^4z\, d^4w\, \frac{\delta}{\delta A_\alpha(z)}\, D_+(z-w)\, \frac{\delta}{\delta A_\alpha(w)}\right) \mathcal{G}_A(x,y), \qquad (125)$$

where $\mathcal{G}_A$ satisfies (121). This states that if we know how to construct the Green function $\mathcal{G}_A$ for motion in the absence of vacuum polarization and self action, then we will, by differentiation, be able to extract the full Green function

$\mathcal{G}_A$ for electron motion in an external potential without vacuum polarization effects.

Constructing $\mathcal{G}_A$, of course, is no easy task in general. It is, however, exactly the relativistic equivalent of evaluating the Green function in a potential V that we spent the last sections learning about. To proceed let's introduce an operator notation whereby $\mathcal{G}_A(x,y)$ is regarded as a matrix element of an operator $\mathcal{G}_A$ in an appropriate space

$$\mathcal{G}_A(x,y) = (x|\mathcal{G}_A|y). \tag{124}$$

The operator P is $i\nabla_x$ as usual and the operator X has the commutation relation with it

$$[X_\mu, P_\nu] = -i\, g_{\mu\nu}. \tag{125}$$

The Green function equation reads

$$[\not{P} - m - e\not{A}(X)]\,\mathcal{G}_A = 1 \tag{126}$$

in this notation.

Now suppose we are interested in a process where a very fast electron of momentum p_1 scatters from our external potential into a state of momentum p_1'. If we evaluate $(p_1'|\mathcal{G}_A|p_1)$, then the transition matrix $(p_1')|T(A)|p_1)$ follows in the usual manner

$$(p_1'|T(A)|p_1) = e(p_1'|\not{A}\,\mathcal{G}_A\,\mathcal{G}_o^{-1}|p_1), \tag{127}$$

where $\mathcal{G}_o = \mathcal{G}_{A=0} = (\not{P} - m)^{-1}$.

As in the non-relativistic case it is useful to expand P about the fast momentum p_1 which is a c-number. We write

$$\mathcal{G}_A = 1/(\not{P} - m - e\not{A}) \tag{128}$$

$$= \frac{1}{(m + \not{p}_1)\,(\not{P} - m - e\not{A})}\,(m + \not{p}_1), \tag{129}$$

$$= \frac{1}{2p_1\cdot P - (m^2+p_1^2) - e(m+\not{p}_1)\not{A} + (\not{P}-\not{p}_1)(m-\not{p}_1)}\,(m + \not{p}_1), \tag{130}$$

and begin expanding about p_1 by treating the term $(\not{P}-\not{p}_1)\times$ $\times(m-\not{p}_1)$ as a perturbation. That is, write

$$2p_1\cdot P - (m^2+p_1^2) - e(m+\not{p}_1)\not{A} = \mathcal{G}_1^{-1}(A), \tag{131}$$

and $$\mathcal{G}_A = \frac{1}{\mathcal{G}_1^{-1}(A) - F_1}\ (m+\not{p}_1), \tag{132}$$

with $F_1 = -(\not{P}-\not{p}_1)(m-\not{p}_1)$.

The first eikonal approximation to $\mathcal{G}_A$ becomes

$$\mathcal{G}_1(A)\ (m+\not{p}_1). \tag{133}$$

The next approximation would be

$$\mathcal{G}_1(A)\ F_1\ \mathcal{G}_1(A)\ (m+\not{p}_1), \tag{134}$$

and so forth. If one wished to expand about the final momentum, p_1', then the appropriate object to construct would be

$$\mathcal{G}_1'(A) = 2p_1'\cdot P - (m^2+p_1'^2) - e\not{A}(m+\not{p}_1'), \tag{135}$$

and the first eikonal Green function is

$$(m+\not{p}_1')\ \mathcal{G}_1'(A). \tag{136}$$

Let's consider in some detail the eikonal T-matrix arising from the approximate $\mathcal{G}_A$ in (133). The on mass-shell T matrix will be given by

$$\mathcal{T}_E(p_1\to p_1';A) = \lim_{p_1^2\to m^2} e(p_1'|A(X)\mathcal{G}_1(A)|p_1)(p_1^2-m^2) \tag{137}$$

$$= \lim_{p_1^2\to m^2} e\,\bar{u}(p_1')\int d^4y\ e^{-ip_1'\cdot y}\ \not{A}(y)\ (y|\mathcal{G}_1(A)|p_1)(p_1^2-m^2). \tag{138}$$

Now look at $(y|\mathcal{G}_1(A)|p_1)$ near $p_1^2 = m^2$. We first exponentiate the denominator in $\mathcal{G}_1(A)$

$$\mathcal{G}_1(A) = -i\int_0^\infty d\tau\, e^{i\,\tau[2p_1\cdot P-(m^2+p_1^2)\,-e(m+\not{p}_1)\not{A}(X)\,]}. \tag{139}$$

Next we use the following identity for operators A and B

$$e^{A+B} = \left(e^{\int_0^1 dt\, e^{At}B\, e^{-At}}\right)_- e^A \tag{140}$$

where $(\,)_-$ means anti ordering according to the parameter t. {This identity is proved by exactly the same procedure as one uses to establish the time ordered form for the U-matrix in scattering theory.} Identifying the operators $A = 2p_1\cdot P - (m^2 + p_1^2)$ and $B = -e(m + \not{p}_1)\not{A}(X)$ we find directly from (139) and (140)

$$(y\,|\mathcal{G}_1(A)\,|p_1) = -i\int_0^\infty d\tau\, e^{-i\,e(m+\not{p}_1)\int_0^\tau dt\,\not{A}(y-2p_1t)}$$

$$\times\, e^{-i\,\tau(m^2-p_1^2)}\, e^{i\,p_1\cdot y}\, u(p_1) \tag{141}$$

yielding a T matrix

$$\mathcal{T}_E(p_1\to p_1';A) = \int d^4y\, e^{i(p_1-p_1')\cdot y}\,\bar{u}(p_1)\, e\not{A}(y)\,\times$$

$$\times\exp\left[-ie\int_0^\infty dt\,(m+\not{p}_1)\,\not{A}(y-2p_1t)\right]u(p_1). \tag{142}$$

Upon noting that for any vector V the quantity $\exp[(m+\not{p})\not{V}]\,u(p)$ simplifies by use of the Dirac equation $(m-\not{p})u(p) = 0$ to $\exp[2p\cdot V]\,u(p)$, we may cast (142) into

$$\mathcal{T}_E(p_1\to p_1';A) = \int d^4y\, e^{i(p_1-p_1')\cdot y}\,\bar{u}(p_1')\, e\not{A}(y)\, u(p_1)$$

$$\times\exp\left(-2i\, e\int_0^\infty dt\, p_1\cdot A(y-2p_1t)\right) \tag{143}$$

our result for the eikonal T matrix expanded about p_1.

If we wanted the full eikonal T-matrix including

radiative corrections along the electron line (but, of course, without vacuum polarization effects), our instructions are to operate on $\mathcal{J}_E$ with

$$T_E(A) = \exp\left(-\tfrac{1}{2}\int d^4w\, d^4z\, \frac{\delta}{\delta A_\alpha(w)}\, D_+(w-z)\, \frac{\delta}{\delta A_\alpha(z)}\right)\mathcal{J}_E(A), \tag{144}$$

which action inserts photons at all points z propagates them to w and then absorbs them. Even without computation, we can see that something is terribly amiss in our formula. Note that the spin structure of the amplitude $\mathcal{J}_E(A)$ is such that only the convective part of the electron current $\bar{u}(p')\gamma_\alpha u(p)$ enters, the possible magnetic piece $\bar{u}(p')\sigma_{\alpha\beta}(p'-p)^\beta u(p)$ is absent. The photon insertion operator will not change this fact, yet we well know that even the lowest order radiative correction to the Born approximation (neglecting the modulation factor in (143)) yields a magnetic moment term with the famous $\alpha/2\pi$ as its strength at zero momentum transferred to the potential. You may see where we have lost this contribution by examining the appropriate lowest order Feynman graphs and making on them the high energy approximation we have been discussing. What you will discover is that you have not only thrown away the magnetic moment term but also have arranged that the convective term is no longer renomalizable by the standard (physically acceptable) procedure of subtracting the value of the graph at zero momentum transfer.

The reason we have encountered this difficulty is that we took very seriously the external potential formalism as providing a generating functional for photon loop insertions connecting electron lines not moving rapidly with respect to one another. That is, until we asked that (143) generate T_E we were in no trouble. The proper manner in which to handle radiative corrections to the scattering of a fast electron in an external potential has been given by Bjorken, et al[8] who treat photon insertions into electron lines and photon propagations <u>exactly</u> while using the eikonal on the interaction with an external potential. As they point out, one finds in this way the correct anomalous moment and charge form factors arising from photonic vertex corrections, and one also exhibits the unfortunately very complex structure of the radiative corrections to

the rather simple (physically intuitive) form for $\mathcal{J}_E(A)$ given above. Our lesson in this is to use $\mathcal{J}_E(A)$ as a generating functional only for connecting the external potential to our fast electron or for connecting another electron to the first. Be very wary of photons connected to the same electron line in the eikonal approximation. We will resurrect this warning when we discuss the form factor in eikonal electrodynamics.

With this caveat we may now examine the eikonal approximation to the scattering of two electrons. We can evaluate this in the absence of vacuum polarization corrections by letting each electron move in its own external potential and then picking up all the connections between the individual electrons. So let electron number one go from momentum p_1, helicity λ_1 to p_1', λ_1' in the potential A_1 while number two goes from $p_2, \lambda_2 \to p_2', \lambda_2'$ in A_α. The scattering matrix for this with no self-action or interaction is

$$\mathcal{J}_{A_1A_2}(p_1,\lambda_1+p_2,\lambda_2 \to p_1',\lambda_1'+p_2',\lambda_2') = \mathcal{J}(p_1,\lambda_1 \to p_1',\lambda_1';A_1) \times$$
$$\mathcal{J}(p_2,\lambda_2 \to p_2',\lambda_2';A_2). \qquad (145)$$

We now need an operation to insert photons at all external potentials and connect these insertions by photon propagations. That operator is, of course,

$$\exp -\frac{i}{2}\int d^4w\, d^4z \left(\frac{\delta}{\delta A_{1\alpha}(w)} + \frac{\delta}{\delta A_{2\alpha}(w)}\right) \times$$
$$\times D_+(w-z)\left(\frac{\delta}{\delta A_{1\alpha}(z)} + \frac{\delta}{\delta A_{2\alpha}(z)}\right). \qquad (146)$$

We have received our instructions to discard the self action terms in (146) or treat the self photons more exactly so we take only the A_1, A_2 cross terms and find for the full T-matrix without vacuum polarization and self action

$$T(p_1,\lambda_1 + p_2,\lambda_2 \to p_1',\lambda_1' + p_2',\lambda_2') =$$
$$\exp\left(-i\int d^4w\, d^4z \frac{\delta}{\delta A_{1\beta}(w)} D_+(w-z) \frac{\delta}{\delta A_{2\beta}(z)}\right) \times \qquad (147)$$

$$\mathcal{T}(p_1\lambda_1 \to p_1'\lambda_1';A_1)\ \mathcal{T}(p_2,\lambda_2 \to p_2',\lambda_2';A_2)\Big|_{A_1 = A_2 = 0}. \tag{147}$$

We have built an eikonal approximation to $\mathcal{T}$ by expanding P about a convenient fast momentum. Before we took the indident or final momentum as the point of expansion; now let us choose the average of the two for symmetry and ease. So we expand $\mathcal{T}(A_1)$ about $Q_1 = (p_1+p_1')/2$ and $\mathcal{T}(A_2)$ about $Q_2(p_2+ p_2')/2$. If we further note that

$$\bar{u}(p_1',\lambda_1')\ \gamma_\alpha\ u(p_1,\lambda_1) \approx \frac{Q_{1\alpha}}{m}\ \delta_{\lambda_1',\lambda_1} \tag{148}$$

to leading order in large momenta, we may write for the eikonal approximation to $\mathcal{T}(A_1)$

$$\mathcal{T}_E(p_1,\lambda_1 \to p_1',\lambda_1';A_1) =$$

$$\frac{\delta_{\lambda_1'\lambda_1}}{2m}\int d^4y\ e^{i(p_1-p_1')\cdot y}\ e_1 2Q_1\cdot A_1(y) e^{-ie_1\int_0^\infty dt 2Q_1\cdot A_1(y-2Q_1t)} \tag{149}$$

$$=\frac{\delta_{\lambda_1'\lambda_1}}{2m}\int d^4y_1\ e^{i(p_1-p_1')\cdot y_1}\frac{1}{i}\frac{d}{d\sigma_1}\left(e^{-ie_1\int_{\sigma_1}^\infty dt_1 2Q_1\cdot A_1(y_1-2Q_1t_1)}\right)_{\sigma=0} \tag{150}$$

where we have let the charge on electron one be e_1.

To evaluate T_E we observe that the connecting operator in (147) is just a displacement operator which we apply twice to yield

$$T_E(p_1\lambda_1+p_2\lambda_2 \to p_1'\lambda_1'+p_2'\lambda_2') = \frac{\delta_{\lambda_1'\lambda_1}\delta_{\lambda_2'\lambda_2}}{4m^2}\ \times \tag{151}$$

$$\times\int d^4y_1 d^4y_2\ e^{i(p_1-p_1')\cdot y_1+i(p_2-p_2')\cdot y_2}\ \frac{1}{i^2}\frac{d}{d\sigma_1}\frac{d}{d\sigma_2}\ \times$$

$$\times\Big(\exp(-i)^3 4e_1e_2Q_1\cdot Q_2\int_{\sigma_1}^{\infty}dt_1\int_{\sigma_2}^{\infty}dt_2 D_+(y_1-y_2-2Q_1t_1+2Q_2t_2)\Big)_{\sigma_1=\sigma_2=0}. \quad (151)$$

Since we have turned off the external potential we expect to find in (150) a delta function of energy momentum conservation. To exhibit this, change variables to $W=(y_1+y_2)/2$, $w=(y_1-y_2)$ and do the d^4W integration. This yields a $i(2\pi)^4\delta^4(p_1+p_2-p_1'-p_2')$ which we remove from T_E without taking the trouble to define a new symbol at this point. This leaves us with the on-shell T-matrix

$$\frac{+i\,\delta_{\lambda_1'\lambda_1}\delta_{\lambda_2'\lambda_2}}{4m^2}\int d^4w\, e^{i(p_1-p_1')\cdot w}\frac{d}{d\sigma_1}\frac{d}{d\sigma_2}\Big(\exp-i\int_{\sigma_1}^{\infty}dt_1\int_{\sigma_2}^{\infty}dt_2 \times$$

$$4e_1e_2Q_1\cdot Q_2 D_+(w-2Q_1t_1+2Q_2t_2)\Big)_{\sigma_1=\sigma_2=0}. \quad (152)$$

We will evaluate this in the center of momentum frame in which the spatial parts of Q_1 and Q_2 are equal and opposite

$$\vec{Q}_1+\vec{Q}_2=0. \quad (153)$$

Decompose the integration variable w into a two vector $\underset{\sim}{b}$ orthogonal to $\vec{Q}_1$ and $\vec{Q}_2$ and pieces along the Q's

$$w=\underset{\sim}{b}+\tau_1Q_1+\tau_2Q_2. \quad (154)$$

The Jacobian involved in this change is

$$(Q_{1o}+Q_{2o})\;|\vec{Q}_1|, \quad (155)$$

so the T-matrix reads

$$i(Q_{1o}+Q_{2o})\,|\vec{Q}_1|\,\frac{\delta_{\lambda_1'\lambda_1}\delta_{\lambda_2'\lambda_2}}{4m^2}\int_{-\infty}^{+\infty}d\tau_1\int_{-\infty}^{+\infty}d\tau_2\int d^2b\; e^{+i\underset{\sim}{A}\cdot\underset{\sim}{b}}\times$$

$$\times\frac{d}{d\sigma_1}\frac{d}{d\sigma_2}\Big(\exp-ie_1e_2Q_1\cdot Q_2\int_{-\tau_1+2\sigma_1}^{\infty}d\beta_1\int_{\tau_2+2\sigma_2}^{\infty}d\beta_2\, D_+(\underset{\sim}{b}-\beta_1Q_1+\beta_2Q_2)\Big)_{\sigma_1=\sigma_2=0} \quad (156)$$

$$= +i(Q_{1o}+Q_{2o})\,|\vec{Q}_1|\,\delta_{\lambda_1'\lambda_1}\,\delta_{\lambda_2'\lambda_2}\int d^2b\, e^{i\underset{\sim}{\Delta}\cdot\underset{\sim}{b}}\Big[\exp\Big(-ie_1e_2Q_1\cdot Q_2\ \times$$

$$\int_{-\infty}^{+\infty}d\beta_1\int_{-\infty}^{+\infty}d\beta_2\ D_+(\underset{\sim}{b}-\beta_1Q_1+\beta_2Q_2)\Big)\ -\ 1\Big]. \tag{157}$$

The component of the momentum transfer $p_1'-p_1$ along the direction $\underset{\sim}{b}$ has been denoted by $\underset{\sim}{\Delta}$. The integrals over β_1 and β_2 may also be performed using the representation of $D_+(x)$ in (112):

$$\int_{-\infty}^{+\infty}d\mu_1\int_{-\infty}^{+\infty}d\mu_2 D_+(\underset{\sim}{b}-\mu_1Q_1+\mu_2Q_2) =$$

$$\int_{-\infty}^{+\infty}d\mu_1\int_{-\infty}^{+\infty}d\mu_2\int\frac{d^4q}{(2\pi)^4}\,\frac{e^{-iq\cdot(\underset{\sim}{b}-\mu_1Q_1+\mu_2Q_2)}}{\mu^2-q^2-i\epsilon} = \tag{158}$$

$$= \int\frac{d^4q\ \delta(q\cdot Q_1)\ \delta(q\cdot Q_2)}{(2\pi)^2}\ \frac{e^{-iq\cdot\underset{\sim}{b}}}{\mu^2-q^2-i\epsilon} \tag{159}$$

$$= \frac{1}{|\vec{Q}_1|(Q_{1o}+Q_{2o})}\int\frac{d^2q}{(2\pi)^2}\,\frac{e^{i\underset{\sim}{q}\cdot\underset{\sim}{b}}}{\mu^2+|\underset{\sim}{q}|^2} \tag{160}$$

$$= \frac{1}{2\pi|\vec{Q}_1|(Q_{1o}+Q_{2o})}\ K_o(\mu b)\ , \tag{161}$$

which, by now, should be no surprise.

We may simplify this a bit more by introducing the invariants

$$s = (p_1+p_2)^2 \text{ and } t = (p_1-p_1')^2. \tag{162}$$

To leading order in s for t finite the eikonal T-matrix reads[9]

$$T_E(s,t) = \frac{is}{2m^2}\ \delta_{\lambda_1'\lambda_1}\,\delta_{\lambda_2'\lambda_2}\ \times \tag{163}$$

$$\times \int d^2b\ e^{i\underset{\sim}{\Delta}\cdot\underset{\sim}{b}} \left[\exp - \left(i\,\frac{e_1 e_2}{2\pi}\ K_0(\mu b)\right) - 1\right], \qquad (163)$$

where $t = -\ |\underset{\sim}{\Delta}|^2$.

The characteristic and significant features of this result are (1) no spin flip on the electron lines occurs to this order in the eikonal expansion; (2) the T-matrix is of the form i s F(t) which is standard for the spin orbit form of interaction involved in electrodynamics and which leads to constant total cross sections. (3) This form cannot be correct for a very large range of t since it violates t-channel unitarity.[10]

It is very meaningful to inquire as to what set of Feynman graphs is summed by the eikonal T-matrix (163). The answer is not terribly difficult to come by since in arriving at that answer we explicitly eliminated vacuum polarization graphs, that is graphs where an electron position pair forms a closed loop which connects to the outside world by photons only and we discarded radiative corrections to electron lines. That leaves us with the generalized ladder graphs shown in Figure 1. It is amusing to note by the way that the asymptotic behavior of the first graph is $se^2F_2(t)$ for large s and fixed t. For the second graph it is $(s \log s)e^4F_4(t)$ which seems inconsistent with the expansion of T_E in powers of e. However, for the third graph the asymptotic behavior is $(-s \log(-s))e^4F_4(t)$ with the same $F_4(t)$ as before. These fourth order graphs then combine to yield $-\ i\ s\ \pi e^4F_4(t)$. So there is a neat cancellation which occurs. In order to demonstrate the equivalence of the expansion of T_E term by term with the Feynman graphs it is still necessary to show that the correct functions of t appear. This can be done[9] with a certain amount of labor. That it must work is physically pretty clear.

Having made this connection with Feynman graphs one is immediately led to ask whether one can find other graphs, neglected by our procedure, which have an asymptotic behavior which dominates over i s F(t). There are such graphs. They correspond to "towers" of exchanges in the t-channel. Figure 2 is the prototype of such graphs in which photons shaken off the electron lines interact

via a photon-photon scattering loop. Such a diagram behaves for large s, fixed t as s (logs) f(t) a contribution which is not cancelled by other graphs of the same order in e^2.[11] Adding more loops gives more logarithms - one extra per loop. In a manner familiar from the discussion of the generation of Regge behavior in field theories, the sum of these leading logarithms builds up to a power.

Again from our experience with Reggeology we can forsee the disaster that is approaching from the power behavior due to the sum. When one adds up the contributions of graphs like Fig. 2 at t=0, the positive definite contributions of these various contributions to what is effectively the unitarity relation for elastic scattering will take the original power s^1, the "1" of which we were so proud before, and push the power to $1 + \epsilon$, $\epsilon > 0$. For the particular set of graphs in Fig. 2, $\epsilon = 11\alpha\pi/32$; a number with no dramatic significance in itself. What is significant is that the power is greater than one, and this set of graphs considered leads to a T-matrix which violates the unitarity bound.[12] One may consider curing this by iterating these "towers" down the s-channel arguing that they eikonalize and, following closely the argument given from Eqs. (104) - (109), expect the resultant amplitude to saturate the unitarity bound $s(\log s)^2$. This is what appears to occur, but I will leave it for Professor Sugar to defend the point in detail.

The final eikonalistic topic I would like to treat by the methods we have been discussing is that of the electromagnetic form factor or vertex function. Actually this is such an ancient topic[11] one, perhaps, ought to be embarrassed to raise the issue at this date. Others[12] having overcome this reluctance of late it seems more or less appropriate to include the matter here.

We are interested in the amplitude for an electron to go from momentum p_1 to p_2 emitting a photon of momentum q. Were there no self action the answer to the question would be

$$\Gamma^0_\beta(q) = - \int d^4y \, e^{iq\cdot y} \frac{\delta}{\delta A^\beta(y)} (p_2 | \mathcal{G}_A^{-1} | p_1) \Big|_{A=0} \qquad (164)$$

$$= \int d^4y \; e^{i(p_1-p_2-q)\cdot y} \; e \; \bar{u}(p^2) \; \gamma_\beta \; u(p_1) \tag{165}$$

$$= i(2\pi)^4 \; \delta^4(p_1-p_2-q) \; e \; \bar{u} \; (p_2) \; \gamma_\beta \; u(p_1). \tag{166}$$

We are, of course, interested in the coefficient of the delta function which will give us a measure of the dynamical response to an external source T_α.

When we include self action, but as is our procedure here leave aside vacuum polarization terms, the vertex function becomes

$$\Gamma_\beta(q) = - \frac{\delta}{\delta A^\beta(q)} \; (p_2 | G_A^{-1} | p_1) \Big|_{A=0}. \tag{167}$$

The general form for $\Gamma_\beta(q)$ is well known to be

$$\Gamma_\beta(q) = i(2\pi)^4 \delta^4(p_2+q-p_1) \; \bar{u}(p_2) \; \{\gamma_\beta F_1(q^2) + \sigma_{\beta\tau} \; q^\tau \; F_2(q^2)\} \; u(p_1), \tag{168}$$

and we are interested in an eikonal approximation to the $F_i(q^2)$.

To proceed we write (167) in matrix notation

$$\Gamma_\beta = - \frac{\delta G_A^{-1}}{\delta A^\beta} \Big| A = 0 \tag{169}$$

and using $G_A^{-1} G_A = 1$ note that

$$\Gamma_\beta = G_A^{-1} \; \frac{\delta G_A}{\delta A^\beta} \; G_A^{-1} \Big| A=0 \tag{170}$$

$$= G_A^{-1} \; K\left(\frac{\delta \mathcal{G}_A}{\delta A^\beta}\right) \; G_A^{-1} \Big| A=0, \tag{171}$$

where K is the operator exponential of functional derivatives which as in (125) converts $\mathcal{G}_A$ to G_A.

Now observe that

$$\frac{\delta}{\delta A^\beta}\, G_A - \frac{\delta}{\delta A^\beta}\left(\frac{1}{\not p - m - e\not A}\right) \tag{172}$$

$$= \frac{1}{\not p - m - e\not A}\, e\gamma_\beta\, \frac{1}{\not p - m - e\not A}\ . \tag{173}$$

What we wish to do is to expand the right hand G_A in (173) about the incoming momentum p_1 and the left hand G_A about p_2. For this purpose we employ our eikonal Green function as recorded in (141) and its counterpart for p_2. The resulting vertex function before the connecting operator is applied is given by

$$-\int d^4y \int_0^\infty d\tau_1 \int_0^\infty d\tau_2\ e^{-i\tau_2(m^2-p_2^2)} e^{-ip_2\cdot y}\, \bar u(p_2) e\gamma_\beta e^{-iq\cdot y}$$

$$\times \exp\left[-ie\int_0^{\tau_2} dt_2 \not A(y+2p_2t_2)(m+\not p_2) \ -ie\int_0^{\tau_1} dt_1 (m+\not p_1)\not A(y-2p_1t_1)\right]$$

$$\times\ e^{ip_1\cdot y}\, e^{-i\tau_1(m^2-p_1^2)} u(p_1). \tag{174}$$

The application of the functional operator to insert photons in all possible ways is aided by the elementary result

$$\exp\left[\int d^4w\, d^4z\, \frac{\delta}{\delta A^\alpha(z)} D_+(z-w) \frac{\delta}{\delta A_\alpha(w)}\right] \exp\int d^4x A^\alpha(x) B_\alpha(x)$$

$$= \exp\int d^4x A^\alpha(x)\ \exp\int d^4w\, d^4z\, B^\alpha(w)\, D_+(w-z)\, B_\alpha(z), \tag{175}$$

for any vector function $B_\alpha(x)$. If we use this, employ the G_A^{-1}'s to go to the mass shell and finally turn off the potential A_α, we may extract from (174) the eikonal vertex function

$$\Gamma_\beta^{\,E}(q) = i(2\pi)^4\delta^4(p_2+q-p_1)\ \bar u(p_2) e\gamma_\beta u(p_1) F_1^{\,E}(q^2), \tag{176}$$

where

$$F_1^{\,E}(q^2) = \exp\left\{-\frac{i\alpha}{8\pi^3}\int\frac{d^4k}{\mu^2-k^2-i\varepsilon}\left[\frac{p_1\lambda}{p_1\cdot k} - \frac{p_2\lambda}{p_2\cdot k}\right]\right\}^2 \tag{177}$$

and, as noted,

$$F_2^E(q^2) = 0. \tag{178}$$

The question one must ask about this result is, of course, where we might expect it to be valid. It is clearly not correct for all q^2 since we know quite well that in electrodynamics F_2 is not zero but at $q^2=0$, say, is the famous $\alpha/2\pi + 0(\alpha^2/\pi^2)$ terms. Since we have no F_2 form factor the only way we may reasonably interpret (177) and (178) is that in some limit we are finding F_2 negligible with respect to F_1. Since the only variable around is q^2, the only limit around is large q^2. The asymptotic behavior of $F_1^E(q^2)$ is exp - $(\log^2 q^2)$; an answer which is not inconsistent with the experimental results for large spacelike momentum transfers in elastic electron scattering.

Even in the large q^2 region one must worry seriously about the interpretation I have suggested. First, it is unusual, even for physicists to compare the magnitude of coefficients of distinct tensor characters. Yet there is no instruction in the calculation to indicate how else to interpret (178). Second, to my knowledge no one has investigated contributions like those in Figure 2 to elastic scattering to demonstrate that, at least order by order in e^2, they are not (or are) important. Their predominance in elastic scattering should make them of some concern. Even then, since those contributions alone violate unitarity in scattering processes, their real significance for the form factor has yet to be found. It cannot be said that the eikonal approximation to the vertex function is beyond all reproach. There are interesting questions yet to ask.

Figure 1.

The set of generalized Fenyman graphs summed by the eikonal T-matrix.

Figure 2.

The lowest order diagrams which contribute a large s, fixed t behavior which dominates the basic eikonal graphs of Figure 1. The whole set of these contributions leads to a behavior $s^{1+\epsilon}$, $\epsilon>0$, which violates unitarity.

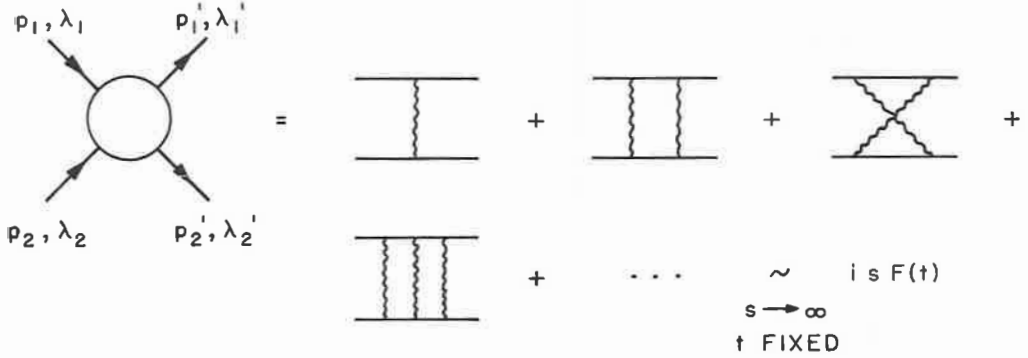


Figure 1

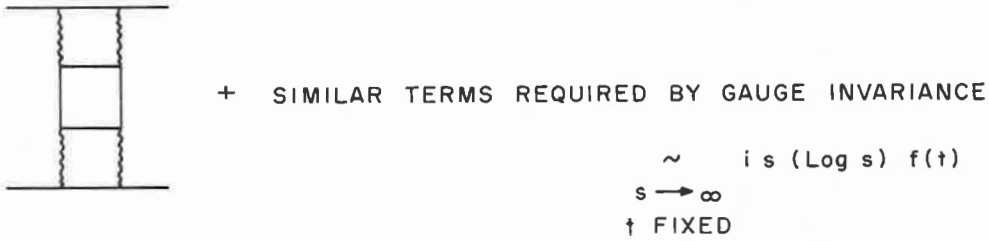


Figure 2

References

1. See Chapter 7 of J.D. Jackson, Classical Electrodynamics (John Wiley and Sons, Inc., New York, 1962).

2. R.J. Glauber in Lectures in Theoretical Physics, edited by W.E. Brittin and L.G. Dunham (Interscience Publishers, Inc., New York, (1959)), Volume I, p. 315.

3. R. B. Blankenbecler and R. Sugar, Phys. Rev. 183, 1387 (1969).

4. See the lectures by C. Itzykson and H.Abarbanel in the 1970 Cargese Summer School of Physics, to be published.

5. E.S. Fradkin has used this method for some years in his discussion of infrared approximations in field theory. His work can be traced from Nuc. Phys. 76, 588 (1966).

6. Their reprinted work in Quantum Electrodynamics (Dover Publications, Inc., N.Y. (1958), J.Schwinger, editor, are especially useful.

7. I use the notation and conventions of J.D. Bjorken an and S.D. Drell, Relativistic Quantum Fields, McGraw Hill Book Co., New York,(1964). Here $P^{\alpha}=i\nu^{\alpha}$.

8. J.D. Bjorken, et al, Phys. Rev. D3, 1382 (1971).

9. This result was derived in many of the lower orders of perturbation theory by H. Cheng and T.T. Wu, Phys. Rev. Letters 22, 666 (1969) and then to all orders by S.J. Chang and S.K. Ma, Phys. Rev. Letters 22, 1334 (1969) by summing Feynman graphs. The functional derivation was given by H.D.I. Abarbanel and C. Itzykson Phys. Rev. Letters 23, 53, (1969). The basic technique involved was exhibited by M. Levy and J. Sucher, Phys. Rev. 186, 1656, (1969) who derived the eikonal T-matrix in a φ^3 theory and more heuristically in QED.

10. V.N. Gribov, Nuc. Phys. 22, 249 (1961)

11. The ancient literature can be found in the book by N.N. Bogoliubov and D.V. Shirkov, Introduction to the Theory of Quantized Fields (Interscience Publishers, 1959). See Section 44.

12. Much of the modern work is referenced in a preprint by P.M. Fishbane and J.D. Sullivan, University of Illinois, May, 1971.

THE RELATIVISTIC EIKONAL MODEL*

Robert L. Sugar
Department of Physics, University of California
Santa Barbara, California 93106

I. Introduction

In his lectures Professor Abarbanel has shown us how the eikonal approximation can be used to discuss a wide range of scattering processes. I shall focus my attention on its application to the scattering of relativistic particles. The goal is to construct a simple model of ultra-high energy hadron interactions.

In non-relativistic potential scattering and in classical electrodynamics one can prove that the eikonal approximation is valid at high energies for a wide class of potentials or indicies of refraction. There is, of course, no such proof in quantum field theory. In fact the simple eikonal picture of high energy particles propagating through the interaction region in a straight line can not be the entire story in field theory, since it neglects the possibility of the high energy particles fragmenting. The best that one has been able to do so far is to study the high energy behavior of classes of Feynman diagrams to see when the eikonal approximation can be expected to hold, and, at least as important, when and why it breaks down.

I shall start by considering the scattering of two high energy particles which interact via the exchange of elementary quanta. These diagrams have already been discussed by Professor Abarbanel. I shall use them to orient our discussion and to introduce the moment space techniques which we shall need when we move on to more complicated diagrams. The elementary particle exchange diagrams are

*Work supported in part by the National Science Foundation.

not a satisfactory starting point for constructing a model of high energy hadron interactions since they predict that scattering becomes pure elastic at high energies. In order to introduce inelasticity into the model we shall be led to consider the exchange of compound objects such a $\lambda\varphi^3$ ladders and quantum electrodynamic (Q.E.D.) towers. These diagrams do lead to interesting models whose predictions I shall discuss in some detail. Finally I shall use the intuition we have gained from our study of Feynman graphs to construct a simple model of production amplitudes. These amplitudes will be used to build a model of the Pomerachuk singularity.

II. ELEMENTARY PARTICLE EXCHANGE

Let us start by considering the elastic scattering of two high energy particles which interact via the exchange of elementary quanta.[1,2,3] In addition to the ladder graph shown in figure 1a we must take into account all of the graphs that can be obtained from it by crossing the lines of the exchanged particles. A typical example is shown in figure 1b. There are n! distinct diagrams associated with the exchange of n quanta.

We are interested in scattering at very high energy for fixed, small values of the momentum transfer. It is convenient to work in the center of mass and to take the z axis along the incident direction of particle 1. We can then write (see figure 1)

$$P_1 \simeq \left(\frac{\sqrt{s}}{2} ; 0,0, \frac{\sqrt{s}}{2} \right)$$

$$P_2 \simeq \left(\frac{\sqrt{s}}{2} ; 0,0,- \frac{\sqrt{s}}{2} \right) \qquad (1)$$

$$\Delta \simeq (0;\Delta_x,\Delta_y,0) \equiv (0;\underset{\sim}{\Delta},0),$$

where $\sqrt{s}$ is the center of mass energy and $t = \Delta^2$ is the invariant momentum transfer. In writing Eq. (1) we have dropped terms of order $m^2/\sqrt{s}$ and $t/\sqrt{s}$. It will be convenient to write a general four-vector, q, in terms of the variables

$$q_{\pm} = q_0 \pm q_z$$
$$\underset{\sim}{q} = (q_x, q_y). \tag{2}$$

In particular $P_{1+} \simeq P_{2-} = \sqrt{s}$ and $P_{2+} \simeq P_{1-} \simeq \frac{m_1^2 + m_2^2}{2\sqrt{s}}$.

The results that we will obtain are essentially independent of the spins of the incident particles, so, for simplicity, I will take them both to be spinless. On the other hand the spins of the exchanged particles are crucial. Let us start with spin zero quanta. The amplitude for single particle exchange is given by

$$M_1 = \frac{\lambda^2}{\underset{\sim}{\Delta}^2 + \mu^2} \tag{3}$$

where μ is the mass of the exchanged particle. The two particle exchange diagrams are shown in figure 2. They give

$$M_2 = -i\lambda^4 \int \frac{d^4q}{(2\pi)^4} \left[q^2-\mu^2 + i\epsilon\right]^{-1} \left[(q-\Delta)^2 -\mu^2 + i\epsilon\right]^{-1} \tag{4}$$
$$\left[(P_2-q)^2-m_2^2+i\epsilon\right]^{-1} \left\{\left[(P_1+q)^2-m_1^2+i\epsilon\right]^{-1}+\left[(P_1+\Delta-q)^2-m_1^2+i\epsilon\right]^{-1}\right\}.$$

The basic idea of the eikonal approximation is that the incident particles propagate through the interaction region in a straight line retaining their large momentum even in the intermediate states. Let us start by assuming that this picture is correct. We shall discuss its validity at the end. Under this assumption the main contribution to the integral in Eq. (4) comes when the components of q are small compared to P_{1+} and P_{2-}, so we can write

$$[(P_1+q)^2 - m_1^2+i\epsilon]^{-1} + [(P_1+\Delta-q)^2 -m_1^2 + i\epsilon]^{-1}$$
$$= [q^2+2P_1\cdot q+i\epsilon]^{-1} + [(\Delta-q)^2 + 2P_1\cdot(\Delta-q) + i\epsilon]^{-1} \tag{5}$$
$$\simeq [\sqrt{s}\, q_- +i\epsilon]^{-1} + [-\sqrt{s}\, q_- + i\epsilon]^{-1} = -2\pi i\delta(\sqrt{s}q_-).$$

Notice that this is the same result that one would have

obtained by replacing the propagator $[(P_1+q)^2-m_1{}^2+i\epsilon]^{-1}$ by its mass shell delta function, since $\delta^+[(P_1+q)^2 - m_1{}^2] \simeq \delta(\sqrt{s}q_-)$.

Figure 2b is obtained from 2a by interchanging the order in which the exchanged quanta are absorbed by particle 1. If in each of these diagrams we interchange the order in which the quanta are emitted by particle 2, we obtain the diagrams of figure 3 which are, of course, identical to the original ones. Thus we can replace the propagator $[(P_2-q)^2 - m_2{}^2 + i\epsilon]^{-1}$ in Eq. (4) by the quantity

$$\tfrac{1}{2}\left\{[(P_1-q)^2 -m_2{}^2 +i\epsilon]^{-1} + [(P_2-\Delta+q)^2 - m_2{}^2 + i\epsilon]^{-1}\right\} \simeq \tfrac{1}{2} \cdot (-2\pi i)\ \delta(\sqrt{s}\, q_+). \tag{6}$$

Again, aside from the factor of $\frac{1}{2}$, this is equivalent to replacing the propagator $[(P_2-q)^2 - m_2{}^2 +i\epsilon]^{-1}$ in figure 2a by its mass shell delta function.

Eq. (4) now becomes

$$M_2 = \frac{1}{2!}\frac{i}{2s}\int \frac{d^2\underset{\sim}{q}}{(2\pi)^2}\ \frac{\lambda^2}{\underset{\sim}{q}^2+\mu^2}\cdot\frac{\lambda^2}{(\underset{\sim}{q}-\underset{\sim}{\Delta})^2+\mu^2}, \tag{7}$$

where we have used the fact that $d^4q = \frac{1}{2}\, dq_+dq_-d^2\underset{\sim}{q}$. Observe that the momenta carried by the exchanged quanta is space-like, lying in the x-y plane. It should also be noticed that an important cancellation has occurred between the two Feynman diagrams of figure 2. It is well known that the box graph of 2a has the asymptotic form[4]

$$M_{2a} \simeq f(t)\ \ell n(-s)/s \tag{8}$$

whereas the crossed box graph of 2b has the form

$$M_{2b} \simeq f(t)\ \ell n\ (-u)/u. \tag{9}$$

Since for large s and fixed t, $u \to -s$ Eq. (7) is recovered and the functional form of f(t) can be read off.

The rather trivial calculation which we have just done contains most of the ideas which we will need for the rest of our work. We must now extend it to the n particle exchange graphs illustrated in figure 1. Again if the incident particles retain their large momenta, their propagators can be linearized by dropping quadratic terms in the loop momenta relative to terms of the form $p_i \cdot q$. For example, in the diagram of figure 1a the propagators of particle 1 become

$$f(q_1 \ldots q_{n-1}) \simeq \frac{1}{2P_1 \cdot q_1 + i\epsilon} \cdot \frac{1}{2P_1 \cdot (q_1+q_2)+i\epsilon} \cdots \frac{1}{2P_1 \cdot (q_1+q_2+\ldots q_{n-1})+i\epsilon}$$

$$\simeq s^{-(n-1)/2} \frac{1}{q_{1-}+i\epsilon} \cdot \frac{1}{q_{1-}+q_{2-}+i\epsilon} \cdots \frac{1}{q_{1-}+\ldots q_{n-1-}+i\epsilon} \cdot \tag{10}$$

There are $n!$ distinct terms corresponding to the different orders in which particle 1 can absorb the exchanged quanta. After summing over all these terms we can obtain a result analogous to Eq. (5). To see this it is convenient to treat all the exchanged particles on an equal footing so we introduce the dummy variable q_{n-} by writing

$$1 = \int dq_{n-} \; \delta\Big(\sum_1^n q_i\Big)$$

and define

$$F(q_1 \ldots q_n) = \frac{1}{q_{1-}+i\epsilon} \cdots \frac{1}{q_{1-}+\ldots q_{n-1-}+i\epsilon} \; \delta\Big(\sum_1^n q_{i-}\Big) \tag{11}$$

so that

$$f(q_1 \ldots q_{n-1}) = \int dq_{n-} \; F(q_1 \ldots q_n). \tag{12}$$

For a diagram in which particle 1 absorbs the exchanged quanta in the order $v_1, v_2, \ldots v_n$, $F(q_1 \ldots q_n)$ is replaced in Eqs. (11) and (12) by $F(q_{v_1} \ldots q_{v_n})$.

We now prove the rather remarkable identity

$$\sum_P F(q_{v_1} \ldots q_{v_n}) = \prod_{i=1}^n \delta(q_{i-}) \; (-2\pi i)^{n-1} \tag{13}$$

where $\sum_P$ indicates the sum over all permutations of $v_1 \ldots v_n$. The fourier transform of F is given by

$$\tilde{F}(\tau_{v_1},\ldots\tau_{v_n}) = \int \prod_{i=1}^{n} dq_{i-}\, e^{i\Sigma\tau_i q_{i-}}\, F(q_{v_1}\ldots q_{v_n}). \tag{14}$$

Writing

$$\begin{aligned} k_1 &= q_{v_1-} \\ k_i &= k_{i-1} + q_{v_i-} \end{aligned} \tag{15}$$

and using Eq. (11) we find

$$\begin{aligned} \tilde{F}(\tau_{v_1},\ldots\tau_{v_n}) &= \int \prod_{i=1}^{n-1} \frac{dk_i}{k_i+i\epsilon}\, e^{ik_1(\tau_{v_1} - \tau_{v_2})} \\ &\cdot e^{ik_2(\tau_{v_2} - \tau_{v_3})} \ldots e^{ik_{n-1}(\tau_{v_{n-1}} - \tau_{v_n})} \qquad (16) \\ &= (-2\pi i)^{n-1}\, \theta(\tau_{v_n} - \tau_{v_{n-1}})\ldots\theta(\tau_{v_2} - \tau_{v_1}). \end{aligned}$$

So

$$\sum_P \tilde{F}(\tau_{v_1}\ldots\tau_{v_n}) = (-2\pi i)^{n-1}, \tag{17}$$

and Eq. (15) is obtained immediately by taking the inverse fourier transform. Thus after summing over all orderings of absorption, the propagators of particle 1 can all be replaced by their mass shell delta functions, i.e. by the quantity

$$(-2\pi i)^{n-1} \prod_{i=1}^{n-1} \delta(\sqrt{s}\, q_{i-}). \tag{18}$$

Homework problem 1: Obtain Eq. (18) directly in momentum space. Hint: It is convenient to first prove the identity

$$\sum_{P(v_1 \ldots v_{n-1})} \frac{1}{q_{v_1}+i\epsilon} \cdot \frac{1}{q_{v_1}+q_{v_2}+i\epsilon} \cdots \frac{1}{q_{v_1}+\ldots q_{v_{n-1}}+i\epsilon}$$

$$= \prod_{i=1}^{n-1} \frac{1}{q_i+i\epsilon} \, .$$

Homework problem 2: In approximating the propagators it is sometimes convenient to drop terms of the form $q_i \cdot q_j$ but retain the q_i^2 terms. i.e. we write

$$\frac{1}{(P_1+\sum_1^{\ell} q_i)^2-m^2+i\epsilon} \cong \frac{1}{\sum_{i=1}^{\ell} q_i^2+2P_1 \cdot \sum_{i=1}^{\ell} q_i+i\epsilon} \, .$$

Show that in this approximation the propagators of particle 1 can be replaced by

$$\prod_{i=1}^{n-1} \left[\frac{1}{q_i^2+2P_1 \cdot q_i+i\epsilon} + \frac{1}{q_i^2-2P_1 \cdot q_i+i\epsilon} \right]$$

which of course reduces to Eq. (18) if we drop the q_i^2 terms.

A result similar to Eq. (18) clearly holds for the propagators of particle 2. However, if we sum over all orderings of emission of the exchanged quanta, we count each Feynman diagram n! times, so we must divide by this factor.

The eikonal approximation thus predicts that the total amplitude for the exchange of n quanta is

$$M_n = \frac{1}{n!} \left(\frac{i}{2s}\right)^{n-1} \int \prod_{i=1}^{n-1} \frac{\underset{\sim}{q}_i}{(2\pi)^2} \left[\prod_{i=1}^{n-1} \frac{\lambda^2}{\underset{\sim}{q}_i^2+\mu^2} \right] \cdot \frac{\lambda^2}{(\sum \underset{\sim}{q}_i - \underset{\sim}{\Delta})^2+\mu^2} \qquad (19)$$

Defining

$$\delta_0(\underset{\sim}{b}) = \frac{1}{2s} \int \frac{d^2q}{(2\pi)^2} e^{-i\underset{\sim}{q}\cdot\underset{\sim}{b}} \frac{\lambda^2}{\underset{\sim}{q}^2+\mu^2}, \qquad (20)$$

Eq. (19) becomes

$$M_n = \frac{-2is}{n!} \int d^2\underset{\sim}{b}\, e^{i\underset{\sim}{\Delta}\cdot\underset{\sim}{b}} \left[i\delta_0(\underset{\sim}{b})\right]^n \tag{21}$$

or

$$M(s,\underset{\sim}{\Delta}) = \sum_1^\infty M_n = 2is \int d^2\underset{\sim}{b}\, e^{i\underset{\sim}{\Delta}\cdot\underset{\sim}{b}} (1-e^{i\delta_0(\underset{\sim}{b})}) \tag{22}$$

which is the familiar eikonal result.

The above calculation can easily be repeated for the case of the massive photon exchange. The vertex shown in figure 4, which contributes a factor λ for the case of scalar exchange, takes the form $\lambda(2P_{1\mu}+q_\mu)$ for vector exchange. Since we are assuming that the components of the loop momenta are small compared to those of P_1 we can drop the q_μ term. Then all vertices along the world line of particle 1 become $\lambda 2P_{1\mu}$. Similarly all those along the world line of particle 2 become $\lambda 2P_{2\nu}$. The only other change is that each propagator of an exchanged particle is now multiplied by $g_{\mu\nu}$. Thus the calculation goes through unchanged and we again obtain Eq. (22) with δ_0 replaced by

$$\delta_1(\underset{\sim}{b}) = 4P_1\cdot P_2\,\delta_0 \cong 2s\,\delta_0(\underset{\sim}{b}), \tag{23}$$

Again the eikonal phase is given by the two dimensional fourier transform of the Born approximation divided by 2s.

Homework problem 3: Consider the scattering of two spin $\frac{1}{2}$ particles which exchange massive photons. Take the interaction to be $\lambda A_\mu \bar{\psi}_i \gamma_\mu \psi_i$. Show that the eikonal amplitude is given by

$$M = i\,\delta_{\lambda_1\lambda_1'}\,\delta_{\lambda_2\lambda_2'}\,\frac{s}{m_1 m_2} \int d^2b\,(1-e^{i\delta(\underset{\sim}{b})})\,e^{i\underset{\sim}{\Delta}\cdot\underset{\sim}{b}}$$

where

$$\delta(\underset{\sim}{b}) = -\lambda \int \frac{d^2q}{(2\pi)^2}\, e^{i\underset{\sim}{q}\cdot\underset{\sim}{b}}\,(q^2+\mu^2)^{-1}$$

and λ_i and λ_i' stand for the initial and final helicities of the spin $\frac{1}{2}$ particles.

Hint: It is helpful to define $\gamma_\pm = \gamma_0 \pm \gamma_3$ and to notice that

$$\gamma_+^2 = \gamma_-^2 = 0$$

$$[\gamma_+, \gamma_-]_+ = 4.$$

Eq. (22) was obtained under the assumption that the eikonal approximation is valid. It has been shown by Tiktopoulos and Treiman that it is for the case of massive photon exchange.[5] In fact the correction terms are down from the leading ones by a full power of 1/s. However, the same authors have also shown that the eikonal approximation does not hold for the exchange of four or more scalar quanta.[5] The reason is not difficult to see. There is a well defined prescription for reading off the high energy behavior of Feynman diagrams made up solely of scalar particles.[4] The easiest terms to consider are those arising from so called end point contributions. One merely searches for the shortest path by which either incident particle can traverse the diagram. The number of distinct paths of minimum length determines the power of $\ell n s$. Thus if the minimum path has length ℓ and can be achieved in m distinct ways, the asymptotic behavior of the diagram arising from end point contributions is $(1/s)^\ell (\ell n s)^{m-1}$. For example, for the diagram of figure 1a we obtain $\ell n s / s^{n-1}$. For diagrams involving the exchange of n quanta the eikonal paths are of length n-1. Unfortunately for $n \geq 4$ there are always diagrams of the type illustrated in figure 5 which have paths of length three. All such graphs have asymptotic behavior $(\frac{1}{s})^3$, so the eikonal paths do not give rise to the leading asymptotic behavior of these graphs. The numerator factors in the diagrams for spin 1 exchange prevent this catastrophe from happening there.

The fact that the eikonal approximation breaks down for scalar exchange is not terribly serious since the scattering amplitude just goes to the Born approximation for large s in any event. Eq. (22) even gives the scattering amplitude correctly to order $(\frac{1}{s})^2$ since there are no short circuit paths for the two and three particle exchange graphs. However, we are advised to be careful. Although the eikonal approximation is intuitively appealing, it is not always correct. It would appear to be necessary to check its validity carefully in each application.

Even though the eikonal approximation is correct for the case of spin 1 exchange, these diagrams are not a useful starting point for constructing a model of high energy hadron scattering since they predict that the scattering is pure elastic at high energies, i.e. that there is no particle production. Recall that after all the cancellations have taken place the propagators of the incident particles can all be replaced by mass shell delta functions. Furthermore, since the plus and minus components of all loop momenta vanish, the momenta of the exchanged particles is entirely space-like and lies in the x-y plane. Thus these particles are never near the mass shell, so there are no contributions from multi-particle intermediate states.

We can also see this point directly. In our normalization, the scattering amplitude, M, is related to the S-matrix by

$$\langle P_1' P_2' | S | P_1 P_2 \rangle = 1 + \frac{(2\pi)^4 \delta^4(P_1+P_2-P_1'-P_2') M(P_1P_2;P_1'P_2')}{[2P_{1o}2P_{2o}2P_{1o}'2P_{2o}']^{\frac{1}{2}}} \tag{24}$$

so the contribution of the two-particle intermediate states to the imaginary part of the forward scattering amplitude is

$$\begin{aligned} \text{Im } M^{(2)}(P_1P_2;P_1P_2) &= \tfrac{1}{2}(2\pi)^4 \int \frac{d^4P_1' d^4P_2'}{(2\pi)^2} |M(P_1P_2;P_1'P_2')|^2 \\ &\cdot \delta^+(P_1'^2-m_1^2)\delta^+(P_2'^2-m_2^2)\delta^4(P_1+P_2-P_1'-P_2') \\ &= \tfrac{1}{2}\frac{1}{(2\pi)^2}\int d^4q\, |M|^2 \delta^+[(P_1+q)^2-m_1^2]\delta^+[(P_2-q)^2-m_2^2] \\ &\simeq \frac{1}{4s}\int \frac{d^2\underset{\sim}{q}}{(2\pi)^2} |M(s,\underset{\sim}{q})|^2 \end{aligned} \tag{25}$$

In the last step we have made our usual high energy approximations. Denoting the eikonal amplitude given in Eq. (22) by $M_E(s,\underset{\sim}{\Delta})$ we see that

$$\text{Im } M_E(s,0) = 2s \int d^2\underset{\sim}{b}[1-\cos\delta(\underset{\sim}{b})] \tag{26}$$

and

$$\frac{1}{4s}\int\frac{d^2q}{(2\pi)^2}\,|M_E(s,\underset{\sim}{q})|^2 = s\int d^2b(1-e^{i\delta})\,(1-e^{-i\delta^*}). \qquad (27)$$

So if δ is real as it always is for elementary particle exchange, the eikonal amplitude satisfies two-particle unitarity exactly. Since the contribution to Im M(s,0) from each type of intermediate state is positive definite, all production amplitudes must be identically zero if δ is real.

III. LADDER EXCHANGE

We have seen that if the eikonal picture is correct, the incident particles will propagate through the interaction region in a straight line, staying on the mass-shell at all times. The quanta that are emitted and absorbed will have space-like momenta and therefore will never be near the mass shell. Thus, if we wish to introduce inelasticity into the model, it seems necessary to consider the exchange of compound objects. Since we are considering strong interactions, a natural choice is Regge pole exchange. If the "Born term" is taken to be the exchange of a single Reggeon, one would guess on the basis of our previous work that the full amplitude is given by

$$M_R(s,\underset{\sim}{\Delta}) = 2is\int d^2\underset{\sim}{b}\; e^{i\underset{\sim}{\Delta}\cdot\underset{\sim}{b}}\,(1 - e^{i\delta_R(\underset{\sim}{b})}), \qquad (28)$$

where

$$\delta_R(\underset{\sim}{b}) = \frac{1}{2s}\int\frac{d^2\underset{\sim}{q}}{(2\pi)^2}\; e^{-i\underset{\sim}{q}\cdot\underset{\sim}{b}}\,\gamma(-\underset{\sim}{q}^2)\; s^{\alpha(-\underset{\sim}{q}^2)} \qquad (29)$$

$\alpha(t)$ is the trajectory function and $\gamma(t)$ is the residue function multiplied by the signature factor. Equations (28) and (29) are usually referred to as the Regge-Eikonal Model (REM). This model was proposed several years ago on the basis of analogy with non-relativistic potential scattering, and it has been used rather successfully in fitting high energy scattering data.[6] If one expands the right-hand side of Eq. (28) in powers of δ_R, the first order term just gives the simple Regge pole amplitude. The higher order terms give rise to cuts in the angular momentum plane

associated with the exchange of two or more Reggeons. The REM provides a particularly appealing model of cuts since no new parameters enter once the trajectory and residue functions of the pole are known.

We are now in a position to investigate the extent to which the REM can be justified in field theory. It is clearly not sufficient to return to the diagrams of figures 1 and 2 and reinterpret the wavy lines as Reggeons. For example, it is well known that the diagrams of figure 2, which are essentially planar, do not give rise to Regge cuts.[4] Their asymptotic behavior is much weaker than that of the diagrams which do. In order to proceed it is necessary to have a specific model of the Regge poles. A simple one can be obtained by considering the ladder and crossed ladder graphs in $\lambda\phi^3$ theory.[7,8] These diagrams are shown in figure 6. The asymptotic behavior of the sum of these graphs has the well known form

$$\beta(t)\ (1 + e^{-i\pi\alpha(t)})\ s^{\alpha(t)} \qquad (30)$$

where $t = \Delta^2 = -\underset{\sim}{\Delta}^2$. The eikonal model suggests that we nowconsider the exchange of an arbitrary number of ladders, summing over all possible ways in which the legs of the ladders are attached to the world lines of the incident particles. Typical examples are shown in figures 7 and 8. Some of these diagrams do give rise to Regge cuts. Our task is to evaluate their high energy behavior and either verify Eqs. (28) and (29) or see where it breaks down.

Before embarking on this program I should mention that there is a second model based on the ideas we have been discussing which has attracted considerable interest. That is the model of Cheng and Wu.[9] In their stupendous study of high energy behavior in quantum electrodynamics, they were led to consider the Q.E.D. analogue of ladder graphs which they call towers. Examples are shown in figure 9. The heavy lines represent electrons and the wavy lines photons. Topologically the only difference between the towers and the $\lambda\phi^3$ ladders is that in Q.E.D. it is necessary to include crossed electron boxes (shown in figure 9b) in order to maintain gauge invariance. The asymptotic behavior of the sum of single tower graphs is found

to be

$$i\delta_{\lambda_1\lambda_1'}\,\delta_{\lambda_2\lambda_2'}\,F(t)\,(\ln s)^{-2}\,s^{1+\frac{11}{32\pi}\alpha^2} \tag{31}$$

where α is the fine structure constant. This amplitude corresponds to a fixed cut in the angular momentum plane in contrast to the ladder exchange amplitude which gives rise to a moving pole. Notice that the tower exchange amplitude by itself violates the Froissart bound since it goes to infinity faster than a single power of s. A similar difficulty would arise for the $\lambda\phi^3$ ladders if the coupling constant were large enough so that $\alpha(t)>1$ for $t\leq 0$. If, as expected, the scattering amplitude takes the form of Eq. (28) after multi-ladder or multi-tower exchanges are included, then there will be no violation of the Froissart bound. I will return to this point and to the very interesting predictions of the Cheng-Wu model after discussing the $\lambda\phi^3$ ladder graphs.

I have no intention of reproducing the rather lengthy calculations necessary to obtain the high energy behavior of the multi-ladder exchange diagrams; however, I would like to briefly review the results of this work.

For the simple ladder and crossed ladder graphs Eq. (30) is valid for all values of the coupling constant.[10] The correction terms are down by at least a fractional power of s. Unfortunately the multi-ladder graphs are so complicated that it is extremely difficult to do more than calculate the leading behavior in $\ln s$ at each order in the coupling constant. It is important to keep in mind the consequences of restricting ourselves to a leading log calculation. For ladder graphs the Regge trajectory and residue functions have the perturbation expansions

$$\begin{aligned}\alpha(t) &= -1 + \lambda^2\, a_2(t) + \lambda^4\, a_4(t) + \ldots \\ \beta(t) &= \lambda^2 + \lambda^4\, b_4(t) + \ldots\end{aligned} \tag{32}$$

so we can write

$$\begin{aligned}\beta(t)\, s^{\alpha(t)} &= [\lambda^2 + \lambda^4 b_4 + \ldots]\,\frac{1}{s}\, e^{\ln s[\lambda^2 a_2+\lambda^4 a_4+\ldots]} \\ &= [\lambda^2 + \lambda^4 b_4+\ldots]\,\frac{1}{s}\sum_{n=0}^{\infty}\frac{(\ln s)^n}{n!}\,[\lambda^2 a_2+\lambda^4 a_4+\ldots]^n\end{aligned} \tag{33}$$

Thus, if at each order in λ^2 we retain only the leading power in $\ell n s$, then in effect we are working to order λ^2 in the perturbation expansion of the residue and trajectory functions.

NowGribov has shown on rather general grounds that the amplitude arising from an n-Reggeon cut will have the asymptotic form[11]

$$\frac{1}{(2s)^{n-1}} \int \prod_{i=1}^{n} \frac{d^2 q_i}{(2\pi)^2} \gamma(-q_i^2) s^{\alpha(-q_i^2)} f_n(q_1, q_2, \ldots q_n) \cdot \delta\left(\sum_{i=1}^{n} q_i - \Delta\right). \tag{34}$$

Substituting Eq. (33) into Eq. (34) we see that in the case of Regge cuts, the leading log approximation is equivalent to working to lowest order in λ^2 for the trajectory and residue functions and for the function f_n. It is f_n, of course, which we wish to calculate. Eq. (28) predicts $f_n(q_1, \ldots q_n) \equiv \frac{1}{n!}$, but we can at best verify this to leading order in λ^2 if we are only able to do a leading log calculation.

Let us start by considering the two ladder exchange diagrams illustrated in figures 7 and 8. The situation here is completely different from the case of elementary particle exchange. In the latter case the two second order graphs had the same asymptotic behavior and a delicate cancellation occurred between them. The two-ladder graphs on the other hand have a wide variety of asymptotic behaviors. For example, the planar diagram shown in figure 7a has only the end point contributions which we have previously discussed. Its asymptotic behavior is $\ell n s / s^3$ independent of the number of rungs in the ladders. This type of diagram does not give rise to Regge cuts and is negligible compared to those that do. The diagram of figure 7b does give rise to a cut, but it too can be neglected in the weak coupling limit. In fact the only diagrams that contribute in the leading log calculation are the four shown in figure 8 and the twelve that can be obtained from them by replacing one or both of the ladders by a crossed ladder.[12,13]

The four diagrams of figure 8 can be redrawn in the more familiar Mandelstam form shown in figure 10. We notice at once that if we blindly follow the eikonal prescription we will over count each of the leading graphs by a factor of four. No such difficulty arises in Q.E.D.

Originally there was an even more alarming problem: the first calculation of the Mandelstam diagram (figure 10), which was widely quoted in the literature,[4] was in disagreement with the eikonal model. Fortunately, this calculation was incorrect. The asymptotic behavior of the Mandelstam graph plus the associated graphs with crossed ladders just gives the second order term in the expansion of Eq. (28).[12,13] Thus, the eikonal model does check to second order, at least in the weak coupling limit.

The eikonal picture can also be restored. There are four contributions to the asymptotic behavior of the Mandelstam graph. The large momentum P_1 can follow either the path ABCD or AEFD. Similarly P_2 can follow either A′B′C′D′ or A′E′F′D′. Returning to figure 8 we see that if we retain all four diagrams, but keep only the contributions arising when the large momenta stay on the eikonal paths (the straight lines in figure 8), then all over counting problems are eliminated and we recover the correct result for the Mandelstam graph. Furthermore, the asymptotic behavior of the non-leading diagrams also comes when the large momenta stay on the eikonal paths so these graphs could be added in for free. Thus a blind application of the eikonal identity would indeed give the correct result!

The calculation becomes much more complicated when more than two ladders are exchanged.[12] From Eq. (32) we see that the leading Regge pole goes to $\ell = -1$ in the weak coupling limit. Eq. (34) then tells us that the cut arising from exchanging this pole n times has its branch point at $\ell = -(2n-1)$ in this limit. Thus terms contributing to this cut will have large s behavior of the form $s^{-(2n-1)}$ x powers of ℓns.

In figure 11a we have drawn a typical three ladder exchange diagram. If the large momenta follow the eikonal paths AEFGHD and A′E′F′G′H′D′ we do indeed get

contributions which go like s^{-5} as expected for the 3-Reggeon cut. However, there are "short circuit" paths ABCD and A'B'C'D' which give contributions to the asymptotic behavior of order s^{-3}. These terms do not contribute to the 3-Reggeon cut, but they may well contribute to the two Reggeon cut. In figure 11b we have redrawn the diagram in the Mandelstam form. Notice that when the large momenta follow the paths ABC and A'B'C', the diagram corresponds to the exchange of a ladder (BCB'C') and a rather complicated object (EHE'H'). As the number of rungs in either ladder making up EHE'H' is increased, there is no corresponding increase in powers of $\ell n s$. So in the weak coupling limit, and only in the weak coupling limit, the contribution of the short circuit path can be neglected compared to the s^{-3} contribution of the two ladder graphs.

It is not difficult to see that all short circuit paths correspond to the exchange of at least one object more complicated than a ladder. Since at each order in the coupling constant, the leading power of $\ell n s$ comes from the exchange of a pure ladder, the contributions from the short circuit paths can be dropped in the weak coupling limit. On the other hand, the contributions from the eikonal paths of the n-ladder exchange graphs just give the n-th order term in the expansion of Eq. (28). It must be emphasized that we are not computing the leading asymptotic behavior of the n-ladder graphs. That comes from the short circuit paths and is of order s^{-3} for $n \geq 2$. What we are doing is to compute the leading contribution in the weak coupling limit to the cut which arises from exchanging the leading Regge pole n times. This contribution comes from the eikonal paths of the n ladder graphs.

The situation in Q.E.D. is much cleaner. There the contributions from the short circuit paths are smaller than those from the eikonal paths. The enhancement of the eikonal paths comes about because of the numerator factors associated with the spin of the photon. Furthermore, since the amplitude for n tower exchange has the asymptotic form $s^{1+n\frac{11}{32}\pi\alpha}$, it makes much more sense to keep this term in the weak coupling limit, than the corresponding term arising from n ladder exchange.

One can use the momentum space techniques which we have developed to extract the contribution to the n-ladder graphs arising when the incident momenta follow the eikonal paths. However, it is necessary to show that the leading asymptotic behavior of the ladder and crossed ladder graphs of figure 6 actually does arise when the incident particles retain their large momenta. This calculation is quite instructive because it shows explicitly why the simple eikonal model must break down away from the weak coupling limit.

Let us start by considering the ladder graph of figure 6a. The amplitude is given by

$$M_n = -\lambda^2\left[\frac{i\lambda^2}{(2\pi)^4}\right]^n \int \prod_{i=1}^{n} d^4k_i \left[k_i^2-m^2+i\epsilon\right]^{-1}\left[(k_i+\Delta)^2-m^2+i\epsilon\right]^{-1}$$
$$\cdot\left[(P_1-k_1)^2-m^2+i\epsilon\right]^{-1} \prod_{j=1}^{n-1}\left[(k_j-k_{j+1})^2-m^2+i\epsilon\right]^{-1} \qquad (35)$$
$$\cdot\left[(P_2+k_n)^2-m^2+i\epsilon\right]^{-1} .$$

We first perform all the k_{i-} integrations. We adopt the convention of closing all the k_{i-} contours in the upper half plane when $k_{i+} > 0$ and in the lower half plane when $k_{i+} < 0$. Since $\Delta_\pm \simeq 0$, we never pick up poles from propagators making up the sides of the ladders. It will be obvious in a minute that the pole arising from the deonominator $[(P_2+k_n)^2-m^2+i\epsilon]$ does not contribute to the leading behavior in ℓns, so we have

$$M_n(s,\underset{\sim}{\Delta}) \simeq -\lambda^2\left[\frac{\lambda^2}{4\pi}\right]^n \int \prod_{i=1}^{n} \frac{d^2k_{i\perp}}{(2\pi)^2} \int_0^{\sqrt{s}} \frac{dk_{1+}}{\sqrt{s}-k_{1+}} \int_0^{k_{1+}} \frac{dk_{2+}}{k_{1+}-k_{2+}}$$
$$\cdots \int_0^{k_{n-1}} \frac{dk_{n+}}{k_{n-1+}-k_{n+}} \cdot \prod_{i=1}^{n} \left[k_i^2-m^2+i\epsilon\right]^{-1}\left[(k_i+\Delta)^2-m^2+i\epsilon\right]^{-1}$$
$$\cdot \left[(P_2+k_n)^2-m^2+i\epsilon\right]^{-1} \qquad (36)$$

It is now convenient to make the change of variables

$$k_{i+} = \sqrt{s}\, y_1 \ldots y_i \qquad i=1,\ldots n. \tag{37}$$

Notice that

$$k_{1-} = \frac{m^2}{\sqrt{s}} - \frac{m^2+k_1^2}{\sqrt{s}(1-y_1)} \tag{38}$$

$$k_{i-} = k_{i-1-} - \frac{m^2+(k_{i-1}-k_i)^2}{\sqrt{s}\, y_1 \circ\circ y_{i-1}(1-y_i)}$$

so that

$$k_{i-} < 0 \tag{39}$$

and

$$k_{i+}\, k_{i-} \xrightarrow[y_i \to 0]{} 0.$$

Also

$$(P_2+k_n)^2 - m^2 = (\sqrt{s} + k_{n-})\left(\frac{m^2}{\sqrt{s}} + \sqrt{s} \prod_1^n y_i\right) - k_n^2 - m^2$$

$$\equiv [s\Pi y_i - \mathbb{m}^2] \tag{40}$$

$\mathbb{m}^2$ is a positive definite quantity whose weak dependence on the k_i and y_i can be neglected for our purposes. Eq. (36) now becomes

$$M_n(s,\Delta) = -\lambda^2\left[\frac{\lambda^2}{4\pi}\right]^n \int_0^1 \prod_{i=1}^n dy_i\, f(y_1,\circ\ldots y_n)[s\pi y_i - \mathbb{m}^2 + i\epsilon]^{-1} \tag{41}$$

where

$$f(y_1,\circ\ldots y_n) = \int \prod_{i=1}^n \frac{d^2k_i}{(2\pi)^2}\, [k_i^2-m^2+i\epsilon]^{-1}\,[(k_i+\Delta)^2-m^2+i\epsilon]^{-1}. \tag{42}$$

Having isolated all of the s dependence of M_n in a single denominator it is an easy matter to study the asymptotic behavior of $M_n(s,\underset{\sim}{\Delta})$ by means of the Mellin transform. We define

$$\tilde{M}_n(\ell,\underset{\sim}{\Delta}) \equiv \int_0^\infty (-s)^{-(\ell+1)} M_n(s,\underset{\sim}{\Delta})\, ds \tag{43}$$

$$\simeq \lambda^2\left[\frac{\lambda^2}{4\pi}\right]^n \int_0^1 \prod_{i=1}^n dy_i\, f(y_1 \ldots y_n) \left(\frac{1}{m^2}\right)^{\ell+1} \int_0^\infty \frac{dx\, x^{-(\ell+1)}}{[x\pi y_i + 1]}$$

$$\simeq \lambda^2\left[\frac{\lambda^2}{4\pi}\right]^n \left(\frac{1}{m^2}\right)^{\ell+1} \int_0^1 \prod_{i=1}^n dy_i\, y_i^{\ell}\, f(y_1 \ldots y_n) \frac{\pi}{\sin\pi(\ell+1)} .$$

Notice that the integral defining $\tilde{M}_n(\ell,\underset{\sim}{\Delta})$ only exists for $-1 < \mathrm{Re}\ \ell < 0$. Thus the inverse transform is given by

$$M_n(s,\underset{\sim}{\Delta}) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} d\ell (-s)^{\ell}\, \tilde{M}_n(\ell,\underset{\sim}{\Delta}) \tag{44}$$

with $-1 < c < 0$.

It is convenient to expand $f(y_1 \ldots y_n)$ in a Taylor series about the point $y_i = 0$.

$$f(y_1 \ldots y_n) = f(0,\ldots 0) + \sum_{i=1}^n y_i \frac{\partial f(0,\ldots 0)}{\partial y_i} + \ldots \tag{45}$$

so

$$M_n(s,\underset{\sim}{\Delta}) = \lambda^2\left[\frac{\lambda^2}{4\pi}\right]^n \frac{1}{2\pi i} \frac{1}{m^2} \int_{c-i\infty}^{c+i\infty} d\ell \left(-\frac{s}{m^2}\right)^{\ell} \frac{\pi}{\sin\pi(\ell+1)} \tag{46}$$

$$\cdot\left[f(0,\ldots 0)\left(\frac{1}{\ell+1}\right)^n + \sum_{i=1}^n \frac{\partial f(0,\ldots,0)}{\partial y_i}\left(\frac{1}{\ell+1}\right)^{n-1} + \ldots\right] .$$

For large s the leading behavior of $M_n(s,\underset{\sim}{\Delta})$ clearly comes from the pole at $\ell = -1$. We have

$$M_n(s,\underset{\sim}{\Delta}) \simeq \lambda^2\left[\frac{\lambda^2}{4\pi}\right]^n \left(\frac{1}{-s}\right) \left[f(0,\ldots 0) \frac{\ell n\left(\frac{-s}{m^2}\right)^n}{n!} + \right.$$

$$+ \sum_i \frac{\partial f(0,\ldots,0)}{\partial y_i} \frac{\ell n\left(\frac{-s}{m^2}\right)^{n-1}}{(n-1)!} + \ldots . \Big] \qquad (47)$$

From Eqs. (39) and (42) we see that

$$f(0,\ldots 0) = \left[\int \frac{d^2k}{(2\pi)^2} [(k^2+m^2)]^{-1} [(k+\Delta)^2+m^2]^{-1}\right]^n$$

$$\equiv [I(\Delta)]^n , \qquad (48)$$

so that the leading log approximation

$$M(s,\Delta) = \sum_{n=1}^{\infty} M_n(s,\Delta) = \lambda^2 \left(\frac{1}{-s}\right) \sum_{n=1}^{\infty} \left[\frac{\lambda^2}{4\pi} I(\Delta) \ell n(-s)\right]^n \frac{1}{n!}$$

$$\simeq \lambda^2 (-s)^{\alpha(\Delta)} \qquad (49)$$

with

$$\alpha(\Delta) = -1 + \frac{\lambda^2}{4\pi} I(\Delta)$$

in agreement with Eq. (32).

Notice that the leading log term comes entirely from the region of phase space in which all the y_i are infinitesimal. In particular we have

$$k_{1+} = \sqrt{s}\, y_1 << P_{1+} = \sqrt{s}. \qquad (50)$$

Furthermore, it is clear from Eq. (41) that the leading behavior comes from the region $\prod_{i=1}^{n} y_i \simeq m^2/s$, so

$$|k_{n-}| \simeq [(k_{n-1}-k_n)^2 + m^2]/\sqrt{s} \prod_{i=1}^{n-1} y_i(1-y_n)] \simeq y_n\sqrt{s} <<$$

$$<<P_{2+} = \sqrt{s}. \qquad (51)$$

Thus for this term the incident particles do retain their large momenta in the intermediate states.

Homework Problem 4:

Consider the n-ladder exchange diagrams in the weak coupling limit. Use the eikonal identity in the form suggested in problem 2 to extract the contribution arising when the large momenta stay on the eikonal paths. Show explicitly that your answer agrees with the n-th term in the expansion of Eq. (28).

Let us now consider the second leading log terms in Eq. (47). Here the incident particles do not necessarily retain their large momenta. For example, the term proportional to ∂f arises when y_1 is finite so it corresponds to the ∂y_1 fragmentation of the incident particle into two virtual ones which share the large incident momentum. This term makes a contribution of order λ^4 to the Regge residue function. Clearly terms corresponding to the fragmentation of the incident particles into large numbers of fast virtual particles give rise to lower logarithms and to higher order corrections to the residue function.

It is now clear that the simple eikonal picture can only be correct in the weak coupling limit because it ignores the possibility of fragmentation of the incident particles. The best that a dedicated eikonalist can hope for is that after fragmentation, each of the fast virtual particles moving in one direction will only exchange soft objects (ladders, towers, etc.) with those moving in the opposite direction. In this picture there would be one term in the scattering amplitude for each type of fragmentation. For example, a term in which particle 1 did not fragment, but particle 2 went into two fast virtual particles would be expected to have the form

$$M(s,\Delta) = 2is \int d^2\underset{\sim}{b}\, e^{i\underset{\sim}{\Delta}\cdot\underset{\sim}{b}} \int d^2\underset{\sim}{r} \int_0^1 dx\, |\psi(\underset{\sim}{r},x)|^2$$
$$\left(1 - e^{i\left[\delta(\underset{\sim}{b}+\frac{1}{2}\underset{\sim}{r};sx) + \delta(\underset{\sim}{b}-\frac{1}{2}\underset{\sim}{r};(1-x)s)\right]}\right) \qquad (52)$$

where $\psi(\underset{\sim}{r},x)$ is the probability amplitude for particle 2 to fragment into two virtuals with x-y separation $\underset{\sim}{r}$. x is the fraction of the large incident momentum retained by the incident particle.

This picture is based primarily on analogy with the Glauber formula for the high energy scattering of bound systems. To my knowledge there has been no attempt to derive an expression like Eq. (52) from field theory due to the difficulty in calculating non-leading log terms for the multi-ladder exchange diagrams.

IV, PREDICTIONS OF THE MODEL

Let us put aside questions concerning the fragmentation of the incident particles and ask what the eikonal model in its simplest form predicts. We shall be interested primarily in the energy dependence of the scattering amplitude which is determined by the energy dependence of the eikonal phase. As a result, most of our results will not be changed qualitatively by the addition of terms of the type illustrated in Eq. (52).

We begin with the Regge-eikonal model defined by Eq. (28) and (29). If the trajectory function is analytic at t=0 and if residue function does not vary too rapidly there, then at very high energies we can write

$$\begin{aligned}\delta_R(s,\underset{\sim}{b}) &= \frac{1}{2s}\int \frac{d^2\underset{\sim}{q}}{(2\pi)^2} e^{-i\underset{\sim}{q}\cdot\underset{\sim}{b}}\,\gamma(-\underset{\sim}{q}^2)s^{\alpha(-\underset{\sim}{q}^2)} \\ &\simeq \frac{\gamma(0)}{2s}\,s^{\alpha(0)}\int\frac{d^2\underset{\sim}{q}}{(2\pi)^2}\,e^{-i\underset{\sim}{q}\cdot\underset{\sim}{b}}\,e^{-\underset{\sim}{q}^2\alpha'(0)\ell ns} \qquad (53)\\ &= \frac{\gamma(0)}{8\pi\alpha'(0)\ell ns}\,s^{\alpha(0)-1}\,e^{-b^2/4\alpha'(0)\ell ns}.\end{aligned}$$

Since

$$\gamma(t) = -\beta(t)\left[e^{-i\pi\alpha(t)} \pm 1\right]/\sin\pi\alpha(t), \qquad (54)$$

$$\text{Im}\ \delta(s,\underset{\sim}{b}) > 0 \qquad (55)$$

as required by unitarity.

For an ordinary Regge trajectory $\alpha(0) \lesssim 1$, so we can expand the right hand side of Eq. (28) in a power series in δ_R.

$$M_R^n(s,\underset{\sim}{\Delta}) = -2is \int d^2b\, e^{i\underset{\sim}{\Delta}\cdot\underset{\sim}{b}} [i\delta_R(s,\underset{\sim}{b})]^n/n!$$

$$\simeq -2is \left[\frac{i\gamma(0)}{8\pi\alpha'(0)\ell ns}\right]^n s^{n(\alpha(0)-1)}$$

$$\cdot \frac{1}{n!} \int d^2b e^{i\underset{\sim}{\Delta}\cdot\underset{\sim}{b}} e^{-nb^2/4\alpha'(0)\ell ns} \qquad (56)$$

$$= \frac{\gamma(0)}{nn!} \left[\frac{i\gamma(0)}{8\pi\alpha'(0)\ell ns}\right]^{n-1} s^{n\alpha(0)-n+1} e^{-\Delta^2\alpha'(0)\ell ns/n}$$

$$\simeq \frac{\gamma(0)}{nn!} \left[\frac{i\gamma(0)}{8\pi\alpha'(0)\ell ns}\right]^{n-1} s^{n\alpha(\frac{-\Delta^2}{n^2})-(n-1)}$$

$M_R^1(s,\underset{\sim}{\Delta})$ is, of course, just the amplitude for the exchange of a single Regge pole, while $M_R^n(s,\underset{\sim}{\Delta})$, $n \geq 2$, corresponds to a cut in the angular momentum plane arising from the n-fold exchange of the Regge pole. The leading ℓ-plane singularity of M_R^n is most easily studied by means of the Mellin transform discussed in the last section. Recall that the Sommerfield-Watson representation of the scattering amplitude takes the form

$$M(s,t) = \frac{i}{2\pi} \int_c \frac{d\ell(2\ell+1)}{\sin\pi\ell} [P_\ell(-z_t) \pm P_\ell(z_t)] M_\ell(t), \qquad (57)$$

where the contour c runs from $-i\infty$ to $+i\infty$ staying to the right of all singularities of $M_\ell(t)$. Now $z_t = 1 + \frac{s}{2P_t^2}$, so

$$M(s,t) \xrightarrow[s\to\infty]{} \frac{1}{2\pi i} \int_c \frac{d\ell(2\ell+1)\,(e^{-i\pi\ell} \pm 1)}{\sin\pi\ell} M_\ell(t) \left(\frac{s}{2P_t}\right)^\ell$$

$$\cdot \Gamma(2\ell+1)/2^\ell\Gamma^2(\ell+1) \qquad (58)$$

$$\equiv \frac{1}{2\pi i} \int_c d\ell\, s^\ell f_\ell(t)$$

Thus the right most singularity of $f_\ell(t)$ and therefore of $M_\ell(t)$ can be found by taking the Mellin transform of $M(s,t)$

$$f_\ell(t) = \int_{s_0}^{\infty} ds\ s^{-(\ell+1)}\ M(s,t)\ . \tag{59}$$

Since the leading singularity in $f_\ell(t)$ depends only on the asymptotic behavior in s of $M(s,t)$, the value of s_0 is irrelevant for our purposes. Returning to Eq. (56) we can write

$$f_\ell^{\,n}(\underset{\sim}{\Delta}) = K_n \int_{s_0}^{\infty} ds\ s^{-(\ell+1)}\ s^{A_n(\underset{\sim}{\Delta})}/(\ell n s)^{n-1} \tag{60}$$

where

$$K_n = \frac{\gamma(0)}{nn!}\left[\frac{i\gamma(0)}{8\pi\alpha'(0)}\right]^{n-1} \tag{61}$$

and

$$A_n(\underset{\sim}{\Delta}) = n\alpha(-\underset{\sim}{\Delta}^2/n^2) - (n-1). \tag{62}$$

Making the change of variables $s = e^x$ gives

$$f_\ell^{\,n}(\underset{\sim}{\Delta}) = K_n \int_{x_0}^{\infty} dx\ e^{-x(\ell-A_n)}\ x^{1-n} \tag{63}$$

so

$$f_\ell^{\,1}(\underset{\sim}{\Delta}) = \frac{\gamma_0}{\ell-\alpha(-\underset{\sim}{\Delta}^2)} \tag{64}$$

$$f_\ell^{\,n}(\Delta) = k_n(\ell-A_n)^{n-2}\ \ell n(\ell-A_n) + R_n(\ell,\underset{\sim}{\Delta}), \tag{65}$$

where $n \geq 2$. $R_n(\ell,\underset{\sim}{\Delta})$ contains terms which either vanish more rapidly than $(\ell-A_n)^{n-2}\ \ell n(\ell-A_n)$ as $\ell \to A_n$ or which are analytic there. The position of the branch points given in Eq. (62) and the behavior of the amplitude at the branch point given in Eq. (65) are in agreement with the general results of Gribov and his co-workers.[11,14]

As I mentioned before the eikonal model of cuts has been used to fit the experimental data with reasonable success. Rather than discuss these fits I would like to move on to the rather exotic situation which arises when $\alpha(0)$ becomes larger than one. For $\alpha(0) < 1$, in the ℓ-plane there is a pole and an infinite sequence of branch points. Notice that $A_n(0) > A_{n-1}(0)$ in this case. As $\alpha(0)$ increases to 1 the pole and cuts all collide at $\ell=1$, $t=0$. The ℓ-plane structure of the amplitude changes dramatically if $\alpha(0)$ is increased further. For $\alpha(0) > 1$ it is convenient to rewrite Eq. (53) in the form

$$\delta_R(s,\underset{\sim}{b}) = \frac{\gamma(0)}{8\pi\alpha'(0)\ell n s} e^{\ell n s[\alpha(0)-1-b^2/4\alpha'(0)(\ell n s)^2]} \qquad (66)$$

Now if $b^2 < 4(\alpha(0)-1)\alpha'(0)$, $\delta_R(s,\underset{\sim}{b})$ grows like a power of s, while if $b^2 > 4(\alpha(0)-1)\alpha'(0)$, $\delta_R(s,\underset{\sim}{b})$ goes to zero like a power of s. Combining this information with Eq. (55) we see that

$$M_R(s,\underset{\sim}{b}) \equiv 2is\,(1-e^{i\delta_R(s,b)}) \simeq 2is\,\theta(b_o-b) \qquad (67)$$

where

$$b_o^2 = 4(\alpha(0)-1)\alpha'(0)\,(\ell n s)^2 \equiv (R_o\,\ell n s)^2 \,. \qquad (68)$$

Eq. (67) tell us that in impact parameter space all the scattering takes place inside a disc whose radius increases like $\ell n s$. Since $M_R(s,\underset{\sim}{b})$ is pure imaginary, the disc is perfectly absorbing, i.e. it is black. Eq. (67) should be contrasted with the amplitude for the exchange of a single Regge pole with $\alpha(0) \leq 1$. In our present approximation we have

$$M_{Regge}(s,\underset{\sim}{b}) \simeq \frac{\gamma(0)}{8\pi\alpha'(0)\ell n s}\, e^{-b^2/4\alpha'(0)\ell n s}, \qquad (69)$$

so virtually all of the scattering takes place inside a disc whose radius grows like $\sqrt{\ell n s}$. The disc is obviously not black in this case.

The first model involving a black disc with a logarithmically increasing radius was that of Cheng and Wu.[8] In their case the eikonal phase is given by

$$\delta_{cw}(s,b) = s^{\frac{11}{32}\pi\alpha}\,\tilde{f}(\underset{\sim}{b}) \qquad (70)$$

with $\tilde{f}(\underset{\sim}{b}) \to e^{-\mu b}$ for large b. As a result their scattering amplitude in impact parameter space has the same form as Eq. (67) with b_0 again proportional to $\ell n s$. Therefore, virtually all of the predictions of the Regge-eikonal model with $\alpha(0) > 1$ are identical to those of Cheng and Wu. However, in the Cheng-Wu model the quantity analogous to $\alpha(0)$, $(1 + \frac{11}{32}\pi\alpha)$, is predicted to be greater than one!

All of the predictions of the model can be read off from Eq. (67). The forward scattering amplitude is given by

$$M_R(s,\underset{\sim}{\Delta} = 0) = \int d^2b\; M_R(s,\underset{\sim}{b}) = 2\pi i s b_0^2 \; . \qquad (71)$$

The optical theorem tells us that

$$\sigma_{tot} \simeq \frac{\mathrm{Im}M_R(s,\underset{\sim}{\Delta} = 0)}{s} = 2\pi b_0^2 = 2\pi(R_0 \ell n s)^2, \qquad (72)$$

so the Froissart bound is saturated.

The elastic cross-section is given by

$$\sigma_{el} = \frac{1}{(2\pi)^2}\int d^4P_1'\, d^4P_2'\, \delta^4(P_1+P_2-P_1'-P_2')\,\delta^+(P_1'^2-m^2)\,\delta^+(P_2'^2-m^2)$$

$$\cdot \frac{|M_R(P_1P_2;P_1'P_2')|^2}{2P_{10}2P_{20}} \cdot \frac{1}{2P_1/P_{10}} \qquad (73)$$

$$\simeq \frac{1}{4s^2}\int \frac{d^2\underset{\sim}{\Delta}}{(2\pi)^2}\,|M_R(s,\underset{\sim}{\Delta})|^2 = \frac{1}{4s^2}\int d^2b\,|M_R(s,\underset{\sim}{b})|^2$$

$$= \pi(R_0 \ell n s)^2 .$$

The inelastic cross-section is equal to the elastic cross-section since

$$\sigma_{in} = \sigma_{tot} - \sigma_{el} = \pi(R_0 \ell n s)^2 . \qquad (74)$$

Notice that in the REM R_0 depends on $\alpha(0)$ and $\alpha'(0)$, but not on the Regge residue function γ. As a result, Eq. (72) predicts that all total cross-sections become equal at infinite energies. Similarly according to Eqs. (73) and (74) all elastic and inelastic cross-sections should become equal at infinite energies. These predictions are of course not in agreement with present experiments, nor is there any evidence that total cross-sections increase with energy. However, one can always argue that we have not yet reached asymptotic energies and that R_0 is small. Furthermore, if one includes effects associated with the fragmentation of the incident particles, the cross-sections for different processes will almost certainly become unequal, even though the energy dependence is likely to remain the same. The above remarks apply equally to the Cheng-Wu model.

Away from the forward direction the scattering amplitude is given by

$$M_R(s,\underset{\sim}{\Delta}) = 2is \int d^2b\, e^{i\underset{\sim}{\Delta}\cdot\underset{\sim}{b}}\, \theta(b_0-b)$$

$$= (2is)\,(2\pi) \int_0^\infty bdb\, J_0(b\Delta)\,\theta(b_0-b) \qquad (75)$$

$$= 2\pi is(R_0 \ell ns)^2 \left[J_1(R_0 \ell ns\, \sqrt{-t})/\tfrac{1}{2}R_0 \ell ns\sqrt{-t}\right]$$

where $\Delta = |\underset{\sim}{\Delta}| = \sqrt{-t}$. For large values of s the t dependent term in the square bracket is sharply peaked at $t = 0$. The width of the forward peak is given by

$$-t_0 = c/(R_0 \ell ns)^2 \qquad (76)$$

where c is a constant depending on the properties of J_1. The $1/(\ell ns)^2$ shrinkage of the width of the diffraction peak should be contrasted with the $1/\ell ns$ shrinkage predicted by simple Regge pole exchange. There is no evidence for the more rapid shrinkage in the Serpukhov data on proton-proton elastic scattering, and the preliminary ISR results indicate even a slower shrinkage than $1/\ell ns$. In my opinion the models that we are discussing are quite likely to founder on this point.

The ℓ-plane structure of the amplitude given in Eq. (75) can be studied by use of the Mellin transform,

$$f_\ell(t) = \int_{s_0}^{\infty} s^{-(\ell+1)} M_R(s,t)$$

$$= \int_{x_0}^{\infty} dx\, e^{-(\ell-1)x}\, x\, J_1(R_0\, x\sqrt{-t}\,) \left[4\pi i\, R_0/\sqrt{-t}\,\right]$$

$$= \int_{0}^{\infty} dx\, e^{-(\ell-1)x} x\, J_1(R_0\, x\,\sqrt{-t}\,) \left[4\pi i\, R_0/\sqrt{-t}\,\right] \qquad (77)$$

$$+ C(\ell,t)$$

$$= \frac{4\pi i\, R_0^2}{[(\ell-1)^2 - R_0^2 t]^{3/2}} + C(\ell,t)\ .$$

$C(\ell,t)$ is clearly an entire function of ℓ. The pole and infinite sequence of cuts present in the ℓ-plane for $\alpha(0) \lesssim 1$, have coalesced into a single cut with branch points at

$$\ell = 1 \pm iR_0\sqrt{-t}\,. \qquad (78)$$

At $t = 0$ this cut collapses into a third order pole at $\ell = 1$, as is necessary if the total cross section is to grow like $(\ell n s)^2$.

As a final example of the exotic predictions made by the REM with $\alpha(0) > 1$, consider a process such as p-n charge exchange scattering where quantum numbers are exchanged. We expect the amplitude to be dominated by the exchange of a single ρ meson, $\alpha_\rho(0) < 1$. However, it is also necessary to exchange our leading trajectory an arbitrary number of times. By now the counting should be straightforward. Neglecting the spin of the incident particles we find

$$M_{C.E.}(s,\underset{\sim}{\Delta}) = 2is \int d^2b\, e^{i\delta(s,\underset{\sim}{b})} [-i\delta_\rho(s,\underset{\sim}{b})]\, e^{i\underset{\sim}{\Delta}\cdot\underset{\sim}{b}}\,. \qquad (79)$$

In the forward direction the answer is extremely simple.

$$M_{C.E.}(s,0) = \frac{\gamma_\rho(0)}{4\pi\alpha'_\rho(0)\,\ell ns}\, s^{\alpha_\rho(0)} \int d^2b\; e^{-b/4\alpha'_\rho(0)\,\ell ns}\,\theta(b-b_o)$$

$$= \gamma_\rho(0)\; s^{[\alpha_\rho(0)-R_o^2/4\alpha'_\rho(0)]}. \qquad (80)$$

Clearly if our model is correct and if R_o is appreciably different from zero, the energy dependence of n-p charge exchange amplitude near the forward direction will be quite different from what one would predict by looking at the Chew-Frautschi plot.

V. PRODUCTION AMPLITUDES

It has long been assumed that in order to have a real understanding of high energy hadron interactions one will have to have a satsifactory model of production amplitudes. Let us see whether the insight that we have gained from studying the eikonal model of elastic scattering can help us with this problem.[16]

At present the most popular model of production amplitudes is the multiperipheral model. Let us consider a world with only one type of particle which has spin zero. The multiperipheral amplitude for the production of n-1 particles is shown schematically in figure 12. The variables have been chosen to be the same as in our discussion of the ladder graph, and as there the independent variables will be taken to be the k_i and y_i, $i = 1,\ldots n$. The high energy behavior of the ladder graph came when all the produced particles except the one carrying momentum P_2+k_n were on the mass shell. In the present case, this particle is also on the mass shell and we see from Eq. (40) that this imposes the constraint

$$s \prod_{i=1}^{n} y_i = m^2. \qquad (81)$$

The physical significance of the y_i can be seen by noticing that the sub-energy between any two adjacent particles is given by

$$s_i = (k_{i-1} - k_{i+1})^2 =$$

$$= \sqrt{s}\, y_1 \cdots y_{i-1}(1-y_i y_{i+1}) \left[\frac{(k_i - k_{i+1})^2 + m^2}{\sqrt{s} y_1 \cdots y_i (1-y_{i+1})} + \frac{(k_{i-1} - k_i)^2 + m^2}{\sqrt{s} y_1 \cdots y_{i-1}(1-y_i)} \right]$$

$$- (k_{i-1} - k_i)^2, \qquad i = 1, \ldots n \qquad (82)$$

with $k_0 \equiv P_1$ and $k_{n+1} \equiv P_2$. For small values of the y_i we have

$$s_i \simeq [(k_i - k_{i+1})^2 + m^2]/y_i \equiv m_i^2/y_i. \qquad (83)$$

For large values of the sub-energies Eq. (81) can be re-written in the form

$$\prod_{i=1}^{n} s_i = sA \qquad (84)$$

In the simplest version of the multiperipheral model one assumes that only nearest neighbor interactions along the chain are important. One then writes the amplitude shown in figure 12 in the form

$$T_{2n} = \prod_{i=1}^{n} M(s_i, k_i^2) \prod_{j=1}^{n-1} g(k_j, k_{j+1}) \qquad (85)$$

where M is the two-body scattering amplitude. For simplicity I shall approximate the vertex functions, g, by constants, although this is by no means crucial.

It is well known that one of the major problems associated with the multiperipheral model is that it leads to a violation of the Froissart bound if the leading t-plane singularity of the elastic scattering amplitude reaches 1.[15] We shall see this explicitly in a few minutes. This difficulty can of course be avoided by assuming that the Pomeranchuk pole has a t = 0 intercept slightly less than one. Let us take an alternative approach and try to construct a model of production amplitudes whose form guarantees that unitarity will be satisfied independent of the structure of the two-body amplitude.

I will only consider the region of phase space in which all sub-energies are large. The amplitude will be taken to be zero elsewhere. The region of small sub-energies is of course far from negligible, but the question that I would like to ask is whether it is possible to tame the high sub-energies tail of the amplitude which leads to the violation of the Froissart bound in the multiperipheral model.

If $M(s,t) \to s^a$ for large s, then we see from Eqs. (84) and (85) that $T_{2n} \to s^a$ when all the sub-energies are large. Our experience with the eikonal model tells us that for $a \approx 1$, it will be necessary to consider the exchange of more than one chain. Consider the two chain diagrams shown in figure 13. If the incident particles retain a large fraction of their momenta, then we expect to be able to replace their propagators by their mass shell delta functions. Thus the constraint of Eq. (84) holds for the sub-energies along each chain

$$\prod_{j=1}^{n_i} s_{ij} = sA_i \qquad i = 1,2 \qquad (86)$$

where s_{ij} is the jth sub-energy on the ith chain. We can use the delta functions to do the integrals over the plus and minus components of the loop momenta, q, picking up a factor of $\frac{i}{s^2}$ as usual. The amplitude will thus have an s dependence of the form s^{2a-1}, so it will be just as important as the single chain diagram for $a \approx 1$. Clearly it will be necessary to consider the exchange of an arbitrary number of chains.

Another lesson we learned from the eikonal model is that scattering amplitudes are sometimes simpler in impact parameter space. From Eq. (39) we see that for large sub-energies the momentum transfer variables $k_i^2 \to -\underset{\sim}{k}_i^2$, i.e. they can be expressed in terms of two-dimensional variables in the x-y plane. It is therefore convenient to take two-dimensional fourier transforms with respect to the $\underset{\sim}{k}_i$ and write Eq. (85) in the form

$$T_{2n}(s_i, \underset{\sim}{b}_i) = g^{n-1} \prod_{i=1}^{n} M(s_i, \underset{\sim}{b}_i) \ . \qquad (87)$$

For the two-chain diagrams of figure 13 the integrations over the loop momenta $\underset{\sim}{q}$ just introduce a delta function which sets the total impact parameter of chain 1 equal to that of chain 2, i.e.

$$\sum_{i=1}^{n_1} \underset{\sim}{b}_{1i} = \sum_{i=1}^{n_2} \underset{\sim}{b}_{2i} \, . \tag{88}$$

This is exactly analogous to the result of Eq. (21).

