Proof of the Mandelstam Representation for Every Order in Perturbation Theory*

RICHARD J. EDEN
Clare College, Cambridge University, Cambridge, England
(Received October 27, 1960)

It is proved that every term in the perturbation series for a scattering amplitude satisfies the Mandelstam representation when there are no anomalous thresholds. The absence of anomalous thresholds can be investigated from a few low-order diagrams, or reduced diagrams. Under certain conditions it is shown that their absence in fourth order ensures their absence in every order.

1. INTRODUCTION

(a) Statement of the Problem

IN this paper I will describe a proof of the Mandelstam representation for a scattering amplitude for every order in perturbation theory. The proof applies to any collision process that does not have anomalous thresholds. It is shown that under certain conditions the absence of anomalous thresholds in fourth order ensures their absence to all orders, and more generally that only certain lower order diagrams need be considered. For simplicity the main account of the proof will be described for equal-mass scalar particles, and special features for more general particles will be discussed afterwards.

The starting point of the proof is the scattering amplitude for a general Feynman diagram. This is a function of any two of the three invariant energies squared, s, t, and u. It is defined for real values of these variables in physical scattering regions by an equation of the form

$$F(s,t)=\lim_{\epsilon\to 0} c_1\int dk_1\cdots\int dk_l \frac{1}{\prod_{j=1}^{n}(q_j^2-m^2+i\epsilon)}, \quad (1.1)$$

where the internal four-momenta are denoted $k_1\cdots k_l$. The four-momentum q_j for any internal line of the diagram will be a linear function of the k_i and the external four-momenta p_a. The small imaginary part $-i\epsilon$ associated with each mass m^2 ensures that causality is properly included in the scattering process. The definition (1.1) is valid for each of the three physical scattering regions, which are denoted I, II, and III in Fig. 1. A change of variables can be effected by means of the relation

$$s+t+u=4m^2. \quad (1.2)$$

The problem involves two objectives. The first is to construct a function $F(z_1,z_2)$ of two complex variables z_1, z_2, that tends to the Feynman amplitude $F(s,t)$ when a suitable limit is taken. The same function must also give the amplitude in the other physical scattering regions when we use a relation similar to Eq. (1.2) in three complex variables,

$$z_1+z_2+z_3=4m^2, \quad (1.3)$$

and again take a suitable limit. The second objective is to show that the function $F(z_1,z_2)$ satisfies the double dispersion relation proposed by Mandelstam.[1]

(b) Outline of the Proof

The amplitude given by Eq. (1.1) can be written[2]

$$F(s,t)=\lim_{\epsilon\to 0} c_1\int_0^1 d\alpha_1\cdots d\alpha_n\int dk_1\cdots dk_l \times\frac{\delta(1-\sum\alpha_i)}{[\sum\alpha_j(q_j^2-m^2+i\epsilon)]^n} \quad (1.4)$$

$$=\lim_{\epsilon\to 0} c_2\int_0^1 d\alpha_1\cdots d\alpha_n \frac{n(\alpha)}{[D_\epsilon(\alpha,s,t)]^p}, \quad (1.5)$$

where p is a positive integer and where

$$D_\epsilon(\alpha,s,t)=sf(\alpha)+tg(\alpha)-m^2K(\alpha)+i\epsilon\sum_1^n\alpha_iC(\alpha). \quad (1.6)$$

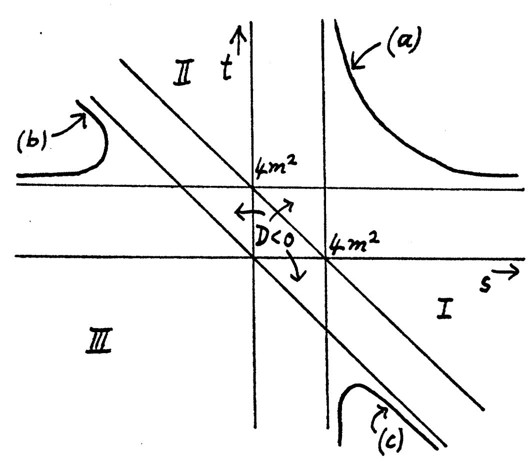

FIG. 1. The real s, t plane with equal masses. Physical scattering regions are denoted I, II, and III. The relation $D<0$ is first proved for the inner triangle and then extended to the larger triangle.