The above ideas can be put together to write down a concrete model for the production amplitude. We assume that only interactions between nearest neighbors are important. This is a popular, but probably bad assumption. My only excuse for it is that it makes the model simple enough to solve. Before any particles are produced the incident particles are nearest neighbors so I will include interactions between them. I assume that the basic interaction comes from the exchange of a simple object such as a ladder or tower. A typical diagram is shown in figure 14. The wavy lines are of course to be crossed in all possible ways. Notice that unlike the case of elastic scattering there are diagrams in which the incident particles interact directly zero times. Their interaction introduces a factor

$$S(s,\underset{\sim}{b}) \equiv \sum_{n=0}^{\infty} [i\delta(s,\underset{\sim}{b})]^n/n! = 1 + \frac{i}{2s} M(s,\underset{\sim}{b}) \, . \tag{89}$$

The amplitude for the exchange of N chains with n_1-1 particles produced from the first chain, n_2-1 from the second, etc., is given by

$$M_{2;n_1 \cdots n_N}(s,\underset{\sim}{b};\, s_{ij},\, \underset{\sim}{b}_{ij}) \tag{90}$$
$$= S(s,\underset{\sim}{b}) \left(\frac{i}{2s}\right)^{N-1} \prod_{i=1}^{N} g^{n_i-1} \prod_{j=1}^{n_i} M(s_{ij},\, \underset{\sim}{b}_{ij})$$

with the constraints

$$\prod_{j=1}^{n_i} s_{ij} = sA_i, \quad \sum_{j=1}^{n_i} \underset{\sim}{b}_{ij} = \underset{\sim}{b} \qquad i = 1,\ldots N \tag{91}$$

In the process of crossing the lines of all exchanged objects we interchange all the chains. So to avoid overcounting we take $n_1 \geq n_2 \ldots \geq n_N$.

Having constructed a model of production amplitudes we can calculate directly the imaginary part of the elastic amplitude using the unitarity equation. Let us first recall how the calculation goes in the multiperipheral model. We call $\mathrm{Im}\, T^n(s,\underset{\sim}{\Delta})$ the contribution to the imaginary part of the elastic amplitude arising from intermediate states of n+1 particles. In our normalization

$$\mathrm{Im}\, T^n(s,\underset{\sim}{\Delta}) = \pi \int \prod_{i=1}^{n} \frac{d^4k_i}{(2\pi)^3}\, \delta^+[(P_1-k_1)^2 - m^2]$$

$$\cdot \prod_{j=1}^{n-1} \delta^+[(k_j-k_{j+1})^2 - m^2]\, \delta^+[(P_2+k_n)^2 - m^2] \tag{92}$$

$$\cdot\, T_{2n}(s_i,k_i)\, T^*_{2n}(s_i,k_i+\Delta)$$

$$\simeq \frac{1}{4s}\left(\frac{g^2}{4\pi}\right)^{n-1} \int \prod_{i=1}^{n} \frac{d^2\underset{\sim}{k}_i}{(2\pi)^2} \int_0^1 \prod_{i=1}^{n} dy_i\, \delta\Big(\prod_1^n y_i - m^2/s\Big)$$

$$\prod_{i=1}^{n} M(m_i^2/y_i,\underset{\sim}{k}_i)\, M^*(m_i^2/y_i,\underset{\sim}{k}_i+\underset{\sim}{\Delta})$$

In Eq. (92) we have performed all the $k_{i\perp}$ integrations by means of the delta functions, and have made the same change of variables as in Eq. (37). Since we are setting the production amplitude equal to zero when any of the subenergies are not large, we have made approximations appropriate for small values of the y_i. Taking the two-dimensional fourier transform with respect to $\underset{\sim}{\Delta}$ gives

$$\mathrm{Im}\, T^n(s,\underset{\sim}{b}) = \frac{1}{4s}\left(\frac{g^2}{4\pi}\right)^{n-1} \int \prod_{i=1}^{n} d^2\underset{\sim}{b}_i \int_0^1 \prod_{i=1}^{n} dy_i\, \delta\Big(\prod_1^n y_i - n^2/s\Big)$$

$$\cdot \prod_{i=1}^{n} |M(m_i^2/y_i,\underset{\sim}{b}_i)|^2\, \delta^2\Big(\underset{\sim}{b} - \sum_{i=1}^{n} \underset{\sim}{b}_i\Big) \tag{93}$$

We now return to the production amplitudes defined by Eqs. (90) and (91). When we square the amplitude for the production of n particles we find two types of terms: diagonal ones in which chains of equal length come together and off diagonal ones in which three or more chains are joined together. The latter clearly correspond to non-nearest neighbor interactions and must be discarded to be consistent with our earlier neglect of such interactions. The contribution to the imaginary part of the elastic amplitude arising from the production amplitude with N chains of length $n_1, n_2, \ldots n_N$ is easily seen to be

$$\text{Im } M^{n_1 \cdots n_N}(s,\underset{\sim}{b}) = |S(s,\underset{\sim}{b})|^2 \; |\tfrac{1}{s}|^{N-1} \prod_{i=1}^{N} \text{Im } T^{n_i}(s,\underset{\sim}{b}), \quad (94)$$

where $\text{Im } T^n(s,\underset{\sim}{b})$ is defined by Eq. (93). The total contribution to $\text{Im } M(s,\underset{\sim}{b})$ coming from the N chain diagrams is given by

$$\begin{aligned} \text{Im } M^N(s,\underset{\sim}{b}) &\equiv \sum_{n_1 \geq n_2 \geq \ldots \geq n_N} \text{Im } M^{n_1 \cdots n_N}(s,\underset{\sim}{b}) \\ &= \frac{1}{N!} \sum_{n_i=1}^{\infty} \text{Im } M^{n_1, \ldots n_N}(s,\underset{\sim}{b}) \qquad (95) \\ &= |S(s,\underset{\sim}{b})|^2 \; |\tfrac{1}{s}|^{N-1} \; C(s,\underset{\sim}{b})/N! \quad , \end{aligned}$$

where

$$C(s,\underset{\sim}{b}) \equiv \sum_{n=1}^{\infty} \text{Im } T^n(s,\underset{\sim}{b}) \qquad (96)$$

is the multiperipheral model's prediction for the imaginary part of the elastic scattering amplitude. In our model the total imaginary part of the elastic amplitude is given by

$$\begin{aligned} \text{Im } M(s,\underset{\sim}{b}) &= \frac{1}{4s} \; |M(s,\underset{\sim}{b})|^2 + \sum_{n=1}^{\infty} \text{Im } M^N(s,\underset{\sim}{b}) \\ &= \frac{1}{4s} \; |M(s,\underset{\sim}{b})|^2 + |S(s,\underset{\sim}{b})|^2 \; s(e^{C(s,\underset{\sim}{b})/s} - 1). \qquad (97) \end{aligned}$$

We have consistently written $M(s,b)$ in the form

$$M(s,\underset{\sim}{b}) = 2is(1-e^{i\delta(s,\underset{\sim}{b})}). \tag{98}$$

However we are presently interested in diffraction scattering so we expect M to be pure imaginary which means the δ is also pure imaginary. It is therefore convenient to write

$$\delta(s,\underset{\sim}{b}) = ia(s,\underset{\sim}{b}), \tag{99}$$

Eq. (97) then becomes

$$2s(1-e^{-a}) = s(1-e^{-a})^2 + e^{-a}(e^{C/s}-1), \tag{100}$$

so unitarity is satisfied exactly provided

$$a(s,\underset{\sim}{b}) = C(s,\underset{\sim}{b})/2s. \tag{101}$$

Eq. (101) is a bootstrap equation for the elastic scattering amplitude. One guesses an input form for $a(s,\underset{\sim}{b})$ or equivalently for $M(s,\underset{\sim}{b})$ and uses Eqs. (93) and (96) to compute $C(s,\underset{\sim}{b})$. If the output value of $a(s,\underset{\sim}{b})$ obtained from Eq. (101) matches the input a we have a self consistent model in which the elastic amplitude satisfies unitarity exactly.

As a first example let us take the input amplitude to be dominated by a single Regge pole with intercept $\alpha(0)=1$. In this case it turns out to be simplest to work in moment rather than impact parameter space. We write

$$M_{in}(s,\underset{\sim}{\Delta}) = \gamma(\underset{\sim}{\Delta})s^{\alpha(\underset{\sim}{\Delta})} \tag{102}$$

Substituting into Eq. (92) and taking the Mellin transform gives

$$\mathrm{Im}T^n(\ell,\underset{\sim}{\Delta}) \equiv \int_{s_0}^{\infty} ds\ s^{-(\ell+1)}\ \mathrm{Im}T^n(s,\underset{\sim}{\Delta}) \simeq$$

$$\simeq \frac{1}{4}\left(\frac{g^2}{4\pi}\right)^{n-1} \int \prod_{i=1}^{n} \frac{d^2k_i}{(2\pi)^2}\, ds_i\, \gamma(\underset{\sim}{k}_i)\,\gamma(\underset{\sim}{k}_i + \underset{\sim}{\Delta}) \qquad (103)$$

$$\cdot\, s_i^{\alpha(\underset{\sim}{k}_i) + \alpha(\underset{\sim}{k}_i + \underset{\sim}{\Delta}) - \ell - 2} \,.$$

In the last line of Eq. (103) we have made the change of variables $s_i = m_i^2/y_i$. We must cut off the lower end of the s_i integrations at some large number, call it $\bar{s}_0$, since we have consistently assumed that T_{2n} vanishes if any of the s_i are not large. Defining

$$I(\ell, \underset{\sim}{\Delta}) \equiv \int \frac{d^2k}{(2\pi)^2}\, \gamma(\underset{\sim}{k})\,\gamma(\underset{\sim}{k}+\underset{\sim}{\Delta}) \int_{s_0}^{\infty} ds\; s^{\alpha(\underset{\sim}{k})+\alpha(\underset{\sim}{k}+\underset{\sim}{\Delta})-\ell-2}$$

$$= \int \frac{d^2k}{(2\pi)^2}\, \frac{\gamma(\underset{\sim}{k})\,\gamma(\underset{\sim}{k}+\underset{\sim}{\Delta})\,(\bar{s}_0)^{\alpha(\underset{\sim}{k})+\alpha(\underset{\sim}{k}+\underset{\sim}{\Delta})-\ell-1}}{\ell-\alpha(\underset{\sim}{k})-\alpha(\underset{\sim}{k}+\underset{\sim}{\Delta})+1} \qquad (104)$$

we see that for $\ell \approx 2\alpha(\frac{1}{2}\underset{\sim}{\Delta})-1$

$$I(\ell, \underset{\sim}{\Delta}) \simeq \frac{-\gamma(\frac{1}{2}\underset{\sim}{\Delta})}{8\pi\alpha'(0)}\, \ln[\ell - 2\alpha(\tfrac{1}{2}\underset{\sim}{\Delta}) - 1] \qquad (105)$$

so

$$C(\ell, \underset{\sim}{\Delta}) \equiv \sum_{n=1}^{\infty} \mathrm{Im}T^n(\ell, \underset{\sim}{\Delta})$$

$$\simeq \frac{1}{4}\, I(\ell, \underset{\sim}{\Delta}) \left[1 - \frac{g^2}{4\pi} I(\ell, \underset{\sim}{\Delta})\right]^{-1} . \qquad (106)$$

Since $\alpha(0) = 1$, $C(\ell, 0)$ must have a pole in the ℓ-plane for $1 < \ell < 2$.

In the multiperipheral model where C is the imaginary part of the elastic scattering amplitude this pole leads directly to a violation of the Froissart bound. In the present model there is no such violation provided we take $a_{out} = C$; however, it is clear that Eq. (102) does not

lead to a self consistent solution to the model.

The fact that an input Regge pole with $\alpha(0) = 1$ leads to a function $C(s,\underset{\sim}{b})$ which grows with s suggests that we try an input amplitude of the form

$$M_{in}(s,\underset{\sim}{b}) = 2is\ \theta(b_o - b) \tag{107}$$

with

$$b_o = R_o\ \ell n s .$$

Substituting into Eq. (93) and again taking the Mellin transform yields

$$\begin{aligned} \mathrm{Im}T^n(\ell,\underset{\sim}{b}) &\simeq \frac{1}{4}\left(\frac{g^2}{4\pi}\right)^{n-1} \int \prod_{i=1}^{n} d^2\underset{\sim}{b}_i\ ds_i\ 4s_i^{-\ell}\theta(b_{io}-b_i) \\ &\quad \delta\Big(\underset{\sim}{b} - \sum_{i=1}^{n} \underset{\sim}{b}_i\Big) \\ &= \frac{1}{4}\left(\frac{g^2}{4\pi}\right)^{n-1} \int \frac{d^2q}{(2\pi)^2}\, e^{-i\underset{\sim}{q}\cdot\underset{\sim}{b}} \int \prod_{i=1}^{n} d^2\underset{\sim}{b}_i\ ds_i e^{i\underset{\sim}{q}\cdot\underset{\sim}{b}_i} \\ &\quad \cdot 4s_i^{-\ell}\ \theta(b_{io} - b_i) \end{aligned} \tag{108}$$

where $b_{io} = R_o\ \ell n s_i$. The integrals over the $\underset{\sim}{b}_i$ and s_i can be done as in Eqs. (75) and (77). Retaining only the leading ℓ-plane singularities we find

$$\begin{aligned} \mathrm{Im}T^n(\ell,\underset{\sim}{b}) &\simeq \frac{1}{4}\left(\frac{g^2}{4\pi}\right)^{n-1} \frac{1}{2\pi} \int_0^\infty q\ dq\ J_o(qb) \\ &\quad \cdot\Big[8\pi R_o^2/[(\ell-1)^2 + R_o^2q^2]^{3/2}\Big]^n . \end{aligned} \tag{109}$$

The inverse Mellin transform of Eq. (109) is surprisingly simple.[17]

$$\mathrm{Im}T^n(s,\underset{\sim}{b}) = \frac{1}{4}\left(\frac{g_a}{4\pi}\right)^{n-1} \frac{1}{2\pi} \int_0^\infty q\, dq\, J_0(qb)$$

$$\cdot (8\pi R_0^2)^n\, s(\ell ns)^{\frac{3}{2}n-\frac{1}{2}} \frac{\Gamma(\frac{1}{2}) J_{\frac{3}{2}n-\frac{1}{2}}(R_0 \ell ns\ q)}{\Gamma(\frac{3n}{2})\ (2R_0 q)^{\frac{3}{2}n-\frac{1}{2}}}$$

$$= \frac{s\ (2g^2/R_0)^{n-1}\left[(b_0^2-b^2)^{\frac{1}{2}}\right]^{3n-2}}{(b_0^2-b^2)^{\frac{1}{2}}\quad (3n-2)!}\ \theta(b_0-b) \tag{110}$$

For virtually the entire internal $0 \le b \le b_0$, the quantity $(b_0^2-b^2)^{\frac{1}{2}}$ is large, so the series can be summed to give

$$C(s,\underset{\sim}{b}) = \sum_{n=1}^{\infty} \mathrm{Im}T^n(s,\underset{\sim}{b})$$

$$\simeq \frac{1}{3}\ \frac{s\ e^{\alpha(b_0^2-b^2)^{\frac{1}{2}}}}{\alpha(b_0^2-b^2)^{\frac{1}{2}}}\ \theta(b_0-b) \tag{111}$$

where

$$\alpha = (2g^2/R_0)^{\frac{1}{3}}\ .$$

Writing

$$a_{out} = C(s,\underset{\sim}{b})/2s \tag{112}$$

we see that a_{out} grows like a power of s for $b < b_0$ so

$$M_{out}(s,\underset{\sim}{b}) = 2is(1-e^{-a_{out}}) \simeq 2is\ \theta(b_0 - b) = M_{in}(s,\underset{\sim}{b}). \tag{113}$$

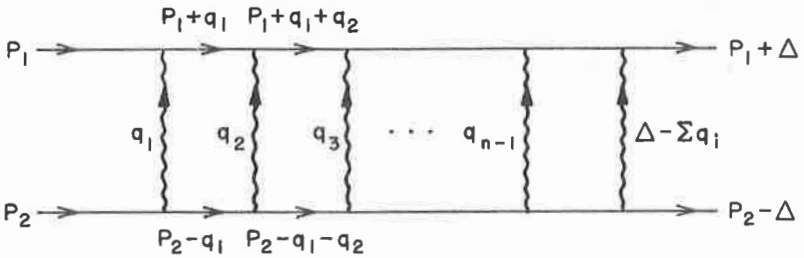
We therefore have a self consistent two-particle scattering amplitude which satisfies unitarity exactly.

The above calculation does not constitute too strong

an argument in favor of the black disc model of diffraction scattering. Even within the context of the model there is no good reason for ignoring interactions between non-nearest neighbors. Such interactions would clearly introduce more factors of the S-matrix into the production amplitude. This final state could therefore lead to a self-consistent elastic amplitude for which the total cross-section does not increase with energy. In my opinion the main lesson of this calculation is that when the leading ℓ-plane singularity of the elastic amplitude approaches one, it is probably necessary to consider production from more than one multiperipheral chain. How this idea can be used to construct a realistic model of diffraction scattering remains a challenging problem.

Figure 1

a)


P_1
P_1+q_1
$P_1+q_1+q_2$
$P_1+\Delta$
q_1
q_2
q_3
q_{n-1}
$\Delta-\Sigma q_i$
P_2
$P_2-\Delta$
P_2-q_1
$P_2-q_1-q_2$

b)

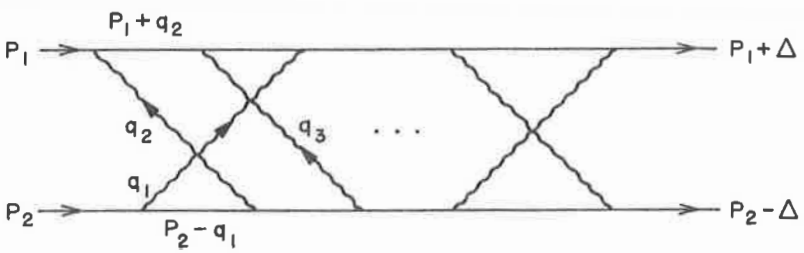

P_1
P_1+q_2
$P_1+\Delta$
q_2
q_3
q_1
P_2
$P_2-\Delta$
P_2-q_1

Figure 2

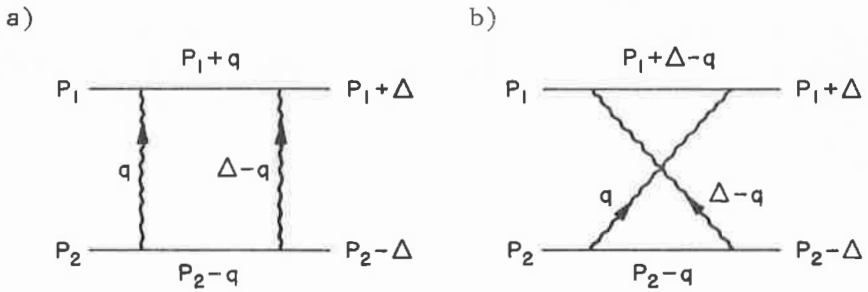


Figure 3

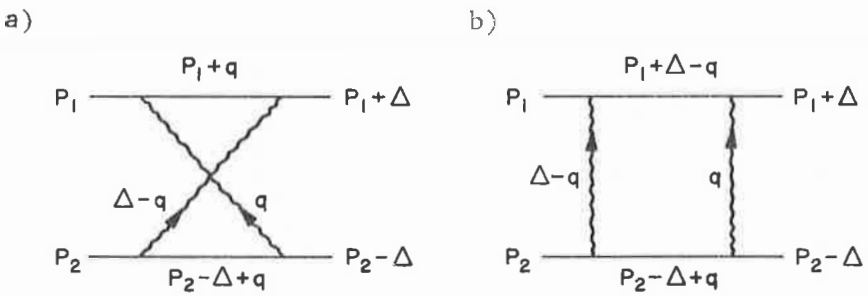


Figure 4

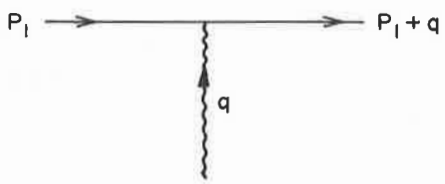


Figure 5

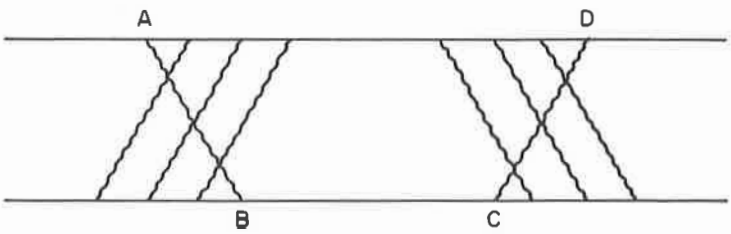
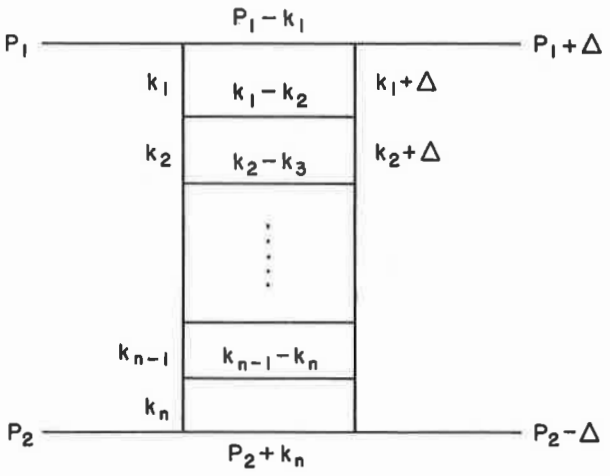


Figure 6

a)


$P_1 - k_1$
P_1
$P_1 + \Delta$
k_1
$k_1 - k_2$
$k_1 + \Delta$
k_2
$k_2 - k_3$
$k_2 + \Delta$
k_{n-1}
$k_{n-1} - k_n$
k_n
P_2
$P_2 - \Delta$
$P_2 + k_n$

b)

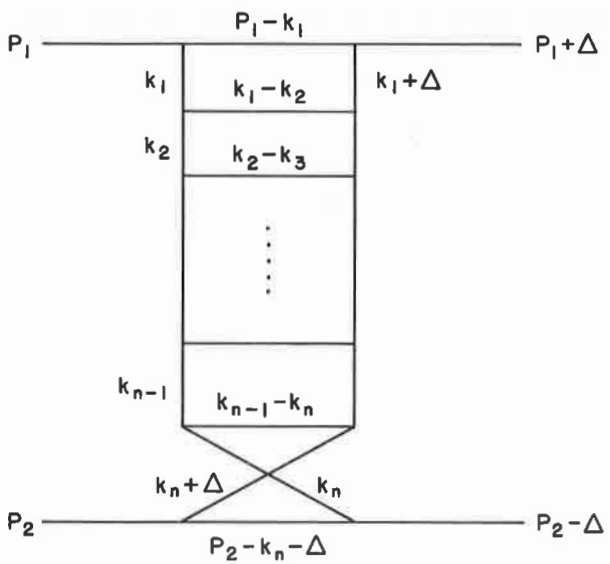
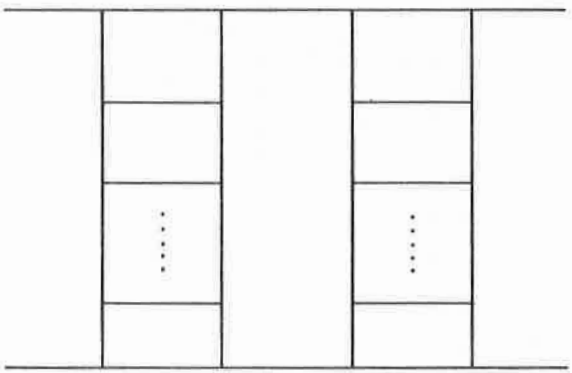

$P_1 - k_1$
P_1
$P_1 + \Delta$
k_1
$k_1 - k_2$
$k_1 + \Delta$
k_2
$k_2 - k_3$
k_{n-1}
$k_{n-1} - k_n$
$k_n + \Delta$
k_n
P_2
$P_2 - \Delta$
$P_2 - k_n - \Delta$

Figure 7

a)



b)

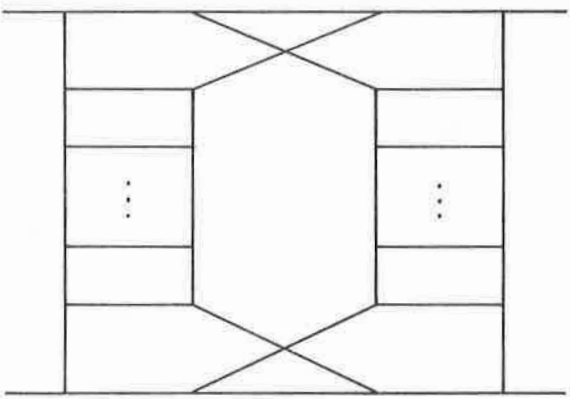
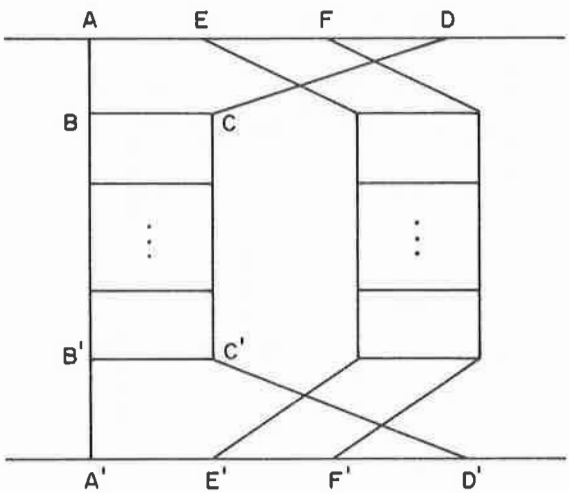


Figure 8

a)



b)

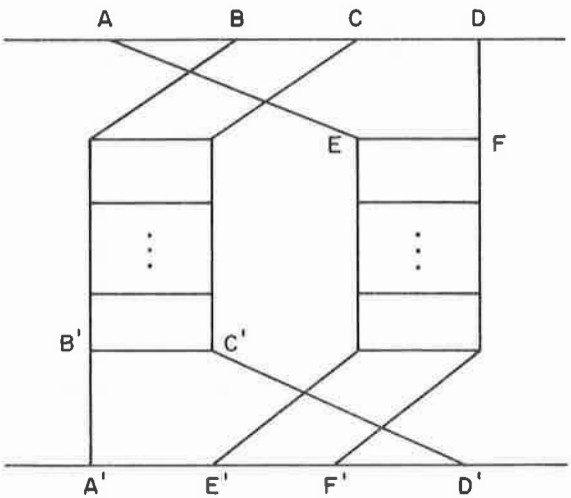
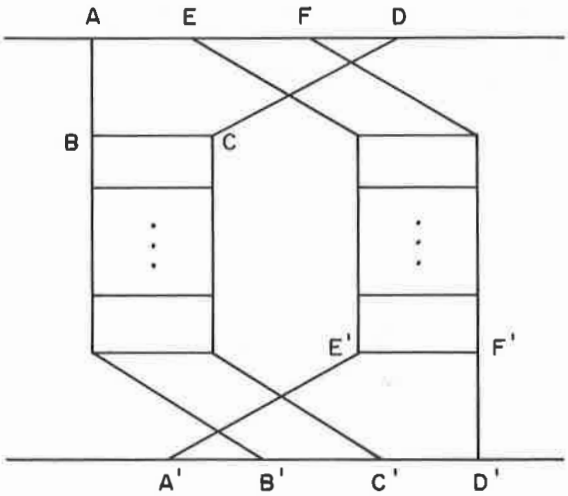


Figure 8

c)



d)

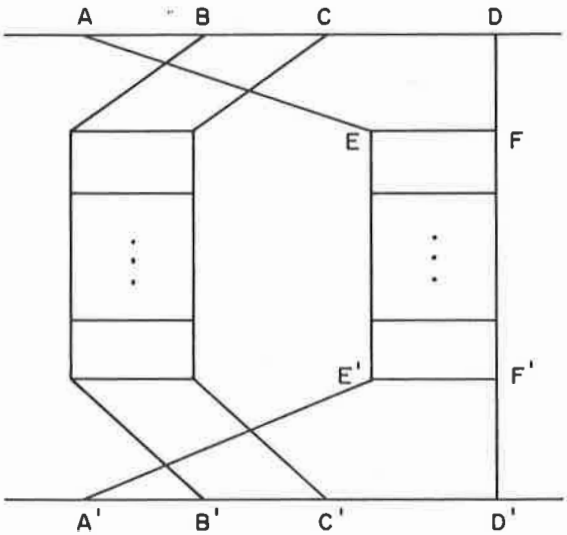
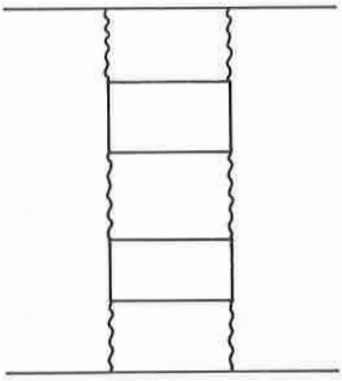


Figure 9

a)



b)

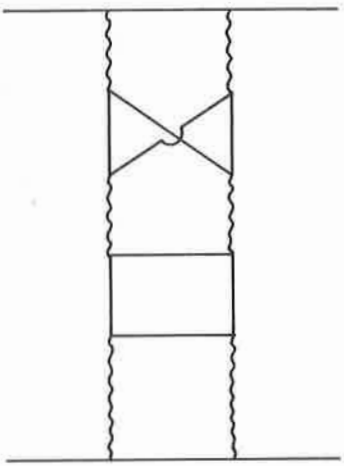


Figure 10

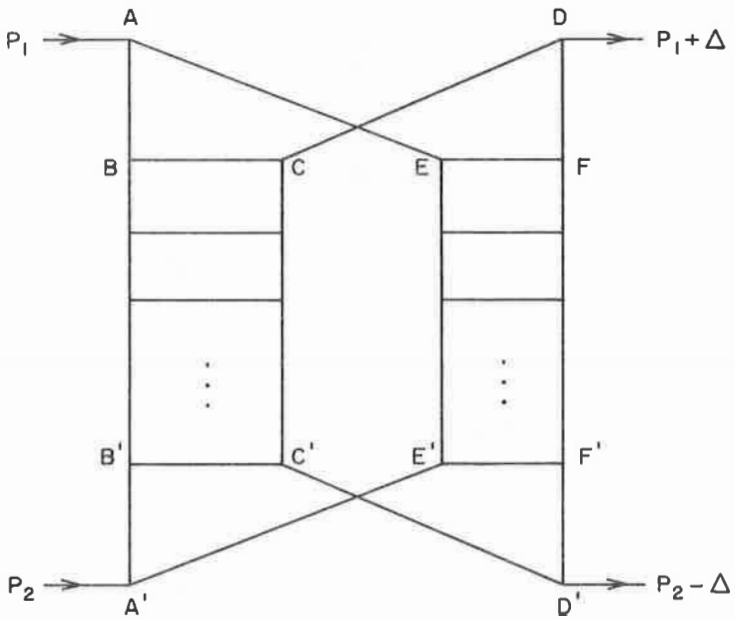
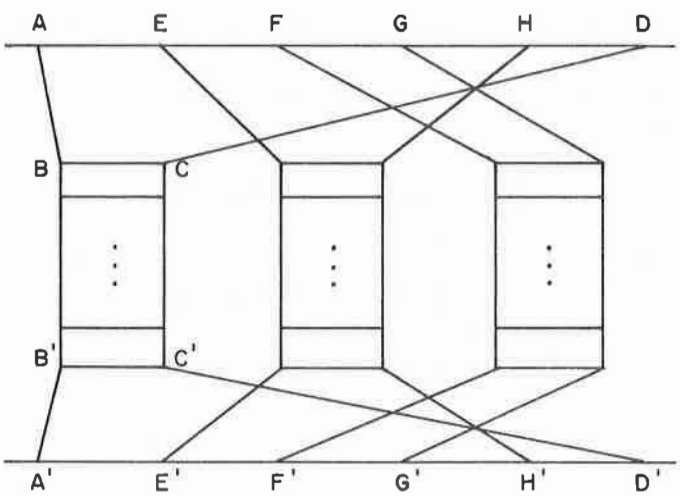


Figure 11

a)



b)

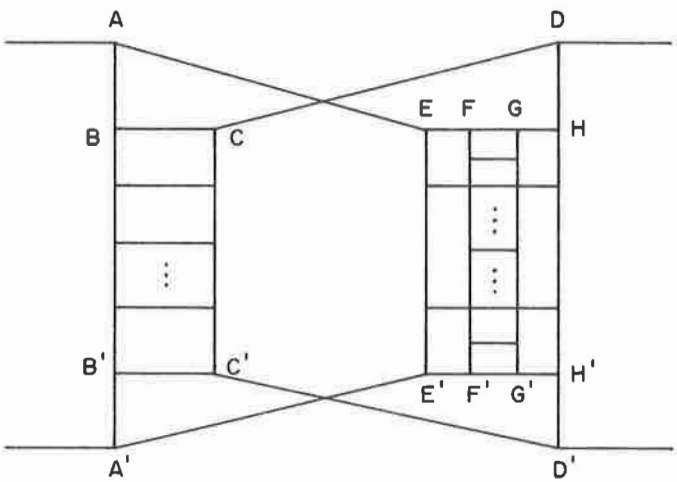


Figure 12

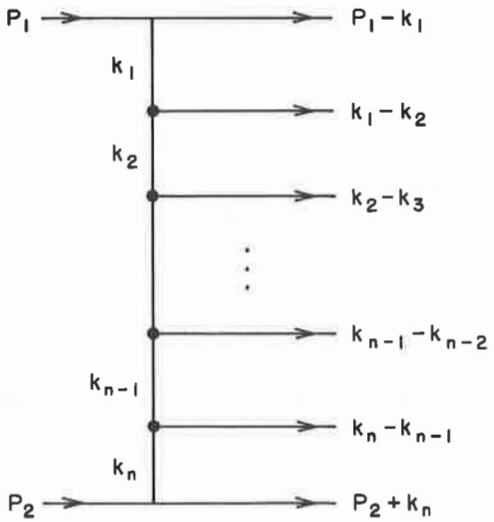


Figure 13

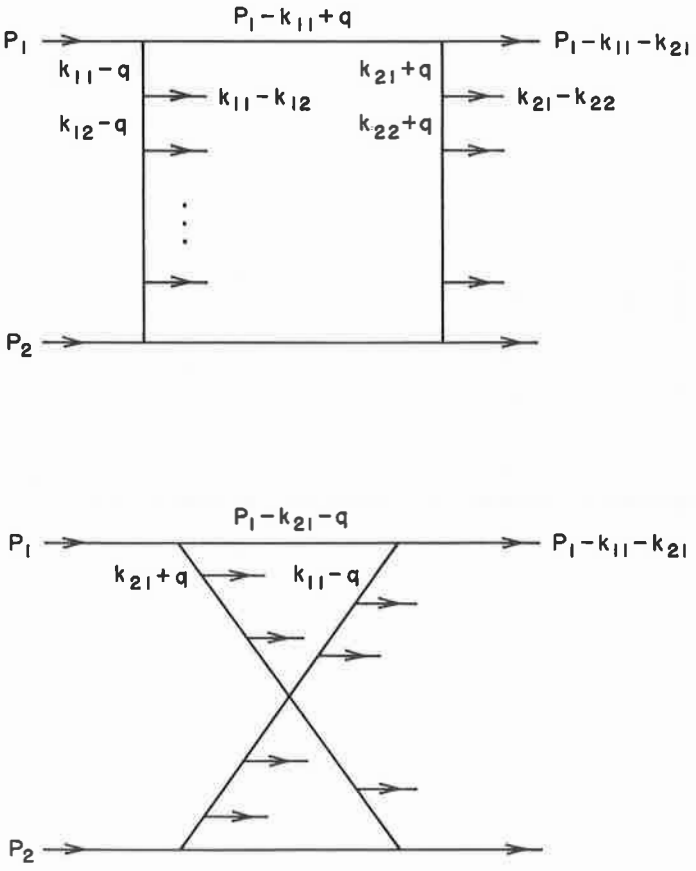
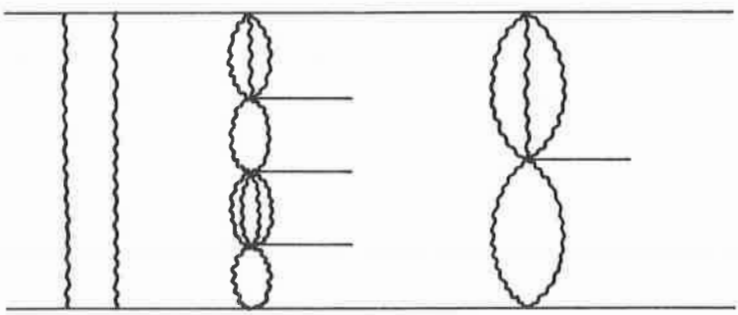

P_1
$P_1 - k_{11} + q$
$P_1 - k_{11} - k_{21}$
$k_{11} - q$
$k_{11} - k_{12}$
$k_{21} + q$
$k_{21} - k_{22}$
$k_{12} - q$
$k_{22} + q$
P_2
P_1
$P_1 - k_{21} - q$
$P_1 - k_{11} - k_{21}$
$k_{21} + q$
$k_{11} - q$
P_2

Figure 14



REFERENCES

1. S.J. Chang and S. Ma, Phys. Rev. Lett. 22, 1334 (1969); Phys. Rev. 188, 2389 (1969).

2. M. Levy and J. Sucher, Phys. Rev. 186, 1656 (1969).

3. H.D.I. Abarbanel and C.I. Itzykson, Phys. Rev. Lett. 23, 53 (1969).

4. R.J. Eden, et. al., "The Analytic S-Matrix" Cambridge University Press, London (1966).

5. G. Tiktopoulos and S.B. Treiman, Phys. Rev. D3, 1037 (1971).

6. R. Arnold, Phys. Rev. 153, 1523 (1967); S. Frautschi and B. Margolis, Nuovo Cimento 56A, 1155 (1969).

7. S.J. Chang and T.M. Yan, Phys. Rev. Lett. 25, 1586 (1970).

8. B. Hasslacher, et. al., Phys. Rev. Lett. 25, 1591 (1970).

9. H. Cheng and T.T. Wu, Phys. Rev. Lett. 24, 1456 (1970).

10. B.W. Lee and R.F. Sawyer, Phys. Rev. 127, 2266 (1962).

11. V.N. Gribov, JETP 26, 414 (1968).

12. G.M. Cicuta and R.L. Sugar, Phys. Rev. D3, 970 (1971).

13. B. Hasslacher and D.K. Sinclair, Phys. Rev. D3, 1770 (1971).

14. V.N. Gribov, I.Pomeranchuk, and K. Ter-Martorosyan, Phys. Rev. 139, B184 (1965).

15. J. Finkelstein and K. Kajantie, Phys. Lett. 26B, 305 (1968).

16. J.R. Fulco and R.L. Sugar, Phys. Rev. (To be published.)

17. A. Erdelyi, et. al., "Tables of Integral Transform" McGraw-Hill (1954).

EIKONAL DESCRIPTION OF HIGH-ENERGY PARTICLE SCATTERING

M. M. Islam
Department of Physics, University of Connecticut
Storrs, Connecticut 06268

1. Introduction

A few years ago at this institute I gave a set of lectures on the same subject.[1] My primary motivation at that time was to generate interest in this field by pointing out that the eikonal description provides not only a useful but probably a powerful tool to explore hadron interactions. An added motivation was that new accelerators were being built, and we needed a suitable theoretical framework to discuss experimental results in an energy region where the partial wave expansion is hopelessly impractical. It is with some satisfaction that one can look back, and notice that considerable work has been done in this area during the last couple of years. Various field theory models have been studied - a number of which have indeed led to the eikonal description of high energy scattering. Phenomenologically the eikonal description has been used to obtain full scattering amplitudes after assuming suitable input amplitudes, and then compared with the ever increasing volume of experimental data. To these efforts we can now add the fact that Serpukhov and CERN ISR have begun producing experimental results, and NAL will have results in the near future. It is therefore an appropriate time to try to understand in a simple physical way some of the important aspects of the eikonal description of high energy scattering, and in these lectures we shall be mainly concerned with the following aspects:

1) How does a field theory model lead to an eikonal description?

2) What does it mean to take the input amplitude in

the eikonal representation as a Regge pole amplitude?

3) How can a phenomenological complex energy-dependent optical potential be interpreted in the language of the S-matrix theory?

2. Eikonal Amplitude from Field Theory

By summing the generalized ladder diagrams of field theory (Fig. 1), the eikonal description of high-energy elastic scattering has been obtained by many authors. Apart from elaborate and rigorous mathematical methods[2] various approximate methods such as (i) functional techniques with soft-meson approximation,[3,4] (ii) infinite momentum technique,[5] (iii) modification of the Feynman propagator[6-8] have been used. We shall adopt the last method for our discussion. Even though mathematically not very rigorous, this method provides a straightforward graphical way of obtaining the desired result, and also lends itself easily to the description of large momentum transfer scattering and to the investigation of radiative corrections.[9] An added feature of the method is that it is very similar to the graphical analysis of infrared divergence in quantum electrodynamics (QED),[10] and we shall have occasions to go over to QED to check with results known there. We shall also be interested in discussing radiative corrections to the eikonal amplitude, mainly for two reasons: (1) a number of authors have already investigated the question of radiative corrections to all orders to the eikonal amplitude in QED;[11,12,13] (2) such corrections have been proposed as an explanation for the precipitous fall of high-energy large-angle pp elastic scattering.[14]

2(a). Summing an Infinite Set of Diagrams

We began with the diagram in Fig. 2 which is somewhat more general than the generalized ladder diagram in Fig. 1. The shaded blob in this diagram represents some unspecified interaction; we consider this interaction as not known to us at the beginning. The wavy lines represent mesons which are being exchanged between two particles a and b. The initial 4-momenta of the particles are p_a and p_b, and their

final 4-momenta are p_a' and p_b'. For simplicity we shall first consider the particles $\underline{a}$ and $\underline{b}$ as spinless and the mesons as scalar mesons. Let $k_1, k_2, \ldots, k_{n+1}$ are the 4-momenta emitted along the $\underline{a}$ line, and $\bar{k}_1, \bar{k}_2, \ldots \bar{k}_{n+1}$ are the 4-momenta absorbed along the $\underline{b}$ line. The unknown interaction carries off 4-momentum k_r from the $\underline{a}$ line, and will be represented by a Feynman amplitude $M^h(k_r)$.[15] The Feynman amplitude[16] for the process in Fig. 2 is now

$$M = \left[\frac{ig^2}{(2\pi)^4}\right]^n \int \prod_{\substack{i=1 \\ (\neq r)}}^{n+1} \frac{d^4k_i}{k_i^2-\mu^2+i\epsilon}\; d^4k_r\, M^h(k_r)\; \delta\Big(q - \sum_{i=1}^{n+1} k_i\Big)$$

$$\times \frac{1}{(p_a-k_1)^2-m^2+i\epsilon}\; \frac{1}{(p_a-k_1-k_2)^2-m^2+i\epsilon} \cdots \frac{1}{(p_a'+k_{n+1})^2-m^2+i\epsilon}$$

$$\times \frac{1}{(p_b+\bar{k}_1)^2-m^2+i\epsilon}\; \frac{1}{(p_b+\bar{k}_1+\bar{k}_2)^2-m^2+i\epsilon} \cdots \frac{1}{(p_b'-\bar{k}_{n+1})^2-m^2+i\epsilon}, \qquad (2.1)$$

where $q = p_a - p_a'$. We use the standard notation: $s=(p_a+p_b)^2$, $t=(p_a-p_a')^2$, $u=(p_a-p_b')^2$. The δ - function in (2.1) is eliminated by integrating over k_r, and the following approximation is then made. For propagators along $\underline{a}$ line before k_r

$$\frac{1}{(p_a-K)^2 - m^2+i\epsilon} \simeq \frac{1}{\sum_i (k_i^2 - 2p_a\cdot k_i)+ i\epsilon}, \qquad (2.2)$$

where $K = \sum_i k_i$; this means we drop the $k_i k_j (i \neq j)$ terms in the denominator. This is our basic approximation. The reasoning behind is that we are interested in the limit where the external 4-momentum p_a has components (p_{ao} and p_{az}) going to infinity, so that the part of the 4-momentum k_i that can contribute significantly is in the neighborhood

of $k_i = 0$. The reason for keeping, however, the k_i^2 terms is that no unknown cut-off is needed, as will be seen later, for integrals over d^4k_i. Mesons whose 4-momenta are such that (2.2) is justifiable will be referred to by us as soft mesons. In the above approximation we have for propagators after k_r

$$\frac{1}{(p_a' + K)^2 - m^2 + i\epsilon} \simeq \frac{1}{\sum_i (k_i^2 + 2p_a' \cdot k_i) + i\epsilon} ; \tag{2.3}$$

similar approximations are made along the b line before and after absorption of k_r. From now on, we shall refer to $M^h(k_r)$ as the hard interaction. From (2.2) and (2.3) it is seen that the hard interaction as defined here specifies the propagator approximations along a and b lines, and need not correspond to large values of the momentum transfer k_r.

Eq. (2.1) with the propagator approximation becomes

$$M = \left[\frac{ig^2}{(2\pi)^4}\right]^n \int \prod_{\substack{i=1 \\ (\neq r)}}^{n+1} \frac{d^4k_i}{k_i^2 - \mu^2} M^h \left(q - \sum_{\substack{i=1 \\ (\neq r)}}^{n+1} k_i\right) \rho_a \rho_a' \rho_b \rho_b' , \tag{2.4}$$

where

$$\rho_a = \frac{1}{k_1^2 - 2p_a \cdot k_1} \frac{1}{k_1^2 + k_2^2 - 2p_a \cdot (k_1 + k_2)} \cdots \frac{1}{\sum_{i=1}^{r-1} (k_i^2 - 2p_a \cdot k_i)} , \tag{2.5}$$

$$\rho_a' = \frac{1}{k_{n+1}^2 + 2p_a' \cdot k_{n+1}} \frac{1}{k_{n+1}^2 + k_n^2 + 2p_a' \cdot (k_{n+1} + k_n)} \cdots \frac{1}{\sum_{i=r+1}^{n+1} (k_i^2 + 2p_a' \cdot k_i)} ; \tag{2.6}$$

$\mathcal{P}_b$ and $\mathcal{P}'_b$ are similar products of propagators along p_b line and p'_b line.[17] Each propagator has a $+i\epsilon$ term which is no longer explicitly written.

Let us now assume that out of the 4-momenta $k_1, k_2 \ldots, k_{r-1}$ emitted by the p_a line, ℓ momenta denoted by $k'_1, k'_2, \ldots, k'_\ell$ are absorbed by the p_b line and the rest $r-1-\ell$ momenta $k'_{\ell+1}, k'_{\ell+2}, \ldots, k'_{r-1}$ by the p'_b line. Similarly, we assume that out of the 4-momenta $k_{r+1}, k_{r+2}, \ldots, k_{n+1}$ emitted by the p'_a line, m momenta denoted by $k''_1, k''_2, \ldots, k''_m$ are absorbed by the p_b line, and the rest $n+1-r-m$ momenta $k''_{m+1}, k''_{m+2}, \ldots, k''_{n+1-r}$ by the p'_b line. Obviously, the $\bar{k}_i$'s introduced earlier are a particular labeling of the k''s, k'''s, and k_r. We now want to sum over all diagrams obtained by all possible ways of attaching the k''s and k'''s. This summation is easily carried out because of the following identity:

$$\frac{1}{a}\prod_{i=1}^{n}\frac{1}{a_1+a_2\ldots+a_i+a}+\sum_{m=1}^{n}\prod_{i=1}^{m}\frac{1}{a_1+a_2+\ldots+a_i}\prod_{i=m}^{n}\frac{1}{a_1+a_2+\ldots+a_i+a}$$

$$=\frac{1}{a}\prod_{i=1}^{n}\frac{1}{a_1+a_2+\ldots+a_i}. \qquad (2.7)$$

To see explicitly the remarkable simplification resulting from this identity, let us suppose that a set of 4-momenta are attached to the p_b line as shown in Fig. 3. Let us take the product of all propagators and sum over all possible ways of attaching the 4-momentum k. Then because of (2.7) the sum is

$$\frac{1}{k^2+2p_b\cdot k}\,\frac{1}{q_1^2+2p_b\cdot q_1}\,\frac{1}{q_1^2+q_2^2+2p_b\cdot(q_1+q_2)}\cdots\frac{1}{\sum_{i=1}^{j}(q_i^2+2p_b\cdot q_i)}\cdot$$

This means that in the sum the term due to k factorizes out and the rest remains unchanged. Therefore, the scattering amplitude obtained from Eq. (2.4) by summing over

all possible ways of attaching k′'s and k″'s to the p_b and p'_b lines is

$$M = \left[\frac{ig^2}{(2\pi)^4}\right]^n \int \prod_{\substack{i=1 \\ (\neq r)}}^{n+1} \frac{d^4k_i}{k_i^2-\mu^2}\; M^h\left(q - \sum_{\substack{i=1 \\ (\neq r)}}^{n+1} k_i\right) \not{p}_a \not{p}'_a$$

$$\times \prod_{i=1}^{\ell} \frac{1}{k_i'^2+2p_b\cdot k_i'} \prod_{j=\ell+1}^{r-1} \frac{1}{k_j'^2-2p_b'\cdot k_j'}$$

$$\times \prod_{i=1}^{m} \frac{1}{k_i''^2+2p_b\cdot k_i''} \prod_{j=m+1}^{n+1-r} \frac{1}{k_j''^2-2p_b'\cdot k_j''} \quad . \tag{2.8}$$

Now the ℓ momenta $k_1', k_2', \ldots, k_\ell'$ is a subset of the set of momenta $k_1, k_2, \ldots, k_{r-1}$. Let us call this subset s'. The m momenta $k_1'', k_2'', \ldots, k_m''$ are also a subset of the set of momenta $k_{r+1}, k_{r+2}, \ldots, k_{n+1}$, and let us call this subset s''. Eq. (2.8) then corresponds to a particular choice of the subsets s' and s''. To include all possible topologically different Feynman diagrams where ℓ momenta from the p_a line and m momenta from the p'_a line are absorbed by the p_b line, we have to sum over all such subsets.

At this point we observe that if the sum over all possible permutations of $k_1, k_2, \ldots, k_{r-1}$ and of $k_{r+1}, k_{r+2}, \ldots, k_{n+1}$ are considered in Eq. (2.8), then the products of propagators along the p_a and p'_a lines factorize. This is because of the identity

$$\sum_P \frac{1}{a_{i1}} \frac{1}{a_{i1}+a_{i2}} \cdots \frac{1}{a_{i1}+a_{i2}+\ldots+a_{ij}} = \frac{1}{a_1}\frac{1}{a_2}\cdots\frac{1}{a_j}\,, \tag{2.9}$$

where i1, i2,...,ij is a permutation of 1, 2, ...,j and

$\sum_P$ represents sum over all permutations. Using this result we can now write the desired scattering amplitude as

$$M = \left[\frac{ig^2}{(2\pi)^4}\right]^n \int \prod_{\substack{i=1 \\ (\neq r)}}^{n+1} \frac{d^4k_i}{k_i^2-\mu^2}\; M^h\left(q - \sum_{\substack{i=1 \\ (\neq r)}}^{n+1} k_i\right)$$

$$\times \frac{1}{(r-1)!} \prod_{i=1}^{r-1} \frac{1}{k_i^2-2p_a\cdot k_i} \quad \frac{1}{(n+1-r)!} \prod_{j=r+1}^{n+1} \frac{1}{k_j^2+2p_a'\cdot k_j}$$

$$\times \sum_{s'} \left\{ \prod_{i=1}^{\ell} \frac{1}{k_i'^2+2p_b\cdot k_i'} \prod_{j=\ell+1}^{r-1} \frac{1}{k_j'^2-2p_b'\cdot k_j'} \right\}$$

$$\times \sum_{s''} \left\{ \prod_{i=1}^{m} \frac{1}{k_i''^2+2p_b\cdot k_i''} \prod_{j=m+1}^{n+1-r} \frac{1}{k_j''^2-2p_b'\cdot k_j''} \right\}. \qquad (2.10)$$

The division by $(r-1)!$ and $(n+1-r)!$ above is to compensate for the overcounting introduced by summing over all possible permutations. Introducing the Fourier transform of $M^h(k_r)$:

$$M^h(k_r) = \int d^4x\; e^{-ik_r\cdot x}\; M^h(x), \qquad (2.11)$$

Eq. (2.10) can be written in the form

$$M = i^n \int d^4x\; e^{-iq\cdot x}\; M^h(x)\; \frac{1}{(r-1)!} \sum_{s'} (U_1)^{\ell}\; (U_2)^{r-1-\ell}$$

$$\times \frac{1}{(n+1-r)!} \sum_{s''} (U_3)^m\; (U_4)^{n+1-r-m}, \qquad (2.12)$$

where the U's are defined by

$$U_1 = U(x;p_a,p_b) \quad , \quad U_2 = U(x;p_a,-p_b') \ ,$$

$$U_3 = U(x;-p_a',p_b) \quad , \quad U_4 = U(x;-p_a',-p_b') ;$$

$$U(x;p,p') = \frac{g^2}{(2\pi)^4}\int \frac{d^4k\ e^{ik\cdot x}}{k^2-\mu^2+i\epsilon}\ \frac{1}{k^2-2p\cdot k+i\epsilon}\ \frac{1}{k^2+2p'\cdot k+i\epsilon}. \tag{2.13}$$

We note that the sums over s′ and s″ introduce the following factors:

$$\sum_{s'} = \binom{r-1}{\ell} \quad , \quad \sum_{s''} = \binom{n+1-r}{m} \quad ,$$

so that Eq. (2.12) becomes

$$M^{n+1,r}_{\ell,m} = i^n \int d^4x\ e^{-iq\cdot x}\ M^h(x)\ \frac{(U_1)^\ell}{\ell!}\ \frac{(U_2)^{r-1-\ell}}{(r-1-\ell)!}$$

$$\times \frac{(U_3)^m}{m!}\ \frac{(U_4)^{n+1-r-m}}{(n+1-r-m)!} \ . \tag{2.14}$$

The superscripts and the subscripts for the amplitude are put in to indicate (i) that it corresponds to the exchange of n soft mesons and one hard interaction, (ii) that the hard interaction occurs after r-1 soft mesons have been emitted by the p_a line, (iii) that ℓ mesons emitted from the p_a line and m mesons from the p_a' line are absorbed by the p_b line. We now sum over all the possible values of ℓ and m for fixed n and r. This leads to the amplitude

$$M^{n+1,r} = \sum_{\ell=0}^{r-1} \sum_{m=0}^{n+1-r} M^{n+1,r}_{\ell,m}$$

$$= i^n \int d^4x\, e^{-iq\cdot x}\, M^h(x)\, \frac{(U_1+U_2)^{r-1}}{(r-1)!}\, \frac{(U_3+U_4)^{n+1-r}}{(n+1-r)!} \qquad (2.15)$$

In Eq. (2.15) the hard interaction has been taken to occur at the r-th position. To obtain the scattering amplitude for n soft exchange and one hard interaction we want to sum (2.15) over all values of r. However, at this point we have to distinguish between two cases: (i) the hard interaction is the same process as the soft exchange; (ii) the hard interaction is different from the soft exchange. In the first case, the sum over r has to be divided by n+1, since all diagrams with different values of r are topologically the same Feynman diagram.[18] In the second case, diagrams with different values of r are different, and so no division by n+1 is required. We thus obtain

$$M^{n+1} = \int d^4x\, e^{-iq\cdot x}\, M^h(x)\, \frac{(i\chi)^n}{(n+1)!} \qquad \text{for case (i)}, \qquad (2.16a)$$

$$= \int d^4x\, e^{-iq\cdot x}\, M^h(x)\, \frac{(i\chi)^n}{n!} \qquad \text{for case (ii)}, \qquad (2.16b)$$

where

$$\chi = \chi(x) = U_1 + U_2 + U_3 + U_4 \qquad (2.17)$$

This leads to two different results for the total scattering amplitude:

$$M = \sum_{n=0}^{\infty} M^{n+1} = \int d^4x\, e^{-iq\cdot x}\, M^h(x)\, \frac{e^{i\chi}-1}{i\chi}\,, \qquad \text{(case i)} \qquad (2.18a)$$

$$= \int d^4x \; e^{-iq\cdot x} \; M^h(x) \; e^{i\chi}. \quad \text{(case ii)} \qquad (2.18b)$$

Eq. (2.18a) is the formula derived by Lévy and Sucher[6] which holds if the hard interaction is the same as the soft interaction, while Eq. (2.18b) is the formula that corresponds to Schiff s large-angle result[19] and holds when the hard interaction is different from the soft interaction.[20]

We should notice that Eq. (2.18a) is exact in fourth order (i.e., exchange of two mesons) since the approximation of dropping $k_i k_j$ $(i \neq j)$ terms do not come in till we go to sixth order.[j] On the other hand, if Eq. (2.18b) is used in the same order the amplitude will be larger by a factor of 2. This result has been noted by Lévy and Sucher in the asymptotic limit $s \to \infty$, $|t|$ fixed.[21,22] Another point worth commenting on is that, contrary to popular belief, the eikonal description can be used for large momentum transfer scattering. Suppose we know about a mechanism that can carry off large momentum transfer. Now, notice that in the derivation of Eq. (2.18b) no restriction was made on the magnitude of k_r, the requirement was that the k_i's are small $(i \neq r)$. Therefore, (2.18b) retains its validity for large q. This indicates that the analysis of large-momentum transfer scattering in terms of Fig. 2 and Eq. (2.18b) will provide insight into interactions at small distances. In this context we should note that Eq. (2.18b) essentially corresponds to the distorted wave Born approximation or absorptive correction formula, if $\chi(x)$ becomes a function of the impact parameter only.

2(b). Radiative Corrections

We shall investigate radiative corrections to all orders due to soft mesons. These corrections are of two types: (i) vertex type where a soft meson emitted from an external line is absorbed by it after the hard interaction; (ii) self-energy type where the soft-meson emitted by an external line is absorbed by it before the hard interaction. A typical diagram to be considered is shown in Fig. 4. For simplicity we shall first discuss

vertex type radiative corrections and ignore self-energy corrections.

Let us begin with Fig. 2 and make a vertex type insertion connecting p_a line and p_a' line by a meson line. Let q_1 be the 4-momentum of this meson line. Then apart from the obvious modification of the propagator products $\mathcal{P}_a$ and $\mathcal{P}_a'$, the Feynman amplitude will be multiplied by the factor

$$\frac{i}{(2\pi)^4} \quad g^2 \int \frac{d^4 q_1}{q_1^2 - \mu^2 + i\epsilon} \quad .$$

We now sum over all possible ways of attaching q_1 to p_a' line and then sum over all possible ways of attaching it to p_a line. Because of the factorization of the q_1 propagator part arising from the identity (2.7), the net result is multiplying the original Feynman amplitude of Fig. 2 by the factor

$$\frac{i}{(2\pi)^4} \quad g^2 \int \frac{d^4 q_1}{q_1^2 - \mu^2 + i\epsilon} \frac{1}{q_1^2 - 2p_a \cdot q_1 + i\epsilon} \frac{1}{q_1^2 - 2p_a' \cdot q_1 + i\epsilon} \quad .$$

This is the radiative correction due to a single vertex type insertion. Notice that this factor does not depend on x since q_1 does not enter into the equation for k_r.

Next consider two vertex insertions, and let q_1 and q_2 be the 4-momenta of the meson lines connecting p_a and p_a' lines. Keeping the initial positions and q_1 and q_2 fixed, we first sum over all possible ways of attaching them to the p_a' line. Then we sum over their initial positions along the p_a line. The result is to multiply the original Feynman amplitude by the factor

$$\frac{i}{(2\pi)^4} \quad g^2 \int \frac{d^4 q_1}{q_1^2 - \mu^2 + i\epsilon} \quad \frac{1}{q_1^2 - 2p_a \cdot q_1 + i\epsilon} \quad \frac{1}{q_1^2 - 2p_a' \cdot q_1 + i\epsilon}$$

$$\times \frac{i}{(2\pi)^4} \quad g^2 \int \frac{d^4 q_2}{q_2^2 - \mu^2 + i\epsilon} \quad \frac{1}{q_2^2 - 2p_a \cdot q_2 + i\epsilon} \quad \frac{1}{q_2^2 - 2p_a' \cdot q_2 + i\epsilon} \quad .$$

This factor however does not represent correctly the effect of two vertex type insertions since contributions from all diagrams with q_1 and q_2 interchanged have been counted, while topologically they represent the same Feynman diagram. To compensate for this we have to divide the above factor by 2!. Thus the radiative correction due to two vertex insertions connecting p_a and p_a' lines is to multiply the Feynman amplitude of Fig. 2 by the factor $\frac{1}{2!}(iV_1)^2$, where

$$V_1 = \frac{g^2}{(2\pi)^4}\int \frac{d^4k}{k^2-\mu^2+i\epsilon}\,\frac{1}{k^2-2p_a\cdot k+i\epsilon}\,\frac{1}{k^2-2p_a'\cdot k+i\epsilon} \qquad (2.19)$$

The generalization of this result when there are n lines connecting p_a and p_a' is $\frac{1}{n!}(iV_1)^n$. Therefore, if we sum over all these vertex type radiative corrections the result will be multiplying the original Feynman amplitude by the factor e^{iV_1}. Similarly, the sum over all vertex type insertions connecting p_b and p'_b will be e^{iV_2}, where V_2 is obtained from V_1 by replacing p_a by p_b and p_a' by p_b'.

We now want to find how the scattering amplitude is modified due to self-energy insertions. To this end let us consider the external line p_a and make a self-energy insertion (Fig. 5). The propagator product along p_a line will be modified in the usual way and the scattering amplitude will be multiplied by the factor

$$i\,\frac{g^2}{(2\pi)^4}\int \frac{d^4q_1}{q_1^2-\mu^2+i\epsilon}\,.$$

We now regard both ends of the radiative correction as corresponding to emission of 4-momenta, q_1 and q_1', so that $q_1' = -q_1$. Keeping the q_1 end fixed, we sum over all possible ways in which q_1' can be attached to the p_a line. Then we sum over all possible ways in which q_1 can be attached to the p_a line. The net result will be multiplying the original Feynman amplitude by the factor

$$\frac{ig^2}{(2\pi)^4}\int \frac{d^4q_1}{q_1^2-\mu^2+i\epsilon}\,\frac{1}{q_1'^2-2p_a\cdot q_1'+i\epsilon}\,\frac{1}{q_1^2-2p_a\cdot q_1+i\epsilon}\,.$$

The radiative correction due to one self-energy insertion is then the above multiplicative factor divided by 2. The division by 2 is necessary, since the same Feynman diagram has been considered twice while summing over both ends.

Next we consider the case where two self-energy insertions are made to the p_a line. As before both ends of each self-energy insertions are taken as corresponding to emission of 4-momenta, so that we have $q_1'=-q_1$, $q_2'=-q_2$. We sum over all possible positions of q_2' end and then over q_2 end. The same thing is repeated for q'_1 and q_1 ends. The result is that the sum over all radiative corrections due to two self-energy insertions to the p_a line is the multiplicative factor

$$\frac{1}{2!}\,\frac{1}{2}\,\frac{ig^2}{(2\pi)^4}\int\frac{d^4q_1}{q_1^2-\mu^2+i\epsilon}\quad\frac{1}{q_1^2+2p_a\cdot q_1+i\epsilon}\quad\frac{1}{q_1^2-2p_a\cdot q_1+i\epsilon}$$

$$\times\,\frac{1}{2}\,\frac{ig^2}{(2\pi)^4}\int\frac{d^4q_2}{q_2^2-\mu^2+i\epsilon}\quad\frac{1}{q_2^2+2p_a\cdot q_2+i\epsilon}\quad\frac{1}{q_2^2-2p_a\cdot q_2+i\epsilon}.$$

The factor 1/2! included above is to take into account the fact that diagrams with q_1 and q_2 interchanged correspond to the same Feynman diagram. Generalization of the above result is now straightforward. If there are n_a number of self-energy insertions to the p_a line, then the Feynman amplitude is to be multiplied by

$$\frac{1}{n_a!}\left(\frac{1}{2}\,iS_1\right)^{n_a},$$

where

$$S_1=\frac{g^2}{(2\pi)^4}\int\frac{d^4k}{k^2-\mu^2+i\epsilon}\quad\frac{1}{k^2-2p_a\cdot k+i\epsilon}\,\frac{1}{k^2+2p_a\cdot k+i\epsilon}. \tag{2.20}$$

If self-energy corrections to all orders are taken into account, then this means multiplying the scattering amplitude by the factor $\exp(\frac{1}{2}\,iS_1)$.

The above arguments can be equally applied for the self-energy corrections to the other external lines. Hence the net effect of including self-energy and vertex type radiative corrections to all orders is to modify the amplitude (2.18a) to

$$M = e^{i\left(\sum_{i=1}^{2} V_i + \frac{1}{2}\sum_{i=1}^{4} S_i\right)} \int d^4x\, e^{-iq\cdot x}\, M^h(x)\, \frac{e^{i\chi}-1}{i\chi}\,, \tag{2.21a}$$

and the amplitude (2.18b) to

$$M = e^{i\left(\sum_{i=1}^{2} V_i + \frac{1}{2}\sum_{i=1}^{4} S_i\right)} \int d^4x\, e^{-iq\cdot x}\, M^h(x)\, e^{i\chi}\,. \tag{2.21b}$$

S_2, S_3, S_4 in the above equations are obtained from S_1 by replacing p_a by p_a', p_b and p_b' respectively. It is worthwhile to note that the radiative corrections do not depend on x and therefore factor out without any new approximations. This factorization has also been obtained by Barbashov et al.[13] using functional integration technique.

2(c). Nucleon-Nucleon Scattering via Vector Meson Exchange

We now consider particles a and b as spin ½ nucleons, and the meson they exchange to be a massive neutral vector meson. Physically this case is more realistic than the scalar case discussed earlier in section 2(a). The reason is in the scalar case when $s \to \infty$ the Born amplitude dominates,[23] so that our summing of an infinite set of diagrams is of no practical value. On the other hand, in the present case the eikonal amplitude for $s \to \infty$ does retain contributions from all higher order diagrams, and shows that the eikonal description provides in a compact form the sum of an infinite set of diagrams.

Let us first define the soft-meson approximation in the present case. We shall assume the vector meson to be coupled to a conserved source, so that in the propagator

of the vector meson the $k_\mu k_\nu/m^2$ term can be dropped. Examining Fig. 2 we note that, because of the fermion nature of the particles and the vector nature of the meson, the propagator part of the p_a line will now be of the form

$$\cdots\gamma_{\mu_3}\frac{\not p_a-\not k_1-\not k_2+m}{(p_a-k_1-k_2)^2-m^2+i\epsilon}\gamma_{\mu_2}\frac{\not p_a-\not k_1+m}{(p_a-k_1)^2-m^2+i\epsilon}\gamma_{\mu_1}u(p_a).$$

The denominators are exactly the same as we encountered in the scalar case, and the same approximation, namely, dropping k_ik_j $(i\neq j)$ terms will be made for them. As for the numerators, we approximate them by neglecting meson 4-momenta compared to the 4-momenta of the external particles. The rationale like before is that at high energy the 4-momentum of an external particle is very large, while the dominating contribution to the scattering amplitude comes from small values of the meson 4-momenta. The above approximation has the following consequences:

$$\gamma_{\mu_2}\frac{\not p_a-\not k_1+m}{(p_a-k_1)^2-m^2+i\epsilon}\gamma_{\mu_1}u(p_a)\simeq\gamma_{\mu_2}\frac{\not p_a+m}{k_1^2-2p_a\cdot k_1+i\epsilon}\gamma_{\mu_1}u(p_a)$$

$$=\gamma_{\mu_2}u(p_a)\frac{(2p_a)_{\mu_1}}{k_1^2-2p_a\cdot k_1+i\epsilon}, \qquad (2.22)$$

and

$$\cdots\gamma_{\mu_3}\frac{\not p_a-\not k_1-\not k_2+m}{(p_a-k_1-k_2)^2-m^2+i\epsilon}\gamma_{\mu_2}\frac{\not p_a-\not k_1+m}{(p_a-k_1)^2-m^2+i\epsilon}\gamma_{\mu_1}u(p_a)$$

$$\simeq\cdots\gamma_{\mu_3}u(p_a)\frac{(2p_a)_{\mu_2}}{k_1^2+k_2^2-2p_a\cdot(k_1+k_2)+i\epsilon}\frac{(2p_a)_{\mu_1}}{k_1^2-2p_a\cdot k_1+i\epsilon}. \qquad (2.23)$$

Similar results are obtained for the outgoing particles. Thus for the p_a' line

$$\bar{u}(p_a') \gamma_{\mu_{n+1}} \frac{\not{p}_a' + \not{k}_{n+1} + m}{(p_a' + k_{n+1})^2 - m^2} \gamma_{\mu_n} \frac{\not{p}_a' + \not{k}_{n+1} + \not{k}_n + m}{(p_a' + k_{n+1} + k_n)^2 - m^2} \gamma_{\mu_{n-1}} \cdots$$

$$\simeq \frac{(2p_a')_{\mu_{n+1}}}{k_{n+1}^2 + 2p_a' \cdot k_{n+1}} \frac{(2p_a')_{\mu_n}}{k_{n+1}^2 + k_n^2 + 2p_a' \cdot (k_{n+1} + k_n)} \bar{u}(p_a') \gamma_{\mu_{n-1}} \cdots \tag{2.24}$$

Eqs. (2.23) and (2.24) show that all the γ_μ's along any external line can be replaced by the corresponding components of the external 4-momentum vector. This leads to the result that if a meson is attached to two external lines, say p_1 and p_2, then we have a factor $4p_1 \cdot p_2$ in the amplitude. As seen above the denominators along the external lines are precisely the same as those in the case of scalar particles (sec. 2(a)). Therefore, the results obtained there regarding sum over all final positions and average over all initial permutations, can be used here without any change. Thus the Feynman amplitude in the present case corresponding to Fig. 2 and Eq. (2.14) is

$$M_{\ell,m}^{n+1,r} = (-i)^n \int d^4x \, e^{-iq\cdot x} \, \bar{u}(p_b') \, \bar{u}(p_a') \, M^h(x) \, u(p_a) u(p_b)$$

$$\times \frac{(4p_a \cdot p_b U_1)^\ell}{\ell!} \frac{(4p_a \cdot p_b' U_2)^{r-1-\ell}}{(r-1-\ell)!}$$

$$\times \frac{(4p_a' \cdot p_b U_3)^m}{m!} \frac{(4p_a' \cdot p_b' U_4)^{n+1-r-m}}{(n+1-r-m)!} \, . \tag{2.25}$$

The extra $(-1)^n$ factor in Eq. (2.25) arises because the vector meson propagator has a negative sign relative to the scalar propagator. M^h as before represents the hard interaction occurring at the r-th position along a line.

From Eq. (2.25) we find that the scattering amplitude for the sum over all diagrams of the type Fig. 2 is

$$M = \int d^4x\, e^{-iq\cdot x}\, \bar{u}(p_b')\bar{u}(p_a')\, M^h(x)\, u(p_a)u(p_b) \times \frac{e^{i\chi(x)}-1}{i\chi(x)} \quad , \tag{2.26a}$$

if the hard interaction corresponds to the soft exchange, i.e.,

$$M^h(k_r) = g^2 \frac{\gamma_\mu\, g_{\mu\nu}\, \gamma_\nu}{k_r^2-\mu^2+i\epsilon} \quad .$$

On the other hand, the sum is

$$M = \int d^4x\, e^{-iq\cdot x}\, \bar{u}(p_b')\bar{u}(p_a')\, M^h(x)\, u(p_a)u(p_b)\, e^{i\chi(x)} \quad , \tag{2.26b}$$

if the hard interaction is different from the soft exchange. The function $\chi(x)$ in (2.26a,b) is given by

$$\chi(x) = -4p_a\cdot p_b U_1 - 4p_a\cdot p'_b U_2 - 4p'_a\cdot p_b U_3 - 4p'_a\cdot p'_b U_4 \quad , \tag{2.27}$$

where the U's are as defined in section 2(a) earlier.

Radiative corrections to all orders for the present case can now be easily determined from the corresponding result for the spin-zero case. Namely, the effect of taking into account all radiative corrections (both vertex and self-energy types) will be to modify the amplitude (2.26a) to

$$M = e^{-\frac{1}{2}A} \int d^4x\, e^{-iq\cdot x}\, \bar{u}(p_b')\ \ \bar{u}(p_a')\ M^h(x) u(p_a) u(p_b) \times \frac{e^{i\chi(x)}-1}{i\chi(x)}\ , \qquad (2.28a)$$

and (2.26b) to

$$M = e^{-\frac{1}{2}A} \int d^4x\, e^{-iq\cdot x}\, \bar{u}(p_b')\bar{u}(p_a')\ M^h(x) u(p_a) u(p_b)\ e^{i\chi(x)}\ , \qquad (2.28b)$$

where

$$-\tfrac{1}{2}A = -i\left[4p_a\cdot p_a' V_1 + 4p_b\cdot p_b' V_2 + 2p_a\cdot p_a S_1 + 2p_a'\cdot p'_a S_2 + 2p_b\cdot p_b S_3 + 2p_b'\cdot p_b' S_4\right]. \qquad (2.29)$$

It is worthwhile to note the following implication of Eq. (2.28b) in quantum electrodynamics. Since $\mu^2 \to 0$ in QED the integrals V_1, S_1, S_2 etc. become infrared divergent. To examine this divergence let us introduce a cut-off Λ, and drop k^2 terms compared to $p_a.k$ and $p_a'.k$ in the denominators of the above integrals. Then

$$-i\left[4p_a\cdot p_a' V_1 + 2p_a\cdot p_a S_1 + 2p_a'\cdot p_a' S_2\right] = \frac{ig^2}{2(2\pi)^4} \int^{\Lambda} \frac{d^4k}{k^2+i\epsilon} \left[\frac{p_a}{p_a\cdot k} - \frac{p_a'}{p_a'\cdot k}\right]^2\ , \qquad (2.30)$$

so that when $p_a' \to p_a$ the right-hand side vanishes. This means that the infrared divergence due to vertex corrections and self-energy corrections cancel exactly for forward scattering. In QED this indeed should happen because of Ward identity.

Let us at this point consider a possible approximation of Eq. (2.28b). Suppose the hard interaction is such that only very small values of x are important for the factor $e^{i\chi(x)}$. We may then put x=0 in the argument of χ and obtain

$$M \simeq e^{-\frac{1}{2}A+i\chi(0)}\ \bar{u}(p_b')\bar{u}(p_a')\ M^h(q)u(p_a)u(p_b). \qquad (2.31)$$

The above approximation is equivalent to saying that $M^h(k_r)\simeq M^h(q)$. In this approximation the contribution of all soft vector meson exchanges becomes a known exponential factor that multiplies the unknown hard interaction. This is precisely the result obtained in the infrared divergence analysis of QED.[10] There the approximation of replacing $\chi(x)$ by $\chi(0)$ can of course be independently justified because the photon mass goes to zero ($\mu \to 0$). The soft meson factor $\exp[-\frac{1}{2}A + i\chi(0)]$ in Eq. (2.31) can provide a rapid fall-off for high energy fixed angle scattering (s,t, u all going to infinity). This has been suggested by Fried and Gaisser[14] to explain large angle pp scattering.

Let us now try to see how our relativistic analysis is connected with the conventional eikonal representation of the scattering amplitude.[24,1] For this purpose we first take the scalar case with the hard interaction same as the soft. Then

$$M = \int d^4x\ e^{-iq\cdot x}\ M^h(x)\ \frac{e^{i\chi(x)}-1}{i\chi(x)}\ , \qquad (2.18a)$$

and

$$\chi(x) = \frac{g^2}{(2\pi)^4}\int \frac{d^4k e^{ik\cdot x}}{k^2-\mu^2+i\epsilon}\left[\frac{1}{k^2-2p_a\cdot k+i\epsilon}+\frac{1}{k^2+2p_a'\cdot k+i\epsilon}\right]$$

$$\times\left[\frac{1}{k^2+2p_b\cdot k+i\epsilon}+\frac{1}{k^2-2p_b'\cdot k+i\epsilon}\right]. \qquad (2.32)$$

We express the external 4-momenta in terms of the average

4-momenta $p=\frac{1}{2}(p_a + p_a')$ and $p'=\frac{1}{2}(p_b + p_b')$. Then in the c.m. system $\vec{p} = -\vec{p}'$, $p_0=p_0'=p_{a^0}$. For $x \neq 0$ we take the limit $|\vec{p}| \simeq p_0 \to \infty$, drop the k^2 and $q\cdot k$ terms in the denominator compared to $p\cdot k$ and $p'\cdot k$, and obtain

$$\chi(x) \simeq -\frac{g^2}{(2\pi)^4}\int \frac{d^4k\, e^{ik\cdot x}}{k^2-\mu^2+i\epsilon}\left[\frac{1}{2p\cdot k-i\epsilon} - \frac{1}{2p\cdot k+i\epsilon}\right]$$

$$\times\left[\frac{1}{2p'\cdot k+i\epsilon} - \frac{1}{2p'\cdot k-i\epsilon}\right]$$

$$= -\frac{g^2}{16\pi^2}\int \frac{d^4k\, e^{ik\cdot x}}{k^2-\mu^2+i\epsilon}\,\delta(p\cdot k)\,\delta(p'\cdot k)$$

$$= \frac{g^2}{32\pi^2}\,\frac{1}{p_0^2}\int \frac{d^2k_\perp}{\underset{\sim}{k}_\perp^2+\mu^2}\, e^{-i\underset{\sim}{k}_\perp\cdot\underset{\sim}{x}_\perp}$$

$$= \frac{g^2}{4\pi s}\, K_0(\mu b). \qquad (|\underset{\sim}{x}_\perp|=b) \tag{2.33}$$

$K_0(z)$ is the modified Bessel function of the second kind. Inserting (2.33) in (2.18a) and noting that the momentum transfer vector q is normal to the z-axis (taken along $\vec{p}$), we obtain

$$M = \int d^2x_\perp e^{i\underset{\sim}{q}\cdot\underset{\sim}{x}_\perp}\int dx_0\, dx_3 M^h(x)\left[\frac{e^{i\chi(b)}-1}{i\chi(b)}\right], \tag{2.34}$$

where $\underset{\sim}{q}$, $\underset{\sim}{x}_\perp$ are vectors lying in the xy plane and $\chi(b) = \frac{g^2}{4\pi s}K_0(\mu b)$. Now

$$\int dx_0\, dx_3 M^h(x) = \int dx_0\, dx_3 \frac{1}{(2\pi)^4}\int d^4k\, e^{ik\cdot x}\,\frac{-g^2}{k^2-\mu^2+i\epsilon} =$$

$$= \frac{g^2}{(2\pi)^2} \int d^2k_\perp \frac{e^{-i\underset{\sim}{k}_\perp \cdot \underset{\sim}{x}_\perp}}{k_\perp^2 + \mu^2}$$

$$= \frac{g^2}{2\pi} K_0(\mu b). \qquad (2.35)$$

From (2.34) and (2.35) we get

$$M = -2is \int d^2x_\perp \, e^{i\underset{\sim}{q} \cdot \underset{\sim}{x}_\perp} \left[e^{i\chi(b)} - 1\right]$$

$$= -4\pi is \int_0^\infty bdb \quad J_0(b|\underset{\sim}{q}|) \left[e^{i\chi(b)} - 1\right], \qquad (2.36)$$

which is precisely the conventional eikonal form for the scattering amplitude. We notice further that for $s \to \infty$, $\chi(b) \sim \frac{1}{s}$, so that (2.36) becomes

$$M \simeq 4\pi s \int_0^\infty b \, db \, J_0(b|\underset{\sim}{q}|) \, \chi(b)$$

$$= \frac{g^2}{\underset{\sim}{q}^2 + \mu^2}$$

$$= -\frac{g^2}{q^2 - \mu^2}, \qquad (2.37)$$

which is the Born amplitude.

What happens when we have nucleon-nucleon scattering via vector meson exchange? In this case $\chi(x)$ is given by Eq. (2.27). Now,

$$4p_a \cdot p_b = 2(p_a + p_b)^2 - 4m^2 \simeq 2s \text{ for } s \to \infty.$$

Similarly, $4p_a\cdot p_b' = 4p_a'\cdot p_b = 4p_a'\cdot p_b' = 2s$. Thus $\chi(x)_{\text{vector}} = -2s\,\chi(x)_{\text{scalar}}$. Eq. (2.33) then leads to the result that for vector meson exchange

$$\chi(x) = -\frac{g^2}{2\pi} K_0(\mu b). \tag{2.38}$$

Therefore the scattering amplitude (2.26a) becomes

$$\begin{aligned} M &= \bar{u}(p_b')\,\gamma_\mu u(p_b)\,\bar{u}(p_a')\,\gamma_\mu u(p_a) \\ &\quad \times \int d^4x\, e^{-iq\cdot x}\, \frac{1}{(2\pi)^4}\int \frac{g^2 e^{ik\cdot x}}{k^2-\mu^2+i\epsilon}\, d^4k \left[\frac{e^{i\chi(b)}-1}{i\chi(b)}\right] \\ &= -i\,\bar{u}(p_b')\,\gamma_\mu u(p_b)\,\bar{u}(p_a')\gamma_\mu u(p_a) \\ &\quad \times \int d^2x_\perp\, e^{i\underset{\sim}{q}\cdot\underset{\sim}{x}_\perp}\left[e^{i\chi(b)}-1\right], \end{aligned} \tag{2.39}$$

where $\chi(b) = -\frac{g^2}{2\pi} K_0(\mu b)$. Let us now use the following helicity representation of the spinors:[25]

$$u(\lambda_a,p_a) \simeq \left(\frac{w}{2m}\right)^{\frac{1}{2}}\begin{pmatrix}1\\ 2\lambda_a\end{pmatrix} e^{i\theta\sigma_2/4}\,\chi_{\lambda_a},\quad \bar{u}(\lambda_a',p_a')$$

$$\simeq \left(\frac{w}{2m}\right)^{\frac{1}{2}}\chi^\dagger_{\lambda_a'}\, e^{\frac{i\theta\sigma_2}{4}}\,(1,-2\lambda_a'), \tag{2.40a}$$

$$u(\lambda_b,p_b) \simeq \left(\frac{w}{2m}\right)^{\frac{1}{2}}\begin{pmatrix}1\\ 2\lambda_b\end{pmatrix} e^{i\theta\sigma_2/4}\,\chi_{-\lambda_b},\quad \bar{u}(\lambda_b',p_b')$$

$$\simeq \left(\frac{w}{2m}\right)^{\frac{1}{2}}\chi^\dagger_{-\lambda_b'}\, e^{\frac{i\theta\sigma_2}{4}}\,(1,-2\lambda_b'); \tag{2.40b}$$

here $\lambda_a(\lambda_b)$ and $\lambda_a'(\lambda_b')$ are the initial and final helicities of particle a (b); θ is the c.m. scattering angle; $\chi_{\frac{1}{2}} = \begin{pmatrix}1\\0\end{pmatrix}$, $\chi_{-\frac{1}{2}} = \begin{pmatrix}0\\1\end{pmatrix}$ and $w = (p_a)_0$. From (2.39) and (2.40) we then obtain for $s \to \infty$ and t fixed

$$M = -i\frac{s}{2m^2}\delta_{\lambda_a'\lambda_a}\delta_{\lambda_b'\lambda_b}\int d^2x_\perp e^{i\underset{\sim}{q}\cdot\underset{\sim}{x}_\perp}\left[e^{i\chi(b)}-1\right]. \qquad (2.41)$$

A number of features of Eq. (2.41) are worth noticing. (1) It shows that the high-energy momentum-transfer amplitude for spin -½ scattering via neutral soft vector meson exchanges is a helicity conserving amplitude. (2) Since Im $M(t=0) \propto s\,\sigma_{tot}$, Eq. (2.41) predicts a constant asymptotic total cross section. (3) As the eikonal $\chi(b)$ is completely real, the inelastic cross section vanishes at high energy. This of course represents an unrealistic feature of the model. (4) Since $\chi(b) = -\frac{g^2}{2\pi} K_0(\mu b)$ is energy independent, diagrams of all orders in coupling constant contribute to the high energy behavior of the full amplitude. This may reflect an important feature of strong interactions where perturbative calculations are not very meaningful.

3. Regge-Eikonal Model

The Regge-eikonal model corresponds to taking the "Born amplitude" or the single scattering amplitude as that due to the exchange of a single Regge pole, and then generating the full amplitude by means of the eikonal representation. Let us now explore the consequences of such a model.[26] The full elastic amplitude[27] has the eikonal representation[1]

$$T(s,t) = ikW\int_0^\infty b\,db\,J_0(b\Delta)\left[1 - e^{i\chi(s,b)}\right], \quad (t=-\Delta^2) \qquad (3.1)$$

where the eikonal $\chi(s,b)$ is related to the single scattering amplitude $T_1(s,t)$ by

$$\chi(s,b) = \frac{1}{k\,W}\int_0^\infty \Delta d\Delta\,J_0(b\Delta)\,T_1(s,t). \qquad (3.2)$$

If $T_1(s,t)$ corresponds to a single Regge pole exchange amplitude, then for s large it has the form

$$T_1(s,t) = b(t)\ \zeta(t)\left(\frac{s}{s_o}\right)^{\alpha(t)}\ ; \tag{3.3}$$

$\zeta(t)$ is the signature factor and $b(t)$ is proportional to the Regge residue. We assume for simplicity an even signatured pole throughout, so that

$$\zeta(t) = \frac{1 + e^{-i\pi\alpha(t)}}{\sin \pi\alpha(t)} = \frac{e^{-i\frac{\pi}{2}\alpha(t)}}{\sin \frac{\pi}{2}\alpha(t)}\ .$$

The n-th term in Eq. (3.1) obtained by expanding the exponential is

$$T_n(s,t) = -\ (i)^{n+1}\frac{kW}{n!}\int_o^\infty b\ db\ J_o(b\Delta)\ \chi^n(s,b), \tag{3.4}$$

which we refer to as the n-th multiple scattering term. Notice that the full amplitude is simply a sum over all the multiple scattering terms:

$$T(s,t) = \sum_{n=1}^{\infty} T_n(s,t). \tag{3.5}$$

When the single scattering term is given by (3.3), the n-th multiple scattering term becomes, using Eq. (3.2),

$$T_n(s,t) = -\tfrac{1}{2}(i)^{n+1}\frac{s}{n!}\int_o^\infty b\ db\ J_o(b\Delta)\frac{1}{s^n}$$

$$\times \prod_{i=1}^{n}\int_{-\infty}^{\infty} dt_i J_o(b\Delta_i)b(t_i)\zeta(t_i)\left(\frac{s}{s_o}\right)^{\alpha(t_i)}\ , \tag{3.6}$$

where $t_i = -\Delta_i^2$.

Let us write

$$\tau^{(n)}(t;t_1,t_2,\ldots,t_n) \equiv \int_0^\infty b\, db\, J_0(b\Delta) \prod_{i=1}^{n} J_0(b\Delta_i). \tag{3.7}$$

Then,

$$T_n(s,t) = \frac{i^{n-1}}{2n!s_0^{n-1}} \int dt_1\, dt_2 \ldots dt_n\, \tau^{(n)}(t;t_1,t_2,\ldots,t_n)$$

$$\times \left(\frac{s}{s_0}\right)^{\sum_i \alpha(t_i)-n+1} \prod_{i=1}^{n} b(t_i)\, \zeta(t_i). \tag{3.8}$$

Eq. (3.8) can be written in the form

$$T_n(s,t) = \int dj \left(\frac{s}{s_0}\right)^j \zeta(j)\, \rho_n(t;j), \tag{3.9}$$

where

$$\zeta(j) = \frac{e^{-i\frac{\pi}{2}j}}{\sin\frac{\pi}{2}j}$$

and

$$\rho_n(t;j) = \frac{1}{2n!s_0^{n-1}} \sin\tfrac{\pi}{2}j \int dt_1\, dt_2 \ldots dt_n\, \tau^{(n)}(t;t_1,\ldots,t_n)$$

$$\times \delta\left(j-\sum_i \alpha(t_i) + n-1\right) \prod_{i=1}^{n} \frac{b(t_i)}{\sin\frac{\pi}{2}\alpha(t_i)}. \tag{3.10}$$

$\rho_n(t;j)$ given above can be expressed as integrals over two dimensional vectors:

$$\rho_n(t;j) = \frac{1}{n!\, s_0^{n-1}\pi^{n-1}} \sin\frac{\pi}{2}j \int d\underset{\sim}{\Delta}_1 d\underset{\sim}{\Delta}_2 \ldots d\underset{\sim}{\Delta}_n\, \delta^2\Big(\underset{\sim}{\Delta} - \sum_{i=1}^{n}\underset{\sim}{\Delta}_i\Big)$$

$$\times\, \delta\Big(j - \sum_i \alpha(t_i) + n-1\Big) \prod_{i=1}^{n} \frac{b(-\Delta_i^2)}{\sin\frac{\pi}{2}\alpha(-\Delta_i^2)} \,. \tag{3.11}$$

This result is easily obtained by using the representation

$$J_0(b\Delta) = \frac{1}{2\pi}\int_0^{2\pi} d\Phi\; e^{\pm ib\cdot\underset{\sim}{\Delta}}$$

in (3.7). The upper limit of j-integration in (3.9) is determined by the maximum value of $\sum_i \alpha(t_i)$ because of the δ-function in (3.10). However, the integrand $\rho_n(t;j)$ vanishes unless $\sum_{i=1}^{n}\Delta_i \geq \Delta$. This follows from the fact that $\tau^{(n)}$ given by (3.7) vanishes for $\Delta > \sum_{i=1}^{n}\Delta_i$ $(n\geq 2)$.[28] Therefore, to determine the upper limit of integration in (3.9) we have to maximize $\sum_i \alpha(t_i)$ subject to the condition that $(\sum_i \Delta_i - \Delta)$ is positive.

Let us define

$$\chi_n \equiv \sum_{i=1}^{n}\Delta_i - \Delta = \sum_{i=1}^{n}(-t_i)^{\frac{1}{2}} - (-t)^{\frac{1}{2}}. \tag{3.12}$$

We want to maximize $\sum_{i=1}^{n}\alpha(t_i)$ keeping χ_n fixed at some positive value, say c. The problem is equivalent to maximizing the function

$$\Phi_n = \sum_{i=1}^{n}\alpha(t_i) - \lambda\chi_n$$

where λ is an undetermined Lagrange multiplier. This

gives

$$\frac{d\alpha(t_i)}{dt_i} = \lambda \frac{\partial \chi_n}{\partial t_i}$$

$$= - \frac{\lambda}{2(-t_i)^{\frac{1}{2}}} \quad . \tag{3.13}$$

We notice that the same equation is satisfied for all values of i (i = 1,2,...,n). Therefore, if t_o is the solution of Eq. (3.13), then $t_i = t_o$ for all i. From (3.12) we then obtain

$$\chi_n = c = n(-t_o)^{\frac{1}{2}} - (-t)^{\frac{1}{2}}$$

or

$$t_o = - \frac{(c+\Delta)^2}{n^2} \quad . \tag{3.14}$$

Therefore, for $\chi_n = c$,

$$\Big[\sum_i \alpha(t_i) \Big]_{max} = n\, \alpha(t_o) = n\alpha \left(- \frac{(c+\Delta)^2}{n^2} \right) . \tag{3.15}$$

Now, the smallest value of χ_n is c = 0. Since $\alpha(t)$ is considered to be a monotonically increasing function of t, the value given by (3.15) increases as c decreases. Hence the maximum possible value of the left-hand side of (3.15) is given by

$$\Big[\sum_i \alpha(t_i) \Big]_{max} = n\alpha \left(- \frac{\Delta^2}{n^2} \right). \tag{3.16}$$

This in turn gives the upper limit of integration in (3.9)

to be

$$j_{max}(t) = n\alpha\left(\frac{t}{n^2}\right) - n + 1. \tag{3.17}$$

Let us at this point try to determine the asymptotic contribution to the amplitude T(s,t) due to a branch cut in the complex angular momentum plane. To this end we first write T(s,t) as a countour integral (Watson-Sommerfeld transform) in the j-plane:

$$T(s,t) = -\frac{1}{4i}\int_C dj(2j+1)\,\frac{\left[P_j(-z_t)+P_j(z_t)\right]}{\sin \pi j}\,a(j,t); \tag{3.18}$$

a(j,t) is the analytically continued even signatured partial wave amplitude.[29] (For simplicity we have kept only the even partial waves). The contour C encloses all the poles at even integers on the real axis (Fig. 6). Next we distort the contour C to a line parallel to the imaginary axis and passing through $j = -\frac{1}{2}$. We thus pick up the contributions due to the Regge poles. If now there is a branch point in the j-plane with the branch cut going left as in Fig. 6, the contribution due to this cut to T(s,t) will be of the form

$$T_{cut}(s,t) = \tfrac{1}{2}\int dj^{\alpha_c(t)}\left[\frac{1+e^{-i\pi j}}{\sin\pi j}\right]\left(\frac{s}{s_o}\right)^j \Delta f_j(t), \tag{3.19}$$

where $\Delta f_j(t)$ is proportional to the discontinuity of $a(j,t)$ across this branch cut, and $\alpha_c(t)$ is the position of the branch point. For t negative $\alpha_c(t)$ is real and the branch cut extends along the negative real axis.[30] That cuts in the angular momentum plane should occur when two or more Regge poles are exchanged was first argued by Amati, Fubini and Stanghellini.[31] However, it was Mandelstam[32] who showed from Feynman diagram models with third double spectral functions that such cuts contribute in an important way to the asymptotic behavior of the scattering

amplitude. This was further confirmed by Polkinghorne.[33] The positions of these cuts are given by

$$\alpha_c(t) = n\alpha\left(\frac{t}{n^2}\right) - n+1 \ , \qquad (3.20)$$

when n Regge poles each with trajectory α are exchanged in the t-channel. If we now compare the above expression for the cut contribution with the expression for $T_n(s,t)$ given in Eq. (3.9), we find that they are of the same form. Also, $\alpha_c(t)$ given by (3.20) is identical to the upper limit of integration in (3.9) as given by Eq. (3.17). We may therefore consider $T_n(s,t)$ as a Regge cut contribution arising from the exchange of n Regge poles, provided there are reasons to believe that $\rho_n(t;j)$ in (3.9) can be equated with $\Delta f_j(t)$ in (3.19). Now, the form of the partial wave discontinuity across n-Reggeon exchange cut has been investigated by Gribov, Pomeranchuk and Ter-Martirosyan[34] using multiparticle unitarity conditions analytically continued in the j-plane, and by Gribov[35] using Reggeon diagrams. For $t < 0$, their result is

$$\Delta f_j(t) = (-1)^n \frac{\pi}{s_o^{n-1}} \sin\frac{\pi}{2}j \int d\underset{\sim}{k}_1\, d\underset{\sim}{k}_2 \ldots d\underset{\sim}{k}_n (2\pi)^2 \delta(\underset{\sim}{\Delta} - \sum_i \underset{\sim}{k}_i)$$

$$\times\ \delta\Big(j - \sum_i \alpha(-\underset{\sim}{k}_i^2) + n-1\Big) \frac{N^2_{\alpha(-\underset{\sim}{k}_1^2)\ldots\alpha(-\underset{\sim}{k}_n^2)}}{\prod_{i=1}^{n} \sin\frac{\pi}{2}\alpha(-\underset{\sim}{k}_i^2)} \qquad (3.21)$$

where $N_{\alpha(-\underset{\sim}{k}_1^2)\ldots\alpha(-\underset{\sim}{k}_n^2)}$ is the n-Reggeon production amplitude. Comparing (3.21) with (3.11) we notice that $\rho_n(t;j)$ can indeed be equated with $\Delta f_j(t)$ if $(-1)^n N^2_{\alpha\ldots\alpha}$ can be obtained from the product of n Regge residues. We conclude that the n-th multiple scattering term in the Regge-eikonal model can be interpreted as a Regge-cut contribution with a definite prescription for the discontinuity across the cut.