* This work was completed except for minor revisions while the author was a visiting physicist at the Lawrence Radiation Laboratory, Berkeley. It was reported at the Rochester Conference on High-Energy Physics in August, 1960.

[1] S. Mandelstam, Phys. Rev. **112**, 1344 (1958); **115**, 1741 (1959); **115**, 1752 (1959).
[2] R. J. Eden, Phys. Rev. **119**, 1763 (1960).

The value of D_ϵ at $\epsilon=0$ will be denoted $D(\alpha,s,t)$. When s and t are real and have physical values, D_ϵ is nonzero for real values of α in the range of integration, since by inspection $C(\alpha)$ is positive when the α variables are positive.[2] Then the integral is well defined, but it may become singular in the limit as ϵ tends to zero, if the integrand has either coincident or end point singularities in each integration variable $\alpha_1\cdots\alpha_n$ within the (real) range of integration.[2–6]

The first step in the proof is to show that the function $F(z_1,t)$ defined by Eq. (1.8) satisfies a dispersion relation in z_1 when t is real and in the range

$$-4m^2<t<4m^2. \tag{1.7}$$

$$F(z_1,t)=c_2\int_0^1 d\alpha_1\cdots d\alpha_n \frac{n(\alpha)}{[D(\alpha,z_1,t)]^p}. \tag{1.8}$$

It is shown in Sec. 2 that $F(z_1,t)$ satisfies a dispersion relation and that when

$$-4m^2<t<0,\quad s>4m^2-t, \tag{1.9}$$

the function $F(z_1,t)$ tends to the Feynman amplitude $F(s,t)$ for process I as $z_1=s+i\epsilon$ tends to s. It is also noted that for $s<0$ the Feynman amplitude for process III is obtained as $z_1=s-i\epsilon$ tends to s.

In step two of the proof the analogous single variable dispersion relation for a function $F(s,z_2)$ is written down.[7] This contains an integrand

$$F(s,\,t+i\epsilon')-F(s,\,t-i\epsilon'), \tag{1.10}$$

where t may be in either of the ranges $t>4m^2$, or $t<-s$. In order to express this integrand by a dispersion relation in z_1, we must show that there exist analytic functions $F(z_1,\,t+i\epsilon')$ and $F(z_1,\,t-i\epsilon')$. Their difference must be shown to tend to the function (1.10) in a suitable limit and to satisfy a dispersion relation in z_1 in the limit as ϵ' tends to zero. These functions are defined by the formula (1.8) in which t is replaced by $t\pm i\epsilon'$. The proof that they have the required analytic properties forms the third step in the proof and is described in Sec. 3.

The basis of this third step is the method of analytic completion[8] by which a domain of analyticity in the z_1 plane for a particular value of z_2 can be extended through a tube in z_1, z_2 space. The starting point is the upper half z_1 plane, when $z_1=t$ is in the range (1.7), the boundary of the domain being a large semicircle. It is then shown that the semicircle can be displaced first to $z_2=t+i\epsilon'$ and then to larger values of t without meeting any singularities of $F(z_1,z_2)$. The same result is true for $z_2=t-i\epsilon'$ provided $\epsilon'<\epsilon$ where $z_1=s+i\epsilon$ along the edge of the semicircle. This ensures the required analyticity for $t>4m^2$. The region $t<-s$ is most simply handled by making a change of variable and using oblique axes.[7] We note in Sec. 4 that this completes the proof of the Mandelstam representation for equal-mass particles.

In Sec. 5 it is noted that a key point in the validity of the proof depends on the absence of anomalous thresholds. When the leading normal threshold bounds a physical region for the scattering process that is being considered, the fourth order diagram (or reduced diagram) determines whether anomalous thresholds are absent to all orders. This applies to pion-pion scattering and to pion-nucleon scattering for example. When there is a gap between the region where the amplitude is real and the physical region, it is necessary to consider a limited number of other lower order diagrams, in order to verify that there are no anomalous thresholds to any order. This has been done for nucleon-antinucleon scattering as an example and is discussed in Sec. 5. In Sec. 6 some concluding remarks are made which note the particular points where knowledge from perturbation theory is explicitly used. Two alternative procedures for studying the validity of the Mandelstam representation are also briefly noted.

2. SINGLE VARIABLE DISPERSION RELATIONS

In this section the analytic properties of $F(z_1,t)$ defined by Eq. (1.8) will be studied and it will be shown to satisfy a dispersion relation in z_1 when t is in the range (1.7). In an earlier paper[2] it has been shown that

$$D(\alpha,s,t)<0,\quad \text{for}\quad \alpha_i\geqslant 0,\quad \textstyle\sum\alpha_i=1, \tag{2.1}$$

when s and t are in the Euclidean region defined by Eq. (1.2), and

$$s\geqslant 0,\quad t\geqslant 0,\quad u\geqslant 0, \tag{2.2}$$

with the exception of the three points $s=4m^2$, $t=0$, $u=0$, etc. We will first show that the function $F(z_1,t)$ exists, and is analytic for z_1 in the upper half plane, when

$$0<t<4m^2. \tag{2.3}$$

Let $z_1=s_1+is_2$; then

$$D(\alpha,z_1,t)=(s_1+is_2)f(\alpha)+tg(\alpha)-m^2K(\alpha). \tag{2.4}$$

We will assume that the α variables are all real and then show that D is never zero, which justifies the assumption. Since f, g, and K are each real for real α, D cannot be zero unless f is zero. If $f(\alpha)=0$, then

$$D(\alpha,s,t)=tg(\alpha)-m^2K(\alpha). \tag{2.5}$$

The right-hand side of Eq. (2.5) is a value of $D(\alpha,s,t)$ at $s=0$ for a particular set of values of the α variables. But for t in the range given by Eq. (2.3) and with $s=0$, D is negative for all real positive values of the α vari-

[3] R. J. Eden, Proc. Roy. Soc. (London) **A210**, 388 (1952).
[4] L. D. Landau, Nuclear Phys. **13**, 181 (1959).
[5] J. Tarski, J. Math. Phys. **1**, 154 (1960).
[6] J. C. Polkinghorne and G. R. Screaton, Nuovo cimento **15**, 289 (1960).
[7] R. J. Eden, Phys. Rev. **120**, 1514 (1960).
[8] G. Källen and A. Wightman, Kgl. Danske Videnskab. Selskab, Mat.-fys. Skrifter **1**, No. 6 (1958); and S. Bochner and W. T. Martin, *Several Complex Variables* (Princeton University Press, Princeton, New Jersey, 1948).

ables, including those that give $f(\alpha)=0$. Hence the right-hand side of Eq. (2.5) is nonzero, and $D(\alpha,z_1,t)$ is nonzero for z_1 in the upper half plane. Also for

$$0<s<4m^2-t, \tag{2.6}$$

$D(\alpha,s,t)$ is nonzero and real. This proves that $F(z_1,t)$ defined by Eq. (1.8) exists, and that it is analytic in the upper half plane and in a region including part of the real axis given by Eq. (2.6). It therefore satisfies a dispersion relation.

The region in which D is negative can be enlarged to the region

$$s<4m^2,\quad t<4m^2,\quad u<4m^2. \tag{2.7}$$

Let us consider first the part of this region in which s and t are positive. If D was not negative in this region (for α positive), there would be a singularity of $F(z_1,t)$ at the smallest real value s of z_1 for which the change of sign occurs. By varying t, a straight line or a curve of singularities will be obtained. Any straight line of singularities would necessarily enter the Euclidean region (2.2), or one of the physical regions indicated I and II in Fig. 1. There are no such lines since there are no singularities of the function $F(z_1,t)$ in the Euclidean region, and the only straight lines of singularities entering these physical regions are the normal thresholds $s=(nm)^2$, or $t=(nm)^2$, $n=2, 3, \cdots$. There cannot be any curves of singularities between $t=0$ and $t=4m^2$. To prove this we assume continuity of curves of singularities of $F(z_1,t)$. The continuity assumption will be justified later. A continuous curve must have a turning point at which

$$dt/ds=0. \tag{2.8}$$

Otherwise the curve would enter either the Euclidean region or a physical region, neither of which is allowed. Now it has been shown[2] that

$$dt/ds=-f(\alpha)/g(\alpha), \tag{2.9}$$

in which the α variables are given the numerical values for coincident or end-point singularities that lead to the curve of singularities under discussion. These will be called their "critical" values. Since $g(\alpha)$ is bounded, $f(\alpha)$ must become zero at a turning point. But if $f(\alpha)$ is zero, for this particular set of values of the α variables D is given by Eq. (2.5). We have already seen that the right-hand side of Eq. (2.5) is nonzero when t is in the range (2.3), for all positive α. Hence D cannot be zero at a turning point, and the function F cannot be singular along such a curve. This argument excludes also the possibility that a curve has zero slope only asymptotically.