Let us now investigate some of the consequences of the above interpretation. First we examine the j-plane

structure of the partial wave amplitude when the discontinuity is given by the multiple scattering term. For this purpose consider only the double scattering term (n = 2). Eq. (3.11) gives in this case

$$\Delta f_j(t) = \frac{1}{\pi s_o} \sin\frac{\pi}{2} j \int d\underset{\sim}{\Delta}_1 \; \delta\left(j-\alpha(-\underset{\sim}{\Delta}_1^2)-\alpha(-(\underset{\sim}{\Delta}-\underset{\sim}{\Delta}_1)^2) + 1\right)$$

$$\times \frac{b(-\underset{\sim}{\Delta}_1^2)}{\sin\frac{\pi}{2}\alpha(-\underset{\sim}{\Delta}_1^2)} \quad \frac{b(-(\underset{\sim}{\Delta}-\underset{\sim}{\Delta}_1)^2)}{\sin\frac{\pi}{2}\alpha(-(\underset{\sim}{\Delta}-\underset{\sim}{\Delta}_1)^2)} \tag{3.22}$$

Therefore,

$$\frac{f_j(t)}{\sin\frac{\pi}{2}j} = \frac{1}{\pi}\int^{2\alpha(\frac{t}{4})-1} \frac{dj'}{j'-j} \quad \frac{\Delta f_{j'}(t)}{\sin\frac{\pi}{2}j'}$$

$$= -\frac{1}{\pi^2 s_o} \int d\underset{\sim}{\Delta}_1 \frac{b(-\underset{\sim}{\Delta}_1^2)\; b(-(\underset{\sim}{\Delta}-\underset{\sim}{\Delta}_1)^2)}{j-\alpha(-\underset{\sim}{\Delta}_1^2) - \alpha(-(\underset{\sim}{\Delta}-\underset{\sim}{\Delta}_1)^2)+1}$$

$$\times \frac{1}{\sin\frac{\pi}{2}\alpha(-\underset{\sim}{\Delta}_1^2)} \quad \frac{1}{\sin\frac{\pi}{2}\alpha(-(\underset{\sim}{\Delta}-\underset{\sim}{\Delta}_1)^2)} \; . \tag{3.23}$$

This is in agreement with Reggeon diagram and Feynman diagram models.[35,36] It is worth mentioning that strictly speaking $f_j(t)$ above is not the partial wave amplitude, but essentially the Mellin transform of the scattering amplitude T(s,t). Notice that if the cut contribution is given by Eq. (3.19), then T(s,t) is related to $f_j(t)$ by

$$T(s,t) = -\frac{1}{4i}\int_C dj \; \zeta(j) \left(\frac{s}{s_o}\right)^j f_j(t), \tag{3.24}$$

where C is the contour given in Fig. 6. For $t < 0$, if all the j-plane singularities of $f_j(t)$ are to the left of $j = 0$, then (3.24) can be written as a Mellin transform:

$$T(s,t) = -\frac{1}{4i}\int_{\sigma-i\infty}^{\sigma+i\infty} dj\ \zeta(j)\left(\frac{s}{s_o}\right)^j f_j(t), \qquad (3.25)$$

where σ is determined by the right-most singularity of $f_j(t)$.

Let us now examine the asymptotic behavior of the n-th multiple scattering term when the exchanged Regge pole is the Pomeranchon pole ($\alpha(0) = 1$). The reason for special interest in this case is that all the branch points lie on the right of the pole in the physical scattering region ($t < 0$) and, therefore, the cut contributions should be quite important in determining the high-energy behavior of the full scattering amplitude. Notice that if the Pomeranchuk trajectory is linear,

$$\alpha(t) = 1 + \alpha' t \quad \left(\alpha' = \frac{d\alpha}{dt}\Big|_{t=0} > 0\right),$$

and

$$j_n(t) = n\alpha\left(\frac{t}{n^2}\right) - n + 1$$

$$= 1 + \alpha'\frac{t}{n}, \qquad (3.26)$$

so that the branch points are on the right of the pole ($t < 0$) and for $t = 0$ or $n \to \infty$ they accumulate at the point $j = 1$.

The n-th multiple scattering term for the linear Pomeranchuk trajectory is given by Eq. (3.6) to be

$$T_n(s,t) = -\frac{i2^{n-1}}{n!\, s_o^{n-1}}\left(\frac{s}{s_o}\right)\int \Delta_1 d\Delta_1 \Delta_2 d\Delta_2 \ldots \Delta_n d\Delta_n$$

$$\times \int_o^\infty b db\ J_o(b\Delta) \prod_{i=1}^{n} J_o(b\Delta_i) \quad \times$$

$$\times\, e^{-\alpha' \sum_{i=1}^{n} \Delta_i^2 \ell n\left(\frac{s}{s_o} e^{-i\frac{\pi}{2}}\right)} \prod_{i=1}^{n} \frac{b(-\Delta_i^2)}{\sin\frac{\pi}{2}\alpha(-\Delta_i^2)}. \qquad (3.27)$$

The dominant contribution in Eq. (3.27) for $s \to \infty$ occurs when $\sum_{i=1}^{n} \Delta_i^2$ is minimum and $\sum_{i=1}^{n} \Delta_i = \Delta$, that is, from values of $\Delta_i = \frac{\Delta}{n}$. (We can also arrive at this result by examining Eq. (3.9) which shows that the dominant contribution to the integral comes from the upper limit $j = j_n(t)$ and, therefore, from values of $t_i = \frac{t}{n^2}$). We can now approximate (3.27) by

$$T_n(s,t) = -\frac{i}{n!} \frac{2^{n-1}}{s_o^{n-1}} \left(\frac{s}{s_o}\right) \left[\frac{b\left(-\frac{\Delta^2}{n^2}\right)}{\sin\frac{\pi}{2}\alpha\left(-\frac{\Delta^2}{n^2}\right)}\right]^n$$

$$\times \int_0^\infty b\,db\, J_o(b\Delta) \prod_{i=1}^{n} \int_0^\infty \Delta_i\, d\Delta_i\, J_o(b\Delta_i)\, e^{-\alpha' \Delta_i^2 \ell n\left(\frac{s}{s_o} e^{-i\frac{\pi}{2}}\right)}. \qquad (3.28)$$

The integrals can be easily done and the result is

$$T_n(s,t) = \left(\frac{s}{s_o} e^{-i\frac{\pi}{2}}\right)^{1-\alpha'\frac{\Delta^2}{n}} \frac{1}{\xi^{n-1} n!\, n\, s_o^{n-1}} \left[\frac{b(-\frac{\Delta^2}{n^2})}{\sin\frac{\pi}{2}\alpha(-\frac{\Delta^2}{n^2})}\right]^n. \qquad (3.29)$$

$$\left(\xi = \alpha' \ell n\left(\frac{s}{s_o} e^{-i\frac{\pi}{2}}\right)\right)$$

Using (3.26) we can rewrite (3.29) in the form

$$T_n(s,t) = \left(\frac{s}{s_o}\right)^{j_n(t)} \frac{e^{-i\frac{\pi}{2} j_n(t)}}{\left[\alpha'\left\{\ell n\left(\frac{s}{s_o}\right) - i\frac{\pi}{2}\right\}\right]^{n-1}} \frac{(-1)^n}{n!\, n\, s_o^{n-1}} \left[\frac{-b(-\frac{\Delta^2}{n^2})}{\sin\frac{\pi}{2}\alpha(-\frac{\Delta^2}{n^2})}\right]^n. \qquad (3.30)$$

The reason for extracting a $(-1)^n$ factor is that for small Δ^2 $b(\Delta^2)$ should be negative. This follows from the following argument. The single Pomeranchon term gives us

$$T_1(s,t=0) = -i\left(\frac{s}{s_0}\right) b(0). \tag{3.31}$$

Therefore, when $s \to \infty$,[37]

$$\begin{aligned} \text{Im } T(s,0) &= \text{Im } T_1(s,0) \\ &= -\left(\frac{s}{s_0}\right) b(0). \end{aligned} \tag{3.32}$$

But $\text{Im } T(s,0) = \frac{kW}{4\pi}\sigma_{tot}$ is a positive definite quantity. Hence $b(0)$ is negative. So if $b(-\Delta^2)$ varies say exponentially with Δ^2 $(b(-\Delta^2)=b(0)e^{-a\Delta^2})$, it has to be negative. This equation then exhibits the following properties of the Regge cut terms:

i) Asymptotically the n-Reggeon cut term behaves as

$$\left(\frac{s}{s_0}\right)^{j_n(t)} \qquad \frac{1}{\left[\ell n\left(\frac{s}{s_0}\right)\right]^{n-1}}.$$

Thus the cut contribution is similar to the pole contribution except for the logarithmic factor that depends on n. For $t < 0$ the cut dominates the pole, and asymptotically the phase of the cut is given by $e^{-i\frac{\pi}{2}j_n(t)}$.

ii) The contributions from different cuts arising from Pomeranchon exchanges alternate in sign due to the factor $(-1)^n$ and, therefore, will tend to cancel each other.[38]

The first result implies that the partial wave discontinuity should vanish like $(j-j_n(t))^{n-2}$ near the branch point $(n > 2)$.[39] The proof is simple. Suppose $\rho_n(t;j)$ has no zero at the branch point. Then from Eq. (3.9)

$$T_n(s,t) = \left(\frac{s}{s_o}\right)^{j_n(t)} \int^{j_n(t)} dj \left(\frac{s}{s_o}\right)^{j-j_n(t)} \zeta(j)\rho_n(t;j)$$

$$\simeq \left(\frac{s}{s_o}\right)^{j_n(t)} \zeta(j_n)\rho_n(t;j_n) \frac{1}{\ell n\frac{s}{s_o}}, \qquad (3.33)$$

when $s \to \infty$. But this will contradict our explicit calculation of $T_n(s,t)$ in Eq. (3.30) as far as $\ell n s$ behavior is concerned. If now $\rho_n(t;j)$ has a zero of the order p at the branch point, i.e., $\rho_n(t;j) = (j-j_n)^p \tilde{\rho}_n(t)$ when $j \to j_n$, then instead of (3.33) we obtain

$$T_n(s,t) = \left(\frac{s}{s_o}\right)^{j_n(t)} \zeta(j_n) \tilde{\rho}_n(t) \frac{(-1)^p p!}{(\ell n\frac{s}{s_o})^{p+1}} . \qquad (3.34)$$

Comparing the powers of logarithm in Eqs. (3.30) and (3.34), we get $p + 1 = n - 1$ or $p = n - 2$, which proves our assertion about the vanishing of the partial wave discontinuity. This result again agrees with that of Gribov et al.[34] Furthermore, on the basis of the Reggeon unitarity condition Ansel'm and Dyatlov[40] have concluded that successive Regge cut terms alternate in sign — a result we have already obtained.

Summarizing, we can say that if the single scattering term in the eikonal description is taken to be a Regge pole amplitude, then the n-th multiple scattering term can be regarded as a Regge cut. It leads to the same branch point, the same functional form for the discontinuity and the same high-energy behavior as expected for the Regge cut arising from n-Reggeon exchange. However, the question whether the n-th multiple scattering term provides the full contribution due to the n-Reggeon exchange cut remains open. In this context we should note that a number of authors have found the Regge-eikonal model to yield correctly in $\lambda\Phi^3$ theory the full amplitude obtained by exchanging ladders in all possible ways between two high-energy particles, when the coupling constant is small.[41]

4. Optical Potential in S-Matrix Theory

We shall now try to see how a phenomenological complex energy-dependent optical potential may be interpreted in the language of S-matrix theory. Why are we interested in this question? To answer that let us consider high energy elastic hadron-hadron scattering, in particular pp scattering. Two descriptions have been often used to discuss this scattering. In one, the optical model description, hadrons are pictured as extended objects having certain matter distributions which interact in some way during collision, the strength of the interaction being in general energy-dependent.[42] In the other, the S-matrix description, one draws a Feynman-like diagram indicating how the fundamental process occurs, and one asymptotically evaluates the diagram to determine what type of singularity it represents in the j-plane. In both cases the basic process is iterated to obtain the full amplitude.[43] In the optical model the iteration is done using the eikonal description. This leads to the multiple scattering terms. We have realized that if the iteration is also done in the S-matrix approach by using the eikonal representation, say with the single Regge pole amplitude as input, then the multiple scattering terms are Regge cuts. However, notice that the basic ingredient, i.e., the single scattering term in one description is different from that of the other. So one does not see any dynamical interrelation between them. Models which are in between the two, the hybrid models,[44] combine bodily one part of the optical model with another part of the S-matrix. We are thus faced with the following problem. Is it possible to see how the basic mechanism in one description, say the optical model, can be understood in the language of the other, the S-matrix theory?[45] We shall now study a model in which the basic process can indeed be discussed using both the optical model and the S-matrix descriptions.[46,47] The model further shows wherein lies the advantage of one description over the other and how they fit in together. We hope study of this model will provide valuable insight into other hadron scattering models - where one description is used to the exclusion of the other.

The model I have in mind is a model of high-energy pp elastic scattering which has been studied over a period

of years.[48] Originally the model was formulated as an optical model,[49] and so I shall begin by presenting it as such. What was assumed is that the two protons have finite hadronic matter distributions and that they interact by exchanging a meson. The strength of the interaction was taken to be complex and energy dependent. In other words, we are assuming that the basic interaction between the protons is an optical potential of the form $V(t,r) \sim g(t)u(r)$, where the radial dependence is connected with the hadronic distributions of the protons and $g(t)$ is a phenomenological complex energy-dependent coupling constant (we use t as the square of the c.m. energy in this chapter, and s as the negative square of the momentum transfer.) In terms of the optical Born amplitude, we are assuming it to be

$$A_1(t,s) = - t^{\frac{1}{2}} g(t) \frac{F^2(s)}{m^2 - s} . \qquad (4.1)$$

We may now try to picture this phenomenological amplitude in terms of the diagram in Fig. 7. The two blobs represent the fact that the two protons have finite hadronic matter distributions corresponding to $F^2(s)$ in (4.1). However, we immediately realize that Fig. 7 is not a Feynman diagram, since if it were the coupling constant $g(t)$ should be real. So we are faced with the problem of understanding (4.1) in the S-matrix language.

I shall first give away the answer and then present the technical arguments showing how one arrives at it. The answer is that a Born term like (4.1) corresponds to a Mandestam diagram of the type shown in Fig. 8.[47] The solid lines represent nucleons, the short dashed lines π mesons, the long dashed lines a vector meson, say ω, and the wiggly line represents a Regge trajectory, say P'.[50] The blobs represent $N + \omega \to \pi + N$ amplitudes. To see what is happening physically, let us replace each blob with a nucleon line, and then look at the diagram as in Fig. 9. A high energy nucleon comes in and breaks up into a "core nucleon" and a "cloud pion". The core nucleon interacts with the core nucleon of the other incoming nucleon via the vector meson ω, while the pion interacts with the other pion by exchanging the Regge

trajectory P′. Finally, the nucleons absorb respective cloud pions and emerge as outgoing nucleons. Obviously, Fig. 8 is a generalization of Fig. 9. The form factor that Fig. 8 implies is the ωNN form factor given by the diagram in Fig. 10.

Let us now come to the technical part of showing how a Feynman-like diagram such as Fig. 9 can provide a Born term like (4.1). We shall start following a paper by Rothe.[51] We consider all particles to be equal mass and spinless, and try to evaluate the Mandelstam cut diagram with a Regge pole and a single particle exchange (Fig. 11). That diagrams of this type must occur in nature even when all particles lie on Regge trajectories has been noted by a number of authors.[52] The Feynman amplitude corresponding to this diagram (Fig. 11) is

$$A(s,t) = -i\frac{g^6}{(2\pi)^{12}} \int d^4\xi_1\, d^4k_1\, d^4\eta_1 \quad \frac{1}{\eta_1^2 - m^2 + i\epsilon}$$

$$\times \prod_{i=1}^{4} \frac{1}{\xi_i^2 - m^2 + i\epsilon} \qquad \frac{1}{k_i^2 - m^2 + i\epsilon}$$

$$\times \quad R(\eta_2^2, U; \xi_3^2, \xi_4^2, k_3^2, k_4^2) \quad ; \qquad (4.2)$$

R is the Regge pole amplitude and $U = (\xi_3 + k_3)^2$. Let us introduce the Lorentz invariant variables $s = (p_1 - q_1)^2$, $s' = \eta_2^2$, $s'' = \eta_1^2$, $t = (p_1 + p_2)^2$, $t' = (p_1 + \eta_1)^2$ and $t'' = (p_2 - \eta_1)^2$. We now express the integration over $d^4\eta_1$ in terms of the variables s', s'', t', t'':

$$d^4\eta_1 \simeq \frac{1}{8t}\, \tau(s,s',s'')\, ds'\, ds''\, dt'\, dt'', \qquad (4.3)$$

for $t \to \infty$;[53] $\tau(s,s',s'')$ is given by

$$\tau(s,s',s'') = \frac{\pi}{2} \int_0^\infty bdb\, J_0(b\Delta)\, J_0(b\Delta')\, J_0(b\Delta'') =$$

$$= \frac{\theta(-s^2-s'^2-s''^2+2ss'+2ss''+2s's'')}{[-s^2-s'^2-s''^2+2ss'+2ss''+2s's'']^{\frac{1}{2}}} \quad . \tag{4.4}$$

$$(s=-\Delta^2, s'=-\Delta'^2, s''=-\Delta''^2)$$

Eq. (4.2) now becomes

$$A(s,t) \underset{t\to\infty}{\simeq} -\frac{i}{8}\left(\frac{g}{4\pi^2}\right)^6 \frac{1}{t}\int ds'ds'' \frac{\tau(s,s',s'')}{s''-m^2+i\epsilon}$$

$$\times \int_{-\infty}^{\infty} dt' \int_{-\infty}^{\infty} dt''\ F(s',s'',t',t''\ ;\ s,t), \tag{4.5}$$

where

$$F(s',s'',t',t'';s,t) = \int d^4\xi_1\, d^4k_1\ \frac{R(s',U;\xi_3^2,\xi_4^2,k_3^2,k_4^2)}{\prod_{i=1}^{4}\left[\xi_i^2-m^2+i\epsilon\right]\left[k_i^2-m^2+i\epsilon\right]}.$$

(4.6)

To carry out the integrals over F, we have to examine its singularities in t′ and t″. To this end, we first examine the function $\tilde{F}$ obtained from F by taking out the Regge pole amplitude:

$$\tilde{F}(s',s'',t',t'';s) = -\frac{g^8}{(2\pi)^8}\int \frac{d^4\xi_1}{\prod_{i=1}^{4}(\xi_i^2-m^2+i\epsilon)} \frac{d^4k_1}{\prod_{i=1}^{4}(k_i^2-m^2+i\epsilon)}$$

$$= A^c(s,t';s',s'')\ A^c(s,t'';s',s'') \quad . \tag{4.7}$$

The "crossed"amplitude $A^c(s,t';s',s'')$ is given by the diagram in Fig. 12. Since this is a fourth order box diagram

it obeys the Mandelstam representation; namely,

$$A^c(s,t';s',s'') = \frac{1}{\pi^2}\iint \frac{dt''\,du''\;\rho(t'',u'')}{(t''-t'-i\epsilon)(u''-u'-i\epsilon)}\,, \tag{4.8}$$

where $u' = 2m^2 + s' + s'' - t' - s$. Notice that $\rho(t'',u'')$ is the third double spectral function with respect to the variable s. Since s is fixed, (4.8) shows that the cuts of A^c in t' plane are those given in Fig. 13. The integration contour in t'-plane runs between the two cuts, so that no matter how we try to distort this contour we always pick up a cut contribution. Therefore, we expect a nonvanishing result when t' integration is carried out in (4.5). If we did not have the crosses, then the integration contours in t' and t'' planes would not have been trapped, and the amplitude A(s,t) given by (4.5) could have vanished.

To evaluate A(s,t) we now make two assumptions. (1) The dominant contributions to the t' and t'' integrations in (4.5) come from theshold singularities. (2) The four-momentum η_2 carried by the Regge pole is very nearly zero.[54] Since at $t' = t'' = 4m^2$ Landau analysis gives $\xi_2 = -\xi_4$, $k_2 = -k_4$,[55] we obtain $U = \frac{t}{4}$. Furthermore at these values of t' and t'', $\xi_4^2 = m^2$, $k_4^2 = m^2$, so that the second assumption implies ξ_3^2, k_3^2 are close to their mass shell values. We can then approximate $R(s',U;\xi_3^2,\xi_4^2,k_3^2,k_4^2)$ in (4.6) by the mass shell Regge amplitude $R(s',\frac{t}{4})$. This leads to

$$A(s,t) \simeq \frac{i}{g^2 8(2\pi)^4}\,\frac{1}{t}\int ds'ds''\,\frac{\tau(s,s',s'')}{s''-m^2}R(s',\tfrac{t}{4})\left[\int_{-\infty}^{\infty} dt' A^c(s,t';s',s'')\right]^2. \tag{4.9}$$

The integral over $A^c(s,t';s',s'')$ can be converted into an integral over its absorptive part:

$$\int_{-\infty}^{\infty} dt' A^c(s,t';s',s'') = 2i\int_{4m^2}^{\infty} dt'\; A_t(t',u',s). \tag{4.10}$$

We approximate this integral by replacing $A_t(t',u',s)$ by a function $\tilde{A}_t(t')$ which is equal to $A_t(t',u',s)$ at threshold and asymptotically $(t' \to \infty)$.[56] It can be shown that

$$\tilde{A}_t(t') = -2g \frac{\mathrm{Im}F(t',s'',m^2)}{t' - m^2}, \tag{4.11}$$

where $\mathrm{Im}F(t',s'',m^2)$ is the imaginary part of the vertex diagram in Fig. 14. Thus the left-hand side of (4.10) becomes

$$\int_{-\infty}^{\infty} dt' \, A^c(s,t';s',s'') \simeq -4ig \int_{4m^2}^{\infty} dt' \frac{\mathrm{Im}F(t',s'',m^2)}{t'-m^2}$$

$$= -4\pi \, i \, g \, F(m^2,s'',m^2)$$

$$= -4\pi \, i \, g \, F(s''). \tag{4.12}$$

From (4.9) we now obtain

$$A(s,t) \simeq -\frac{i}{8\pi^2} \frac{1}{t} \int ds'ds'' \frac{\tau(s,s',s'')}{s''-m^2} R(s',\tfrac{t}{4}) \, F^2(s''). \tag{4.13}$$

Since $\frac{t}{4} \to \infty$, the Regge amplitude can be written as

$$R(s',\tfrac{t}{4}) = \beta(s') \frac{e^{-i\pi\alpha(s')} \pm 1}{\sin \pi\alpha(s')} (\tfrac{t}{4})^{\alpha(s')}. \tag{4.14}$$

Because of the factor $(\frac{t}{4})^{\alpha(s')}$, the dominant contribution in (4.13) comes from the neighborhood of $s' = 0$. The triangle function then gives $s'' \simeq s$. Therefore, (4.13) can be approximated by

$$A(s,t) \simeq -\frac{i}{8\pi^2} \frac{F^2(s)}{s-m^2} \beta(0) \frac{e^{-i\pi\alpha(0)} \pm 1}{\sin\pi\alpha(0)} \frac{1}{t} \int ds'ds'' \tau(s,s',s'') (\tfrac{t}{4})^{\alpha(s')} =$$

$$= -\frac{i}{32\pi}\beta(0)\left[\frac{e^{-i\pi\alpha(0)}\pm 1}{\sin\pi\alpha(0)}\right]\frac{F^2(s)}{s-m^2}\left(\frac{t}{4}\right)^{\alpha(0)-1}\frac{1}{\alpha'(0)\ell n(\frac{t}{4})} . \qquad (4.15)$$

The asymptotic behavior $\frac{t^{\alpha(0)-1}}{\ell nt}$ shows that A(s,t) is a fixed cut with a branch point at $j = \alpha(0) - 1$. Comparing (4.15) with our Born term

$$A_1(t,s) = -t^{\frac{1}{2}}g(t)\frac{F^2(s)}{m^2-s} , \qquad (4.1)$$

we notice that they are essentially identical. In both cases, we have (i) a particle propagator, (ii) each propagator accompanied by form factors at the two vertices, (iii) separation between the energy dependence and the momentum transfer dependence. We infer that the Born amplitude (4.1) can be identified with a fixed Regge-cut amplitude like (4.15). In other words, the phenomenological optical potential (4.1) corresponds in the S-matrix language to a Mandelstam cut diagram with a Regge pole and a single particle exchange.

Well, what have we learned about the interrelation between the optical model and the S-matrix descriptions? We have actually learned a great deal. We have realized that the phenomenological optical potential does correspond to a Feynman-like diagram which provides the detailed mechanism of the process, and that the complex energy dependence of the optical potential comes about via the Regge pole exchange. Does this mean we can now do away with the optical model? The answer is no as soon as we recognize that the general form factor diagram involves a hadronic blob (Fig. 10), and therefore cannot be calculated in the S-matrix theory. On the other hand, the radial dependence of the optical potential provides us the knowledge about the form factor. We see that the S-matrix and the optical model descriptions are complementing each other. Another way of putting the above result is that the diagram in S-matrix theory provides us the branch point in the j-plane and tells us that the discontinuity across it is proportional to the square of

the form factor. But the actual information about the form factor comes from the intuitive picture of spatial extension implicit in the optical model.

We conclude that the optical model and the S-matrix descriptions together give us a more complete picture of hadron interactions than they separately do. It will be stimulating to see whether such interplay between these two descriptions can be established in the realm of diffraction scattering.

Figure Captions

Fig. 1. A generalized ladder diagram.

Fig. 2. A generalized ladder diagram with an unspecified hard interaction.

Fig. 3. Absorption of soft mesons of 4-momenta $q_1, q_2, \ldots, k, \ldots, q_j$ by the p_b line. The cross represents the onset of the hard interaction.

Fig. 4. A generalized ladder diagram with radiative corrections and a hard interaction.

Fig. 5. Emission of soft mesons by the p_a line together with a single self-energy insertion.

Fig. 6. The integration contour for Watson-Sommerfeld transform together with singularities in the j-plane. $\alpha_R(t)$ represents the position of a Regge pole, while $\alpha_c(t)$ that of a branch point.

Fig. 7. Born amplitude in the optical model.

Fig. 8. Mandelstam diagram with a Regge pole (P′) and a vector meson (ω) exchange. Solid lines represent nucleons, short-dashed lines the pions. The blobs represent $N + \omega \to \pi + N$ amplitudes.

Fig. 9. Same as Fig. 8 with each blob replaced by a nucleon line.

Fig. 10. ωNN form factor diagram.

Fig. 11. Mandelstam cut diagram with a Regge pole (wavy line) and a single particle exchange. All particles (solid lines) are of equal mass.

Fig. 12. Box diagram for the amplitude $A^c(s,t';s',s'')$.

Fig. 13. Branch cut of the amplitude $A^c(s,t';s',s'')$ in the t'-plane.

Fig. 14. Diagram representing the vertex function $F(t',s'',m^2)$.

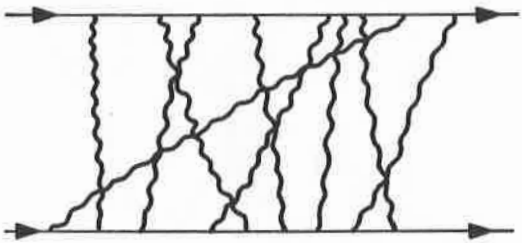


Figure 1

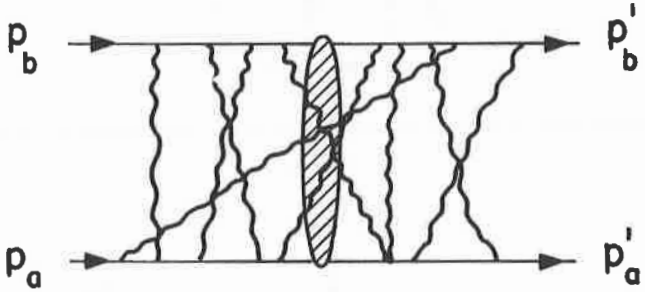


Figure 2

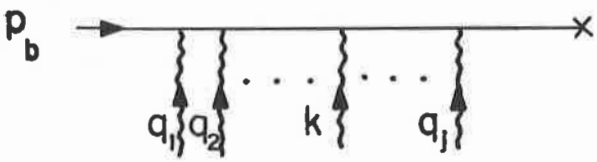


Figure 3

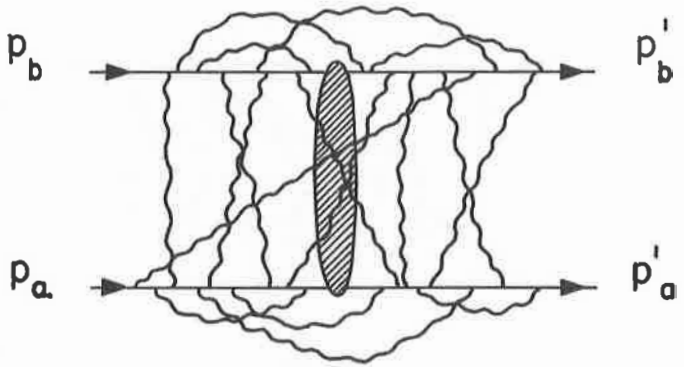


Figure 4

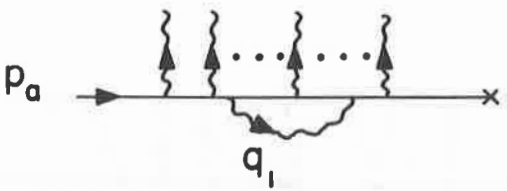


Figure 5

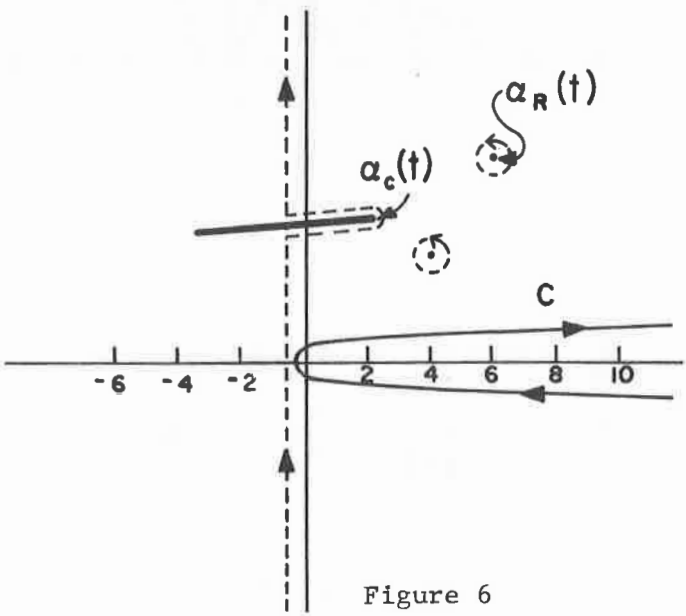


Figure 6

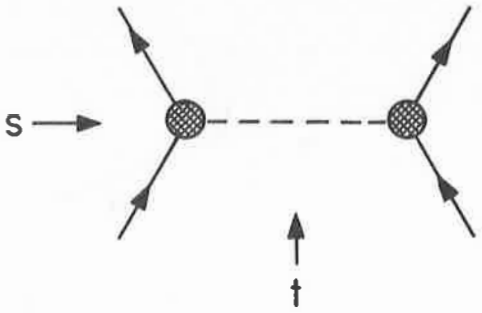


Figure 7

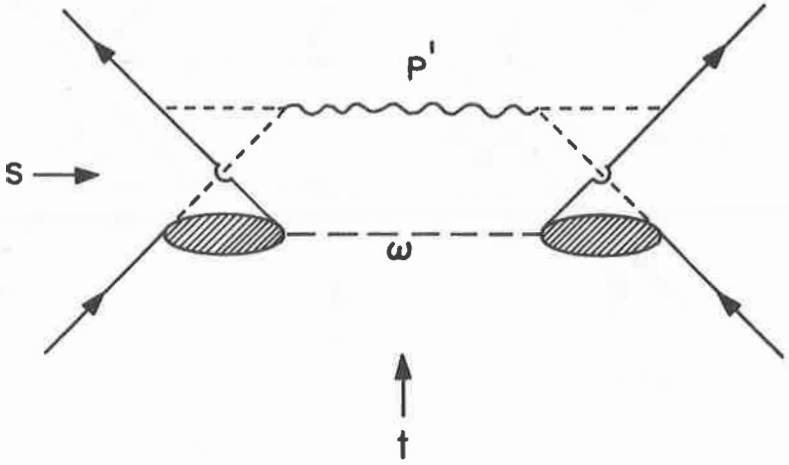


Figure 8

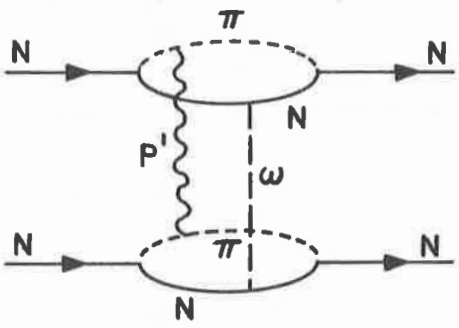


Figure 9

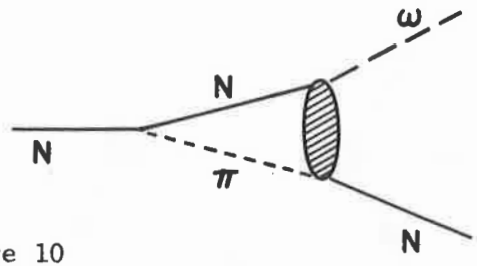


Figure 10

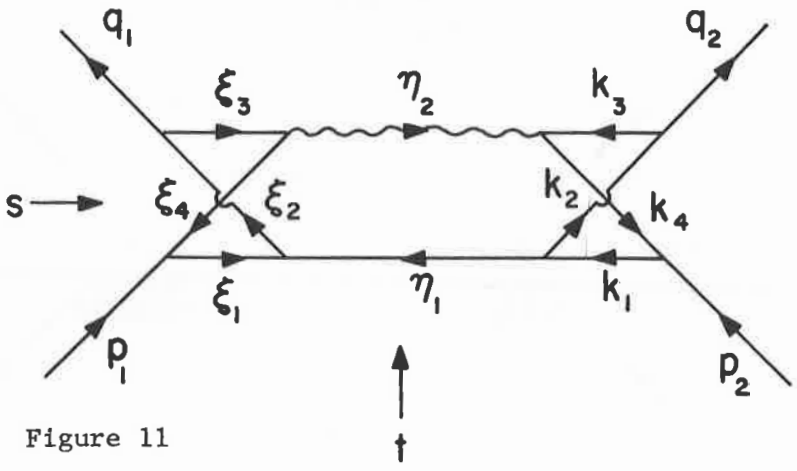


Figure 11

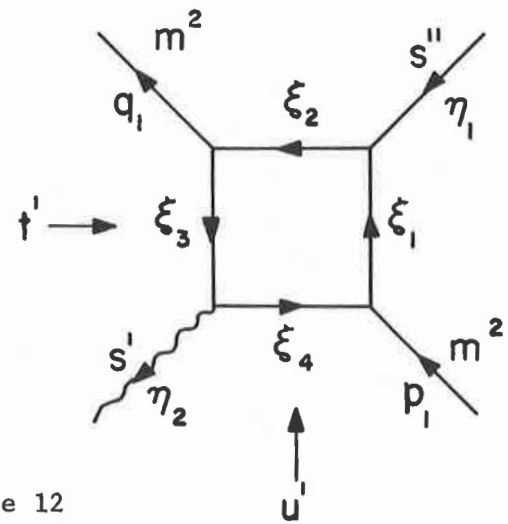


Figure 12

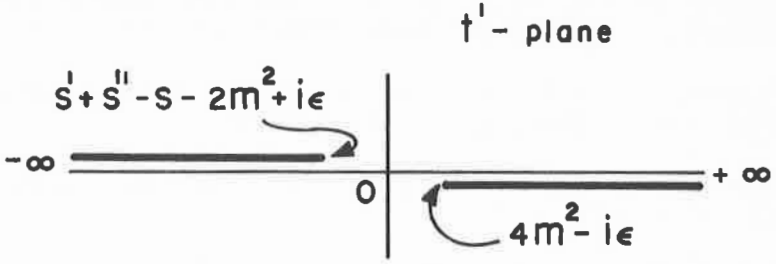


Figure 13

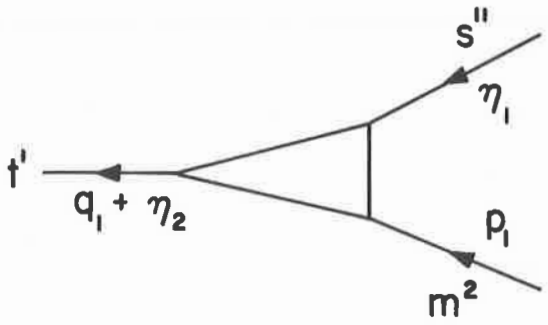


Figure 14

REFERENCES

1. M.M. Islam, in Lectures in Theoretical Physics, edited by A.O. Barut and W.E. Brittin (Gordon and Breach, Science Publishers, 1968) Vol. XB, p. 97.

2. H.Cheng and T.T. Wu, Phys. Rev. 182, 1852 (1969); 182, 1868 (1969); 186, 1611 (1969).

3. H.D.I. Abarbanel and C. Itzykson, Phys. Rev. Letters 23, 53 (1969).

4. B.M. Barbashov, S.P. Kuleshov, V.A. Matveev, and A.N. Sissakian, JINR preprint E2-4692, Dubna (1969).

5. S.J. Chang and S.K. Ma, Phys. Rev. Letters 22, 1334 (1969); Phys. Rev. 188, 2385 (1969); H. Cheng and T.T. Wu, Phys. Rev. 182, 1899 (1969).

6. M. Lévy and J. Sucher, Phys. Rev. 186, 1656 (1969).

7. F. Englert, P. Nicoletopoulos, R. Brout, and C. Truffin, Nuovo Cimento 64A, 561 (1969).

8. E.A. Remler, Phys. Rev. D1, 1214 (1970).

9. M.M. Islam, Nuovo Cimento, 5A, 315 (1971).

10. S. Okubo, Nuovo Cimento 18, 70 (1960); K.E. Erikson, Nuovo Cimento 19, 1010 (1961); S. Weinberg, Phys. Rev. 140, B516 (1965). For general treatment of infrared divergence, see D.R. Yennie, S.C. Frautschi, and H. Suura, Ann. of Phys. (N.Y.) 13, 379 (1961); K.T. Mahanthappa, Phys. Rev. 126, 329 (1962).

11. Y.P. Yao, Phys. Rev. D1, 2971 (1970).

12. S.J. Chang, Phys. Rev. D1, 2977 (1970).

13. B.M. Barbashov, S.P. Kuleshov, V.A. Matveev, and A.N. Sissakian, JINR preprint E2-4983, Dubna (1970).

14. H.M. Fried and T.K. Gaisser, Phys. Rev. 179, 1491 (1969); T.K. Gaisser, Phys. Rev. D2, 1337 (1970).

15. Thus if the unknown corresponds to the exchange of a scalar meson, then $M^h(k_r) = -\frac{g^2}{k_r^2-m^2+i\epsilon}$.

16. For the scattering of spin-zero particles, the Feynman amplitude is related to the S-matrix in the following way:

$$S_{fi}=\delta_{fi}+i(2\pi)^4\delta(p_a+p_b-p_a'-p_b')\frac{1}{(16\,p_{a0}p_{a0}'p_{b0}p_{b0}')^{\frac{1}{2}}}M.$$

The differential scattering cross section is given by $\frac{d\sigma}{d\Omega}=\frac{1}{(8\pi W)^2}|M|^2$, where W is the c.m. energy.

17. p_b line is the part of b line before the hard interaction, and p_b' line is the part after the hard interaction. The p_a line and p_a' line are defined similarly.

18. This has been regarded by Lévy and Sucher as an averaging over the n + 1 ways by which 4-momentum conservation can be imposed on the internal momenta k_i's.

19. L.I. Schiff, Phys. Rev. 103, 443 (1956); D.S. Saxon and L.I. Schiff, Nuovo Cimento 6, 614 (1957).

20. On the basis of this argument Schiff's original result is not correct, since the same potential was used for both soft and hard interactions. However, the physical picture employed, namely - large angle scattering of order n + 1 occurs as a single scattering through a large angle accompanied by n small angle scattering, is in complete agreement with Eq. (2.18b).

21. M. Lévy and J. Sucher, Phys. Rev. D2, 1716 (1970).

22. Numerical evaluation of impact parameter amplitudes in potential scattering and comparison with exact phase shift analysis show that even for large-angle scattering the Lévy-Sucher formula is better and provides a more accurate imaginary part of the

amplitude than the Schiff formula (Y. Hahn, Phys. Rev. C2, 775 (1970)).

23. This is shown later.

24. R.J. Glauber, in Lectures in Theoretical Physics, edited by W.E. Brittin and L.G. Dunham (Interscience Publishers, New York, 1959) vol. 1, p. 315.

25. S. Gasiorowicz, Elementary Particle Physics (John Wiley and Sons, Inc., 1966) p. 463; J.J. Sakurai, Advanced Quantum Mechanics (Addison-Wesley Publishing Company, 1967) p. 195.

26. M. Martinis, Nuovo Cimento 59A, 490 (1969).

27. T(s,t) has the partial wave expansion

$$T(s,t)=\frac{W}{k}\sum_{\ell=0}^{\infty}(2\ell+1)P_{\ell}(\cos\theta)\frac{\left[e^{2i\delta_{\ell}}-1\right]}{2i},$$

and is related to the differential elastic cross section by

$$\frac{d\sigma}{d\Omega}=\frac{1}{W^2}\left|T(s,t)\right|^2 \qquad (s=W^2).$$

28. Tables of Integral Transforms, edited by A. Erdélyi (McGraw-Hill Book Co., 1954) Vol. 2, p. 54.

29. For a comprehensive treatment of the Regge pole theory, see P.D.B. Collins and E.J. Squires, "Regge Poles in Particle Physics", edited by G. Hohler (Springer Tracts in Modern Physics, 1968) Vol. 45.

30. Asymptotically the dominant cut contribution comes from the neighborhood of the branch point, so that the lower limit in (3.19) does not have to be specified.

31. D. Amati, S. Fubini and A. Stanghellini, Phys. Letters 1, 29 (1962).

32. S. Mandelstam, Nuovo Cimento 30, 1127 (1963) and 30,

1148 (1963).

33. J.C. Polkinghorne, J. Math. Phys. 4, 1396 (1963).

34. V.N. Gribov, I. Ya. Pomeranchuk, and K.A. Ter-Martirosyan, Phys. Rev. 139, B184 (1965); Sov. J. Nucl. Phys. 2, 258 (1966).

35. V.N. Gribov, JETP (Sov. Phys.) 26, 414 (1968).

36. J.C. Polkinghorne, J. Math. Phys. 6, 1960 (1965).

37. In the forward direction (t = 0) the pole contribution dominates, since the cut contributions are down by the lns factors.

38. A number of authors have proposed this as an explanation of the rapid fall of high-energy large-angle pp elastic scattering. See A.A. Ansel'm and I.T. Dyatlov, Sov. J. Nucl. Phys. 6, 439 (1968); S.C. Frautschi and B. Margolis, Nuovo Cimento 56A, 1155 (1968).

39. For n = 2 this result is not correct, since t-channel unitarity will require the cut discontinuity to vanish at the tip of the cut (J.B. Bronzan and C.E. Jones, Phys. Rev. 160, 1494 (1967).

40. A.A. Ansel'm and I.T. Dyatlov, Sov. J. Nucl. Phys. 6, 430 (1968).

41. B. Hasslacher et al., Phys. Rev. Letters 25, 1591 (1970); G.M. Cicuta and R.L. Sugar, Phys. Rev. D3, 970 (1971); B. Hasslacher and D.K. Sinclair, Phys. Rev. D3, 1770 (1971).

42. T.T. Chou and C. N. Yang, Phys. Rev. Letters 20, 1213 (1968); Phys. Rev. 170, 1591 (1968).

43. Work of many authors belong to the second category; for example, the work of Ansel'm and Dyatlov, and of Frautschi and Margolis quoted in reference 38. The work of Cheng and Wu (Phys. Rev. Letters 24, 1456 (1970),) and of Chang and Yan (Phys. Rev. Letters 25, 1586 (1970) also belong to this category.

44. C.B. Chiu and J. Finkelstein, Nuovo Cimento 57A, 649 (1968); R.C. Arnold and M.L. Blackmon, Phys. Rev. 176, 2082 (1968).

45. Other authors have also raised the same question. For example, L. Durand, III, and R. Lipes, Phys. Rev. Letters 20, 637 (1968).

46. M.M. Islam, Phys. Letters 31B, 313 (1970).

47. M.M. Islam, Nuovo Cimento Letters 4, 447 (1970).

48. M.M. Islam and Joe Rosen, Phys. Rev. Letters 22, 502 (1969); Phys. Rev. 185, 1917 (1969).

49. M.M. Islam, Nuovo Cimento Letters 48, 251 (1967).

50. Reasons for choosing ω and P' come from the phenomenological analysis.

51. R.J. Rothe, Phys. Rev. 159, 1471 (1967).

52. J.H. Schwarz, Phys. Rev. 162, 1671 (1967). D.I. Olive and J.C. Polkinghorne, Phys. Rev. 171, 1475 (1968).

53. An extra factor of 2 has to be included in the denominator of the Jacobian since the transformation of variables is not one-to-one. See T.W.B. Kibble, Phys. Rev. 117, 1159 (1959).

54. Assumption (2) can be mathematically justified by using a different set of integration variables. See M.M. Islam, Phys. Letters 36B, 586 (1971). Physically assumption (1) implies that the particles produced in the intermediate state are essentially at rest in the rest frame of the incoming particle.

55. R.J. Eden, High Energy Collisions of Elementary Particles (Cambridge University Press, 1967), p. 92.

56. For proofs of these statements and others made below, see Ref. 46. The dependence of A_t and $\tilde{A}_t$ on s' and s'' have been suppressed to simplify notation.

Part II

DUAL MODELS IN STRONG INTERACTIONS

AN INTRODUCTION TO DUAL RESONANCE MODELS

K. Bardakci
Department of Physics
University of California
Berkeley, California 94720

I. Elementary Duality and Consequences

Dual resonance models are the natural outcome of the concept of duality, which I want to review briefly here. In a nutshell, a dual amplitude is an analytic and crossing symmetric amplitude that has Regge asymptotic behavior. Historically, however, various constraints naturally satisfied by dual amplitudes, the so-called superconvergence relations[1] and finite energy sum rules[2,3] (F.E.S.R.) were studied first, and so I will start with them. Consider the equal mass scalar two particle scattering amplitude of Fig. 1., and define the usual variables $s=(k_1+k_2)^2$, $t=(k_2+k_3)^2$, $u=(k_1+k_3)^2$, $\nu=\frac{s-u}{4m}$ (m is the mass). Odd and even parts of the scattering amplitude A are defined by $A^{\pm}(\nu,t)=\frac{1}{2}[A(\nu,t)\pm A(-\nu,t)]$. If A completely reggeizes for large ν, one can write,

$$A^{(\pm)}(\nu,t) \to \sum_i \beta_i^{(\pm)}(t)\ \frac{\pm 1+e^{-i\pi\alpha_i^{(\pm)}(t)}}{\mathrm{Sin}(\pi\alpha_i^{(\pm)}(t))}\ \nu^{\alpha_i^{(\pm)}(t)}, \quad (1.1)$$

where α's are the trajectories, and β's are the residues. The summation includes both the leading trajectory and the non-leading trajectories; cuts, fixed singularities etc. are assumed to be absent. If A is analytic in ν except on the real axis for fixed real t, then it satisfies a dispersion relation. In addition, suppose that the leading regge trajectory α satisfies $\mathrm{Re}(\alpha(t)<-n$ for some t, where n is an integer. Then $\nu^{\ell}A(\nu,t)$ satisfies an unsubtracted dispersion relation for integer ℓ satisfying $0\leq \ell \leq n$. This immediately leads to a series of super convergence relations:

$$\int_{-\infty}^{\infty} d\nu(\nu)^{\ell-1} \mathrm{Im}\{A(\nu,t)\} = 0, \qquad (1.2)$$

where $1 \leq \ell \leq n$.

If the leading regge trajectory does not satisfy the condition $\alpha < -n$, then we define an auxiliary amplitude $\bar{A}$,

$$\bar{A}^{(\pm)}(\nu,t) \equiv A^{(\pm)}(\nu,t) - \sum_i \beta_i^{(\pm)}(t) \frac{\pm 1 + e^{-i\pi\alpha_i^{(\pm)}(t)}}{\mathrm{Sin}(\pi\alpha_i^{(\pm)}(t))} \nu^{\alpha_i^{(\pm)}(t)} \qquad (1.3)$$

where the sum over i includes all Regge trajectories that satisfy $\mathrm{Re}(\alpha_i) > -n$, where n is an arbitrary positive integer. (For simplicity, we assume there is no α with $\mathrm{Re}(\alpha) = -n$.) $\bar{A}$, so defined, satisfies the superconvergence relations given by equation (1.2). As in reference 2, one then chooses a cutoff energy N, and uses the form given by eq. (1.3) for the imaginary part of $\bar{A}$ for $|\nu| \leq N$. For $|\nu| > N$, we set,

$$\bar{A}^{(\pm)}(\nu) \to \sum_j \beta_j^{(\pm)} \frac{\pm 1 + e^{-i\pi\alpha_j^{(\pm)}}}{\mathrm{Sin}(\pi\alpha_j^{(\pm)})} \nu^{\alpha_j^{(\pm)}}, \qquad (1.4)$$

where the sum over j now includes only those trajectories that satisfy $\mathrm{Re}(\alpha_j) < -n$. This set is complimentary to the set of trajectories of eq. (1.3). After performing some trivial integrations, and combining terms, we get,

$$\int_0^N d\nu\ (\nu)^{\ell} |\mathrm{Im}[A^{(\pm)}(\nu,t)]| = \sum_i \frac{\beta_i^{(\pm)} N^{\alpha_i^{(\pm)}+\ell+1}}{\alpha_i^{(\pm)}+\ell+1}, \qquad (1.5)$$

where N is an arbitrary cutoff point, ℓ is even for (+) amplitude and odd for (-) amplitude, and the sum on the right hand side goes over all the trajectories that contribute to A. The set of equations given by (1.5) can be taken to be the definition of duality; they imply complete regge asymptotic behaviour and crossing. However, if one does not specify Im(A) any further, the equations lack content. We supplement them by postulating that Im{A} is entirely built out of narrow (to be precise, zero width)

resonances. Since these resonances also lie on their regge trajectories, we then have a set of bootstrap equations for the regge trajectories. In the preceding discussion, the contribution of the pomeranchon and the non-resonant part of Im{A} have been neglected. At this point, it is customary to lump these contributions together and throw them both away, invoking the Harari-Freund hypothesis[4]. This treatment may ultimately prove unsatisfactory, but at this stage of the development of dual resonance models, nothing better seems to be available.

Let us now restate duality in a weaker form: A two particle scattering amplitude is dual if it satisfies the following condition,

$$A(s,t) = \sum_{n=0}^{\infty} \frac{C_n^{(s)}(t)}{s-s_n} + \sum_{n=0}^{\infty} \frac{C_n^{(u)}(t)}{u-u_n} \qquad (1.5)$$

$$= \sum_{n=0}^{\infty} \frac{C_n^{(t)}(s)}{t-t_n} + \sum_{n=0}^{\infty} \frac{C_n^{(u)}(s)}{u-u_n} \; .$$

Here the C's are polynomials of degree n in their arguments, n being the maximum allowable angular momentum of the resonance at $s = s_n$. The existence of such an upper limit is the same as the absence of the so-called ancestors. Equation (1.5) follows from analyticity, regge asymptotics, and narrow resonance approximations. It does not, however, conversely imply regge asymptotics; cuts and fixed singularities in the angular momentum may be compatible with (1.5).

II. Quark Diagrams and Planar Duality

So far we have not attached any isospin and SU(3) indices to our amplitude. Let i_1, i_2, i_3 and i_4 be internal symmetry indices of particles 1,2,3 and 4 of Fig. 1.

The amplitude can be expressed as follows:

$$A^{i_1 i_2 i_3 i_4}(s,t) = \sum_{\lambda} I^{i_1 i_2 i_3 i_4}_{(s)\lambda} \; A_{\lambda}^{(s)}(s,t)$$

$$+ \sum_{\lambda} I^{i_1 i_2 i_3 i_4}_{(u),\lambda} A_\lambda^{(u)}(s,t), \qquad (2.1)$$

where the I's are appropriate numerical coefficients, labeled by the SU(3) (or isospin) quantum numbers λ in the s and u channels respectively, and the A_λ's satisfy an equation similar to (1.5),

$$A_\lambda^{(s)} = \sum_n \frac{C_n^{(s)}(t)}{s-s_n} = \sum \frac{C_n^{(t)}(s)}{t-t_n}, \qquad (2.2)$$

and a similar equation for $A_\lambda^{(u)}$. Note that the s and u channels have to be treated separately, since they have different SU(3) properties.

The statement of duality given by eq. (1.1) is usually supplemented by the further assumption of the absence of exotic resonances. Any meson resonance whose quantum numbers are given by a quark-antiquark system, or baryons composed of three quarks are non-exotic, everything else is exotic. This imposes rather stringent conditions on the A_λ's; they must be so chosen that no exotic resonances appear in (2.1). The solution to this problem is given by Harari and by Rosner[6], and is graphically expressed by the well-known quark diagrams of Fig. 2. The external mesons are taken to belong to 1 and 8 representations of SU(3), therefore they can be represented, in the SU(3) space, by a pair of quark-antiquark indices. The quark lines represent Kroenecker delta's in the quark indices, so that Fig. 2 corresponds to an SU(3) factor of the following form:

$$I \sim \delta_{\alpha\alpha'}\, \delta_{\beta\beta'}\, \delta_{\gamma\gamma'}\, \delta_{\eta\eta'}.$$

The index i_1 of eq. (2.1) is now the pair of indices η and γ, etc. These rules can easily be generalized to meson-baryon scattering, as in Fig. 3, but extension to baryon-antibaryon scattering is impossible without the introduction of exotics, as is well known by now. From the quark content of the s and t channels in Fig. 2, and from the symmetry of the diagram with respect to the two channels, it is clear that it provides a solution to the problem at hand. The question is, is it the only solution? For the simpler case of pion-pion scattering, we shall

show that this is indeed the case. In this case, λ can be taken to run over three values, $\lambda = 0,1,2$, and to label the total isospin in the s-channel. The t-channel isospin wave functions are then given by the crossing matrix:

$$\{C_{\lambda\lambda'}\} \equiv \begin{pmatrix} 1/3 & 1 & 5/3 \\ 1/3 & \frac{1}{2} & -5/6 \\ 1/3 & -\frac{1}{2} & 1/6 \end{pmatrix}$$

through the relation,

$$I_{(t),\lambda} = \sum_{\lambda'} C_{\lambda\lambda'} I_{(s),\lambda'}. \qquad (2.3)$$

The matrix C has three eigenvectors, one with eigenvalues - 1, the other two with eigenvalues + 1. (There can only be two kinds of eigenvalues, ± 1, since $C^2 = 1$.) The s and t channels are identical in this problem; so the eigenvector with the negative eigenvalue, which involves a sign change between s and t channels, cannot be present. The eigenvectors with + 1 eigenvalue are the following:

$$\chi_1 = \begin{pmatrix} 3 \\ 2 \\ 0 \end{pmatrix}, \quad \chi_2 = \begin{pmatrix} 5 \\ 0 \\ 2 \end{pmatrix}.$$

Clearly, χ_1 is the quark solution, since the exotic I=2 component is absent in this solution. This demonstrates the uniqueness of the solution for the simple example. The more general case of SU(3) multiplets can be treated similarly using SU(3) crossing matrices[6].

A large number of experimental predictions follow from eq. (2.1) combined with the Harari-Rosner quark hypothesis, and they are in general well satisfied, especially in the case of mesons. The main result is exchange degeneracy; ρ,w,f and A_2 trajectories all turn out to be exchange degenerate, and so do φ and f′, and K^x and K^{xx} trajectories. Assuming only isospin invariance, one can also get some information about SU(3) breaking; for example, the ω-φ mixing angle turns out to satisfy $tg^2\theta = \frac{1}{2}$.

We postpone at this point, however, any detailed discussion of the various applications, and instead turn to the extension of the quark diagrams to many particle processes. This extension turns out to be very simple. One first labels the N particle amplitude of Fig. 4 in a definite (and for time being, arbitrary) order. This defines a set cyclic planar channels, namely, the channels made of particles labeled by successive integers. (Here we naturally use a system modular N; N+1 is identified with 1, N+2 with 2, etc.) Next comes the definition of overlapping and non-overlapping channels. Two planar channels are overlapping if they share a proper set of legs (particles); they are non-overlapping if one includes the other or if they are totally disjoint. For an N point function, one can always choose a maximal set of non-overlapping channels; for example, the channels (1,2), (1,3),...,(1,N-3), where we label the channel that couples to the external legs i,i+1,....,; by simply (i,j). There are clearly various different ways of choosing a maximal set of non-overlapping channels, but the number of such channels is always equal to N-3. The generalization of eq. (2.2) is then the following: The dual amplitude A can be written as a sum over the poles in any maximal set of non-overlapping planar channels,

$$A = \sum_{n_{ij}} \prod_{i,j} \frac{C_{n_{ij}}}{s_{ij} - s_{n_{ij}}} , \qquad (2.4)$$

where $s_{ij} = (p_i + p_{i+1} + \cdots + p_j)^2$, s_{nij} is the square of the mass of a given resonance, and the pair of indices i and j run over a maximal set of non-overlapping planar channels. (The equation should hold for all possible choices of such a set.) The C's are polynomials in various momentum transfer variables.

The dual "skeleton" amplitude of eq. (2.4) is then multiplied by the appropriate SU(3) factors derived from the quark diagrams of Fig. 4. The resulting amplitude is then symmetrized between identical bosons and antisymmetrized between identical fermions. For identical particles it does not therefore matter how one initially chooses the set planar channels; the final symmetrization eliminates the arbitrariness of the initial choice. When all the particles are not identical, however, there is an

ambiguity in the choice of the planar channels. This ambiquity is usually resolved by requiring absence of exotic channels, signature requirements, etc. Finally, a compact form of SU(3) quark factor for mesons due to the Chan and Paton[7] can easily be written down. If the external mesons are SU(3) octets, we have;

$$I \cong \mathrm{Tr}\left\{\lambda_{i_1}\lambda_{i_2}\cdots\lambda_{i_N}\right\}, \tag{2.5}$$

where λ's are usualy SU(3) matrices, indices $i_1, i_2, \cdots i_N$ label the respective external mesons. For singlet mesons, the corresponding λ is replaced by the unit matrix. Clearly, eq. (2.5) is a restatement of the Harari-Rosner quark rules.

III. The Veneziano Formula and Some Applications

An explicit solution to eq. (2.2) is provided by Veneziano's celebrated ansatz[8]:

$$A(s,t) \equiv B[-\alpha(s), -\alpha(t)] = \frac{\Gamma[-\alpha(s)]\,\Gamma[-\alpha(t)]}{\Gamma[-\alpha(s)-\alpha(t)]}$$

$$\equiv \int_0^1 dx\, x^{-\alpha(s)-1}\,(1-x)^{-\alpha(t)-1}, \tag{3.1}$$

where $\alpha(s) = a + bs$, the trajectory function, has to be a linear function of s, in order to avoid non-polynomial residues at the poles. (Absence of ancestors). For simplicity, we take the s and t channels to be identical. The properties of the Veneziano formula can be investigated either by considering it as a ratio of gamma functions, or from the integral representation. I prefer the latter approach at this point, since it generalizes readily to the N point amplitudes. Expanding $(1-x)^{-\alpha(t)-1}$ term in the integral representation of eq. (3.1) in power series in x, and integrating term by term, we get:

$$B[-\alpha(s),-\alpha(t)] = \sum_{n=0}^{\infty} \frac{[\alpha(t)+1][\alpha(t)+2]\cdots[\alpha(t)+n]}{n!\;(n-\alpha(s))}, \tag{3.2}$$

Of course, a similar expression with s and t interchanged is also valid. This verifies the duality of the amplitude and the absence of ancestors for linear trajectories explicitly. Actually, the amplitude satisfies a stronger duality condition; it reggeizes in both channels. It is easy to prove this statement in the unphysical limit $|s| \to \infty$, with $\mathrm{Re}[\alpha(s)] < 0$. Changing variables by $x = \exp\{\frac{z}{\alpha(s)}\}$, we have,

$$B = -\frac{1}{\alpha(s)} \int_0^\infty dz\, e^{-z} \left\{1 - \exp\left(\frac{z}{\alpha_{(s)}}\right)\right\}^{-\alpha(t)-1}$$

$$\tilde{=} [-\alpha(s)]^{\alpha(t)} \int_0^\infty dz\, e^{-z}\, z^{-\alpha(t)-1} \qquad (3.3)$$

$$= [-\alpha(s)]^{\alpha(t)}\, \Gamma[-\alpha(t)].$$

The second step follows by treating $z/\alpha(s)$ as a small quantity for $|s| \to \infty$, expanding the exponential in powers of $z/\alpha(s)$ and keeping only the lowest non-vanishing term. Since z ranges between 0 and ∞, one has to show that only those values of z small compared to $\alpha(s)$ contribute to the asymptotic limit. This can be established by a more careful analysis.

The integral representation of eq. (3.1) does not exist for $\mathrm{Re}[\alpha(s)] > 0$, so that it is not possible to demonstrate reggeization in the right half plane directly. One can, however, transform the original integral representation into a form[9] that exists for $\mathrm{Re}[\alpha(s)] > 0$. This is done by first converting the path of integration in (3.1) into a closed curve around the real interval from 0 to 1, and then wrapping this closed curve around the real axis from 1 to ∞. The details are given by Mandelstam in his Brandeis notes and will not be repeated here. The end result is the following alternative integral representation:

$$B[-\alpha(s), -\alpha(t)] = \frac{\mathrm{Sin}[\pi[\alpha(s)+\alpha(t)+1]]}{\mathrm{Sin}[\pi[\alpha(s)+1]]} \times \int_1^\infty dx\, x^{-\alpha(s)-1}\, (x-1)^{-\alpha(t)-1}. \qquad (3.4)$$

This representation is well-defined for $\mathrm{Re}[\alpha(s)]>0$, and reggeization can be proved by the same method that led up to eq. (3.3).

At this point, a few general comments are in order:

a) The requirement of duality is not sufficient to fix the amplitude uniquely. One can always add satellite terms of the form $\sum_{n,m} C_{nm} B[-\alpha(s)-n,-\alpha(t)-m]$ to the leading beta function formula.

b) The regge trajectories in s and t channels need not be the same. In particular, they can have different slopes and intercepts. Of course, we have nothing against different intercepts. Different slopes in the s and t channels, however, give rise to amplitudes that blow up exponentially for fixed angle and large energy.[9] To avoid this unphysical behavior, all the slopes have to be taken equal. In what follows, we set this universal slope equal to one for simplicity.

Now, some applications:

a) The process $\pi + \pi \to \pi + \omega$. This is the process that led up to the discovery of the Veneziano formula.[10,8] The simplest ansatz for the amplitude is the following:

$$A \tilde{=} \epsilon^{i_1 i_2 i_3} \epsilon_{\mu_1\mu_2\mu_3\mu_4} k_1^{\mu_1} k_2^{\mu_2} k_3^{\mu_3} \times \epsilon^{\mu_4} B[-\alpha(s)+1, -\alpha_\rho(t)+1], + \text{Perm.} \tag{3.5}$$

required by Bose symmetry, where i_1, i_2 and i_3 are the isospin indices of the pions, the k's are the momenta, ϵ is the polarization vector of the ω, and α_ρ is the ρ trajectory. The isospin and the spin factors that multiply the B function are in this case uniquely determined by isospin and parity conservation. However, one still has the freedom of adding satellite terms, and only some vague requirements of simplicity prohibit us from doing so. The spin factor and the arguments of the B function are so arranged that the first particle appears at $\alpha_\rho=1$ with spin one, and the "ghost" at $\alpha_\rho=0$ is eliminated.

b) $\pi + \pi \to \pi + \pi$. This is one of the earliest interesting applications of the Veneziano formula.[11]

$$A \tilde{=} \mathrm{Tr}(\tau^{i_1}\tau^{i_2}\tau^{i_3}\tau^{i_4})\ [\alpha_\rho(s)+\alpha_\rho(t)-1] \times B[\alpha_\rho(s)+1,\ -\alpha_\rho(t)+1]. \tag{3.6}$$

The i's are the pion isospin labels, and the τ's are the usual Pauli matrices. As usual, it is understood that the expression for A is to be symmetrized with respect to the external lines.

Eq. (3.6) has several nice and at the same time puzzling features. The isospin factor is precisely the Harari-Rosner quark recipe. The kinematical factor kills the scalar ghost on the ρ trajectory and at the same time satisfies the Adler P.C.A.C. condition if $\alpha_\rho(o)=\frac{1}{2}$. The ρ-f exchange degeneracy is a by product, as promised earlier.

The pomeranchuk trajectory is absent from (3.6), in accordance with the Harari-Freund philosophy. One then perhaps considers (3.6) as a solution to a bootstrap equation involving the normal trajectories, rather than as a phenomenological description of the π - π amplitude. Finally, the lack of uniqueness is again obvious.

The ansatz of (3.6) can be extended to the scattering amplitudes for the whole pseudoscalar octet[12]. The τ's in the trace are then replaced by λ's, and the ρ trajectory is then replaced by K^x or Φ trajectories depending on the quantum numbers. For example, $k + \pi \to k + \bar{u}$ amplitude is given by,

$$A_{k+\bar{u}\to k+\bar{u}} \tilde{=} \chi_1^{+}(\tau^{i_2}\tau^{i_3})\chi_2\ [\alpha_\rho(s)+\alpha_{k^x}(t)-1] \times B[-\alpha_\rho(s)+1,\ -\alpha_{k^x}(t)+1] + \text{Terms required by symmetry.} \tag{3.7}$$

χ_1 and χ_2 are the isospin wave functions of the kaons, and the corresponding quark picture is given by Fig. 5. The horizontal lines carry n or p type quarks; vertical lines carry λ type quarks.

c) Pseudoscalar-vector scattering. As an example, let us consider $\pi + \rho \to \pi + \rho$. The isospin factor is still given by the quark factors. It then remains to determine the spin factor. The trajectories that can couple to the $\rho\,\pi$ channel are π, φ, A_1 and A_2, whereas the $\pi\,\pi$ channel has only ρ and f trajectories. Furthermore, since this is an elastic process, the residues at the poles corresponding to particle exchanges must be positive, since these residues are squares of various (real) coupling constants. A negative residue would imply negative metric objects (ghosts). After a few tries, the amateur spinologist can easily convince himself (if not others) no simple factor analogous to the one in eq. (3.6) will do. The basic reason is the conspiracy theory which in general requires parity doublets and among other things, forces ω and A_1 to be degenerate. It is possible to avoid some of these problems, but only at the cost of introducing a large number of satellite terms and forfeiting the general factorization properties of the amplitude.[13]

d) Meson-Nucleon scattering. This is a process which provides a good example for parity doubling. In the baryon channel, we have the McDowell symmetry between the positive parity partial wave $f_\ell^{(+)}$ and negative parity partial wave $f_\ell^{(-)}$,

$$f_\ell^{(+)}(\sqrt{s}) = -\, f_{\ell+1}^{(-)}(-\sqrt{s}). \tag{3.8}$$

So except in the case of vanishing residue, a pole in $f_\ell^{(+)}$ implies a pole in $f_{\ell+1}^{(-)}$, leading to parity doubling. Another way to see this is to note that one needs the Dirac projection operator $\Lambda = \not{p} + m_J$ to project out the negative parity states, where p is the momentum in the given channel and m_J is the mass of the state with angular momentum J. If m is taken to be independent of J, only one parity doublet at most is eliminated and all the others survive. This explains why most dual resonance models with baryon trajectories are plagued with parity doublets. The only way of eliminating the parity doublet trajectory we know of so far[14] introduces a cut in the complex J plane by taking $m_J = \sqrt{J-a}$. However, dual models with cuts in angular momentum plane are much more complicated than the ordinary kind

and they do not seem to factorize. The interested reader should consult a review by Berger and Fox.[15]

Finally, I would like to mention a model due to Mandelstam[16], which introduces the spin factors in a way analogous to the isospin factors. One simply extends the quark rules of Harari and Rosner to include spin. The spectrum of the model so obtained is a parity doubled SU(6) ⊗ O(3). In addition, the parity doublets turn out to be ghosts. This simple model highlights the difficulties that lie in the path of a future satisfactory model.

4. Many Particle Amplitudes

An explicit N point function which satisfies the conditions of generalized duality stated in section II and in particular eq. (2.4) can be explicitly constructed in the form of an integral representation. Following the notation of section II and Fig. 4, we ansatz the following expression:

$$B_N = \int_0^1 \frac{du_{1,2}\cdots du_{1,N-2}}{W} \prod_{i,j} (u_{i,j})^{-\alpha_{ij}-1}, \qquad (4.1)$$

where W is a suitable volume measure and the pair of indices i and j run over all planar channels. The poles in various channels come about when the corresponding u vanishes. Since overlapping channels cannot develop simultaneous poles, the corresponding u's cannot vanish simultaneously. One simple way of ensuring this is to impose the following conditions:

$$u_{i,j} = 1 - \prod_{m,n \subset C_{ij}} (u_{m,n}), \qquad (4.2)$$

where m and n range over the set C_{ij} of all channels that overlap with the channel (ij). The eq's given by (4.2) can be used to determine all the u's in terms of a complete non-overlapping set; we have,

$$u_{ij} = \frac{(1 - u_{\ell,i}\cdots u_{\ell,j-1})\,(1 - u_{\ell,i-1}\cdots u_{\ell,j})}{(1 - u_{\ell,i}\cdots u_{\ell,j})\,(1 - u_{\ell,i-1}\cdots u_{\ell,j-1})}, \qquad (4.3)$$

where the conventions $u_{i,i} = 0$, $u_{i,j+N} \equiv u_{i,j}$ etc. are understood, and ℓ is arbitrary. For example, taking $\ell+1$, i=2, j=3,4---,N-1, all other u's are expressed in terms of the complete non-overlapping set $u_{1,2},\ldots,$ $u_{1,N-3}$. That (4.3) satisfies (4.2) can be verified by direct substitution. Eq. (4.3) can also be used to show the invariance of (4.1) under a cyclic transformation of variables; in fact, (4.3) itself defines such a transformation. The second factor on the right hand side of eq. (4.1) is cyclic symmetric as it stands. The volume element also turns out to be invariant under cyclic change of variables if W is chosen as follows:

$$W = \prod_{i<j} (u_{i,j})^{j-i-1} , \qquad (4.4)$$

where in the product i ranges over values 2,3,...,N-2, and j ranges over 3,---, N-1, subject to the condition $i < j$. With the definition of B_N completed, it is possible to verify the duality condition of eq. (2.4), by expanding the integrand of (4.1) in Taylor series in a linearly independent set of u's and integrating term by term. Before plunging into a dicussion of the properties of B_N, it will prove useful to reexpress it in two equivalent but formally distinct ways. By repeatedly using the equation $\alpha_{ij} \equiv a+(p_i+\cdots+p_j)^2 = -(i-j)\, a+2\sum_{m,n} p_m \cdot p_n$, eq. (4.1) can be recast into the following form,

$$B_N = \int_0^1 du_{1,2}\cdots du_{1,N-2}\ (u_{1,2})^{-\alpha_{12}-1}\cdots(u_{1,N-2})^{-\alpha_{1,N-2}-1}$$

$$(1-u_{1,2})^{-\alpha_{23}-1}\cdots(1-u_{1,N-2})^{-\alpha_{N-1,N-1}} \qquad (4.5)$$

$$\times \prod_{\substack{i>1 \\ j>i}} (1-u_{1,i}\cdots u_{1,j-1})^{-2\, k_i\cdot k_j}$$

This expression for B_N is useful in factorizing the amplitude. An alternative form, which is very elegant and of importance for some of the deeper aspects of the theory, is due to Koba and Nielsen[18]. Although it is not needed in what follows, I cannot resist describing it briefly. Consider general projective transformations in

the complex plane,

$$z' = \frac{A + Bz}{C + Dz} \tag{4.6}$$

where A,B,C and D are constants. Given four distinct points z_1, z_2, z_3, and z_4, the cross ratio is an invariant under the projective mapping,

$$\frac{(z_1-z_3)\ (z_2-z_4)}{(z_1-z_4)\ (z_2-z_3)} = \frac{(z_1'-z_3')\ (z_2'-z_4')}{(z_1'-z_4')\ (z_2'-z_3')} \tag{4.7}$$

One of the standard results from projective geometry implies that cross ratios satisfy an identity of the form eq. (4.2). We can therefore make the following identification:

$$u_{ij} = \frac{(z_i - z_j)\ (z_{i-1} - z_{j+1})}{(z_i - z_{j+1})\ (z_{i-1} - z_j)}, \tag{4.8}$$

where u's are restricted to the interval between 0 and 1 if the z's are real and ordered on the index i they carry. One can verify the identity (4.2) directly from (4.8), and everything works more easily and elegantly in terms of the new set of variables. Because of projective invariance, it is possible to set $z_1 = 0$, $z_{N-1} = 1$, $z_N = \infty$, since any three points can be mapped into the points 0,1 and ∞ by a suitable projective transformation. With this choice, the relation between the u's and the z's is simple:

$$u_{1,2} = \frac{z_2}{z_3},\ u_{1,3} = \frac{z_3}{z_4},\ \text{--},\ u_{1,N-2} = \frac{z_{N-2}}{z_{N-1}}. \tag{4.9}$$

It then follows that in the integral representation, one fixes three adjacent z's, multiplies the integrand by a suitable weight factor and integrates over all the z's that are not fixed. The range of the integration is any projective image of the real axis, with the ordering of z's preserved. The reader is referred to the original article of Koba and Nielsen for more details.

I would like now to illustrate various properties of B_N for the simple (but non-trivial) case of N = 5. Choosing $u_{1,2}$ and $u_{1,3}$ as independent variables, we have,

$$u_{2,3} = \frac{1 - u_{1,2}}{1 - u_{1,2}u_{1,3}} \qquad u_{3,4} = \frac{1 - u_{1,3}}{1 - u_{1,2}u_{1,3}}$$

$$u_{2,4} = 1 - u_{1,2}u_{1,3} \qquad w = u_{2,4} \tag{4.10}$$

and

$$B_5 = \int_0^1 \int_0^1 du_{1,2}\, du_{1,3}\, (u_{1,2})^{-\alpha_{1,2}-1}\, (u_{1,3})^{-\alpha_{1,3}-1}$$

$$\left(\frac{1 - u_{1,2}}{1 - u_{1,2}u_{1,3}}\right)^{-\alpha_{2,3}-1} \left(\frac{1 - u_{1,3}}{1 - u_{1,2}u_{1,3}}\right)^{-\alpha_{3,4}-1}$$

$$\times \left(1 - u_{1,2}\, u_{1,3}\right)^{-\alpha_{2,4}-2} .$$

Suppose we wish to show that this integral representation looks the same with another choice of independent variables; for example, $u_{1,5}$ and $u_{2,5}$. The following change of variables accomplishes this end,

$$u_{1,5} = 1 - u_{1,2}u_{1,3}$$

$$u_{2,5} = \frac{1 - u_{1,3}}{1 - u_{1,2}u_{1,3}} \tag{4.11}$$

$$\frac{\partial(u_{1,2}u_{1,3})}{\partial(u_{1,5}u_{2,5})} = \frac{u_{1,3}}{1 - u_{1,2}u_{1,3}}$$

Similarly, various other properties of B_N can be explicitly verified in this special case.

I would now like to examine the multiregge limit of the five point function. As an interesting byproduct, we shall also obtain an expression for the coupling of an external scalar particle into two reggeons. The multiregge limit we shall consider is given by:

$$|s_{2,3}| \to \infty \quad |s_{1,5}| \to \infty, \quad \frac{s_{2,3}s_{3,4}}{s_{1,5}} = \kappa = \text{constant}.$$

$$s_{1,2} = \text{const.} \qquad s_{4,5} = \text{const.}$$

In (4.10), make the following change of variables:

$$u_{1,5} = \exp\left(\frac{-xy}{\alpha_{2,3}\alpha_{3,4}}\right), \quad u_{3,4} = \exp\left(\frac{y}{\alpha_{3,4}}\right),$$

which gives, (4.12)

$$B_5 = \int_0^\infty \int_0^\infty dx\, dy\, \frac{y}{\alpha_{2,3}(\alpha_{3,4})^2} \exp\left(\frac{\alpha_{1,5} x\, y}{\alpha_{2,3}\alpha_{3,4}}\right)$$

$$\times \exp(-y) \left\{ \frac{1 - \exp\left(\frac{-xy}{\alpha_{2,3}\alpha_{3,4}}\right)}{1 - \exp\left(\frac{y}{\alpha_{3,4}} - \frac{xy}{\alpha_{2,3}\alpha_{3,4}}\right)} \right\}^{-\alpha_{12}-1}.$$

In the limit $|\alpha_{2,3}| \to \infty$, $|\alpha_{3,4}| \to \infty$,

we expand everything in inverse powers of $\alpha_{2,3}$, $\alpha_{3,4}$ and keep the leading terms only as usual.

This gives us the following leading asymptotic behavior,

$$B_5 \sim (-\alpha_{2,3})^{\alpha_{1,2}} (-\alpha_{3,4})^{\alpha_{4,5}} \qquad (4.13)$$

$$\times \int_0^\infty \int_0^\infty dx\, dy \exp\left\{-x-y + \frac{xy}{\kappa}\right\} x^{-\alpha_{1,2}-1} y^{-\alpha_{4,5}-1}.$$

Of course, all the following manipulations are valid only in the unphysical region $\mathrm{Re}(\alpha_{2,3}) < 0$, $\mathrm{Re}(\alpha_{3,4}) < 0$, $\mathrm{Re}(\kappa) < 0$. The double integral on the right hand side of (4.13) can, however, be continued analytically to the physical region in κ, and surprisingly, one finds a cut on the positive real axis starting at zero. This cut, originally discovered by summing suitable sets of Feynman graphs, seems, therefore, to be a universal feature of the double reggeon-scalar vertex. How does this cut effect factorization, and the definition of signature for regge poles? For this and some other related questions, I refer the interested reader to a recent preprint by J. Weis.[19]

5. Factorization

Consider the multiperipheral graph of Feg. 6, which is to be factorized into two clusters; one cluster consisting of particles numbered from 1 to ℓ, and the other cluster including all the rest. Expanding the integrand of eq. (4.5) in powers of the variable $u_{1,\ell}$, we obtain the following:

$$B_N = \int_0^1 du_{1,2} \cdots du_{1,N-2}\; I_{\ell\text{eft}}(k_1, \cdots k_\ell j\; u_{1,2} \cdots, u_{1\,\ell-1})$$

$$\times I_{\text{right}}(k_{\ell+1}, \cdots, k_N j\; u_{1\,\ell+1} \cdots, u_{1\,N-2}) \times u_{1,\ell}^{-\alpha_1 \ell - 1} (1-u_{1,\ell})^{a-1}$$

$$\times \exp\left\{2 \sum_{i=2}^{\ell} \sum_{j=\ell+1} \sum_{n=1}^{\infty} \frac{k_i \cdot k_j}{n} (u_{1,i} \cdots u_{1,j-1})^n\right\}. \quad (5.1)$$

The factors I_{left} and I_{right} consist of those factors in the integrand which depend only on the momenta and internal variables of the left cluster or the right cluster respectively. The factors which depend on both clusters are written as the exponential of a logarithm, and the logarithm is expanded in a power series in the variable $u_{1,\ell}$.

The next step is to split the last factor in eq. (5.1) into subfactors that depend only on the two clusters separately. To this end, we introduce an infinite set of harmonic oscillator operators[20], a_n^μ and $a_n^{+\mu}$, where the vector index μ takes on the values 0,1,2,3 and n ranges from 1 to ∞. The commutation relations are the following:

$$\left[a_n^\mu,\; a_m^{+\upsilon}\right] = g^{\mu\upsilon}\, \delta_{n,m}. \quad (5.2)$$

In effecting the factorization, the following two identities are crucial,

$$u^{-R} a_n^\mu\, u^R = u^n a_n^\mu, \quad u^{-R} a_n^{+\mu}\, u^R = u^{-n}\, a_n^{+\mu}, \quad (5.3a)$$

where $R = \sum_1^\infty n\; a_n^{+\mu}\; a_{n,\mu}$.

$$\langle 0|\ \exp\ (P.a_n)\ \exp\ \left\{Q.a_n^+\right\}|0\rangle = \exp\ (P.Q) \qquad (5.3b)$$

for any two four vectors P and Q, with a vacuum that satisfies $a_n|0\rangle=0$.

The two identities can be combined to yield the following:

$$\exp\left\{2\sum_{i=2}^{\ell}\sum_{j=\ell+1}^{N-1}\sum_{n=1}^{\infty}\frac{k_i k_j}{n}\ (u_{1,i}\cdots u_{1,j-1})^n\right\}$$

$$= \langle 0|\ \exp\left\{\sum_{n=1}^{\infty}\sqrt{\frac{2}{n}}\ P_n^{(left)}.a_n\right\}\ u^R_{1,\ell}$$

$$\exp\left\{\sum_{m=1}^{\infty}\sqrt{\frac{2}{m}}\ P_{(m)}^{(right)}\ .a_m^+\right\}|0\rangle, \qquad (5.4)$$

where

$$P_{(n)}^{(left)} \equiv \sum_{i=2}^{\ell} k_i (u_{1,i}.u_{1,i+1}\cdots u_{1,\ell-1})^n$$

$$P_{(n)}^{(right)} \equiv \sum_{j=\ell+1}^{N-1} k_j (u_{1,\ell+1}\cdots u_{1,j-1})^n .$$

Substituting eq. (5.4) intc eq. (5.1), and carrying out the integration over $u_{1,\ell}$, we obtain,

$$B_N = \int_0^1 du_{1,2}\cdots du_{1,\ell-1}\ du_{1,\ell+1}\cdots du_{1,N-2} \times I_{left} \times I_{right}$$

$$\times\langle 0|\exp\left\{\sum_{1}^{\infty}\sqrt{\frac{2}{n}}\ P_{(n)}^{(left)}.a_n\right\}\ B(R-\alpha_{1,\ell},a)$$

$$\exp\left\{\sum_{1}^{\infty}\sqrt{\frac{2}{m}}\ P_{(m)}^{(right)}\ .a_m^+\right\}|0\rangle. \qquad (5.5)$$

The two exponentials are the vertices that go with the two clusters, and $B(R-\alpha_{1,\ell},a)$ serves as the propagator. In the special case of unit intercept, a = 1, the

propagator simplifies to $\frac{1}{R-a_1,\ell}$. From eq. (5.5), it is easy to see that there is a great deal of degeneracy at each level. In the simple case of unit intercept, the particle degeneracy for $\alpha=n$ is the same as the number of states of the form $a_{n_1}^{+\mu_1} a_{n_2}^{+\mu_2} \cdots a_{n_k}^{+\mu_k} |0>$, with $n_1+n_2+\ldots+n_k = n$. (Here I am ignoring the complications due to spurious states.) The problem, therefore, reduces to a problem partitioning n into integers, and no simple analytic solution can be given. However, an asymptotic solution exists[20], and shows that the degeneracy goes like $\exp\left\{\frac{2\pi}{\sqrt{6}}\sqrt{n}\right\}$ for large n.

Another property of the spectrum is the existence of ghosts. The lack positive definite metric can be traced to the indefinite sign on the right hand side of the commutation relations (5.2). It can immediately be verified that states that contain an odd number of time like excitation have negative norm and are, therefore, ghosts. The hope, so far, is that these states may be spurious and decouple from physical states. But this is an extensive and complicated topic into which I shall not enter.

Finally, it is possible to establish multiple factorization for an arbitrary number of clusters. Split the left hand cluster of Fig. 6 into two new clusters as in Fig. 7. The identity that corresponds to this split is the following:

$$B_\ell^{(r)} \equiv \int_0^1 du_{1,2} \cdots du_{1,\ell-1} \; I(1,2,\ldots,\ell-1)$$

$$\times \langle 0| \exp\left\{\sum_1^\infty \sqrt{\frac{2}{n}}\, P_{(n)}^{\ell} \cdot a_n\right\} |1r\rangle \qquad (5.6)$$

$$= \int_0^1 du_{1,2} \cdots du_{1,\ell'-1} \; du_{1,\ell+1} \sim \cdot du_{1,\ell-1} \times I(1,2,\ldots,\ell'-1)$$

$$\times I(\ell'+1,\ldots,\ell-1) \langle 0| \exp\left\{\sum_1^\infty \sqrt{\;} \sum_{i=2}^{\ell'} (k_i \cdot a_n)(u_{1,i} \cdots u_{1,\ell'-1})^n\right\}$$

$$\times B(R-\alpha_{1,\ell'},a)\ \exp\Big\{\sum_{1}^{\infty}\sqrt{\frac{2}{m}}\sum_{j=\ell}^{\ell'+1}(k_j\cdot a_m^+)\ (u_{1,j}\text{---}u_{1,\ell'+1})^m\Big\}$$

$$(u_{1,\ell'+1}\text{---}u_{1,\ell-1})^R \qquad \text{(5.6 continued)}$$

$$\exp\Big\{\sum_{1}^{\infty}\sqrt{\frac{2}{p}}\sum_{h=\ell'+1}^{\ell}(k_h\cdot a_p)(u_{1,h}\text{--}u_{1,\ell-1})^p\Big\}|r\rangle.$$

In the above equation, $I(1,2,-,\ell-1)$ is the integrand of the beta function of the first ℓ - 1 variables, and similarly for the other I's. P^{ℓ} is the same as P^{left} of eq. (5.4), and $|r\rangle$ is an arbitrary state. The equation is derived by factorizing through the use of identities (5.3) and carrying out the integration over the variable $u_{1,\ell'}$. This process can be repeated until each cluster consists of only one leg. This gives us a completely factorized expression for B_N,

$$B_N = \langle 0|\ V(k_2)\Delta(\alpha_{1,2})\ V(k_3)\text{---}\Delta(\alpha_{1\,N-2})V(k_{N-1})\ |0\rangle, \qquad (5.7)$$

where

$$\Delta(\alpha) \equiv B(-\alpha,a),$$

$$V(k) = \exp\Big\{\sqrt{2}\ k\cdot\sum_{1}^{\infty}\frac{a_n^+}{n}\Big\}\exp\Big\{\sqrt{2}\ k\cdot\sum_{1}^{\infty}\frac{a_m^+}{m}\Big\}.$$

Eq. (5.7) is the starting point for the investigation of many of the deeper properties of dual resonance models. To the interested reader, who wants to pursue any one of the topics barely touched upon here, I recommend various more complete review articles; among them Mandelstam's lectures at the Brandeis summer school in 1970, lectures by Jacob in the same summer school and the review article by Sivers and Yellin to appear in Reviews of Modern Physics.

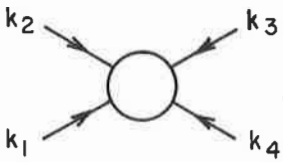


Figure 1

Two Particle Scattering Amplitude

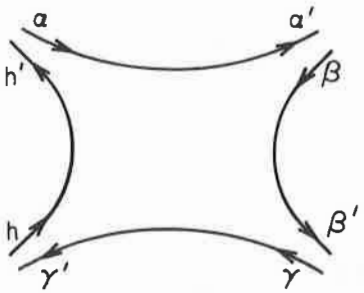


Figure 2

Quark Diagram for
Meson-Meson Scattering

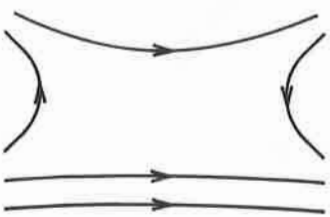


Figure 3

Quark Diagram for
Meson-Baryon Scattering

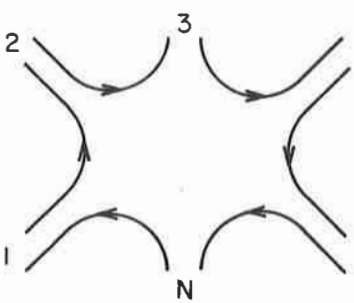


Figure 4

Quark Diagram for
the N-Point Amplitude

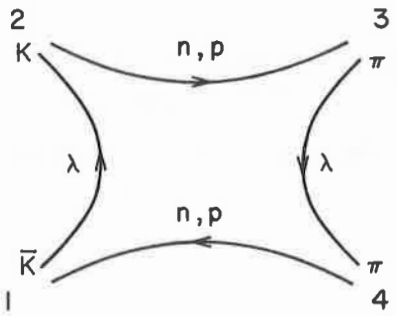


Figure 5

$K+\pi \to K+\pi$ Amplitude

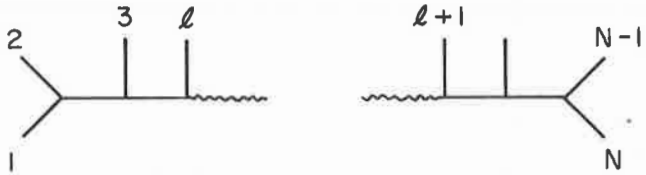


Figure 6

Factorization into Two Clusters

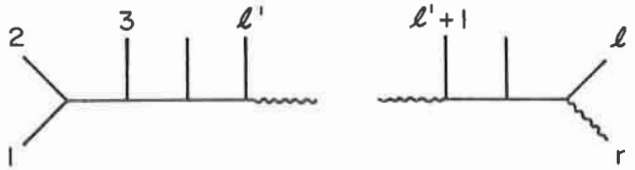


Figure 7

Multiple Factorization

References

1. V. de Alfaro, S. Fubini, G. Furlan and G. Rosetti, Phys. Letters 21, 576 (1966).

2. R. Dolen, D. Horn and C. Schmid, Phys. Rev. 166, 1768 (1968).

3. K. Igi and S. Matsuda, Phys. Rev. Letters 18, 625 (1967). A. Logunov, L.D. Soloviev and A.N. Tarkelidze Phys. Letters 24 B, 181 (1967).

4. H. Harari, Phys. Rev. Letters 20, 1385 (1968). P.G.O. Freund, Phys. Rev. Letters 20, 235 (1968).

5. H. Harari, Phys. Rev. Letters 22, 562 (1969). J. Rosner, Phys. Rev. Letters 22, 689 (1969). T. Matsuoka, K. Ninomiya and S. Sawada, Prog. Theor. Phys. 42, 56 (1969).

6. R. Capps, Phys. Rev. Letters 22, 1215 (1969). J. Mandula, J. Weyers and G. Zweig, Ann. Rev. Nucl. Science (To be published 1970).

7. H.M. Chan and J. Paton, Nuclear Phys. B 10, 519 (1969).

8. G. Veneziano, Nuovo Cimento 57A, 190 (1968).

9. S. Mandelstam, Phys. Rev. Letters 21, 1724 (1968).

10. M. Ademollo, H.R. Rubinstein, G. Veneziano and M.A. Virasoro, Phys. Letters 27 B, 99 (1968) and Phys. Rev. 176, 1904 (1968).

11. C. Lovelace, Phys. Letters 28 B, 265 (1968). J. Shapiro, Phys. Rev. 179, 1345 (1969).

12. K. Kawarabayashi, S. Kitsakado and H. Yabuki, Phys. Letters 28 B 432 (1968).

13. See, for example, H. Rubinstein and V. Rittenberg and J. Doren, Nuclear Phys. 20B, 663 (1970).

14. R. Carlitz and M.Kislinger, Phys. Rev. Letters 24, 186 (1970).

15. E.L. Berger and G. Fox, Argonne Preprint.

16. S. Mandelstam, Phys. Rev. 184, 1625 (1969).

17. For a review of this subject, see H.M. Chan, Royal Society Meeting (1969).

18. Z. Koba and H.B. Nielsen, Nuclear Physics B12, 517 (1969).

19. J. Weis, M.I.T. preprint.

20. S. Fubini, D. Gordon and G. Veneziano, Phys. Rev. Letters 29B, 679 (1969).

THE PRESENT STATUS OF DUAL MODELS

E. Squires
University of Durham
Durham, England

I. Introduction

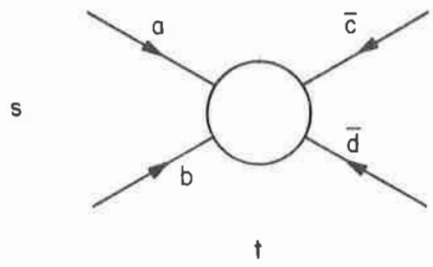
It is the purpose of these lectures to discuss how far the ideas of duality, particularly as expressed through dual models, are supported by experiment and also to what extent the ideas can be formulated in an internally consistent way.

In the next chapter we discuss the concept of duality, and see how far the various ingredients of duality are experimentally supported. Dual models are discribed briefly in the third chapter where in particular some of the difficulties of extending dual models to realistic particles are discussed. The fourth chapter contains the confrontation of the predictions of dual models with an experiment. We endeavor to see how far the constraints inherent in dual models are supported by the data. Finally, in the last chapter, we treat some particular topics in the application of duality ideas to'inclusive reactions' - a topic which is new and offers exciting prospects.

In these lectures we do not discuss the problem of dual models as a starting point for a complete theory- thus dual loops and twisted loops etc. will only be mentioned briefly in the last chapter as a model for the Pomeron. These topics are treated elsewhere in the volume by several authors. Our approach to dual models is rather that of regarding them as a simple approximation to physical amplitudes - an approximation that, in contrast to most models, is 'global', i.e. expected to apply, at least qualitatively, over the whole range of the variables.

II. The Ingredients of Duality

We consider the process a+b → c+d depicted as:



This is described by the amplitude A(s,t) normalized so that

$$\frac{d\sigma}{d\Omega} = \frac{1}{(8\pi)^2}\,\frac{1}{s}\,\frac{q_{cd}}{q_{ab}}\,|A(s,t)|^2 \qquad 2.1$$

where q_{cd} and q_{ab} are the momenta in the ab and cd center-of-mass systems respectively. "Duality" is a set of assumptions about A(s,t), which we now discuss.

First Ingredient: "Resonance-saturation"

We assume, in the s-channel physical region,

$$\text{Im } A \simeq \text{Im}A^{\text{Res}} + \text{Im}A^{\text{Pom}}, \qquad 2.2$$

where the two forms on the r.h.s. have the following properties:

(i) $\text{Im } A^{\text{Pom}} = 0$ unless the additive quantum numbers of a are equal to the additive quantum numbers of c. (Note that we are here considering the "forward peak" corresponding to $|t| < |u|$; if we wanted to consider the backward region then the roles of c and d would be

interchanged). Also we require

$$\frac{\text{Im } A^{\text{Pom}}}{s} \underset{s \to \infty}{\to} \text{constant} \qquad 2.3$$

In fact this statement has no experimental significance and, in application, is interpreted as

$$\frac{\text{Im } A^{\text{Pom}}}{s} \simeq \text{constant}, \; s \gtrsim \text{few } (\text{GeV})^2 . \qquad 2.4$$

(ii) $$\frac{\text{Im } A^{\text{Res}}}{s} \underset{s \to \infty}{\to} 0$$

(iii) $$\text{Im } A^{\text{Res}} = \sum_{\text{poles}} \text{Im} \left[\frac{Ri}{(s - s_i + i\sqrt{s_i}\Gamma_i)} \right] (2\ell+1) \, P_\ell(t) \qquad 2.5$$

where we have written this expression for the case where the external particles are spinless. Note that the particular form for the pole makes Γ_i the "width" in the conventional sense; provided the Γ_i are small the precise form of the pole is not too important.

This expression of course has no content unless we say something about the values of s_i and Γ_i since <u>any</u> function $\text{Im}A^{\text{Res}}$ can be respresented by a sum of poles to arbitrary accuracy! We can try to give it content by adding

(iv) The poles really exist. This however is a statement without significance since we cannot do experiments off the real s-axis and we have no theory to tell us where the poles are (see our second ingredient however). Of more use is

(v) The resonances have narrow widths ($\Gamma \lesssim 100$ rev), are well separated for each J($|s_i - s_i + 1| \ll \Gamma^2$ for poles of the same J), and there are no "ancestors", i.e. in a plot of J against s all poles lie below a line $J = a\,s + b$; see fig. 2.1.

Before we consider the experimental tests of eq. 2.1 we note that the equation is only assumed to be approximate. We know it cannot be exact since if it were exact we could use a dispersion relation to compute ReA and we would then know A completely. It would not satisfy unitarity even in this elastic region, still less in the region of s for which $a + b \to x + \bar{x}$ would be possible where for x the reader can choose any macroscopic object which takes his fancy. (If we abandoned the possibility of writing a dispersion relation then maybe we could take eq. 2.1 as an exact equation).

It is rather unsatisfactory to have to begin with an equation which is approximate since we cannot be sure how good the approximation should be and experimental tests are hard to interpret. In the absence of any theory underlying eq. 2.1, we do not know of any "limit" in which the equation should be exact, and of course have no idea how to construct higher order corrections.

Turning to the tests, there are essentially only three.
(a) All cross-sections where the quantum numbers of a are not equal to the quantum numbers of b should lead to zero. This of course only tests (i) and (ii) above. It is true in all cases.
(b) The only quantitative test is due to Harari and Zarmi (1969) who considered πN elastic scattering and, using partial wave analyses of the data, plotted the Argand diagrams for $I_t = 0$ and 1, where I_t is the t-channel isotopic spin. Their results are shown in fig. 2.2. We see that for $I_t = 1$ (for which there is no A^{Pom} contribution) the partial-waves behave very much like a series of circles which close on themselves (this is the behavior expected for a set of narrow, well spread resonances), whereas for $I_t = 0$ this is not the case, thus showing the presence of the A^{Pom} background. In a later paper (Harari and Zarmi (1970)) they combined the $I_t = 0$ partial-waves into forms for which the s-channel nucleon helicity is changed and those for which it is unchanged:

$$F^{J}_{\pm +} = \tfrac{1}{2}\left[f_{(J - \frac{1}{2})+} \pm f_{(J+\frac{1}{2})-} \right]. \qquad (2.6)$$

The new Argand plots are shown in fig. 2.3 and we see that Im A^{Pom} appears to be absent also in the helicity flip term. This has some support from other processes although this is disputed.

(c) Provided the resonances are narrow, well separated and not too small in comparison to Im A^{Pom}, then each particular resonance will dominate the appropriate partial wave in the neighborhood of $s \approx s_i$. Then, by unitarity for elastic processes we will deduce $R_i > 0$. Thus, using the optical theorem:

$$\sigma^{Tot} = \frac{1}{2q_{ab}\sqrt{s}} \quad \text{Im } A(s,o), \qquad (2.7)$$

all total cross-sections should approach their high-s limit from above. This is true for the "major" effects (there is evidence that at Serpukov energies the K^+p total cross-section rises by about 1 mb - we regard this as a "minor" effect!)