We deduce that there are no singularities in the region (2.7) by using the symmetry between s, t, and u. Hence D is negative throughout this region. We can now check that $F(z_1,t)$ does indeed define the physical branch of the Feynman amplitude in an appropriate limit. We have

$$D(\alpha,\, s+i\epsilon,\, t)=(s-4m^2+\eta)f(\alpha)+(4m^2-\eta)f(\alpha)+tg(\alpha)-m^2K(\alpha)+i\epsilon f(\alpha). \tag{2.10}$$

When $f(\alpha)$ is positive, this expression has the same form as D_ϵ given by Eq. (1.6) from which the Feynman amplitude is defined. When $f(\alpha)$ is zero or negative, and when $s>4m^2$, and $\epsilon=0$, we have

$$D(\alpha,s,t)=(s-4m^2+\eta)f+(4m^2-\eta)f+tg-m^2K \tag{2.11}$$

$$\leqslant(4m^2-\eta)f+tg-m^2K. \tag{2.12}$$

This expression is negative for η small and real, when t satisfies

$$-4m^2+\eta<t<4m^2. \tag{2.13}$$

Hence in the limit as ϵ tends to zero, either D is nonzero so that it equals D_ϵ in the same limit; or it has exactly the same form as D_ϵ and leads to the same distortion of the contours of α integration in this limit.

When s is negative, the point $z_1=s-i\epsilon$ corresponds to

$$z_3=u+i\epsilon. \tag{2.14}$$

This provides the correct limit giving the Feynman amplitude when u is the energy squared. The cut z_1 plane of the function $F(z_1,t)$ is indicated in Fig. 2 when $-4m^2<t<0$. This cut plane defines the physical sheet of the amplitude $F(z_1,t)$ in the variable z_1, and is valid for t in the range (1.7).

The analogous function $F(s,z_2)$ can similarly be shown to satisfy a dispersion relation in z_2 when $-4m^2<s<4m^2$. This is

$$F(s,z_2)=\frac{1}{2\pi i}\int_{4m^2}^{\infty}\frac{[F(s,\,t+i\epsilon')-F(s,\,t-i\epsilon')]dt}{t-z_2}+\frac{1}{2\pi i}\int_{-\infty}^{-s}\frac{[F(s,\,t+i\epsilon')-F(s,\,t-i\epsilon')]dt}{t-z_2}. \tag{2.15}$$

We see from this formula that in order to obtain a double dispersion relation we must define a function $F(z_1,\,t\pm i\epsilon')$ in the region $t>4m^2$. This will be done by

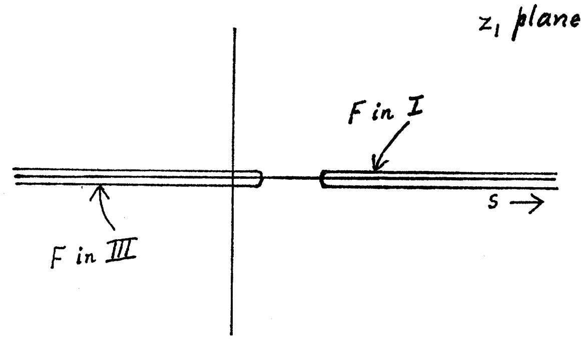

FIG. 2. The complex z_1 plane with $-4m^2<t<0$. The limits giving the amplitude for the physical regions I and III are indicated.

analytic continuation from the region $z_2=t<4m^2$ where a single-variable dispersion relation has been proved. The second integral on the right of Eq. (2.15) will be similarly considered after changing variables.[7] The formula (2.15) defines the physical sheet in the variable z_2. A similar formula can be used to define the physical sheet in the variable z_3.

3. ANALYTIC COMPLETION IN THE PHYSICAL SHEET

In Sec. 2 it has been shown that $F(z_1,t)$ defined by Eq. (1.8) has no singularities in the upper half z_1 plane for t in the range (1.7). Hence if C is a contour defined by a large semicircle in the upper half z_1 plane, we can write $F(z_1,t)$ in the form

$$F(z_1,t)=\frac{1}{2\pi i}\int_C \frac{F(z,t)dz}{z-z_1}. \tag{3.1}$$

In the four-dimensional space of the complex variables z_1, z_2 the contour C in Eq. (3.1) lies in a plane $z_2=$constant.

The method of analytic completion[8] consists of displacing the contour C by varying z_2 in Eq. (3.1). Provided the integrand does not become singular on the contour C, the formula (3.1) shows that $F(z_1,z_2)$ must be analytic everywhere inside the contour. It is therefore necessary only to prove that $F(z,z_2)$ is not singular when $z=s+i\epsilon$, and $z=Re^{i\theta}$, $(0<\theta<\pi)$, for arbitrarily large values of R. The initial value of z_2 is $t<4m^2$, and we will vary it to $t\pm i\epsilon'$, and then increase t past the threshold $4m^2$. Higher thresholds can be considered in a similar manner.

The function $F(z_1,z_2)$ is defined by

$$F(z_1,z_2)=c_2\int_0^1 d\alpha_1\cdots d\alpha_n\frac{n(\alpha)}{[D(\alpha,z_1,z_2)]^p}. \tag{3.2}$$

This definition has been shown in Sec. 2 to be valid with real integration paths for the α variables when z_1 and z_2 are near the real region (2.7). Its analytic continuation consists in varying z_1 and z_2 from this region and distorting the α paths of integration so that singularities of the integrand [zeros of $D(\alpha,z_1,z_2)$] do not cross them. For $z_2=t$, with $t\lesssim 4m^2$ we have also seen in Sec. 2 that $F(z_1,t)$ is analytic in the z_1 plane cut along the real axis as shown in Fig. 2. We now consider displacement of the contour C as indicated in Fig. 3.

The first displacement of C is to

$$z_2=t+i\epsilon'. \tag{3.3}$$

We will be concerned in particular with values of z_1, on C near the real axis,

$$z_1=s+i\epsilon. \tag{3.4}$$

Since ϵ and ϵ' can be arbitrarily small, it is only necessary to consider those surfaces of singularities in the z_1, z_2 space that intersect the real s, t plane. There, they will appear as curves of singularities or straight lines of singularities. This follows from the fact that the critical values of α are real when s and t are real and hence the derivative dz_2/dz_1 is real at the intersection with the real s,t plane. Accidental degeneracy might lead to a point singularity but this can be avoided by a slight variation in one or more masses. The reality of the critical α values for real s and t is most directly proved by the following argument given by Landshoff and Polkinghorne (private communication). The Landau-Bjorken equations $\partial D/\partial\alpha_i=0$ are real for real s and t. The self-consistency of these equations determines a set of curves or points for real s and t (not necessarily related to a physical branch). Given s and t at one of these points or on a curve, the corresponding critical values of the α variables will be unique (except for accidental degeneracy such as two curves crossing). Since the Landau-Bjorken equations are real the complex conjugate values of the α's also provide a critical set and by uniqueness these two sets are the same. Hence the critical values of the α variables are real, and the surfaces of singularities have real derivatives for s and t real.

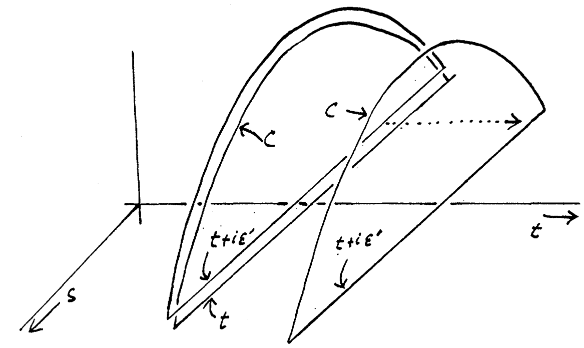

FIG. 3. The displacement of the contour C in the complex z_1, z_2 space. Only their real axes, s and t, are shown explicitly.