In conclusion on the first ingredient of duality we note two points. First, there are really no direct tests (i.e. not employing the other ingredients) for inelastic processes. Secondly, it is not surprising that one only makes this assumption for the imaginary parts because the imaginary part of a pole contribution falls off much faster, as one moves away from the pole, than the real part, so the latter will in general have significant contribtuions from u-channel poles.

Second Ingredient: No "exotics."

An exotic particle is one that cannot be made of either $(q\bar{q})$ or (qqq), where the quarks have the usual quantum numbers. There are three tests.

(a) No particles or resonances with exotic quantum numbers have been observed. In particular exotic channels are free from significant peaks in their mass distributions. Clearly the deuteron and all objects with $B > 1$ are counter examples. Their existance means that we should qualify this ingredient by asserting that all exotic particles are "low-lying" and therefore, hopefully,

unimportant.

If we combine ingredients 1 and 2 then we obtain two other predictions which can be tested.

(b) If we have a process for which the s-channel is exotic, Im $A^{Res} = 0$, and if the t-channel is such that Im $A^{Pom} = 0$, then Im $A = 0$. Such a process is $K^{\circ}p \rightarrow K^{+}n$ for which we have the result [Firestone et al. (1971)]:

$$\frac{\text{Im } A(t=o)}{\text{Re } A(t=o)} < 4\ \%$$

This is a remarkable result which seems better than we could have expected!

(c) If we consider elastic processes for which the s- channel is exotic then, by the optical theorem, they should have constant total cross-sections. The behaviors of known total cross-sections is given in Fig. 2.4 and is in accordance with this prediction.

Third Ingredient: Regge Asymptotic Behaviour for A^{Res}

We shall not discuss this in detail here (see Squires (1970) and Collins (1971) for recent reviews) but some points are worth emphasizing.

The simple predictions of Regge theory are very well confirmed, namely, those that depend on the existence of linear, approximately parallel, trajectories corresponding to the known particles. In particular, where the t-channel is exotic, peaks are either absent or are very small (consistent with low lying exotic particles.) Also the shrinkage expected from linear trajectories is evident even out to large $|t|$. This is illustrated by the analysis of Daum, Michael and Schmid (1970) (and further unpublished work by the same authors discussed by Schmid (1971)) on $K^{\pm}p$ scattering (see fig. 2.5), by the work of Barger and Phillips (1971) on the ρ trajectories in $\pi\ \rho \rightarrow \pi^{\circ}n$ (fig. 2.6) and by the data of Brabson et al. (1970) on the "effective α" for $\pi^{\pm}$ scattering out at least to $|t| \sim 3(\text{GeV})^2$ (fig. 2.7).)

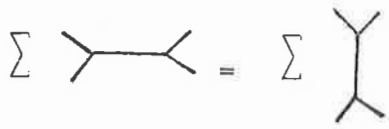
There are several places where the known poles,

without conspiracy and with factorisation (i.e. assuming the poles are simple) fail badly, in particular, $\pi^- p \to \pi^0 n$ polarisation, π contribution in forward direction of $np \to pn$ and $\gamma p \to \pi^{\pm} n$ and cross-over zeros.

3. If one breaks factorisation by including "cuts" we can trivially solve these problems. However, if we use the conventional model of cuts, namely the absorptive model, one fails to fit the data without introducing free parameters (which, in the opinion of the author, make the model equivalent to the data.) Particular failures are that the absorptive model cannot reproduce the large shrinkage shown in figs. 2.5, 2.6, and 2.7 ($\sigma_{cut}(t)$ has a much smaller slope than $\sigma_{pole}(t)$), it cannot explain the new $\pi p \to \pi^0 n$ polarisation data (see Coleman Johnson (1971)) and it cannot fit the π contribution to $\gamma p \to \pi^+ n$ without having to boost the absorption to an unreasonable level.

4. Experiment requires certain Regge contributions to cancel. This leads to "exchange degeneracy." For example, the reality of $K^0 p \to K^+ n$ noted above requires that the ρ and A_2 have the same trajectories and residues in the process. Similarly the experimental observation that exotic total cross-sections are constant requires cancellation among the contributing Regge trajectories. Using the first two ingredients of duality we can also make predictions without recourse to experiments. For example, in $\pi^+\pi^+ \to \pi^+\pi^+$, since the s-channel is exotic, there must be no Regge contribution to Im A^{Res}, hence the ρ and f must exactly cancel and therefore be exchange degenerate in trajectory and residue. All these exchange degeneracy requirements are compatible with the masses of the known particles in the trajectories.

The relation between s-channel resonances and t-channel Regge poles which comes from combining ingredients 1 and 3 is linear - in contrast to bootstrap type theories. It can be depicted diagramatically as:



Clearly this can in principle run into difficulties with factorisation unless there are constraints on the trajectories (hence the exchange degeneracy predictions). If we restrict ourselves to Meson-Meson and Meson-Baryan scattering then the conventional SU(3) $\underline{1} + \underline{8}$ mesons and $\underline{1} + \underline{8} + \underline{10}$ baryon forms a compatible "solution," with a well defined F/D ratio. However, if we include Baryon-Baryon scattering, we cannot have a solution without exotics (in $B\bar{B}$ scattering we have $\underline{10}$, $\overline{\underline{10}}$ and $\underline{27}$ which are exotic and the cross-channel trajectories should cancel these three states - there are not enough parameters to make this possible).

Even in the M-B case there are problems. For example consider $KM \to KM$. Since this is exotic we predict exchange degeneracy in the cross-channels, in particular for the Y* baryons in $\bar{K}M \to \bar{K}M$. However these same Y* baryons will occur in, for example, $\pi\Lambda \to \pi\Lambda$ for which neither cross-channel is exotic and for which we therefore should not have exchange degeneracy. Schmid (1970) has shown how this problem is solved in an approximate way. It appears that there Y* which are strongly coupled to the KN system are exchange degenerate, and that in addition there are further Y* particles which are coupled to $\pi\Lambda$ but have only weak coupling to $\bar{K}N$.

Although some of the consequences of exchange degeneracy, e.g. reality of $K^0p \to K^+n$ mentioned above, are experimentally confirmed, there are others that fail. For example, exchange degeneracy of trajectories (regardless of residues) is sufficient to ensure the equality of $K^-n \to \pi^-\Lambda$ and $\pi^-p \to K^0\Lambda$, whereas in fact the cross-section for the former is about twice that of the latter. Other "line-reversed" reactions show the same difficulty.

Absorptive cuts are of no help here since they have the opposite effect to that required. It is worth noting that, as with many of the problems of Regge theory [Squires,(1971)], this difficulty is caused by factorisation; in those cases where one channel is exotic, so that exchange degeneracy follows directly from duality, the line reversal predictions are well satisfied by the data.

We have not so far discussed Regge-Regge cuts, i.e. cuts caused by the exchange of two Reggeon. Harari (1971) has discussed evidence for these and given reasons for believing they are important. The experimental evidence for shrinkage at large $|t|$ suggests that their effective α must be not too different from the single Regge pole, at least at present energies. They should show up where the t-channel is exotic, but in some cases there may be problems. For example (A. Martin, private communication), because of exchange degeneracy one would expect the exotic line-reversed reactions $K^-p \to \pi^+\Sigma^-$ and $\pi^-p \to K^+\Sigma^-$ to have equal cross-sections if they are due to Regge-Regge cuts - the data however suggests a factor 10 between them.

Fourth Ingredient : Analyticity

So far we have not explicitly used this. If we combine it with Regge behaviour we obtain finite-energy-sum-rules (FESR). To obtain these we subtract from an amplitude with no A^{Pom}, the Regge pole contributions with $\alpha_i(t)>-1$, so that the resulting function goes to zero with s faster than s^{-1}, i.e.

$$A(s,t) - \sum_{\alpha_i>-1} \beta_i(t)\,(s-s_0)^{\alpha_i(t)} = 0(s^{-1-\epsilon}),\ \epsilon>0 \qquad (2.8)$$

It follows that this expression satisfies a superconvergence relation:

$$\int ds\ \mathrm{Im}\left[A(s,t) - \sum_{\alpha_i>-1} \beta_i(t)\,(s-s_0)^{\alpha_i(t)}\right] = 0 \qquad (2.9)$$

The integral is over the cut in the s-plane, which in general includes the right-hand, physical, s-cut, a left-hand cut, due to the u-channel, and bound-state poles. We consider for simplicity just the right-hand cut (the other is similarly treated), and we separate this into a cut from threshold s_0 to some value s_1 and a remainder, where s_1 is chosen so that

$$A(s,t) \simeq \sum_i \beta_i(t)(s-S_0)^{\alpha_i(t)},\ s>s_1 \qquad (2.10)$$

where we include a selected number of the leading Regge-pole terms (how many are required will depend upon s_1 of course). Then eq. (2.9) becomes

$$\int_{s_0}^{s_1} \mathrm{Im}\left[A - \sum_{\alpha_i > -1} \beta_i(t)\,\alpha_i(t)\right] ds + \int_{s_1}^{\infty} \mathrm{Im} \sum_{\alpha_i < -1} \beta_i(t)\,(s-s_0)^{\alpha_i(t)}\, ds \simeq 0 \qquad (2.11)$$

The integrals over the Regge terms can be evaluated and we obtain finally

$$\int_{s_0}^{s_1} \mathrm{Im}\, A(s,t)\, ds \simeq \mathrm{Im} \sum_i \beta_i(t)\, \frac{(s-s_0)^{\alpha_i(t)+1}}{\alpha_i(t)+1} \qquad (2.12)$$

This calculation could have been performed equally well with A replaced by $(s-s_0)^n A(s,t)$, for n=0,1,2 -- This leads to moment sum rules

$$\int_{s_0}^{s_1} (s-s_0)^n\, \mathrm{Im}\, A(s,t)\, ds \simeq \mathrm{Im} \sum_i \beta_i(t)\, \frac{(s-s_0)^{\alpha_i+n+1}}{\alpha_i+n+1} \qquad (2.13)$$

Finally, if we add the first ingredient we obtain

$$\int_{s_0}^{s_1} (s-s_0)^n\, \mathrm{Im}\, A^{\mathrm{Res}}(s,t)\, ds \simeq \mathrm{Im} \sum_i \beta_i(t)\, \frac{(s_1-s_0)^{\alpha_i+n+1}}{\alpha_i+n+1} \qquad (2.14)$$

Great care is required in using these relations. If they were interpreted as equalities (rather than approximate relations), then the fact that they hold for all n would imply equality of the Regge term and the resonance term at <u>all</u> s. This is known as "<u>local duality</u>" - in its extreme

form it is clearly false, and some averaging is necessary. If one averages up to some value s_1, where Regge behavior is good, and takes the lowest non-trivial value of n (o or 1) then one obtains what is normally referred to as "global duality."

Since equations of the form of eq. (2.14) relate s-channel poles to t-channel poles they clearly can be used to form a "bootstrap" theory - a linear bootstrap in contrast to the conventional one using unitarity. In using these relations it has been usual to begin by assuming $\alpha(t)$ to be linear and $\beta(t)$ to contain a kinematic factor but to be otherwise constant. A particular interesting case is the $\pi\pi\to\pi w$ system which has been studied extensively by Ademollo et al.(1967) (and further references stated in this paper.) By putting in the ρ trajectory they obtained a satisfactory self-consistent solution over a small range of $|t|$. To go further they needed other trajectories ("daughters").

The question arises in "bootstraps" of this sort as to how far the "solutions" are unique. The answer was given by Veneziano (1968) who showed that the FESR have a class of solutions which can be written in closed form and which have an infinite number of free parameters. This led to "dual models".

III. Dual (Veneziano) Models

The solution to the $\pi\pi\to\pi w$ FESR was given by Veneziano (1968) as

$$A + \frac{\Gamma(1-\alpha_s)\,\Gamma(1-\alpha_t)}{\Gamma(1-\alpha_s-\alpha_t)} + (t,u) + (s,u) \qquad (3.1)$$

where

$$\alpha = a + bs \qquad (3.2)$$

This has the correct Regge behavior and it has poles at $\alpha_s = 1$ (spins 0 and 1), $\alpha_s = 2$ (spins 0, 1, and 2), $\alpha_s = 3$ (spins 0, 1, 2, and 3), etc. Thus it has a linear trajectory with parallel, integer spaced daughters. Clearly however it is not unique and can be replaced by the general

form

$$A = \sum_{n,m\geq 1} C_{nm} \frac{\Gamma(n - \alpha_s)\,\Gamma(n - \alpha_t)}{\Gamma(n + m - \alpha_s - \alpha_t)} + (t,u) + (s,u) \qquad (3.3)$$

The terms with n,m ± 1 are called 'satellites'. A satellite contributes only to poles with $\alpha_s \geq n$, and only to those trajectories below the $(m - 2)^{th}$ daughter level. It is clear that this non-uniqueness means that the residues of all poles are independent, i.e. given a set of residues one can choose a unique set of Cnm's to fit them. Thus without some extra restrictions the duality constraints on residues are non-existent.

Under certain conditions, in particular meromorphy, crossing and a suitable bound for large values of the variables, the general form given in eq. 3.3 is unique, i.e. the arbitrary set of parameters Cnm gives the maximum ambiguity (Tiketopolous 1970).

At this stage we generalize the dual model to a general N-point function. To do this we consider not the amplitudes in eq. 3.1 but the β-function:

$$B_4\ (\alpha_s,\alpha_t) = \frac{\Gamma(-\alpha_s)\ \Gamma(-\alpha_t)}{\Gamma(-\alpha_s - \alpha_t)} \qquad (3.4)$$

$$\equiv \int_o^1 u^{-\alpha_s-1} (1 - u)^{-\alpha_t-1}\, du \qquad (3.5)$$

$$\equiv \int_o^1 du \int_o^1 dv\, u^{-\alpha_s-1}\ v^{-\alpha_t-1}\, S(u+r-1) \qquad (3.6)$$

We label the external lines of our N-point function i = 1,2, -- N_1 corresponding to a particular ordering, and define

$$s_{ij} = (p_i + p_j)^2 \qquad (3.7)$$

$$\alpha_{ij} = a + b\, s_{ij} \qquad (3.8)$$

Then the generalisation of eq. 3.6 is

$$B_N\ (\alpha_{12}, \text{- -}\quad) = \int_0^1 \prod_{\substack{\text{All}\\ \text{pairs}}} \left[du_{ij}\ u_{ij}^{-\alpha_{ij}-1}\right]$$

$$\times \prod_{\substack{k\ell\\ \text{(a set of non-dual pairs)}}} S\Big(u_{k\ell} + \prod_{\substack{\text{(All pairs dual to } k\ell)}} u_{mn} - 1\Big) \qquad (3.9)$$

Variables are dual if it is not possible to have poles simultaneously in both. Thus at a pole in one variable the dual variables correspond to momentum transfers and the requirement of no simultaneous poles is necessary in order to insure that the residue of a pole is a polynomial in the momentum transfer (i.e. there are no ancestors).

In the particular case of N=5 we have

$$B_5 = \int_0^1 du_{12}\ du_{45}\ u_{12}^{-\alpha_{12}-1}\ u_{45}^{-\alpha_{45}-1}\ (1 - u_{12})^{-\alpha_{23}-1}$$

$$(1 - u_{45})^{-\alpha_{34}-1}\quad (1 - u_{12}\ u_{45})^{-\alpha_{15}+\alpha_{23}+\alpha_{31}+1} \qquad (3.10)$$

Here we have chosen one particular two non-dual pairs, i.e. (12) and (45), so that eq. 3.10 appears to be non-cyclic. However, one obtains exactly the same result whichever set one chooses so that B_5 and in general B_N are cyclically invariant. Of course, to obtain the amplitude it is necessary to add B_N's with other non-cyclic orderings of the external lines (cf. eq. (3.1)).

We now discuss some of the important properties of B_N, considering in particular B_5. To do this we first expand

$(1-u_{12}\ u_{45})^{-\alpha_{15}+\alpha_{23}+\alpha_{31}+1}$ in equation 3.10 by

by the binomial theorem:

$$B_5 = \sum_{k=0}^{\infty} (-1)^k \binom{-\alpha_{15}+\alpha_{23}+\alpha_{34}+1}{k} B_4(\alpha_{12} - k, \alpha_{23})$$

$$B_4 (\alpha_{45}-k, \alpha_{34}) \qquad (3.11)$$

Near $\alpha_{12} = n$ we have the pole of the Γ function in B_4. This gives:

$$B_5 \simeq \frac{(-1)^n}{n - \alpha_{12}} \sum_{k=0}^{n} \binom{-\alpha_{15}+\alpha_{23}+\alpha_{34}+1}{k} \binom{-\alpha_{23}-1}{n-k}$$

$$B_4(k-\alpha_{45}, \alpha_{34}) \qquad (3.12)$$

It is easy to see that this has exactly the same structure of poles as B_4, namely that shown in fig. 3.1.

Further, the residue of a pole factorises as shown in fig. 3.2, the amplitude for the intermediate state at $\alpha_{12}=n$ interacting with particles 3,4 and 5 being a B_4 function. This property ("bootstrap consistency") is of course vital for the self-consistency of the theory since we wish to identify the intermediate states with the external particles. However we note that, except for particles on the leading trajectory (where only k=o contributes) one obtains a sum of B_4 terms, i.e. satellites are essential. Further, if we considered not B_5 but B_N, the residue of a given state would not factorise but would contain a <u>finite</u> sum of factored terms. Thus the daughters correspond not to single particles but are degenerate - the degree of degeneracy increases as we go down the daughter sequence, but is always finite (independent of the number of external lines!).

As with B_4, B_N has the correct Regge behavior in all the possible Regge limits. For example, in the limit s_{34}, s_{45} large with s_{12} = const. $\times s_{34}\, s_{45}$ we have (see fig. 3.3)

$$B_5 \sim s_{34}^{\alpha_{23}}\, s_{45}^{\alpha_{15}} \qquad (3.13)$$

In making use of these dual models to compare with data there are some troubles.

(i) The satellite ambiguity. As we noted above we can add arbitrary satellites - this is true for B_N as well as for B_4. It is clearly not possible to demand that we have no satellites for any external particles because as we have seen no-satellites for B_5 produces satellites for B_4. However we can hope to postulate that there will be no satellite when external particles lie on the leading trajectory. Even this turns out to be not completely possible and it is replaced by the notion of "maximal duality", which we define below.

(ii) The amplitude is not unitary. This shows itself in the fact that it has real axis poles, whereas in reality the poles should be on the second sheet. Also some of the particles in elastic amplitudes turn out to have negative residues (i.e. are "ghosts"). A possible cure is to replace $\alpha(s)$ by a function having a right hand cut. Since the residue of a t-pole is a polynomial in $\alpha(s)$, it will not then be a polynomial in s, i.e. we will have "ancestors". This is not usually considered acceptable.

A more sophisticated cure is to regard the Veneziano amplitude as the "Born term" of a pertubation series. This topic is treated elsewhere in this book and we will not discuss it further except to note that this programme has not yet been successful and it is not clear whether it is possible without destroying the successes of the "Born term".

(iii) We do not know how to include spin $\frac{1}{2}$ particles. One trouble here is that any model of a Fermion linear trajectory naturally gives parity doublets. Since some of these do not exist it is necessary to add satellites to arrange that their residues are zero.

(iv) For many trajectories, $\alpha(o) > 0$. Now B_4 given in eq. 3.4 has a pole at $\alpha(s) = 0$, which, if $\alpha(o) > 0$, will have an imaginary mass ($s < o$), i.e. be a "tachyon." These do not exist and must be removed. For the four point function we did this already in the $\pi\pi \to \pi w$ case in eq. 3.1. However, the functions in eq. 3.1 are <u>not</u> B_4 functions

$\left(B_4\ (\alpha_s - 1, \alpha_t - 1) = \frac{\Gamma(1-\alpha_s)\ \Gamma(1-\alpha_t)}{\Gamma(2 - \alpha_s - \alpha_t)}\right)$, so it is not clear how to generalise this.

To see what is required we consider first the 5 point ($\pi\pi\pi\pi\sigma$) amplitude. If we call the σ particle 1 then the simplest choice for the amplitude would be

$B_5\ (\alpha_\pi^{12}, \alpha_\rho^{23}, \alpha_\rho^{34}, \alpha_\rho^{45}, \alpha_\rho^{15})$, but this would give tachyons at $\alpha_\rho = 0$. There is however a simple solution (Waltz 1970), namely, the function

$$\alpha_\rho^{34} \times B_5(\alpha_\pi^{12}, \alpha_\rho^{23}-1, \alpha_\rho^{34}, \alpha_\rho^{45}-1, \alpha_\pi^{15}).$$

The tachyons in (23) and (45) are removed by the replacement of α_ρ by $\alpha_\rho - 1$ in these variables; that in (34) is removed by the factor α_ρ^{34}. In addition this amplitude still has the leading trajectory in all variables. Although we have apparently lowered the ρ trajectory by one unit in (23), the factor α_ρ^{34} outside provides the extra one power in the momentum transfer to restore it. Similarly in (45). Note however that since the σ is a daughter we would expect to have to add satellites in general.

As a second example we consider the 6π amplitude. Here we can have either abnormal parity (π) states, normal parity states (w, A_2) in the 3π channels. To obtain the latter we need a negative purity factor outside. A possible solution proposed by Dorren et al (1970) is

$$\epsilon_{\mu\nu\rho\sigma}\epsilon_{\alpha\beta\gamma\sigma}P_{1\mu}P_{2\nu}P_{3\rho}P_{4\alpha}P_{5\beta}P_{6\gamma} \times$$

$$B_6(\alpha_{12}-1, \alpha_{23}-1, \alpha_{34}-2, \alpha_{45}-1, \alpha_{56}-1, \alpha_{61}-2, \alpha_{123}-1, \alpha_{234}-2, \alpha_{345}-2).$$

The factor "$\epsilon_{123}\epsilon_{456}$" behaves as a 1^- particle in the 123 channel and a $2^-, 2^+$ combination in the 234, 354 channels. Hence each of the three terms in eq. 3.14 is leading in all variables. To verify this for the high energy behavior requires care; it is necessary to utilise the fact that the 9 variables explicitly shown in eq. 3.16 are not independent - because of the Gram determinant conditions there are only 8 independent variables for the 6 point function [see Dorren et al (1970)].

Note that it is necessary to use a factor outside which is not cyclically symmetric, so it is essential to add the cyclically rotated terms as in eq. 3.14. The non-cyclic permutations (different orderings of the external lines must of course also be added).

Unfortunately, as noted by Dorren et al (1970) eq. 3.16 will not do since it has parity doublets along the leading (A_2,w) trajectory, e.g. the first term has parity doublets in the 234 and 345 channels. We can remove these, from the leading trajectory only by replacing $\alpha_{234}-2$ and $\alpha_{345}-2$ by $\alpha_{234}-3$ and $\alpha_{345}-3$ respectively in the first term of eq. 2.14, and similarly in the other terms. This means that the individual terms are not leading in all variables, and duality now holds in the sense that the leading direct channel resonances are dual to non-leading Regge-pole contributions. As far as the leading trajectories are concerned our model is thus more like an "interference-model" where we add direct channel poles and cross-channel Regge poles.

A feature that we might look for in models of this type is possible non-factorisation of the leading trajectory (Rittenberg and Rubinstein (1970)) leading perhaps to a split A_2. For example in eq. 3.14 it is not obvious that the (w,A_2) trajectory is the (123) channel of the first term, and that in the (234) and (345) channels are the same. The need to remove the leading trajectory as noted above eliminates this possibility, however, for the 6π amplitude.

Further work on the $N\pi$ amplitude (e.g. Dorren et al. (1971)) has uncovered many additional problems. In particular it does not appear possible to maintain finite factorisation of the daughters, i.e. the degree of degeneracy increases without limit as the number of external particles is increased. From the formal point of view this is a serious problem, but we shall not discuss it further but shall continue with an attempt to compare the simple predictions with experiment.

IV. Comparison with Experiment.

Dual models, in constrast to many phenomenological models are very ambitious in that they apply to the whole range of the variables. Thus they contain many experimental predictions and comparisons with experiment should be easy and might be expected to lead unambiguously to their acceptance or rejection.

However, in practice, it is not so simple because

(i) The model in its simple form clearly predicts Regge behavior - with all the troubles that brings

(ii) The model contains real axis poles, which must be moved into the complex plane before comparison with experiment is made. This introduces some free parameters.

(iii) All experiments involve nucleons, with spin $\frac{1}{2}$, and we do not know how to include spin $\frac{1}{2}$ particles in the model.

(iv) We have only a model for A^{Res} (in eq. 2.1), so to fit data we need to add A^{Pom}. We do not know the precise form of this.

The first two problems are related to the lack of unitarity in dual models and various prescriptions to incorporate unitarity lead to absorptive corrections (which as we have noted do not help much) and to poles off the real axis.

It is convenient to group the predictions of dual models into the following categories:

A. Relations between couplings $\begin{cases} \beta(t), t>0 \\ \text{parent} \mid \text{daughter} \end{cases}$

B. Relations between 'crossed' processes

C. Regge-residues: $\beta(t)$, $t<0$

D. Daughter structure.

We consider tests of these predictions in turn.

A. First we consider this for $\pi\pi \to \pi\pi$ for which the amplitude has the form:

$$\begin{aligned} A^{I_s=2} &= V(t,u) \\ A^{I_s=1} &= V(s,t) - V(s,u) \\ A^{I_s=0} &= \tfrac{3}{2}\,[V(s,t) + V(s,u)] - \tfrac{1}{2}\,V(t,u) \end{aligned} \qquad (4.1)$$

with

$$V(s,t) = \beta\,\frac{\Gamma(1-\alpha_s)\,\Gamma(1-\alpha_t)}{\Gamma(1-\alpha_s-\alpha_t)}\,, \qquad (4.2)$$

α being the ρ trajectory. If we use SU(3) then we can extend this to include all members of the meson octets. It is possible to read off coupling constants of the various mesons directly from this expression. However, one obtains corrected values if one first "unitarises" the formula. Lovelace (1969a) does this by projecting out the partial-waves, $V_\ell(s)$, from eq. (4.1) and then writing for the unitarised partial-wave amplitudes:

$$A_\ell(s) = V_\ell(s)\,[1 + \rho(s)\,V_\ell(s)\,]^{-1} \qquad (4.3)$$

This is a matrix equation if one includes $K\bar{K}$, $n\bar{n}$ states in addition to $\pi\pi$ states. Im ρ is determined from unitarity. Lovelace also gives ρ a left-hand-out determined to give $A_\ell(s)$ the correct threshold behavior and then obtains Re ρ from a dispersion relation. The predictions, and comparison with experiment, are shown in Table 4.1. Perhaps the main problem is that there is no evidence for the ρ' (the spin 1 daughter of the A_2).

This formula also gives scattering lengths and phase-shifts in accordance with the accepted values. In addition it has the correct Adler condition and, with a suitable prescription agrees with off-mass shell extrapolation to the π-pole in production experiments (see Lovelace (1969a) for further details.)

For the meson-baryon case the situation is more

complex. As a particular example we quote the expressions given by Lovelace (1969b) for the $\bar{K}N$ amplitudes:

$$A^0 = \underset{(92)}{102}\,\frac{\Gamma(3/2 - \alpha_s)\,\Gamma(1 - \alpha_t)}{\Gamma(3/2-\alpha_s-\alpha_t)} - 36\,\frac{\Gamma(\frac{1}{2}-\alpha_s)\,\Gamma(1-\alpha_t)}{\Gamma(3/2-\alpha_s-\alpha_t)}$$

$$B^0 = \underset{(215)}{203}\,\frac{\Gamma(\frac{1}{2}-\alpha_s)\,\Gamma(1-\alpha_t)}{\Gamma(3/2-\alpha_s-\alpha_t)} - \underset{(-19)}{29}\,\frac{\Gamma(3/2-\alpha_s)\,\Gamma(2-\alpha_t)}{\Gamma(5/2-\alpha_s-\alpha_t)}$$

$$A^1 = \underset{(-31)}{-27}\,\frac{\Gamma(3/2-\alpha_s)\,\Gamma(1-\alpha_t)}{\Gamma(3/2-\alpha_s-\alpha_t)} + 17\,\frac{\Gamma(3/2-\alpha_s)\,\Gamma(1-\alpha_t)}{\Gamma(5/2-\alpha_s-\alpha_t)}$$

$$B^1 = \underset{(-13)}{-65}\,\frac{\Gamma(3/2-\alpha_s)\,\Gamma(1-\alpha_t)}{\Gamma(5/2-\alpha_s-\alpha_t)} - \underset{(-19)}{11}\,\frac{\Gamma(3/2-\alpha_s)\,\Gamma(2-\alpha_t)}{\Gamma(5/2-\alpha_s-\alpha_t)} \qquad (4.4)$$

with

$$\alpha_t = 0482 + 0.9t$$
$$\alpha_s^0 = -0.62 + 0.9s \quad \text{(for the } I_s = 0 \text{ amplitudes)}$$
$$\alpha_s^1 = -0.22 + 0.9s \quad \text{(for the } I_s = 1 \text{ amplitudes)} \qquad (4.5)$$

These are chosen to remove unwanted parity doublets and to fit the known resonances. The predictive power of the model is poor since few resonance widths are known. This model has many ghosts among its daughters. The numbers in brackets in eq. 4.4 refer to the best values for fitting high energy scattering. It will be seen that the extrapolation from the positive s to the negative s region works quite well in this case. In view of other known failures (eq. the Δ in backward πN scattering) this should perhaps be viewed with caution.

B. The most interesting examples that have been discussed here involve 5-point functions. As an example (Chan et al. (1970)) we consider the $(p\bar{p}K^+K^0\pi^-)$ amplitude. This describes the following experimental processes: $K^+p \to K^0\pi^+p$; $K^-p \to \bar{K}^0\pi^-p$; $\pi^-p \to K^0K^-p$; $\pi^+p \to \bar{K}^0K^+p$; $p\bar{p} \to K^-K^0\pi^+$.

There are three possible orderings of the external lines for which we can have dual expressions (i.e. for which the channels are not exotic). These are shown in fig. 4.1 together with the trajectories used (these were chosen to obtain agreement with experiment). Note that this process is a good one to consider since no P exchange is possible. We ignore spin and then, because of signature it turns out that the three terms must be combined in a fixed ratio so that there is only one remaining parameter (the overall magnitude). Some results are shown in figs. 4.2, 4.3 and 4.4.

The general agreement is good but the model gives incorrect crossing. This is shown in fig. 4.2 where we see that crossing from πp to Kp channels is not successful. The result is even worse if we cross to the $\bar{p}p$ channel but this is probably due to the neglect of spin.

Another recent example is due to Schreiner et al. (1971) who discuss $K^+p \to \Lambda K^+K^-$ and the related processes. Again crossing fails.

C. When we use the Veneziano amplitude at high s to test the form of the Regge residue, it is necessary to smooth out the s-poles on the real axis. The smoothing which one obtains by introducing a reasonable imaginary part into $\alpha(s)$ is not enough - the predicted form has rapid oscillations well beyond where they are seen experimentally. These are due to the daughters and it may be that this indicates that narrow daughters are not in fact present. On the other hand, it is possible that the peaks in $A(s,t)$ are in practice smoothed out by the fact that the degeneracy of levels with different spin is broken - a realistic treatment of unitarity would almost certainly give rise to such a breaking of degeneracy. However, the presence of these oscillations means that one should use the high energy limit of the Veneziano amplitudes, rather than the amplitude itself in fitting the data.

The model predicts nonsense-wrong-signature-zeros in the residue. If absorptive cuts are introduced then it essentially becomes the 'weak-cut' or 'Argonne' model. As such it has difficulties as we have already noted. However the general features of the t dependence seem to be correct. This is seen for example in fig. 4.2 for the

production process. For K^-p elastic scattering fig. 4.5 shows the fit of Lovelace (1969b) (the model here has a fixed pole Pomeron and absorptive cuts so it is not clear how much of the t-dependence is really testing the Veneziano residue).

The correlation between the residue for negative t and positive t is good in some cases (e.g. for $\bar{K}N$ scattering as shown in eq. 4.4), but there is one serious failure - namely the Δ is backward π - N scattering. The discrepancy here is not removed by absorptive cuts.

Gunion and Roberts (1971) give detailed fits to 12 meson-baryon inelastic processes using Veneziano residues and absorptive cuts (calculated according to a rather more sophisticated prescription than usual.)

In conclusion we can say that the asymptotic Veneziano model has all the troubles of the weak-cut Regge theory, but no more! Thus the specifically 'dual' aspects appear satisfactory in so far as they are being tested.

D. In order to look at the daughter structure we require to look in the region of large mass and low spin. The obvious place in BB anihilation. In particular $\bar{p}n \to \pi^-\pi^+\pi^-$ at threshold has been extensively studied. The data is due to Annino et al. (1968). Its principle feature is a very deep hole in the middle of the Dalitz plot corresponding to $\alpha_s \approx \alpha_t \approx 1.5$ (see figures 4.6 and 4.7). We should first note that the existence of the hole or zero is not too surprising and might be expected regardless of duality. To see this we assume that resonance poles occur at equally spaced intervals in s and t (see fig. 4.8). Now in order that the residue of an s-pole, for example, should be a polynomial in t it is necessary that the double poles, where these lines cross, should be removed. Thus lines of zeros must cross each intersection. An obvious pattern for such lines is that indicated in fig. 4.8. (Note that it is not the only possible pattern and there is no reason why the lines should be straight). Odorico (1970) has pointed out that there is remarkable evidence for approximate linearity of the lines of zeros, and that these zeros do seem to show up outside the nest of poles, i.e. away from the $s,t > 0$ quadrant. For example, in $\pi^-p \to \pi^0 n$ the Δ

pole in the s and the ρ pole in the u-channel give rise to a zero at

$$\alpha_{\Delta}(s) + \alpha_{\rho}(u) = 1$$

as t = -o.6. This agrees with the position of the familiar dip. In constrast in $\pi^- p \to \eta n$ the Δ is absent in the s-channel (which is purely $I=\frac{1}{2}$), so this dip is not expected, in agreement with experiment.

It is interesting to compare this explanation of the dips with the Regge pole one in terms of Nwsz of ρ trajectories. First we note that the Regge pole model by itself does not give an immediate explanation of the difference between the $\pi^0 n$ and ηn final state. Secondly, the dip in the former case comes from the Γ factor in the denominator of the (s,t) and (u,t) Veneziano terms (i.e. the ones which contribute to the Regge behavior). Now it is apparently an accident that this dip occurs at exactly the same value of t as the dip in the (s,u) term due to the Odorico mechanism. In particular the coincidence of these dips does not follow from duality in its most general form since the Veneziano model, which is dual, only has this property because of the particular relation between the masses of the N, Δ, and ρ. This result would however follow from "local duality", so it appears that in this case at least the experimental situation is consistent with a stronger form of duality than that implied simply by the existance of a Veneziano representation for the amplitude.

To return to $\bar{p}n \to 3\pi$, Lovelace suggested that, since at threshold only the $0^-(\bar{p}n)$ state can contribute, one should fit this with the $\pi \to 3\pi$, i.e. 4π Veneziano function. In fact this gives too much ρ resonance, so he tried instead of eq. 4.2, the satellite term $\Gamma(1-\alpha_s)\,\Gamma(1-\alpha_t) \,/\, \Gamma(2-\alpha_s-\alpha_t)$. This immediately gives the dip around $\alpha_s=\alpha_t=1.5$, since at the point the Γ function in the denominator is infinite. However, the detailed fit to the data is poor. Following Altarelli and Rubinstein (1969) we can fit the data well with a sum of Veneziano like terms (cf. eq. 3.3). The best fit (L. Nicholas, private communication) is
$C_{10}=1$ (fixed), $C_{11}= 2.9\pm 0.5$, $C_{20}= 2.1\pm 1.0$, $C_{21}= 7\pm 2$, $C_{30}=-3 \pm 1$, $C_{22}= -0.1 \pm 0.4$. This does not agree with Altarelli and Rubinstein's fit - but this is presumably because these authors only fitted projections of the Dalitz

plot. Adding further terms to the series does not significantly improve the agreement.

It is not surprising that several terms are needed, and indeed a fit with one term would have suggested that something serious is wrong with the model. This is because at $J=0^-$, $s=4M_N^2$, we are (on a linear π trajectory) very close to the 0^- particle on the third daughter trajection of the π. Thus, as we saw in section 3 we expect several satellites. Clearly there now arises what ought to be the most exciting test of the Veneziano model, namely, does B_5 applied to $\bar{p}n \to \pi^-\pi^+\pi^-$ give the above values for the C_{mn}. Alas we meet again the spin problem and we do not know which B_5 to use.

Rubinstein, Chaichian and Squires (1969) took the following expression:

$$A = \alpha_{12}{}^{\rho}\, B_5(\alpha_{12}{}^{\rho},\ \alpha_{23}{}^{\rho}-1,\ \alpha_{34}{}^{B}-\tfrac{1}{2},\ \alpha_{45}{}^{\pi},\ \alpha_{15}{}^{B}-3/2)$$

$$+ c(\alpha_{34}{}^{B}-\tfrac{1}{2})\ B_5(\alpha_{12}{}^{\rho}-1,\ \alpha_{23}{}^{\rho}-1,\ \alpha_{34}{}^{B}-\tfrac{1}{2},\ \alpha_{45}{}^{\pi}-1,\ \alpha_{15}{}^{\pi}-\tfrac{1}{2}) \qquad (4.6)$$

where particles 4 and 5 are nucleons and 1,2, and 3 are pions. The first term in eq. 4.6 is leading in all variables. It does not however have a pole at $\alpha_{15}=\frac{1}{2}$, and it is to restore this that the second term is added. For further justification of this term we refer to the original article. By a suitable choice of the free parameter a good agreement with the Altarelli-Rubinstein fit was obtained. However it is not possible to fit the parameters given above (in particular eq. 4.6 given $C_{21}=-C_{22}$) and in fact eq. 4.6 does not give a good fit to the experimental data.

The overall magnitude of the amplitude can be obtained from the $N\bar{N}\pi$ and $\rho\pi\pi$ couplings - see fig. 4.9. It is then possible to evaluate the $\bar{p}n \to 3\pi$ decay rate at threshold and compare with experiment. Agreement is obtained provided the decay width of the third 0^- daughter of the pion (i.e. the intermediate state in $\bar{p}n \to \pi^+\pi^+\pi^-$ at threshold) is not larger than about 100 MeV (see Rubinstein, Chaichian and Squires (1970).

The conclusion at present is that B_5 is not well supported by the data but that the test is unreliable since we do not know how to write down unambiguously the $\bar{p}n \to 3\pi$ dual model.

At higher energy data on this process is also available (Bettini et al 1970), and is given in fig. 4.10. There is the suggestion in this data that the deep holes appear only in alternate squares of the mesh of resonances. This led Oderico (1971) to suggest alternate models in which the pattern of zeros had the form shown in fig. 4.11. However, in order to prevent ancestors from occurring due to the zeros at fixed (s-t) crossing the pole lines, it is necessary to have additional poles at fixed (s+t). These correspond to poles at fixed u whose mass varies with the external masses - a feature which is hardly acceptable.

Attempts to understand fig. 4.10 in terms of dual models for $\bar{p}n \to 3\pi$ are in progress but, due to the fact that there are now many amplitudes, there are many problems.

V. Duality and Inclusive Reactions

1. Generalised "optical theorems"

An inclusive cross-section is one which is summed over a set of unobserved particles. Clearly inclusive processes are natural things to measure experimentally. The simplest example is a total cross-section

$$\sigma^{Tot}_{ab} = \sum_{X} |A_{ab\to X}|^2 \qquad 5.1$$

where the sum is over all sets of stable particles (X). Kinematic factors will be ignored. The next most simple example is where one final particle is observed, a+b→c+X for which we write

$$\sigma^{Tot}_{ab\to c} = \sum_{X} |A_{ab\to cX}|^2 \qquad 5.2$$

Unitarity makes these processes theoretically simple.

Thus for the total cross-section σ_{ab}^{Tot} we have the "optical theorem":

$$\sigma_{ab}^{Tot} = \sum_X |A_{ab\to X}|^2$$

$$= \text{Im } A_{ab\to ab} \qquad \text{(unitarity)}$$

$$= (2i)^{-1}[A_{ab\to ab}(s+) - A_{ab\to ab}(s-)]$$

(hermitian analyticity)

$$=$$

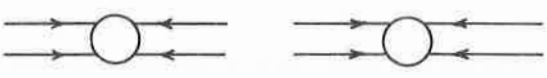
5.3

We ignore the $(2i)^{-1}$ factor in the pictorial form of the equation.

Mueller (1970) showed that this idea could usefully be extended to other inclusive processes. There are in fact several ways of doing this, which we shall discuss. All the relevant expressions were obtained in the formal development of S-matrix theory and are given, for example, in Eden et al. (1966) Chapter IV. First we write

$$\sigma_{ab}^{Tot} = \sum_X$$

$$= \sum_X$$

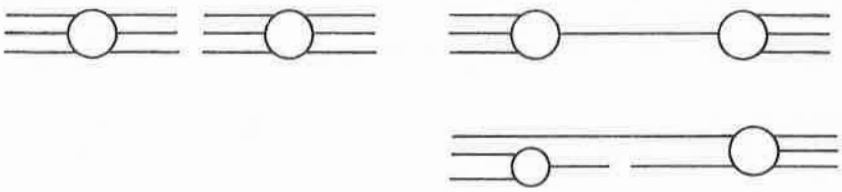


$$= \sum_X$$

5.4

where in the last equation the notation is that $\pm$ refers to the sign of the $(i\epsilon)$ in the left sub-energies, the total energy and the right sub-energies respectively.

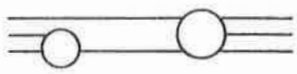
We try using simple physical unitarity:



5.5

Combining this with eq. 5.4 we find

$$\sigma^{Tot}_{ab\to c} = \text{Im } A_{ab\bar{c}\to ab\bar{c}} - \Big\{ \sum_X$$



5.6

This is not useful because of the partially disconnected pieces in the bracket.

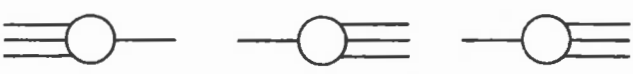
We can remove the disconnected pieces if we take only the discontinuity across the cut in $s_{ab\bar{c}}$ (which is not of course equal to the imaginary part). Using Eden et al. (1966), eq. 5.7.8, we have

5.7

The r.h.s. of this equation is not quite the same as the r.h.s. of eq. 5.4 so one obtains an additional term:

$$\sigma^{Tot}_{ab\to c} =$$

$$+ \sum_X$$



5.8

This equation is given in Ellis et al. (1971); the correction term spoils its usefulness. A much simpler form can however be obtained if we start with 5.7 and continue around the cut in the left sub-energies, thus obtaining

5.9

from which we obtain

$\sigma^{Tot}_{ab\to c}$ =

5.10

Thus expressing the total inclusive cross-section as the discontinuity across the three particle ($s_{ab\bar{c}}$) cut of a particular $3 \to 3$ amplitude in the forward direction. This equation is given in Stapp (1971).

2. Duality and Exotic $ab\bar{c}$

As a first application of eq. 5.10 we consider the case (Chan et al. (1971)) whose $ab\bar{c}$ is an exotic channel. In this case there are no dual diagrams with $ab\bar{c}$ adjacent, i.e. no poles in $s_{ab\bar{c}}$, so, in the meromorphic approximation, the r.h.s. of eq. 5.10 is zero and $\sigma^{Tot}_{ab\to c}$ is zero because there are Pomeranchon-exchange contributions which are not included in dual models. Chan et al. (1971) deduced that if $ab\bar{c}$ is exotic then $\sigma^{Tot}_{ab\to c}$ should be independent of energy (i.e.) of s_{ab}) in the high energy limit; objections to this were raised by Ellis et al. (1971). This question cannot be settled simply on the basis of dual models since these, by their definition, do not include Pomeron contributions. What is needed is a generalisation of eq. 1.1 to $3 \to 3$ reactions. In order to obtain such a plausible generalisation we consider the model in which the Pomeron arises from dual twisted loops.

To explain this we discuss first a + b → c + d where we suppose ab is exotic. Then we cannot draw dual diagrams with a and b adjacent, so dual diagrams must have the form shown in fig. 5.1, where X_1 and X_2 are sets of stable particles. We combine this with a similar diagram for c + d to make the twisted loop shown in fig. 5.2a and fig. 5.2b. The quark-duality diagram corresponding to this is shown in fig. 5.3, and we see from this that only vacuum quantum numbers can be exchanged in the cross-channel, i.e. "a" = "c" and "b" = "d". Thus we can regard this diagram as giving a model for the Pomeron (for calculation of these diagrams we refer to articles by Alessandrini et al. (1970), Lovelace (1970), Alessandrini (1970), etc.; they do not give the correct intercept for the Pomeron but give the reasonable slope of $\frac{1}{2}$).

The corresponding twisted loop for $ab\bar{c}$ elastic scattering is shown in fig. 5.4. [Veneziano (1971), Chan and Hoyer (1971)]. However, there are several possible labellings of this diagram. Suppose 'a' is the central one, then this means that 'a' is coupled to the other particles only through the Pomeron, i.e. we have fig. 5.5 where the $b\bar{c}$ scattering includes Regge exchange. This diagram behaves as $(s_{a\bar{b}})^1$ and so gives a contribution to the inclusive cross-section $\sigma^{Tot}_{ab\to c}$ which is constant with energy (s_{ab}). Note that in the fragmentation region of particle b, with $s_{b\bar{c}}$ negative and fixed $s_{ab\bar{c}}$ is proportional to s_{ab}. This is the origin of the claim made by Chan et al. (1971) that includes cross-sections which are independent of energy if the $ab\bar{c}$ channel is exotic.

However, other orderings of the particles of fig. 5.4 are possible. For example, exchanging a and c we have the diagram of fig. 5.6 which is not constant in s_{ab}. The crucial question now is the magnitude of this term, which is governed by the magnitude of the Pomeron contribution to $b\bar{c} \to b\bar{c}$ at $s_{b\bar{c}}$ small and negative! The physical idea of Pomeron contribution as a manifestation of direct channel unitarity (i.e. giving the shadow of all the inelastic terms) is clearly inappropriate here. In general loop diagrams should be smaller than the 'Born terms' (i.e. the Veneziano amplitudes) so we expect diagrams like this to be small corrections, except in the conditions where the energy dependence of the Pomeron term makes it dominate

over non-Pomeron terms. This is clearly the case in fig. 5.5 (we are considering s_{ab} large) but not in fig. 5.6. Thus the prediction is that the energy dependence of $\sigma^{Tot}_{ab\to c}$ should be small if $ab\bar{c}$ is exotic.

Since this prediction involves both the duality assumption and the nature of the P contribution it will be interesting to see how well it turns out experimentally. There are many consequences (e.g. Chan et al. (1971)) but few good experimental tests. Chan and Hoyer (1971) compare the following processes:

(i) $K^+p \to \pi^+$ [$ab\bar{c}$ exotic, ab also exotic].

(ii) $K^-p \to \pi^-$

(iii) $\pi^+p \to \pi^+$ [$ab\bar{c}$ exotic].

(iv) $\pi^-p \to \pi^-$.

They define R_{K+}, R_{K^-}, $R_{\pi+}$, $R_{\pi-}$, by

$$R_{K+} = \frac{\sigma^{Tot}_{K^+p\to\pi^-}}{\sigma^{Tot}_{K^+p}} \qquad 5.11$$

etc.

The Pomeron contribution to the R's are all equal,

e.g. $$R^P_{K+} = \frac{\gamma_P(p\pi^-)(p\pi^-)}{\gamma_{Ppp}} \qquad 5.12$$

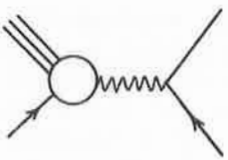
(see fig. 5.7). Experimentally it turns out that $\frac{R_{K+} - R_{K^-}}{R_{K+}}$ $\frac{R_{K+} - R_{\pi-}}{R_{K+}}$ are both of the order of unity, whereas $\frac{R_{K+} - R_{\pi+}}{R_{K+}}$ is consistent with zero. This suggest that R_{K+} and $R_{\pi+}$ are Pomeron dominated in accordance with the fact that in both cases $ab\bar{c}$ is exotic. Note that in one of them (K^+), ab is

also exotic, so the equality of the R's suggests that this particular condition is irrelevant.

As with $2 \to 2$ processes, the factorisation assumption strengthens the predictions of duality. Thus (Ellis et al. (1971) if one accepts that $ab\bar{c}$ exotic implies only a Pomeron contribution and uses factorisation one can show for example that there is never any Regge contribution to the $p \to K^-$ fragmentation. Hence, for example, $\sigma^{Tot}_{p\to K^-}$ is constant in energy even though in this case $ab\bar{c}$ is not exotic.

3. Triple Regge Limits

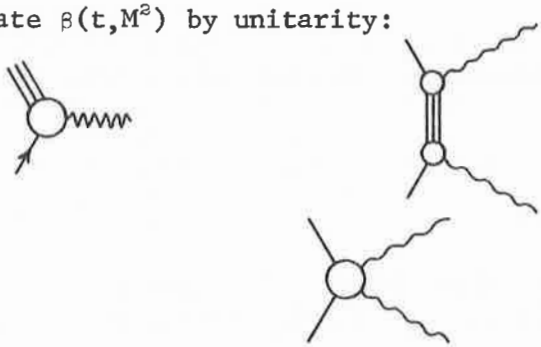
As another example of duality applied to inclusive processes we consider the triple Regge limit (DeTar et al. (1971). Here one is interested in the region $t = s_{b\bar{c}}$ fixed, $M^2 = s_{ab\bar{c}}$ large and (s/M^2), $s = s_{ab}$, also large. Then, for the cross-section, we write



$$= \beta(t,M^2) \left(\frac{s}{M^2}\right)^{2\alpha_1(t)} |\gamma^{(1)}_{b\bar{c}}(t)|^2 \qquad 5.13$$

for large $\left(\frac{s}{M^2}\right)$. Here α_1 is the leading trajectory with the quantum numbers of $b\bar{c}$. In general of course one should sum 5.13 over several such Regge pole contributions.

We evaluate $\beta(t,M^2)$ by unitarity:



5.14

on applying two-particle unitarity to the elastic scattering of particle a and the Regge-pole α_1. The energy here is M^2 so if M^2 is large we can then use Regge theory to write eq. 5.14 as

$$\beta(t,M^2) \simeq (M^2)^{\alpha_2(0)} \quad \gamma_{aa}^{(2)}(0) \; \gamma_{11}^{(2)}(0) \qquad 5.15$$

where α^2 is the leading trajectory in (1,a) elastic scattering - including of course the Pomeron.

Thus we finally obtain

$$\frac{d^2\sigma_{ab\to\bar{c}}}{dt\, dM^2} = s^{-2}\gamma_{aa}^{(2)}(0)\gamma_{11}^{(2)}(0)\,|\gamma_{b\bar{c}}^{(1)}(t)|^2\,(M^2)^{\alpha_2(0)}(s/M^2)^{\alpha_1(t)} \qquad 5.16$$

(see fig. 5.8)

An interesting application of this formula is to study the triple-Pomeron coupling (certain versions of the multiperipheral bootstrap would like this to be zero). Consider a process for which $b\bar{c}$ permits the Pomeron. Then eq. 5.15 is a P elastic scattering which contains resonances dual to Regge-pole terms and also a non-resonant background dual to Pomeron exchange. The latter is proportional to $\gamma_{PP}^{P}(0)$, i.e. the required triple Pomeron coupling. If this is zero then there should be no background when we take the Pomeron for α_1, i.e. the background should tend to zero with s, whereas resonances should remain constant with s (for small t). The expected behaviour for no triple P coupling is illustrated in fig. 5.9 a and b.

With the available data there are conflicting reports in the literature. Wang and Wang (1971) study $p\bar{\pi} \to \pi^- X$ ($b = c = \pi^-$) and $pp \to pX$ ($b = c = p$), and claim that the data is consistent with very little PPP coupling. On the other hand in a more detailed study, using additional data, Edelstein et al. (1971) claim that appreciable PPP coupling is required at least for $|t| \geq 0.264$ (Gev/c)^2.

A case where $b\bar{c}$ does not permit the P is considered by Chliapnikov et al. (1971), namely $pK^+ \to K^0$. They find

that the s-dependence is approximately independent of M^2, i.e. "background" and"resonances" vary in approximately the same way with s. According to duality this means that P α_1 α_1 and R α_1 α_1 couplings are appreciable.

Much further work remains to be done, and is being done, on these topics. The ideas can of course be extended to inclusive processes where two or more are observed. A start on this has been made by Chian-Li et al. (1971).

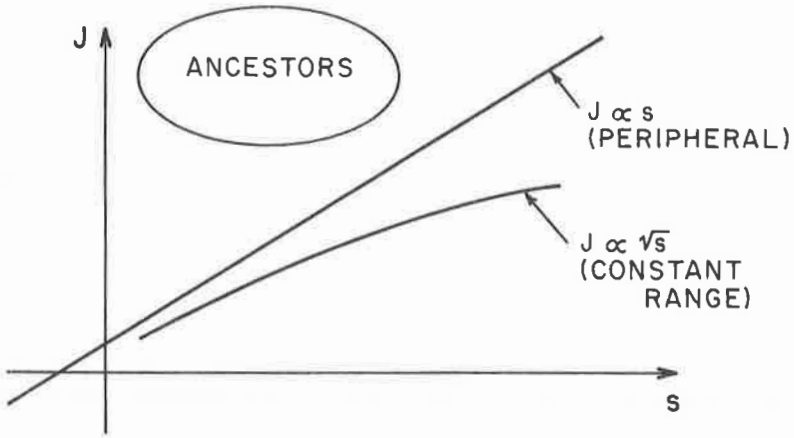


Fig. 2.1

Plot of spin against $(\text{Mass})^2 = s$ for resonances.

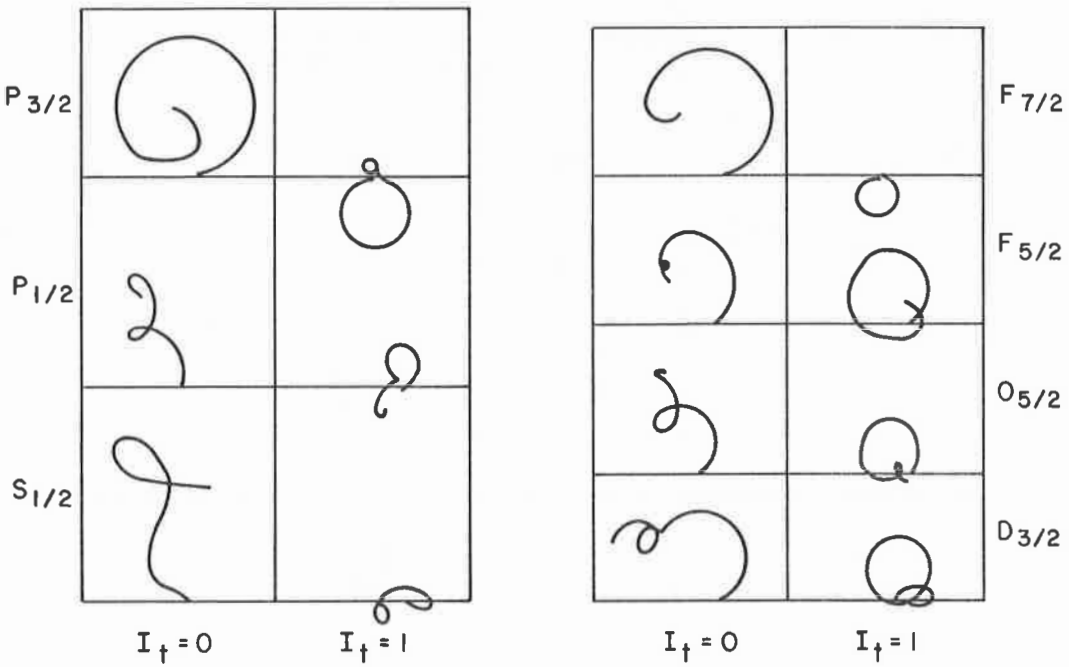


Fig. 2.2

Argand plots of s-channel π-N amplitudes

from Harari and Zarmi.

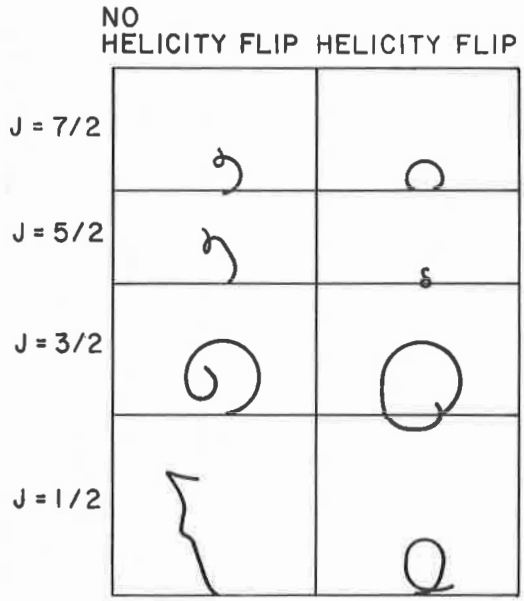


Fig. 2.3

Argand plots of $I_t = 0$ MN amplitudes for helicity non-flip and helicity flip.

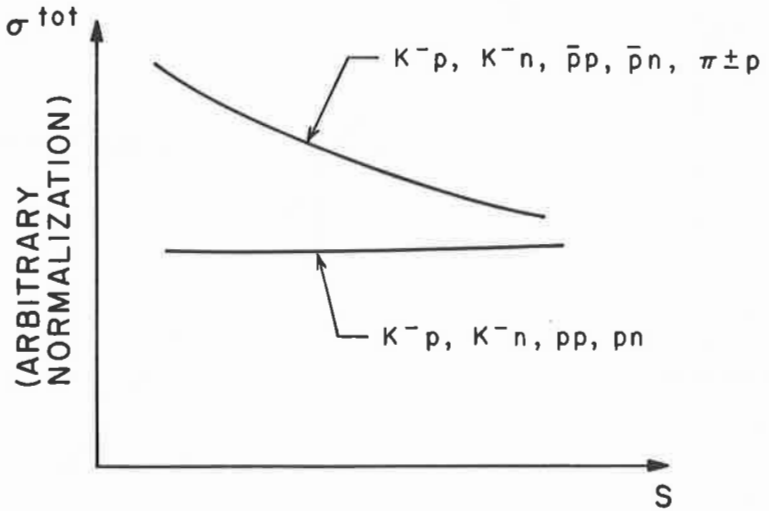


Fig. 2.4

Qualitative behaviour of total cross-sections.

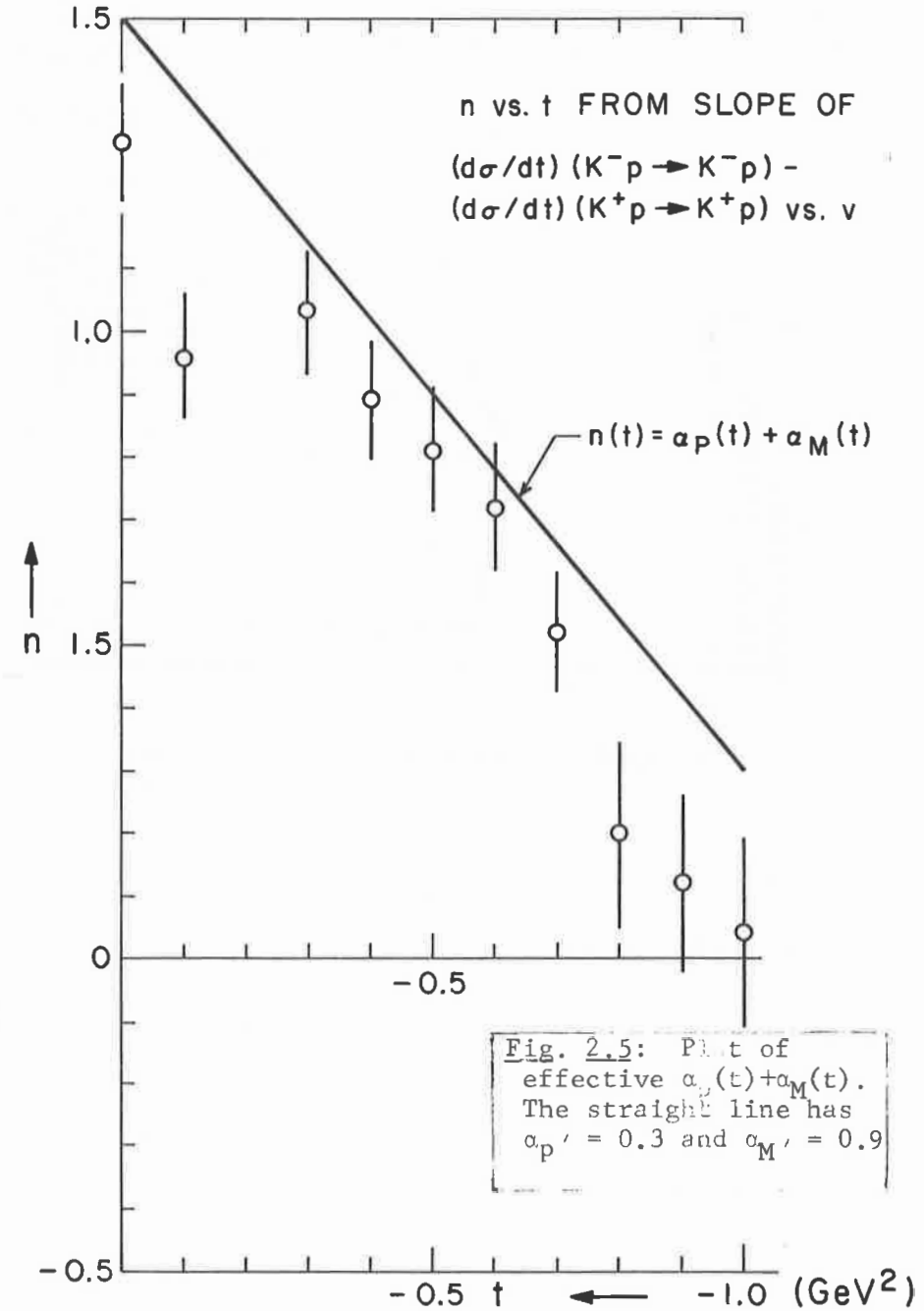
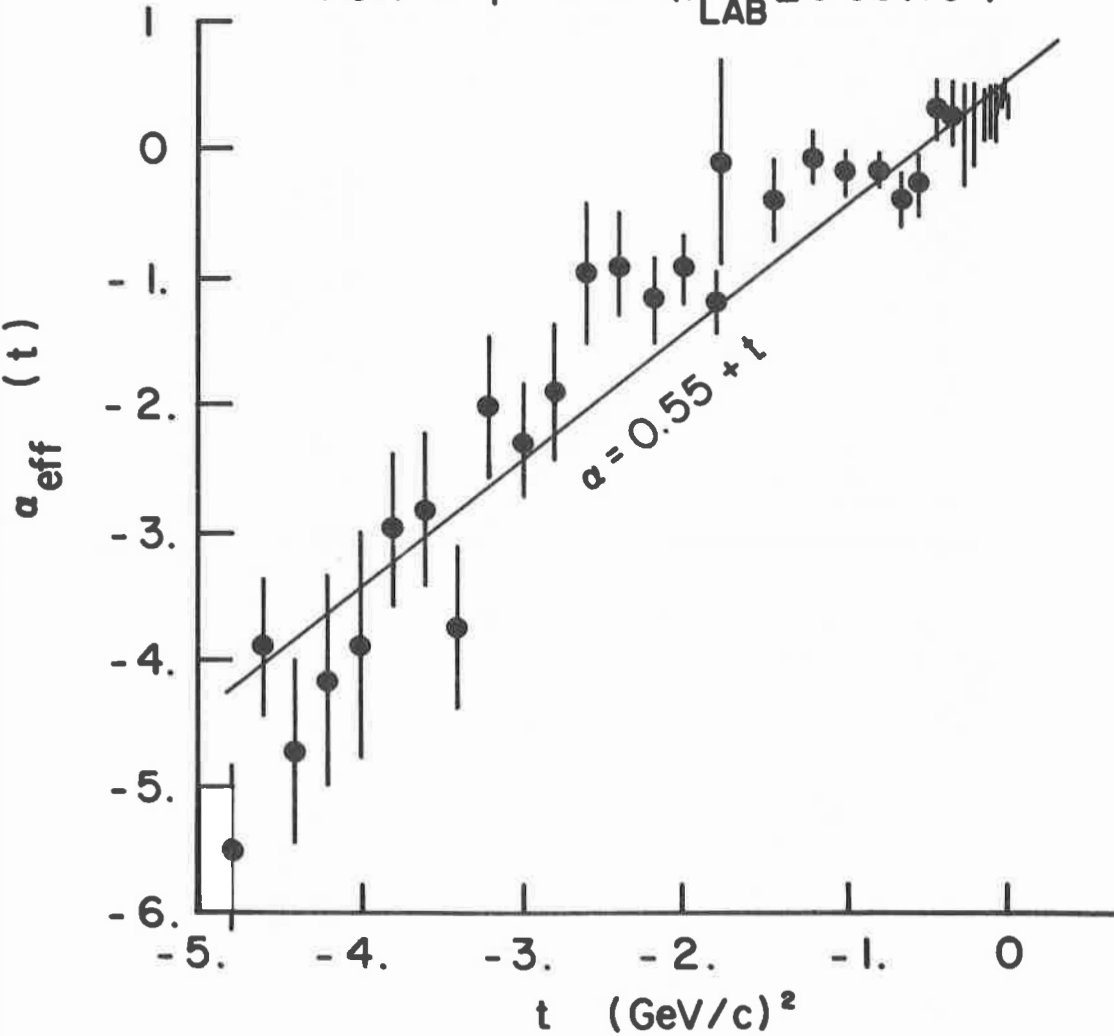


Fig. 2.5: Plot of effective $\alpha_P(t)+\alpha_M(t)$. The straight line has $\alpha_P' = 0.3$ and $\alpha_M' = 0.9$

Fig. 2.6: EFFECTIVE REGGE TRAJECTORY FOR $\pi^- p \to \pi^0 n$ ($P_{LAB} \leq 5$ GeV/c)



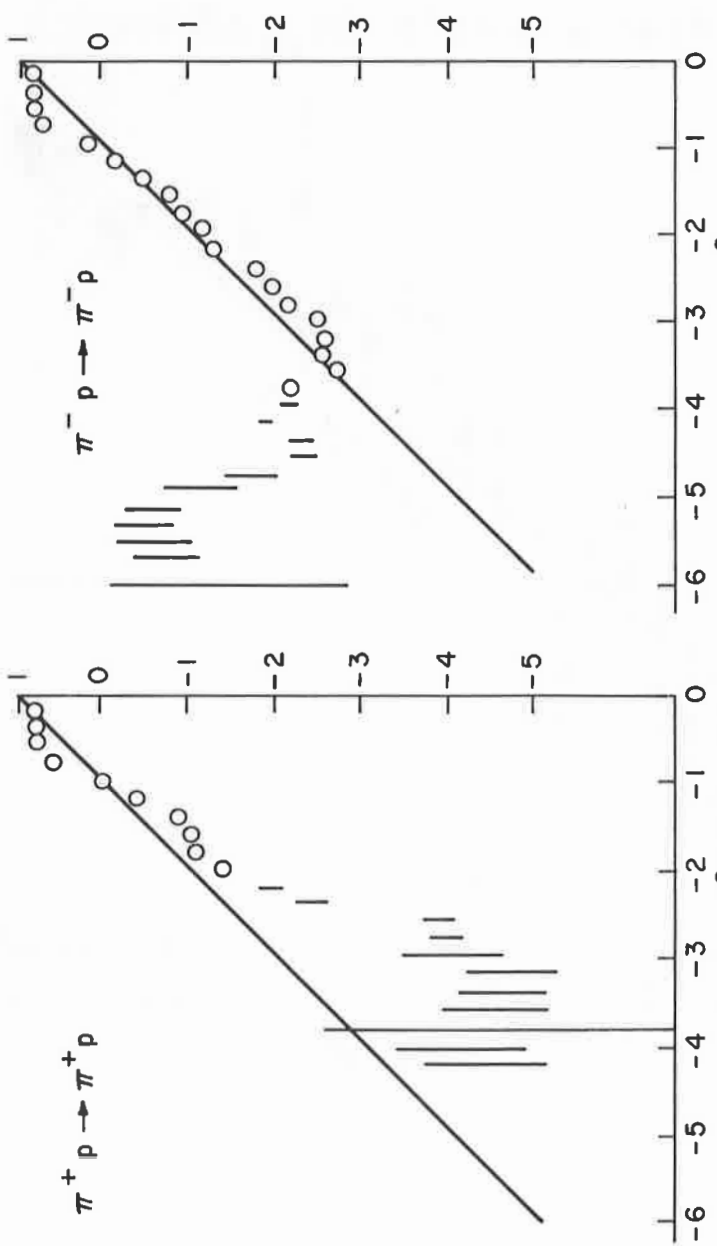


Fig. 2.7

Plots of α^{eff} for $\pi^{\mp}$p elastic scattering.

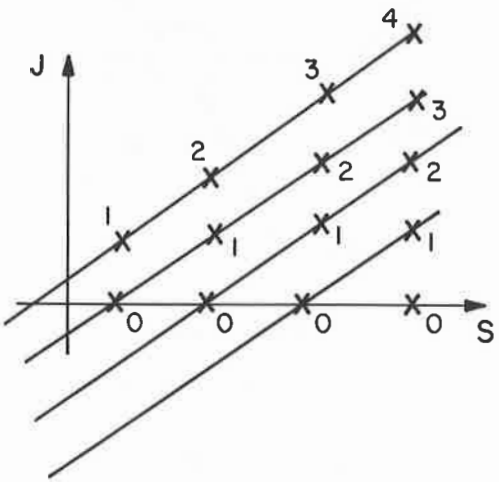


Fig. 3.1

Showing the structure of resonance-poles in the Veneziano model.

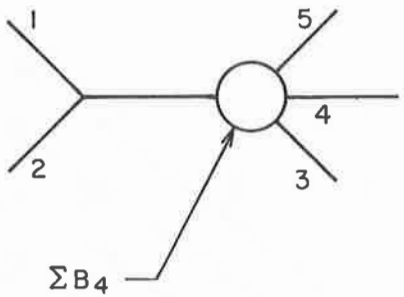


Fig. 3.2

Showing the factorisation property of B_6 at a pole of S_{12}

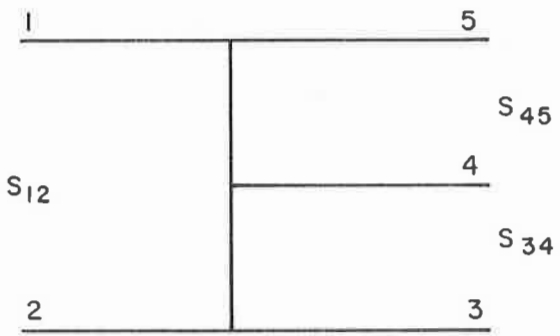


Fig. 3.3

Multi-Regge limit for S_{45}, S_{34} large.

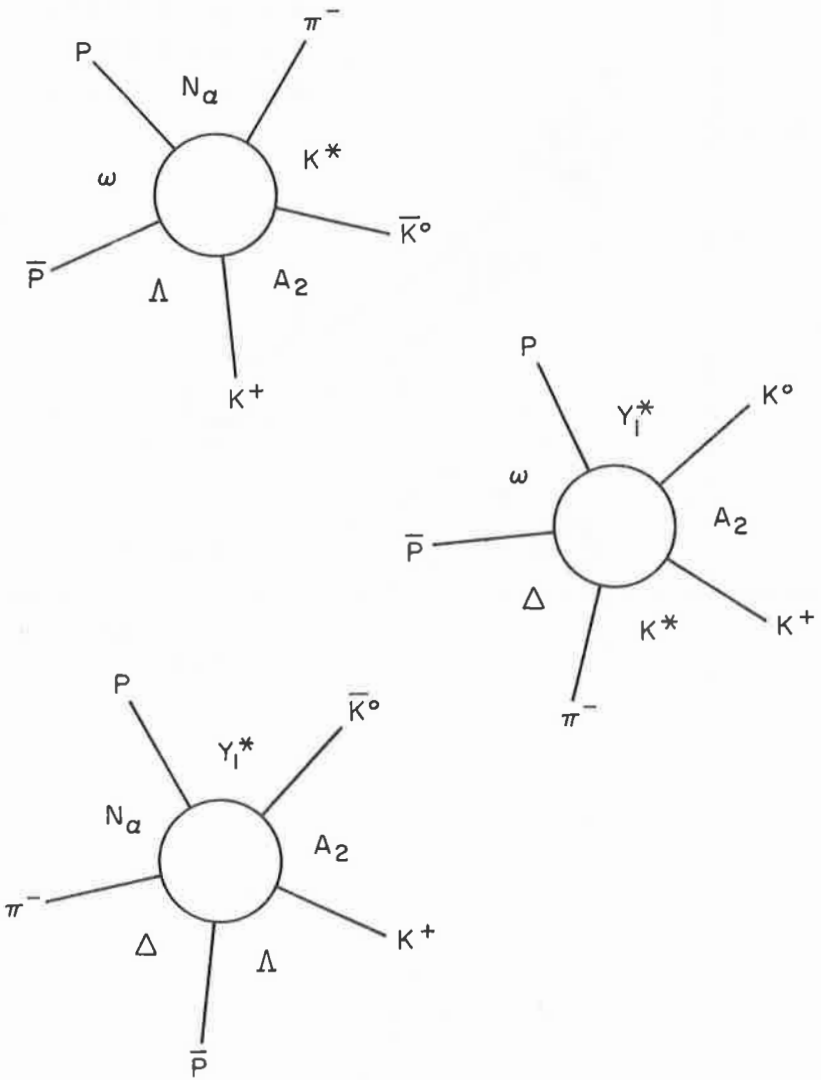


Fig. 4.1

The possible Veneziano diagrams for $(p\bar{p}K^+\bar{K}^\circ\pi^-)$ amplitude.

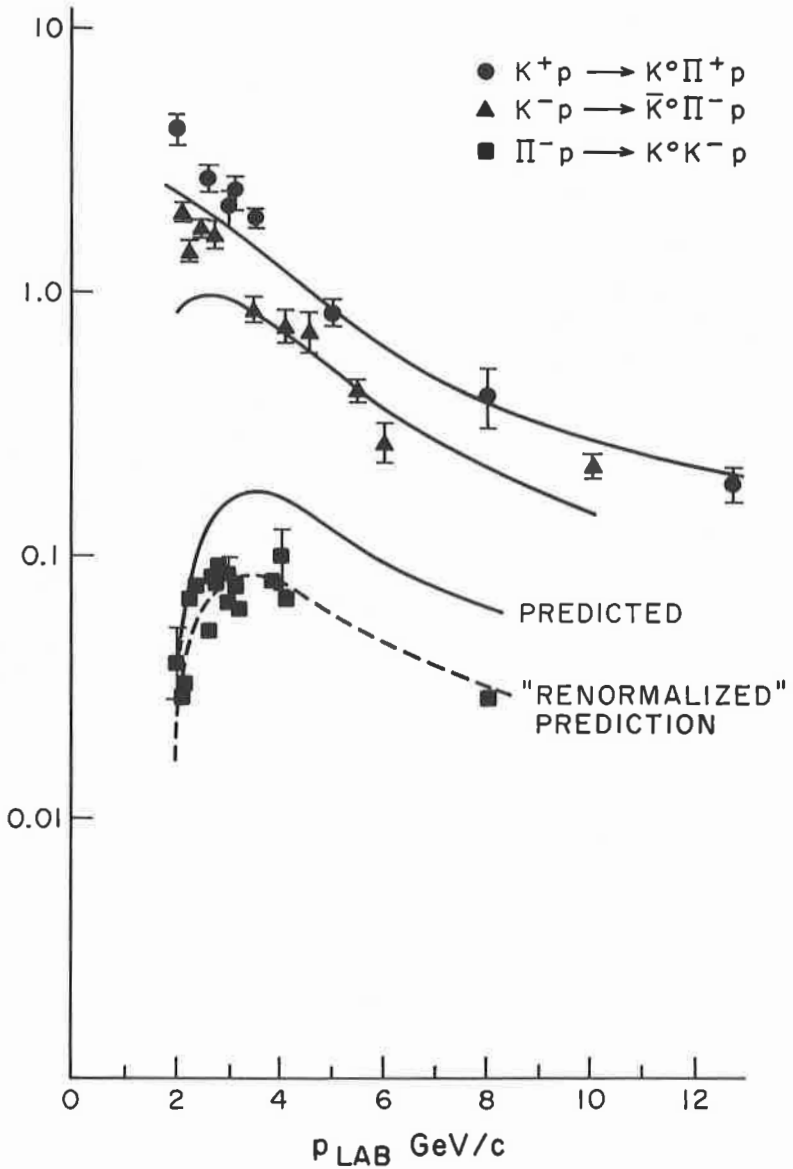


Fig. 4.2

Cross-sections for $K^{\pm}P \to K^0(\bar{K}^0)\pi^{\pm}P$ and $\pi^-P \to K^0K^-P$, from Chan et al.

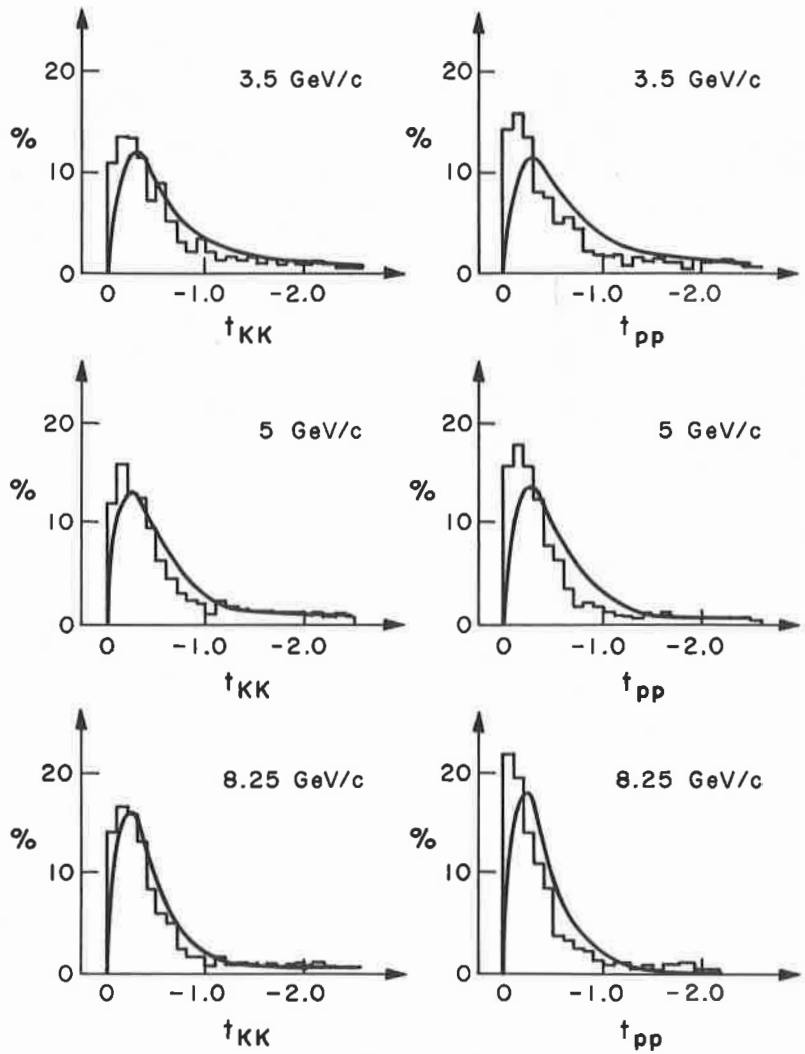


Fig. 4.3

Proton-proton and kaon-kaon momentum transfer distributions in $K^+P \rightarrow K^0\pi^+P$, from Chan et al.

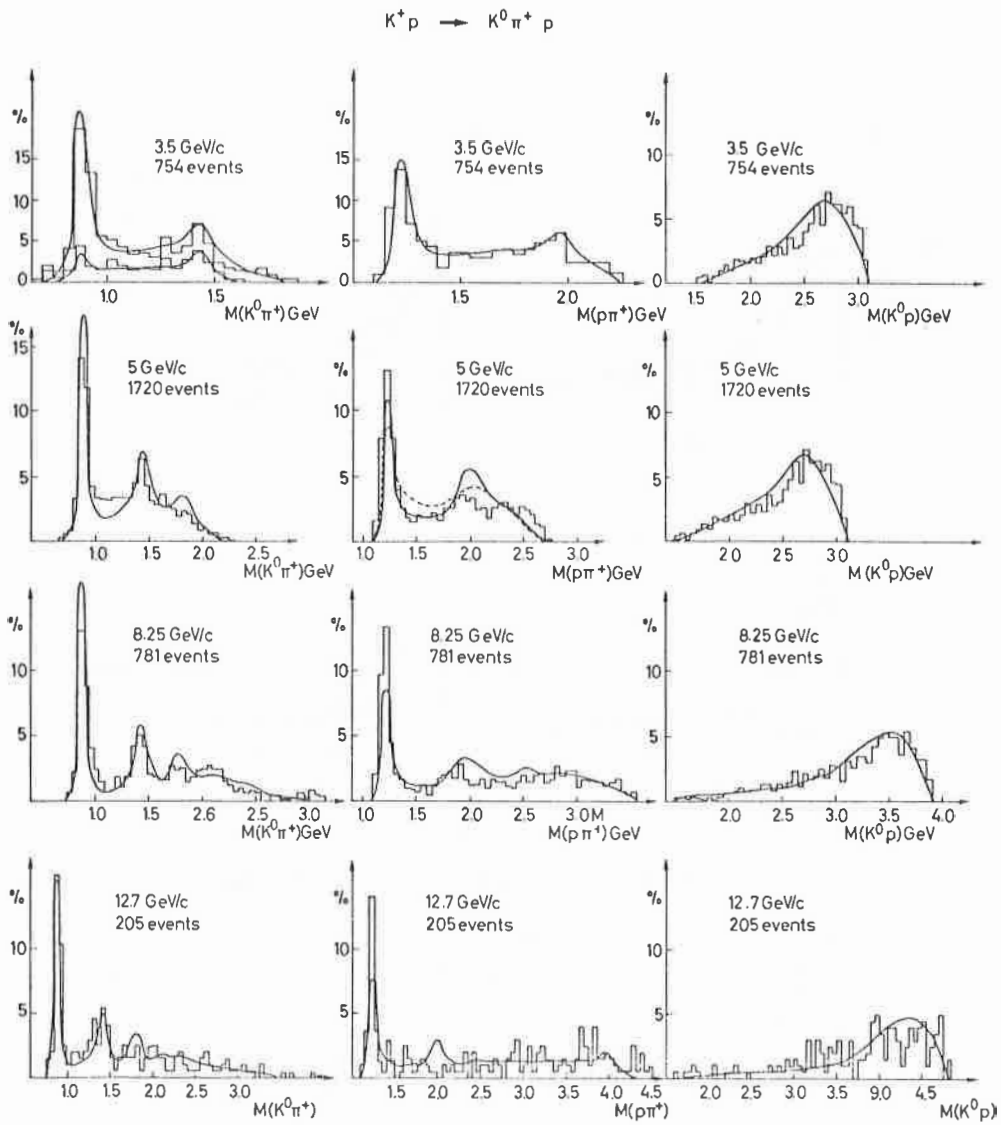


Fig. 4.4

Invariant mass distributions in $K^+P \rightarrow K^0\pi^+P$, from Chan et al.

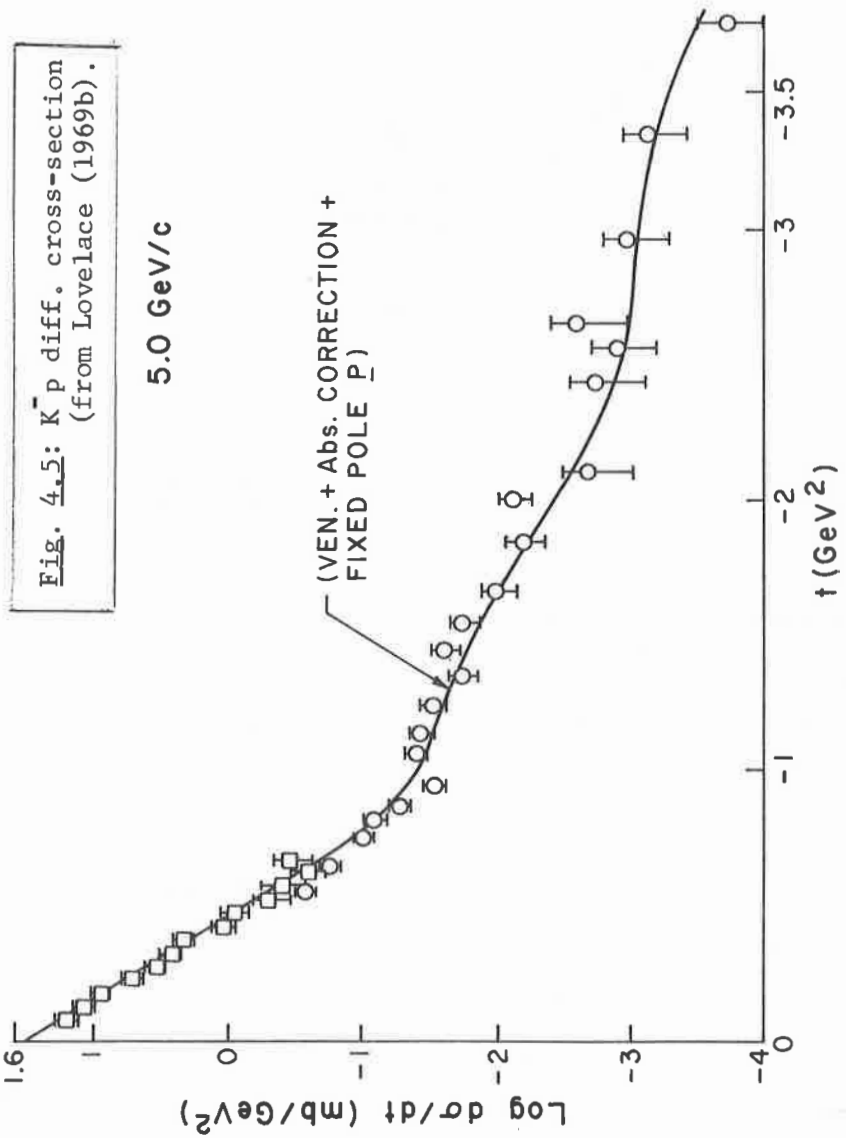


Fig. 4.5: K^-p diff. cross-section (from Lovelace (1969b)).

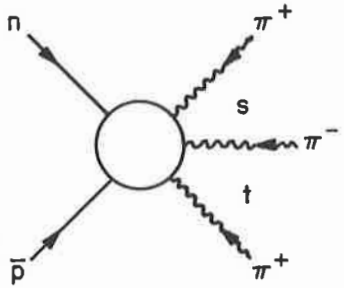


Fig. 4.6

The Veneziano diagram for $\bar{p}n \to \pi^-\pi^+\pi^-$.

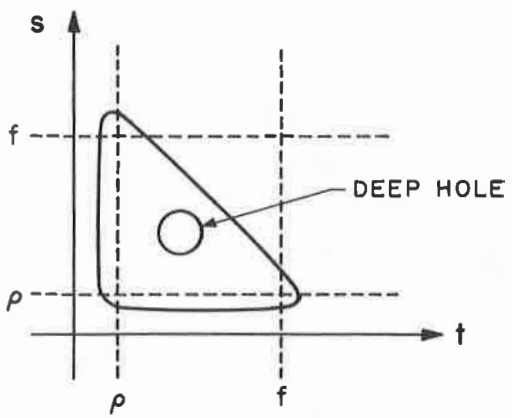


Fig. 4.7

Dalitz plot for $\bar{p}n \to \pi^-\pi^+\pi^-$ at threshold.

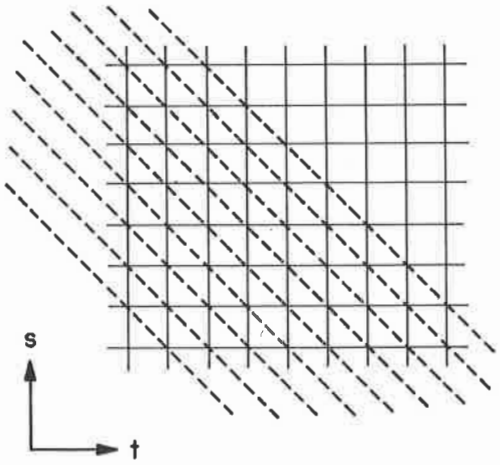


Fig. 4.8

A simple pattern of zeros (dashed lines) which removes ancestors at the poles (full lines).

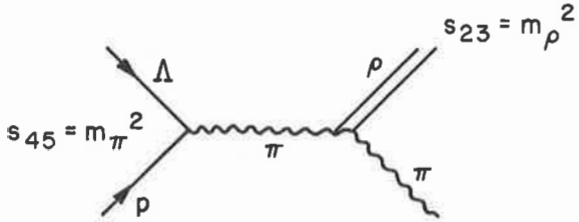


Fig. 4.9

Showing how the normalisation of the $\bar{p}n \to 3\pi$ amplitude is obtained from the NN_π and $\rho\pi\pi$ couplings.

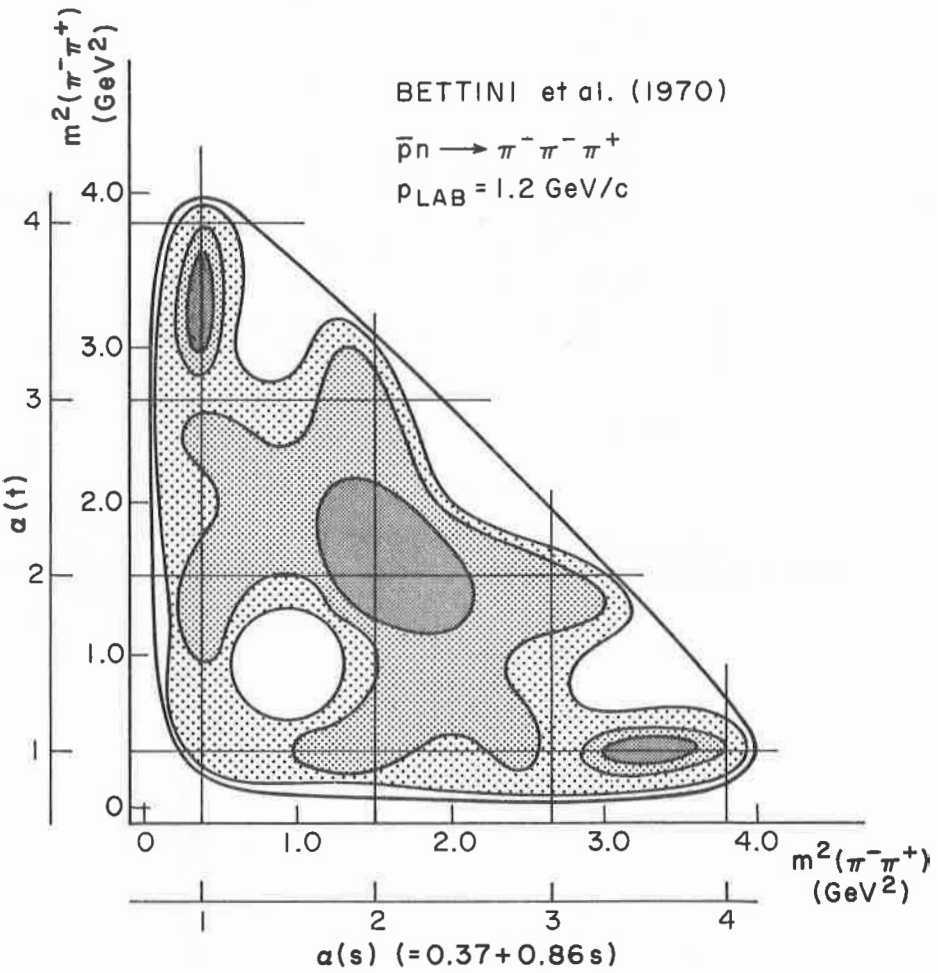


Fig. 4.10

Dalitz plot for $\bar{p}n \to \pi^-\pi^-\pi^+$ at 1.2 GeV/c.

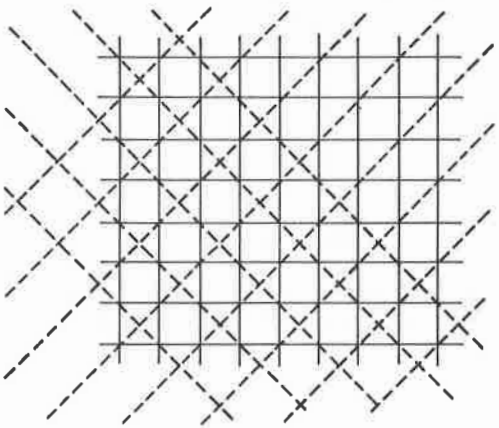


Fig. 4.11

Alternative pattern of zeros suggested by Odorico (1971).

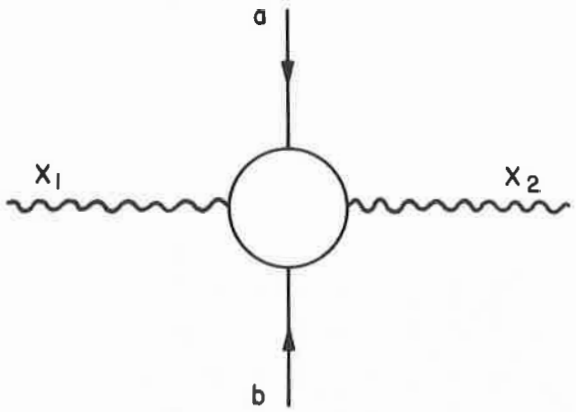


Fig. 5.1

A dual diagram for $ab \to X_1 + X_2$ with a,b non-adjacent.

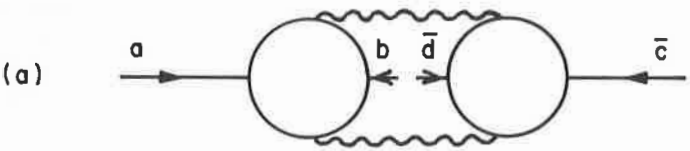
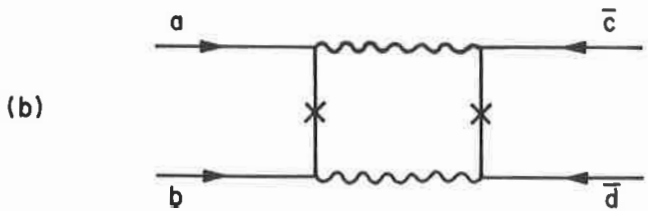


Fig. 5.2: Two ways of depicting the twisted-loop which is suggested as the mechanism which produces the Pomeron contributions to two-body processes.



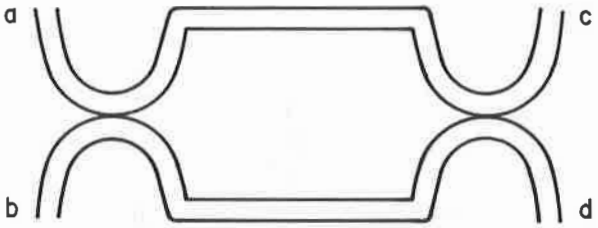
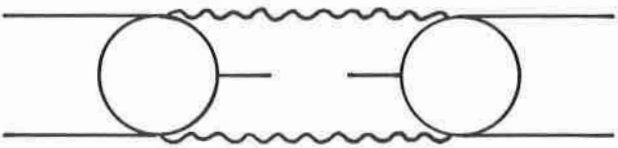


Fig. 5.3

The quark diagram
corresponding to fig. 4.2.



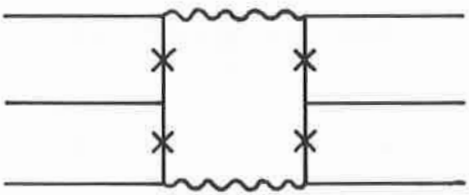


Fig. 5.4

Twisted loops for the 3 → 3 process.

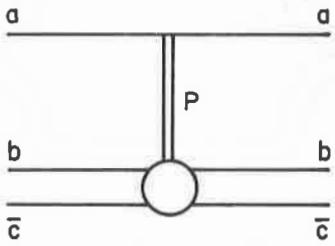


Fig. 5.5

A contribution to $ab\bar{c} \to ab\bar{c}$, which behaves as $(s_{ab})^1$ for large s_{ab}.

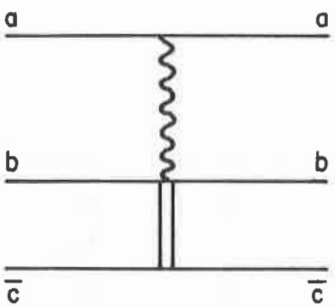


Fig. 5.6

A contribution to $ab\bar{c} \to ab\bar{c}$, which does not behave as $(s_{ab})^1$ but which is probably small in the kinematic limit considered.

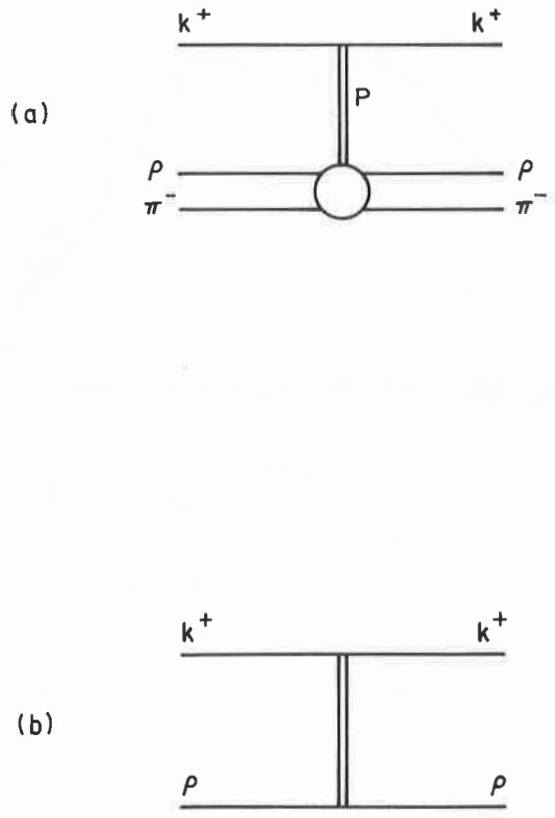


Fig. 5.7

Showing the Pomeron contribution to (a) $K^+p\pi^- \to K^+p\pi^-$ and (b) $K^+p \to K^+p$.

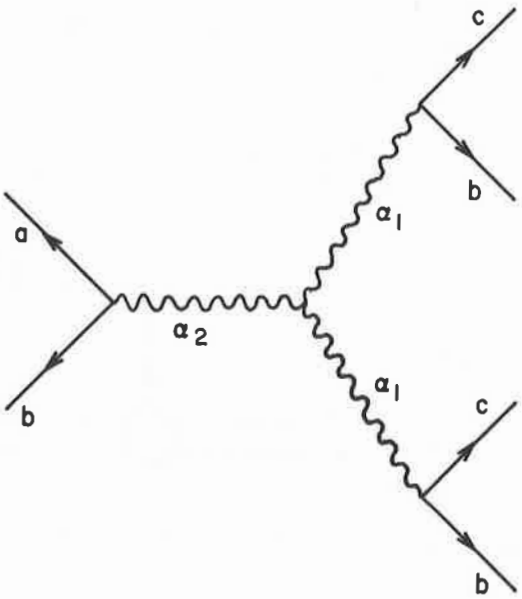


Fig. 5.8: The triple Regge limit.

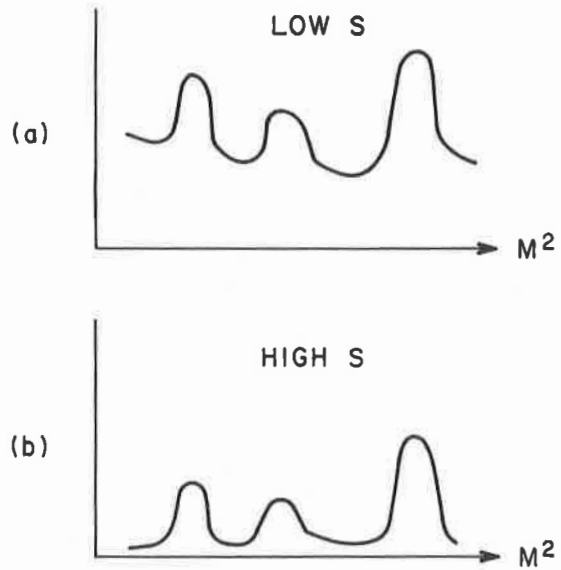


Fig. 5.9

Illustrating the expected behaviour of an inclusive cross-section in the triple Regge limit if the PPP coupling is zero.

References

M. Ademollo, H.R. Rubinstein, G. Veneziano and M.A. Virasoro (1967), Phys. Rev. Letters, 19, 1402.

V. Alessandrini, D. Amati, M. Le Bellac and D. Olive (1970), CERN Th. 1160, to be published in Physics Reports.

V. Alessandrini (1970), CERN th. 1215, to be published in Nuovo Cimento.

G.A. Altarelli and H.R. Rubinstein (1969), Phys. Rev., 183, 1469.

P. Annino, et al. (1968), Phys. Rev. Letters, 20, 402.

V. Barger and R.J.N. Phillips (1971), Wisconsin preprint, "Deductions from new $\pi^- p \to \pi^0 n$ data at large momentum transfers", to be published in Nuc. Phys. B.

A. Bettini, et al. (1970), Padova preprint INFN-PD 70/2.

B.B. Brabson, et al. (1970), Phys. Rev. Letters, 25, 553.

H.M. Chan, C.S.Hsue, C.Quigg and J.M.Wang (1971), Phys. Rev. Letters, 26, 672.

H.M. Chan and P. Hoyer, CERN Th. 1339, to be published in Physics Letters B.

P. Chliapnikov, O. Czyzewski, J. Finkelstein and M. Jacob (1971), CERN Th. 1336, to be published in Physics Letters B.

J.W. Coleman and R.C.Johnson, (1971), Durham preprint, to be published in Nuc. Phys. B.

C. Daum, C. Michael and C.Schmid (1970), Phys. Letters 31B, 222.

C.E. DeTar, C.E. Jones, F.E. Low, J.H. Weiss and J.E.Young, (1971), Phys. Rev. Letters 26, 675.

J.D. Dorren, V. Rittenberg, H.R. Rubinstein, M. Chaichian and E.J.Squires (1970), Nuovo Cimento 1A, 149.

J.D. Dorren, V. Rittenberg and H.R. Rubinstein (1970), CERN Th. 1191.

R.M. Edelstein, V. Rittenberg and H.R. Rubinstein (1971), Weizmann preprint (WIS-71/19/Ph).

R.J. Eden, P.V. Landshoff, D.I. Olive and J.C. Polkinghorne (1966), "The Analytic S-Matrix" (Cambridge).

J. Ellis, J. Finkelstein, P.H. Frampton and M. Jacob (1971), CERN Th. 1316.

A. Firestone et al. (1971), Phys. Rev. Letters 25, 958.

J.F. Gunson and R.G. Roberts (1971), Nuc. Phys. B28, 210.

H. Harari (1971), Phys. Rev. Letters 26, 1079.

H. Harari and Y. Zarmi (1969), Phys. Rev. 187, 2230.

H. Harari and Y. Zarmi (1970), Phys. Letters 32B, 291.

C.L. Jen et al. (1971), preprint NYO-226 2TA.

C. Lovelace (1970), Physics Letters 32B, 703.

C. Lovelace (1969a), CERN Th. 1041 (Argonne conference report).

C. Lovelace (1969b), CERN Th. 1047.

R. Odorico (1970), Nuovo Cimento Letters 3, 61.

R. Odorico (1970), Physics Letters 33B, 489.

V. Rittenberg and H.R. Rubinstein (1970), Phys. Rev. Letters 25, 191.

H.R. Rubinstein, M.Chaichian and E.J.Squires (1969), Physics Letters 30B, 189.

H.R. Rubinstein, M. Chaichian and E.J.Squires (1970), Nuc. Phys. B21, 283.

C. Schmid (1970), Nuovo Cimento Letters 1, 165.

C. Schmid (1971), CERN preprint Th. 1343 "Phenomenology at Intermediate Energies".

P.A. Schreiner et al. (1971), Nuc. Phys. B28, 85.

E.J. Squires (1971), "Regge-Pole Phenomenology" Lectures at the 1970 Heidelberg Summer School, published in Springer Tracts in Nuclear Physics Vol. 57.

H. Stapp (1971), UCRL 20623.

G. Tiktopolous (1970), Phys. Letters 31B, 138.

G. Veneziano (1968), Nuovo Cimento 57A, 190.

G. Veneziano (1971), Nuovo Cimento Letters 1, 681.

R.E. Waltz (1970), Nuc. Phys. B18, 61.

J.M. Wang and L.L. Wang (1971), Phys. Rev. Letters 26, 1287.

GROUP THEORETICAL PROPERTIES

OF DUAL RESONANCE MODELS

P. Ramond
National Accelerator Laboratory
Batavia, Illinois

I. INTRODUCTION

One of the remarkable things about duality is that it leads to the formulation of very esthetic theoretical ideas although it has its roots in the structure of strong-interaction data. Surely this marriage of conceptual beauty with experimental observation is no accident. The first steps towards the construction of amplitudes that were "dual" have been excellently described in several review articles[1]; in these lectures we would rather like to show the emergence of a very fundamental group theoretical structure that seems to underlie all dual resonance models (DRM) built to date. Since no DRM duplicates the data closely enough, we would like to understand how to add the missing ingredients without affecting the properties we like about the more primitive models (like factorization, crossing, Regge behavior, etc.).

As we are only at the beginning of our understanding of duality, we can only talk at the moment about mesons and ask the more pragmatic reader to bear with us while we try to unravel this very mysterious concept.

The other purpose of these notes is to familiarize the reader with the mathematical techniques used in deriving DRM's. Hence the character of what follows will be rather technical as it must be at this stage of the art.

II. MATHEMATICAL PRELIMINARIES

The work of Koba and Nielsen[2] has shown the relevance of projective transformations in dual resonance models

(DRM). These transformations are generated in the complex plane by real Mobius transformations which are locally isomorphic to the more familiar SU(1,1) group, the non-compact partner of SU(2). We concentrate from now on in the study of SU(1,1) so as to understand its role in DRMs in greater detail.[3]

If h is an element of SU(1,1) it is in one-to-one correspondence with the pseudounitary unimodular 2x2 matrix

$$h \to \begin{pmatrix} \alpha & \beta \\ \beta^* & \alpha^* \end{pmatrix} \qquad |\alpha|^2 - |\beta|^2 = 1, \tag{1}$$

where α and β are complex numbers and the star denotes complex conjugation. Its Lie algebra is generated by the operators L_0, L_+, and L_- which obey

$$\left[L_0, L_\pm\right] = \pm L_\pm; \qquad \left[L_+, L_-\right] = -L_0 \tag{2}$$

and has a Casimir operator

$$L^2 = L_+L_- + L_-L_+ - L_0^2 . \tag{3}$$

In the complex z-plane, h corresponds to

$$z \to (hz) \equiv z' = \frac{\alpha z+\beta}{\alpha^*+\beta^* z} . \tag{4}$$

In particular it maps any point on the unit circle onto the unit circle. In order to construct the representations of the SU(1,1) algebra, we choose a certain representation for the generators

$$L_0 = z\frac{d}{dz}; \qquad L_\pm = \frac{1}{\sqrt{2}} z^{\pm 1}\left(z\frac{d}{dz} - J\right) \tag{5}$$

for which

$$L^2 = -J(J+1) \tag{6}$$

is automatically a c-number. It can be shown that there exists basically two types of unitary representations of the

algebra: those for which the spectrum of L_0 is unbounded and those for which it is bounded. For reasons that will become clear later we concentrate on the latter ones. There we again have two subdivisions since L_0 can be bounded either from above or below. These unitary irreducible representations (UIR) of the algebra are

1. $D_J^{(+)}$, where J is a real negative number, as required by unitarity, and the spectrum of eigenvalues of L_0 is bounded below

$$L_0 = -J, -J+1, -J+2, \ldots,$$

and it is spanned by the states

$$|J,m)_+ = \left(\frac{(m-1-2J)!}{m!}\right)^{\frac{1}{2}} z^{m-J} \tag{7}$$

Note that

$$L_- |J,0)_+ = 0, \tag{8}$$

and the states are generated by successive application of L_+ on $|J,0)_+$.

2. $D_J^{(-)}$, where again J is real and negative and the spectrum of L_0 is bounded above

$$L_0 = J, J-1, J-2, \ldots .$$

It is spanned by the basis

$$|J,m)_- = \left(\frac{(m-1-2J)!}{m!}\right)^{\frac{1}{2}} z^{-m+J} \tag{9}$$

and

$$L_+ |J,0)_- = 0 \tag{10}$$

so that the states are generated by the successive application of L_- on $|J,0)_-$.

The connection of these representations to the DRM's is achieved in the following way. Introduce the operator functions[4]

$$F_\rho(z) = \sum_{m=0}^{\infty} a_\rho^{(m)} |J,m)_+ \qquad \epsilon D_J^{(+)} \tag{11a}$$

$$\tilde{F}_\rho(z) = \sum_{m=0}^{\infty} a_\rho^{(m)\dagger} |J,m)_- \qquad \epsilon D_J^{(-)} \tag{11b}$$

where the coefficients of the basis vectors are harmonic oscillator operators[5] obeying

$$\left[a_\rho^{(m)}, a_\sigma^{(n)}\right] = \left[a_\rho^{(m)\dagger}, a_\sigma^{(n)\dagger}\right] = 0$$

$$\left[a_\rho^{(m)}, a_\sigma^{(n)\dagger}\right] = g_{\rho\sigma}\delta_{n,m} \qquad n,m = 0,1,\ldots \tag{12a}$$

we use the metric $g_{oo} = -g_{ii} = -1$ so that we immediately see that the $a_0^{(n)\dagger}$ will introduce negative norm states in the theory. (This disease plagues all relativistic theories.) The vacuum state $|0>$ is defined by

$$a_\mu^{(n)} |0> = 0 \qquad n=0,1,\ldots \tag{12b}$$

Furthermore consider the case $n = 0$ to be describing a translational mode,[6] that is, let $J = -\epsilon/2$ where ϵ is a positive infinitesimal. Then, when written in terms of the canonical coordinates,

$$q_\rho = \frac{1}{\sqrt{\epsilon}}\left[a_\rho^{(0)\dagger} + a_\rho^{(0)}\right] \tag{13a}$$

$$p_\rho = \frac{i\sqrt{\epsilon}}{2}\left[a_\rho^{(0)\dagger} - a_\rho^{(0)}\right] \tag{13b}$$

$F_\rho(z)$ and $\tilde{F}_\rho(z)$ are separately singular as $\epsilon \to 0$; however, this singularity is absorbed by taking their sum

$$Q_\rho(z) = F_\rho(z) + \tilde{F}_\rho(z) \tag{14}$$

$$= q_\rho + ip_\rho \ln z + \sum_{n=1}^{\infty} \frac{1}{\sqrt{n}} \left[a_\rho{}^{(n)\dagger} z^{-n} + a_\rho{}^{(n)} z^{n} \right], \tag{15}$$

which can be loosely interpreted as the dual generalization of a coordinate. Another quantity of interest is the "generalized momentum"

$$P_\rho(z) = iz \frac{d}{dz} Q_\rho(z) \tag{16}$$

$$= p_\rho + i \sum_{n=1}^{\infty} \sqrt{n} \left[a_\rho{}^{(n)\dagger} z^{-n} - a_\rho{}^{(n)} z^{+n} \right]. \tag{17}$$

The relevant representation of the SU(1,1) operators is now obtained by taking the matrix elements of the operators Eq. (5) between the states Eq. (11a) or equivalently Eq. (11b):

$$L_0 = (F|L_0|F) = \sum_{m=0}^{\infty} (m + \tfrac{\epsilon}{2})\, a^{(m)\dagger} \cdot a^{(m)} \tag{18a}$$

$$L_+ = (F|L_+|F) = \sum_{m=0}^{\infty} \left(\tfrac{1}{2}(m+\epsilon)(m+1) \right)^{\frac{1}{2}} a^{(m+1)\dagger} \cdot a^{(m)} \tag{18b}$$

$$L_- = (F|L_-|F) = \sum \left(\tfrac{1}{2}(m+\epsilon)(m+1) \right)^{\frac{1}{2}} a^{(m)\dagger} \cdot a^{(m+1)}. \tag{18c}$$

Another more elegant way of obtaining the representation of the SU(1,1) operators is to consider the Fourier coefficients of the square of the "generalized momentum"[7]

$$L_{-m} = \frac{1}{2\pi} \int_{-\pi}^{+\pi} \frac{dz}{z} z^{-m} : P_\mu(z) P^\mu(z) : \tag{19}$$

where z is on the unit circle and the normal ordering applies to the periodic modes only. Specializing expression (19) to m = 0, ±1, we obtain the usual representation of the SU(1,1) generators, namely

$$L_0 = 2\left\{\frac{1}{2}p^2 + \sum_{m=1}^{\infty} m\, a^{(m)\dagger}\cdot a^{(m)}\right\} \tag{20a}$$

$$L_{+1} = 2\left\{ip\cdot a^{(1)\dagger} + \sum_{m=1}^{\infty} \sqrt{m(m+1)}\; a^{(m+1)\dagger}\cdot a^{(m)}\right\} \tag{20b}$$

$$L_{-1} = 2\left\{-ip\cdot a^{(1)} + \sum_{m=1}^{\infty} \sqrt{m(m+1)}\; a^{(m)}\cdot a^{(m+1)\dagger}\right\} \tag{20c}$$

Here, unlike the previous representation we have already taken ϵ to zero. However, for calculational purposes we prefer to use Eqs. (18) and let $\epsilon \to 0$ only at the end of all calculations.[8]

For general integer m, Eq. (19) yields

$$L_m = 2\left\{i\sqrt{m}\; p\cdot a^{(m)\dagger} + \sum_{n=1}^{\infty} \sqrt{n(m+n)}\, a^{(n+m)\dagger}\cdot a^{(n)} - \frac{1}{2}\sum_{n=1}^{m-1} \sqrt{n(m-n)}\, a^{(m-n)\dagger} \cdot a^{(n)\dagger}\right\} \tag{21}$$

These operators were first found by Virasoro[9] in conjunction with the ghost compensation mechanism that occurs in the DRM's. They form among themselves the so-called Virasoro algebra[10]

$$\left[L_m, L_n\right] = 2(n-m)L_{n+m} + \frac{4}{3}n(n^2-1)\delta_{n,-m}\,. \tag{22}$$

The generators $1/2n\; L_{\pm n}$ and $1/n\; L_0 + (n^2-1)/3n$ form, for a given n, an SU(1,1) algebra and generate finite transformations of the form

$$z \xrightarrow[n]{} z' = \left[\frac{\alpha z^n + \beta}{\alpha^* + \beta^* z^n}\right]^{1/n} \qquad n = 1,2,\ldots \tag{23}$$

At the present moment, however, the relevance of this algebra to duality has not been clarified although it is suspected to be very deep., All we can say is that it acts as a gauge group for dual models. More will be said on this in the course of these lectures

The major part of the mathematical equipment needed in dual "modelry" has now been presented, and we turn our attention to the problem of the construction of dual factorizable tree amplitudes.

III. GROUP THEORETICAL RULES FOR THE CONSTRUCTION OF DUAL AMPLITUDES

We wish to emphasize that the rules we will enunciate in this section[11] are not the product of very deep insight but rather of a detailed analysis of the N-point generalization of the Veneziano amplitude. In addition, we believe them to be necessary but not sufficient.

1. Associate with the absorption of a particle of momentum k_μ, with various quantum numbers collectively labelled by $\{\lambda\}$, a vertex operator $V(k_\mu,\{\lambda\};z)$, where $z = e^{-i\tau}$.

2. In order to preserve the correct selection rules at each vertex, we require that V transforms under the groups which generate $\{\lambda\}$ as the field of the absorbed particle.

3. At this stage, the dynamical assumption of duality is expressed in terms of an additional transformation requirement. Namely we demand that

$$[L_0,\ V(k,\{\lambda\};z)] = -z\,\frac{d}{dz}\,V(k,\{\lambda\};z) \tag{24}$$

$$[L_\pm,\ V(k,\{\lambda\};z)] = -\frac{z^{\pm 1}}{\sqrt{2}}\left(z\frac{d}{dz} \mp J_s\right)V(k,\{\lambda\};z), \tag{25}$$

where J_s in this case is a scalar function depending on the various quantum numbers of the particle

$$J_s = J_s(m^2, j, c^{\{\lambda\}}); \tag{26}$$

here j is the spin and $c^{\{\lambda\}}$ represents the Casimir operators of the groups which generate $\{\lambda\}$. This means that the additional feature of dual vertices is that they are labelled by the Casimir operator of SU(1,1).

If T is a finite unitary transformation of SU(1,1), it follows that

$$T\, V(k,\{\lambda\};z)T^{\dagger} = |\alpha^{*}+\beta^{*}z|^{2J_s} V(k,\{\lambda\};z') \tag{27}$$

with

$$z' = \frac{\alpha z+\beta}{\alpha^{*}+\beta^{*}z} . \tag{28}$$

4. An arbitrary number of particles can interact in a dual manner only if their dual vertices have the same SU(1,1) spin, i.e.,

$$J_s(m_1^2, j_{(1)},\ldots) = J_s(m_2^2, j_{(2)},\ldots), \tag{29}$$

which implies, as we shall see later, relations between the various quantum numbers of the particles. The origin of this requirement becomes clear when one tries to build amplitudes out of these dual vertices.

5. The factorizable dual amplitude for the scattering of an arbitrary number of particles in a given order is just given in the tree approximation by the vacuum expectation value of the product of their dual vertices taken so as to make an SU(1,1) invariant.[12] The amplitude corresponding to Fig. 1 is then given by

$$A_N(k_1,\ldots,k_N) = \int\ldots\int dx_1\ldots dz_N K_N(z_1,\ldots,z_N)\,\delta^{(4)}\left(\sum_1^N k_1\right)$$

$$\langle 0|V(k_1,\{\lambda\}_1;z_1)V(k_2,\{\lambda\}_2;z_2)\ldots V(k_N,\{\lambda\}_N;z_N)|0\rangle . \tag{30}$$

The requirement that A_N be SU(1,1) invariant imposes severe restrictions on the kernel function $K_N(\{\lambda\})$. In fact, given the transformation properties (27) of the dual vertices, we have been able to find such a kernel only when all the external particles had the same SU(1,1) spin J_s, which explains the previous requirement. We now show how to build K_N up to any SU(1,1) invariant function.

It is easy to see by using the projective invariance of the vacuum and inserting $T^{\dagger}T$ between the vacuum and $V(k_1,\{\lambda\}_1,z_1)$ and pushing T to the right by means of Eq. (27) that any such kernel must obey

$$dz_1\ldots dz_N K_N(\{\lambda\}) \prod_{i=1}^{N} |\alpha^*+\beta^* z_i|^{2J_s^{(i)}} =$$

$$= dz_1'\ldots dz_N' K_N(\{z'\}), \tag{31}$$

where

$$z_i' = \frac{\alpha z_i+\beta}{\alpha^*+\beta^* z_i} \qquad i=1,2,\ldots,N. \tag{32}$$

From the last equation, it is straightforward to see that

$$\frac{dz_i'}{z_i'} = \frac{dz_i}{z_i} \quad \frac{1}{|\alpha^*+\beta^* z_i|^2}, \tag{33}$$

as well as

$$\frac{z'_{i+1}-z'_i}{z_{i+1}-z_i} = \frac{1}{(\alpha^*+\beta^* z_{i+1})(\alpha^*+\beta^* z_i)}. \tag{34}$$

We find a solution to Eq. (31) when all $J_s^{(i)}$ are equal, say

$$J_s^{(i)} \equiv J_s \qquad i = 1,\ldots N, \tag{35}$$

namely

$$K_N(z_1,\ldots,z_N) = \frac{1}{z_1\ldots z_N} \prod_{i=1}^{N} |z_{i+1}-z_i|^{-J_s-1}, \tag{36}$$

where we have defined $z_{N+1}\equiv z_1$. We point out that this solution is not unique as it can be multiplied by any SU(1,1) scalar function of the z_i's. In particular, we can put an ordering condition on the arguments of the z_i's according to the order in which the vertices appear. As we shall show

by example, this condition is necessary for the factorization of the amplitude. Hence one factorizable amplitude is given by

$$A_N(k_1,\ldots,k_N) = \int \prod_{i=1}^{N} \frac{dz_i}{z_i} |z_{i+1}-z_i|^{-J-1} \theta(\arg z_{i+1}-\arg z_i)$$

$$\langle 0| \prod_{i=1}^{N} V(k_i,\{\lambda\}_i;z_i) |0\rangle \delta^{(4)} \left(\sum_i^N k_i . \right) \qquad (37)$$

Since the integrand is invariant under a three parameter group, it is really a function of N-3 variables. So far we have not said how duality comes about. The fact is that all the vertices we shall consider give rise to cyclic invariant amplitudes. All we can say is that the covariance under SU(1,1) does not seem to be sufficient to insure cyclic invariance. It may be that covariance of the vertices under the Virasoro algebra is a requirement for it.[13] However, in the following we shall not concern ourselves with such highbrow considerations; rather we aim to show in detail how the various ideas discussed above come into being when one considers specific vertices which obey our criteria.

IV. CONSTRUCTION OF THE N-POINT VENEZIANO FUNCTION

In order to give content to the preceding section, we give in great detail the derivation of the N-point function for external scalar particles using as a starting point the dual vertex for the absorption of a scalar particle. We first observe that

$$\left[L_0, F_\rho(z) \right] = -z \frac{d}{dz} F_\rho(z) \qquad (38a)$$

$$\left[L_\pm, F_\rho(z) \right] = -\frac{z^{\pm 1}}{\sqrt{2}} \left(z\frac{d}{dz} \pm \frac{1}{2} \epsilon \right) F_\rho(z) \qquad (38b)$$

where we have used representation (18) for the generators and the commutation relations (12). This means that $F_\rho(z)$ transforms with an SU(1,1) spin $J_s = -\epsilon/2$. The same holds for $\tilde{F}_\rho(z)$. This is a very important point, and it will be

used for functions for which J_s is not infinitesimal when we want to add extra quantum numbers to the model.

Introduce the vertex for the absorption of a scalar particle[14]

$$V(k_\mu,\{0\};z)\equiv V_0(k,z)=e^{-\frac{k^2}{2\varepsilon}}e^{ik\cdot\tilde{F}(z)}e^{ik\cdot F(z)} \tag{39}$$

where the factor appearing in front cancels the infinity that appears in $\tilde{F}_\rho$ and F_ρ. In fact we can rewrite it as

$$V_0(k_\mu;z) = :e^{ik\cdot Q(z)}: \tag{40}$$

where the normal ordering :: only applies to the periodic modes. In order to see if this is a suitable dual vertex, we must check its commutation relations with the SU(1,1) generators.

Since we now approach the realm of detailed calculations, it is good to quote a well-known and probably forgotten identity: if A and B are any two operators, then

$$e^{A}Be^{-A}=B + [A,B] + \frac{1}{2!}\Big[A,[A,B]\Big] + ..., \tag{41}$$

where the other terms are left to the imagination of the reader. Then, it is easy to check that

$$\Big[L_0,\ V_0(k,z)\Big] = -z\frac{d}{dz}\,V_0(k,z) \tag{42a}$$

$$\Big[L_\pm,\ V_0(k,z)\Big] = -\frac{z^{\pm1}}{\sqrt{2}}\left(z\frac{d}{dz}\pm\frac{1}{2}k^2\right)V_0(k,z) \tag{42b}$$

where we have used the mathematically ambiguous[15] form

$$\lim_{\varepsilon\to0}\ \varepsilon\sum_{m=0}^{\infty}\frac{(m+\varepsilon-1)!}{m!} = 1. \tag{43}$$

Nevertheless, the end result is the same whether or not one chooses to calculate using a representation where ε is

not yet equal to zero. The use of Eq. (43) yields consistent results and we shall keep with the use of the representation (18) for the generators.

This means that for a scalar particle $J_s=-\frac{1}{2}k^2$. Introduce the trajectory function

$$\alpha(x) = \alpha_0 + \frac{1}{2}\, x \tag{44}$$

which means that $J_s=-\alpha_0$, the intercept of the mother trajectory. Now that we have a respectable vertex we can try to calculate an amplitude for the absorption of any number of scalar.

Consider the vacuum expectation value of N scalar vertices, the computation of which is made easy by realizing t that the commutator between F and $\tilde{F}$ is a c-number, namely

$$\left[F_\rho(z_j),\, \tilde{F}_\sigma(z_1)\right] = g_{\rho\sigma}\, \left(\frac{1}{\epsilon} - \ln|z_j-z_1| - \frac{1}{2}\, i\pi\phi_{j1}\right) \tag{45}$$

where

$$\phi_{j1} = \begin{cases} +1 & \arg z_j > \arg z_1 \\ -1 & \arg z_j < \arg z_1. \end{cases} \tag{46}$$

Needless to say the last equation is obtained by using Eq. (11) and noting that

$$\frac{z_1-z_j}{\sqrt{z_1 z_j}} = -i\phi_{1j}|z_1-z_j| \tag{47}$$

where z is a point on the unit circle, as well as the expansion for the logarithm

$$\ln(1-x) = -\sum_{n=1}^{\infty} \frac{x^n}{n}\,. \tag{48}$$

We can then use the identity

$$e^A e^B = e^B e^A e^{[A,B]}, \tag{49}$$

which holds only when [A,B] is a c-number. It then follows from Eqs. (45) and (46) that

$$\langle 0|\prod V_0(k_i,z_i)|0\rangle = e^{-\frac{1}{2\epsilon}\sum_{i=1}^{N} k_i^2} e^{-\sum_{j<1}\left[k_j\cdot F(z_j),k_1\cdot\tilde{F}(z_1)\right]} \tag{50a}$$

$$= e^{-\frac{1}{2\epsilon}\sum_{i=1}^{N} k_i^2 - \frac{1}{\epsilon}\sum_{j<1} k_j\cdot k_1 + \frac{i\pi}{2}\sum_{j<1} k_j\cdot k_1\phi_{j1}} \prod_{j<1}|z_j-z_1|^{k_j\cdot k_1} \tag{50b}$$

By noting that

$$\sum_{i<j} k_i\cdot k_j = \frac{1}{2}(k_1+\ldots+k_N)^2 - \frac{1}{2}\sum_{i=1}^{N} k_i^2, \tag{51}$$

we get rid of the infinite factor by using in addition conservation of momentum. The phase factor can be absorbed only if ϕ_{j1} does not change sign, which shows the need for the ordering condition on the angles. As stated before we can form an amplitude only when all the SU(1,1) spin are equal, i.e. in this case only when all the scalar particles have the same mass. In this case the amplitude is given by

$$A_N(k_1,\ldots,k_N) = \int \prod_{i=1} \frac{dz_i}{z_i}|z_{i+1}-z_i|^{\alpha_0-1}\theta(\arg z_{i+1}) \prod_{j<1}|z_j-z_1|^{k_j\cdot k_1} \tag{52}$$

which is, up to a factor, the Koba-Nielsen[2] form. Note the disappearance of the kernel when $\alpha_0=1$. We have stated above that, due to the invariance of the integrand under a three parameter group, three integration variables are superfluous. So as to give meaning to this statement, we now proceed to show how the above reduces to the well-known B-function in the case N=4 (see Fig. 2).

Introduce the anharmonic ratio

$$x = \frac{(z_1-z_2)(z_3-z_4)}{(z_1-z_3)(z_2-z_4)} \tag{53}$$

which is real when all the z's are on the unit circle. Then

$$(1-x) = \frac{(z_1-z_4)(z_2-z_3)}{(z_1-z_3)(z_2-z_4)} . \tag{54}$$

Together with the use of the kinematical relations

$$k_1\cdot k_2 = k_3\cdot k_4 = -\alpha(s) - \alpha_0$$

$$k_1\cdot k_4 = k_2\cdot k_3 = -\alpha(t) - \alpha_0$$

$$k_1\cdot k_3 = k_2\cdot k_4 = \alpha(s) + \alpha(t)$$

we obtain

$$\prod_{j<1} |z_j-z_1|^{k_j\cdot k_1}$$

$$= x^{-\alpha(s)}(1-x)^{-\alpha(t)}(|z_1-z_2||z_3-z_4||z_2-z_3||z_1-z_4|)^{-\alpha_0} . \tag{56}$$

The last factor is cancelled by the measure, leaving us with

$$\int \frac{dz_1 dz_3 dz_4}{(z_1-z_3)(z_3-z_4)(z_4-z_1)} \theta(\arg z_1-\arg z_3)\ \theta(\arg z_3-\arg z_4)$$

$$\times\ \theta(\arg z_4-\arg z_1)\int_0^1 dx\ x^{-1-\alpha(s)}(1-x)^{-1-\alpha(t)} . \tag{57}$$

Hence all the kinematics are contained in one integrand while a three-fold integration separates. Furthermore it can be checked that the range of integration of x is from 0 to 1 only because of the θ-functions. Thus we have

$$A_4(k_1,\ldots,k_4) = \int dH \int_0^1 dx\ x^{-1-\alpha(s)}(-x)^{-1-\alpha(t)} . \tag{58}$$

The differential dH is known as the Haar measure.[16] It is an infinite constant and should be divided out of the

amplitude. Calculations performed for an arbitrary number of legs show it to be independent of the number of external particles. It acquires a certain physical significance when one realizes that it is equal to the three-point function between scalar particles and could thus be interpreted as a bare coupling constant. This interpretation, however, is not consistent with factorization of an amplitude with many external particles.

We should emphasize that the ordering condition is necessary for factorization in the sense that it allows for the correct range of x. The amplitude (58) factorizes and corresponds to Fig. 2. We should, in order to obtain the full amplitude, add all the inequivalent penetrations of the external legs, as shown in Fig. 3.

When $\alpha_0=1$, however, the amplitude we calculate by omitting the ordering conditions is automatically equal to the sum of all the inequivalent penetrations. This remarkable property is true independent of the number of legs.[17]

We went through this calculational section to acquaint the reader with the mathematical techniques used in dual resonance models. Although the calculations were performed using the scalar vertex, much of the "meat" is the same when considering amplitudes (or equivalently vertices) with more complicated external particles. We feel that it is always good to give meaning to the abstruse concepts of the previous section by showing explicitely how they lead to familiar results.

V. GAUGE CONDITIONS

In constructing the scalar vertrices, it seems that we got more than we bargained for. Indeed, using the representation (19) for the Virasoro operators (of which the SU(1,1) generators form a subset), we find the following commutation relations[18]

$$\left[L_{\pm n}, V_0(k;z)\right] = -\frac{1}{\sqrt{2}}\, z^{\pm n}\left(z\frac{d}{dz} \pm n\frac{k^2}{2}\right)V_0(k;z) \qquad (59)$$

which means that the scalar vertices are covariant under

a much more general algebra. The physical meaning of this peculiar equation becomes clear in the case $\alpha_0=k^2/2=1$. We can then rewrite the commutator as

$$\left[L_{\pm n},\ V_0^{\alpha_0=1}(k,z)\right] = -\frac{1}{\sqrt{2}}\, z\, \frac{d}{dz}\left[z^{\pm n} V_0^{\alpha_0=1}(k;z)\right] \qquad (60)$$

remembering that z is on the unit circle, this means that the commutator is a perfect differential. In the case $\alpha_0=1$, the amplitude can be written as

$$A_N(k_1\ldots,k_N) = \langle 0|\prod_{i=1}^{N} \upsilon(k_i)|0\rangle\, \delta^{(4)}\left(\sum_1^N k_i\right) \qquad (61)$$

where

$$\upsilon(k_i) = \int_0^{2\pi} \frac{dz_i}{z_i}\, V_0^{\alpha_0=1}(k_i,z_i). \qquad (62)$$

Then, using the periodicity condition, we obtain

$$\left[L_{\pm n},\ \upsilon(k)\right] = 0. \qquad (63)$$

Since $L_{-n}|0\rangle=0$, it is clear that the above means that there are subsidary conditions between the states appearing in the factorization of the model. A quick look back at the explicit representation of the L_m Eq. (21) shows that in the rest frame, they relate the "ghost-like" mode introduced by $a_0^{(n)}$ (n=1,....) which gives rise to negative norm states to other states in the theory. In fact explicit calculations show that they act as "ghost compensators" in such a way that there does not seem to be any negative norm state when $\alpha_0=1$ although there is a tachyon at the nonsense point of the mother trajectory.[19] We note that such a tachyon will appear whenever the mother trajectory has positive intercept as it does in the real world for the ρ! We wish to stress that although this condition for ghost elimination so far holds only when the mother trajectory has unit intercept, the DRM's are the only model to have such a mechanism. Indeed it may be argued that a similar mechanism will always be needed for relativistic theories of strong interactions where an

infinite number of negative norm states will be introduced by the nature of the Lorentz metric.

At this stage of our understanding of the dual resonance model, it is well to review what we have. First of all, the model we have considered has no internal quantum numbers so that we are really dealing (at best) with what is hopefully the skeleton of a strong interaction theory. The nature of the commutation relations (12a) introduces negative norm states which are compensated only when the mother trajectory has unit intercept, thereby introducing a tachyon (which we would have had in any case in any channel involving the ρ trajectory). The model has an additional difficulty because the J_s of the dual vertices depends on the mass of the particle; since they have to be the same to obtain SU(1,1) invariance, it is not clear how to go off mass shell and keep SU(1,1)[20] This problem is, of course, acutely felt when one wants to introduce electromagnetic interactions.

Although the above remarks make it clear that the model must be improved, we have found a surprising group theoretical structure which seems to be at the origin of all the esthetic properties of the model. In addition we have found the existence of a gauge-like algebra, which seems to eliminate unwanted negative norm states. It is clear then that at least one of these features must be kept in devising more physical dual models.

In order to gain more familiarity with the SU(1,1) aspect of the bare model, we will try to build dual vertices for the excited states of the theory. Then we will try to add quantum numbers to the bare theory and will examine several models that were recently proposed.

VI. DUAL VERTICES FOR EXCITED STATES

We now try to construct dual vertices for the excited particles that appear in the bare dual model. We start by constructing the vertex for the emission of a vector particle. Recall that P_μ is like a generalized momentum, which suggests the form[11]

$$V_\mu(k,\ j=1;z) = e^{ik\cdot\tilde{F}(z)} P_\mu(z)\ e^{ik\cdot F(z)} \qquad (64)$$

which is written in a normal ordered form and where we disregard the factor $e^{-k^2/2c}$ from now on. Using the representation (18) for the generators, we find that

$$\left[L_0, V_\mu(k,\ j=1;z)\right] = -z\frac{d}{dz}\ V_\mu(k,1;z) \tag{65a}$$

$$\left[L_\pm,\ V_\mu(k,1;z)\right] = -\frac{z^{\pm 1}}{\sqrt{2}}\left[z\frac{d}{dz} \pm \ 1+\frac{k^2}{2}\ \right]$$
$$\times\ V_\mu(k,1;z) + \frac{z^{\pm 1}}{2\sqrt{2}}\ k_\mu V_0(k,z)\,. \tag{65b}$$

There are two important things to notice about the last equation. First, as written $V_\mu(k,1;z)$ is not covariant under SU(1,1) because of the second term appearing in Eq. (65b) and that this extra term is along k_μ and appears with the same sign for both L_+ and L_-. The only way to get rid of this term is to put a spin 1 projection operator! Thus the requirement of covariance under SU(1,1) forces the addition of the projection operator. The dual vertex for a vector meson is then

$$V_\mu(k,z) = \left(g_{\mu\nu} - \frac{k_\mu k_\nu}{k^2}\right) e^{ik\cdot\tilde{F}(z)} P_\nu(z)\ e^{ik\cdot F(z)}. \tag{66}$$

The second thing to notice is that the J_s of this vector dual vertex is

$$J_s = -\left(1+\frac{k^2}{2}\right) = -\left(1-\frac{M_v^{\,2}}{2}\right) \tag{67}$$

Using our criterion (4), this vector meson will interact in a dual way with scalars only if

$$\frac{M_s^{\,2}}{2} = -1 + \frac{M_v^{\,2}}{2} \tag{68}$$

where M_s and M_v are the masses of the scalar and the vector particles, respectively. This means that our vector meson lies on the first recurrence of the mother trajectory.

Similarly, we can find the dual vertex for a particle of spin two.

Introduce the notation

$$\hat{P}_\mu(z) \equiv \left(g_{\mu\nu} - \frac{k_\mu k_\nu}{k^2}\right) P_\nu(z). \tag{69}$$

The obvious choice is

$$V_{\mu\nu}(k,\ j=2;z) = e^{ik\cdot\tilde{F}(z)} :\ \hat{P}_\mu(z)\hat{P}_\nu(z):\ e^{ik\cdot F(z)}. \tag{70}$$

Explicit calculation shows that as it stands this vertex transforms covariantly under SU(1,1) with a dual spin

$$J_s^{(2)} = -\left(2 + \frac{k^2}{2}\right). \tag{71}$$

However, before interpreting it as the dual vertex for a spin two particle, we must subtract the traces. It turns out that this procedure is enforced by requiring covariance under $L_{\pm 2}$!!.[21] Hence, as hinted at in the previous section, the Virasoro operators play the role of projecting the vertices into definite spin states. The relation (71) shows that the spin 2 particle we are talking about lies on the second recurrence of the mother trajectory. It is rather straightforward to generalize our procedure to take into account all the states on the mother trajectory. (See Fig. 4)

It turns out that dual vertices can be written for some of the daughter states that appear in the theory, as has been shown by Fubini and collaborators.[22] This construction is relevant only where the propagator is diagonal, i.e. $\alpha_0 = 1$. We quote the result for the spin 1 daughter which has a dual vertex

$$V_\mu^D(k,z) = e^{ik\cdot\tilde{F}(z)} :\left(z\frac{d}{dz} + \frac{2}{k^2}k\cdot P\right)\hat{P}_\mu + k_\mu\ \{....\}:\ e^{ik\cdot F(z)}. \tag{72}$$

This vertex has $J_s^D = -(2+k^2/2)$ so that it lies under the spin 2 state on the mother trajectory. We note, however, that it has a component along k_μ so that in this case covariance under SU(1,1) is not sufficient to eliminate this

troublesome component. Indeed we need covariance under $L_{\pm 2}$ to handle it satisfactorily. However, the problem is technically very complicated and as of this writing not entirely solved, i.e. although one finds all the possible daughter vertices covariant under SU(1,1) it is very hard to separate the correct linear continuation which are also covariant under $L_{\pm 2}$, although considerable effort in this direction has been spent.

We wish to point out, however, that the spin 1 part of this daughter vertex can be rewritten as a perfect differential when $\alpha_0=1$, so that the first daughter decouples according to the mechanism outlined in the previous section. Actually this phenomenon occurs all along the first daughter trajectory when $\alpha_0=1$ so that it decouples entirely from the problem. This was first pointed out by Di Vecchia and Del Giudice[19] by a close analysis of the spectrum.

The main conclusion of this section is that the dual daughter vertices for definite spin states have not all been constructed as they must be covariant under the Virasoro algebra.[23] This problem seems hopelessly complicated at the moment, and we have nothing to add to it; rather we turn our attention to the inclusion of internal quantum numbers in the bare model.

VII. DUAL MODELS WITH ADDITIONAL QUANTUM NUMBERS

As stated earlier, another direction of research is to incorporate additional degrees of freedom into the bare model without upsetting the group theoretical properties under at least SU(1,1). In this section we describe three such models in their chronological appearance in the literature. These are the models proposed by Bardakci Halpern[24](I), Clavelli[25](II), and Neveu and Schwarz[26](III). Their common feature is that they start by introducing new operators as coefficients of the basis functions of the bounded representation of SU(1,1). They all have a G-parity operator and display a spectrum which, although not yet the physical one, shows great improvement over that of the bare model. The last model (III) has a new feature which is responsible for decoupling the tachyon appearing on the mother trajectory although another

appears in the model. These statements will be clarified by considering the models in detail.

A. Bardakci-Halpern Model

The new degrees of freedom are introduced through the quark-like operator function

$$\psi_{(r)} = \sum_{m=0}^{\infty} b_r^{(m)} |-\tfrac{1}{2},m)_+ + d_r^{(m)\dagger} |-\tfrac{1}{2},m)_- \quad r=1,2,3, \tag{73}$$

where the notation for the states is that of Section II and the coefficient obeys the following anticommutation relations

$$\left\{b_r^{(m)}, d_s^{(n)}\right\} = \left\{b_r^{(m)\dagger}, d_s^{(n)}\right\} = 0$$

$$\left\{b_r^{(n)}, b_s^{(m)\dagger}\right\} = \delta_{rs}\delta_{nm} = \left\{d_r^{(n)}, d_s^{(m)\dagger}\right\} \quad \begin{matrix} r,s=1,2,3 \\ m,n=0,1,\ldots \end{matrix} \tag{74}$$

The point of this construction is that one can define new SU(1,1) and Virasoro operators

$$L'_{-m} = -\frac{i}{4\pi}\int_0^{2\pi}\frac{dz}{z}\, z^{-m} : \left[\psi_{(r)}, z\,\frac{d}{dz}\psi_{(r)}\right] : \tag{75}$$

such that the quark-like functions $\psi_{(r)}$ transform under these new operators with a SU(1,1) spin -1/2. It follows that, if we define the new SU(1,1) operators as the sum of those appearing in Section II and the operators defined by Eq. (75)

$$L_{-m}^{T} = L_{-m} + L'_{-m}, \tag{76}$$

then the following vertices transform covariantly under $L_{\pm m}^{T}$, L_0^{T}:

(a) a vertex without internal quantum number "Pomeron" vertex which is just the scalar vertex of the bare theory

$$V^P(k,z) = V_0(k,z) \text{ with } J_s^P = -\frac{k^2}{2} \tag{77a}$$

(b) a vertex representing a quark-like state

$$V_{(r)}^{Q}(k,z)=\psi_{(r)}(z)V_0(k,z) \text{ with } J_s^{Q}= -\frac{1}{2}-\frac{k^2}{2} \tag{77b}$$

(c) a meson-like vertex

$$V_\alpha^{M}(k,z)=:\psi^\dagger(z)\lambda_\alpha\psi(z):V_0(k,z) \text{ with } J_s^{M}= -1-\frac{k^2}{2} \tag{77c}$$

here the λ_α are the SU(3) matrices.

As implied by the definition (75), the operators L_{-m}^{T} satisfy among themselves a Virasoro algebra. In addition, since the above quoted vertices are also covariant under the Virasoro operators, there is a decoupling scheme at work where $J_s=-1$ which conveniently yields a mass zero meson. In this case, the spectrum is shown in Fig. 5. The model suffers from certain diseases, namely the lack of half integer spaced trajectories, existence of tachyons and exotic "quark" states.

B. Clavelli Model

In this case the new degrees of freedom are introduced by a scalar function belonging to $D_J^{(+)}$ and $D_J^{(-)}$ with $J=-\frac{1}{4}$

$$H(z) = \sum_{m=0}^{\infty} b^{(m)}\left|-\tfrac{1}{4},m\right\rangle_- + d^{(m)\dagger}\left|-\tfrac{1}{4},m\right\rangle_+ \tag{78}$$

where

$$\left\{b^{(m)},d^{(n)}\right\} = \left\{b^{(m)},d^{(n)\dagger}\right\} = 0$$
$$\left\{b^{(m)},b^{(n)\dagger}\right\} = \left\{d^{(m)},d^{(n)\dagger}\right\} = \delta_{n,m} \quad n,m=0,1,\ldots \tag{79}$$

The new SU(1,1) generators are introduced by sandwiching the representation (5) of the generators between the states H in the same way as was done

earlier for the bare model. However, it does not seem to be possible to build the Virasoro operators.

As should be obvious by now, the H(z) transform covariantly under these new SU(1,1) operators with $J_S = -\frac{1}{4}$. The dual vertex for a pseudo-scalar meson is

$$V^M(k,z) = : H^\dagger(z)H(z):V_0(k,z), \tag{80}$$

which has

$$J_s = -\frac{1}{2} - \frac{1}{2}k^2. \tag{81}$$

There is a G-parity operator

$$G = e^{i\pi \sum_{m=0}^{\infty} (b^{(m)\dagger}\cdot d^{(m)} + d^{(m)\dagger}\cdot b^{(m)})}. \tag{82}$$

The obvious choice is to take the dual vertex $V^M(k,z)$ to represent the pion. Then one fixes J_s so that its mass vanishes through Eq. (81). The spectrum of states one obtains this way is shown in Fig. 6. It has the virtue of having a ρ trajectory with the correct intercept, i.e., half integer spacing between meson trajectories. Although many particles have their correct mass value, (π,ρ,A_1), the model has no room for abnormal parity coupling $(\omega, A_2,\ldots)$. In addition, negative norm states appear on the fifth trajectory. Another model was considered by the same author to include SU(3) breaking by introducing the quark-like function

$$H_{(r)}(z) = \sum_{m=0}^{\infty} b_r^{(m)} |\eta_r,m)_- + d_r^{(m)} |\eta_r,m)_+ \quad r=1,2,3 \tag{83}$$

where the η_r act as the breaking parameters. However, the model leads to $\pi^o\eta$ degeneracy.

C. Neveu-Schwarz Model

The authors consider the function

$$H_\mu(z) = \sum_{m=0}^{\infty} c_\mu^{(m)\dagger} |-\tfrac{1}{2},m)_- + c_\mu^{(m)} |-\tfrac{1}{2},m)_+ \qquad (84)$$

where $c_\mu^{(n)}$ and $c_\mu^{(n)\dagger}$ are four-vector operators obeying the anticommutation relations

$$\left\{c_\mu^{(n)}, c_\rho^{(m)\dagger}\right\} = \delta_{nm}\delta_{\mu\rho} \qquad (85)$$

from which one can build the Virasoro operators

$$L_{-m}^{(c)} = \frac{1}{4\pi}\int_0^{2\pi}\frac{dz}{z} z^{-m} :H(z)\ z\frac{d}{dz}H(z): \qquad (86)$$

which obey the usual algebra.

The new operators are the sum, as in the previous models

$$L_{-m}^{T} = L_{-m}^{(a)} + L_{-m}^{(c)}. \qquad (87)$$

It is no wonder that under these operators, $H_\mu(z)$ transforms covariantly with $J_s=-1/2$. At this point, a new feature of this particular model emerges. Since H_μ is a four vector, it can be coupled to P_μ, which leads us to consider the operators[27]

$$G_{-m} = \frac{1}{\sqrt{2\pi}}\int_0^{2\pi}\frac{dz}{z} z^{+m} H_\mu(z) P^\mu(z) \qquad m=\pm\tfrac{1}{2},\pm\tfrac{3}{2},\ldots \qquad (88)$$

which satisfy

$$\left[L_m^{T}, G_n\right] = \left(\frac{m}{2} - n\right) G_{n+m} \qquad (89)$$

$$\left\{G_n, G_m\right\} = -2L_{n+m}^{T}. \qquad (90)$$

As we shall see later these new operators act as additional decouplers when $\alpha_0=1$. The dual pion vertex is

$$V^\pi(k,z) = k\cdot H(z) V_0(k,z) \text{ which has } J_s = -\frac{1}{2} - \frac{k^2}{2}, \qquad (91)$$

which can be rewritten as

$$V^{\Pi}(k,z) = -\frac{z^{-\frac{1}{2}}}{\sqrt{2}}\left[G_{\frac{1}{2}}, V_0(k,z)\right] \tag{92}$$

then it follows that

$$\left\{G_{\frac{1}{2}}, V^{\Pi}(k,z)\right\} = -\sqrt{2}\, z^{-\frac{1}{2}}\left[L_1, V_0(k,z)\right] \tag{93a}$$

$$= -\sqrt{2}\, z^{\frac{1}{2}}\left(\frac{d}{dz} + \frac{k^2}{2}\right)V_0(k,z). \tag{93b}$$

Since $\alpha_0=1$, $k^2/2=1/2$ so that in fact we have

$$\left\{G_{\frac{1}{2}}, V^{\Pi}(k,z)\right\} = -\sqrt{2}\, z\, \frac{d}{dz}\left[z^{\frac{1}{2}} V_0(k,z)\right]. \tag{94}$$

We have a perfect differential on the right-hand side of Eq. (94). Such a perfect differential eliminates an integration variables in the amplitude and this eliminates a propagator, thus giving zero for the amplitude. These are the new gauges introduced by Neveu and Schwarz, and they serve to decouple the tachyon lying on the mother trajectory. More detail is to be found in Ref. 26.

The virtues of this model are quite remarkable since it allows for abnormal parity couplings, the first dual factorizable model to do so. Also since there are two decoupling schemes at work, it is not likely that negative norm states will appear in the model.

Although the discovery of these new gauges allows for the construction of a more "real life" model (see Fig. 7), it is clear that one still has a long way to go. It should be noted that by adding a fifth mode[28] to the Neveu Schwarz model, one can eliminate the tachyon,[29] but the price is the loss of half integer spacing between meson trajectories and an increase in the ω-ρ mass difference.

VII. CONCLUSION

These lectures have been delivered with the aim of familiarizing the reader with what seems to many to be an exotic field of physics; we hope they have been successful in this respect. For the sake of completeness, however, we should point out the existence of more fundamental approaches to dual theories that have been sparked by Y. Nambu[30] and H. B. Nielsen[31] as well as many other people.[32] The most exciting aspect of these works is the understanding of the Virasoro algebra as a gauge group, not unlike that found in general relativity.[33] There is little doubt that if duality has anything to do with strong interactions,[34] these can be considered as the strong gauges, pretty much on the same footing as the electromagnetic gauge for electromagnetic interactions.

In summary we can say that we are at the beginning of an understanding of duality in terms of strong interactions and that the theories we discussed are necessarily very elementary, but there are group theoretical concepts that seem to transcend any given dual model--when we understand their origin we shall undoubtedly be able to build more satisfactory (in the sense that they reproduce the observed spectrum) dual models.

ACKNOWLEDGMENTS

The author wishes to thank Profs. Brittin and Barut for their kind hospitality. He also would like to thank Profs. G. Domokos, S. Kovesi-Domokos, L. Girardello, and W. Wyss for challenging discussions at the summer school. Discussions with Dr. Clavelli were useful and continuous throughout the preparation of these lectures.

FIGURE CAPTIONS

Fig. 1. Dual N-point factorizable amplitude.

Fig. 2. Four point amplitude.

Fig. 3. Total dual amplitude for four external scalars

Fig. 4. Particle spectrum in the $\alpha_0=1$ case of the bare dual model. The dotted line means that the particles lying on it are decoupled from the rest.

Fig. 5. Particle spectrum in the Barkakci-Halpern model.

Fig. 6. Particle spectrum in the Clavelli model.

Fig. 7. Particle spectrum in the Neveu-Schwarz model.

Fig. 1

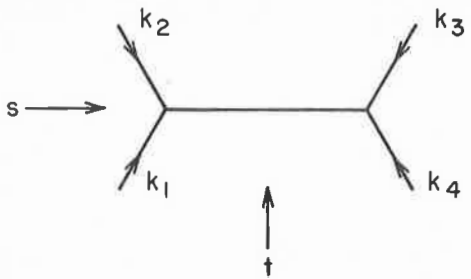


Fig. 2

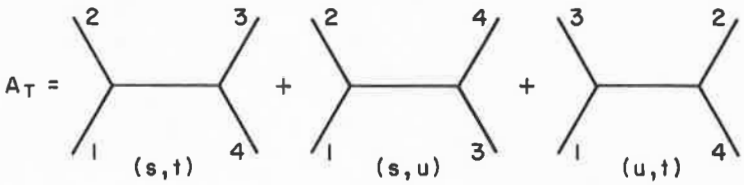

A_T =
2
3
1
4
(s,t)
+
2
4
1
3
(s,u)
+
3
2
1
4
(u,t)

Fig. 3

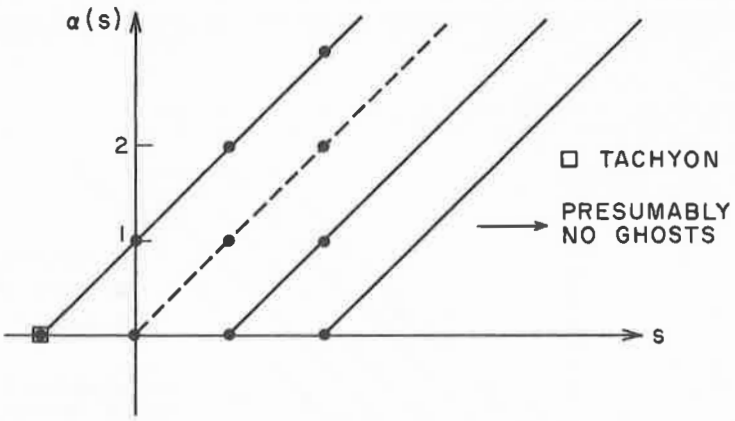

α(s)
2
1
s
□ TACHYON
PRESUMABLY
NO GHOSTS

Fig. 4

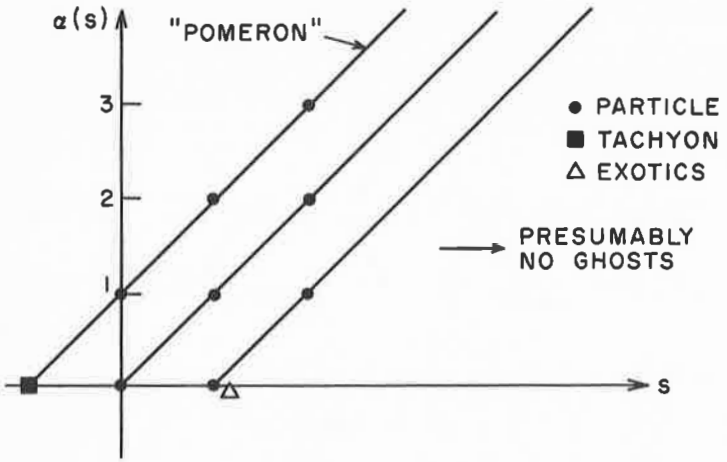

α(s)
"POMERON"
3
2
1
s
● PARTICLE
■ TACHYON
△ EXOTICS
PRESUMABLY
NO GHOSTS

Fig. 5

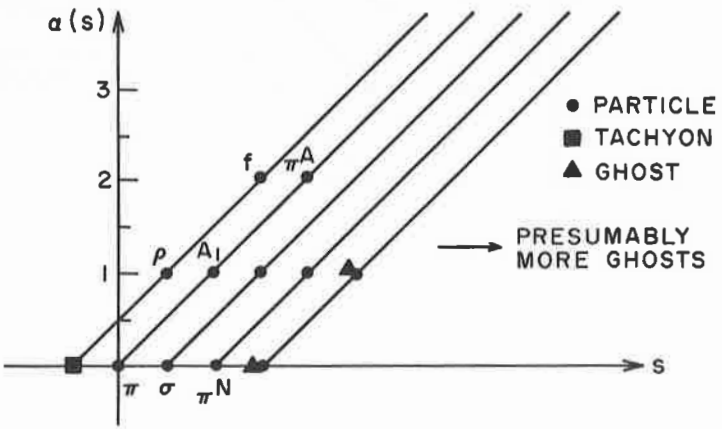

α(s)
3
2
1
f
πA
ρ
A1
π
σ
πN
s
● PARTICLE
■ TACHYON
▲ GHOST
PRESUMABLY
MORE GHOSTS

Fig. 6

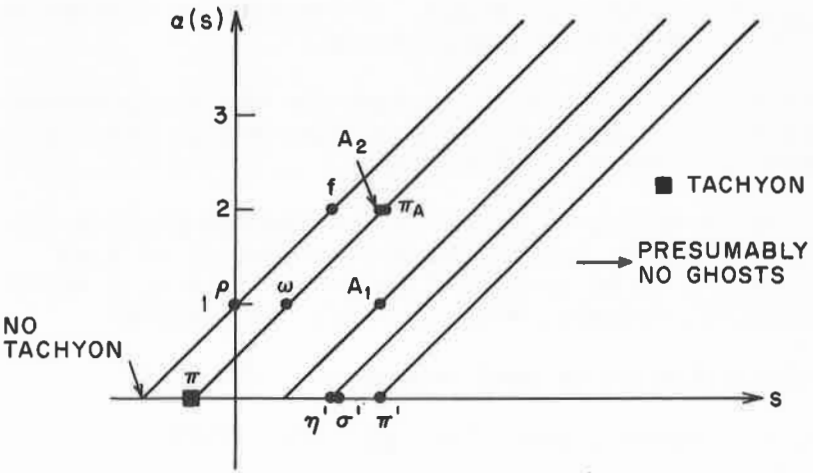

α(s)
3
2
1
A2
f
πA
ρ
ω
A1
π
NO
TACHYON
TACHYON
PRESUMABLY
NO GHOSTS
s
η'
σ'
π'

Fig. 7

REFERENCES

1. See, for example, G. Veneziano, Lectures at the Erice Summer School (1970).

2. Z. Koba and H. Nielsen, Nucl. Phys. B12, 517 (1969).

3. For greater detail concerning the representations of SU(1,1) see the appendix of L. Clavelli and P. Ramond, Phys. Rev. D2, 973 (1970) and all references cited therein.

4. L. Clavelli and P. Ramond, op.cit.

5. S. Fubini, D. Gordon, and G. Veneziano, Phys. Letters 29B, 679 (1969); Y. Nambu University of Chicago Report No. EFI 69-64 (unpublished).

6. The relevance of the zero mode for the group theory was first pointed out by S. Fubini and G. Veneziano, Nuovo Cimento 67A, 29 (1970).

7. A. Della Selva, S. Saito, Nuovo Cimento Letters IV, 689 (1970); P. Ramond, An Interpretation of Dual Theories, to be published in Nuovo Cimento; K.Bardakci and M. B. Halpern, Phys. Rev. D3, 2493 (1971).

8. This procedure is used throughout Ref. 3.

9. M. A. Virasoro, Phys. Rev. D1, 2933 (1970).

10. This algebra was first derived by S. Fubini and G. Veneziano in Ann. Phys. (N.Y.) 63, 12 (1971) without the c-number. Its existence was pointed out independently by J. Weiss and L. Clavelli. An independent derivation is found in Y. Nambu, University of Chicago Preprint #EFI-70-07 (unpublished).

11. L. Clavelli and P. Ramond, Phys. Rev. D3, 987 (1971).

12. This was first shown by L. Clavelli and P. Ramond, Ref. 3, for scalar external legs.

13. The author wishes to thank Prof. K. Bardakci for a discussion on this point.

14. S. Fubini and G. Veneziano, Nuovo Cimento 67A, 29 (1970); L. Clavelli and P. Ramond, Ref. 3.

15. This ambiguity may be avoided by considering representations of SL(2,R); thanks are due to Dr. I. Bars for this remark.

16. See the appendix of Ref. 2.

17. D. B. Fairlie and K. Jones, Nucl. Phys. B15, 323 (1970).

18. See Ref. 10.

19. E. Del Giudice and P. Di Vecchia, MIT Preprint 167 (1970); R. Brower and C. Thorn have shown that the first nine trajectories are free of ghosts when $\alpha_0=1$ in a CERN preprint (1971).

20. See in this respect the recent works of C. Lovelace, Phys. Letters 34B, 500 (1971), I. T. Drummond, CERN preprint Th. 1301 (1971), C. Rebbi, University of Trieste preprint (1971), and L. Clavelli and P. Ramond NAL preprint (1971).

21. The author thanks Prof. P. Di Vecchia for a discussion on this point.

22. P. Campagna, E. Napolitano, S. Sciuto, and S. Fubini, Nuovo Cimento 2A, 911 (1971). Amplitude with excited vertices were built by C. Rosenzweig and Sukhatme (to be published in Nuovo Cimento).

23. See in this respect P. Di Vecchia and S. Fubini, MIT preprint #194 (1971). The problem shows its full complexity in the papers by Brower and Thorn, op. cit., and by Galzenati, F. Gliozzi, R. Musto, and F. Nicodemi, Naples preprint (1970).

24. K. Bardakci and M. Halpern, Phys. Rev. D3, 2493 (1971).

25. L. Clavelli, Phys. Rev. D3, 3166 (1971).

26. A. Neveu and J. Schwarz, Factorizable Dual Model of Pions, (to be published in Nucl. Phys. B); A. Neveu, J. Schwarz, and C. Thorn, Reformulation of the Dual Pion Model, Princeton University preprint.

27. This type of "square-root" algebra was first considered in the process of deriving a dual analog to the Dirac equation of P. Ramond, Phys. Rev. D3, 2415 (1971).

28. This trick of adding a fifth mode was first introduced independently by Y. Nambu and S. Fubini and G. Veneziano, op. cit.

29. M. B. Halpern and C. B. Thorn, Phys. Letters 35B, 441 (1971).

30. Y. Nambu, as cited in Refs. 5 and 10.

31. H. B. Nielsen, NORDITA preprint (1969), unpublished.

32. Y. Miyamoto, Prog. Theor. Phys. 43, 564 (1970); P. Olesen, CERN preprint 1280 (1971); L. N. Chang F. Mansouri, (in preparation); P. Ramond, Phys. Rev. D3, 2415 (1971) and Ref. 7.

33. The author thanks Profs. L. Girardello and W. Wyss for enlightening him on this point. See A.O. Barut, L. Girardello, and W. Wyss (in preparation).

34. See the lectures of Prof. Squires in this volume for the experimental tests of duality.

SELF-CONSISTENT ELECTROMAGNETIC "DUAL" AMPLITUDES AND THE ELECTRON-POSITRON SYSTEM

James McEnnan*
University of Florida
Gainesville, Florida 32601

I. INTRODUCTION

Quantum electrodynamics is almost the only part of elementary particle physics in which there is no doubt about how to procede to a definite answer to any problem which lies within its scope. It is also fair to say that fundamentally QED is no better understood than any other elementary particle theory. It is merely the case that in QED the simplest algorithms seem to work. However, no one seems to know exactly why they work. In fact, one can almost immediately generate several reasonable arguments why they should not work. This applies in particular to perturbation theory, which has become so identified with the relativistic quantum mechanics of electromagnetically interacting particles that virtually no distinction is made. Since there is no "real" theory of electromagnetic interactions, just as there is no strong or weak interaction theory, only a set of rules which, in the case of QED, give reasonable answers, it is of some interest to see if the pure electromagnetic interactions of elementary particles can be treated without reference to the usual formulations. For this reason, I have made some attempts to do the classic QED calculations in a manner which is independent of perturbation theory. This has led me to an analytic S-matrix approach. There, however, one runs into problems almost immediately. Consider, for example, the calculation of the electron anomalous magnetic moment. We would like to calculate the (unitarity) diagram of fig. (1). In the helicity representation, the vertex can be written in the form $a_{\lambda\epsilon} D^{J=1}_{\lambda\epsilon}(\theta',\phi')$, where the

*Present address: University of Pittsburgh, Pittsburgh, Pennsylvania 15213

$a_{\lambda\varepsilon}$ are linear combinations of the electron form factors. The electron-positron amplitude can also be expanded in a Jacob-Wick expansion, so that

$$H_{\lambda\mu}(s,t,u)=\sum(2J+1)h^{J}_{\lambda\mu}D^{J}_{\lambda\mu}(\theta'',\phi''). \qquad (1)$$

With this, we find

$$\mathrm{Im}a_{\lambda\varepsilon}=2\pi\rho\sum h^{J=1}_{\lambda\mu}a_{\mu\varepsilon}, \qquad (2)$$

so that only the J=1 angular momentum state of the electron-positron amplitude contributes to the electron form factors. Everyone knows this. However, the J=1 phase shifts are not known. The perturbation amplitude is useless since it does not have a finite Jacob-Wick expansion. Actually, it is less than useless since one can prove that the perturbation expansion does not converge to the correct amplitude when the process has bound states. So, what do we do? Well, we read the gospels according to Chew and company and note that elementary particle physics should be self-consistent and this is good. Of course, since we are interested in electromagnetic interactions we have to introduce a minor heresy--that self-consistency also applies (in a certain way) to QED. (Note that I mean by self-consistency the same sort of thing one attempts in the bootstrap program in strong interaction physics.) This brings us to the subject of my lectures--the construction and properties of self-consistent dispersion-theoretic electromagnetic scattering amplitudes. (Actually, for the benefit of the purists, I should point out that what results is only a semi-boostrap. The electron pole and residue are inserted a priori, but the amplitude is self-consistent only if it has positronium poles at the position determined by the usual Coulomb bound state formulas.) I got into this because I needed usable amplitudes for dispersion calculations in electrodynamics. It has turned out, however, that the means have become more interesting that the ends.

II. THEORETICAL CONSIDERATIONS

The procedure for constructing self-consistent electromagnetic scattering amplitudes is actually very simple:

1. Unitarity, which gives a singular, inhomogeneous, non-linear integral equation for the amplitude is used to evaluate the double spectral functions (dsf) directly. Essentially, only the diagrams of fig. (2) contribute through second order in the fine structure constant, α.

   a. To evaluate the inhomogeneous terms we employ the usual pole approximations to the amplitudes which appear.

   b. To evaluate the elastic unitarity we use an initial trial amplitude which is based on the Coulomb amplitude.

2. We use the calculated dsf to construct a new trial amplitude which is then reinserted into the elastic unitarity to obtain a new estimate of the dsf.

3. The process is iterated until self-consistency is achieved and the calculated dsf remain unchanged, to the desired order, when the new amplitude is reinserted.

Because of our particular choice for the initial trial amplitude, our second order amplitude does not require iteration. The self-consistent electromagnetic scattering amplitude which results from our procedure will be cutoff independent, analytic and will satisfy the requirements of crossing. It will have the correct dsf and will reduce to the usual Born approximation in lowest order. (Actually, the requirement that the amplitude reduce to the Born approximation is useful in determining its form.) Self-consistency requires the correct Regge asymptotic behavior and the Coulomb bound state poles. Also, the amplitude must have a Jacob-Wick expansion.

We should point out why we emphasize the role of the double spectral functions. Actually, there are two reasons:

1. Simplicity
2. The dsf have no poles, so that an expansion in a power series in α may actually be possible.

To elucidate the second point, let us consider the Schrodinger equation with a Coulomb potential. The scattering amplitude can be obtained in closed form using parabolic coordinates. For a particle of mass m and momentum p in an attractive potential, we find

$$A(E,z)=\frac{\alpha m}{2p^2}\frac{\Gamma(1-i\eta)}{\Gamma(1+i\eta)}\left(\frac{-t}{4p^2}\right)^{-1+i\eta}-\frac{1}{ip}\delta(1-z), \tag{3}$$

where $\eta=\alpha m/p$, $t=-2p^2(1-z)$, $E=p^2/2m$ and $z=\cos\theta$. $\delta(x)$ is the Dirac delta-function. (Note that the delta-function is necessary to satisfy unitarity. For the Coulomb potential, the S-matrix is the connected part--not the scattering amplitude.) We see that A(E,z) has Regge asymptotic behavior with only the leading trajectory contributing and, due to the gamma-function, the correct Coulomb bound state poles. The partial wave projections of the amplitude(3) have the form

$$a_\ell(E)=\frac{1}{p}e^{i\delta_\ell}\sin\delta_\ell, \tag{4}$$

where $\delta_\ell=\arg\Gamma(\ell+1-i\eta)$, so that A(E,z) satisfies elastic unitarity for all $E>0$.

If we now consider the perturbation expansion of the Coulomb amplitude, we can write

$$A(E,z)\simeq A_{(1)}(E,z)+A_{(2)}(E,z)+\ldots, \tag{5}$$

where the Born term is given by $A_{(1)}(E,z)=-2m\alpha/t$ and

$$A_{(2)}(E,z)=\frac{(2m\alpha)^2}{4\pi p^2}\int_0^\infty\frac{dk}{k^2-p^2-i\varepsilon}\frac{1}{\sqrt{K}}\ln\frac{\gamma^2-z+\sqrt{K}}{\gamma^2-z-\sqrt{K}}. \tag{6}$$

In the integral, $K=(\gamma^2-z)^2-(\gamma^2-1)^2$, $\gamma=(k^2+p^2)/2kp$, and the angular integrations for the second Born

term have been carried out. We note the folowing: (1) The first Born term is identical to the lowest order term in the expansion of the exact Coulomb amplitude. (2) The integral (6) for the second Born term is infinite. However, if we consider $A_{(2)}(E,z)$ for a Yukawa potential, $\lambda e^{-\mu r}/r$, in the limit $\mu \to 0$, then the integral has a branch cut in z for real $z>1$. The discontinuity across this cut is given by

$$\mathrm{Im}_z A_{(2)}(E,z) = \frac{(2m\alpha)^2}{2p^2}\int_0^\infty \frac{dk}{k^2-p^2-i\varepsilon}\frac{1}{\sqrt{K}}. \tag{7}$$

This integral (7) exists and has a branch cut for real $E>0$. The discontinuity across this cut is just

$$\mathrm{Im}_E \mathrm{Im}_z A_{(2)}(E,z) = \rho^{(2)}_{E,z}(E,z) = \pi\eta(2m\alpha/t)\theta(E)\theta(z-1). \tag{8}$$

This expression (8) is identical to the second order term in the expansion of the exact dsf obtained from (3). Thus, although the perturbation series does not converge to the correct amplitude, the dsf which can be extracted from the perturbation series does correspond to the exact dsf, at least through second order in α. Since it can be shown[1] that each term in the perturbation expansion for a superposition of Yukawa potentials has the correct cut structure, it seems reasonable to assume that even though the perturbation series may not converge, the dsf obtained from a perturbation expansion may be correct to all orders. At least, we will make that assumption in what follows.

III. SPIN-0--SPIN-0 ELECTROMAGNETIC SCATTERING

We would now like to illustrate our procedure for the construction of self-consistent electromagnetic scattering amplitudes by a relatively simple example. We will consider the elastic scattering of two fictitious spinless particles which only interact electromagnetically. The kinematics are as in fig. (3). Particles 1 and 3 have mass m, particles 2 and 4 have mass μ. For spinless scattering there is a

single Lorentz covariant amplitude, $A(s,t,u)$, where $s=(k_1+k_2)^2$, $t=(k_1+k_3)^2$ and $u=(k_1+k_4)^2$ are the usual Mandelstam variables. $s+t+u=2m^2+2\mu^2$, so that only two are independent. Through terms of second order in α, unitarity is saturated by the five diagrams of fig. (4)--(5). Of these, the only diagrams which contribute to the dsf are the two photon intermediate state (4d) in the t-channel and the u-channel elastic unitarity (5). We will consider these diagrams in some detail shortly. The remaining diagrams are relatively simple. Diagram (4a) represents the one-photon intermediate state and is just the Born term. It can be written

$$A(s,t,u)_{Born}=\tfrac{1}{4}(s-u)(4\pi\alpha/t), \tag{9}$$

where the factor $\frac{1}{4}(s-u)$ is due to the spin of the exchanged photon. Diagrams (4b) and (4c) represent vacuum polarization and vertex corrections to the Born term. They are essentially equivalent to the Feynman diagrams of fig. (6) plus renormalization. Although the Feynman integrals are infrared divergent, the unitarity diagrams can be evaluated without the introduction of a cutoff if the photon pole terms are replaced by the appropriate generalizations of the Coulomb amplitude (3). This substitution, which will be discussed in detail when we evaluate the elastic unitarity contribution to the dsf, results in the following correction to the imaginary part of the amplitude in the t-channel:

$$\mathrm{Im}_t A(s,t,u)=\tfrac{1}{4}(s-u)\,\mathrm{Im}_t\{(4\pi\alpha/t)[\Gamma(t)-1]\}, \tag{10}$$

where $\Gamma(t)=1+\gamma(t,m^2)+\gamma(t,\mu^2)$, and

$$\gamma(t,M^2)=\frac{t}{\pi}\int_{4M^2}^{\infty}\frac{dx}{x(x-t)}\,\mathrm{Im}\gamma(x,M^2). \tag{11}$$

The imaginary part of the form factor can be written

$$\mathrm{Im}\gamma(t,M^2)=(\alpha/q_M W_t)[\tfrac{1}{2}(2M^2-t)\psi(2)-\frac{2}{3}q_M^4/t], \tag{12}$$

where $q_M=\frac{1}{2}(t-4M^2)^{1/2}$, $W_t=2E_t=\sqrt{t}$ and $\psi(z)$ is the digamma function. Note that since we are only dealing with the second order terms here, this result is consistent with the

conjecture that the radiative corrections to the leading term of the amplitude, excluding vacuum polarization, can be obtained by multiplying by the appropriate form factors.[2] However, only the leading term factors in this manner. We see finally, that the net result of the addition of diagrams (4b) and (4c) to (4a) is that the Born term is multiplied by the vertex function, $\Gamma(t)$.

A. Double Spectral Functions

The first non-vanishing contribution to the dsf that we will consider is due to the two photon intermediate state in the t-channel. We have

$$\mathrm{Im}_t A(s,t,u)_{2\gamma} = \tfrac{1}{2}\rho_{2\gamma} \sum_{\text{spins}} \int d\Omega' R(K'') R^{\dagger}(K'), \qquad (13)$$

where $\rho_{2\gamma}$ is the two photon phase space factor, the sum is over the polarizations of the intermediate state photons and $R(K)$ is the two photon annihilation amplitude in pole approximation. With kinematics as in fig. (7),

$$R(K') = M(s')(k_1\cdot\varepsilon_5)(k_3\cdot\varepsilon_6) + M(u')(k_1\cdot\varepsilon_6)(k_3\cdot\varepsilon_5) + 2\pi\alpha(\varepsilon_5\cdot\varepsilon_6), \qquad (14)$$

where $M(s') = 4\pi\alpha/(s'-m^2) = 4\pi\alpha/(-2k_1\cdot k_5)$, $M(u') = 4\pi\alpha/(u'-m^2) = 4\pi\alpha/(-2k_1\cdot k_6)$. [$R(K'')$ can be obtained from $R(K')$ by the substitutions, $1\leftrightarrow 4$, $3\leftrightarrow 2$.] This amplitude (14) is gauge invariant and satisfies the requirements of Bose statistics. Using the explicit form of $R(K)$ in the unitarity integral (13), we can do the spin sums and the angular integration. We find

$$\mathrm{Im}_t A(s,t,u)_{2\gamma} = \theta(t)\{\tfrac{1}{4}(u-m^2-\mu^2)^2 I_0(s,t,u) + \frac{1}{8}(t-2m^2) I_1(t) + \frac{1}{8}(t-2\mu^2) I_2(t) + \tfrac{1}{4} I_3(t)\} + \{s\leftrightarrow u\}, \qquad (15)$$

where $\theta(z)$ is the Heaviside function. The only term in (15) which contributes to the dsf is that proportional to $I_0(s,t,u)$, which is essentially the Mandelstam integral; i.e.,

$$I_o(s,t,u)=\frac{4\pi\alpha^2}{t(2q_uW_u)}\ln\frac{m^2+\mu^2-u+2q_uW_u}{m^2+\mu^2-u-2q_uW_u}, \tag{16}$$

where

$$2q_uW_u=\{[u-(m+\mu)^2][u-(m-\mu)^2]\}^{1/2}. \tag{17}$$

The cut in u of $I_o(s,t,u)$ extends from $(m+\mu)^2$ to infinity along the real u-axis, and the discontinuity is that of the logarithm. We find, then, that the two photon contribution to the dsf can be written

$$\begin{aligned}\rho_{tu}(s,t,u)_{2\gamma}&=\rho_{ts}(u,t,s)_{2\gamma}\\ &=\tfrac{1}{4}(u-m^2-\mu^2)^2\ \frac{8\pi^2\alpha^2}{t(2q_uW_u)}\,\theta(t)\theta[u-(m+\mu)^2].\end{aligned} \tag{18}$$

Equation (18) represents the inhomogeneous part of the dsf.

u-channel elastic unitarity [diagram (5)] accounts for the remainder of the second order dsf. We have

$$\mathrm{Im}_uA(s,t,u)_{\text{elastic}}=\tfrac{1}{2}\rho_2(u)\int d\Omega'A(s,t,u)A^*(s,t,u), \tag{19}$$

where $\rho_2(u)$ is the two body phase space factor. If we were to proceed in the usual manner of perturbation theory, we would insert the Born term into the right-hand side of (19) and use it to determine the second order contribution to $\mathrm{Im}_uA(s,t,u)$. However, the result would be infra-red divergent and a cutoff would be necessary. In order to circumvent this particular difficulty we will, instead, use an initial trial amplitude, $A_o(s,t,u)$, which is derived from the Born term (9) by the substitution

$$\frac{4\pi\alpha}{t}\ \rightarrow\ f_o(s,t,u)=\frac{-4\pi\alpha}{4q_u^2}\,\frac{\Gamma(1-i\eta(u))}{\Gamma(1+i\eta(u))}\left(\frac{-t}{4q_u^2}\right)^{-1+i\eta(u)}$$

$$+\left|\frac{\alpha}{2q_uW_u\eta(u)}\right|\ \frac{\delta(-t/4q_u^2)}{2\pi i\rho_2(u)}. \tag{20}$$

$f_0(s,t,u)$ has essentially the same form as the Coulomb amplitude (3), with the substitution of the correct relativistic two body kinematical factors. The trajectory function, $\eta(u)$, is arbitrary in (20); it will be determined by self-consistency. We note here that the amplitude, $A_0(s,t,u)$, which finally results can also be derived from a consideration of the infinite dimensional unitary representations of $O(4,2)$[3] and, in addition, a similar form can be extracted from the eikonal approximation.[4] Thus, the use of this amplitude as a basis for iteration is consistent with other, more familiar, treatments of electromagnetic scattering processes. Inserting the amplitude, $A_0(s,t,u)$ defined above, into the unitarity integral (19), we find

$$\mathrm{Im}_u A(s,t,u)_{\text{elastic}}$$

$$=\theta[u-(m+\mu)^2]\ \tfrac{1}{2}(u-m^2-\mu^2)\left(\frac{\alpha}{q_u W_u \eta(u)}\right)\mathrm{Im}_u\{A_0(s,t,u) \qquad (21)$$

$$+\tfrac{1}{4}(u-m^2-\mu^2)f_1(s,t,u)+\tfrac{1}{2}f_2(s,t,u)\},$$

where

$$f_1(s,t,u)=4\pi\alpha(u-m^2-\mu^2)^{-1}\,|\Gamma(1-i\eta)|^2[P_{i\eta}(-z_u)-1],$$

$$f_2(s,t,u)=4\pi\alpha q_u^2(u-m^2-\mu^2)^{-1}\,|\Gamma(1-i\eta)|^2[i\eta P_{i\eta}(-z_u)], \qquad (22)$$

$P_\nu(z)$ is a Legendre function of the first kind and $z_u=1+t/2q_u^2$. We note the following: (1) If $\eta(u)=\alpha(u-m^2-\mu^2)/2q_u W_u$, the original trial amplitude is reproduced in (21).[5] (2) For that value of $\eta(u)$, the $f_i(s,t,u)$ in (22) are of order α^2 and do not have any spurious poles. The elastic unitarity contribution to the dsf can be obtained relatively simply from (21) and (22) if we remember that $P_\nu(z)$ is analytic in the z-plane cut along the negative real axis from $-\infty$ to -1. The discontinuity across the cut is $-2i\sin\pi\nu P_\nu(-z)$. We find that to second order

$$\rho_{ut}(s,t,u)_{\text{elastic}}=\rho_{tu}(s,t,u)_{2\gamma},$$

$$\rho_{us}(s,t,u)_{elastic}=0. \tag{23}$$

The result (23) is independent of the actual value of $\eta(u)$ since the discontinuity in t of $\mathrm{Im}_u A(s,t,u)$ introduces an additional multiple of $\eta(u)$ which is cancelled by the leading factor of (21). Adding (18) and (23), the complete second order dsf associated with the unitarity diagrams (4d)--(5) can be written

$$\rho_{tu}(s,t,u)=2\rho_{ts}(u,t,s)=2\rho_{tu}(s,t,u)_{2\gamma}, \tag{24}$$

where $\rho_{tu}(s,t,u)_{2\gamma}$ is defined by (18).

B. Self-Consistent Scattering Amplitude

Using the information presented above [particularly eqs. (21), (23) and (24)], it is possible to construct an amplitude which has the dsf terms indicated in (24) and which also has the appropriate vacuum polarization and vertex corrections. We will present the amplitude as an ansatz; it is then trivial to verify that it has the requisite properties. We write

$$A(s,t,u)=2F(s,t,u)+F(u,t,s), \tag{25}$$

where

$$F(s,t,u)=F_0(s,t,u)+\tfrac{1}{4}(u-m^2-\mu^2)f_1(s,t,u)+\tfrac{1}{4}f_2(s,t,u), \tag{26}$$

and

$$F_0(s,t,u)=\tfrac{1}{4}(s-u)\Gamma(t)f_0(s,t,u)=\Gamma(t)A_0(s,t,u). \tag{27}$$

By construction, A(s,t,u) defined by (25) is cutoff independent. One can show explicitly[5] that the delta-function terms in (25) conspire to produce exactly the matrix elements, for each channel, of -iI, where I is the identity operator. Thus, the S-matrix corresponding to this transition amplitude is analytic. We note that (25) exhibits the proper dsf terms required by (24). If we set

$$\eta(u)=\alpha(u-m^2-\mu^2)/2q_u W_u, \qquad (28)$$

then (25) will be self-consistent. Proof: For this value of $\eta(u)$, $A_0(s,t,u)$ reproduces itself in the elastic unitarity integral (21). Since the additional terms in (25) are all of order α^2, their introduction into the unitarity can have no effect on the second order dsf. Thus, the reintroduction of (25) into the elastic unitarity integral will produce no additional second order contributions to the dsf. (We note that (25) has a well defined partial wave expansion so that it will admit a process of iteration.) We also note that A(s,t,u) displays Regge asymptotic behavior and the correct Coulomb bound state Regge poles. The spectrum generated by the trajectory function (28) is given by

$$E_B = -\tfrac{1}{2}\frac{\alpha^2\mu}{n^2} + \frac{11}{32}\frac{\alpha^4\mu}{n^4} + \ldots \quad (n=1,2,3\ldots) \qquad (29)$$

(μ is now the reduced mass). The second term represents the recoil corrections to the spectrum and was first calculated by Breit and Brown.[6] As an additional check on the correctness of our ansatz, we note that (25) reduces to the usual Born term in lowest order. As a matter of fact, since $(u-m^2-\mu^2)/2W_u$ is equal to the reduced mass at threshold, (25) actually reproduces the non-relativistic Coulomb amplitude at low energy. Finally, we point out that, in addition to the unitarity cuts which we have examined, A(s,t,u) exhibits a left-hand cut which is required by self-consistency. This cut is due to the factor $W_u=\sqrt{u}$ which appears in the trajectory function and which originates in the two body phase space factor. It is probably an inescapable feature of relativistic scattering. We note that both the Klein-Gordon and Dirac Coulomb scattering amplitudes have a cut structure of this type. We conclude that (25) represents a satisfactory electromagnetic scattering amplitude for two, spinless, nonidentical particles, and that it should be accurate through second order in α.

IV. ELECTRON-POSITRON SCATTERING

A self-consistent electron-positron scattering amplitude can be constructed using essentially the same procedure

as that outlined above. The complications introduced by spin are merely algebraic (but exceedingly tedious). Accordingly, we will only indicate the final result and then discuss some of its properties. The amplitude for the process described by fig. (8) can be written

$$T(k_4,k_2;k_1,k_3)=\sum_i A_i(s,t,u)Y_i(k_4,k_2;k_1,k_3)$$
$$+\tilde{A}_i(s,t,u)\tilde{Y}_i(k_4,k_2;k_1,k_3), \qquad (30)$$

where $\tilde{A}_i(s,t,u)=-A_i(s,u,t)$ (we set $m_e=1$, so that $s+t+u=4$) and $\tilde{Y}_i(K)$ is obtained from $Y_i(K)$ by the exchange $k_4 \leftrightarrow -k_3$. The $Y_i(K)$ (i=V,S,P,A,T) are essentially the five spinor basis functions used in the N-N problem by GGMW.[7] We have (in the t-channel)

$$Y_V \equiv V=\bar{u}(k_4)\gamma^\mu v(k_2)\bar{v}(k_3)\gamma_\mu u(k_1),$$
$$Y_S \equiv S=\bar{u}(k_4)v(k_2)\bar{v}(k_3)u(k_1),$$
$$Y_P \equiv P=\bar{u}(k_4)\gamma_5 v(k_2)\bar{v}(k_3)\gamma_5 u(k_1), \qquad (31)$$
$$Y_A \equiv A=\bar{u}(k_4)\gamma_5\gamma^\mu v(k_2)\bar{v}(k_3)\gamma_5\gamma_\mu u(k_1),$$
$$Y_T \equiv T=\bar{u}(k_4)i\sigma^{\mu\nu}v(k_2)\bar{v}(k_3)i\sigma_{\mu\nu}u(k_1),$$

where the spinor normalization is that of Bjorken and Drell.[8] The invariant amplitudes can be written

$$A_i(s,t,u)=2F_i(s,t,u)+(-1)^i F_i(u,t,s), \qquad (32)$$

where

$$(-1)^i = \begin{cases} -1, & i=V,T \\ +1, & i=S,P,A \end{cases} \qquad (33)$$

and

$$F_i(s,t,u)=F_i^o(s,t,u)+\sum_{j=1}^{3} P_j^i(s,t,u)f_j(s,t,u). \qquad (34)$$

In eq. (34), the $P^i_j(s,t,u)$ are simply ratios of polynomials; their exact form is given in ref. (5). The $f_j(s,t,u)$ are linear combinations of Legendre functions. We find

$$f_1(s,t,u) = -4\pi\alpha(u-2)^{-1}\,|\Gamma(1-i\eta)|^2[P_{i\eta}(-z_u)-1],$$

$$f_2(s,t,u) = 4\pi\alpha(4q_u^2)(u-2)^{-1}\,|\Gamma(1-i\eta)|^2\{(1+i\eta)[P_{i\eta}(-z_u) + P_{1+i\eta}(-z_u)]-\tfrac{1}{2}(1-z_u)\}, \tag{35}$$

$$f_3(s,t,u) = -4\pi\alpha(4q_u^2)(u-2)^{-1}\,|\Gamma(1-i\eta)|^2\, i\eta P_{i\eta}(-z_u),$$

where $z_u = 1+t/2q_u^2$. The $F^o_j(s,t,u)$ are the invariant amplitudes of our original lowest order approximation to the electron-positron amplitude, $T_o(K)$, with the appropriate vacuum polarization and vertex corrections. Our trial scattering amplitude, which we inserted into the elastic unitarity to generate the higher order corrections, is obtained from the usual Born term by the substitution

$$(4\pi\alpha/t) \to f_o(s,t,u), \tag{36}$$

where $f_o(s,t,u)$ is defined in (20). For electron-positron scattering, the trajectory function which is required by self-consistency is given by

$$\eta(u) = \alpha(u-2)/2q_uW_u, \tag{37}$$

which gives the correct positronium bound state spectrum, including reduced mass and recoil corrections, and the appropriate Regge asymptotic behavior. By construction, the amplitude (30) exhibits the correct second order dsf and reduces to the Born term in lowest order. Moreover, it is cut-off independent, analytic and has a well defined Jacob-Wick expansion. We also note that $T_o(K)$ has been applied to calculations of the Lamb shift in hydrogen[9] and the electron anomalous magnetic moment,[10] so that the second order amplitude (30) should prove to be of considerable theoretical utility. However, its relation to the actual physical problem can only be resolved by a confrontation with experiment. So far, the perturbation calculation of the electron-positron elastic cross section has been able to account for the

experimental situation,[11] although the tests have not been exhaustive. Since there are significant differences between our self-consistent amplitude and the one obtained from perturbation theory, further experiments would seem appropriate.

V. "DUALITY" AND ELECTROMAGNETIC SCATTERING

I should point out immediately that the "duality" which appears in the title of this lecture is not necessarily identical to that which appears in any of the other lectures presented this summer. However, our self-consistent electron-positron scattering amplitude does appear to exhibit a symmetry which is as close to the usual form of duality as may be possible outside the field of hadron physics. At first, there does not seem to be room for duality of any kind. The electron-positron amplitude describes fermion-antifermion scattering. The t- and u-channels are identical so that interchanging t and u only introduces a minus sign. This is the Pauli principle and has nothing to do with duality. The s-channel describes electron elastic scattering and thus is "exotic"; there are no bound states of the two electron system. However, the electron-electron scattering amplitude does exhibit poles at the positions associated with the usual Coulomb bound states. This circumstance is analogous to that which obtains in the Schrodinger amplitude for a repulsive $1/r$ potential. In that case, the amplitude has the form

$$A(E,z)=\frac{-\alpha m}{2p^2}\,\frac{\Gamma(1+i\eta)}{\Gamma(1-i\eta)}\left(\frac{-t}{4p^2}\right)^{-1-i\eta}-\frac{1}{ip}\,\delta(1-z). \qquad (38)$$

We see that this amplitude differs from (3) by the fact that the trajectory is replaced by its complex conjugate. Thus, for a repulsive potential, the poles are on the second sheet. In our electron-positron amplitude, these second sheet poles appear explicitly due to the permutation of the Mandelstam variables in (32). If $F(s,t,u)$ has poles in u, then $F(u,t,s)$ will have the corresponding poles in s. We find, then, that the amplitude (30) has poles in t and u which are due to the positronium bound states, and poles in s which correspond to the second sheet poles of the electron-electron scattering amplitude. The "dual" nature of the amplitude arises from the following consideration. If we use the fact that the trajectory function, $\eta(x)$, is symmetric about $x=2$ (it is

necessary to pay attention to phases), then it is possible to show that

$$F(u,t,s)=e^{i\phi}F(s,t,u)+B(s,t,u), \tag{39}$$

where $B(s,t,u)$ is a possible background term and $F(s,t,u)$ is given by (34). We note the following: $B(s,t,u)$ has no poles--only the branch cut singularities of the amplitude--and is at most $0(\alpha)$ with respect to $F(s,t,u)$. (Actually, it is possible to argue another way and show that $\phi=B=0$ for our second order amplitude.) The point we wish to make is that eq. (39) represents a limited permutation symmetry of the electron-positron amplitude. Using this property, it is possible to write the amplitude as a function which has only the physical positronium poles (and corresponding asymptotic behavior), or as a function which exhibits only the second sheet electron-electron poles, and without sacrificing any of the other properties which we require of the amplitude. It is this behavior which we call "duality" in the case of electromagnetic scattering amplitudes. (We also find that the amplitude for spinless non-identical particles possess a similar symmetry.) The actual significance of this result for the strong interactions is probably nil. Nevertheless, the fact that unphysical sheet Regge poles seem to be of an importance in electromagnetic scattering which equals that of the physical sheet singularities is interesting. Also, the existence of a symmetry in electromagnetic scattering which is not simply a result of crossing may be of some value. At least, the problems raised here may prove amusing to contemplate.

VI. CONCLUDING REMARKS

Using virtually the same procedure as that which was outlined in the original hadron bootstrap programs, we have constructed self-consistent electromagnetic scattering amplitudes having most of the properties one assumes would be found in a result of a "future correct theory." Moreover, the procedure is simple enough that it may supplant perturbation theory for certain purposes--particularly the construction of electromagnetic scattering amplitudes for processes in which there are bound states and, possibly, the calculation of radiative corrections to bound state energy levels. The question naturally arises--of what relevance are our results to hadron physics? One simple possibility

is that a solution to the Schrodinger equation with a Yukawa potential may be found. In that case, one could attempt to do nucleon-nucleon elastic scattering in essentially the same manner as electron-positron scattering; hopefully, the results would be equally agreeable. Failing a breakthrough in potential theory, one could certainly use the eikonal approximations to the Yukawa amplitude as a basis for iteration. The problem of convergence would undoubtedly be more difficult, but it is certainly worth investigating. In any case, we feel that the procedure outlined above will prove to be of some value within the realm of electromagnetic interactions. It remains to be seen if this work can be extended into a larger domain.

FIGURE CAPTIONS

Fig. (1). Electron-positron intermediate state contribution to the electron form factors.

Fig. (2). Unitarity diagrams which contribute to electromagnetic scattering amplitudes through second order in α. Diagrams (a) and (b) (and all higher order diagrams) constitute the inhomogeneous contributions to the amplitude. Diagram (c) is the elastic unitarity contribution.

Fig. (3). Diagram of the scattering process. Unbroken line indicates mass, m, dashed line, mass, μ.

Fig. (4). t-channel unitarity through second order introduces the following contributions to the imaginary part of the scattering amplitude: (a) one photon exchange (b) two boson exchange (mass m) (c) two boson exchange (mass μ) (d) two photon exchange.

Fig. (5). u-channel elastic unitarity.

Fig. (6). Feynman diagrams which contribute to the spin-0 electromagnetic form factor. Diagrams (a) and (c) represent vacuum polarization contributions, diagrams (b) and (d) are vertex corrections.

Fig. (7). Two photon annihilation amplitude (t-channel).

Fig. (8). Diagram for electron-positron scattering. The s-channel is appropriate for electron-electron scattering.

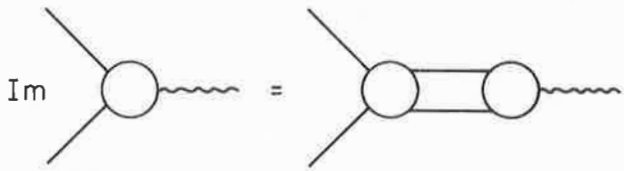


Figure 1

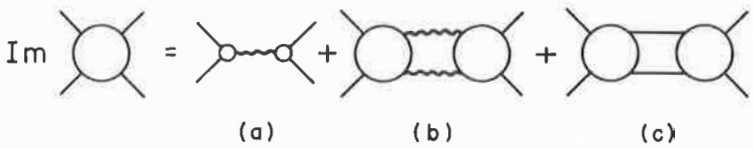


Figure 2

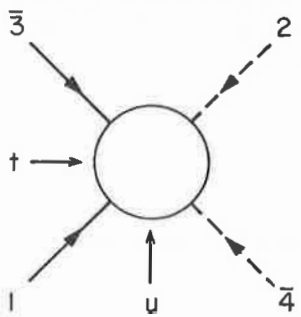


Figure 3

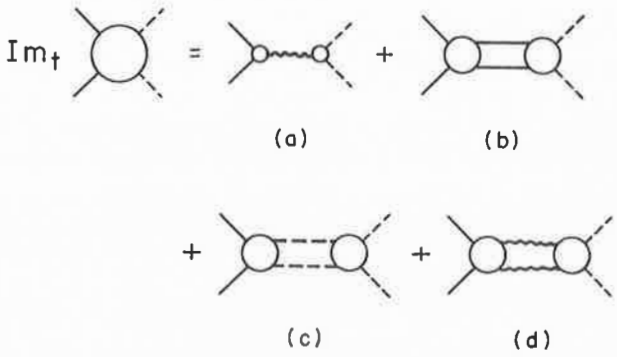


Figure 4

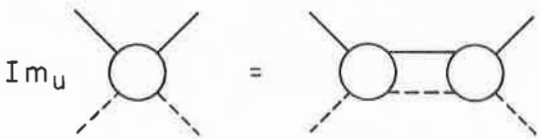


Figure 5

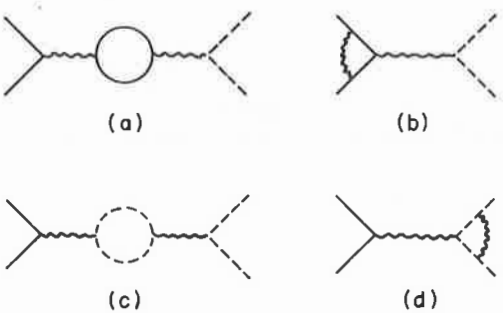


Figure 6

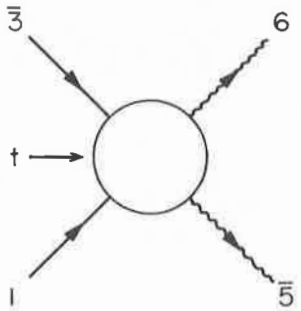


Figure 7

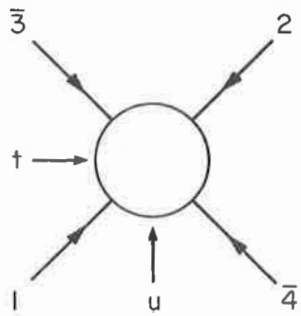


Figure 8

REFERENCES

1. Blankenbecler, R., M. L. Goldberger, N. N. Khuri and S. B. Treiman, Annals of Physics 10, 62 (1960).

2. Yao, Y. -P., Phys. Rev. D 1, 2316 (1970).

3. Barut, A. O., and W. Rasmussen, Phys. Rev. D 3, 956 (1971).

4. Lévy, M., and J. Sucher, Phys. Rev. D 2, 1716 (1970).

5. McEnnan, J., Phys. Rev. D 3, 1935 (1971).

6. Breit, G., and G. E. Brown, Phys. Rev. 74, 1278 (1948).

7. Goldberger, M. L., M. T. Grisaru, S. W. MacDowell and D. Y. Wong, Phys. Rev. 120, 2250 (1960).

8. Bjorken, J. D., and S. D. Drell, Relativistic Quantum Mechanics (McGraw-Hill, New York 1964).

9. McEnnan, J., Phys. Rev. 181, 1967 (1969).

10. McEnnan, J., (unpublished).

11. Hearn, A. C., P. K. Kuo and D. R. Yennie, Phys. Rev 187, 1950 (1969).

DUAL LOOPS AND MULTIPLY PERIODIC FUNCTIONS

Stanley Fenster
Argonne National Laboratory
Argonne, Illinois 60439

The implementation of unitarity in the factorized dual resonance model requires the addition of dual loops to dual trees. The expression for a single dual loop contains a doubly periodic function and the general theory of M loops brings in multiply-periodic functions of a complex variable. These functions are connected with automorphic functions for Schottky and Fuchsian groups. The mathematical development goes back to Riemann and involved all the great names in mathematics of the last half of the nineteenth century. The related physics problem is the calculation of the electrostatic potential in a two-dimensional surface with any number of conductors embedded. The mathematics teaches us how to generalize the method of images which works for the case of one conductor to the case of many conductors. The analogue model of Nielsen makes the connection with the dual resonance model.

The way automorphic functions arise in the dual resonance model has been explained in several papers. These lectures constitute a supplementation, taking more time on certain topics, such as Abelian integrals, than would have been appropriate in these papers.

We study domains of the complex plane $z=x+iy$ and harmonic functions $u(x,y)$ defined on them. In general these domains D have boundaries which may be simply visualized as circles whose insides are exterior to D. A crosscut of a domain is an arc which, apart from its endpoints, lies entirely in the domain; the end points of the crosscut coincide with boundary points of the domain. A simply-connected domain D is defined by the property that all points in the interior of any closed curve (which doesn't

cross itself) which consists of points of D are also points of D. A crosscut connecting two different boundary points of a simply-connected domain D divides D into two simply-connected domains without common points. If there exist closed curves which consist entirely of points of D and in whose interior there are points not belonging to D, then D is multiply-connected. D is said to be of connectivity n if there exist no more than n-1 such curves whose interiors have no points in common and which contain in their interiors points not belonging to D. The circular ring $1< x^2 + y^2 < 4$ is of connectivity two (doubly connected). A bounded domain of connectivity n is a bounded domain with n - 1 "holes". By n - 1 appropriately chosen crosscuts, a domain of connectivity n can be transformed into a simply-connected domain. For example, the crosscut $1 < x < 2$, $y = 0$ transforms the circular ring $1 < x^2 + y^2 < 4$ into a simply-connected domain. The restriction to bounded domains is removed by introducing the point at infinity in the usual way.

Harmonic functions u(x,y) may be defined on a domain D. They satisfy the Laplace equation $\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}\right)u(x,y)=0$. There are two significant formulae that hold for more general functions u and v; these formulae become particularly simple and useful when u and v are harmonic. They are called Green's first and second formula:

$$\text{(I)}\quad \iint_D u\Delta v\,dx\,dy + \iint_D \left(\frac{\partial u}{\partial x}\cdot\frac{\partial v}{\partial x} + \frac{\partial u}{\partial y}\,\frac{\partial v}{\partial y}\right)dxdy = \int_\Gamma u\,\frac{\partial v}{\partial n}\,ds$$

$$\text{(II)}.\quad \iint_D (u\Delta v - v\Delta u)\,dx\,dy = \int_\Gamma \left(u\,\frac{\partial v}{\partial n} - v\,\frac{\partial u}{\partial n}\right)ds.$$

The curve Γ is the boundary of D; it may consist of several separate closed curves. The derivative $\partial u/\partial n$ is the usual normal derivative. If u and v are harmonic, the formulas simplify because $\Delta u = 0$, $\Delta v = 0$. These theorems enable one to solve boundary value problems of the first and second kind and suggest the introduction of Green's function and Neumann's function. Let w(x,y) be harmonic in D and let $r = \sqrt{(x-\xi)^2 + (y-\eta)^2}$; then log r is harmonic in D + Γ except at (ξ,η) and the same is true of the

function $h(x,y) = -\log r + w(x,y)$. The first result is derived from the application of formula (II) to such a function. We find that if $\Delta u = 0$ in D and has no singular behavior, then

$$\text{(III)} \quad u(\xi,\eta) = \frac{-1}{2\pi}\int_\Gamma \left(u\frac{\partial h}{\partial n} - h\frac{\partial u}{\partial n}\right) ds.$$

The function h suggests that we define the Green's function $g(x,y;\xi,\eta)$ of a domain D as

$$g(x,y;\xi,\eta) = -\log\sqrt{(x-\xi)^2 + (y-\eta)^2} + g_1(x,y;\xi,\eta),$$

where g_1 is harmonic in D; if (x,y) tends to <u>any point</u> of the <u>boundary</u> Γ of D, g tends to zero. Then

$$u(\xi,\eta) = -\frac{1}{2\pi}\int_\Gamma u\frac{\partial g}{\partial n}\, ds.$$

Two important properties of g are that it is positive in D and symmetric under the interchange $(x,y) \leftrightarrow (\xi,\eta)$ of its arguments. As an example, one may choose D to be the circle C_R of radius R with center at the origin. If r,θ are polar coordinates in the (x,y) plane and ρ,φ the polar coordinates of a point ζ inside C_R $(\rho < R)$, then

$$g(z,\zeta) = \log\frac{\rho\cdot r_2}{R\cdot r_1},$$

$$r_1^2 = r^2 - 2\rho r\cos(\theta - \varphi) + \rho^2$$

$$r_2^2 = \left(\frac{R^2}{\rho}\right)^2 - 2\left(\frac{R^2}{\rho}\right)\cdot r\cos(\theta - \varphi) + r^2.$$

Green's third formula (III) contains a symmetry between u and $\partial u/\partial n$ on Γ. Consequently one is led to define the Neumann function of D with respect to ζ as

$$N(z,\zeta) = -\log r + N_1(z,\zeta)$$

with the condition $g = 0$ on Γ replaced by

$$\frac{\partial N}{\partial n} = \text{const}, \qquad z \in \Gamma$$

and with normalization

$$\int_\Gamma N(z,\zeta)\, ds_z = 0.$$

Then, if $\int_\Gamma N(z)\, ds = 0$, we can show that

$$u(\zeta) = \frac{1}{2\pi} \int_\Gamma N(z,\zeta) \frac{\partial u}{\partial n}\, ds.$$

For the domain C_R we find

$$N(z,\zeta) = \log \frac{R^2}{\rho^2 r_1 r_2}.$$

The distinctive properties of multiply-connected domains is not evidenced in the foregoing material. At this point we introduce the conjugate harmonic function as a step toward those properties. If the function $u(x,y)$ is harmonic in D, we can associate with it another function $v(x,y)$ by means of the Cauchy-Riemann equations

$$\frac{\partial v}{\partial x} = -\frac{\partial u}{\partial y}, \quad \frac{\partial v}{\partial y} = \frac{\partial u}{\partial x}.$$

Then $\Delta v = 0$ and v is called the harmonic conjugate of u. If a curve C connects two points (x,y) and (x_0,y_0) we can show that

$$v(x,y) - v(x_0,y_0) = \int_C \left(-\frac{\partial u}{\partial y}\, dx + \frac{\partial u}{\partial x}\, dy\right).$$

This line integral is independent of the curve C; its value depends only on the terminals (x,y) and (x_0,y_0). It is important to note that the conjugate harmonic function is only determined up to an arbitrary additive constant which plays the part of an integration constant and was expressed as the value $v(x_0,y_0)$.

In the results concerning harmonic functions so far, no reference was made to the connectivity of the domains D involved. There are certain properties of harmonic conjugates, however, which are decisively influenced by the connectivity of D. Consider the simply-connected case. The line integral formula for v then shows that v is a

single-valued function in D, that is, it has a uniquely defined value at each point of D (provided $v(x_0,y_0)$ is held fixed). This follows from the independence from C. Since in a simply-connected domain any two curves connecting the same two points can be deformed into each other, all curves C are included and v is unique. Suppose now we do not require $\Delta u = 0$ everywhere in D. Let C_1 and C_2 be two curves giving the values v_1 and v_2 to the line integral. We can prove that

$$v_1 - v_2 = \iint_B \Delta u \, dx \, dy,$$

where B is enclosed by C_1 and C_2. If D is multiply-connected it has at least one hole. By properly choosing C_1 and C_2 to go from (x_0,y_0) to (x,y) along different shores of the hole one can be assured that $\Delta u \neq 0$ inside B, namely inside the hole. Then $v_1 \neq v_2$ and v is multiple-valued. This property depends on the fact that C_1 cannot be continuously deformed into C_2 so that all the intermediate curves lie completely in D. (The interior of the hole is considered to be outside of D). Green's formula shows that

$$v_1 - v_2 = \iint_B \Delta u \, dxdy = \int_b \frac{\partial u}{\partial n} \, ds \equiv p,$$

where b is a closed curve surrounding the hole and $p \neq 0$. We now ask: what are the possible different values v can take at the same point? Since the only way to change v is to integrate an extra time completely around the hole, we find

$$v_1 = v_2 + mp, \qquad m = 0, \pm 1, \pm 2, \ldots$$

The various values which the harmonic conjugate v of a single-valued harmonic function u can take at one point differ by integral multiples of a number p, the modulus of periodicy of v. For example, consider the conjugate of $u = \log r$, $r = \sqrt{x^2 + y^2}$; it is $v = \theta = \tan^{-1}(\frac{y}{x})$. Both functions are harmonic in the ring $1 < x^2 + y^2 < 4$, since the only singularity is at $(x,y) = (0,0)$. Now θ is not

single-valued in the ring; indeed, the angle is only determined up to an integral multiple of 2π, and a closed positive circuit about the origin adds 2π to θ. The function $v = \theta$ thus has $p = 2\pi$.

Turning now to a domain D of connectivity n, we see that the different values of v defined by

$$v(x,y) = v(x_0,y_0) + \int_{s_0}^{s} \frac{\partial u}{\partial n}\, ds$$

depend on the number of times each of the n-1 holes is surrounded by the integration path. There will therefore be n-1 independent moduli of v, corresponding to a complete circuit about each of these holes. Attaching to these holes the subscripts $1,2,\ldots,n-1$, we thus have the independent moduli $p_1,p_2,\ldots,p_{n-1}$, where the word "independent" means that none of these moduli is, in general, expressible as the sum of integral multiples of the other moduli. If b_ν is a curve which surrounds the hole of subscript ν but does not contain in its interior points of any other hole, we have

$$p_\nu = \int_{b_\nu} \frac{\partial u}{\partial n}\, ds, \quad \nu = 1,2,\ldots,n-1.$$

Since, apart from continuous deformations inside D, the various integration paths connecting two points of D differ only by the number of times they surround the various holes, we thus obtain the result that the various values of v(x,y) are of the form

$$v(x,y) = v(x_0,y_0) + \sum_{\nu=1}^{n-1} m_\nu p_\nu,$$

where the m_ν are arbitrary integers and $v(x_0,y_0)$ is any one of the values of v(x,y). If all moduli p_ν vanish, v will be single-valued in D. As a final point, one may call the outer boundary of D the n^{th} circle. Then a modulus p_n is defined, and Green's formula gives

$$\sum_{\nu=1}^{n} p_\nu = 0.$$

An important function defined on the domain D with n

holes is the harmonic measure. Let the n bounding holes be denoted by Γ_ν, $\nu = 1,2,\ldots,n$. The harmonic measure $w_\nu(z)$ is harmonic in D and has the boundary values unity on Γ_ν and zero on Γ_μ, $\mu \neq \nu$. This function is given by

$$w_\nu(\zeta) = -\frac{1}{2\pi}\int_{\Gamma_\nu} \frac{\partial g(z,\zeta)}{\partial n_z}\, ds_z$$

and is equal to $-\frac{1}{2\pi}$ times the modulus p_ν of the Green's function $g(z,\zeta)$. Since the Green's function has a singularity at the point $z=\zeta$ we do not find the sum of the moduli to be zero, but rather

$$\sum_{\nu=1}^{n} w_\nu(z) = 1$$

for all values of z in D.

We may take a further step and find the moduli of the set of $w_\nu(z)$. They are given by

$$p_{\mu\nu} = \int_{\Gamma_\mu} \frac{\partial w_\nu}{\partial n}\, ds$$

and, as can be shown with Green's formula, are symmetric. Furthermore the same method shows that the symmetric quadratic form

$$\sum_{\mu=1}^{n-1} \sum_{\nu=1}^{n-1} p_{\nu\mu}\, \lambda_\nu \lambda_\mu$$

always positive-definite. Therefore the set of linear equations

$$\sum_{\mu=1}^{n-1} p_{\nu\mu}\, \lambda_\mu = \kappa_\nu$$

always has a non-trivial solution. As an example, one may show that the harmonic measures associated with the outer and inner boundaries of the ring $a^2 < x^2+y^2 < b^2$ are

$$w_1(z) = \frac{\log(r/a)}{\log(b/a)},$$

$$w_2(z) = \frac{\log(r/a)}{\log(a/b)},$$

$$r = \sqrt{x^2 + y^2}.$$

Another result, relevant for the calculation of asymptotic forms of loop integrals, is the Hadamard formula for the variation of the Green's function with a small variation in boundary. Suppose Γ is the boundary of D, and Γ^* is a new boundary obtained with a normal variation $\delta n(s)$. Let g* be the new Green's function; then

$$g^*(z,\zeta)-g(z,\zeta) = \frac{1}{2\pi}\int_\Gamma \frac{\partial g(\eta,z)}{\partial n_\eta} \cdot \frac{\partial g(\eta,\zeta)}{\partial n_\eta}\, \delta n(s) ds_\eta.$$

Similarly,

$$\delta w_\nu(z) = \frac{1}{2\pi}\int_\Gamma \frac{\partial w_\nu(\eta)}{\partial n_\eta} \cdot \frac{\partial g(\eta,z)}{\partial n_\eta}\, \delta n(s)\, ds_\eta$$

$$\delta p_{\nu\mu} = -\int_\Gamma \frac{\partial w_\nu(\eta)}{\partial n_\eta} \frac{\partial w_\mu(\eta)}{\partial n_\eta}\, \delta n(s)\, ds_\eta$$

As a step toward finding harmonic measures let us examine harmonic conjugate functions in D and note a suggestive property. Let D contain two circles C_1 and C_2. Join C_1 and C_2 with a curve L. Consider a harmonic conjugate function v evaluated at P (Figure 1). This value is not unique but suppose we choose one definite admissible value $v(P_1)$. If we follow v as it varies along a curve Γ_1 surrounding C_1 we find that it varies continuously up to the point P_2 where it has the value $v(P_2)$ with

$$v(P_2) = v(P_1) + \int \frac{\partial u}{\partial n} ds$$
$$= v(P_1) + p_L .$$

Here P_2 is infinitesimally distant from P_1, for we have split the line L into L_1 and L_2 which play the role of

the upper and lower lips of a branch cut. Now $v(Q) \neq v(P)$ but the foregoing arguments showed that $v(Q_2)=v(Q_1)+ p$ so that the difference $\delta v_L = v_1 - v_2$ is constant along L. Let us for the moment ignore the closeness of L_1 to L_2 and make a topological deformation of the closed curve $C_1 L_1 C_2 L_2$ so that L_1 and L_2 form circles and C_1 and C_2 are close and parallel lines. (Fig. 3). It is a remarkable fact that a class of harmonic conjugates v exists with the property that points such as R_1, R_2 may be found which are identified, in the sense that $\delta v_C = v(R_2) - v(R_1) = p_C = v(S_2) - v(S_1)$, where p_C is a modulus associated with traverses about L_1. Furthermore, there exists a δv which has both the properties of δv_L and δv_C. All harmonic conjugates in D have the (L) property; but very few have the (C) property as well; those with both properties are similar to the harmonic measures w. Proof of their existence is a major mathematical achievement and carries the names of Riemann, Schwarz and Neumann. One approach depends on the knowledge that harmonic functions are electrostatic potentials and that the infinite plane may be visualized as the surface of a sphere (we have gone to three-dimensional space). The sphere has two holes C_1 and C_2 cut out. The topological deformation joining C_1 to C_2 is equivalent to pushing out the rims of the holes and joining rim to rim to form a torus. Then C_1 is identified with C_2 and pairs of points (R_1, R_2) and (S_1, S_2) fuse to form R and S on the torus surface. The curves L and C are then closed curves on the surface which intersect each other once. If the torus is visualized as a rubber tire lying on the ground, then L runs around the top and C loops through the tire. Now it is known that a potential without sources can exist on a torus - the potential lines follow L and the field lines follow C. Knowing this we may reverse the process and cut the tire open along C and expect to find harmonic conjugates δv described above to exist.

Our task then, is to find real harmonic functions which identify the points on C_1 in a one-to-one manner with the points on C_2 and have the property that the difference $v_1 - v_2$ is a constant (the same for $P_1 P_2$ and $Q_1 Q_2$) around the circles. These functions will contain parameters which depend on the coordinates of the centers of the circles and their radii. Since the torus generalizes to a sphere with n handles it can be proved that 2n circles (n=1,2,...) may be cut out without losing the possibility

of a solution. The circles will be identified in pairs by the harmonic function. The curves L and C are called crosscuts on the surface and are denoted by $K_{2\mu-1}$ and $K_{2\mu}$, $\mu = 1,2,\ldots,n$.

Since the values of a harmonic function inside a closed curve are determined by its values on the boundary, we expect that the harmonic function for D will be determined by the 2n real moduli p for the 2n crosscuts K. There are two solutions on the torus, (n=2), the L-type potentials and the C-type, having orthogonal equipotential line families. The question arises as to whether there are any more, and a theorem states that there are 2n and no more that are linearly independent of each other. Every other such function can be expressed, up to an additive constant, as a linear combination of multiples of these functions with constant coefficients. Note that up to now the potentials considered have no singularities; they are called functions of the first kind, and their only discontinuities are the jumps across the crosscuts. In the future we will assume that each crosscut has two lips and that all functions considered have jump discontinuities across these cuts which are constant as one moves along the cut. Note that the cut surface is simply-connected.

Having proved the existence of real functions u of the first kind we may form complex functions $w = u + iv$ out of two of them. The complex functions w will be analytic functions if u and v satisfy a pair of Cauchy-Riemann equations. Some theorems can be proved about the moduli of w. We will perform integrations over a two-dimensional surface embedded in three-dimensional space using crosscuts K. We can always recover the complex plane with holes by cutting, since we are not dealing with arbitrary harmonic functions, but only with functions of the first kind. The μ^{th} pair of crosscuts has an odd member $K_{2\mu-1}$ which we denote by a_r and an even member b_r. Their crossing is indicated — in Fig. 4. as the crossing of two strips. The curves a_r and b_r are closed on a portion of the surface not shown. According to Green's formula we have

$$\iint_D dxdy \left(\frac{\partial u}{\partial x}\frac{\partial v}{\partial y} - \frac{\partial v}{\partial x}\frac{\partial u}{\partial y}\right) = \iint_D \left[\left(\frac{\partial v}{\partial x}\right)^2 + \left(\frac{\partial v}{\partial y}\right)^2\right] dxdy = \int_\Gamma u\, dv$$

where Γ is the boundary of D considered to be composed of a single curve, composed of both edges of each of the cross-cuts; the positive directions of integration are shown. The positive side of a_r is FD and the negative side is CE, etc. Each crosscut a_r or b_r has associated with it a constant modulus A_r^u or B_r^u with respect to $u(x,y)$ and A_r^v or B_r^v with respect to $v(x,y)$. Then we have

$$u_C - u_E = u_D - u_F = B_r^u$$

$$v_C - v_E = v_D - v_F = B_r^v$$

$$u_F - u_E = u_D - u_C = A_r^u$$

$$v_F - v_E = v_D - v_C = A_r^u \ .$$

Consider the integral $\int u\, dv$ taken along the two edges of the cross-cut a_r: let u_- and u_+ denote the functions along the negative and the positive edges so that $u_+ - u_- = A_r^u$. The value of the integral for the two edges is

$$\int_F^D u_+ \, dv + \int_C^E u_- \, dv = \int_F^D (u_+ - u_-) dv$$

$$= A_r^u \int_F^D dv = A_r^u (v_D - v_F) = A_r^u B_r^v \ .$$

Similarly, when the value of the integral for the two edges of the crosscut b_r is taken, we have

$$\int_D^C u_+ \, dv + \int_E^F u_- \, dv = \int_D^C (u_+ - u_-) \, dv$$

$$= B_r^u \int_D^C dv = B_r^u (v_C - v_D) = -B_r^u A_r^v \ .$$

Summing for the whole boundary of the resolved surface, we have

$$\int udv = \sum_{1}^{n} (A_r^u B_r^v - A_r^v B_r^u)$$

and therefore

$$\sum_{1}^{n} (A_r^u B_r^v - A_r^v B_r^u) = \iint \left[\left(\frac{\partial v}{\partial x}\right)^2 + \left(\frac{\partial v}{\partial y}\right)^2 \right] dxdy.$$

The double integral is always positive so the sum is always positive. Recall that w=u+iv so that the complex moduli of w are given by $A_r^u + iA_r^v$ and $B_r^u + iB_r^v$. We see that an analytic function w(z) of the first kind cannot have its moduli of periodicy for all a_r equal to zero. Second, the function w cannot have its moduli for b_r all zero; it cannot have its moduli all purely real, or all purely imaginary. Any function of the first kind has the form

$$w + \sum_{1}^{n} m_r \left(A_r^u + iA_r^v\right) + \sum_{1}^{n} n_r \left(B_r^u + iB_r^v\right)$$

where m_r and n_r are integers and w is a value, at any point on the surface, of a function of the first kind with the same moduli.

It has been mentioned that there are n linearly independent functions w_s of the first kind, s = 1,2,...,n. Their moduli are

$$A_{sr}^u + iA_{sr}^v \quad \text{and} \quad B_{sr}^u + iB_{sr}^v.$$

We can form a 2n × 2n matrix

$$M = \begin{pmatrix} A^u & A^v \\ B^u & B^v \end{pmatrix}$$

out of them. Using the fact that a function of the first kind with all zero real periods is a constant we may show that the determinant of M is non-zero. Another determinant which cannot vanish is $|A^u + iA^v|$. Unlike M it is composed of complex numbers and does not contain the B moduli. The non-singular property of these matrices allows a choice of normal forms for the w_s. Starting with

any linearly independent set we form linear combinations such that $A^u_{rs} = 0$ and $A^v_{rs} = i\pi\delta_{rs}$. These functions $w_s^{(N)}$ are called normal functions of the first kind. The quantities B are then not completely independent of each other. To show this we return to the integral

$$\iint_D dxdy \left(\frac{\partial w_1}{\partial x} \frac{\partial w_2}{\partial y} - \frac{\partial w_2}{\partial x} \frac{\partial w_1}{\partial y} \right) = \int_\Gamma w_1 \, dw_2$$

$$= \sum_1^n \left(A_{r_1} B_{r_2} - A_{r_2} B_{r_1} \right),$$

which is valid for all complex $w_1(z), w_2(z)$ which are analytic with singularities on the surface. If $w_1(z)$ and $w_2(z)$ are first kind functions they have no singularlities and everywhere in D we have

$$i \frac{\partial w_1}{\partial x} = \frac{\partial w_1}{\partial y} \quad \text{and} \quad i \frac{\partial w_2}{\partial x} = \frac{\partial w_2}{\partial y} ;$$

therefore $\sum_1^n \left(A_{r_1} B_{r_2} - A_{r_2} B_{r_1} \right) = 0$. For the $w_s^{(N)}$ we find then $B_{rs} = B_{sr}$. Furthermore the demonstrated positivity of

$$\sum_1^n \left(A_r{}^u B_r{}^v - A_r{}^v B_r{}^u \right)$$

for each $w_s = u_s + iv_s$ implies the positivity of the matrix B_{rs}.

We now introduce singularities which are logarithmic and define functions of the third kind. (We omit the second kind, which have poles and are obtainable by differentiation of functions of the third kind.) Green's functions and Neumann functions are of this type. Let the logarithmic singularities be at P_1 and P_2 with coefficients chosen unity. This process introduces very little that is new, for suppose we draw tiny circles C_1 and C_2 about P_1 and P_2 and join them by a crosscut (branch cut); the modulus of periodicy for this crosscut is just $2\pi i$. If the other

moduli are chosen exactly as they were for the $w_s^{(N)}$, we have $\Omega^{(E)}$, an elementary function of the third kind. We may add linear combinations of the $w_s^{(N)}$ to $\Omega^{(E)}$ to form $\Omega^{(N)}$, the normal function of the third kind, which has zero A-moduli. This is the so-called complex normalization. An alternative normalization is possible, the real normalization, in which the A^u and B^u are all zero, i.e. all moduli are pure imaginary. Note that in both of these normalizations there remain 2n real free constants; they are distributed over A^u, B^u, A^v and B^v in different ways in the two cases. For the complex normalization we can prove the "Vertauschung" theorem or the law of interchange of argument and parameter. It involves two normal functions of the third kind which differ in the location of their logarithmic singularities. Let Ω_P and Ω_Q have their singularities respectively at P_1 and P_2 and Q_1 and Q_2; then integration of $\int \Omega_P d\Omega_Q$ shows that

$$\Omega_P(Q_2) - \Omega_P(Q_1) = \Omega_Q(P_1) - \Omega_Q(P_2).$$

The simplest example of a third kind function is the one relevant to the sphere with no handles or the sphere or the complex plane with no holes. We find

$$\Omega_{\zeta_1 \zeta_2}(z) = \log \frac{z - \zeta_1}{z - \zeta_2} + \text{const.}$$

So far we have dealt with harmonic functions on surfaces with n handles and no boundaries. Starting with such a surface, we can, with one fell swoop sever it so as to create two surfaces, each with no handles and n holes. Each of these surfaces may then be flattened out to resemble the complex plane with the point at infinity included (Figs. 5,6,7). If we started with a simple sphere (no handles) the slice would give only C_1 on the lower hemisphere and C_2 on the upper hemisphere as boundaries respectively of surfaces B_1 and B_2, where the original sphere is $D = B_1 + B_2$. Knowing the third kind functions Ω on D, we can construct Green's and Neumann functions on B_1. In the zero handle case we imagine C_1 mapped conformally on to the real axis, B_1 into the upper half plane and B_2 into the lower half plane. The Green's function of B_1 is defined to be the single valued harmonic function on B_1 which vanishes on the boundary of B_1 ($C_1+C_3+C_5+C_7$) and which has a logarithmic pole at the point ζ such that the difference

$$g(z,\zeta) - \log \frac{1}{|z-\zeta|}$$

remains regular at ζ. Let $\zeta=\zeta_1$ and $\zeta_2=\bar{\zeta}$ (complex conjugate). Then

$$g(z,\zeta) = \tfrac{1}{2} \log \frac{(z-\zeta)}{(z-\bar{\zeta})} \cdot \frac{(\bar{z}-\bar{\zeta})}{(\bar{z}-\zeta)}$$

$$= \tfrac{1}{2} \log \frac{(z-\zeta)}{(z-\bar{\zeta})} - \tfrac{1}{2} \log \frac{(\bar{z}-\zeta)}{(\bar{z}-\bar{\zeta})}$$

$$g(z,\zeta) = \tfrac{1}{2} \left[\Omega_{\zeta\bar{\zeta}}(z) - \Omega_{\zeta\bar{\zeta}}(\bar{z}) \right]$$

is equal to zero on the real axis $z = \bar{z}$, has only one logarithmic singularity ζ in B_1, the upper half plane, and is expressed as a difference of two third kind functions. We may reexpress g as

$$g(z,\zeta) = \tfrac{1}{2} \log \left(\frac{z-\zeta}{z-\bar{\zeta}} \right) + \tfrac{1}{2} \log \left(\frac{\bar{z}-\bar{\zeta}}{\bar{z}-\zeta} \right)$$

$$= \tfrac{1}{2} \left[\Omega_{\zeta\bar{\zeta}}(z) + \bar{\Omega}_{\zeta\bar{\zeta}}(z) \right]$$

$$= R_e \, \Omega_{\zeta\bar{\zeta}}(z)$$

Therefore Im $[\Omega_{\zeta\bar{\zeta}}(z) - \Omega_{\zeta\bar{\zeta}}(\bar{z})]$ is identically zero.

To generalize this result to the case of more circles we consider the <u>symmetrical</u> case in which $C_{2\mu}$ is obtained by reflection in the real axis of $C_{2\mu-1}$. We call the real axis C_0. The Green's function is required for the domain B_1 consisting of the upper half plane bounded by $C_0, C_1, C_3, \ldots, C_{2n-1}$. The insides of the circles are considered to be exterior to B_1. The domain D consists of the whole complex plane with the interiors of all circles C removed. The Green's function of B_1 and the third kind functions of D depend on the positions of the centers of the $C_{2\mu-1}$, and on their radii. Therefore our task becomes one of writing down a real harmonic function which is zero on C_0 and

constant at all the circles $C_{2\mu-1}$, $C_{2\mu}$ in terms of parameters locating these circles and defining their radii. The question as to the best way to parametrize the circles has a deep answer which leads to the theory of discontinuous groups of Mobius transformations. Recall that a Möbius transformation T given by the correspondence

$$w(z) = \frac{az + b}{cz + d}$$

maps circles in the complex z- plane into circles in the complex w-plane. Here a,b,c,d are constants independent of z, and a straight line is considered to be a circle with center located at the point at infinity. Therefore one can find a set (a_r,b_r,c_r,d_r) such that

$$w(z) = \frac{a_r z + b_r}{c_r z + d_r}$$

maps C_0 into C_r. The n Möbius transformations which carry out these mappings are called generators. It will be shown that the specification of the domain B_1 by the set of $w_r(z)$ is the propitious choice. Before proceeding with the construction of Green's function, we must explain several salient geometric properties of Möbius (projective) transformations in general. We will henceforth assume ad-bc = 1. The inverse transformation T^{-1} is given by

$$w = \frac{-dz + b}{cz - a} .$$

A 2×2 unimodular matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ may be associated with T; then the inverse matrix is associated with T^{-1}. The conformal mapping accomplished by T maps any given three points a,b,c into three variable points α,β,γ. Introduce the cross-ratio symbol $(z_1\ z_2\ z_3\ z_4)$, where

$$(z_1\ z_2\ z_3\ z_4) = \frac{(z_1 - z_3)(z_2 - z_4)}{(z_1 - z_4)(z_2 - z_3)} ;$$

then

$$(w\ 1\ 0\ \infty) = (z\ a\ b\ c)$$

reads

$$w = \frac{(z-b)(a-c)}{(z-c)(a-b)} = \frac{(a-c)z + b(c-a)}{(a-b)z + c(b-a)} ,$$

and has the form of a Möbius transformation. The mapping $(a,b,c) \rightarrow (\alpha,\beta,\gamma)$ is given by

$$(w\ \alpha\ \beta\ \gamma) = (z\ a\ b\ c).$$

Its inverse transformation comes out to be

$$(w\ a\ b\ c) = (z\ \alpha\ \beta\ \gamma).$$

Since three points determine a circle and the mapping is conformal, we see that circles are always mapped into circles or lines. A Mobius transformation $w=\frac{az+b}{cz+d}$ has two <u>fixed points</u> f_1, f_2, which satisfy the quadratic equation

$$f = \frac{af + b}{cf + d}.$$

When we have found f_1 and f_2 we can write $w = \frac{az + b}{cz + d}$ in the equivalent form

$$\frac{w - f_1}{w - f_2} = K \frac{z-f_1}{z-f_2}$$

or $(w\ \infty\ f_1\ f_2) = (z\ \frac{f_2-Kf_1}{1 - K}\ f_1\ f_2)$,

which show immediately that $f_1 \rightarrow f_1$ and $f_2 \rightarrow f_2$. The constant K is called the <u>multiplier</u> ; a Möbius transformation is determined by giving its two fixed points and multiplier. The inverse transformation has the same fixed points f_1 and f_2.

In an analytic transformation $w = f(z)$, a lineal element $dz = z_2-z_1$ connecting two points in the infinitesimal neighborhood of a point z is transformed into the lineal element dw in the neighborhood of w. We have $dw = f'(z)dz$; hence, the length of the element is multiplied by $|f'(z)|$. Infinitesimal lengths in the neighborhood of a point z are

are multiplied by $|cz+d|^{-2}$, since

$$w = \frac{az+b}{cz+d}\ ,\ \frac{dw}{dz} = f'(z) = \frac{1}{(cz+d)^2}\ .$$

The circle I,

$$I: \quad |cz+d| = 1$$

which is the complete locus of points in the neighborhood of which lengths and areas are unaltered in magnitude by the transformation, is called the isometric circle of T. It is easy to see that lengths and areas within the isometric circle are increased in magnitude, and lengths and areas without the isometric circle are decreased in magnitude by the transformation. For, if z is within I, $|z+\frac{d}{c}| < \frac{1}{|c|}$, or $|cz+d| < 1$, and $|f'(z)| > 1$. A length or area within I is thus magnified in all its parts. Similarly, if z is without I, $|f'(z)|<1$; and a length or an area is diminished in all its parts.

The inverse transformation

$$w = \frac{-dz + b}{cz - a}$$

has the isometric circle I′, where

$$I': \quad |cz - a| = 1.$$

The center of I is at $-\frac{d}{c}$ and its radius is $\frac{1}{|c|}$; while the center of I′ is at $\frac{a}{c}$ with the radius of I′ being the same as the radius of I. It is a theorem that a transformation carries its isometric circle into the isometric circle of the inverse transformation. For T carries I into a circle I_0 without alteration of lengths in the neighborhood of any point, hence T^{-1} carries I_0 back to I without alteration. But I′ is the complete locus of points in the neighborhood of which T effects no change of length; hence I_0 coincides with I′. We can show this algebraically by computing $|cw(z) - a|$ when $|cz+d| = 1$ and $w = \frac{az+b}{cz+d}$. We find $(-cw+a)(cz+d) = 1$ or

$$|cw-a| = \frac{1}{|cz+d|} .$$

Therefore, if $|cz+d| = 1$ then also $|cw-a| = 1$. Furthermore, from $|cw-a||cz+d| = 1$ we can see that the outside $|cz+d| > 1$ of I maps into the inside $|cw-a| < 1$ of I′ and has all lengths shrunk in order to fit in.

If the multiplier K is a real number, the transformation is called hyperbolic. There are four geometric facts which can be proved about hyperbolic transformations:

(i) any circle through the fixed points $f_1 f_2$ is transformed into itself, each of the two arcs into which the circle is separated by the fixed points being transformed into itself;

(ii) the interior of a circle through the fixed points is transformed into itself;

(iii) any circle orthogonal to the circles through the fixed points is carried into some other such circle;

(iv) the fixed points are inverse with respect to each circle of (iii). This point has an obvious electrostatic interpretation.

Möbius transformations, whose singularities are poles may seem quite different from Green's functions, whose singularities are logarithms. Actually Green's functions are obtained from infinite sums of Möbius transformations by indefinite integration on z; the integral contains the logarithms. These integrals are called Abelian Integrals of the Third Kind, and represent third kind functions discussed above. The differential under the sign of integration has poles, and is called an Abelian differential of the second kind. The infinite sums are sums over image points when given a potential-theoretic interpretation.

The matrix representation $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ of a Möbius transformation T makes it evident that compound transformations can be effected and that a group structure will emerge. The groups appearing in dual tree theory are SU(1,1) and

$SU(1,1)\times SU(1,1)$. They are Lie groups, containing infinitesimal transformations in which the constants independent of z are infinitesimally different from zero (or unity). But there is another way to generate groups starting with Möbius transformations; and that is to specify say n transformations $T_1, T_2 \ldots, T_n$, where T_r has a definite form

$$T_r: \qquad w_r = \frac{a_r z + b_r}{c_r z + d_r} , \qquad r=1,2,\ldots,n$$

with constants (a_r, b_r, c_r, d_r) quite different from those required to make w infinitesimally close to z. These <u>generators</u> are given once and for all, and the infinity of elements of the group are obtained by <u>compounding</u> Möbius transformations discretely out of them. Thus, for n=1, the group elements are obtained by taking all possible <u>powers</u> of T_1:

$$T_1: \quad \frac{w-f_1}{w-f_2} = K \frac{z-f_1}{z-f_2} ,$$

$$(T_1)^2: \quad \frac{w-f_1}{w-f_2} = K^2 \cdot \frac{z-f_1}{z-f_2} ,$$

$$\vdots$$

$$(T_1)^\nu: \quad \frac{w-f_1}{w-f_2} = K^\nu \frac{z-f_1}{z-f_2} , \qquad \nu=0, \pm 1, \pm 2, \ldots$$

The compounding is done in the usual sense of a function of a function; $(T_1)^2(z) = T_1(T_1(z))$, etc. It is easy to show that $(T)^\nu \equiv T^\nu$, the ν^{th} power of T_1 is in fact given by the above formula. (It is this fact that makes the fixed-point representation convenient; every power of T is a transformation with <u>the same fixed points</u>, and a multiplier which is the K of the generator raised to the corresponding power.) This set of transformations constitutes an Abelian group with an infinite number of commuting elements and one <u>generator</u> T. This group is called a <u>discontinuous</u> group since, although it contains the identity ($\nu=0$), it caontians no elements in an infinitesimal neighborhood of the identity, as does a Lie group such as SU(1,1). We have already

discussed the geometry of T and T^{-1} in terms of their fixed points and isometric circles; we now go on to the geometry of the multiply-infinite set of Möbius transformations generated by taking all possible powers and products of powers of the T_r, r=1,2,...,n such as $(T_1)^3(T_2)^{-5}(T_3)^2(T_1)^6(T_3)^{-5}$. Because the T_r do not commute with one another a given T_r may appear as often as we like in the product (which contains always a finite number of factors). Once the T_r are specified, a set of n pairs of isometric circles $I_{2\mu-1}$, $I_{2\mu}$ are defined which divide up the complex z- plane into (i) fundamental region R_0 outside all the circles and (ii) the complement of the fundamental region consisting of the sum of the insides of the isometric circles. A transformation of the group carries the points of the fundamental region R_0 out of the fundamental region into another region R; this region R is quite equivalent to R_0 and R may be called a fundamental region for the group. Two configurations (points, curves, regions) are said to be congruent with respect to a group if there is a transformation of the group which carries one configuration into the other.

A major theorem is contained in the statement: If no two points of a region are congruent, the transforms of the region by two distinct transformations of the group do not overlap. A corollary is that the transforms of a fundamental region by two distinct transformations of the group do not overlap. Each group element has its own isometric circle. By considering the isometric circle of an element which is the product of two given ones (two successively performed Möbius transformations) we find that the radii of the isometric circles are bounded; the number of isometric circles with radii exceeding a given positive quantity is finite; and given any infinite sequence of distinct isometric circles $I_1, I_2, I_3, \ldots$, of transformations of the group, the radii being $r_1, r_2, r_3, \ldots$, then $\lim_{n\to\infty} r_n \equiv 0$. In studying the arrangement of isometric circles, the following theorem is significant. Let I_1, I_2, I_1' and $T_{21} \equiv T_2T_1$ where T_1 operates first. If I_1' and I_2 are exterior to one another, then I_{21} is contained inside I_1. The proof is simple. Let z be a point outside I_1. Then $z_1 = T_1(z)$ is inside I_1' and therefore, by hypothesis, outside I_2. Regions around z are shrunk by T_1. Next apply T_2, mapping z_1 into z_{21}. Regions around z_1 are shrunk into regions around z_{21}. Now T_{21} maps z into z_{21} and produces the shrinking of T_1 plus the shrinking of T_2, so T_{21}

shrinks regions around z into regions around z_{21}. Therefore z is external to I_{21} and every point z external to I_1 is also external to I_{21}. Therefore, I_{21} is contained in I_1.

In the dual resonance model the fixed points of all the generating transformations $T_{2\mu-1}$, $T_{2\mu}$, $\mu=1,2,\ldots,n$ lie on the Koba-Nielsen unit circle, which by Möbius invariance can be taken to be the real or imaginary axes. If the fixed points of the generators be on the real axis the group is called a Fuchsian group, and the fixed points of all the transformations of the group will also lie on the real axis. Then the real axis is called the principal circle for the group and it contains, in addition to the fixed points, also the centers of the isometric circles. The isometric circles of a pair $T_{2\mu-1}$, $T_{2\mu}$ of generators have their centers on the real axis, have equal radii and are congruent. The fundamental region R is the whole complex plane minus the insides of the T_r. We designate by R_0 the part of R lying within the principal circle (or R_0 is the upper half plane). R_0 is carried by all the transformations of the group into a region in the interior of the principal circle. A theorem states that any closed region lying entirely within the principal circle is covered by a finite number of transforms of R_0. These regions fit together without lacunae. Also, the transforms of R_0 fill up, without lacunae, the whole interior of the principal circle. They cluster in infinite number about the limit points of the group (which lie on the principal axis and inside the isometric circles of the generators or outside R_0.) Groups of this type are called simultaneously Schottky groups and Fuchsian groups of the second kind.

We are now ready to introduce automorphic functions with respect to a discontinuous group. Automorphic functions are the generalization of the circular, hyperbolic and elliptic functions of ordinary analysis. A circular function, such as sin z, has the property that it remains unchanged in value if z is replaced by $z+2m\pi$, where m is any integer; that is the function is unaltered in value if z be subjected to a transformation of the group $w = z + 2\pi m$. A hyperbolic function, such as sinh z, is unchanged in value if z be subjected to a transformation of the group $w = z + 2\pi i m$. An elliptic function, such as

the Weierstraussian Function $\wp(z)$, retains its value under transformations of the group $w = z + mw + m'w'$.

The automorphic function is an extension of this concept to the more general discontinuous group. Roughly speaking, a function is automorphic with respect to such a group if it has the same value at congruent points. If $T_i(z)$ is any transformation of the group, then $f(z)$ is an automorphic function with respect to the group if $f(z)$ is a single-valued analytic function inside the principal circle and $f[T_i(z)] = f(z)$ there. A theorem states that each limit-point of the group is an essential singularity of the function. For in the neighborhood of a point at which a function is analytic or has a pole, the function can take on any value only a finite number of times. In the neighborhood of a limit point, there is an infinite number of congruent points at which the function takes on the same value. Hence, the limit point is an essential singularity. We shall now actually set up an automorphic function by means of a series.

Let the transformations of the group be

$$z_i = T_i(z) = \frac{a_i z+b_i}{c_i z+d_i}, \quad a_i d_i - b_i c_i = 1$$

$$i = 0,1,2,\ldots$$

the identical transformation being $z_0 = T_0(z) = z$. For simplicity, we write $z_{ij} = T_i(z_j) = T_i T_j(z)$. Suppose, first, that the group is finite. Let the group contain m transformations $(i=0,1,\ldots,m-1)$. Let $H(z)$ be any <u>rational</u> function of z and form the function

$$\varphi(z) = H(z) + H(z_1) + H(z_2) + \cdots + H(z_{m-1}).$$

This function has no singularities other than poles. If we apply a transformation T_k of the group to z, we have

$$\varphi(z_k) = H(z_k) + H(z_{1k}) + \cdots + H(z_{m-1,k}).$$

Now, $z_k, z_{1k}, \ldots, z_{m-1,k}$ are the set of transforms of z and, since z_k is congruent to z, this set coincides with $z, z_1, \ldots, z_{m-1}$. The terms of the sum are the same as before, their order being merely interchanged; hence $\varphi(z_n)=\varphi(z)$. The function is thus automorphic. It has no other singularities than poles. If the group contains an infinite number of transformations and we extend the sum to an infinite number of terms and add convergence factors $(c_i z+d_i)^{-2}$, we get the Poincaré theta series

$$\theta(z) = \sum_{o}^{\infty} (c_i z + d_i)^{-2} H(z_i)$$

which is not quite an automorphic function, but rather satisfies

$$\theta(z_j) = (c_j z + d_j)^2 \theta(z).$$

It is easy to form an automorphic function from two theta series with different $H_1(z)$ and $H_2(z)$. The ratio $\theta_1(z)/\theta_2(z)$ is, in fact, automorphic. We now show how Abelian differentials with poles (second kind) may be formed with the Poincaré theta series for a Fuchsian Schottky group.

We take

$$H(z) = \frac{1}{z-a}$$

and define $\theta(z,a)$ by

$$\theta(z,a) = \Sigma \frac{1}{(\gamma_i z+\delta_i)^2} \cdot \frac{1}{w_i(z) - a},$$

$$w_i(z) = \frac{\alpha_i z + \beta_i}{\gamma_i z + \delta_i}.$$

We define R_i as the region which R_o is mapped into by T_i. The totality of R_i fill the plane inside the principal circle. The function $\theta(z,a)$ has two simple infinities

inside any one of the regions R_i, namely the homologues of $\underline{a}$ and $\underline{\infty}$ that lie inside R_i. The same statement is true of $\theta(z,w_j(a))$. It follows that constant multiples of the two functions can be chosen such that their difference shall be independent of $\underline{a}$. We can show, using the fact that the product of two matrices

$$\begin{bmatrix} a_i & \beta_i \\ \gamma_i & \delta_i \end{bmatrix} \cdot \begin{bmatrix} a_j & \beta_j \\ \gamma_j & \delta_j \end{bmatrix}$$

gives another matrix which defines an element of the group and a term in the series, that

$$\theta(z,w_j(a)) - \theta(z,a) = \theta(z, -\frac{\delta_j}{\gamma_j}) ,$$

which is independent of a. Recall that $-\frac{\delta_j}{\gamma_j}$ is the center J_j of the isometric circle for T_j. The function

$$\theta(z, -\frac{\delta_j}{\gamma_j})$$

has no poles, its only singular points being the singular points of the group. The points $-\frac{\delta_j}{\gamma_j}$ and $\frac{\alpha_j}{\gamma_j}$ are the

homologues of infinity, and there are clearly an infinite number of functions $\theta(z,J_i)$. However, only n of them are linearly independent, since we can show that

$$\theta(z,J_p) - \theta(z,J_q) = -\theta(z, J_{p^{-1}q})$$

where $J_{p^{-1}q}$ is the center of the isometric circle for the transformation $T_p^{-1}T_q$, which is given by

$$J_{p^{-1}q} = \frac{\delta_p \alpha_q - \beta_p \gamma_q}{-\gamma_p \alpha_q + \alpha_p \gamma_q} .$$

The same process of direct substitution giving this result also can show that

$$\theta(z,J_p) = -\theta(z,J_{p-1}).$$

Hence, if T_p and T_q are two of the fundamental substitutions (generators), and J the homologue of infinity formed by any combination of these two, then $\theta(z,J)$ can be expressed in the form $m\theta(z,J_p) + n\theta(z,J_q)$, where m and n are positive integers. It follows from this at once that if $T_1, T_2 \ldots T_n$ are the n fundamental substitutions, then, whatever homologue of infinity J may be, $\theta(z,J)$ can always be expressed in the form

$$\sum_1^n m_i \ \theta(z,J_i).$$

We now carry out integrals of θdz of the form

$$\int_{z_0}^{z} \theta(z,J)dz = \Sigma \log \frac{w_i(z) - J}{w_i(z_0) - J},$$

when the integral path is at first confined to the generating polygon (fundamental region). The generating polygon consists of the space outside n pairs of circles C_1 and C_1', C_2 and C_2', ..., C_n and C_n', each external to all the others. The substitution T_p transforms the circle C_p' into the circle C_p, and the space outside the circle C_p' into the space inside the circle C_p, so that the point J_p is inside C_p and the point J_{p-1} is inside C_p'. An element of the integral of $\theta(z,J_p)$ may be expressed as follows:

$$\theta(z,J_p)dz = \Sigma d. \log (w_i(z) - J_p)$$

$$= \Sigma d \log \frac{\alpha_i \left(z - \dfrac{\delta_i J_p - \beta_i}{-\gamma_i J_p + \alpha_i}\right)}{\gamma_i \left(z + \dfrac{\delta_i}{\gamma_i}\right)}$$

$$= \Sigma d \cdot \log \frac{\alpha_i (z - J_{i-1}p)}{\gamma_i (z - J_{i-1})} .$$

When T_i is the identical substitution $w = z$, the corresponding term on the right hand side is of a slightly different form, namely $d \log(z-J_p)$; hence

$$\theta(z,J_p)dz = d \cdot \log(z - J_p) + \Sigma\, d \cdot \log \frac{\alpha_i(z - J_{i-1}p)}{\gamma_i(z - J_{i-1})}$$

where now the identical substitution is not included under the sign of summation. Now, as before stated, J_p is inside C_p, and it is easy to see from the theorem on the location of isometric circles that J_{i-1} and $J_{i-1}p$ are either both inside or both ourside C_p; hence

$$\int_{C_p} \theta(z,J_p)\, dz = 2\pi i,$$

where the integral is taken in the positive direction around a closed curve within the generating polygon surrounding C_p once, and surrounding no other circle.

Again, J_p is outside C_p', and J_{i-1} and $J_{i-1}p$ are either both inside or both outside C_p', except when $i = p$, and then J_{i-1} is inside and $J_{i-1}p$ is outside, hence

$$\int_{C_p'} \theta(z,J_p)\, dz = -2\pi i,$$

where the integral is taken positively around C_p'. For any other circle C_q or C_q', J_p is outside and J_{i-1} and $J_{i-1}p$ are always either both inside or both outside. Hence

$$\int_{C_q} \theta(z,J_p)dz = \int_{C_q'} \theta(z,J_p)dz = 0.$$

These results prove incidentally that the n functions $\theta(z,J_1)$, $\theta(z,J_2),\ldots,\theta(z,J_n)$ are linearly independent of each other, and that they are not mere constants.

If A_p,A_p' are two corresponding points on C_p and C_p', and B_p,B_p' any other pair of corresponding points on the same two circles, and if A_pA_p', B_pB_p' are joined by paths A_pMA_p', B_pNB_p', which do not enclose any of the circles, then $\int\theta(z,J)dz$ around $A_pMA_p'B_p'NB_pA_p$ vanishes, since the integrand is finite and continuous at all points within the contour. Therefore

$$\left(\int_{A_pMA_p'} - \int_{B_pNB_p'} + \int_{A_p'B_p'} - \int_{A_pB_p}\right) \theta(z,J)\, dz = 0.$$

Now, if z_1,z_2 are the corresponding points on the circular arcs $A_p'B_p'$, A_pB_p,

$$z_2 = \frac{\alpha_p z_1 + \beta_p}{\gamma_p z_1 + \delta_p} \quad ;$$

therefore

$$\theta(z_2,J) = (\gamma_p z_1+\delta_p)^2\ \theta(z_1,J),$$

and

$$dz_2 = (\gamma_p z_1+\delta_p)^{-2}\ dz_1\,;$$

therefore

$$\theta(z_2,J)\ dz_2 = \theta(z_1,J)\ dz_1.$$

It follows that

$$\int_{A_pB_p} \theta(z,J)\ dz = \int_{A_p'B_p'} \theta(z,J)\ dz,$$

the integrals being taken along the circular arcs, and therefore

$$\int_{B_p N B_p'} \theta(z,J)dz = \int_{A_p M A_p'} \theta(z,J)dz.$$

Hence the integral

$$\int_{A_p'}^{w_p(A_p')} \theta(z,J_q)\, dz,$$

where A_p' is any point on the circle C_p', and q is equal or unequal to p, is independent of the particular position of A_p', the path varying continuously without passing outside the fundamental polygon.

Now let the points A_1A_1',A_2A_2', etc. be joined by lines of any form, which neither cut themselves nor each other, and which do not leave the generating polygons; and regard these lines (crosscuts), as well as the circles, as part of the boundary of the polygon. In the figure so formed, the integral of $\theta(z,J)$, when the lower limit has an assigned value, is a one-valued finite continuous function. Let

$$\int_{A_p'A_p} \theta(z,J_q)\, dz = a_{qp}$$

$$\int_{A_p'A_p} \theta(z,J_p)\, dz = a_{pp},$$

so that the quantities a_{pp}, a_{pq} are the constant values of the integrals just discussed. When the variable paths between the corresponding points are chosen so as to be reconcilable with the barriers A_1A_1',A_2A_2',$\cdots$, it can be shown that

$$a_{pq} = a_{qp}.$$

By now it should be dawning that $\theta(z,J_q)dz$ are first Abelian differentials for the domain with the circles excised. The first Abelian integrals are given by

$$\Phi_p = \int^z \theta(z,J_p)\, dz,$$

when J_p is the center of a generator circle; there are n linearly independent $\Phi_p(z)$. The constants a_{qp} are just the moduli of periodicy previously discussed. The normalization here is such that

$$\Phi_p(w_i(z)) - \Phi_p(z) = 2m_p\pi i + \sum_1^n n_i\, a_{pi}, \quad p=1,2,\ldots,n.$$

The real part of $\Sigma\, \lambda_p^2 a_{pp} + 2\Sigma\Sigma\lambda_p\lambda_q\, a_{pq}$ is essentially positive for all real λ_p.

The Abelian integrals of the second and third kinds are now easily found. In place of $H_a(z) = \frac{1}{z-a}$ we put

$$-\frac{\partial H_a(z)}{\partial a} = \frac{1}{(z-a)^2}$$

to obtain the series

$$-\Sigma\, \frac{1}{(\gamma_i z + \delta_i)^2} \cdot \frac{1}{(w_i(z) - a)^2},$$

$$w_i(z) = \frac{\alpha_i z + \beta_i}{\gamma_i z + \delta_i},$$

which has a double infinity at the point a and its homologues; while the homologues of $z = \infty$ are not infinities of the function. Its integral will therefore be a single-valued function, finite and continuous everywhere except at the point a and its homologues; at these points it will have a simple infinity. This function will be denoted by $\psi_a(z)$, so that

$$\psi_a(z) = \Sigma\left(\frac{1}{w_i(z)-a} - \frac{1}{w_i(z_0)-a}\right),$$

where z_o represents the origin of the integration, which can be chosen at convenience. We can show by manipulation that

$$\psi_a(z) = \theta(a, z_o) - \theta(a, z),$$

and that

$$\psi_a(w_p(z)) - \psi_a(z) = -\theta(a, J_p)$$

is a constant. Thus, these functions have one infinity inside each homologue R_i and R_o, are single valued (i.e., they do not change when the variable describes a closed path which cuts the barriers), and, when one of the transformations of the group is performed on the variable, they increase by integral multiples of definite constants.

Again the function

$$\Sigma \frac{1}{(\gamma_i z+\delta_i)^2} \left(\frac{1}{w_i(z)-a} - \frac{1}{w_i(z)-b} \right)$$

is finite everywhere except at a and b and their homologues, which points are simple poles. If the integral from an arbitrary origin z_o is written $\Omega_{ab}(z)$, then

$$\Omega_{ab}(z) = \Sigma \log \frac{(w_i(z)-a)(w_i(z_o)-b)}{(w_i(z)-b)(w_i(z_o)-a)}$$

is an Abelian integral of the third kind, where the branch of the logarithm is that which makes

$$\Omega_{ab}(z_o) = 0.$$

It is easy to show that

$$\Omega_{ab}(z_2) - \Omega_{ab}(z_1) = \Omega_{z_2 z_1}(a) - \Omega_{z_2 z_1}(b).$$

Green's functions can now be explicitly constructed,

and we will choose a particularly propitious arrangement of generator circles. A given pair of circles C_p, C_p' will lie symmetrically with respect to the real axis, and an extra circle C_0 will coincide with the real axis. (C_0 may be imagined as two circles C_0 and C_0' pasted together and mapped conformally on to the real axis.) The generating transformations all have the form

$$\frac{w - a_p}{w - a_p'} = K_p \frac{z - a_p}{z - a_p'},$$

where K_p is real, and a_p, a_p' are conjugate imaginaries; and it therefore follows that the substitutions of the group may be taken in pairs

$$w_i = \frac{\alpha_i z + \beta_i}{\gamma_i z + \delta_i} \quad \text{and} \quad w_i = \frac{\alpha_i' z + b_i'}{\gamma_i' z + \delta_i'}$$

such that $\alpha_i, \beta_i, \gamma_i, \delta_i$ and $\alpha_i', \beta_i', \gamma_i', \delta_i'$ are respectively complex conjugates.

If z, z' and a, a' are complex conjugates, so also are

$$\frac{(\gamma_i z + \delta_i)^{-2}}{w_i(z) - a} \quad \text{and} \quad \frac{(\gamma_i' z' + \delta_i')^{-2}}{w_i'(z') - a'}$$

and therefore also $\theta(z,a)$ and $\theta(z',a')$. Now J_p and $J_{p^{-1}}$ are in this case complex conjugates, and it has been shown that in any case

$$\theta(z, J_{p^{-1}}) = -\theta(z, J_p),$$

whence it follows that

$$\theta(z, J_p) \text{ and } -\theta(z', J_p)$$

are complex conjugates. Hence, when z is real, $\theta(z, J_p)$ is pure imaginary.

Now suppose that z' is any point on C_q', so that z is the corresponding point on C_q; then

$$dz = dr \cdot e^{i\Phi}, \quad dz' = dr\, e^{-i\Phi};$$

but

$$z = w_q(z') = \frac{\alpha_q z' + \beta_q}{\gamma_q z' + \delta_q}$$

and therefore

$$dz = \frac{dz'}{(\gamma_q z' + \delta_q)^2};$$

and

$$(\gamma_q z' + \delta_q)^2 = e^{-2i\Phi}.$$

Also, in consequence of the relation between z and z',

$$\theta(z,J_p) = (\gamma_q z' + \delta_q)^2\, \theta(z',J_p) = e^{-2i\Phi}\theta(z',J_p),$$

or

$$\arg \theta(z,J_p) = -2\Phi + \arg \theta(z',J_p);$$

but, since $\theta(z,J_p)$ and $-\theta(z',J_p)$ are complex conjugates, $\arg \theta(z,J_p) = \pi - \arg \theta(z',J_p)$; therefore

$$\arg \theta(z,J_p) = \frac{\pi}{2} - \Phi.$$

The ratio of any two of the functions $\theta(z,J)$ is therefore real at each of the circles $C_1, C_1', \ldots, C_n, C_n'$ as well as at C_0.

The barriers (crosscuts) in this case may be taken as straight lines perpendicular to the x-axis; and since it was shown that, when z and z' are complex conjugates, so also are $\theta(z,J_p)$ and $-\theta(z',J_p)$, it follows at once that

$$\int_{A_q A_q'} \theta(z,J_p)dz \ (=a_{pq})$$

is <u>real</u>, where a_{pq} is any one of the $n(n-1)$ constants.

Finally, since for points on any one of the bounding circles, including A_o,

$$\arg \theta(z,J_p) = \frac{\pi}{2} - \Phi$$

and

$$\arg dz = \Phi,$$

the variable part of $\int\theta(z,J_p)dz$, or of Φ_p, is a pure imaginary at all the circles.

Sticking to the symmetrical case, let us find a series which will represent, in the space external to $n + 1$ circles, each of which is external to all the others, a function with a single pole at a, and whose imaginary part has the Green's function property of having constant values at the circles. Consider the function $A\psi_a(z) + A'\psi_{a'}(z)$, where A,A' and a,a' are conjugate imaginaries. Regarded as a function of z, this expression has entirely real coefficients, and therefore will take complex conjugate values when z does. Now let A_p be any point on the circle C_p, so that A_p' is the corresponding point on C_p'. Then, if

$$A\psi_a(A_p) + A'\psi_{a'}(A_p) = P + iQ$$

$$A\psi_a(A_p') + A'\psi_{a'}(A_p') = P - iQ;$$

but

$$A_p = \frac{\alpha_p A_p' + \beta_p}{\gamma_p A_p' + \delta_p},$$

and therefore

$$A\psi_a(A_p) + A'\psi_a(A_p) - [A\psi_a(A_p') + A'\psi_{a'}(A_p')] =$$

$$= A\theta(a, J_p) + A'\theta(a', J_p),$$

or

$$iQ = \tfrac{1}{2}A\theta(a, J_p) + \tfrac{1}{2}A'\theta(a', J_p).$$

It follows that, at each separate circular bounding curve of the generating polygon (fundamental region), the imaginary part of $A\psi_a(z) + A'\psi_{a'}(z)$ has constant values.

FIGURE CAPTIONS

1. Region of definition of multiple-valued harmonic conjugate functions; the interiors of C_1 and C_2 are outside of D.

2. The line L is split into L_1 and L_2 where values of v differ by a line integral.

3. Another topological form of the curve of Fig. 2; L_1 and L_2 are looped.

4. Intersection of crosscuts.

5. D is a two-dimensional surface in three-dimensional space. This example contains three handles.

6. The surface D is severed into B_1 and B_2.

7. B_2 is a sphere with four holes (windows) cut out; its sterographic projection on the complex plane is shown.

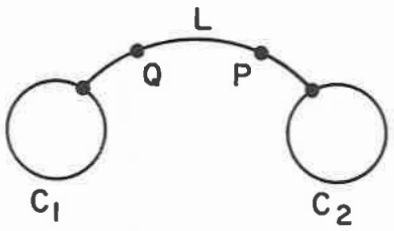


Figure 1

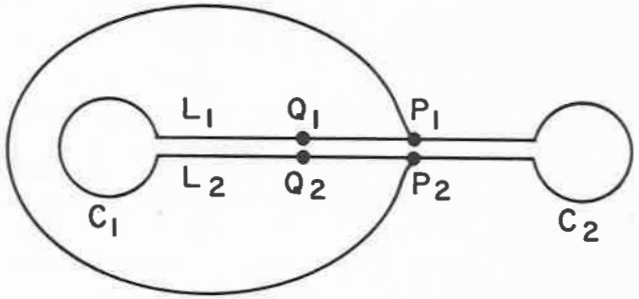


Figure 2

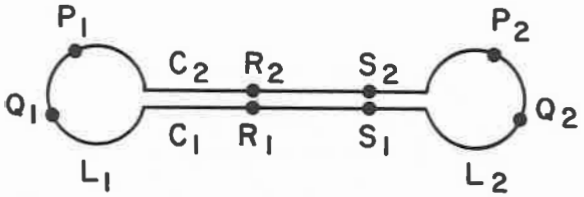


Figure 3

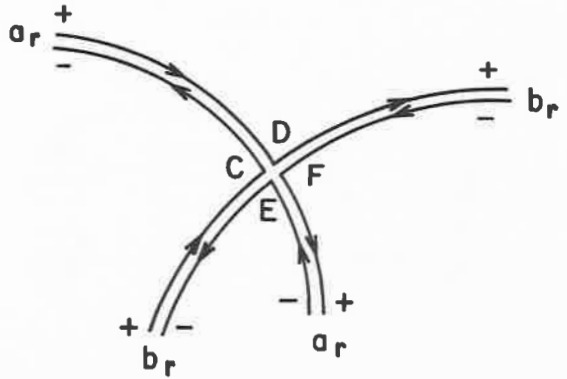


Figure 4

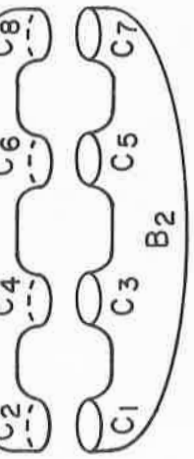


Figure 6

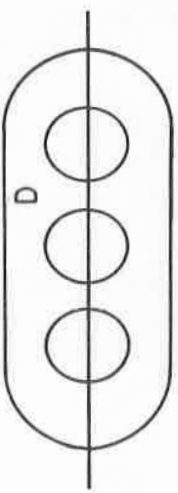


Figure 5

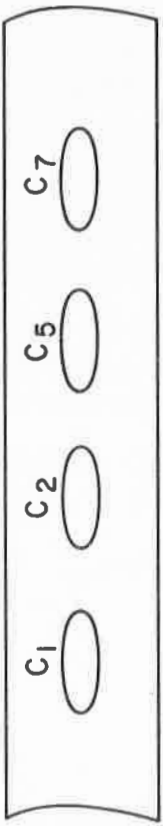


Figure 7

References

1. Z. Nehari, "Conformal Mapping," McGraw-Hill Book Company, Inc., New York (1952).

2. A.R. Forsyth, "Theory of Functions of a Complex Variable," vols. 1 and 2, third ed., Dover Publications, New York (1918,1965).

3. M. Schiffer and D.C. Spencer, "Functionals of Finite Riemann Surfaces," Princeton University Press, Princeton (1954).

4. L.R. Ford, "Automorphic Functions," Chelsea Publishing Company, New York (1929,1951).

5. W. Burnside, Proceedings of the London Mathematical Society XXIII, (1891).

6. C. Lovelace, Phys. Letters 32B, 703 (1970).

7. V. Alessandrini, Nuovo Cemento 2A, 321 (1971).

8. V. Alessandrini and D. Amati, Nuovo Cimento 4A (1971) 793.

9. G. Sansone and J. Gerretsen, "Lectures on the Theory of Functions of a Complex Variable," Vol. II, Wolters-Noordhoff Publishing Co., Groningen (1969).

10. A.D. Karpf, and H. J. Liehl, University of Freiburg, UF 1206-71, Dec. 1971.

Part III

NEW THEORETICAL MODELS IN STRONG INTERACTIONS

A POSSIBLE PHYSICAL BASIS FOR SOME MATHEMATICAL MODELS OF STRONG INTERACTIONS: ATOMS WITH MAGNETIC CHARGES

A. O. Barut
Institute for Theoretical Physics
University of Colorado
Boulder, Colorado 80302

I. INTRODUCTION

A physical picture of strong interaction phenomena based on established fundamental entities and their laws of interaction, as in the case of atomic and molecular physics, would be very important and useful, because the phenomenological or mathematical models of strong interaction-amplitudes have too much freedom in them, and do not in general enhance our knowledge very much. There are in principle no arbitrary parameters in atomic physics, for example, whereas abstract theories either contain enough parameters, or can always be modified to fit new experimental situations. Thus in models based on singularities in the angular momentum variable, or in models with symmetry breaking of abstract symmetry or dynamical groups, there are always additional terms that one can add, thereby modifying the values of the previously established parameters. The phenomenological models are undoubtedly not without foundation, for they contain some general features in agreement with experiment. It would be therefore of great value if these features could be derived or understood from a physical model. One should then also obtain necessarily a unifying basis of the seemingly very different concepts underlying the different abstract models such as quarks, partons, Regge poles, hadronic matter, etc.

Models for Scattering versus Models for Internal Structure

It is appropriate to emphasize the difference between the two types of modelmaking for strong interactions:

(A) One postulates models for scattering amplitudes (e.g. Regge-pole models, dual models, additive quark models,) and then deduces, eventually, from the systematics of all scattering processes the structure of hadrons and why they interact the way they do.

(B) One postulates models for single hadrons and their constitutions and then calculates or deduces the scattering amplitudes.

These two approaches should be complementary, because one asks in general different types of questions, and obtains different types of answers in each case.

Clearly, in atomic and molecular physics the method (B) was, and is, employed. Had we tried to deduce complicated atomic structures from scattering experiments we would have had a difficult task. In hadron physics, the quark model and the parton model maybe, at first, thought to be of type (B), but this is not quite so, because, although one makes a model of the internal constitution, the interaction between these entities is unknown and arbitrary. Consequently, one does not know whether these quarks or partons are really new objects, or simply mathematical "excitations" which appear in a simple additive form.

Contrary to some beliefs that one cannot make models of hadrons as one makes models of atoms and nuclei, there is, we believe, a definite physical picture leading to explicit models and where one knows precisely the interaction. Certain fundamental results derived from this model and not obtainable at present with any other approach, compels one to study further consequences of this physical picture. In this lecture I shall point out to relations of hadron models with magnetic charges with other approaches discussed at this year's Institute. What makes the model even more attractive is the intuitive possibility that not only electromagnetic and strong

interactions are unified, but also weak interactions can also be accomodated into the conceptual framework of the theory.

II. Review of Previous Results

Several aspects of the model has been discussed in detail elsewhere.[1] We consider magnetically neutral bound states made up of positive and negative magnetic charges. Thus total magnetic charge is zero. The two-body systems of two spinless particles with electric and magnetic charges (e_1, g_1) and (e_2, g_2) are characterized by two invariants

$$\alpha = e_1 e_2 + g_1 g_2$$

$$\mu = e_1 g_2 - e_2 g_1 = 0, \pm\tfrac{1}{2}, \pm 1, \pm\tfrac{3}{2}, \ldots . \qquad (1)$$

The quantum number μ turns out to represent the spin of the ground state. We now summarize the consequences of the model.

(i) Spin, charge and Baryon Number

The first result can be stated as follows:

$$(-1)^{2j} = (-1)^{Q} = (-1)^{B}, \qquad (2)$$

where j and Q are spin and charge of the states of the system. Thus, for $\mu = 0$ ($j_{min} = 0$), charge Q must be zero, for $\mu = \frac{1}{2}$ ($j_{min} = \frac{1}{2}$), charge Q must be different from zero. This is a striking simple explanation of the fact that the lowest mesonic state π^0 is spinless and chargless, and the lowest baryonic state, $j = \frac{1}{2}$, is charged.

We emphasize the rather unusual fact that, when magnetic charges are present, a spin $\frac{1}{2}$ state (e.g. proton) can be explicitely constructed out of two spin zero particles.

(ii) Stability

We have also the result that π^0 must be a particle-anti particle system hence annihilates into 2γ, whereas proton must be stable. The quantum number μ also plays the role of the baryon number.

(iii) Form factors

A third simple striking result is that a spin $\frac{1}{2}^+$-particle made out of two magnetic charges (proton) has a magnetic form factor of the dipole form as now verified up to momentum transfers of 25 $(\text{GeV}/c)^2$.

(iv) Principal quantum number

The next result is the appearance of a new quantum number n to distinguish states with the same spin and parity, such as nucleon and N* (1470), both $\frac{1}{2}^+$. The quantum number n is the standard principal quantum number of the two-body problem.

(v) Mass Spectrum and Upper Limit for Resonances

A further result is the explanation of the empirical hydrogen-like spectrum of the excited states of the proton and the prediction that there are no nucleon resonances with masses beyond about 5 GeV.

(vi) Limiting Diffraction Peak for Elastic Scattering

Another result, on the basis of a vector coupling of two protons, is the explanation of the sharp diffraction peak in the p-p scattering, and the prediction of a universal elastic differential cross-section $d\sigma/dt$ at $s \to \infty$ with a slope at $t = 0$ of 10.90 $(\text{GeV}/c)^{-2}$ (except at the very forward point $t \approx 0$, because of optical theorem).

(vii) The Constants in the Theory. Further Estimates.

The value of the magnetic charge g is fixed by Eq. (1). Consequently, the only parameters in the theory are $e^2 = (137)^{-1}$ and the masses of the constituent magnetic charges. The invariant α in Eq. (1), is the new and, in

principle, the only coupling constant for the two-body problem. Because its value is now large, $\alpha = (137/2)^2$, the dynamical two-body problem cannot be solved by any of the standard methods. The form factor and mass spectrum calculations cited above have been made on the basis of global infinite-component wave equations and relativistic dynamical groups in which the strongly bound two-particle system is treated as a single particle with internal degrees of freedom. Therefore, the masses of the magnetic charges are not known. However, assuming that their masses is purely electromagnetic, i.e. $m_g = g^2/e^2\, m_e \approx 2.4$ BeV, one can make a number of reasonable estimates: (a) saturation point of resonances = $2m_g \approx 4.8$ BeV, (b) magnetic moment of the proton $\approx$ 3 nuclear magneton. One has also a qualitative understanding of π^0 mass and lifetime.

(viii) Model for Neutron and charged Pions; (n-p) mass difference

Sofar we have discussed the lowest possible states π^0 and proton. Neutron and charged pions cannot be built as a simple two-body systems, because they have the weak decays into their ground state (proton or π^0, respectively) with the emission of a lepton-neutrino pair. Thus weak interactions may also be viewed as a consequence of the internal structure of hadrons similar to α-decay. Hence one can try to consider models of neutrons and $\pi^\pm$ as bound states of the previous basic proton and π^0-structures, with a purely electrically charged borons $B^\pm$ decaying into $(\ell^\pm\nu)$-pair. The model accounts qualitatively for the properties of neutron and charged pions. The electrically charged particle B can be tightly bound to a magnetic moment at distances of the order of 10^{-13}cm, radius of the magnetic charge. Note that in this model the n-p mass difference (or π^+-π^0 mass difference) is not purely of electromagnetic origin.

(ix) Dynamical Group and Infinite Component Wave Equation

Finally, we mention the mathematical result that the two-body system characterized by μ, (dyonium) has as it dynamical group the group O(4,2) in its unitary

representation which is also characterized by μ. The states in this class of representations are labeled by $|\mu;njm>$. Note the new quantum number μ in addition to the usual quantum numbers. In this space of states $|\mu njm>$ one constructs a conserved current operator j_μ to describe the electromagnetic properties of the model, as well as mass spectrum (see subsection (iii) and (v), and Sect.III). The conserved current j_μ is equivalent to an infinite component wave equation.

(x) Matter with Magnetic Charges

Table I shows schematically a parallel picture of the structure of matter based on electric charges alone (ordinary matter), on the left, and on the right, based on both electric and magnetic charges, (new matter), which we identify with the hadronic and nuclear matter. We remark that the magnetic field on the surface of the magnetic charge in our model is about 10^{12} Gauss, the same as estimated on the surface of the neutron star.

III. Implications for Scattering Processes

The model discussed above gives us now a method to visualize hadron scattering processes. We discuss some examples.

(i) Pion Production

As we have seen pions (and hadrons in general) are, in the present model, pairs of magnetic charges with total magnetic charge zero. Thus pion production will be viewed as pair-production. The subsequent pair, because $g_{tot} = 0$, can escape the production region. On this basis it is possible to account for the energy dependence of the average multiplicity and total production cross-section by multiperipheral-pair production calculations.[2] Because pair production is more favored than the "ionization" of the proton, almost all the energy will go into pions.

(ii) Deep Inelastic Electron-Nucleon Scattering

In the inelastic electron-nucleon scattering

experiments[3]

$$e + p \rightarrow e + \text{anything},$$

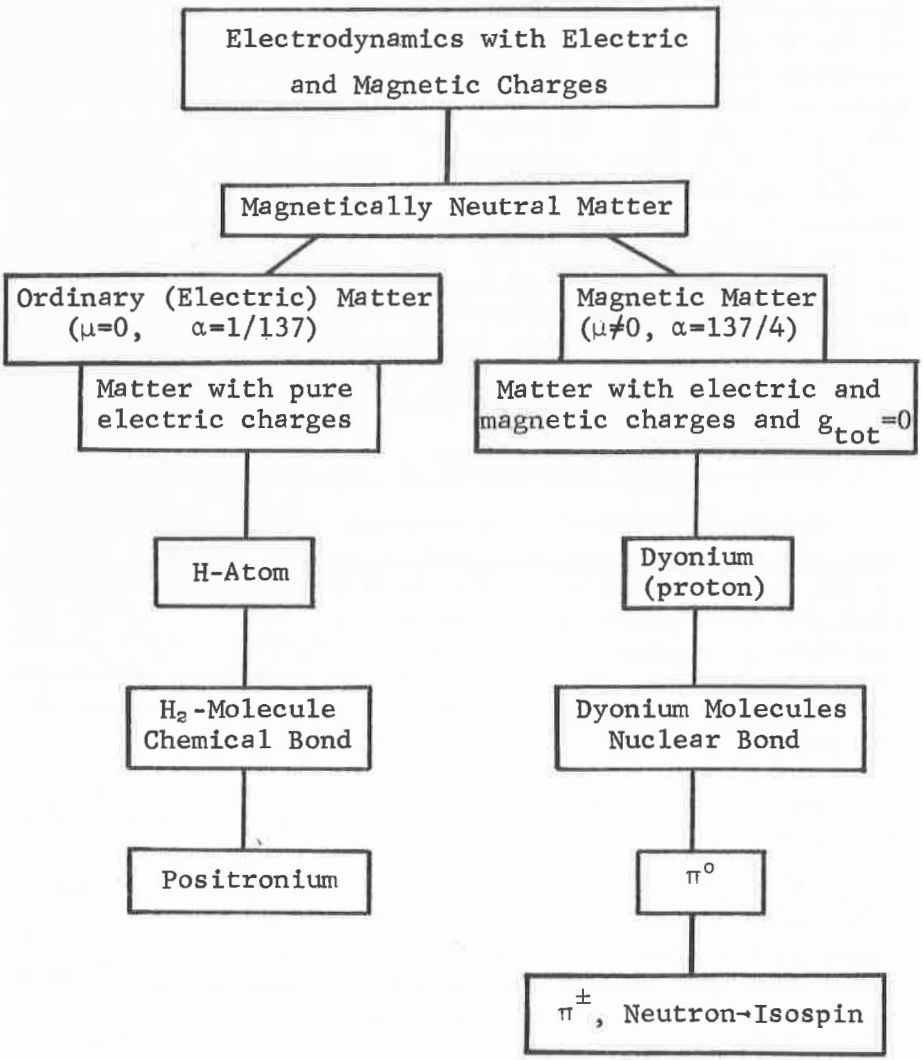
one measures the scattered electron only (so-called inclusive reaction). To lowest order in electromagnetic interactions, one has thus an excitation of the structure of the proton into its excited states, and into possible particle production. The cross-section for this process gives us information on the structure of the proton in the following way: we have seen in previous Section that a conserved current operator j_μ describes the form-factor as well as the excited states of the proton in our model. Because j_μ is known we can evaluate its matrix elements. In the inelastic scattering, the cross-section is related to the products of the matrix elements of the current. Although this process has not been explicitly calculated in the model, it seems on the basis of simple-infinite-component wave equations, that a satisfactory explanation of the experiments is to be expected.[4]

(iii) "Dual" Properties[5] of the Scattering Amplitudes

Consider first e^-e^--scattering. The complete amplitude can be expressed as a sum of photon exchange diagrams in the t-channel (e^+e^-). The resultant amplitude has infinitely many poles in the s-channel, and the amplitude can also be expressed as a sum of infinitely many poles in the s-channel. But we do not sum t-channel exchanges and s-channel poles simultaneously; one can take one or the other. This is a type of duality.

Next consider positronium -positronium scattering. Now in both s- and t-channels we can have positronium intermediate states, as well as photon lines. Thus, we can take, for examples, photon lines in both channels, and obtain poles in both channel or vice versa. Or photon lines in one-channel and poles in the other. This is a quite rich and interesting model to visualize dual amplitudes.[6] In the magnetic charge model of the hadrons we have actually a situation close to positronium case ; the forces are again of Coulomb-type, but with two differences: superstrong coupling constant for the 1/r-potential, and, in addition, the magnetic vector potential which becomes very strong at short distances.

TABLE I


Electrodynamics with Electric and Magnetic Charges
Magnetically Neutral Matter
Ordinary (Electric) Matter (μ=0, α=1/137)
Magnetic Matter (μ≠0, α=137/4)
Matter with pure electric charges
Matter with electric and magnetic charges and $g_{tot}=0$
H-Atom
Dyonium (proton)
H_2-Molecule Chemical Bond
Dyonium Molecules Nuclear Bond
Positronium
π^0
$\pi^{\pm}$, Neutron→Isospin

REFERENCES

1. A discussion of the model in the context and in comparison with other models is given in A.O. Barut, Review of Hadron Symmetries: Multiplets, Supermultiplets and Infinite Multiplets, Rapporteur's Talk given at the XVth International Conference on High Energy Physics, Kiev, 1970, and references therein. Results summarized in Sect. II can be found in detail in Proc. Coral Gables Conference on Fundamental Interactions at High Energies II (Gordon and Breach, 1970, pp. 199-220); in Phys. Rev. D3, 1747 (1971); J. Math. Phys. 12, 841 (1971); in Lectures in Theoretical Physics, Vol. 13 (Assoc. Colorado Univ. Press, 1971); in Topics in Modern Physics: A Tribute to E.U. Condon, (Assoc. Colorado Univ. Press, 1971).pp.15-45.

2. I.H. Duru, Univ. of Colorado Thesis, 1972.

3. See, for example, the review by E.D. Bloom et al, (SLAC-PUB-796), given at the XVth International Conference on High Energy Physics, Kiev 1970.

4. H. Bebié, V. Gorgé and H. Leutwyler, Bern preprint 1971.

5. See Part II. These Proceedings.

6. For more details, see Z.Z. Aydin and A.O. Barut, Coupling of Three or More Positroniums, Univ. of Colorado preprint, 1971.

PARTICLES AS NORMAL MODES OF AN UNDERLYING GAUGE FIELD THEORY*

Fred Cooper
Belfer Graduate School, Yeshiva University
New York, N.Y. 10033

and

Alan Chodos
Department of Physics, University of Pennsylvania
Philadelphia, Pa. 19104

I. Introduction

We would like to suggest that the physical properties of the existing bosons and fermions are well approximated by considering the particles as normal modes of an underlying field theory possessing gauge invariance of the second kind. We envision that at each point in space-time there exists an infinite number of irreducible tensorial fields $\varphi^{\mu_1 \cdots \mu_k}$ and Rarita-Schwinger fields $\psi_\alpha{}^{\mu_1 \cdots \mu_k}$ with bilinear interactions between fields which are nearest neighbors in the Lorentz index space k.[1] The gauge invariance requirements generate these bilinear interactions in a natural way, and the fact that the physical particles are normal modes allows them to have intrinsic structure. In the spirit of the narrow resonance approximation to scattering, we conceive of scattering in a manner analogous to the "rubber band" interpretation of the Veneziano Model.[2] That is, a particle starts out in a given normal mode and absorbs and emits external quanta with corresponding excitation and de-excitation of the underlying field theory. The vertices are obtained by postulating that the external quanta couple to appropriate currents inherent in the Lagrangian; that is, the ρ couples to the vector current, the π to the divergence of

*Work supported in part by the National Science Foundation.

the axial vector current, and the nucleon to the baryonic current. In this way we can give rules for constructing N-point functions in the narrow resonance approximation. For a system consisting of only the π and ρ trajectories, we show it is possible to construct currents obeying the chiral SU2 ⊗ SU2 algebra.

II. Gauge Invariance of the Second Kind as a Generator of Infinite Component Field Theories with Bilinear Interactions.

Consider the massless spin zero Lagrangian

$$L^{(o)} = \tfrac{1}{2}\partial_\mu\varphi\partial^\mu\varphi. \tag{1}$$

We notice that L is invariant under the transformation $\varphi(x)\to\varphi(x)+\alpha$, with α constant. This is gauge invariance of the first kind. Associated with this transformation, there is a conserved current

$$j^\mu = \frac{\delta L^{(o)}}{\delta\partial_\mu\alpha(x)} = \partial^\mu\varphi, \quad \partial_\mu j^\mu = 0 \tag{2}$$

whose integrated fourth component is the generator of the transformation.

Suppose we ask that L is invariant under the transformation $\varphi(x) \to \varphi(x) + \alpha(x)$ (gauge transformation of the second kind). Since $\delta L^{(o)} = \partial_\mu\varphi\partial^\mu\alpha(x)$, we see we need to introduce a new field φ^μ, with bilinear interaction

$$L^{\text{int.}} = m(\partial_\mu\varphi)\varphi^\mu \tag{3}$$

such that

$$\varphi^\mu(x) \to \varphi^\mu(x) - \frac{1}{m}\partial^\mu\alpha(x),$$

$$\varphi \to \varphi + \alpha(x) \tag{4}$$

leaving the quantity

$$G^\mu \equiv \partial^\mu\varphi + m\varphi^\mu \tag{5}$$

invariant.

Thus $L^{(0)'} = \frac{1}{2}(\partial^\mu\varphi + m\varphi^\mu)^2$ is gauge invariant. We notice $L^{(0)'}$ contains a mass term for φ^μ, so we must add a kinetic energy term. If we add an automatically gauge invariant kinetic energy term

$$L^{(1)} = -\tfrac{1}{4}(\partial^\mu\varphi^\nu - \partial^\nu\varphi^\mu)(\partial_\mu\varphi_\nu - \partial_\nu\varphi_\mu) \tag{6}$$

the simple replacement

$$mA^\mu = m\varphi^\mu + \partial^\mu\varphi \tag{7}$$

reduces this theory to an ordinary spin one field theory for a particle of mass m.[3] Also in terms of A^μ the gauge invariance disappears. If we want to maintain gauge invariance, we instead introduce the symmetric kinetic energy term

$$-\tfrac{1}{4}(\partial^\mu\varphi^\nu + \partial^\nu\varphi^\mu)(\partial_\mu\varphi_\nu + \partial_\nu\varphi_\mu) \tag{8}$$

or, so that we have an irreducible representation of the Lorentz group,

$$L^{(1)'} = -\tfrac{1}{4}(\partial^\mu\varphi^\nu + \partial^\nu\varphi^\mu - \tfrac{1}{2}g^{\mu\nu}\partial_\lambda\varphi^\lambda)^2 \equiv -(\delta\partial^\mu\varphi^\nu)^2 \tag{9}$$

where δ is the symmetric traceless projection operator.

Under

$$\varphi^\mu(x) \to \varphi^\mu(x) - \frac{1}{m}\partial^\mu\alpha(x), \tag{10}$$

$$\delta L^{(1)'} = -\frac{2}{m}\partial^\mu\varphi^\nu\delta^{\lambda\sigma}_{\mu\nu}\partial_\lambda\partial_\sigma\alpha(x) \tag{11}$$

which requires the introduction of a new field $\varphi^{\mu\nu}$ such that

$$G^{\mu\nu} \equiv \delta\partial^\mu\varphi^\nu + m_2\varphi^{\mu\nu} \tag{12}$$

is gauge invariant.

Using induction, we generate a Lagrangian with an infinite number of fields $\varphi^{\mu_1\cdots\mu_k} \equiv \varphi^k$:

$$W = \int d^4x\left[\sum_{k=0}^{\infty}\eta_k G^{\nu_1\cdots\nu_k}\left(\delta^{\mu_1\cdots\mu_k}_{\nu_1\cdots\nu_k}\partial_{\mu_k}\varphi_{\mu_1\cdots\mu_{k-1}} + \alpha_k\varphi_{\nu_1\cdots\nu_k}\right)\right.$$

$$-\tfrac{1}{2}\eta_k\, G^{\nu_1\cdots\nu_k}\, G_{\nu_1\cdots\nu_k}\Big]. \tag{13}$$

Here $\eta_k=(-1)^k$, and $\delta^{\mu_1\cdots\mu_k}_{\nu_1\cdots\nu_k}$ is the unit matrix of the $(k/2, k/2)$ representation of the Lorentz group. L is invariant under

$$\varphi_{\mu_1\cdots\mu_k} \to \varphi_{\mu_1\cdots\mu_k} + \gamma_k \delta^{\nu_1\cdots\nu_k}_{\mu_1\cdots\mu_k} \partial_{\nu_1}\cdots\partial_{\nu_k}\Lambda(x)$$

if $\gamma_k + \alpha_{k+1}\gamma_{k+1} = 0$.

This Lagrangian leads to the field equations

$$\Box^2\varphi^k + (k+1)\alpha_k^2\varphi^k = \frac{-k^2}{k+1}\,\delta\partial(\partial\cdot\varphi^k) - (k+1)\alpha_{k+1}\partial\cdot\varphi^{k+1} - \alpha_k\,\delta\partial\varphi^{k-1} \tag{14}$$

where $\delta\partial\varphi_1 = \tfrac{1}{2}\big[\partial_\mu\varphi_\nu + \partial_\nu\varphi_\mu - \tfrac{1}{2}g_{\mu\nu}\partial_\lambda\varphi^\lambda\big]$, etc.

and $\partial\cdot\varphi_\nu = \partial^\mu\varphi_{\mu\nu}$, etc.

We notice that $(k+1)\alpha_k^2$ is the "bare" mass of the field φ^k and that the source of φ^k is the field $\partial\varphi^{k-1}$ and $\partial\cdot\varphi^{k+1}$. Thus gauge invariance produces Lagrangians with an infinite number of fields with nearest neighbor interactions in Lorentz index space.

For fermions we want a gauge invariant Lagrangian of first order which is to be the analog of Eq. (13). To find the appropriate Lagrangian we first replace φ^k by ψ^k_α, (the Rarita-Schwinger field $\psi^{\mu_1\cdots\mu_k}_\alpha$) and consider the quantity

$$G_{k\mu\alpha} \equiv \delta\partial_\mu\psi_{k\alpha} + \bar{\alpha}_k\,\psi_{k\mu\alpha}. \tag{15}$$

(The notation $\psi_{k\mu\alpha}$ means a field with the spinor index α, and with $k+1$ Lorentz indices, k of which are understood and the $(k+1)^{st}$ equal to μ.) Since for fermions γ^μ as well as ∂^μ is a vector we notice that $G_{k\mu\alpha}$ is invariant not only under

$$\psi_{k\alpha} \to \psi_{k\alpha} + \lambda_k \delta\partial_{\mu_1} \cdots \partial_{\mu_k} u_\alpha \tag{16a}$$

but also under

$$\psi_{k\alpha} \to \psi_{k\alpha} + \lambda_k \delta\gamma_{\mu_1} \partial_{\mu_2} \cdots \partial_{\mu_k} u_\alpha \tag{16b}$$

if $k \geq 1$. Now the quantity

$$\delta\gamma_k \equiv \delta^{\mu_1 \cdots \mu_{k+1}}_{\nu_1 \cdots \nu_{k+1}} \gamma^{\nu_{k+1}} \psi^{\nu_1 \cdots \nu_k} \tag{17}$$

is automatically invariant under the second type of gauge transformation of the second kind (Eq. 16b) since $\delta\gamma\gamma A = 0$ ($\gamma_\mu \gamma_\nu$ has no symmetric traceless piece). Therefore the appropriate first order gauge invariant fermion Lagrangian is

$$L = \sum_{k=1}^{\infty} \eta_k \bar{\psi}^k \gamma^\mu \left(i\delta\partial_\mu \psi_k + \bar{\alpha}_k \psi_{k\mu} \right) + \text{h.c.} \tag{18}$$

which leads to the field equations

$$i(\gamma \cdot \delta\partial \psi_k + \partial \cdot \delta\gamma \psi_k) + \bar{\alpha}_k \gamma \cdot \psi_{k+1} + \frac{\eta_{k-1}}{\eta_k} \bar{\alpha}_{k-1} \delta\gamma \psi_{k-1} = 0. \tag{19}$$

For the boson Lagrangian, the gauge invariance tells us that

$$\partial^\mu \left[\partial_\mu \varphi + \alpha_1 \varphi_\mu \right] = 0. \tag{20}$$

This implies that the theory contains no spin-zero particles. Thus the gauge invariant Lagrangian, while able to describe the ρ trajectory, cannot be used to describe the π trajectory. By only requiring that L is invariant under restricted gauge transformation ($\Lambda(x)$ must obey $(\Box^2 + m^2)\Lambda(x) = 0$) we can generalize our Lagrangian to describe the π trajectory. Specifically, we consider

$$L = \sum_{k=0}^{\infty} \left\{ \eta_k \left[G^k \left(\delta\partial \varphi_{k-1} + \alpha_k \varphi_k + \beta_k \partial \cdot \varphi_{k+1} \right) \right] - \tfrac{1}{2} \eta_k G^k G_k \right\} \tag{21}$$

which is invariant under

$$\varphi_k \to \varphi_k + \gamma_k^j \, \delta\partial^{(k-j)} V_j(x)$$

if $$(\Box^2 + m_j^2) \, V_j(x) = 0, \quad \partial \cdot V^j = 0 \tag{22}$$

and $$\frac{-m_j^2}{2} \frac{(k-j+1)(k+j+2)}{(k+1)^2} \beta_k \gamma_{k+1}^j + \alpha_k \gamma_k^j + \gamma_{k-1}^j = 0.$$

For fermions, we can construct a similar generalization:

$$L = \sum_{k=1}^{\infty} \eta_k \bar{\psi}^k \gamma^\mu (i\delta\partial_\mu \psi_k + \bar{\alpha}_k \psi_{k\mu} + i\bar{\beta}_k \partial^\lambda \psi_{k\lambda\mu}) + \text{h.c.} \tag{23}$$

which is invariant under the transformation

$$\psi_k \to \psi_k + \lambda_k^j \delta\gamma\partial^{(k-j-1)} u_j(x)$$

if $$(i\gamma \cdot \partial - m_j) u_j = 0, \quad \gamma \cdot u_j = \partial \cdot u_j = 0 \tag{24}$$

and $$2(k+2)(i\lambda_k^j + \bar{\alpha}_k \lambda_{k+1}^j) - i m_j^2 \, \bar{\beta}_k \lambda_{k+2}^j \frac{(k+1-j)(k+3)(k+j+3)}{(k+2)^2} = 0.$$

The boson Lagrangian leads to the field equations

$$\partial \cdot G^{k+1} + \left(\frac{\eta_{k-1}}{\eta_{k+1}} \right) \beta_{k-1} \delta\partial G^{k-1} = \left(\frac{\eta_k}{\eta_{k+1}} \right) \alpha_k G^k \tag{25}$$

$$G^k = \delta\partial \varphi^{k-1} + \alpha_k \varphi^k + \beta_k \partial \cdot \varphi^{k+1} .$$

The fermion Lagrangian leads to the field equations

$$i\gamma \cdot \delta\partial \psi_k + i\partial \cdot \delta\gamma \psi_k + \alpha_k \gamma \cdot \psi_{k+1} + i\beta_k \partial \cdot \gamma \psi_{k+2}$$

$$+ \frac{\eta_{k-1}}{\eta_k} \bar{\alpha}^*_{k-1} \delta\gamma\psi_{k-1} + i \frac{\eta_{k-2}}{\eta_k} \bar{\beta}_{k-2} \delta\partial\gamma\psi_{k-2} = 0. \qquad (26)$$

III. Mass Spectrum

In order to determine the mass spectrum we assume the existence of a set of normal mode free fields $\tilde{\varphi}^j_N$ of mass m_N, ($\tilde{\psi}^j_{\alpha N}$ for fermions), and spins j (j+½). These fields obey

$$(\Box^2 + m^2_{Nj})\, \tilde{\varphi}_{Nj} = 0 \; ; \; \partial\cdot\tilde{\varphi}_{Nj} = 0,$$
$$(i\gamma\cdot\partial - m_{Nj})\, \tilde{\psi}_{Nj} = 0 \; ; \; \partial\cdot\tilde{\psi}_{Nj} = \gamma\cdot\tilde{\psi}_{Nj} = 0. \qquad (27)$$

The underlying fields have the following decomposition in terms of the normal modes

$$\varphi^k = \sum_{N,j} a^{(Nj)}_k \delta\partial^{(k-j)} \tilde{\varphi}^j_N \; ; \; G^k = \sum_{N,j} b^{(Nj)}_k \delta\partial^{(k-j)} \tilde{\varphi}^j_N$$
$$\psi^{k\alpha} = \sum_{N,j} \left\{ \bar{a}^{(Nj)}_k \delta\partial^{(k-j)} \tilde{\psi}^{\alpha j}_N + \bar{b}^{(Nj)}_k \delta\gamma\partial^{(k-j-1)} \tilde{\psi}^{\alpha j}_N \right\}. \qquad (28)$$

The $a^{(Nj)}_k$ are related to the wave functions; $a^{(Nj)}_k$ and $\bar{b}k^{(Nj)}$ obey second order inhomogeneous difference equations, whereas $b_k^{(Nj)}$ and $\bar{a}_k^{(Nj)}$ obey second order homogeneous equations. The allowed wave function must be normalizeable, and this leads to the eigenvalue equation of the masses m_{N_j} of the normal modes. The homogeneous equations are

$$-\frac{m^2}{2}{}_{Nj} \frac{(k+j+2)(k-j+1)}{(k+1)^2} b^{(Nj)}_{k+1} \eta_{k+1} = -\eta_{k-1}\beta_{k-1} b^{(Nj)}_{k-1} + \eta_k \alpha_k b^{(Nj)}_k$$

$$+\frac{m^2 N_j}{2}\frac{(k-j)(k+j+2)(k+2)}{(k+1)^3}\bar{G}_k^{(Nj)} \qquad (29)$$

$$+i\bar{\alpha}^*_{k-1}\bar{G}^{(Nj)}_{k-1}\left(\frac{\eta_{k-1}}{\eta_k}\right)-\left(\frac{\eta_{k-2}}{\eta_k}\right)\bar{\beta}_{k-2}\bar{G}^{(Nj)}_{k-2} = 0.$$

We notice that although the fermion wave equation is first order, the mass enters the eigenvalue equation as m^2. Thus the gauge invariant Lagrangian automatically leads to Fermion Regge trajectories which are a function of m^2 and not m, in agreement with experiment.

By a rescaling of $\bar{a}_k$ and b_k we can convert both equations to almost identical difference equations. Letting

$$\eta_k b_k^{(Nj)} = B_k^j \tilde{b}_k^{(Nj)}, \quad \eta_k \bar{G}_k^{(Nj)} = A_k^j \gamma_{k+1}^{(Nj)} \qquad (30)$$

with $\beta_{k-1}\frac{B_{k-1}}{B_{k+1}} = -\frac{k^2}{2}$, $\bar{\beta}_{k-2}\left(\frac{k+1}{k+2}\right)\frac{A_{k-2}}{A_k} = -\frac{k^2}{2}$,

we obtain

$$-\frac{m^2 N_j}{2}\frac{(k-j+1)(k+j+2)}{(k+1)^2}\tilde{b}^j_{k+1} = \frac{k^2}{2}\tilde{b}^j_{k-1} - m_k\tilde{b}^j_k \qquad (31a)$$

$$-\frac{m^2 N_j}{2}\frac{(k-j)(k+j+2)}{(k+1)^2}\gamma^{(Nj)}_{k+1} = \frac{k^2}{2}\gamma^{(Nj)}_{k-1} + m_k\gamma^{(Nj)}_k \qquad (31b)$$

where

$$m_k m_{k-1} \equiv M_k = \begin{cases} -\frac{k^2}{2}\frac{\alpha_k\alpha_{k-1}}{\beta_{k-1}} & \text{in (31a)} \\ \frac{k^3}{k+1}\frac{\bar{\alpha}_{k-1}\bar{\alpha}_{k-2}}{2\bar{\beta}_{k-2}} & \text{in (31b)} \end{cases}$$

In reference 4 Chodos and Haymaker have shown that the choice $M_k \sim k^3$ as $k\to\infty$ leads to asymptotically linearly rising Regge trajectories.

The boundary conditions on the difference equation which lead to the eigenvalue condition are that

$$b_k^{(Nj)} = 0 \quad \text{for } k < j \quad \text{(origin condition)}$$

and that the wave function is normalizable, which restricts the behavior of $b_k^{(Nj)}$ as $k \to \infty$. Specifically, using the electric charge matrix element as the norm we have for the boson field theory

$$\langle p' Nj\sigma | Q | pNj\sigma \rangle = 2p_o \delta^3(p-p') \tag{32}$$

which implies

$$1 = \sum_{k=j}^{\infty} \frac{(-1)^{k+j+1}}{2^{k-j+1}} \frac{(j!)^2}{(2j+1)!} \frac{(k+j+2)!}{[(k+1)!]^2 (k-j+1)!} \left(b_{k+1}^{(Nj)} G_k^{(Nj)} + \beta_k b_k^{(Nj)} G_{k+1}^{(Nj)} \right) m^{2(k-j)} .$$

For this series to converge, one must take the faster dying of the two possible asymptotic behaviors (large k) of eqs. (31a,b). If one chooses

$$M_{nj_o}^{-1} = \frac{\gamma}{(n+j_o+a)(n+j_o+a-1)(n+2j_o+1)} \tag{33}$$

for a particular value j_o of the spin, then the difference equations (31a,b) become analogous to the Laguerre polynomial differential equation, and one obtains the eigenvalue condition

$$m^2_{Nj_o} = \frac{4}{\gamma} (N+j_o+a) \tag{34}$$

Thus we find that the spin j_o satellites are evenly spaced in m^2. The mass spectrum depends only on M_{nj_o} which is a function only of the ratio $\frac{\alpha_k \alpha_{k-1}}{\beta_{k-1}}$.

IV. Form Factors

When we write the normal mode fields in terms of the gauge fields, we find the expression contains terms having an infinite number of derivatives acting on the field of spin $j \to \infty$. Thus the normal mode fields are not localized. This is most easily seen by calculating the matrix element of the electromagnetic current between two physical particle states and showing it is an infinite power series in q^2. For simplicity we restrict ourselves to the boson field theory. Letting φ become complex so it describes charged mesons, we have:

$$L = \tfrac{1}{2}\sum_{k=0}^{\infty} \eta_k \left[G^{k+}(\delta\partial\varphi_{k-1} + \alpha_k \varphi_k + \beta_k \partial\cdot\varphi_{k+1}) + \text{h.c.} - G_k^+ G^k \right]. \tag{35}$$

To construct a conserved vector current we consider the phase transformations

$$\begin{aligned} \varphi_k &\to e^{-i\alpha}\varphi_k & G_k &\to e^{-i\alpha} G_k \\ \varphi_k^+ &\to e^{i\alpha}\varphi_k^+ & G_k^+ &\to e^{i\alpha} G_k^+ \end{aligned} \tag{36}$$

The Gell-Mann-Levy equations are

$$j_\mu = \frac{\delta L}{\delta \partial_\mu \alpha} \quad ; \quad \partial_\mu j^\mu = \frac{\delta L}{\delta \alpha} = 0$$

with

$$j^\mu(x) = \frac{i}{2}\sum_{k=0}^{\infty} \eta_k \left\{ G^{+k\mu}\varphi_k + \beta_k \varphi^{+k\mu} G_k - \varphi_k^+ G^{k\mu} - \beta_k G_k^+ \varphi^{k\mu} \right\} \tag{37}$$

Sandwiching $j^\mu(x)$ between the $j = 0$ physical particle state $|N\ j{=}0\ p\rangle$ and using the decompositions of Eq. (28), we obtain:

$$\langle N\ j{=}0\ p' | j^\mu(x) | N\ j{=}0\ p\rangle = \frac{1}{(2\pi)^3}(p^\mu + p'^\mu)\ e^{i(p'-p)\cdot x}$$

$$\times \sum_{k=0}^{\infty} \eta_k \frac{m_N^{2k}}{2^k} \binom{k+2}{k+1} [(k+1)P_{k+1}(\gamma) + P'_{k+1}(\gamma)(1-\gamma)]$$
$$\left(b_{k+1}^{(N)} a_k^{(N)} + \beta_k a_{k+1}^{(N)} b_k^{(N)}\right)$$

where

$$\gamma = \frac{p\cdot p'}{mm'} = \cosh\theta, \quad P_k(\gamma) = \frac{\sinh(k+1)\theta}{(k+1)\sinh\theta} \tag{38}$$

and

$$a_k^{(N=0)} = \left(\frac{1}{k+1}\right) \frac{\lambda}{\Gamma(k+\alpha)} \left(\prod_{\ell=0}^{k-1} \frac{\alpha_\ell}{\beta_\ell}\right) \sum_{n=1}^{k} \left(\prod_{\ell=0}^{n-1} \beta_\ell\right);$$

$$b_k^{(N=0)} = \left(\frac{\gamma}{2}\right)^k \frac{\lambda}{(k+1)\Gamma(k+\alpha)} \left(\prod_{\ell=0}^{k-1} \alpha_\ell\right).$$

We notice that $F(q^2)$ is a power series in $q^2=2m^2-2p\cdot p'$ whose coefficients depend on α_ℓ as well as β_ℓ and not in the ratio $\frac{\alpha_k}{\beta_{k-1}}\alpha_{k-1}$ which determined the mass spectra. Thus we can choose the form factor independently of the mass spectra and hopefully there is a simple choice of α_k and β_k which leads to a dipole form factor as well as a linearly rising Regge trajectory. The expressions for the transition form factors are found in the second of reference (3).

V. Electromagnetic Interactions

We treat electromagnetic interactions by considering the photon as an external field which merely causes transitions from one state (normal mode) of the underlying field to another, in the process exciting and de-exciting all the normal modes. This excitation gives an infinite number of narrow resonance poles in the s and u channels, and leads to the narrow resonance approximation to Compton scattering (elastic or quaielastic, on or off shell). We assume that the external electromagnetic field couples to the usual current generated by the minimal coupling hypothesis.

For the boson Lagragian letting $\varphi \to e^{-i\alpha(x)}\varphi$, and $\varphi^+ \to \varphi^+ e^{i\alpha(x)}$ we find

$$j^\mu(x) = \frac{i}{2}\sum_{k=0}^{\infty} \eta_k (G^{+k\mu}\varphi_k + \beta_k \varphi^{+k\mu} G_k) \qquad (39)$$

$$+\text{h.c.} + \text{seagull terms}$$

For fermions, letting $\psi \to e^{-i\alpha}\psi$, $\bar{\psi} \to \bar{\psi}e^{i\alpha}$ yields

$$j^\mu(x) = \sum_{k=1}^{\infty} \eta_k \left[\delta(\bar{\psi}^k \gamma^\mu)\psi_k + \beta_k \bar{\psi}_k \gamma_\lambda \psi^{k\lambda\mu}\right] + \text{h.c.} . \qquad (40)$$

Using as the effective Hamiltonian

$$H_I = e\int j_\mu(x)\, A^{\mu\ \text{ext}}\, d^3x \qquad (41)$$

we find that the expression for Compton scattering to order e^2 is

$$\langle \alpha(k',\lambda')N'j's'p';\text{out}|\gamma(k,\lambda)Njsp;\text{in}\rangle = \delta_{fi}$$

$$+ \frac{i^2}{2} e^2\ \epsilon^\mu(k,\lambda)\ \epsilon^{\nu *}(k',\lambda')\int d^4x d^4y e^{-i(k\cdot x - k'\cdot y)}$$

$$\times \langle N'j's'p'|T\left(j_\mu(x)\ j_\nu(y)\right)|Njsp\rangle . \qquad (42)$$

Using the decompositions of Equation (28) we find that there are an infinite number of "narrow resonances" in the s and u channel corresponding to figure 1.

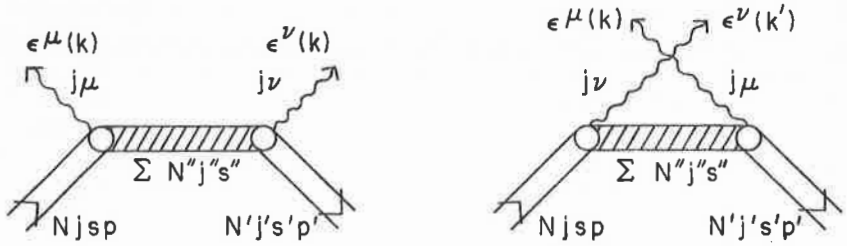


Figure 1.

Since we know the diagonal and off-diagonal $j^{\mu}(x)$ for all q^2, we can also calculate $W_{\mu\nu}$ (and therefore the structure functions W_1 and νW_2 using the fermion current of equation (40):

$$W_{\mu\nu} = \tfrac{1}{2}\sum_{s}\sum_{N'} \langle Njsp|j_{\mu}(o)|N'j's'p_{N'}\rangle \langle N'j's'p_{N'}|j_{\nu}(o)|Njsp\rangle$$

$$\times(2\pi)^3\ \delta^4\ (p+q-p_{N'}). \qquad (43)$$

Since it is likely that we can choose $\bar{\alpha}_k$ and $\bar{\beta}_k$ so that $F(q^2) \sim_q {}^{-4}$ and the mass spectrum is linearly rising, we should be able to reproduce the successes of Domokos and Schoenberg[5] in obtaining scaling; however here we have a real Lagrangian field theory giving this result.

VI. Strong Interaction Dynamics.

We handle strong interactions similarly to electromagnetic interactions, in that we use external quanta to excite the underlying field, with the external quanta coupling to the various currents inherent in the Lagrangian. The prescription for scattering is to treat any one particle as the underlying field in a normal mode, absorb and emit external quanta consistent with the scattering process, with the underlying field being excited and de-excited, finally returning to a normal mode state. To obtain crossing symmetry, we sum over the various ways of choosing each particle as the underlying field in a normal mode. We postulate that the external quanta couples to a current generated from the Lagrangian. For example an external ρ couples to the isospin current, an external π to the divergence of the axial current, an external nucleon to the baryonic current. In order to insure that the whole trajectory is exchanged, these currents must be bilinear in the infinite component fields. To get an idea of the structure of these currents, we look at the trilinear couplings of ordinary field theory, treat each particle as the external particle and try to generalize the resulting bilinear object to the infinite component case. In particular cases, such as the vector and axial vector currents, algebraic constraints (such as chiral SU2 $\otimes$ SU2) tell us what transformations to make on the Lagrangian to obtain

the appropriate infinite component generalizations of these bilinear currents.

Let us first turn to a simple system to illustrate these ideas. Consider a world where there are only the π and ρ trajectories; $\pi^+\rho^-$ scattering in that world is approximated by ρ trajectory exchange in the t channel and the π trajectory in the s channel. In an ordinary field theory, the s channel pole term requires knowledge of the $\rho\pi\pi$ vertex, which is given by

$$g_{\rho\pi\pi}\epsilon^{abc}\rho^{\mu}{}_{a}\varphi_b\partial_\mu\varphi_c \tag{44}$$

Saying that the external ρ couples to the isospin current implies that

$$H_I=\rho^{ext}_{\mu a}j^{\mu a} \quad ; \quad j^{\mu a}=\epsilon^{abc}\varphi_b\partial^{\mu}\varphi_c . \tag{45}$$

Saying the external π couples to the divergence of the axial current is equivalent to

$$H_I=-\partial_\mu\varphi^{ext}_c A^{\mu c} \quad ; A^{\mu c}=-\epsilon^{abc}\rho^{\mu}{}_{a}\varphi_b . \tag{46}$$

To obtain an expression for $\pi^+\rho^-$ scattering in the narrow resonance approximation we first have to write down a Lagrangian which will yield a set of normal modes with quantum numbers, masses and form factors appropriate to the particles on the π and ρ trajectories. Then we will have to generate currents which are the appropriate generalizations of the $V^{\mu i}$ and $A^{\mu i}$ of the equations (45-6). Finally, to calculate the scattering amplitude, we must treat each particle in $\pi^+\rho^- \to \pi^+\rho^-$ as the underlying field in the appropriate normal mode and absorb and emit external π's and ρ's with the effective interactions given by $H_I= -\partial_\mu\varphi_a^{ext} A^{\mu a}$ and $H_I=\rho^{\mu ext}_a V^a_\mu$. This produces the generalized Feynman graphs of Figure 2.

To describe the family of particles on the π and ρ trajectories, we use the fields $\varphi^k_{a\eta}$, where a is the isospin of the trajectory (a= 1,2,3) and $\eta(=0,1)$ is the normality of the trajectory (parity = $(-1)^{j+\eta}$) and is zero for the ρ trajectory and 1 for the π trajectory. The

ensuing Lagrangian that describes both the π and ρ trajectories is

$$L = \sum_{k=0}^{\infty} \eta_k G^k{}_{a\eta} \left[\left(\delta \partial \varphi^{a\eta}_{k-1} + \alpha^{\eta}_k \varphi^{a\eta}_k + \beta^{\eta}_k \partial \cdot \varphi^{a\eta}_{k+1} \right) - \tfrac{1}{2} G^{a\eta}_k \right]. \tag{47}$$

We probably want to choose $\beta_k^{\eta=0} = 0$ to descibe the ρ trajectory. We construct currents in the usual way. For the vector current we let

$$\varphi_a \to \varphi_a + \epsilon_{abc} \Lambda_b \varphi_c \tag{48}$$

$$V^{\mu}_a = \frac{\delta L}{\delta \partial_{\mu} \Lambda^a} = -\sum_{k=0}^{\infty} \eta_k \left(G^{k-1\mu} \times \varphi_{k-1} + \beta_k G_k \times \varphi^{k\mu} \right)_a . \tag{49}$$

Similarly letting

$$\varphi_a \to \varphi_a + \epsilon_{abc} \Lambda_b \epsilon \varphi_c$$

with $\epsilon_{\eta\eta_1} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ exchanging π and ρ fields we find

$$A^{\mu}_a = -\sum_{k=0}^{\infty} \eta_k \left\{ G^{k-1\mu} \epsilon \times \varphi_{k-1} + \beta^{\eta}_k G^{\eta}_k \epsilon_{\eta\eta'} \times \varphi^{k\mu}_{\eta'} \right\}_a . \tag{51}$$

To demonstrate that these currents satisfy the chiral SU2 $\otimes$ SU2 slgebra, one merely notices that

$$V^a_o = \sum \left(\Pi^k \times \varphi_k \right)_a \tag{52a}$$

$$A^a_o = \sum \left(\Pi^k \epsilon \times \varphi_k \right)_a \tag{52b}$$

where π^k is the momentum canonical to φ_k.

Thus $\left[V^a_o(x) , V^b_o(y) \right]_{x_o = y_o} = i \epsilon^{abc} V_o{}^c \delta^3(x-y)$ (53a)

$$\left[V_o^a(x)\ ,\ A_o^b(y)\right]_{x_o=y_o} = i\epsilon^{abc}A_o^c\delta^3(x-y) \qquad (53b)$$

$$\left[A_o^a(x)\ ,\ A_o^b(y)\right]_{x_o=y_o} = i\epsilon^{abc}V_o^c\delta^3(x-y). \qquad (53c)$$

Using our rules for scattering, if we want to do

$$\pi^+(p_1) + \rho^-(p_2) \to \pi^+(p_3) + \rho^-(p_4)$$

we first let the underlying field start in the $\pi^+(p_1)$ state and draw all possible diagrams. We get the four diagrams of Figure 2A. Figure 2.2 for example is given by

$$\frac{\epsilon^\mu(p_2)}{(2\pi)^3\sqrt{2p_{2o}2p_{3o}}}\int d^4x d^4y\ e^{-i(p_2\cdot x - p_3\cdot y)}$$

$$\times\langle\rho^-(p_4),\ \eta=0, j=1, \lambda\,|T\left(V_\mu^{(-)}(x), \partial_\nu A^{\nu(-)}(y)\right)|\pi^+(p_1), \eta=1, j=0\rangle. \qquad (54)$$

The matrix elements of V^μ and A^μ are generalizations of equation (38) and are to be found in reference 3b. Choosing the other particles as the underlying field in a normal mode, we get four more diagrams - Fig. 2.5 to 2.8. We notice there are four s-channel and four t-channel diagrams. If at the 8-point function level we impose factorization, this will probably tell us that not all diagrams are independent, and a careful counting scheme will be needed. It is not difficult to see that all the graphs except for the triple Regge vertex can be generated for the N-point function.

π-N Scattering

In this framework, unlike the Veneziano model, the problem of quantum numbers presents little or no difficulty. Thus it is just as easy to calculate π-P elastic scattering (or π+P → ρ+P etc.) as it was to do $\pi^-\rho^+$ scattering.

For simplicity let us restrict ourselves to a world

where there are only the π, ρ and Nucleon trajectories. A Lagrangian describing the particles in that system is given by

$$L = L^{\pi\rho} + \sum_{\substack{k=1 \\ i=1,2}}^{\infty} \eta_k \bar{\psi}_i \gamma^\mu (i\delta\partial_\mu \psi^i_k + \bar{\alpha}_k \psi^i_{k\mu} + i\bar{\beta}_k \partial^\lambda \psi^i_{k\lambda\mu}) + \text{h.c.} \qquad (55)$$

where i = 1,2 corresponds to the isospin index of the nucleon . $L^{\pi\rho}$ is given by equation (47). The Lagrangian, Eq. (55) determines the masses and electromagnetic form factors of the particles; it also contains the currents which are the generalization of the pole term currents. Once we have the currents, we can calculate N point functions from our set of rules.

The field theory pole terms are the neutron pole in the s channel and the p^o pole in the t channel. We interpret these pole terms two ways and generalize:

$$\bar{N}\gamma_\mu\gamma_5 \frac{\vec{\tau}}{2} N \cdot \partial^\mu \vec{\varphi} \Rightarrow \vec{A}^\mu \cdot \partial_\mu \vec{\Phi} \qquad (56a)$$

$$\text{or } (\bar{N}\gamma^\mu)_\alpha j^{\mu\alpha}_{\eta=1} \qquad (56b)$$

$$\bar{N}\gamma^\mu \frac{\vec{\tau}}{2} N \cdot \vec{\rho}_\mu \Rightarrow \vec{V}^\mu \cdot \vec{\rho}_\mu \qquad (57a)$$

$$\text{or } (\bar{N}\gamma_\mu) j^{\mu\alpha}_{\eta=0} \qquad (57b)$$

$$(\vec{\rho}^\mu \times \vec{\varphi}) \cdot \partial_\mu \vec{\varphi} \Rightarrow \vec{\rho}_\mu \cdot \vec{V}^\mu \qquad (58a)$$

$$\text{or } \partial_\mu \vec{\varphi} \cdot \vec{A}^\mu \qquad (58b)$$

To obtain the infinite component generalization of the nucleon contribution to the vector and axial vector currents we consider the transformations

$$\psi \to e^{-i\vec{\alpha}\cdot\vec{\tau}/2}\psi \quad ; \quad \bar{\psi} \to \bar{\psi}\, e^{i\vec{\alpha}\cdot\vec{\tau}/2}$$

and $\psi \to e^{-i\vec{\alpha}\cdot\vec{\tau}/2\gamma_5}$; $\bar{\psi} \to \bar{\psi}\, e^{-i\vec{\alpha}\cdot\vec{\tau}/2\gamma_5}$ (59)

and find

$$V_a^\mu = \sum_k \frac{\eta_k}{2}\left[\delta(\bar{\psi}^k\gamma^\mu)\frac{\tau^a}{2}\psi_k + \bar{\beta}_k\bar{\psi}_k\gamma_\lambda\frac{\tau^a}{2}\psi^{k\lambda\mu}\right] + \text{h.c.}$$

$$A_a^\mu = \sum_k \frac{\eta_k}{2}\left[\delta(\bar{\psi}^k\gamma^\mu)\gamma_5\frac{\tau^a}{2}\psi_k + \bar{\beta}_k\bar{\psi}_k\gamma_\lambda\gamma_5\frac{\tau^a}{2}\psi^{k\lambda\mu}\right] + \text{h.c.} \tag{60}$$

These are the simple generalizations of the usual vector and axial vector currents, and obey the chiral

To obtain a simple generalization of

$$j_{\mu\alpha} = g_1(\gamma_5\tfrac{\tau}{2}p)_\alpha\partial_\mu\varphi + g_0\left(\tfrac{\tau}{2}p\right)_\alpha\rho_\mu \tag{61}$$

we consider the set of transformations

$$\varphi_\eta^i(x) \to \varphi_\eta^i(x) + \bar{f}(x)\lambda_\eta\frac{\tau^i}{2}\psi(x) + \bar{\psi}(x)\lambda_\eta\frac{\tau^i}{2}f(x)$$

($\varphi^i(x)$ is hermitean).

$$\psi(x) \to \psi(x) + i\lambda_\eta\frac{\tau^i}{2}f(x)\varphi_{i\eta}$$

$$\bar{\psi}(x) \to \bar{\psi}(x) - i\,\bar{f}(x)\lambda_\eta\frac{\tau^i}{2}\varphi_{i\eta} \tag{62}$$

here $f(x)$ is a spinor in isospin and Lorentz space and λ_η is a Dirac matrix defined to be $\begin{cases} g_0 I & \text{if } \eta = 0 \\ g_1\gamma_5 & \text{if } \eta = 1. \end{cases}$

These transformations lead to the following expression for the baryonic current:

$$j_\alpha^\mu = \frac{\delta L}{\delta\partial_\mu\bar{f}_\alpha(x)} = \sum_k \eta_k\left\{\vec{G}^{k-1\mu}\cdot\left(\frac{\vec{\tau}}{2}\lambda_\eta\psi_{k-1}\right)_\alpha + \beta_k\vec{G}_k\cdot\left(\frac{\vec{\tau}}{2}\lambda_\eta\psi^{k\mu}\right)_\alpha\right.$$

$$+\vec{\varphi}_k \cdot \left(\frac{\vec{\tau}}{2} \lambda_\eta \delta(\gamma^\mu \psi^k)\right)_\alpha + \bar{\beta}_k \vec{\varphi}^{k\mu\lambda} \cdot \left(\frac{\vec{\tau}}{2} \lambda_\eta \delta(\gamma_\lambda \psi_k)\right)_\alpha \Big\}. \qquad (63)$$

This is a natural generalization of Eqs. (56b, 57b).

Now that we have the appropriate vector, axial vector and baryonic currents, we can calculate π-p $\rightarrow$ π-p scattering in the narrow resonance approximation in the world where there is only the nucleon trajectory in the s channel and the ρ trajectory in the t channel. Specifically we get the diagrams of Figure 3. (Dashed line is external π, coupling to A^μ and solid single line is external nucleon coupling to $j^{\mu\alpha}$, the baryonic current).

For example, Figure 3b is given by the expression

$$\left[\frac{\bar{u}(p_2)\gamma^\mu}{(2\pi)^3}\right]_\alpha \int d^4x d^4y \, \frac{e^{-i(k_1 x - p_2 y)}}{\sqrt{2k_{1o}}\sqrt{p_{2o}/m}}$$

$$x\langle \pi^-(k_2) N=0, j=0, \eta=1 \,|\, T\left(\partial_\lambda A^\lambda(x), j_{\mu\alpha}(y)\right) |\, n(p_1), N=0, j=\tfrac{1}{2}, \lambda \rangle. \qquad (64)$$

We hope in the future to find out which α_k, β_k lead to dipole-like form factors, and then evaluate these expressions numerically.

Figure Captions

Fig. 1. Compton scattering with the exchange of a Regge trajectory.

Fig. 2. Diagrams contributing to $\pi^+\rho^-$ scattering. Two particles (dashed line for the pion, wavy line for the ρ) are considered to be external quanta, while the other two are projected out from the appropriate currents. The shaded line denotes a propagating sum of normal modes; the value of η is explicitly given for each diagram.

Fig. 3. Diagrams contributing to $\pi^- p$ elastic scattering. Dashed line is external π^- coupling to A^μ, and solid single line is external proton coupling to the baryonic current $j^{\mu\alpha}$.

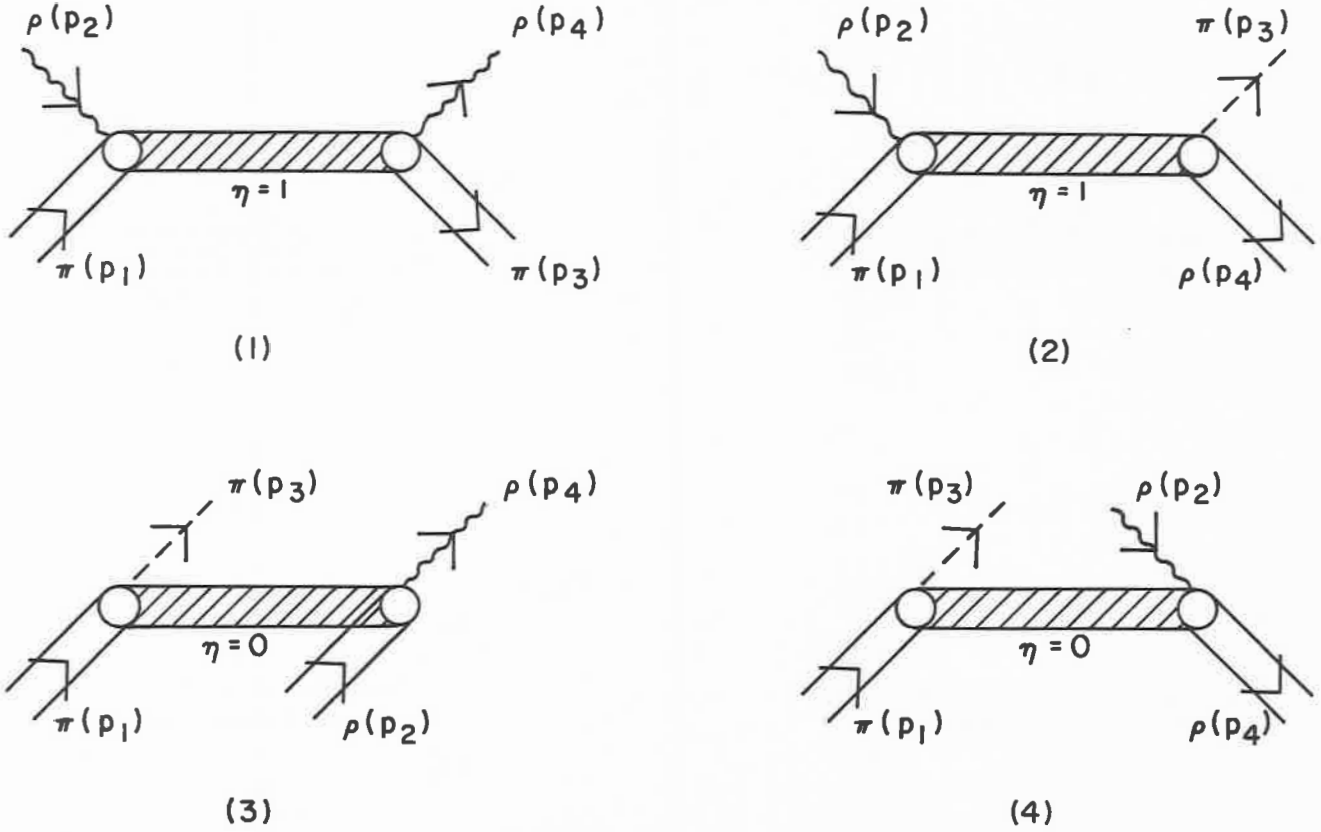

ρ(p2)
ρ(p4)
η = 1
π(p1)
π(p3)
(1)
ρ(p2)
π(p3)
η = 1
π(p1)
ρ(p4)
(2)
π(p3)
ρ(p4)
η = 0
π(p1)
ρ(p2)
(3)
π(p3)
ρ(p2)
η = 0
π(p1)
ρ(p4)
(4)

Figure 2A

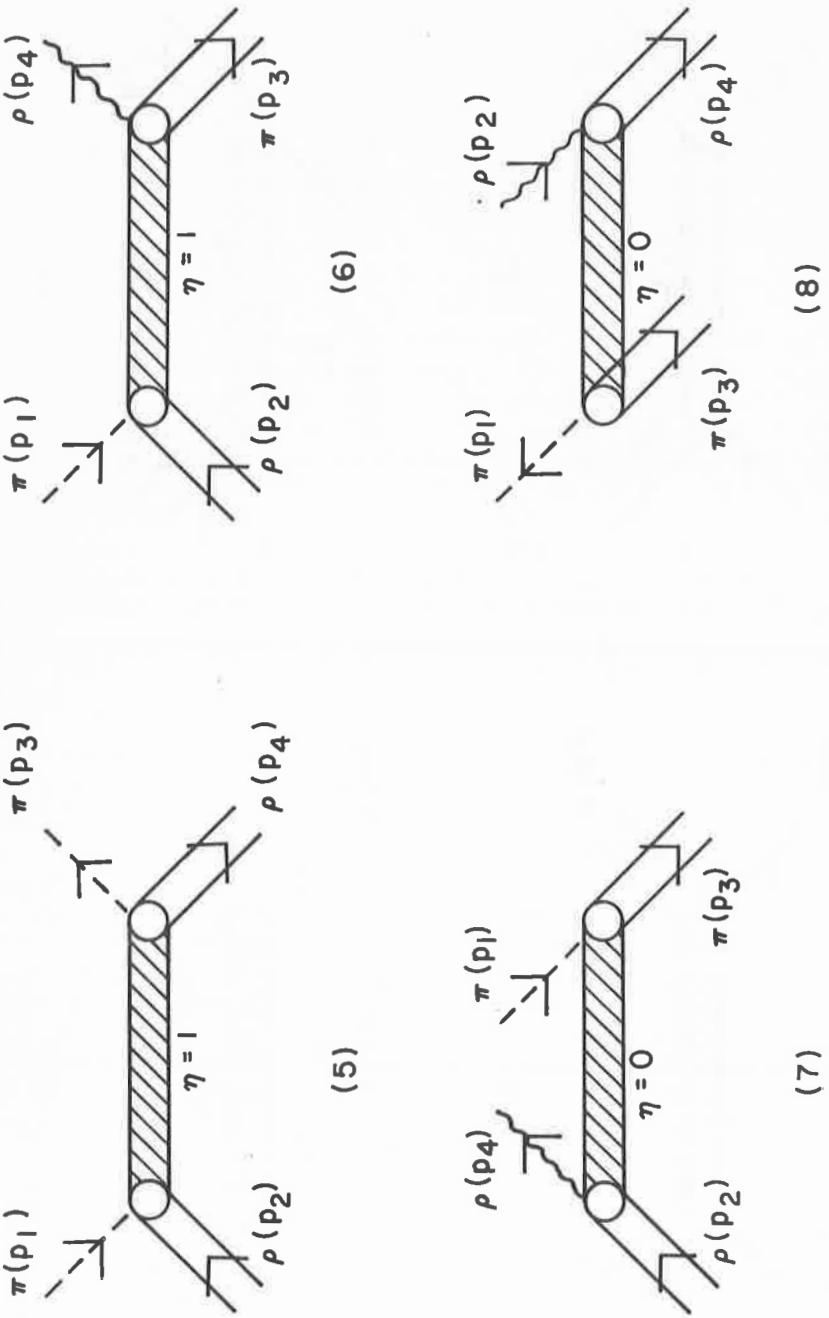

ρ(p4)
π(p3)
η = 1
π(p1)
ρ(p2)
(6)
ρ(p2)
ρ(p4)
η = 0
π(p1)
π(p3)
(8)
π(p3)
ρ(p4)
η = 1
π(p1)
ρ(p2)
(5)
π(p1)
π(p3)
η = 0
ρ(p4)
ρ(p2)
(7)

Figure 2B

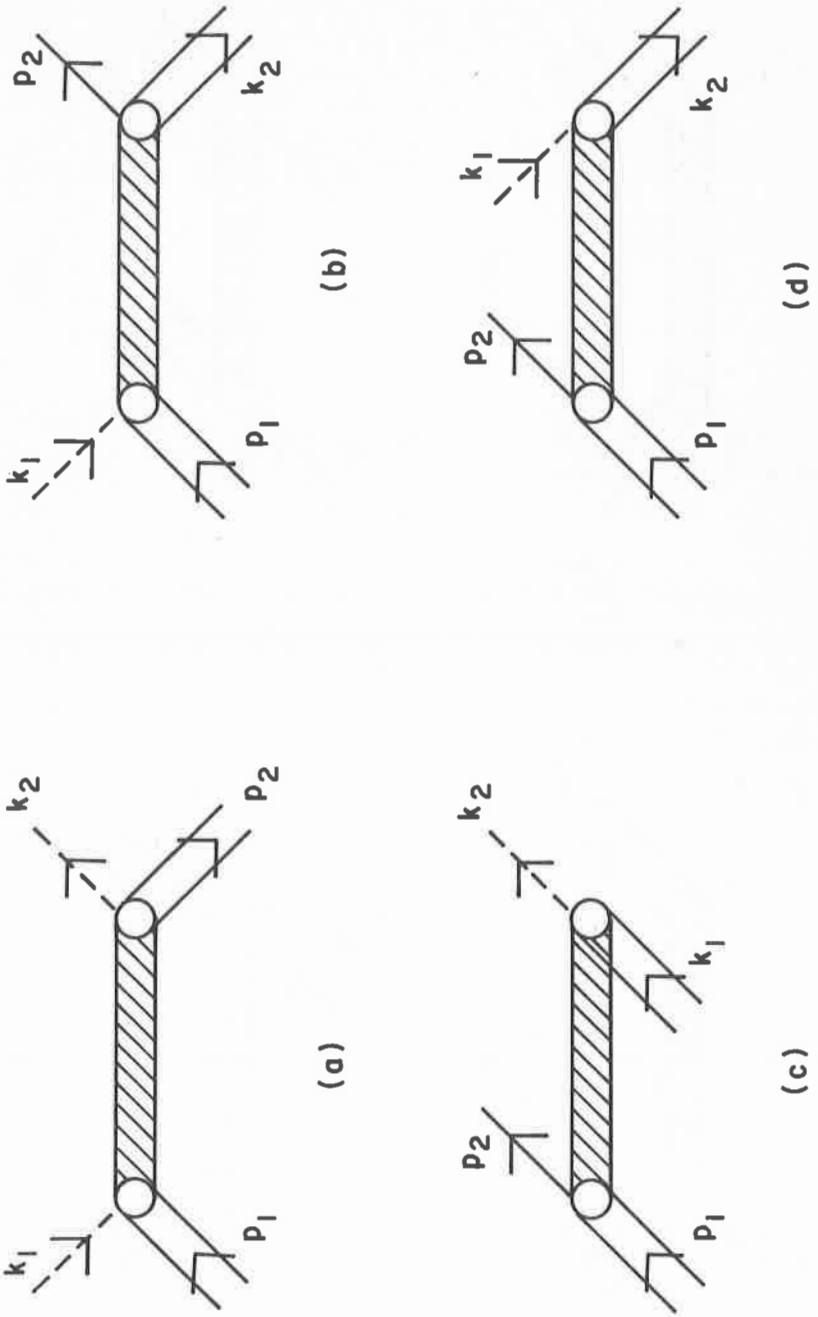


Figure 3

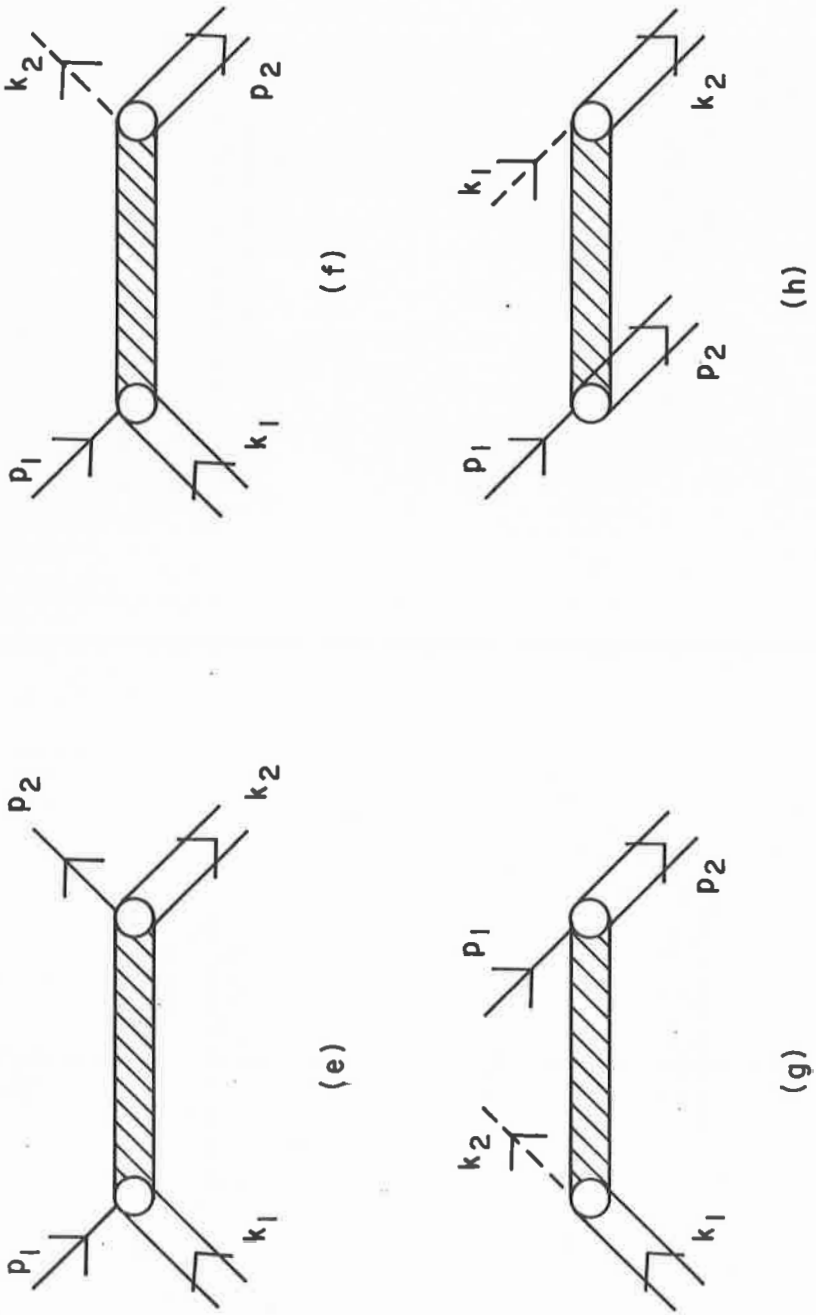


Figure 3 (cont'd)

References

1. Wave equations which were based on similar ideas were first suggested by Y. Nambu, Progr. Theor. Phys. (Kyoto) Suppls. 37-38, 368 (1966).

2. L. Susskind, Nuovo Cimento 69, 457 (1970).

3. See (a) A. Chodos and F. Cooper, Phys. Rev. D3, 2461 (1971) and (b) Phys. Rev. D (to be published) for details.

4. A. Chodos and R. Haymaker, Phys. Rev. D2, 793 (1970).

5. G. Domokos, S. Kovesi-Domokos, and E. Schonberg, Phys. Rev. D3, 1184 and 1191 (1971).

INTRODUCTION TO THE CHARACTERISTIC INITIAL VALUE PROBLEM IN QUANTUM FIELD THEORY.†

G. Domokos
Johns Hopkins University
Baltimore, Maryland 21218

Introduction:

The purpose of these lectures is to give a brief introduction to a new approach to quantum field theory. The physical motivation comes from an attempt to formulate the intuitively appealing "parton"-models and "additive quark models" in a consistent way and - if possible - to extend them to a full dynamical scheme. The method itself is known in the classical theory of partial differential equations (e.g. the equation of the vibrating string is solved by specifying initial data on characteristic lines). In the general theory of relativity, this method has been widely used to study gravitational waves. Quantum electrodynamics has been studied in this framework by the Stanford and Syracuse groups. Last but not least, it should be mentioned that the precursor of this method has been used under the name of "infinite momentum technique" (Cf. S.L. Adler and R. Dashen, Current Algebras, Benjamin, New York, 1964, where also the original papers are reprinted and S. Weinberg, Phys. Rev. 150, 1313 (1966).)

These lectures are divided into the following chapters.

1. Intuitive considerations.
2. Elementary theory of characteristics.
3. Kinematics, spinor technique.

---

†Research supported in part by the U.S. Atomic Energy Commission under Contract No. AT (30-1)-4076.

4. Some elementary properties of the characteristic initial value problem (CIVP) in quantum field theory; physical interpretation.
5. Lightlike quark model.

1.

Electron and neutrino scattering experiments at SLAC and at CERN can be interpreted "as if" the nucleon consisted of elementary constituents, or partons. (Cf. R.P. Feynman, Phys. Rev. Letters, 23, 1415 (1969) and the recent review: S. D. Drell and T-M. Yan, SLAC-PUB-808 (1970). The now-standard argument is somewhat similar in spirit to the old Weizsäcker-Williams method. Imagine that I view the electron scattering process from a reference frame where the nucleon is moving very fast (ideally with the velocity of light, although this situation can never be achieved for a real physical system). If (and that's a big if!...) the nucleon is composed of something else (Feynman's partons?) then in this frame the nucleon appears as a "beam" of the constituents, all of them moving with approximately the velocity of light in the same direction. (Internal motions of the constituents should not matter much in this frame.) Therefore, according to the well-known argument, the inelastic scattering cross section of electrons on the nucleon appears as the incoherent sum of the elastic scattering cross sections on the constituents. There is another (hopefully related) picture for hadronic reactions, the additive quark model. (Cf. J.J.J. Kokkedee, The Quark Model, Benjamin, New York, 1969.) Again assume that hadrons consist of "elementary constituents" (this time Gell-Mann's quarks) and that at high energies a "real" hadron can be represented as a beam of the constituents. This assumption (and a few more technical ones) lead to the remarkable "quark counting" relations for the total cross sections at high energies. Typical of these is:

$$(\sigma_{\pi N})^2 \cong \sigma_{\pi\pi}\sigma_{NN} \qquad (E \to \infty),$$

which (if one believes in the extrapolation of measured $\pi\pi$ cross sections...) is quite well satisfied. The qualitative success of these pictures is quite remarkable.

However, it raises at least two important questions. (These two only under the assumption that these qualitative pictures have something to do with reality. For example, quarks "exist" in some sense...)

a.) Can these pictures be reformulated in a Lorentz-covariant way? If the pictures are realistic, the "infinite momentum" frame should not be essential.

b.) Other models, based on the physical assumption that the nucleon is a strongly interacting system, have been at least as successful as the parton picture. (Cf. G. Domokos, S. Kovesi-Domokos and E. Schonberg, NAL preprint THY-12, 1971 and references quoted there.) Can these-seemingly contradictory-assumptions be reconciled somehow?

We claim that the answer to both questions is in the affirmative. (Cf. in particular reference 11)). Presumably, hadrons can be described by fields. In order to be able to formulate a field theory adapted to the spirit of a parton picture, we have to overcome several problems. Any quantum field theory has to be supplemented with a particle interpretation of the fields (usually stated as an asymptotic condition, or naively - i.e. apart from the delicate problem of wave function renormalisation constants -by studying the field theory in the formal limit of vanishing coupling constants.) In a conventional field theory, one identifies essentially the Fourier coefficients of the free fields as particle creation and annihilation operators. The corresponding particle states can be transformed to rest by a finite Lorentz transformation. In a space-time description this corresponds to specifying the initial condtions on a spacelike surface, say x^0 = const. If, however, I want to construct a field theory of "partons", I have to specify free "particle" states in a reference frame moving with the velocity of light. This is not a Lorentz-frame, hence the "particles" so defined cannot be transformed to rest by a finite Lorentz transformation even though they may carry a rest mass parameter. Hence they cannot even be observed as "ordinary" particles. (Refs. 6),7).) What initial conditions does this picture correspond to? Try to generate the "lightlike" frame as the limit of ordinary Lorentz

frames, and imagine that in some frame I gave the initial conditions on a surface $x^0 = \text{const}$. Now "boost up" in - say - the 3-direction. Then the "new" time-coordinate will be:

$$x^{0\prime} = \cosh\beta x^0 - \sinh\beta x^3 \sim \frac{e^{\beta}}{2}(x^0 - x^3) \quad (\beta \to \infty),$$

where - of course - $\tanh\beta = v$. Thus as $v \to 1$ the plane $x^{0\prime}=\text{const}$. asymptotically approaches a plane with a light-like normal and we must face the problem of specifying initial values on such a plane. It is immediately clear that one cannot prescribe initial conditions quite freely on such a characteristic plane, since it contains a light-like direction (in our case, x^0+x^3) and thus signals can propagate in the plane.

We shall see that in fact there are constraints among the initial data if we want to solve the Cauchy problem with the initial data specified on a characteristic surface. In the next chapter we briefly summarise the relevant facts from the theory of partial differential equations (PDE). (No attempt is made at any mathematical rigor.)

2.

Given a PDE in 1+n-dimensional space, (where I shall call the $(n+1)^{st}$ dimension the time, t,), the Cauchy-problem is the following. I prescribe the function(s) and an appropriate number of their normal derivatives on some surface (e.g. a plane) as initial data. Find the function(s) satisfying the PDE and the initial conditions everywhere in the 1+n dimensional space.

Does this problem always have a solution? The method of answering this question goes back to Cauchy.

In order to simplify matters, take n = 1 and let the coordinates by x_k (k=1,2). Consider a PDE of 2^{nd} order, linear in the highest derivatives. Its general form is

$$\sum_{i,k=1,2} \alpha_{ik} u_{,ik} = \Phi\,(u,\, u_{,i}, x_i) \qquad (2.1)$$

where the a_{ik} are constants, and Φ is some (smooth) function of its arguments. (Also, $u_{,i} = \frac{\partial u}{\partial x_i}$, etc.)

Now, I prescribe u and $u_{,i}$ (the function and its first derivatives) on a curve, C. Try to solve the Cauchy problem. In principle, I can find the solution as a power series. I have around C:

$$\begin{aligned} du_{,1} &= u_{,11}dx_1 + u_{,12}dx_2 \\ du_{,2} &= u_{,21}dx_1 + u_{,22}dx_2 \end{aligned} \qquad (2.2)$$

The lhs of (2.2) is known by definition on and around C. Now (2.1) and (2.2) gives a system of linear equations for $u_{,ik}$ (and by differentiation I find similar equations for the higher derivatives.)

Clearly, eqs. (2.1) and (2.2) have a unique solution, if the determinant

$$D = \begin{vmatrix} a_{11} & 2a_{12} & a_{22} \\ dx_1 & dx_2 & 0 \\ 0 & dx_1 & dx_2 \end{vmatrix} \neq 0 \qquad (2.3)$$

If - a piece of - C is such that D = 0, the Cauchy problem as stated has no solution. (For example, there are restrictions among the initial data.) Such a C is a <u>characteristic surface</u>.

<u>Example</u>. Wave equation in two dimensions:

$$a_{11} = 1, \quad a_{12} = 0, \quad a_{22} = -c^{-2}, \quad \Phi = 0.$$

From (2.3) we find the characteristic surface:

$$D = (dx_1)^2 - \frac{1}{c^2}(dx_2)^2 = 0, \text{ i.e.}$$

$dx_1 = \pm \frac{1}{c} dx_2$ (the two branches of the "light cone" in two dimensional Minkowski space). Similar procedure can be used for systems of partial differential equations; for details the reader is referred to the literature, e.g. ref. 1)

<u>Exercises</u>. 2.1) Find the characteristic surfaces of a Klein - Gordon equation in an external electromagnetic field:

$$[g^{\mu\nu}(\partial_\mu - ieA_\mu)(\partial_\nu + ieA_\nu) - m^2]\,\Phi = 0$$

2.2) Find the characteristic surfaces of the Dirac equation in an external electromagnetic field.

2.3) A spinless particle is coupled to an external tensor field, $H^{\mu\nu}$. The system is described by the Lagrangean density:

$$L = \tfrac{1}{2}(\Phi_{,\mu}\Phi^{,\mu} + m^2\Phi^2 + g\Phi_{,\mu}\Phi_{,\nu}H^{\mu\nu})$$

(g is a coupling constant.) Find the characteristic surfaces of the resulting equations of motion.

In what follows, we shall be interested in characteristic planes (since they can play the rôle - in some sense - of a surface on which Cauchy data can be prescribed. In a "decent" relativistic theory these are planes which are tangent to the light cone. We shall not consider pathological theories (involving couplings with high derivatives, like the one in Exercise 3) above.)

3.

In order to investigate the characteristic initial value problem (CIVP) further, it is convenient (although not necessary - contrary to some statements made in the literature) to introduce a basis in Minkowski space which is adapted to the surfaces chosen. Let us choose a surface $x^0 - x^3 = \text{const.}$ to specify the initial conditions on, and correspondingly, I use as coordinates:

$$\begin{aligned} x^t \equiv t &= \frac{1}{\sqrt{2}}(x^0 - x^3) \\ x^z \equiv z &= \frac{1}{\sqrt{2}}(x^0 + x^3) \\ \vec{x} &= \{x^1, x^2\} \end{aligned} \tag{3.1}$$

Notice that in this basis the nonvanishing components of the metric tensor are:

$$\begin{aligned} g_{ik} &= g^{ik} = \delta_{ik} \quad (i,k = 1,2) \\ g_{zt} &= g_{tz} = g^{zt} = g^{tz} = -1. \end{aligned} \tag{3.2}$$

Thus the scalar product of two vectors, say x^μ and y^μ are written as follows.

$$x.y = x_k\, y_k + x_z\, y^z + x_t\, y^t$$
$$= x_k\, y_k - x_z\, y_t - x_t\, y_t \qquad (3.3)$$

If - as usual - I want to characterise a momentum vector by its <u>covariant</u> components, ($\vec{p}$, $p_z \equiv k$, $p_t \equiv h$), then, for example, the mass shell condition,

$$p_\mu p^\mu + m^2 = 0,$$

can be written as

$$h = \frac{\vec{p} + m^2}{2k.} \qquad (3.4)$$

It is also easy to check that the various volume - and surface elements used in Minkowski space become:

$$d^4x = d^2\vec{x}\, dz\, dt = d\sigma^z\, dt, \qquad (3.5)$$

where $d\sigma^z$ is the (vectorial) element of the surface $t = \text{const.}$

$$d^4p\ \theta(p_o)\ \delta(p^2 + m^2)$$
$$= d^2\vec{p}\ \frac{dk}{2k}\ \Theta(k)\ \delta(h - \frac{\vec{p}^2 + m^2}{2k})\ dh \qquad (3.6)$$

Let us now write out the generators of the Poincaré group in this "lightlike" basis (L-basis).

The generators are: $M_{\mu\nu}$ (homogeneous Lorentz transformations) and P_μ (translations). The subscripts μ, ν now run through the values 1,2,z,t. Introduce the notation:

$$M_{ik} = \epsilon_{ik} M\ ,\ M_{zt} = N$$
$$M_{zi} = E_i\ ,\ M_{ti} = F_i \qquad (3.7)$$
$$P_z = K\ ,\ P_t = H$$

Here ϵ_{ik} is the Levi-Civita' tensor in two dimensions

$(\epsilon_{12} = -\epsilon_{21} = 1,\ \epsilon_{11} = \epsilon_{22} = 0)$.

In the notation (3.7), the nonvanishing commutators of the Poincaré generators read as follows:

$$
\begin{aligned}
&[M, E_i] = i\,\epsilon_{ij}E_j \qquad [M, F_i] = i\,\epsilon_{ij}F_j \\
&[N, E_i] = -i\,E_i \qquad [N, F_i] = i\,F_i \\
&[E_i, F_j] = -i\,\delta_{ij}N + i\epsilon_{ij}M \\
&[E_i, P_j] = i\delta_{ij}K \qquad [F_i, P_j] = i\delta_{ij}H \\
&[F_i, K] = i\,P_i \qquad [E_i, H] = i\,P_i \\
&[N, K] = i\,K \qquad [N, H] = i\,H
\end{aligned}
\tag{3.8}
$$

All other commutators vanish. Notice in particular that M, E_i, P_j, N generate the two dimensional Galilei group with dilations. The operators H and K play the rôles of the "Hamiltonian" and "mass operator", in this Galilean kinematics, respectively.[6] (There is another, "complementary" Galilei group, generated by M, F_i, P_j, N_4. The rôles of H and K are interchanged.) Thus in the transverse space (the kinematics is "nonrelativistic", i.e. governed by a Galilei group.

The remarkable fact about an L-basis is that one can introduce spinor-projections corresponding to the light-like directions. Indeed, on introducing the standard Pauli-matrices in the L-basis, viz.

$$
\begin{aligned}
\sigma^{1\dot{a}b} &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \sigma^{2\dot{a}b} = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \\
\sigma^{z\dot{a}b} &= \begin{pmatrix} \sqrt{2} & 0 \\ 0 & 0 \end{pmatrix} \qquad \sigma^{t\dot{a}b} = \begin{pmatrix} 0 & 0 \\ 0 & \sqrt{2} \end{pmatrix},
\end{aligned}
\tag{3.9}
$$

one immediately verifies that

$$
\nu^{\mu} = \xi^{\dot{a}}_{\dot{a}}\ \xi_b\ \sigma^{\mu\dot{a}b} \tag{3.10}
$$

is a null vector. (Indeed, $\nu_\mu\nu^\mu = \xi_{\dot{a}}\xi^{\dot{a}}\ \xi_b\xi^b = 0$). Hence, a null vector can be characterised entirely by a spinor. (The overall phase of the spinor is arbitrary; a spinor

contains "more information" than a vector, cf. in particular ref. 5). Thus introduce two spinors, ξ_a, η_a, corresponding to the two conjugate lightlike directions in the L-basis, with the phases adjusted in such a way that

$$\xi^a \eta_a = 1 \tag{3.11}$$

Then one verifies that the bilinears:

$$P_a{}^b = \xi_a \eta^b \ , \ Q_a{}^b = -\eta_a \xi^b \tag{3.12}$$

are orthogonal projectors, i.e.

$$\begin{aligned} P_a{}^c P_c{}^b &= P_a{}^b \ , \ Q_a{}^b Q_c{}^b = Q_a{}^b \\ P_a{}^c Q_c{}^b &= Q_a{}^c P_c{}^b = 0 \\ P_a{}^b + Q_a{}^b &= \delta_a{}^b \end{aligned} \tag{3.13}$$

(The phases can be so chosen that P_{ab}, Q_{ab} are Hermitean matrices). Analogous projectors, say $P_{\dot{a}\dot{b}}$, can be constructed by taking the complex conjugates of the spinors ξ_a, η_a. Hence I can construct projectors for spinors of higher rank. In the case I want to include reflections, I have to "double" the spinors in a well-known way. For instance, the Dirac spinor, ψ, is the direct sum of two spinors transforming according to the representations $(\frac{1}{2},0)$ and $(0,\frac{1}{2})$ of SL (2,C):

$$\psi = \begin{pmatrix} u_a \\ v^{\dot{a}} \end{pmatrix}.$$

Exercise. 3.1) Verify that the projectors acting on a Dirac spinor, which correspond to the t and z directions, are:

$$P_t = \tfrac{1}{2}\gamma_t \gamma^t = -\tfrac{1}{2}\gamma_t \gamma_z$$

$$P_z = \tfrac{1}{2}\gamma_z \gamma^z = -\tfrac{1}{2}\gamma_z \gamma_t$$

respectively, where γ_μ ($\mu = 1,2,z,t$) are the usual gamma-matrices in the L-basis:

$$\{\gamma_\mu \ , \ \gamma_\nu\} = 2g_{\mu\nu}.$$

(Hint: use the Van der Waerden representation, with γ_5 diagonal.)

3.2) Construct all the possible projectors for

a) a four-vector

b) an antisymmetric tensor.

(Note. A concise introduction to spinor calculus can be found e.g. in the first chapters of D. R. Corson: "Introduction to Tensors, Spinors and Relativistic Wave Equations", Blackie & Son, London., or in ref. 5.)

4.

We are now ready to discuss the CIVP in quantum field theory. As a warming - up exercise, consider a simple classical equation:

$$(-\Box + m^2)\ \Phi = 0$$

for a free, spinless particle. In the L-basis I have:

$$-\Box = -\partial_k\partial_k + 2\partial_z\partial_t\ ,$$

where - of course - $\partial_k \equiv \frac{\partial}{\partial x^k}$, etc. This equation is of first order in the normal derivative of the characteristic surface t = const. (Cf. Exercise 2.1). Hence, if I want to solve a CIVP by giving initial values on - say - t = 0, I may specify $\Phi(\vec{x},z,o)$, but not $\partial_t\Phi(\vec{x},z,o)$; the number of initial data for the CIVP is smaller than for the "ordinary" Cauchy problem.

Now integrate this equation with respect to z. Let g(z) be such that

$$\partial_z g(z) = \delta(z), \tag{4.1}$$

For example,

$$g(z) = \Theta(z) = \frac{1}{2\pi i}\int_{-\infty}^{\infty} d\tau\ \frac{e^{i\tau z}}{\tau - io} \tag{4.2}$$

or

$$g(z) = \tfrac{1}{2}\epsilon(z) = \frac{1}{2\pi i} P\int_{-\infty}^{\infty} d\tau \frac{e^{i\tau z}}{\tau} \qquad (4.3)$$

Then I get

$$\partial_t \Phi(\vec{x},z,t) = \tfrac{1}{2}(\partial_k\partial_k - m^2)\int_{-\infty}^{\infty} dz' g(z-z')\Phi(\vec{x},z',t), \qquad (4.4)$$

which can be solved trivially (e.g. by a Laplace-Fourier transformation.)

The lesson to be learned is that the solution of the CIVP is ambiguous, since I could have chosen either (4.2) or (4.3) in obtaining (4.4). This ambiguity is the consequence of the fact that signals can propagate in the surface where I specify the initial data. Later in this chapter I shall show that by invoking an additional physical requirement the choice of g(z) becomes unique; the correct choice is (4.3), which corresponds to a "standing wave" in the surface t = const.

If we are dealing with relativistic field equations for a multicomponent field, there are further restrictions on the characteristic initial data. Such equations can always be written in terms of spinors. Consider the example of a self-interacting field, transforming according to the representation $(n/2,0) \oplus (0,n/2)$ of SL(2,C), and thus described by the pair of spinors:

$$u_{a_1\ldots a_n},\ v^{\dot{b}_1\ldots\dot{b}_n},$$

where u,v are symmetric in their indices. The field equations can be written in the form:

$$\partial^{\dot{a}a_1} u_{a_1,\ldots a_n} + F^{\dot{a}}{}_{a_2,\ldots a_n}(u,v) = 0 \qquad (4.5)$$

$$\partial_{a\dot{b}_1} v^{\dot{b}_1,\ldots,\dot{b}_n} + G_a{}^{\dot{b}_2,\ldots,\dot{b}_n}(u,v) = 0,$$

where F, G are (usually algebraic) spinor functions of the fields u,v; the spinor differential operator $\partial^{\dot{a}b}$ is given by:

$$\partial^{\dot{a}b} = \sigma^{k\dot{a}b}\partial_k + \sigma^{z\dot{a}b}\partial_z + \sigma^{t\dot{a}b}\partial_t \qquad (4.6)$$

Now consider the projections,

$$u^{(t)}_{a_1 \ldots a_n} = P_{ta_1}{}^{b_1} \ldots P_t{}^{a_n b_n} u_{b_1, \ldots, b_n},$$

etc. of the fields.

Now we have the following

Lemma. The equation for $u^{(t)}$ $(u^{(z)})$, does not contain the derivative ∂_z (∂_t), respectively.

Proof. (trivial). Use (4.6) and notice that

$$\sigma t = \sqrt{2}\, P_t,$$

$$\sigma z = \sqrt{2}\, P_z$$

(projectors in the lightlike directions, provided the phases of the spinors ξ, η are adjusted appropriately.) The relations

$$P_t P_z + P_z P_t = 0,$$

proved in the last chapter give the statement, Q.E.D. The significance of this lemma is that the "z-projections" of the fields do not satisfy an "equation of motion" (since they do not contain the "time"-derivative), hence they give constraints between the initial data on the planes t = const.

Exercise 4.1.) Verify these statements for the free Dirac equation, both with two- and four-spinors. (Use the results of Exercise 3.1.)

Can the CIVP formulated and solved (apart from the ambiguity in (4.1), which I promised to resolve) for every field theory? The answer is NO!

Theorem. The CIVP is undetermined for field theories containing massive fields of spin higher than $\frac{1}{2}$.

I am not going to give a "general proof" of this theorem. It can be constructed by using a spinor formalism as indicated in eqs. (4.5). At any rate, one has to

be careful not to construct field theories which have "pathological" characteristic surfaces, see Exercise 2.3. This last restriction already excludes most of the high-spin field theories one would like to construct. Nevertheless, the theorem is not empty, as shown by the following

Example. Consider a free, massive field of spin one. (For a change, I use not spinors, but the more conventional vector formalism.) I take the conventional Lagrangean density:

$$L = \tfrac{1}{2}F_{\mu\nu}F^{\mu\nu} + \mu^2 A_\mu A^\mu, \qquad (4.7)$$

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu .$$

Since the field is massive, the Lorentz condition

$$\partial_k A_k - \partial_z A_t - \partial_t A_z = 0 \qquad (4.8)$$

is a consequence of the field equations (in a quantised theory it holds as an operator equation.)

In the L-basis the expression of the Lagrangean becomes:

$$\begin{aligned} L = \tfrac{1}{2}\big[&\partial_k A_i \partial_k A_i - 2\partial_z A_i \partial_t A_i - 2\partial_k A_z \partial_k A_t \\ &+ 2\partial_t A_z \partial_z A_t - \partial_k A_i \partial_i A_k + \partial_z A_i \partial_i A_t \\ &+ \partial_t A_i \partial_i A_z + \partial_k A_z \partial_t A_k - (\partial_t A_z)^2 \\ &+ \partial_k A_t \partial_z A_k - (\partial_z A_t)^2 + \mu^2 A_k A_k - 2\mu^2 A_z A_t\big] \end{aligned} \qquad (4.9)$$

The component A_t does not satisfy an equation of motion but a constraint. In fact, variation of (4.9) with respect to A_t gives:

$$\begin{aligned} &\partial_k \partial_k A_z - \partial_z \partial_t A_z - \partial_z \partial_r A_r \\ &+ \partial_z^2 A_t - \mu^2 A_z = 0 \end{aligned} \qquad (4.10)$$

Using (4.8) and (4.10), (4.9) becomes:

$$L = \tfrac{1}{2}\left[\partial_k A_i\,\partial_k A_i - \partial_k A_i\,\partial_i A_k + (\partial_k A_k)^2 - 2\partial_z A_i\,\partial_t A_i + \mu^2 A_k A_k\right] + \partial_k\,(A_z\,\partial_t\,A_k) \tag{4.11}$$

Hence, for example, the canonical Hamiltonean (with respect to the planes t = const.) does not contain either A_z or A_t. Therefore the dependence of these components on t is not determined. I know, however, that for the description of a massive spin-one particle I need 3 dynamical variables (say, A_k and A_z). The light-like formulation "loses" the longitudinal component as a dynamical variable. (Had I considered an interacting field theory, I would have found that A_z is expressible through A_k and with the components of the source of A_μ, hence again it is not independent.)

Remarks.

a) This loss of dynamical degrees of freedom has a clear intuitive meaning. I indicated already that the CIVP has "something to do" with infinite momentum frame considerations. (The exact connection will be clarified by the result of the next Exercise.) In a lightlike frame a particle has its spin completely aligned along the direction of motion. In the case of A_μ this corresponds to the components

$$A_\pm = \frac{1}{\sqrt{2}}\,(A_1 \pm iA_2).$$

The semiclassical probability of having a zero spin projection is zero; this is reflected in the disappearance of the longitudinal components from the Hamiltonean.

b) There is obviously no difficulty with massless fields of higher spin, so, for example, the CIVP for scalar and spinor electrodynamics or for the gravitational field is O.K. However, for example a quark model with vector gluons is in trouble.

Exercise 4.2.) (Connection with the infinite momentum frame technique.). Show that if there is a multispinor

field given (not necessarily irreducible), say $u_{a_1}, \ldots, a_n$, the various possible projections with the projectors P_t, P_z form "eigenvectors" of the spinor representation of the operator N_o. In particular,

$$P_t \otimes P_z \otimes \cdots \otimes P_t u, \quad \text{(M factors)}$$

$$P_z \otimes P_z \otimes \cdots \otimes P_z u \quad \text{(n factors)}$$

are the projections corresponding to the highest (lowest) eigenvalues, and hence they are the "surviving" field components in an infinite momentum frame moving in the positive (negative) 3-direction.

In the remainder of this chapter I shall consider a free Dirac field: this will be sufficient to illustrate the procedures.

Use a Lagrangean formalism, and introduce - for the sake of convenience - separate notation for the projections of the Dirac field, ψ:

$$\varphi = P_t \psi \quad , \quad \chi = P_z \psi .$$

Now rewrite the conventional free Dirac Lagrangean:

$$L_o = \tfrac{1}{2} \bar{\psi} \overleftrightarrow{\partial} \psi + m\bar{\psi}\psi,$$

in terms of φ and χ. Using the result of Exercise 3.1 and the fact that γ_z, γ_t are nilpotent:

$$\gamma_z^2 = \gamma_t^2 = 0,$$

one gets:

$$L_o = -\frac{i}{\sqrt{2}} \left\{ \varphi^+ \gamma_t (m + \tfrac{1}{2} \gamma_k \overleftrightarrow{\partial}_k) \chi + \chi^+ \gamma_z (m + \tfrac{1}{2} \gamma_k \overleftrightarrow{\partial}_k) \varphi \right.$$

$$\left. + \varphi^+ \overleftrightarrow{\partial}_t \varphi + \chi^+ \overleftrightarrow{\partial}_z \chi \right\} \qquad (4.12)$$

Variation of (4.12) with respect to χ, χ^+ gives the constraints:

$$i\sqrt{2} \frac{\delta L}{\delta \chi^+} = \partial_z \chi + \tfrac{1}{2} \gamma_z (m + \gamma_k \partial_k) \varphi = 0 \qquad (4.13)$$

and its Hermitean conjugate. Using (4.1), the solution of (4.13) is found immediately. One can thus eliminate χ from (4.12) and obtain the action in terms of the independent field components. One finds:

$$
\begin{aligned}
W_o &= \int d^4x\, L_o \\
&= \frac{-i}{\sqrt{2}} \int d^2\vec{x}\, dzdz'dt\, \{\varphi^+(z') \overleftrightarrow{\partial}_t \varphi(z)\ \delta(z-z') \\
&\quad + \varphi^+(z')(m + \overleftarrow{\partial}_k \gamma_k)\ g(z'-z)\ (m + \gamma_r \vec{\partial}_r)\varphi(z)\} \\
&+ \frac{i}{\sqrt{2}} \int d^2\vec{x} dzdz'dt \partial_k\, \{\varphi^+(z')\ (m + \overleftarrow{\partial}_r \gamma_r) g(z'-z)\, \gamma_k \varphi(z) \\
&\quad + \varphi^+(z')\, \gamma_k g(z'-z)\, (m + \gamma_r \vec{\partial}_r)\varphi(z)\}
\end{aligned}
\tag{4.14}
$$

In the last equation we have suppressed the arguments $\vec{x},t$ in the fields φ. We observe that the action is <u>non-local</u> in the z- coordinate. It is necessary to retain the two-dimensional divergence in (4.14), since we eventually will want to turn on external gauge fields in order to generate the expressions of currents. (Clearly, the presence of the divergence does not influence the <u>free</u> equations of motion.)

The theory is quantised canonically: one defines the canonical momentum by

$$
\Pi = \frac{\partial L_o}{\partial(\partial_t \varphi)}\ , \tag{4.15}
$$

which gives the equal-t commutation relation:

$$
\{\varphi(\vec{x},z,t),\ \varphi^+(\vec{x}',z',t)\} = \frac{1}{\sqrt{2}} \delta(\vec{x}-\vec{x}')\, \delta(z-z') \tag{4.16}
$$

In order to find a physical interpretation of the theory, one has to construct the generators of the Poincare' group. In the standard way one gets:

$$
\begin{aligned}
P_i &= -\int d\sigma^z T_{zi} \\
H &= -\int d\sigma^z T_{zt} \\
K &= -\int d\sigma^z T_{zz}
\end{aligned}
\tag{4.17}
$$

$$
\begin{aligned}
E_i &= -t\,P_i + \int d\sigma^z T_{zz} x_i \\
F_i &= \int d\sigma^z\,(z T_{zi} + x_i\,T_{zt}) \\
M &= -\int d\sigma^z\,\epsilon_{ik} x_i T_{zk} \\
N &= -tH - \int d\sigma^z\, z\, T_{zz}
\end{aligned}
\tag{4.18}
$$

Here $T_{z\mu}$ are the components of the <u>symmetric</u> energy- momentum tensor. The easiest way to obtain its expression is to turn on an external gravitational field using the "manifestly covariant form" of the Lagrangean and eliminate X afterwards by means of (4.13). This is a standard exercise and I shall not do it here. After a few standard manipulations (integration by parts, etc.), I can write:

$$
\begin{aligned}
P_k &= \frac{1}{i}\int d^2\vec{x}\,dz\;\varphi^+\partial_k\varphi \\
K &= \frac{1}{i}\int d^2\vec{x}\,dz\;\varphi^+\partial_z\varphi
\end{aligned}
\tag{4.19}
$$

$$
H = \frac{i}{2}\int d^2\vec{x}dzdz'\;g(z'-z)\;\varphi^+(z')\;(m^2-\partial_k\partial_k)\varphi(z)
$$

$$
M = \frac{1}{i}\int d^2\vec{x}dz\varphi^+\,[x_1\partial_2 - x_2\partial_1 + \tfrac{1}{2}i\,\gamma]\varphi, \tag{4.20}
$$

where I introduced $\gamma = -i\gamma_1\gamma_2$. The operators (4.19) can be diagonalised by Fourier transformation. Remember that the Fourier transform of g(z) is

$$
?\frac{1}{k}\ ,
$$

where the question mark indicates <u>some</u> way of avoiding the pole at k = 0. Now introducing:

$$
\varphi(\vec{p},k) = \frac{1}{(2\pi)^{3/2}}\int d^2\vec{x}\;d\;z\;e^{-i(\vec{p}\vec{x}\,+\,kz)}\varphi(\vec{x},z,o),
$$

I get for example:

$$
\begin{aligned}
K &= \int d^2\vec{p}\int_{-\infty}^{\infty}dk\;k\;\varphi^+(\vec{p},k)\;\varphi\;(\vec{p},k) \\
H &= \int d^2\vec{p}\;?\int_{-\infty}^{\infty}dk\quad\varphi^+(\vec{p},k)\varphi(\vec{p},k)\;\frac{\vec{p}^2+m^2}{2k}
\end{aligned}
\tag{4.21}
$$

By performing an infinitesimal gauge transformation: $\delta\varphi = i\varphi$, $\delta\varphi^+ + i\varphi^+$, I can also obtain the expression of the fermion number operator, F:

$$F = \int d^2\vec{p} \int_{-\infty}^{\infty} dk \; \varphi^+(\vec{p}, k)\varphi(\vec{p},k) \qquad (4.22)$$

Clearly, I want the operators (4.17), (4.18) to generate a unitary representation of the Poincaré group of the time-like class. This means in particular that H,K are Hermitean operators with nonnegative spectrum. It follows that

i) I must choose the principal value of the integral over k in (4.21) (Hermiticity)

ii) I must interpret

$$\varphi_a(\vec{p},k) \quad (k>0)$$

as an annihilation operator of a fermion of spin projection "a" ($a = \pm\frac{1}{2}$, cf. (4.20), and momentum components $(\vec{p}, k, \frac{\vec{p}^2 + m^2}{2k})$, and

$$\varphi_{-a}(-\vec{p}, -k) \quad (k < 0)$$

as a creation operator of an antifermion (cf. (4.22)) of spin-projection "a", and momentum components $(\vec{p}, -k, \frac{\vec{p}^2 + m^2}{-2k})$. Condition i) removes the ambiguity in the definition of g(z) as I promised earlier: only the choice (4.3) is correct. In particular this condition means that the correct solution contains no Bondi-Metzner "news function", i.e. no information is allowed to propagate in the plane t = 0, cf. ref. 3).

5.

The material presented in this, last part of the lectures has been published in the form of a separate article[11]. Here I restrict myself to a brief qalitative sketch of the problem and the results. The purpose of the calculation in [11] was to resolve the mystery of the free quark (or parton) models, which I discussed in the Introduction. Surprisingly, the answer is almost trivial. Consider an interacting quark model, for example, by adding a term:

$$L_{int} = \frac{F}{2\sqrt{2}} \left[(\bar{\psi}\ \gamma_\mu\ \psi)^2 - (\bar{\psi}\ \gamma_5 \gamma_\mu \psi)^2 \right]$$

$$= \frac{F}{2\sqrt{2}} \left[(\varphi^+\varphi)(\chi^+\chi) + (\varphi^+\gamma\varphi)(\chi^+\gamma\chi) \right], \qquad (5.1)$$

to the free Lagrangean, (4.12). Here F is a coupling constant of dimension (length)^2. One proceeds as before by solving the constraint equation (4.13) for χ (which now contains a contribution from the interaction) and eliminating χ from the action. The surprising result is that after doing all this, the expression of the action is <u>almost</u> of the form of a free action functional. The only difference is that the function

$$g(z-z') = \tfrac{1}{2}\epsilon(z-z')$$

in (4.14) has to be replaced by an expression:

$$G(\vec{x},t;z,z') = \tfrac{1}{2}\epsilon(z-z')\ \exp\ iF\left[B(\vec{x},z,t) - B(\vec{x},z',t)\right] \cdot \qquad (5.2)$$

For the interaction (5.1), $B(\vec{x},z,t)$ is of the form:

$$B(\vec{x},z,t) = \tfrac{1}{2}\int dz'\,\epsilon(z-z') \left[(\varphi^+(\vec{x},z',t)\varphi(\vec{x},z't) + (\varphi^+(\vec{x},z',t)\gamma\varphi(\vec{x},z',t))\gamma \right]. \qquad (5.3)$$

One can verify that the form (5.2) is quite general; it is the expression of B which depends on the specific form of the interaction (or, rather, the expression in square brackets in (5.3).

The quantity FB has a very simple physical meaning: it is nothing but an operator generalisation of an <u>eikonal phase</u> in ordinary Schrodinger theory. (Cf. the lectures of Drs. Abarbanel, Islam and Sugar in this volume.)

Denote for a moment:

$$V(\vec{x},z,t) = F[(\varphi^+\varphi) + (\varphi^+\gamma\varphi)\gamma],$$

and introduce the Fourier transform:

$$\tilde{V}(\vec{x},k,t) = \frac{1}{2\pi}\int d\ z\ e^{-ikz}\ V(\vec{x},z,t).$$

Then (5.3) can be rewritten as

$$B = P\int \frac{dk}{k}\, e^{ikz}\, \tilde{V}(\vec{x},k,t),$$

which makes the analogy clear. (V plays the role of a "potential".) The fact that the interaction appears in an "eikonal form", has important consequences. In order to see this, one derives the expressions of the currents in the usual way by calculating the first order response function to weak external vector and exial vector fields. In particular, the expressions of the "longitudinal" components of the vector and axial currents turn out to be

$$\begin{aligned} V_z &= \frac{1}{\sqrt{2}}\, \varphi^+\varphi \\ A_z &= \frac{1}{\sqrt{2}}\, \varphi^+\gamma\varphi \end{aligned} \tag{5.4}$$

(For the sake of simplicity, I suppress internal symmetry indices in all these formulae.)

Compare (5.3) with (5.4). On introducing the <u>composite fields</u>:

$$\begin{aligned} S &= \frac{\sqrt{F}}{2} \int dz' E(z-z')\ \varphi^+(z')\varphi(z') \\ P &= \frac{\sqrt{F}}{2} \int dz' E(z-z')\ \varphi^+(z')\gamma\varphi(z'), \end{aligned} \tag{5.5}$$

one can see on the one hand that (5.4) can be written as:

$$\begin{aligned} V_z &= \frac{1}{\sqrt{2F}}\, \partial_z S \\ A_z &= \frac{1}{\sqrt{2F}}\, \partial_z P \end{aligned} \tag{5.6}$$

On the other hand, the expression of the eikonal phase operator, exp(iFB), becomes:

$$\exp(iFB) = \exp i\sqrt{F}\ [S + P\gamma] \tag{5.7}$$

Since the longitudinal components of the currents are proportional to the gradients of S and P, the latter can be interpreted as composite fields corresponding to scalar and pseudoscalar particles. (This is the standard reasoning used in "deriving" PCAC.) After Nambu and Jona-Lasinio

(Y. Nambu and G. Jona-Lasinio, Phys. Rev. 122, 345 (1961); ibid. 124, 246 (1961)) many people have argued that the physical vacuum is degenerate and thus the vacuum expectation values of S and P are different from zero. (This is a familiar situation e.g. in the theory of superconductivity: the ground state is filled with Cooper pairs.) Assume now that this is the case.

The first simple question one can ask is: what are the equations of motion satisfied by "single particle" observables? (Current densities are examples of single particle observables in a quark model; they - or rather their commutators - can be measured in the inelastic scattering of electrons and neutrinos, and hence one can extract dynamical information from these experiments.) In the ordinary nonrelativistic many-body problem one knows that if the ground state (the "vacuum") is degenerate, the Hartree-Fock approximation to the equations of motion of single particle observables works quite well. (The Hartree-Fock approximation - a nonperturbative approximation - consists in replacing the operators corresponding to pairs which fill the ground state by their ground state expectation values. The latter are then calculated self-consistently. Qualitatively, Hartree-Fock is expected to work well if there is a large "condensate" in the ground state, since then the relative quantum fluctuations of the operators around their expectation values are small.)

Let us observe that in this formulation of a relativistic field theory one can borrow freely from the methods of the nonrelativistic many-body problem. Indeed, the action (4.14), with the modification (5.2), can be viewed as one describing a Galilei-invariant field theory in two dimensions with variable mass (cf. the remarks made in Ch. 3.) The fact that the interaction is more complicated than in the usual models of, say, superconductivity, does not matter here.

Now if I calculate in the Hartree-Fock approximation, a surprising thing happens. Replace B in (5.2) by its vacuum expectation value. No matter what that is, it is certainly a constant (translation invariance!). However, then (5.2) - and the equations of motion of φ and thus of the current densities - reduce to what I would obtain if

there were no interaction at all!

Gell-Mann and several other authors (cf. H. Fritzsch and M. Gell-Mann, 1971 Coral Gables Conference) have conjectured that the current densities behave "as if the currents were constructed out of free quarks". We now see that in the present version of an interacting quark model this means nothing else but making a Hartree-Fock approximation. In particular, all the results of the "naive" parton model come out in this way. This does not mean, however, that the theory is altogether a free one: in calculating many-body states (for example the nucleon state vector) there will presumably appear a strong effective interaction. (Notice in particular that the formal expression of the action contains highly singular objects: exponentials in the fields S,P...) One can at least hope that the present approach to field theories will explain some of the puzzling features of hadron physics.

6.

Conclusion. (This is Exercise 6.0) to be worked out by the reader for himself.)

REFERENCES

(This is not a review article on the subject. We list only those papers which we personally found useful in preparing these lectures; hence the following list is necessarily biased and incomplete. We offer our apologies to those authors whose works have been omitted from the list; the inclusion or omission of some work is by no means meant as a judgment of quality.) References are listed according to subjects.

Theory of characteristics in partial differential equations. (PED)

1.) (Any good text on PDE would do). The classic is: F. Hadamard, Lectures on Cauchy's problem (Dover, New York, 1952).
R. Courant and D. Hilbert: Methoden des Mathematischen Physik. (Springer, Berlin, 1937), Vol. II. For a recent review, see F. John, in Mathematics Applied to Physics, (Ed. E. Roubine) (Springer, Berlin, 1970).

CIVP in general relativity.

2.) R. Sachs, Phys. Rev. 128, 2851, (1962).

3.) R. Sachs, in Proc. Theory of Gravitation. (Gauthier-Villars, Paris, and PVN, Warszawa, 1964).

4.) R. Sachs in Relativity, Groups and Topology (Les Houches, 1963) Gordon and Breach, New York.

5.) R. Penrose, "Analysis of Space-Time" (Birkbeck College rep. London, 1968).

"Lightlike" kinematics.

6.) H. Bacry and N. P. Chang, Ann. Phys. (N.Y.) 47, 407 (1968).

7.) G. Domokos in Proc. Nobel Symposium VIII, Wiley-Interscience, New York (1968).

Quantum Electrodynamics and general field theory.

8.) F. Rohrlich: Theory of Photons and Electrons off Null Planes. (Syracuse U. preprint, 1970). (Extensive list of references.)

9.) J. B. Kogut and D. E. Soper, Phys. Rev. D 1, 2901 (1970).

10.) H. Leutwyler, F. R. Klauder and L. Streit, Nuovo Cim. 66A, 536 (1970).

Relativistic Quark model.

11.) G. Domokos, Johns Hopkins U. preprint, NYO-4076-18 (1971).

FLUX QUANTIZATION AND PARTICLE PHYSICS

H. Jehle
George Washington University
Washington, D. C.

Introduction

As reported in earlier papers[1] a consistent theory of leptons based on quantized flux loop may be given. The idea is that instead of trying to relate magnetic monopoles to elementary particles, we take the more conservative approach of considering only closed magnetic flux loops. These flux loops, to conform with the Maxwell-Lorentz equations, are assumed to have the forms of magnetic field lines of a dipole source if an electron or a muon is considered. It is assumed that one loop corresponds to one lepton, and that the magnetic field results from a superposition of alternative forms which this flux loop may adopt, a superposition with complex probability amplitudes. This superposition is similar to the superposition of alternative path histories by which a quantum mechanical path is constructed in Feynman's space-time approach to quantum mechanics.

There is nothing special about such an assumption, but it is interesting to note that the very same definition which introduces quantized flux (as a singular line of the phase θ of the ψ function of field particles) also implies an electric field if that flux line moves. In particular, if the magnetic field, represented by the flux loop, has the dipole movement of one Bohr magneton (or one muon magneton) and if the loop spins about the dipole axis with angular velocity $2m_e c^2/\hbar$ (or $2m_\mu c^2/\hbar$), the resulting electric field is the Coulomb field of charge e. (To make that statement precise, the alternative dipole axes $\vec{\zeta}$ would, for a source lepton of spin in $\vec{Z}$ direction, be distributed with probability amplitudes proportional to $(1+\cos(\vec{\zeta},\vec{Z}))$. This theorem is not so astonishing; it is the reverse of the theorem of the Dirac theory of the

electron which derives the magnetic moment of the electron from its charge.

A separate issue in this magnetic theory of the electron or muon is the question how that magnetic moment $\frac{e\hbar}{2mc}$ is related to the quantized flux $\Phi_q = hc/e$. A single particle of mass m is, in the relativistic theory of the electron or muon, non local[2] in ordinary position, by an amount $\hbar/mc$ and it has a Zitterbewegung frequency $2mc^2/\hbar$ which may be interpreted as a spinning frequency. This "quasinonlocality" $\hbar/mc$, as we may call it, relates the quantized flux to the magnetic moment. It was shown that appropriate superposition, with complex probability amplitudes, of the loop-form contributions results in an effectime magnetic moment equal to the Bohr or muon magneton if the electromagnetic interaction constant is of the order of 1/137. And it was shown that under the same circumstances and by the same superposition rules the electromagnetic energy turns out to be of the order mc^2, and the electromagnetic angular momentum of the order of $\hbar/2$. This is interesting in that it shows a consistent theory of the lepton on the basis of the Maxwell-Lorentz fields is possible. It should be noted, however, that the quantum field which describes the lepton is the spinor type probability amplitude field; the electromagnetic field plays the role of the observable.

Neutrino

We also propose a neutrino to be a flux loop, but of the form of a trefoil spinning through space like a coasting three-blade propeller. In this manner, apart from fluctuations, no net electric field seems to be produced by the moving flux loop. A question is to be raised as to whether or not a neutrino has a magnetic moment (in the direction of its spin or opposite); if this is the case, a loop of one such flux orientation is expected to occur; if this is not the case, the proposal in l.c.[1] Phys. Rev. would apply.

The topology implied in a trefoil knot is considered to be important. It is assumed that the crossing of a flux loop over itself (which would be necessary for a trefoil to transform into a circular loop) is a process which is very rare, a weak process. There are two distinct trefoils (a right handed and a left-handed one). They

correspond to neutrino and an antineutrino respectively. Apart from the handedness of the trefoil, a neutrino and an antineutrino are distinguished by opposite signatures of the frequencies of their probability amplitudes.

This two-fold distinction between neutrino and antineutrino is indeed to be considered as the distinction between particle and antiparticle in general.

Hadrons

To proceed to mesons and baryons, we assume the quark picture, in particular SU(6), and consider a quark to be a flux loop if it is interlinked with another (in the case of a meson) or with two other loops (in the case of a baryon). Their spinning is entirely different from that of a free loop (i.e. a lepton) and there may accordingly no meaning be attached to the concept of a single quark.

Before we discuss the detailed structure of quarks, we may, in connection with the neutrino, remark that the λ quark is expected to be a trefoil whose handedness indicates the strangeness +1 or -1. Indeed, with such interpretation we may readily understand strangeness conservation. We assume that interacting loops, ignoring other loops of the same particles, may only reluctantly cross over each other. Thus, two trefoils of opposite handedness (S=+1 and S=-1) may readily annihilate each other,or be pair-created, without the necessity of any loop-crossing; this is not the case for strangeness non-conserving reactions. They are, by virtue of the topological change, not parity conserving which is interesting to notice.

Assumptions about Loops and their Linkage.

We make the following assumptions about the loops, their forms and linkage, and illustrate them by figures. A quark is a quantized flux loop if interlinked with another loop. If, under observance of conservation laws, a loop may get disengaged from linkage, it behaves like a lepton. It becomes, therefore, meaningless to search for individual quarks.

The quarks belonging to one meson or baryon are able to spin independently. This is possible if they are confined to non-overlapping regions. Each of the meson and baryon loops of Figs. 1 may also move independently about their two axes: they spin about the straight (vertical) axis and whirl about the circular axis. In Figs. 1, 1b circular axis is defined as the equator of the spherical core. The dots indicate the intersections of the circular axis with the picture plane. In Fig. 1c the core is not indicated and the circular axis is seen edgewise as a dash-dot-dash line. The domains in which they move are shown separated by dashed toroidal surfaces.

Even though this scheme seems to be the scheme by which mesons and baryons are expected to function, we should remark that the assumptions of straight, circular axes, as regard topology, is rather special. In the following paragraphs we discuss as an alternative scheme the more general case of oval shaped axes as shown in Figs. 2, 3, 5, 7. The loops of Figs. 5, 7 may spin in the regions shown by Figs. 4, 6: the illustrated interfaces between the toroidal regions may shift toward one or the other axis. In order to clarify the spinning of a quark loop, in the alternative case of Figs. 5 and 7, we have drawn single loops in Figs. 2 and 3. One mode of "spinning" is (cf. Fig.2) a rolling, whirling motion about the left axis (donut) which amounts to a kind of translational motion along the right donut. The other mode (again looking at Fig.2) is a similar "spinning" about the right axis (donut) which is equivalent to a kind of translational motion along the left donut.

Apart from a ± signature, the spinning is considered to occur with equal angular velocities.

A question arises as to the nature of "orbital angular momentum" of the higher lying meson and baryon states. In the case of axes shown in Figs. 1, the higher angular momenta are presumably of the nature of giant quarks, i.e. quark loops possessing spin higher than $\frac{1}{2}$. In the case of Figs. 2-7, the axes may move with respect to each other, which might then represent orbital angular momentum.

Because of the quasi-nonlocal nature of a single particle, the position of a point source is smeared out

over a region of linear extension $\bar{h}/mc$; it extends up to the axes of the hadron. This region, called the "core", is the region in which the Maxwell-Lorentz equations have inhomogeneous terms. For the definition of localization we refer to the earlier papers, and to Section V and Appendix I of the Phys. Rev. paper.[1]

The "attachment" of a loop to the core may be considered a topologically meaningful concept. It is as if the axes by which the core is bounded, were a set of rings with which the fluxloops are linked and held together. The reason for this is that a loop, as shown in the figures, is only a single loopform out of a continuous manifold of pairs (meson) or triplets (baryon) of linked loopforms which are spread over all space. They define a fibrated space with two singular lines, namely the "axes". These singularities of the fibrations provide for a topologically invariant characterization of a loop by its winding numbers. In Figs. 2 and 3 these winding numbers are illustrated; for clarity of the illustration we have represented these two axes by two donuts.

The equivalent electric charge of a quarkloop is evidently (cf. the definition of the electric field in the earlier papers) proportional to the number of times a loopline skips over a point in space, per period of spin, times the spinning frequency. This is immediately seen to be proportional to one of the winding numbers times its spinning frequency minus (or plus) the other winding numbers times its spinning frequency. Assuming the two spinning frequencies to be equal, the equivalent electric charge becomes proportional to the difference (or sum) of the winding numbers.

When it comes to establishing a model which gives account for so invariant a quantity as is the electric charge, it is important to have a model which relates it to invariant quantities such as are the winding numbers, rather than to particular geometrical features.

We may be reminded in that context that the starting point of this fluxquantization project was the recognition of still another invariance property of electric charge. It was the recognition that the electric charge of the

electron and of the muon came out to be identically the same by virtue of exact cancellation of the mass m in the product $(e\hbar/2mc)\cdot(2mc^2/h)$, i.e. by virtue of a strict scale invariance.

With that invariant loop characterization by winding numbers, we may suggest the loopforms Figs. 2 (which are identical with those of Figs. 3) to apply to the $\hbar\mathrm{P}$, and λ quarks. If we assume a tendency of spinning to occur so that the resulting electric field (and thus energy) is as small as possible under the constraint of equality of absolute values of spinning frequencies (angular velocities $2mc^2/\hbar$), we may assume that it is preferably the difference between (rather than the sum of) the two winding numbers which characterizes the equivalent electric charge of a quarkloop.

-The relation between dipole magnetic moment and equivalent electric charge was worked out for electron or muon in the earlier papers[1]. In as far as there are now, in the case of the quarks of mesons and baryons, deviations from a dipole structure of magnetic moment, we may expect the magnetic moment to be only approximately related to electric charge.

Let us now consider the probability amplitude distributions for mesons and baryons. Conventional SU(6) models had the immense success of explaining the ratio of magnetic moments of baryons, in particular of neutron to proton[3]. This establishes the appropriateness of the choice of symmetric spin-isospin functions for baryons.

But, as the quark model considers the quarks as spin $\frac{1}{2}$ particles, their complete wave functions should be antisymmetric, i.e. satisfy the Pauli Principle. It seems to us too much a violation of basic principles of quantum mechanics to postulate para-statistics.

The alternative way of staying within established quantum mechanical principles and to satisfy the Pauli Principle is to assume antisymmetric orbital wave functions for baryons. As long as quarks are considered to be SU(6) particles, such antisymmetric orbital wave functions may, however, scarcely be considered to be fit for the lowest lying baryon states.

The quark loop-model avoids these difficulties. Quark loops are considered as localized objects, but in the spirit of SU(6), they should be able to exchange places in their arrangements shown in Figs. 5 and 7. An antisymmetric "spacial" wave-function of quarks (i.e. describing their distribution over the regions between the axes) may therefore readily be formed and the Pauli Principle is satisfied. Such antisymmetric "spacial" wave function also implies all permutations of arrangements of the quark loops over the "toroidal" regions and thus provides for equal distribution over both axes, for any quark loop of a hadron. This again reasures the winding numbers as correctly determining electric charge.

The question of integrity of electric charge in leptons, mesons, and baryons arises. As regards to leptons, the scale invariance, i.e. the rigorous independence of equivalent electric charge e of mass m is the basic point. The entire theory being based on gauge-invariant definition of the electromagnetic fields, it follows that charge conservation is rigorously maintained. Thus, starting with the electron's or the muon's integrity of charge, all subsequent reaction products will have integer charges also.---Why fractionality -1/3, +2/3, -1/2 of quark charge may be expected, has been discussed in App. II of Phys. Rev.[1]; here we have, through the winding numbers, given an explanation of their exact rations ∓ 1 to ± 2 to ∓ 1.

It should be noted that there may be close connection between the present quark proposal and that based on dipoles of positive and negative magnetic charge.[4] Following the idea of our quark model, it may then, however, be appropriate to represent a meson by two dipoles and a baryon by three dipoles. Even though we would like to avoid the concept of magnetic monopoles entirely (and found this to lead to a successful and conservative interpretation of particle physics) we do not want to object to monopole theories altogether, in particular not in the afore mentioned dipole connection. We do not, however, think the identification of a monopole with a quark to be appropriate.

Acknowledgements

I am deeply indebted to the Research Corporation for an individual Research grant and I would like to extend my very best thanks to the George Washington University Committee on Research for their support through an institutional NSF grant. These supports have made it possible to proceed with this work. I am particularly grateful for the help I received from my colleagues through the years, from Professor W. C. Parke, and from Dr. A. Ghovanlou and for the drawings in this and the preceeding paper, by Dr. T. Furlow and by Miss J. Furlow. The present project was initiated at the Summer Institute for Theoretical Physics, University of Colorado at Boulder, Colorado.

Appendix I

We may make a few remarks in regard to quark loops: A λ and a $\bar{\lambda}$, as they have opposite handedness, may annihilate each other without crossing of flux loops when they come to a mutual approach. Why, nevertheless, may they coexist, attached to the same core of a meson? A model (carefully built with rubber catheters) shows that, if mounted on the same axis, their opposite handedness prevents any annihilation without crossing of flux loops. An-Another question has come up as regards to the absence of spin $\frac{1}{2}$ baryons of the type of $\lambda\lambda\lambda$, of the $\rho\rho\rho$, and of the $\lambda\lambda\lambda$. In that case, there are at least one pair of neighboring quarks of opposite spin present. As they are of equal charge, their magnetic field orientation is opposite. One might assume that they cannot coexist as nearest neighbors because they repel each other.

Geometrical characterizations of independent bundles of flux loop forms have been proposed in the earlier papers. This led to a grouping of loop forms into 207 bundles. It was pointed out that this was a heuristic procedure and that a formal treatment would have to rely more on analytic tools. Indeed, one should take use of generalized spherical harmonics expansions and take account of the invariance with respect to scaling when attempting to define the distribution of probability amplitudes for loop forms.

It may finally be remarked that in the comparison of the electron with muon, the angular group and phase velocities of the terms bilinear in the amplitudes are as 1 to 207. The linear velocities of the spinning loops being assumed to be of the order of c to c for the two particles and the sizes stand in the ratios of 207 to 1, their electromagnetic energies are expected to be related as 1 to 207. The bundling of loop-forms may accordingly account for the difference in mass electron versus muon.

Appendix II

In our proposal to describe a lepton in terms of a quantized flux loop we took recourse to geometrical pictures to provide for a simplified model in terms of which we could readily check on the consistency of that theory, in terms of which also the main structure of the theory could be decided upon, and in terms of which approximate numerical results were obtained. It was evident from the start that a more direct analytic, rather than geometric, formulation was to be achieved. It was, however realized that too many possibilities for the structure of a quantum mechanical theory were at hand. The decision upon the appropriate choices of assumptions may now be made as the consistency of the heuristic model is recognized.

In this note a short sketch of a proposed wave equation for the probability amplitudes ψ of the distribution of loop-forms is given. We may be reminded that instead of a description of loop form distributions by functionals, we may take advantage of the similar shape of all magnetic point dipole loop-forms, to characterize them by 3 angle variables (e.g. the Euler angles) and 1 size parameter σ.

In the calculation of the electric field of the spinning flux-loop, the mass of the electron or the muon cancels out rigerously. This implies the equality of the electric charge of these two leptons. The cancellation of the mass means that there is a scale invariance. Considering this we may write

$$r = \sigma/(\hbar/mc) \qquad (1)$$

as a parameter indicating the size of a loop-form.

In what follows we consider certain analogies between the spinning of the manifold of loop-forms and the spinning of a symmetric top. It is well known that the angular momentum operator of the spinning top may be written as

$$\mathfrak{L}^2 = r^2 \left(\frac{\partial^2}{\partial \xi^2} + \frac{\partial^2}{\partial \eta^2} + \frac{\partial^2}{\partial \rho^2} + \frac{\partial^2}{\partial \chi^2}\right) - \frac{1}{r}\frac{\partial}{\partial r} r^3 \frac{\partial}{\partial r} \qquad (2)$$

with eigen values

$$-4\ell(\ell+1), \qquad (3)$$

ℓ being half odd integer or integer; the eigen value problem posed with the operator (2) is actually an eigen value problem on a 3 dimensional hypersurface

$$r^2 = \xi^2 + \eta^2 + \rho^2 + \chi^2 \qquad (4)$$

The 3 ratios $\xi:\eta:\rho:\chi$ are related to the Euler angles.

The point dipole loopforms are alike in respect to the 3 angular parameters (Euler angles) and the size parameter. We may therefore associate with each loop-form a point in the 4 dimensional space of the variables ξ, η, ρ, χ (with the use of (1) and (4)). We now make the assumption that the wave equation for a point dipole source model be given by

$$r^2 \left(\frac{\partial^2}{\partial \xi^2} + \frac{\partial^2}{\partial \eta^2} + \frac{\partial^2}{\partial \rho^2} + \frac{\partial^2}{\partial \chi^2}\right) + \omega^2 - C)\psi = 0 \qquad (5)$$

ω being the frequency of the ψ wave. With the eigen value of the angular part of ψ, i.e. (3), we get for the r dependent factor of ψ

$$\left(\frac{1}{r}\frac{\partial}{\partial r} r^3 \frac{\partial}{\partial r} - 4\ell(\ell+1) + \omega^2 - C\right)R = 0 \qquad (6)$$

As the wave equation (5) is, in consideration of (4), linear homogeneous of degree zero in r, the angular velocity of loop forms is independent of r, in accord with our assumption of the Zitterbewegung frequency of $2mc^2/\hbar$ as giving the angular velocity of a spinning flux loop.

We want the $|\psi|^2$ to represent the probability for a loop to adopt the size r, i.e. to have its aphelion size $\sigma=r(\hbar/mc)$ to fall in a unit cross section at a distance σ from the source point. Accordingly, R(r) should be proportional to

$$R(r) = a\, r^{-3/2} \tag{7}$$

which, inserted in (6) means

$$-4\ell(\ell+1) + \omega^2 - C = \frac{3}{4} \tag{8}$$

Such a solution (7) corresponds to a point dipole source. For the quasi-nonlocal source, the position of the dipole source appears smeared out. Replacing, accordingly, the point dipole source by an extended source (a crude substitute for a transformation from mean position to position by a Pryce-Tani-Foldy-Wouthuysen transformation), we get rid of that singularity. We may effect this by inttroducing into the wave equation (5) or (6) a "potential" U(r) which is positive in the "core" region $0 \leq r \lesssim 1$ and goes to zero at the core surface $r \gtrsim 1$ and is zero outside, $1 \lesssim r < \infty$ which insures the $r^{-3/2}$ behavior of R(r) for large r.

Considering these, Eq. (6) may be written as

$$\left(\frac{1}{r}\frac{\partial}{\partial r}\, r^3\, \frac{\partial}{\partial r} - U(r) - 4\ell(\ell+1) + \omega^2 - C\right)R = 0. \tag{9}$$

This equation might represent the wave equation for the loop form of an electron or a muon. The choice $C=\frac{1}{4}$ gives, by (8),

$$\omega = 2\ell + 1 \tag{10}$$

the commensurabilities of ω permit phase correlated motion of loopform amplitudes in the case of the electron, to distinguish them from random phased muon amplitudes.

The group theoretical analysis of Eq. (5), (9) is particularly promising. The important results are expected to be already obtained by A. O. Barut.

The wave equation is presumably to be written in a linearized, Dirac form.

Appendix III

The handedness of the flux loops (right handed screw or left handed screw in relation to the axes) is, besides the winding numbers, an important characteristic feature of a loop. The assignment of handedness to the types of quarks will still have to be decided according to a general plan. It seems, however, that as in the case of the neutrino, the handedness is determined by the alternative: quark-antiquarks.

From the consideration of models of flux loops, their intrinsic handedness and their link with the axes, it becomes obvious that between a loop (2,3) and a loop (1,3), both of equal handedness, a transition is simple compared with a transition between (2,3) and a (1,2) loop. This fact may provide for an understanding of the ΔS versus ΔQ relationship.

We may also recall that the flux loop model explains that strangeness nonconserving weak processes imply some crossing of flux lines and are, because of the implied topological change, not parity conserving.

FIGURE CAPTIONS

Figs. 1. Rotationally symmetric axes-system (a straight vertical axis, and a circular, core equator axis).

Fig. 1a. It shows an anti-neutrino, i.e. a trefoil loop. It also shows fluxloop forms for a muon or an electron, the extended source (sphere) characterizing the quasi-nonlocality of a stationary single particle.

Fig. 1b. This interlinkage of two loops $\overline{\eta}$ and λ represents a contribution to a $\overline{K}^{\circ}$ meson. To illustrate the topological (knot theoretical) relationships between the two loops, the space is subdivided by a toroidal surface (a flexible, sumetimes small, sometimes large toroid, having no restrictions on its size). One of the loops (the trefoil λ) is entirely outside the toroid; the other ($\overline{\eta}$) is entirely inside to permit independent spinning. The core reaches to the circular axis which is the core equatorial ring.

Fig. 1c. For a baryon, the space is subdivided into three regions by the two toroidal surfaces shown in this figure. The three loops may thus spin independently and they may share the core. These Figs. 1 represent the loops on the basis of a straight central axis. There is spinning about the straight axis and a whirling, rolling motion about the circular axis.

Fig. 2. It shows the alternative setting on the basis of two interlinked axis shown in Figs. 2,3,5,7, drawn as donuts. In this figure quarks are shown in the unrealistic, unlinked state, because a quantized fluxloop is a quark only when interlinked with other loop(s) which drastically influences its modes of spinning. The figure illustrated the winding numbers η: 2-1, ρ: 3-1, λ: 3-2. Quarks are assumed left-handed, antiquarks right-handed; their charge is given by the orientation of their magnetic moment with respect to spin.

Fig. 3. The same quarkloops as in Fig. 2 are shown here in relation to the left axis. As in Fig. 2 these are not representing leptons.

Fig. 4. For the alternative setting the two axes are shown as interlinked rings with dash-dot-dash lines. The domains to which two meson loops may be confined to permit independent spinning, are indicated by the surface separating the domains. One can easily visualize the modes of independent spinning of two fluxloops (i.e. rolling-whirling motions about the two axes respectively). The two fluxloops are drawn in Fig. 5, not in this Fig. 4. To facilitate the graphical representation, this surface is shown bound by one long-winded line; in reality, however; the surface reaches to infinity. One opening (connecting upper to lower with the far left region) is hidden behind the surface; we see the other axis passing through that opening. The surface separating the two domains may move one way or the other, closer to one or the other axis.

Fig. 5. Loop-antiloop contribution to a meson. We have represented in this figure, the two axes by the two donuts because this facilitates the illustration of the two loops $\bar{\lambda}$ (3,2) and η (2,1).

Fig. 6. The two dash-dot-dash lines may represent the axis in this figure of the alternative axes model. The present donuts are drawn to represent, for a baryon, the two toroidal surfaces which tripart space so that the (3,2) loop of Fig. 7 is entirely inside the right donut and (2,1) loops entirely inside the left donut, and the simple, large (1,0) loop, in the space between them which stretches all the way out. This permits independent spinning to the three loops, a necessary condition that quarks should satisfy.

Fig. 7. Interlinkage of the three loops of a baryon in the alternative model. In this figure, as in Fig. 5, the axes are pictured as donuts to facilitate the illustration of the winding numbers. The interlinked quarkloops are a (3,2), a (2,1) and a (1,0) loop.

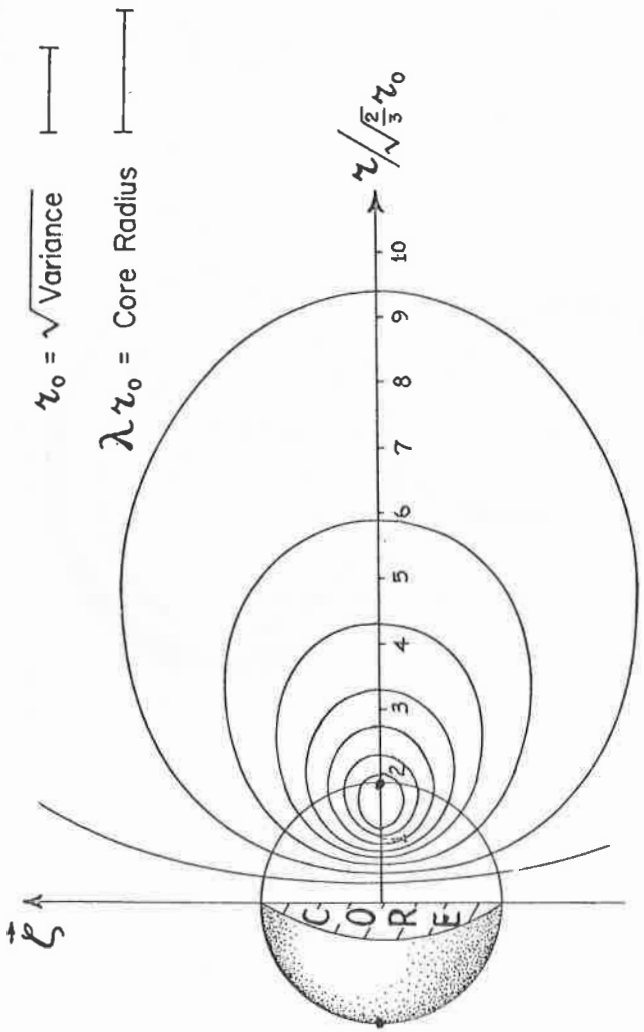

$r_0 = \sqrt{\text{Variance}}$
λr_0 = Core Radius
$r/\sqrt{\frac{2}{3}}r_0$
1
2
3
4
5
6
7
8
9
10
ζ
CORE

Figure 1

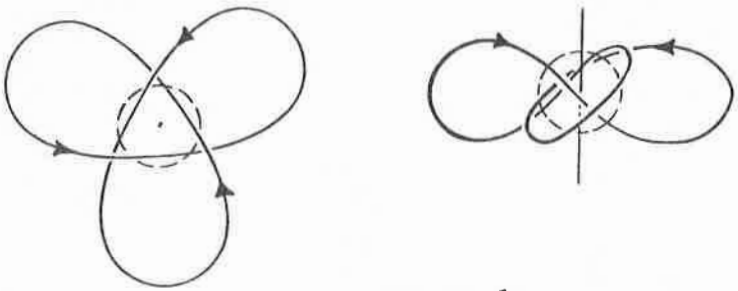


Figure 1a

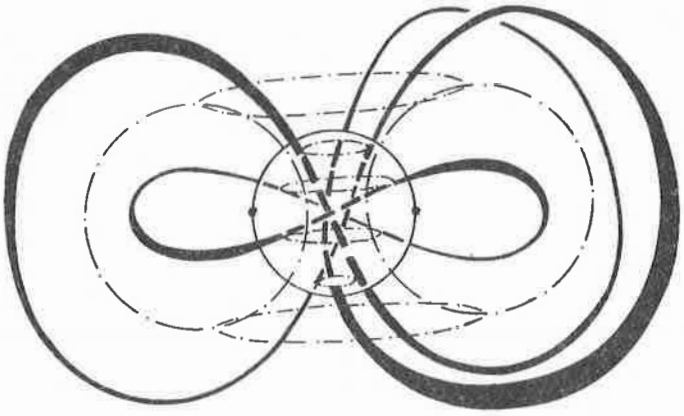


Figure 1b

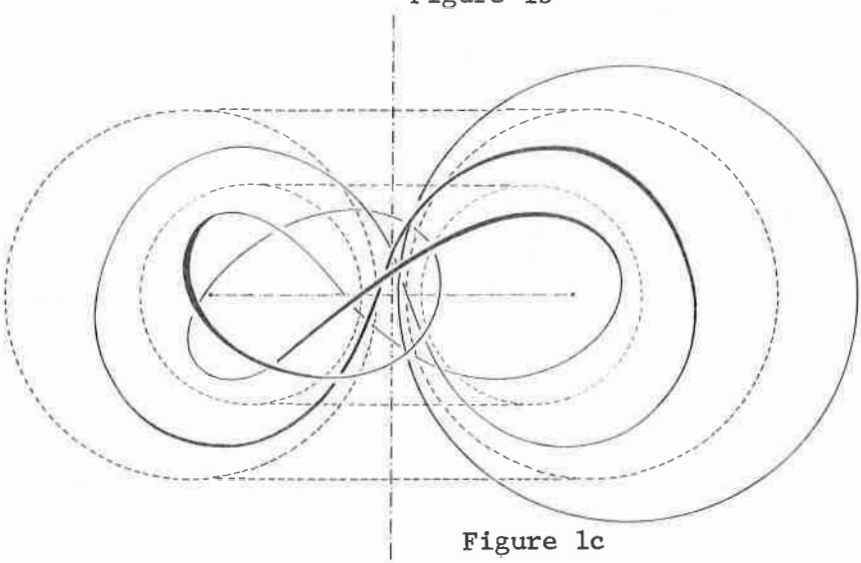
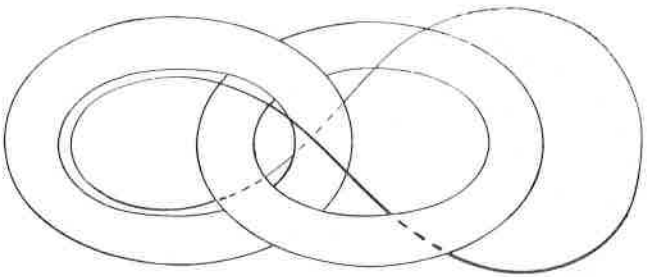
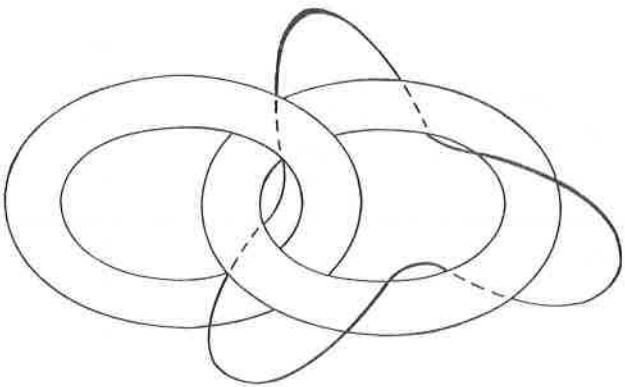


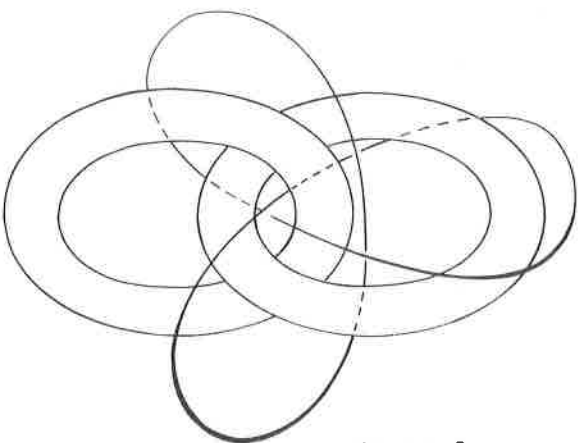
Figure 1c



$\mathfrak{h}(2,1)$

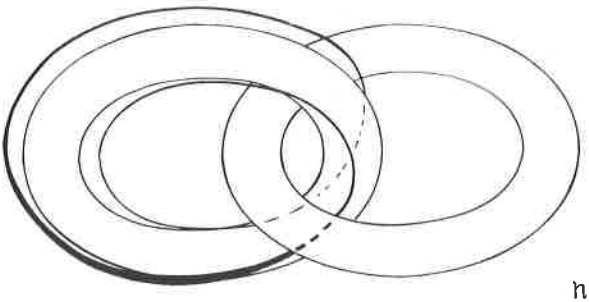


$\mathfrak{p}(3,1)$

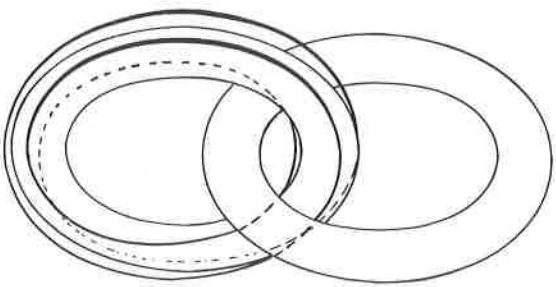


$\overline{\lambda}(3,2)$

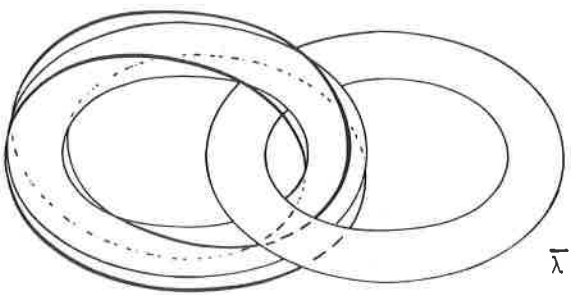
Figure 2



$n(2,1)$



$P(3,1)$



$\overline{\lambda}(3,2)$

Figure 3

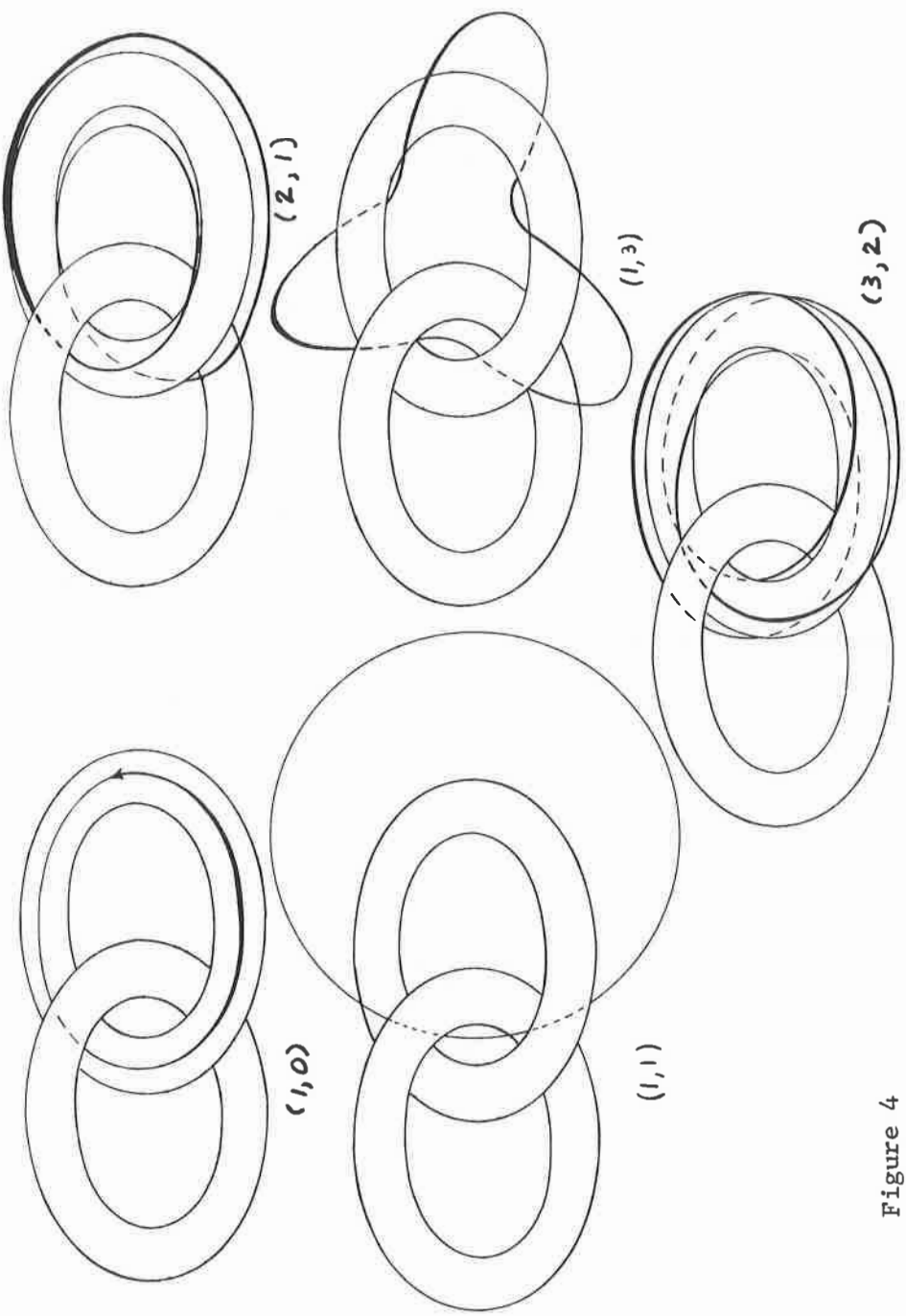


Figure 4

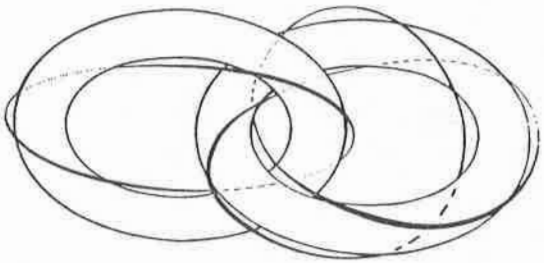


Figure 5

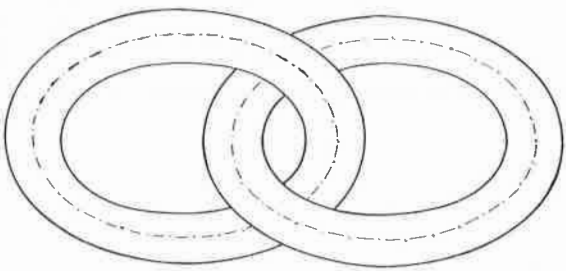


Figure 6

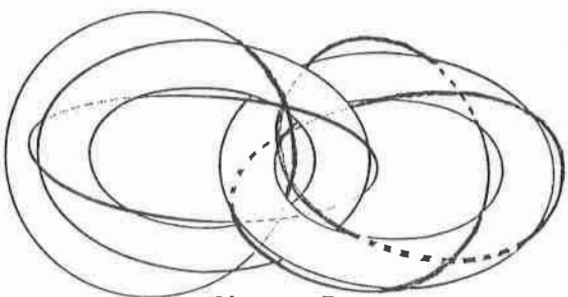


Figure 7

REFERENCES

1. Jehle, H., Boulder Lectures 10 B, (1968); 11 A, 493 (1969); Symm. and Quark Models (Ramesh Chand, ed., Gordon and Breach 1970); Rev. 3 D, 306 (1971).

2. L.Foldy in Quantum Theory, ed. by D.R. Bates (Ac. Pr. 1962); S. Tani, Progr. Th. Ph. 6, 267 (1951); K.M. Case, Phys. Rev. 95, 1323 (1954); M.H.L. Pryce, Pr. Roy. Soc. A 195, 62 (1948); A.O. Barut and S. Malin, Rev. Mod. Phys. 40, 632 (1968); L. Foldy, and S. Wouthuysen, Phys. Rev. 78, 29 (1950); H. Jehle and Wm. C. Parke, Phys. Rev. 137, B760 (1965).

3. M. Beg, Phys. Rev. Let. 13, 514 (1964); W. Thirring, Lecture at 4. Internationale Universitätswochen für Kernphysik, Schladming, February, March 1965; Acta Physica Aust. Supp. II 1965; Some developments of the quark model, Erice lectures 1970 (A Zichichi, ed.); V.F. Weisskopf, CERN Rept. 66-19 (1966); A. Dolgov, Phys. Let. 15, 84 (1965).

4. A.O. Barut, Proc. Second Coral-Gables Conference on Fundamental Interactions at High Energies, (Gordon & Breach, N.Y. 1970), p. 199-220, Phys. Rev. D 3, 1747 (1971); in Topics in Modern Physics-A Tribute to E.U. Condon (Colorado Assoc. Univ. Press, Boulder 1971), p. 15-45.

Part IV

ANALYTIC APPROXIMATION AND EXTRAPOLATION

PRINCIPLES OF ANALYTIC APPROXIMATION THEORY AS APPLIED TO ANALYSIS OF EXPERIMENTAL DATA

R. E. Cutkosky
Carnegie-Mellon University
Pittsburgh, Pennsylvania

Outline

This series of lectures is devoted to an exposition of some mathematical ideas which are unfamiliar to most physicists, but which have an enormous potential use in High Energy Physics as well as in other branches of science.

Section I reviews polynomial approximation theory, and has two purposes: To try to expose something of the beauty of the mathematical ideas contained in this subject, and to illustrate the point that the convergence properties of a sequence of approximating functions depend jointly on the data used and on analyticity.

Section II is an application of the theory of Section I, showing how to use a conformal transformation to obtain the most rapidly convergent polynomial approximation for representation of physical data.

Section III presents further mathematical tools which enable one to go beyond simple polynomial approximations, in order to obtain maximally accelerated expansions in circumstances in which the conformal transformation technique of Section II is too complicated to use effectively, and also to circumvent the apparent limitation in Section II to results based on asymptotic convergence estimates. The fundamental mathematical tool is the Hilbert space of analytic functions. In addition, the ideas of generalized interpolation and of a probability measure in the function space, which are important to physical applications, are presented.

Section IV is devoted to exposition of the application of the mathematical theory of Section III to a reformation of the conventional χ^2-minimization technique for fitting experimental data. It is shown how to construct the best sequence of functions for a linear representation of experimental data, and a convergence test function is constructed which can be used to solve the problem of estimating the truncation point and truncation error of the expansion.

Discussion of the increasingly numerous practical applications of the theory discussed here is omitted because of lack of space. The companion lectures by H. Pfister refer to some of these, and also include a discussion of several aspects of the theory that have been omitted from these lectures.

I. Polynomial Approximation

This section is a review of the theory of polynomial approximation, as discussed by Walsh [1], chapters 3 - 5. The entire subject, as mentioned by Walsh, can be considered as a matter of successive generalizations of the Taylor expansion. We are concerned here with the heuristics. For details, see Walsh.

A. Jacobi series

Consider ν points $z = \beta_1, \ldots, \beta_\nu$, and the polynomial of degree ν: $p(z) = (z-\beta_1) \ldots (z-\beta_\nu)$. The Jacobi series is an expansion of a given function.

$$f(z)=q_0(z)+q_1(z)p(z)+q_2(z)p(z)^2+\ldots+q_n(z)p(z)^n+\ldots \quad (1)$$

where each $q_i(z)$ is a polynomial of degree $\nu-1$. The $q_n(z)$ are determined successively by requiring $f(z)$ and its first n derivatives to be matched at the points points β_i. Let $S_n(z)$ denote the partial sum of the first $n+1$ terms. We have two formulas (where Γ lies inside a domain of analyticity and encloses the β_i):

$$S_n(z) = \frac{1}{2\pi i} \int_\Gamma \frac{p(t)^{n+1} - p(z)^{n+1}}{t-z} \frac{f(t)}{p(t)^{n+1}}\, dt \quad (2)$$

and

$$f(z)-S_n(z) = \frac{1}{2\pi i} \int_\Gamma \frac{p(z)^{n+1}}{p(t)^{n+1}} \frac{f(t)}{t-z}\, dt \qquad (3)$$

To prove these, note first that $S_n(z)$ as given by (2) is a polynomial in z of degree $n=\nu n+\nu-1$, and note from (3) that the first n derivatives of the L.H.S. vanish at $z = \beta_i$.

Now let Γ be a lemniscate curve, i.e. the locus $|p(z)| = \mu$ (a constant). Then on any interior lemniscate $|p(z)| = \mu_1<\mu$, the partial sums S_n converge (according to (3)) as

$$|f - S_n| < M \left(\frac{\mu_1}{\mu} \right)^{n+1}. \qquad (4)$$

We can rewrite this as follows. Let:

$$v(z) = \frac{1}{\nu} \log |p(z)| = \Sigma \frac{1}{\nu} \log |z-\beta_1| \qquad (5)$$

Evidently v(z) is the potential of ν point charges, each of strength-$1/\nu$, and the Γ's are equipotential curves. We have also:

$$(\mu_1/\mu)^n = e^{-(N-1)(v-v_1)} \qquad (6)$$

The Taylor expansion is, of course, the special case $\nu=1$.

B. Possibility of approximation on general curves.

1. Equipotentials

Let C be a given curve (or arc). Let $\varphi(z)=v(z)+iw(z)$ be the complex potential corresponding to a unit charge on C and the boundary condition $v=0$ on C. Let $\zeta(z)=e^{\varphi(z)}$. Then define C_ρ to be the equipotential $|\zeta|=\rho>1$.

2. Approximation of C by a lemniscate.

The idea is to replace the actual continuous charge distribution on C by a (large) number of equal point charges properly distributed. By choosing the number of points ν large enough, we can be sure that there is a lemniscate Γ outside of C but no where more distant from C than some preassigned amount. The curve Γ_ρ then lies slightly outside C_ρ, but by a controlled amount (and thus still in the domain of analyticity of a given function f).

3. Approximation by Polynomials on C.

We know that we can approximate on Γ by a Jacobi series with the error bound

$$|f-P_N| < M\,\rho^{-N} \quad (N \text{ is degree of polynomial}) \qquad (7)$$

Since C lies inside Γ, by the maximum-modulus principle, the error is smaller on C. Remark: This is an existence theorem for polynomial approximation, not a useful constructive technique.

4. Region of uniform convergence.

Lemma: Let $|P_N(z)| \leq L$ for $z \epsilon C$, P_N a polynomial of degree N. Then $P_N(z)/\zeta^N(z)$ is analytic outside C, including ∞. Its maximum modulus therefore occurs on C, where $|\zeta| = 1$. Therefore, generally, $|P_N(z)| \leq L\rho^N$ for z on C_ρ. Now, from (7), we find

$$|P_{N+1} - P_N| \leq (M + \rho M)\,\rho^{-N-1}, \quad z \epsilon C \qquad (8)$$

Therefore, for $\rho_1 < \rho$, we have, by the lemma,

$$|P_{N+1} - P_N| \leq \frac{\rho_1^{N+1}}{\rho^{N+1}}\,[M + M\rho], \quad z \epsilon C_\rho, \qquad (9)$$

so $P_n \to f$ at any point in the interior of C_ρ (and also uniformly on and within C_ρ).

C. Maximal convergence

Given a function f, there is an upper bound R to the values of ρ for which f is analytic on and within C_ρ. Either $R = \infty$, or there is a curve C_R on which $f(z)$ fails to be analytic. Thus, for any $\rho<R$, there are polynomials $P_n(z)$ for which (7) holds. However, there can be no sequence for which (7) holds with $\rho>R$, because this would imply uniform convergence on C_R, and hence analyticity on C_R. Thus, R is said to be the greatest possible rate of geometric convergence. Also, let

$$E_n = \text{Max}_{z\in C} |f(z) - P_n(z)|.$$

Then $\limsup E_n^{1/n} = 1/R$.

D. Best approximation

1. Chebyshef approximation

The Chebyshef polynomial approximant is the polynomial $t_n(z)$ in which the n+1 coefficients are chosen to minimize the error bound

$$E_n = \max_{z\in C} |f(z) - t_n| \tag{10}$$

Since we know that there are polynomials, for any $\rho<R$, satisfying (7), the t_n must a forteriori also satisfy

$$|f(z)-t_n(z)|<M\rho^{-n} \tag{11}$$

and hence converge maximally on C and uniformly on and within C_ρ, $\rho<R$.

2. Least squares approximation

The general case of least pth power approximation is discussed by Walsh, but the proof is a bit lengthier. Here we define the approximants $\pi_n(z)$ by minimizing the integral

$$E_n = \int_C W(z)\ |dz|\ |f(z) - \pi_n(z)|^2 \qquad (12)$$

The proof uses the following lemma: Let $P_n(z)$ be a polynomial of degree n, such that $\int |P_n(z)|^2\ |dz| < L^2$. Let $Q(z) = P_n(z)\zeta^{-n}$. Then on Γ, $|Q| = |P|$, we have (z outside Γ)

$$\frac{Q(z)^2}{\zeta(z)} = \frac{1}{2\pi i} \int \frac{Q(t)^2}{\zeta(t)} \frac{dt}{z-t}, \qquad (13)$$

since the integrand goes to zero as $1/t^2$ at ∞. Hence

$$\frac{|Q(z)|^2}{\rho} < M(\rho) \times L^2, \qquad z \in \Gamma_\rho$$

whence follows the lemma

$$|P(z)| < L \cdot L'(\rho) \times \rho^n, \qquad z \in \Gamma_\rho \qquad (14)$$

Now return to 12. Since maximally converging polynomials exist, and since π_n minimizes the integral, we have $E_n < M/\rho_1^{2n}$, for $\rho < \rho_1 < R$. We assume that W(z) is positive, and bounded from zero. Then we can replace W by its minimum value in the integrand. Furthermore, we use the inequality

$$|a+b|^2 \leq 2|a|^2 + 2|b|^2$$

with $a \sim f - \pi_n$ and $b \sim -f + \pi_{n+1}$, to get

$$\int_C |\pi_{n+1} - \pi_n|^2\ |dz| < M_1 \rho_1^{-2n} \qquad (15)$$

Now let $\rho' = \rho_1/\rho > 1$; on $C_{\rho'}$, by the lemma,

$$|\pi_{n+1} - \pi_n| < M_2 \left(\frac{\rho_1}{\rho}\right)^n \times \frac{1}{\rho_1^n} = M_2/\rho^n \qquad (16)$$

From this point, it is easy to show that the $\pi_n(z)$ converge to f on C maximally, and hence converge uniformly on and within C_ρ, $\rho < R$. It is possible

to drastically weaken the boundedness condition on $W(z)$.

E. Example

Let C be the segment of the real axis $-1 \le z \le 1$. Then the C_ρ are the ellipses with foci at $z = \pm 1$; $\rho = a+b$, where a and b are the semi-axes. Consider least squares approximation with some weight function W. For such an approximation method, it is convenient to construct the set of orthonormal polynomials $q_n(z)$, and write the n^{th} approximant as

$$P_n(z) = \sum_1^n A_m q_m(z) \tag{17}$$

Examples are the Legendre polynomials, etc. In general, we have convergence inside an elliptical domain, and from the error bound on C, we find that $\limsup |A_n|^{1/n} = 1/R$, where C_R is the ellipse passing through the nearest singularity.

II. Polynomial Expansion in an Optimized Variable

A. Acceleration of convergence

Let us suppose we wish to fit a function $f(x)$ on the curve Γ_1 (the "physical region"). We make a conformal transformation to a new variable $z(x)$ and approximate by polynomials in z. The problem is to make the best choice of z.[2,3]

1. Conformal Transformation

First, suppose $x \to z$ maps Γ_1 into the segment $C_1 = (-1, 1)$ and that it maps some curve Γ, within which f is analytic, but on which f has a singularity, onto a unifocal ellipse C_R. The mapping is given explicitly by

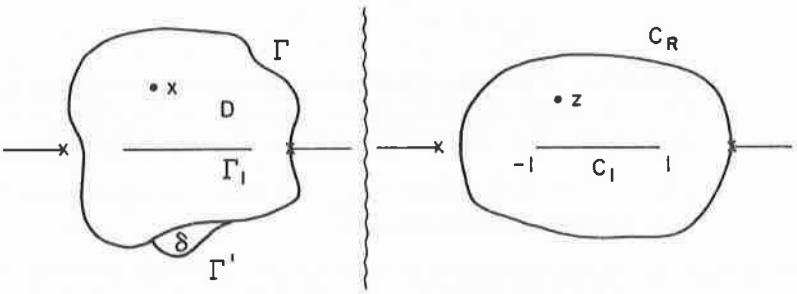
$$z = \sinh \Phi(x) \tag{1}$$

where here $\Phi = v(x) + iw(x)$ and $v(x) = 0$ on Γ and $v(x) = -V$ (a constant) on Γ_1; the value of V is defined by the requirement that the net charge on Γ_1 be -1. Then $\zeta(z) = e^{V+\Phi}$ corresponds to the

function defined in Section I, and $R = e^{V}$ determines the rate of maximal convergence. Furthermore, a polynomial approximation in z converges in the interior of the ellipse C_R, and hence in the (open) interior D of the curve Γ in the x-plane.

2. Optimized Conformal Transformation

Suppose that the domain D enclosed by Γ is enlarged by the addition of a small piece δ as shown:



The boundary Γ' of $D' = D \cup \delta$ is mapped onto a curve C_R, by a modified mapping function involving a new complex potential $\Phi' = V' + iW'$. We expect that enlarging the domain of convergence will increase the rate of convergence and proceed to prove this.

Let $v(x) = V'(x) - V(x)$, then $v(x)$ satisfies the boundary conditions:

$$v = 0 \text{ on } \Gamma_1$$

$$v = V'(x) \text{ on } \Gamma, \quad (V'(x) \leq 0) \tag{2}$$

Let $G(x,x')$ be the Green's function satisfying the following boundary conditions:

$$G = 0 \text{ for } x \text{ on } \Gamma$$
$$G(x,x') = g(x') \text{ for } x \text{ on } \Gamma_1 \tag{3}$$
$$\int_{\Gamma_1} \frac{\partial}{\partial n} G(x,x') |dX| = 0$$

The last integral, taken over both sides of Γ_1, says that there is no net charge on Γ_1.

From Green's theorem

$$v(x)=\int_{\Gamma+\Gamma_1}\left(G(x,x')\frac{\partial}{\partial n'}v(x')-v(x')\frac{\partial}{\partial n'}G(x,x')\right)|dx'| \tag{4}$$

we obtain, using the boundary conditions,

$$v(x)=-\int_{\Gamma} v(x')\frac{\partial}{\partial n'}\,G(x,x')\,|dx'| \tag{5}$$

Since the induced charge density on Γ is everywhere negative, and since $v \leq 0$ on Γ, we have

$$v(x) < 0,\quad x \in D. \tag{6}$$

From (6) we conclude, first:

$$V(\Gamma')-V(\Gamma_1)>V(\Gamma)-V(\Gamma_1)$$

or
$$R' > R \tag{7}$$

so that the convergence rate on Γ_1 is increased, and second, that at an arbitrary point $x_p \in D$

$$V(\Gamma)-V(P)=-V(z_p)>-V'(z_p')=V(\Gamma')-V'(P), \tag{8}$$

so that the convergence rate at x_p is also increased. Thus, the more of the entire domain of analyticity of $f(x)$ that is mapped into the unifocal ellipse, the faster is the convergence, both to the data given on Γ_1 and at an arbitrary point in the domain of analyticity. I call this relation between the full use of analyticity and rapid convergence the "convergence principle".

Finally, we ask whether a further improvement in the convergence rate can be obtained by mapping Γ_1 and Γ, not into a line segment and a unifocal ellipse, but into a pair of simultaneous equipotential curves for some other potential problem. The answer is no, because the potential

difference is a conformal invariant; Any further mapping would leave the convergence rates unchanged.

B. Effects of Errors in Data.

1. A Simple Model of Experimental Uncertainty.

Ciulli [3] represents the experimental function by $h(x)$; and assumes that the true physical function f satisfies

$$|f(x) - h(x)| < \epsilon, \qquad x \in \Gamma_1 \tag{9}$$

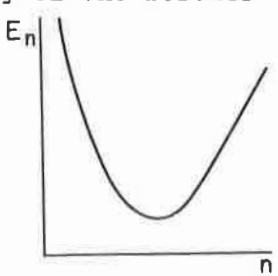
In order to obtain an error estimate for the extrapolated polynomial approximant, it is necessary to assume more about the function f, namely, some boundedness or smoothness property holding also on the curve Γ which contains points of singularity.

By assuming some boundedness condition, it is also possible to sharpen some of the convergence estimates given in Section I. In particular, if $|f| \leq M_0$, we can take M independent of ρ in Eq. (1.11), and also take $\rho = R$.

Under the conditions given above, Ciulli obtains the error bound

$$|f(x) - t_{hn}(x)| \leq \epsilon \rho^n + K(\rho/R)^n \equiv E_n \tag{10}$$

where $t_{hn}(x)$ is the Chebyshef approximant of degree n to the experimental function h. (Actually, this formula doesn't appear in Ciulli's paper, but it is a trivial corollary [4] of the results given.) This formula shows that the extrapolation error behaves qualitatively as shown in the figure: It first decreases with increasing n, when the error bound is dominated by the estimate of truncation error.



If n is too large, it increases again, as a result of instability caused by noise in the experimental data.

There is evidently a best value of n to use; approximating n as a continuous parameter, we have

$$0=\frac{\partial E_n}{\partial n}=\epsilon\rho^n \ln \rho + K(\rho/R)^n \ln(\rho/R), \tag{11}$$

which leads to

$$n = \ln\left[\frac{K \ln(R/\rho)}{\epsilon \ln \rho}\right] / \ln R. \tag{12}$$

Substituting from (11) back into (10), we have for $E = \text{Min } E_n$:

$$E = K\left(\frac{\rho}{R}\right)^n \frac{\ln \rho}{\ln R} \tag{13}$$

According to 12, $n\to\infty$ when $\epsilon\to 0$, therefore $E\to 0$ for $\rho<R$

2. Optimized Stability [4]

In the presence of experimental errors we should use (10) as the error estimate, rather than the corresponding expression with $\epsilon=0$, when we establish optimality. In the variation of the domain considered in section A2, we have now:

$$\delta E = n\epsilon\rho^{n-1}\delta\rho + nK(\rho/R)^{n-1}\delta(\rho/R), \tag{14}$$

where we omit the variation of n because of (11). Using the relation between ϵ, K, and n given by (11) to eliminate K, one obtains after some algebra the expression

$$\delta E = n\epsilon\rho^n \frac{(\ln R)^2}{\ln(R/\rho)} \delta\left(\frac{\ln \rho}{\ln R}\right) \tag{15}$$

The coefficient of $\delta(\ln \rho / \ln R)$ is positive.

However, we can identify the quantity

$$v = \ln \rho / \ln R \tag{16}$$

as a potential (the harmonic measure of Nevanlinna) satisfying the boundary conditions

$$\begin{aligned} v &= 0 \ , \quad x \in \Gamma_1 \\ v &= 1 \ , \quad x \in \Gamma \end{aligned} \tag{17}$$

Now in the variation of the domain we keep the potentials on the boundary fixed; then it is clear that $\delta v < 0$ at all points interior to Γ, from which our theorem follows.

With this simple model of experimental errors we have been able to show explicitly that accelerated convergence and increased stability against noise are closely related. I refer to this relation, in general, as the "stability principle".

III. Approximation by Normed Functions

A. General Remarks

There are many cogent reasons for wanting to improve on the techniques which use polynomial expansions. One reason is that we have made use so far of only asymptotic convergence properties, whereas in practice we want to use a small number of terms. This limitation may be only apparent however - arising because we have so far used rather simple mathematical ideas. A more important reason is that we would like to take account more precisely of the actual distribution of experimental information at discrete points in the physical region. Furthermore, we want to take account of experimental uncertainty in a more exact way than is given by the model of the last section. The treatment of experimental uncertainty is left to section IV of these notes. However, the desire to incorporate Gaussian statistical errors into the formalism is one of the primary motives for treating the boundedness and smoothness limitations on the functions

also in terms of a quadratic norm and for introducing an a priori measure in the function space.[5]

B. Hilbert Space of Analytic Functions

1. Definition and Examples

a. The function space A under consideration consists of functions which are analytic in some suitable open domain D and whose limiting values on the boundary Γ of D possess a suitable Hilbert norm.

b. The simplest example[6]

Let the functions be analytic within the unit circle, and let the inner product be:

$$(f,g)=\frac{1}{2n}\int_{|u|=1} f(u)*g(u)\,|du| \qquad (1)$$

If we write the Taylor expansion of the functions as $f = \Sigma a_n u^n$, $g = \Sigma b_n u^n$, then

$$(f,g) = \Sigma a_n {}^* b_n \qquad (2)$$

c. Generalization to different weight functions.

Let $W(u)$ be real and positive, and suppose that

$$(f,g) = \frac{1}{2\pi}\int_{|u|=1} W(u)\,f(u)^* g(u)\,|du| \qquad (3)$$

Then define $\gamma(u)$ to be analytic in the unit circle, such that on $|u| = 1$

$$\mathrm{Re}\,\gamma(u) = -\tfrac{1}{2}\ln W(u) \qquad (4)$$

Then let

$$\varphi(u) = e^{\gamma(u)}; \qquad (5)$$

$\varphi(u)$ is uniquely determined up to a constant phase factor. On $|u|=1$, $|\varphi|^2=1/w$. With the definitions

$$f = \varphi F, \quad g = \varphi G, \quad \text{etc.} \tag{6}$$

The inner product (3) is transformed into the the form (1) for F and G.

d. To impose smoothness conditions on Γ. We may replace (2) by

$$(f,g) = \Sigma C_n a_n {}^* b_n \tag{7}$$

Now suppose f(u) has on $|u|=1$ a finite number of singularities of the form

$$f(u) \sim (u-u_o)^{\nu}, \quad |u_o|=1, \quad u \sim u_o . \tag{8}$$

Then $a_n \sim n^{-\nu-1}$ for large n, and we may take

$$C_n \sim n^{2\nu+2} \tag{9}$$

for large n, because then the sum (7) is marginally non-convergent. If $p=2\nu+2$ is a positive integer, we may write an integral expression such as (1) in which p derivatives of f or g are taken. (A marginally non-convergent sum is not unreasonable at this point because we will later consider closures of A).

2. The Reproducing Kernel[7]

a. Definition

The reproducing kernel is a function H(x,y), such that for $y \in D$, H(x,y) (as a function of x) lies in A, and possesses the following reproducing property:

$$(H(\cdot,y), f) = f(y) \tag{10}$$

The notation introduced in (10) is that the dot $\cdot$ symbolizes the dummy variable with

respect to which the inner product is evaluated. It follows from (10) that

$$(H(\cdot,x),H(\cdot,y)) = H(x,y), \tag{11}$$

and hence that H is Hermitian,

$$H(x,y) = H(y,x)^*, \tag{12}$$

and also that H(x,y) is an analytic function of y^*.

b. Examples of reproducing kernels

For the Hilbert space (1.b) the reproducing kernel is [6,7]

$$H(u,v) = \frac{1}{1-uv^*} \tag{13}$$

Proof:

$$\frac{1}{2\pi}\int_{|u|=1} \frac{f(u)}{(1-uv^*)^*}\,|du| = \frac{1}{2\pi i}\int_{|u|=1} \frac{du}{u-v} f(u) = f(v) \tag{14}$$

Where we have used the relations holding on $|u|=1$, that $|du|=du/(iu)$ and that $uu^*=1$. More generally, if $||f||^2=\Sigma C_n |a_n|^2$, then

$$H(u,v)=\Sigma\frac{(uv^*)^n}{C_n} \tag{15}$$

c. Elliptical domain

As another example, if the domain of analyticity is the interior of a unifocal ellipse C_R, as defined in section I.E, and if the norm is defined on C_R as the square-integral with a weight function proportional to $|z^2-1|^{-\frac{1}{2}}$, then the reproducing kernel can be written in the form

$$H(z,z') = \Sigma H_n T_n(z) T_n(z')^*, \tag{16}$$

where the T_n are the conventional Chebyshef polynomials, and

$$H_n = (R^{2n}+R^{-2n}+2\delta_{no})^{-1} \tag{17}$$

C. Generalized Interpolation

1. Linear Functionals and Basis Functions

Ordinarily one considers interpolation to the values or values of derivatives of a function at certain points: viz $f_k=f(x_k)$ or $f_k=f'(x_k)$. It is convenient to generalize this idea to the consideration of arbitrary linear functionals of f, for which we introduce the notation I_k: [5]

$$f_k = I_k f \tag{18}$$

Also, we introduce an adjoint I_k^+, acting, for instance, on H(x,y), as follows:

$$H_k(x) \equiv H(x,\cdot)I_k^+ \equiv (I_k H(\cdot,x))^* \tag{19}$$

Eq. (19) also defines an associated basis function $H_k(x)$. Also, we define:

$$H_{\ell k} \equiv I_\ell H_k = I_\ell H(\cdot,\cdot\cdot)I_k^+ \tag{20}$$

We have then a simple theorem:

$$(H_\ell, H_k) = (H(-,\cdot)I_\ell^+,\ H(-,\cdot\cdot)I_k^+)$$

$$= I_\ell H(\cdot,\cdot\cdot)I_k^+ = H_{\ell k}. \tag{21}$$

The $H_k(x)$ are thus convenient basis functions for A, because their inner products can be calculated immediately from the reproducing kernel, let us call this the "natural basis".

2. Interpolating functions

Suppose that we are given the N numbers f_k, $k=1,\ldots,N$. We want to approximate the function f. We say that a function $f_N(x)$ interpolates the f_k, if

$$I_k f_N = f_k, \quad k=1,\ldots,N. \tag{22}$$

Such an $f_N(x)$ is obviously some sort of approximation to the original $f(x)$. In case $f_N(x)$ depends linearily on the f_k, we say that $f_N(x)$ is a linear interpolation. A particular linear interpolation

$$F_N(x) = \Sigma_1^N H_k(x)\beta_k \tag{23}$$

can be constructed from the natural basis functions provided the β_k are chosen appropriately:

$$f_\ell = I_\ell F_N = \sum_{k=1}^{N} H_{\ell k}\beta_k, \tag{24}$$

or, in an obvious matrix notation, with H_N refering to the NxN matrix $H_{\ell k}$,

$$f = H_N\beta, \quad \beta = H_N^{-1} f \tag{25}$$

We now prove an important theorem: Of all functions $f_N(x)$ interpolating the N values f_k, the function $F_N(x)$ given by (23-25) has the smallest norm. In general, we may write

$$f_N(x) = F_N(x) + \varphi(x) \tag{26}$$

where $I_\ell\varphi=0$, $\ell=1,\ldots,N$. Hence

$$0 = I_\ell(H(\cdot,\cdot\cdot),\varphi) = (H_\ell, \varphi) \tag{27}$$

But

$$||f_N||^2 = ||F_N||^2 + ||\varphi||^2 + 2R_e\Sigma(\varphi, H_k)\beta_k,$$

and according to (27), we have

$$||f_N||^2=||F_N||^2+||\varphi||^2, \text{ Q.E.D.} \tag{28}$$

D. Convergence properties

We will now show that $\lim_{N\to\infty} F_N(x)$ exists. Let the functions $\{A_k(x)\}$ denote an orthonormalization of the $\{H_k(x)\}$ according to the Schmidt process. We can write

$$F_N(x) = \sum_1^N a_k A_k(x) \tag{29}$$

We have

$$\sum_1^N |a_k|^2 = ||F_N||^2 \leq ||f||^2 \tag{30}$$

from which it follows that the sequence $F_N(x)$ converges in norm. It is an important consequence of the existence of a reproducing kernel that convergence in norm implies uniform convergence for points in any closed domain $d \subset D$; we sketch the proof.

$$\text{Let } \Delta_M^N(x) = \sum_M^N a_n A_n(x); \tag{31}$$

We have

$$||\Delta_M^N||^2 = \sum_M^N |a_k|^2 < \epsilon$$

provided $M > M(\epsilon)$, by hypothesis. Using the property (10) of the reproducing kernel, we have:

$$\begin{aligned} |\Delta_M^N(x)|^2 &= |(H(\cdot,x),\Delta_M^N)|^2 \\ &=(H(\cdot,x),\Delta_M^N)(\Delta_M^N,H(\cdot,x)) \end{aligned} \tag{32}$$

Now, by the Schwartz inequality, we have

$$|\Delta_M^N(x)|^2 \le ||\Delta_M^N||^2 (H(\cdot,x),H(\cdot,x)) = ||\Delta_M^N||^2 H(x,x) \tag{33}$$

Hence the sequence $\{F_N(x)\}$ converges uniformly in any closed domain in which $H(x,x)$ is bounded uniformly. This of course applies to any interior domain of D, and it may also include the boundary of D for Hilbert spaces which incorporate smoothness conditions on the boundary.

E. Measure in $\bar{A}$

1. Properties of the Wiener-measure

We define a Gaussian relative measure:(5,8)

$$R(f) = \exp(-\tfrac{\mu}{2}||f||^2) \tag{34}$$

which can be interpreted as a relative probability (for the formal construction of a Bayes' hypothesis) of functions in $\bar{A}$, relative to a standard reference function which is taken here as $f_{ref}=0$. In any complete orthonormal basis $\{E_n(x)\}$ we have, formally,

$$H(x,y) = \Sigma_a E_a(x) E_a(y) \tag{35}$$

It is convenient here to specialize to the case of real analytic functions, which includes all important applications, and then the $f_k=I_k f$ can be taken as real, and the $H_k(x)$ and $E_a(x)$ can also be real-analytic. The quantities

$$f_a = (E_a, f) \tag{36}$$

which enter into the representation

$$f(x) = \Sigma_a f_a E_a(x) \tag{37}$$

are then also real.

In such a basis, we have

$$R(f) = \exp(-\tfrac{1}{2}\Sigma f_a^2) \tag{38}$$

where, temporarily, we have set $\mu=1$. Using (38) it is clear that

$$\langle f_a \rangle_A = 0$$
$$\langle f_a f_b \rangle_A = \delta_{ab} \tag{39}$$

Thus, for functions written in the form (37), we have

$$\langle f(x) \rangle_A = 0$$
$$\langle f(x) f(y)^* \rangle_A = H(x,y) \tag{40}$$

Eq. (40) shows that H(x,y) provides a complete description of the Hilbert space A as well as of the measure R defined over it. It is therefore natural to take H(x,y) as the fundamental object, and use it to define A; it may be easier to develop an intuition about the correlation function for boundary values than about the weighting and smoothing that goes into the definition of the norm.

2. Averages over constrained functions

We now consider the restricted Hilbert space A/E of functions f which satisfy the N constraints

$$I_k f = f_k \quad , \quad k = 1, \ldots, N$$

where the f_k are supposed given. This restricted space can be obtained explicitly by orthonormalizing f to the basis H_k. Furthermore, the mean value of f(x) over A/E with respect to the measure (34) or (38) is easily seen to be

$$\langle f(x) \rangle_{A/E} = F_N(x), \tag{41}$$

which just expresses the fact that with a Gaussian measure, the average values and the most likely values are equal.

Now let $f_N(x)$ be any function which interpolates the f_k in the sense defined above, and consider the mean square deviation of $f_N(x)$ from the correct value $f(x)$:

$$\Delta_N(x)^2 \equiv \langle |f_N(x)-f(x)|^2 \rangle_{A/E} = \langle |(f_N-F_N)+(F_N-f)|^2 \rangle_{A/E} \tag{42}$$

Using (41) and the fact that f_N-F_N is just a number independent of $f\in A/E$, we obtain

$$\Delta_N(x)^2 = |f_N-F_N|^2 + \langle (F_N-f) \rangle^2_{A/E} \tag{43}$$

Since $F_N(x)$ is a linear function of the f_k, the second term in (43) is independent of the f_k and is a constant if one performs a further average over the f_k. Likewise, if $f_N(x)$ is also a linear interpolation, i.e. depends linearily on the f_k, it is not hard to perform a further average over the first term also. In any case, it is clear that on the average over A, any arbitrary interpolation will be worse than the minimum norm interpolation $F_N(x)$.

3. Completeness

It was shown in section D that the minimum norm interpolation $F_N(x)$ converges uniformly, but the conditions under which this convergence is to the correct function $f(x)$ have not been discussed. Let us suppose that, associated with the sequence of linear functional operators I_k, there is some sequence of interpolating functions $f_N(x)$ which converges uniformly for $x\in d\subset D$. (Here d is also a continuum.) For example, let I_k be the k^{th} derivative operator at some point $x_o\in D$, and let $f_N(x)$ be the sum of the first N terms of the Taylor expansion around x_o. In this case, $f_N(x)\to f(x)$ for $x\in d$, where d is any circle centered around x_o contained in D. Now, if for all $f\in A$, there is some f_N such that $\Delta_N(x) = |f(x)-f_N(x)|\to 0$, it follows from (43) that for almost all $f\in A$, also $|f(x)-F_N(x)|\to 0$ for $x\in d$ (and indeed, no slower). Thus, in this case,

$F_N(x) \to f(x)$ for $x \in d$; since $F_N(x)$ has a uniform limit (see section D) for all $x \in d' \subset D$, $F_N(x) \to f(x)$ for $x \in d'$, uniformly in every $d' \subset D$.

Stated more s ccinctly, the $\{I_k\}$ are complete if one can find some prescription for always reproducing the uniquely correct $f(x)$ from the sequence of values $\{f_k\}$. For this "test prescription", it is only necessary to prove convergence in some d which is a continuum and which lies inside D.

IV. Treatment of Data With Errors

A. Maximum-Likelihood Method

1. Notation and Definitions

Let us denote the "true physical function" by $F(x)$, and the "true values" of the measured quantities by F_k. It is assumed here, that the function $F(x)$, which lies in the space A defined in Section III, is measured directly. That is, it is like a differential cross section, rather than a scattering amplitude which is not linearly related to the data. The experimental values are $\mathfrak{F}_k$. It is assumed that

$$\langle \mathfrak{F}_k \rangle_E = F_k \tag{1}$$

where $\langle\rangle_E$ means the average over repeated experiments, and that the errors are Gaussian with the covariance matrix

$$\langle (\mathfrak{F}_k - F_k)(\mathfrak{F}_\ell - F_\ell) \rangle_E = v_{k\ell} \tag{2}$$

The classical χ^2 is:

$$\chi^2 = \Sigma (f_k - \mathfrak{F}_k) W_{k\ell} (f_\ell - \mathfrak{F}_\ell) \tag{3}$$

where $f(x)$ is the fitted function, $f_k = I_k f$ are the fitted values, and $W_{k\ell} = (v^{-1})_{k\ell}$. In many simple cases, W is diagonal, but it is convenient to adopt a matrix notation.

Minimization of χ^2 can also be looked upon as maximization of the conditional probability

$$P(\mathfrak{F}/f) = \text{const} \times e^{-\frac{1}{2}\chi^2} \tag{4}$$

The more general maximaum likelihood method involves construction of a Bayes' hypothesis with an a priori probability P(f); then one maximizes $P(\mathfrak{F}/f) \times P(f)$. In the standard approach, one decides on a specific parametrization of f(x), and implicitly assumes that the a priori probability for these explicit parameters is uniform, and also implicitly assumes that the a priori probability for all hidden parameters is a delta-function centered at some value, say, zero. A more natural Bayes' hypothesis is provided by the measure we introduced in Section III. . We then minimize the quantity

$$\begin{aligned} X^2 &= -2\ell n(P(\mathfrak{F}/f) \times R(f)) \\ &= \chi^2 + \Theta \end{aligned} \tag{5}$$

where $\Theta = \mu ||f||^2$. (Any constant factor in the probability can be ignored).

2. Fixed Scale Approximation

Suppose, for illustration, that μ is a known number. In any case we can replace $||f||^2$ by $||F_N||^2$, where F_N is the minimum-norm function interpolating the fitted values f_k. According to (3.25), we have

$$\begin{aligned} ||F_N||^2 &= \Sigma_{k\ell mn}\ f_k\ (H_N)^{-1}_{k\ell}\ H_{\ell m}\ (H_N^{-1})_{mn} f_n \\ &= \Sigma_{k\ell}\ f_k\ (H_N)^{-1}_{k\ell} f_\ell \end{aligned} \tag{6}$$

It is convenient to write (6) and other equations following in a matrix notation, where f denotes an N-component vector, H, v, and W are N×N matrices, etc. Then (6) becomes: $||F_N||^2 = f^{+} H^{-1} f$, and Eq. (5) takes the form

$$X^2 = (f^+ - \mathfrak{f}^+) v^{-1} (f - \mathfrak{f}) + \mu f^+ H^{-1} f \tag{7}$$

Minimizing (7) with respect to f, we have

$$(v^{-1} + \mu H^{-1}) f = v^{-1} \mathfrak{f} \tag{8}$$

B. Eigenvectors and Eigenvalues

1. Simultaneous Diagonalization of W and H.

To write (8) in a more transparent form, as well as to aid in calculating the inverse of H (which in practical cases is nearly singular) and as a preliminary to the further development, we find the eigenvectors $P_{k\alpha}$ and eigenvalues λ^α of H with respect to the weight W of the data:

$$HWP_\alpha = \lambda^\alpha P_\alpha \tag{9}$$

The P_α are orthogonal, and are normalized according to

$$P_\alpha^+ W P_\beta = \delta_{\alpha\beta} \tag{10}$$

This is equivalent to constructing a set of basis functions

$$P_\alpha(x) = \Sigma_{k\ell} H_k(x) (H^{-1})_{k\ell} P_{\ell\alpha} \tag{11}$$

These new basis functions are orthogonal with respect to the inner product defined in A as well as with respect to the weight of the data. They have been normalized by W, and thus have roughly the same size, at the data points. Their norm is λ_α^{-1}.

Let the λ's be ordered: $\lambda^1 \geq \lambda^2 \geq \ldots \geq \lambda^N$. In practice, the λ's decrease like a rapidly convergent geometric series, so that the standard iterative technique for calculating the $P_{k\alpha}$ and λ^α works very well.

2. The Fixed Scale Approximation

We express the experimental data and fitted values in the form

$$\mathfrak{d} = \Sigma_\alpha P_\alpha C_\alpha \quad , \qquad C_\alpha = P_\alpha^+ W\mathfrak{d} \quad ,$$

$$f = \Sigma_\alpha P_\alpha \gamma_\alpha \quad , \qquad \gamma_\alpha = P_\alpha^+ Wf \quad . \tag{12}$$

With this notation, equation (7) takes the form

$$X^2 = \Sigma_\alpha (C_\alpha - \gamma_\alpha)^2 + \mu \Sigma \gamma_\alpha^2 / \lambda_\alpha \tag{13}$$

The best value for γ_α is thus

$$\gamma_\alpha = C_\alpha/(1+\mu/\lambda_\alpha) = \lambda_\alpha C_\alpha/(\mu+\lambda_\alpha) \tag{14}$$

This formula shows that the expansion of f in terms of the P_α is effectively cut off at the point K where $\lambda^K \approx \mu$. However, this cut-off is not sharp, but is spread out over a range of values $\alpha \approx K$.

In contrast to the result obtained in Section II.B, the cut-off point K does not depend on the point to which one wishes to extrapolate. This is a result of using Gaussian distributions for the model of experimental uncertainty as well as for the means of incorporating information on boundary values.

C. Elimination of the Scale Factor

1. Maximum Likelihood [9]

In general, the scale factor μ is not known. A simple way to estimate μ is by extending the maximum likelihood principle to include μ, in which case the μ-dependent terms in the normalization factor of P(f) must be included. If $M \leq N$ terms in the eigenvector expansion of f are included, we have

$$P(f) = \pi_1^M \left(\frac{\mu}{2\pi\lambda_n}\right)^{\frac{1}{2}} \exp\left(-\frac{\mu}{2}\Sigma_1^M \gamma_a^2/\lambda_a\right), \tag{15}$$

so that

$$X^2 = -2\,\ell nP = \chi^2 + \mu\Sigma_1^M \frac{\gamma_n^2}{\lambda_n} - M \log \mu + \text{constant} \tag{16}$$

If we first minimize with respect to μ, we obtain

$$\mu = M/\Sigma\gamma_a^2/\lambda_a, \tag{17}$$

and obtain the expression

$$X^2 = \chi^2 + \Phi \tag{18}$$

where we have dropped constants, and where

$$\Phi = M\,\ell n\,\Sigma_1^M \gamma_a^2/\lambda_a \tag{19}$$

In some cases it may be simpler to first minimize with respect to the γ_a, using the formulas given in Sections A and B, and then vary μ to find the minimum of X^2 as given by Eq. (16).

2. Generalized Analysis of Variance [5]

The method C.1 has two disadvantages: The results depend on the value chosen for M, and the value of Φ weights the coefficients γ_n for small n in the same way as for large n. We shall now introduce a method which explicitly compares the ratios of coefficients for large n to those for smaller n. Let

$$a_n = \gamma_n(\lambda_n)^{-\frac{1}{2}}$$

$$s_n = \frac{1}{n}\Sigma_1^n a_n^2 \tag{20}$$

$$r_n = a_{n+1}^2/s_n$$

We transform

$$P(a^2) \prod_1^M da_n^2 = \left(\frac{\mu}{2\pi}\right)^{M/2} \prod_1^M \frac{da_n^2}{a_n} \; e^{-\frac{\mu}{2}\sum_1^M a_n^2} \tag{21}$$

into the following distribution:

$$P(a^2) \prod_1^M da_n^2 = P(a_1^2, r_1, \ldots, r_{N-1}) da_1^2 \prod_1^{M-1} dr_n$$

$$= \left(\frac{\mu}{2\pi}\right)^{M/2} a_1^{M-2} \prod_1^{M-1} \frac{dr_n}{r_n^{\frac{1}{2}}} \left(\frac{n+r_n}{n+1}\right)^{\frac{M-n-1}{2}} da_1^2 \tag{22}$$

$$\times \; e^{-\frac{\mu a_1^2}{2} \prod_1^{M-1} \frac{n+r_n}{n}}$$

Integrating over da_1^2 we obtain

$$P(r_1 \ldots r_{n-1}) = \prod_1^{M-1} Q_n(r_n) \tag{23}$$

where

$$Q_n(r) = \frac{n^{\frac{1}{2}n} \, \Gamma(\frac{1}{2}n + \frac{1}{2})}{\Gamma(\frac{1}{2})\Gamma(\frac{1}{2}n) \; r^{\frac{1}{2}}(n+r)^{\frac{1}{2}n+\frac{1}{2}}} \tag{24}$$

Thus the statistical distributions of the r_n are independent of each other, as well as of the value of μ.

3. Convergence Test Function [5]

We wish to devise a function which measures, collectively, whether the ratios r_n are larger than one would expect from a Gaussian distribution. Furthermore, we wish to weight somewhat more heavily the r_n for large n. To do this we first

construct the monotone function $\rho_n(r)$, where ρ_n has a χ^2-distribution with one degree of freedom. This function satisfies the following differential equation:

$$Q_n(r)dr = \frac{e^{-\frac{1}{2}\rho_n}}{\sqrt{2\Gamma(\frac{1}{2})}} \frac{d\rho_n}{\sqrt{\rho_n}} \tag{25}$$

For small r, we have:

$$\rho_n/r = 1 + 0(\frac{1}{n}) + 0(\frac{r}{n}) \tag{26}$$

and for large r:

$$\rho = n\ \ell n\ r + 0(\ell n\ \ell n\ r) \tag{27}$$

An approximate formula is thus

$$\rho_n(r) \sim n\ \ell n\ (1 + \frac{r}{n}) \tag{28}$$

(While this formula is good for large n, it may be wrong by as much as 50% for n = 1.) Finally, we define the convergence test function (CTF)

$$\Phi = \sum_1^{M-1} \rho_n \tag{29}$$

which has a χ^2 distribution with M-1 degrees of freedom. This can be combined with the classical χ^2 as in Eq. (18)

$$X^2 = \chi^2 + \Phi \tag{30}$$

where now X^2 has a χ^2 distribution with N-1 degrees of freedom, independent of M.

The logarithmic nature of $\rho_n(r)$ guarantees that if certain γ_n are really required to have large values to fit the data, in the minimization of (30) the effect of χ^2 overwhelms the effect of Φ. On the otherhand, if the values of the γ_n for n greater than some K are all around the noise level ($|\gamma_n| \sim 1$), the minimization of X^2 guarantees

that these γ_n are reduced to extremely small values. The effective cut-off in n occurs at such a value K; it is gradual, although in practice it turns out to be slightly sharper than the cut-off obtained through the use of Eq.(14).

In contrast to the use of method C.1, when the CTF is used the results are completely insensitive to the value chosen for M, provided only that M is somewhat larger than K. In either case, however, the ambiguity is much less than in the conventional method in which only χ^2 is considered.

D. Estimation of Truncation Uncertainty

In addition to the statistical uncertainty, there is always a truncation uncertainty to consider. This component of the uncertainty is usually negligible at the data points, but can become the main component of the uncertainty when one tries to extrapolate too far away from the data region. According to the formalism presented here, with an a priori Gaussian distribution of functions, these two components are combined quadratically.

The "statistical uncertainty" can be calculated by the usual rule: $\Delta X^2 = 1$. This actually is a combination of true statistical uncertainty arising from the terms $n \leq K$, and a part of the truncation uncertainty coming from the terms $K \leq n \leq M$. The residual truncation uncertainty can be calculated as follows. Let us suppose we wish to determine the truncation uncertainty in a real quantity f_o associated with the linear functional operator I_o. We can use the formula

$$\langle f_o^2 \rangle_{A/E(M)} = (H_{oo} - \sum_{\alpha=1}^{M} P_{o\alpha}^2/\lambda_\alpha)/\mu, \tag{31}$$

where $P_{o\alpha} = I_o P_\alpha$. The value of μ to be used in (31) can be estimated by the formula

$$\mu \approx \frac{1 + \Phi}{\sum_1^M a_n^2} \tag{32}$$

which is insensitive to M if $M > K$.

It should be emphasized that the methods presented here have been designed to reduce to a minimum the total uncertainty. However, since it has been customary to neglect the component arising from truncation, the uncertainties obtained by use of these methods may appear to be larger than values sometimes quoted.

REFERENCES

1) Walsh, J.L., Interpolation and Approximation by Rational Functions in the Complex Domain, American Mathematical Society Colloquium Publications, Vol. 20, Fifth Edition (American Mathematical Society, Providence, R. I., 1969)

2) Cutkosky, R. E. and B. B. Deo, Phys. Rev. 174, 1859 (1968)

3) Ciulli, S., Nuovo Cimento 62A, 3DI (1969)

4) Cutkosky, R. E. and S. Ciulli, unpublished

5) Cutkosky, R. E., Annals of Physics 54, 350 (1969)

6) Walsh, J. L. (Reference 1) page 224

7) Davis, P. J. Interpolation and Approximation (Blaisdell, New York, 1963) p. 316

8) Presnajder, P. and J. Pisut, Nuovo Cimento 3A, 603 (1971)

9) Chao, Y. A. and E. Pietarinen, Phys. Rev. Letters 26 1060 (1971)

ANALYTIC EXTRAPOLATION OF SCATTERING AMPLITUDES AND FORM FACTORS †

H. Pfister
University of Tubingen
Tubingen, Germany

1. General Considerations on Extrapolations

As long as there exists no complete elementary particle theory, the extrapolation of physical amplitudes from experimentally known regions into regions which are not yet accessible or cannot be reached in principle, is one of the main tasks of elementary particle physics, comparable to the task of gathering new data. Many of the questions which are in thorough discussion nowadays, are finally questions of extrapolation of amplitudes. As there are: What is a resonance? A pole in the second sheet or only a loop in an Argand diagram? Can all amplitudes be constructed out of resonances? Should an amplitude in some kinematical region be parametrized by resonances or by Regge-terms or by something else?

In spite of this central role of extrapolation of amplitudes, the mainly mathematical problems connected with it are by no means fully solved, and worked out in detail. In times when the experimental material was scarce, it was sufficient to represent for instance the proton form factor by a simple ρ-pole term, or a π-N partial wave by two or three Breit-Wigner resonances. To make, however, full use of the assembling of more and more data of higher and higher precision, it is necessary to develop systematic criterions, which parametrizations of these data are allowed and which not, and even more important, what are the uncertainties in conclusions drawn from those parametrizations.

† Supported in part by Deutsche Forschungsgemeinschaft and National Science Foundation.

The main tool which enables one to get reliable conclusions on the extrapolation of physical amplitudes, is the analyticity of these amplitudes in some domain D. These analyticity properties follow in some cases from axiomatic field theory, in less favourable cases they can still be deduced from perturbation theory. In any case, we will assume for all further discussion that the domain of analyticity of the amplitude in question is known. Unfortunately, this postulate alone does not suffice to solve the extrapolation problem for any real situation in elementary particle physics.

If our amplitude f, analytic in D, would be known in some, however small, continuum inside D, the theorems of complex analysis would guarantee the exact knowledge of f in the whole interior of D. But in practice, f is known experimentally only in discrete points z_i, and in these points only with finite errors ε_i. So, the extrapolation problem is unstable, or, in the mathematicians language, improperly posed: It is easily possible to construct for arbitrary $z_o \in D$ and arbitrary complex number K a function f, analytic in D, which coincides in the points $z_i \neq z_o$ with the experimental result, and fulfills $f(z_o)=K$. So, the extrapolation to any point different from z_i is, strictly spoken, undetermined.

In order to avoid this (unphysical) disaster, one usually introduces two additional postulates on the physical amplitude f:

a) One assumes that f is known along some continuous piece of a curve C_e within some (constant or variable) error:

$$|f(z) - f_e(z)| < \varepsilon(z) \qquad \text{on } C_e \subset D.$$

This smearing of the experimental information is quite plausible and acceptable in cases where the experimental points z_i lie already quite dense and where the relative variation of $f(z_i)$ from one point to the next is small. In many cases, the "point" z_i is also experimentally fixed only within some inaccuracy. In principle, postulate a) can be omitted, as is seen below, but many conclusions and formulas are more evident if we retain it.

b) One assumes that f is bounded on the boundary C_D of D. It is then reasonable that also the class of functions $f_e(z)$, admitted for approximating f, is restricted by this bound, so that we have:

$$|f(z) - f_e(z)| < M \qquad \text{on } C_D.$$

This second postulate is a quite severe limitation from the physical point of view, and by no means plausible. I think, nobody is able to give for any definite physical example a reliable-and not too large-number M, which fulfills the above relation on the whole curve C_D, including in most cases also the point infinity (in energy or momentum transfer), where we cannot even exclude an essential singularity of our amplitude f. Unfortunately, postulate b) cannot be skipped without running in the above mentioned mathematical disaster, but we must, and we can arrange our extrapolation in a way such that the special form of b) and the special constant M influence our results as slightly as possible.

Let us see in short that the disaster of indefiniteness of the extrapolation is really removed by postulates a) and b). This is guaranteed by the Nevanlinna principle: Be D a simply connected domain, and C_e a piece of a curve inside it. If $|f(z)-f_e(z)|<\varepsilon$ on C_e, and $|f(z)-f_e(z)|<M$ on C_D, the boundary of D, then $|f(z)-f_e(z)|<\varepsilon^{1-\omega(z)} \cdot M^{\omega(z)}$ for all $z\in D$, with $\omega(z)$ being the harmonic measure of D: $\Delta\omega(z)=0$; $\Delta\omega(z)=0$; $\omega(z)=1$ on C_D, $\omega(z)=0$ on C_e.
It follows then that the extrapolation is stable inside D: $\lim_{\varepsilon\to 0} |f(z)-f_e(z)|=0$. The stability on C_D can be established by further assumptions on f; for instance, f fulfills some Holder condition on C_D: $|f^{(p)}(z)-f^{(p)}(z_o)|\leq C|z-z_o|^{\alpha}$, or, f has only square root singularities on C_D, or, in a still more physical manner, by calculating only mean values of f over finite arcs on C_D.

2. Historical Remarks on Analytic Extrapolation

The first attempt to use analyticity for a reliable extrapolation of a scattering amplitude seems to have been done by Ciulli and Fischer [1]. They have made a conformal transformation of the p-p scattering amplitude f from the $z=\cos\theta$-plane to some other plane $w=(1-z^2)/(\alpha^2-z^2)$, and have shown that an expansion of f in powers of w shows much

better convergence than an expansion in z. Similar methods of conformal transformation and power series expansion were used by Frazer [2], Lovelace [3], Atkinson [4], Levinger and Peierls [5] and others for extrapolation of p-p, π-p, and π-π scattering amplitudes, and nucleon form factors. But, as stressed by Bertero and Viano [6], these people did not care of the stability problem, and therefore a reliable statement about the extrapolation error could not be made.

3. So Called Optimal Extrapolations

In 1968, independently from each other, Cutkosky, Deo [7] and Ciulli [8] worked out methods of stable extrapolation for physical amplitudes by introducing postulates of the type a) and b). By further idealizing the physical situation (assuming that ϵ is very small and M not too big), Cutkosky and Ciulli could refer to former mathematical work on interpolation and approximation [9], and show that the special conformal transformation w_0, which brings the whole domain D into an ellipse, and the curve C_e of experimental points to the focal line of this ellipse, is optimal in the sense that the series $f=\sum_n a_n {w_0}^n$, with a_n resulting from a X^2 approximation to the experimental data on C_e, converges faster than any other power series approximation of f, and that this expansion also shows the best possible stability, even on the boundary C_D [10]. (With respect to this result, the partial wave expansion is very bad because it converges only in a small part of D. On the other hand, the Cutkosky-Ciulli expansion is not unitary term by term, and it has to be corrected for that.)

The approximation could be further optimized, if also parts of higher Riemann sheets were transformed into the ellipse, as is done in a way in [11], but in general there is not much improvement because the singularities in higher sheets are not well known.

In application to realistic problems, the method of Cutkosky and Ciulli, however, has some shortcomings, as pointed out in [11]: Even for the best known amplitudes (N-N and π-N scattering, nucleon form factors), the error ϵ is so big that only few terms of the series $\sum_n a_n {w_0}^n$ can be calculated with reliability, and then the above mentioned

mathematical theorems, which are of an asymptotic character, are empty. Furthermore, the error of extrapolation, calculated by Cutkosky and Ciulli by some generalized Nevanlinna principle, and depending on ϵ and M, is in realistic cases unacceptably high. In an application to the proton form factor, it has been shown explicitely in [12] that an expansion in powers of the elliptical function w_o is not necessarily superior to other power series expansions, which converge only in a part of D. I admit, there are series expansions in variables w, which, in spite of having all correct analyticity properties, converge only much worse than the w_o-expansion, and give practically no useful information. On the other hand, there are expansions quite different from w_o, not even having the right analyticity structure, which are as good in some examples as w_o. And, to be sure, there exist physical amplitudes, which can only very badly be approximated by a sum of three or four elliptical functions.

As will be discussed below, the analytical extrapolation of physical amplitudes is to a large extent a problem of measure theory in a function space, and results derived with some special conformal transformation depend in a nearly uncontrollable way on this special transformation.

4. The Mean Value Property of Analytic Extrapolation

It has been pointed out first by Bowcock, Cottingham, and Williams [13] that the analytic extrapolation produces not values of f on definite points of C_D, but only mean values over finite arcs on C_D, provided, one does not set severe and physically unjustified limitations on f on the boundary. Mathematically, this mean value character of the extrapolation can be traced back to investigations on the inversion of dispersion relations by Paley and Weiner [14] (see also [15]). You may remember that also in the famous duality-paper by Dolen, Horn and Schmid [16], it is made quite clear that from the Regge asymptotics you get only averages over the resonances at finite energy. Bowcock et al. got their result within a critical discussion- under some simplifying assumptions- of Levingers extrapolations procedure for the proton form factor. It was then shown in [12] that this averaging function, connected with a series expansion of f in powers of a conformal variable w, can be constructed explicitly and uniquely.

Let us shortly present the method for an amplitude f(z) - like the form factor -, having only one cut. (The method can, however, be generalized to amplitudes with two cuts.) Be f(z) measured in K points $z_i<0$, with errors ϵ_i. Then we want to approximate f(z) in the whole plane, cut from 0 to ∞, by a series $\sum_{n=0}^{N} b_n (w(z))^n$, w(z) being some suitable conformal transformation, having the same analytic properties as f(z), and being real for $z<0$. For practical calculations more appropriate is a representation $f(z) \approx \bar{f}(z) = \sum_{n=0}^{N} a_n Q_n(w)$, where the $Q_n(w)$ are the orthonormal polynomials with respect to the measure of experimental information:

$$\sum_{i=1}^{K} \frac{Q_n(w_i) Q_m(w_i)}{\epsilon_i^2} = \delta_{n,m} \qquad (n,m = 0, 1, \ldots N).$$

For finite N both series are equivalent. The coefficients a_n are, as usual, calculated by minimizing

$$\chi^2 = \sum_{i=1}^{K} \frac{[f(z_i) - \sum_{n=0}^{N} a_n Q_n(w_i)]^2}{\epsilon_i^2} .$$

The result is: $$a_n = \sum_{i=i}^{K} \frac{f(z_i) Q_n(w_i)}{\epsilon_i^2} .$$

Here is now the place to insert explicitely the analytic properties of f(z) in form of the (unsubtracted) dispersion relation

$$f(z_i) = \pi^{-1} \int_0^{\infty} \frac{\mathrm{Im}\, f(z')}{z' - z_i} \, dz',$$

Im f(z') being the spectral function of the real physical amplitude. (A subtracted dispersion relation would work equally well.) In this way, we get an integral representation, expressing the approximate function through the complete function:

$$\bar{f}(z) = \int_0^\infty dz' \sum_{n=0}^{N} \sum_{i=1}^{K} \pi^{-1} \frac{Q_n(w(z))Q_n(w(z_i))\ \mathrm{Im}\ f(z')}{(z' - z_i)\ \epsilon_i^2} .$$

Taking the imaginary part of both sides, we end up with:

$$\mathrm{Im}\ \bar{f}(z) = \int_0^\infty dz'\ S_I(z', z)\ \mathrm{Im}\ f(z'),$$

where the averaging function is

$$S_I(z',z) = \pi^{-1} \sum_{n=0}^{N} \sum_{i=1}^{K} \frac{\mathrm{Im}\ Q_n(w(z))\ Q_n(w(z_i))}{(z' - z_i)\ \epsilon_i^2}$$

For $K \to \infty$, $\epsilon_i \to 0$, we can go with $N \to \infty$, and then S_I should approach $\delta(z'-z)$.
The same can be done for the real part of f(z), using a different representation, expressing the analytic properties of f, the so called airfoil equation [17]

$$f(z_i) = \frac{\sqrt{-z_i}}{\pi} \int dz' \frac{\mathrm{Re}\ f(z')}{\sqrt{z'}\ (z'-z_i)}$$

The resulting averaging function $S_R(z',z)$ is in many applications very similar to $S_I(z',z)$.
It can be shown that the averaging functions S are uniquely fixed, once the conformal variable w is given, and that S does not depend on the special way, the analytic properties of f are worked in. This is due essentially to the fact that S does not depend on the experimental values $f(z_i)$ but only on the errors ϵ_i.

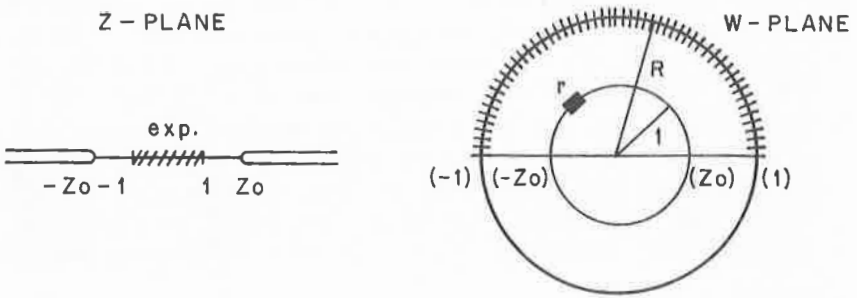
Some remarks should be made with respect to the degree N of approximation, one should choose. It turns out, that χ^2 as a function of N at first falls very rapidly (typically from something like 10^8 at N=0 to 10^2 at N=3), and then at some value N_0 levels off into a value near K, the number of experimental points. Pushing the approximation beyond N_0 would mean that you want to fit also the statistical fluctuations-the noise-in the experimental data. With N=K, you could even construct a polynomial $\bar{f}(z)$ which goes

through the center of all error bars; but surely this does not make any sense. The best (most physical) value N_o can also be found in another way: The separate terms $a_n Q_n(w)$ of the series $\bar{f}(z)$, taken in the extrapolation region ($z>0$ in our example) first fall off (in absolute value), but beyond $n = N_o$ they rise enormously, because fitting also the noise of the data develops fantastic oscillations of $\bar{f}$ in the extrapolation region. $\sum_{n=o} a_n Q_n(w)$ is in fact a semi-convergent series.

Let us say some more words about the averaging functions S (now taken with the optimal value N_o) and the degree of information, they give us about the true physical amplitude f in the extrapolation region $z>0$: By explicit calculation, it can be seen that $S(z',z)$ depends very strongly on z, the special point in the extrapolation region, around which you want to gain information on f. More explicitly: The region $z>0$, in which you can get reliable information, is in a way the mirror-image of the region of good experimental information, taken with respect to the threshold-point $z = 0$, as has been pointed out already some years ago [15, 18]. Near the threshold point and for very large values of z, the information, you can get, is in most cases very scarce. This correspondence between extrapolation region and experimental region is hidden in procedures as Ciulli's, which stick strongly on a unique bound of f for all positive z.

Concerning the error, to which our approximate amplitude $\bar{f}$ is subjected, it is reasonable, to take the absolute value of the last (the minimal) term in the series, $|a_{N_o} Q_{N_o}(w)|$. Once you have accepted, to approximate f by a power series in the special variable w, this can also be proved by well-known χ^2-approximation theory, i.e. by calculating the uncertainties of the coefficients a_n. But, to be sure, there is no strict mathematical proof of this estimate, independent of w. On the other hand, it turns out in practical applications, that approximations $\bar{f}$, expanded in different (but reasonable) variables w, do not differ from each other by more than the above estimate. In any case, it is essential, that by calculating only mean values of f, we get a resonable error estimate, which is independent of some (unphysical) bound on f.

Once it is clear that by analytic extrapolation one can never get f(z) on definite points, but only average values, it is nearby to give up completely the approximation through power series' in some conformal variable w, and to put the averaging function itself on the top of the extrapolation procedure. This point of view has been taken for example by Pisut, Presnajder and Fischer [19]. For convenience only, they first transform the physical z-plane with its two cuts into a ring between two circles.:



Idealizing the situation a little bit, they split the amplitude f into the data and the errors:

$$f(w) = f_e(w) + \varepsilon(w).$$

(Of course, a function $\varepsilon(w)$ does not really exist.) Then, taking a so called focusation function $g(w,w')$, analytic in w withing the ring, real on the unit circle $|w|=1$, and large only along a small arc Γ of it, around $w=w'$, the Cauchy theorem gives:

$$\int_R f_e(w)\ g(w,w')\ dw + \int_R \varepsilon(w)\ g(w,w')\ dw = \int_\Gamma f(w)\ g(w,w')\ dw +$$

$$+\int_{1-\Gamma} f(w)\ g(w,w')\ dw.$$

or

$$\int_\Gamma f(w)\ g(w,w')\ dw =$$

$$= \int_R f_e(w)\ g(w,w')\ dw + \int_R \varepsilon(w)\ g(w,w')\ dw - \int_{1-\Gamma} f(w)\ g(w,w')\ dw$$

The average of f over the arc Γ is given by an integral, which can be calculated from the data, plus two terms, expressing the uncertainty of the extrapolation. These two terms in a way compete with each other: $\int_{1-\Gamma} f(w)\, g(w,w')\, dw$ gets negligible, the more $g(w,w')$ is focussed around w'; but then $g(w,w')$ is large on the R-circle, and the integral over $\epsilon(w)$ grows. In practice, one takes a special class of focusation functions $g(w,w')$, say a kind of Gaussians on the unit circle, and chooses the width in an optimal way, so that the sum of both integrals gets minimal. This optimal choice $g_o(w,w')$ between two competing terms is of course in analogy to the dependence of $|a_n Q_n(w)|$ on n, showing up in the series expansion method. Again you can see, that, due to calculating only mean values, the special assumptions on the amplitude f, for instance to be bounded by M, do not influence the error estimate so strongly, because only the integral $\int_{1-\Gamma} f(w)\, g(w,w')\, dw$ enters into the error estimate. In a very elegant way, Pisut et al. have traced back this increase in stability of the extrapolation, this slighter dependence on assumptions on f, to the fact that the averaged function $F(w') = \int_{|w|=1} f(w)\, g(w,w')\, dw$ has a much larger region of analyticity than the original function $f(w)$.

Comparing the method of Pisut et al. with the power series expansion, it appears, that the former method is more general, in that the focusation functions $g(w,w')$ can be chosen arbitrary (within the analyticity constraints), while the functions $S_I(z',z)$ depend on z' always in the form $\sum_{i=1}^{K} \frac{s_i(z,z_i)}{z' - z_i}$. On the other hand, the second method makes very drastically clear that the averaging (or focusation) functions cannot really be chosen free, but are to a far extent determined by the experimental error-distribution at disposal.

5. The Statistical Nature of Analytic Extrapolation

A shortcoming, common to all extrapolation procedures, discussed till now, is that their results depend to some extent on the special expansion variable w or on the special

type of focusation functions used. In this situation, it has been emphasized by Cutkosky in a fundamental paper [20] that the extrapolation of physical amplitudes in fact is a problem of statistical nature. There is also a subsequent paper along the same reasoning by Presnajder and Pisut [21].

Cutkosky suggests to manage the instability problem not by the unflexible condition $|f(z)|<M$, but by introducing a probability measure for the values of f on the boundary. I, personally, being very sceptical concerning predictions, how an amplitude should probably look like in a hitherto unknown region, would like to suggest another kind of statistical concept in analytic extrapolations: Do the extrapolation simultaneously with all physically and mathematically admissible expansion variables w, or focusation functions g(w,w'), and then average over all results with a suitable measure function, measuring in a way the "probability" of the special parametrization through the width of its smearing function g(w,w') or S(z',z) and the error estimate $|a_{N_o} Q_{N_o}(w)|$ respectively

$$\left| \int_{1-\Gamma} f(w)\, g_o(w,w')\, dw \right| + \int_R |\varepsilon(w)|\, |g_o(w,w')|\, |dw|$$

But I will not go any deeper into this statistical game, because I have not yet done explicit calculations in this direction, and because Prof. Cutkosky surely can tell you much more definite results about that.

6. Survey of Physical Applications

There are a variety of nice applications of the analytic extrapolation techniques by the Cutkosky-group on K-N and N-N scattering. One can for instance get quite reliable values for the N-N and K-N coupling constants [7], and one can represent the corresponding scattering amplitudes even better with less parameters than in the partial-wave method, and can therefore correlate different phase shifts and resolve some ambiguities inherent in these [22]. Just recently there appeared an interesting letter by Deo and Parida [23], in which they show that the high-energy p-p data are better and more uniformly represented by an optimal conformal variable in the sense of Cutkosky and Ciulli than by the Orear- and Krisch-fits.

There are also some applications by other groups: Bowcock and John [24] have taken the π-π p-wave as measured with colliding e^+e^- beams and extracted out of it, using crossing and analyticity, an approximation for the π-π s-wave, showing some broad resonance (σ-meson) near the ρ-mass, but giving no good resolution.

A Nordita-group [25] then has extracted a lot of information out of the πN phase shifts, extrapolating these to the channel $\pi\pi \leftrightarrow N\bar{N}$. For the amplitude T(I=J=0) they get indeed, as above, indications for a broad σ-maximum, but the corresponding phase only very unlikely goes up to 90° as it should be for a real resonance. For T(I=J=1), the imaginary part of the amplitude shows the ρ-peak, but smeared out over something like 4 times the experimental ρ-width. The real and the imaginary part turn out to be quite small a distance above the ρ-mass.

This fact harmonizes excellently with a result, extracted with analytic extrapolation from a quite different amplitude, the proton electromagnetic form factor [11]: By extrapolating the rather extensive and accurate data on the proton form factor G(t) in the spacelike region to the timelike region, it turned out that Im G(t) has no pronounced ρ-maximum at the right place, but a zero a small distance above the ρ-mass. At first, this is surprising, because, taking into account only the 2π intermediate state as usual, we have Im $G(t) \sim F_\pi(t)^* \cdot T(I=J=1, t)$, and it is experimentally absolutely clear that the pion form factor $F_\pi(t)$ has a pronounced ρ-maximum. But now, by our extrapolation procedure, we get only an average of the above expression over a finite t-region (typically of the order of 400 - 500 $(MeV)^2$, i.e. 4 times the ρ-width; the uncertainty in $(\text{Im } G(t))_{av}$ being of the order of 40% compared to 1% -3% error in the experimental data), and therefore the peak in $F_\pi(t)$ can be killed in a way, or be pushed to much smaller t-values by the "zero" in T.

Another interesting result of the proton form factor extrapolation refers to the so called scaling law between the electric and magnetic proton form factor: $G_E(t) = G_M(t)/\mu$. It turned out [12] that the small but systematic deviations from this law in the spacelike region

($G_E > G_M/\mu$ for $0.15\ (\mathrm{GeV})^2 < -t < 0.30\ (\mathrm{GeV})^2$, $G_E < G_M/\mu$ for $0.60\ (\mathrm{GeV})^2 < -t < 1.5\ (\mathrm{GeV})^2$), which have been confirmed by the last experiment [26], produce quite large deviations from this law in the timelike region, and fortunately in the direction as to confirm the relation $G_E = G_M$ at $t = 4M_p^2$, as it is suggested by theoretical arguments [27].

REFERENCES

1) Ciulli, S. and J. Fischer, Nuclear Physics 24, 465 (1961).

2) Frazer, W. R., Phys. Rev. 123, 2180 (1961).

3) Lovelace, C., Nuovo Cim. 25, 730 (1962).

4) Atkinson, D., Phys. Rev. 128, 1908 (1962).

5) Levinger, J. S. and R. F. Peierls, Phys. Rev. 134, B1341 (1964).

6) Bertero, M. and G. A. Viano, Nuovo Cim. 38, 1915 (1965).

7) Cutkosky, R. E. and B. B. Deo, Phys. Rev. Letters 20, 1272 (1968).
Cutkosky, R. E. and B. B. Deo, Phys. Rev. 174, 1859 (1968).

8) Ciulli, S. Nuovo Cim. 61A, 787 (1969), 62A, 301 (1969).

9) Walsh, J. L., Interpolation and Approximation by Ra tional Functions in the Complex Domain, American Math. Soc. Coll. Publ. 20 (1935).
Szego, G., Orthogonal Polynomials, American Math. Soc. Coll. Publ. 23 (1939).

10) Ciulli, S. and R. E. Cutkosky, Nordita-preprint (1969).

11) Pfister, H., Nuclear Phys. B20, 320 (1970).

12) Pfister, H., Nuclear Phys. B22, 327 (1970).

13) Bowcock, J. E., W. N. Cottingham and J. G. Williams, Nuclear Phys. B3, 95 (1967).

14) Paley, R. E. A. C. and N. Wiener, American Math. Soc. Coll. Publ. 19 (1934).

15) Pfister, H., Zeitschr. Phys. 211, 176 (1968).

16) Dolen, R., Horn, D. and C. Schmid, Phys. Rev. 166, 1768 (1968).

17) Tricomi, F. G., Integral Equations, Intersci. Publ. (New York 1957).

18) Pfister, H., Paper presented to the Lund Conference (1969).

19) Pisut, J., Presnajder, P. and J. Fischer, Nuclear Phys. B12 586 (1969).

20) Cutkosky, R. E., Ann. Phys. 54, 350 (1969).

21) Presnajder, P. and J. Pisut, Nuovo Cim. 3A, 603 (1971).

22) Ray, A. et al., Phys. Rev. 183, 1183 (1969).
Cutkosky, R. E. and B. B. Deo, Phys. Rev. D1, 2547 (1970).

Chao, Y. A., Phys. Rev. Letters, 25, 309 (1970).

23) Deo, B. B. and M. K. Parida, Phys. Rev. Letters, 26, 1609 (1971).

24) Bowcock, J. E. and G. John, Nuclear Phys. B11, 659 (1969).

25) Nielson, H., J. L. Petersen and E. Pietarinen, Nuclear Phys. B22, 525 (1970).

26) Berger, C et al., Phys. Letters. 35B, 87 (1971).

27) Bergia, S. and L. Brown, Proc. Int. Conf. Nucl. Structure (Stanford 1963), p. 320.
Anisovich, V. V. et al., JETP Letters 5, 277 (1967).