When t is just below the leading threshold $4m^2$, the only singularities of $F(z_1,t)$ are the normal thresholds,

$$z_1=s=(nm)^2,\quad n=2,3,\cdots, \tag{3.5}$$

$$z_3=u=(nm)^2,\quad n=2,3,\cdots. \tag{3.6}$$

The location of normal thresholds in z_1 or z_3 is independent of the other variables, and they are avoided when

$$z_1=s\pm i\epsilon, \tag{3.7}$$

$$z_3=u\pm i\epsilon''. \tag{3.8}$$

The contour C therefore avoids the singularities given by Eq. (3.5) during the displacement to the point (3.3). Using Eq. (1.2) and (1.3) we have

$$z_3=u-i(\epsilon+\epsilon'). \tag{3.9}$$

Hence the singularities given by the normal thresholds, Eq. (3.6) are not encountered by the contour.

The second displacement of C is obtained by increasing t in Eq. (3.3) past the normal threshold, $t=4m^2$. The singularity corresponding to the threshold itself is avoided by the small imaginary part $i\epsilon'$ in Eq. (3.3). We must also consider curves of singularities and the corresponding surfaces. It was shown in Sec. 2 that there are no curves of singularities just below this threshold. Hence any curve of singularities must either touch the line $t=4m^2$ or tend to it asymptotically. In both cases it must lie above this line except at the tangent point. Near the tangent point (which may be at infinity) the absolute value of the slope of the curve can be made less than any given positive number η, by a suitable restriction on t,

$$|dt/ds|<\eta, \quad \text{for} \quad 4m^2<t\leqslant 4m^2+\delta(\eta). \quad (3.10)$$

The slope of the curve of singularities is also the derivative of the hypersurface in z_1, z_2 space at any point on the curve. If the curve has an equation

$$t=t(s), \quad (3.11)$$

the hypersurface will have an equation

$$z_2=t(z_1). \quad (3.12)$$

Since this is an algebraic equation, we have

$$dz_2/dz_1=dt/ds, \quad (3.13)$$

at any point $z_1=s$, $z_2=t$, on the curve of singularities.

It follows from Eq. (3.13) that when the curve of singularities has a slope $\pm\eta$, a point $z_1=s+i\epsilon$, on the corresponding surface of singularities will have coordinates,

$$z_1=s+i\epsilon, \quad z_2=t\pm i\eta\epsilon. \quad (3.14)$$

Given ϵ and ϵ' in Eq. (3.3) and (3.4), we can always choose δ in Eq. (3.10) so that

$$\eta\epsilon<\epsilon'. \quad (3.15)$$

Hence the contour C can be displaced past the threshold up to the point

$$t=4m^2+\delta, \quad (3.16)$$

either for positive or negative slope of the curve of singularities. If the slope of the curve of singularities is negative, the contour C does not meet the surface of singularities for any value of η. If the slope is negative everywhere then the displacement of the contour C can continue to all positive values of t without meeting any singularities.

It will be our objective now to show that in fact with the analytic continuation defined by Eq. (3.3) and (3.4) the only curves of singularities do have negative slope. That is to say $F(z_1,z_2)$ becomes singular on curves of negative slope only, in the limit,

$$z_1=s+i\epsilon\rightarrow s, \quad z_2=t+i\epsilon'\rightarrow t. \quad (3.14)$$

We will show first that our continuing displacement of the contour C must not be prevented by a curve of

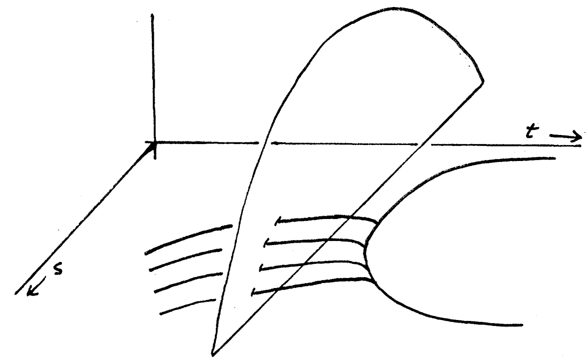

FIG. 4. The surface of singularities from a curve of positive slope and positive second derivative.

singularities of positive slope and positive second derivative. Below such a curve would be a concave region of the real s, t plane near which $F(z_1,z_2)$ is analytic. By considering the intersections of the corresponding surface with the plane

$$z_2=pz_1, \quad (3.15)$$

where p is real and positive, it is clear that near the real curve the surface will extend downwards into the interior of the contour C as indicated in Fig. 4. This would contradict our result that C can be displaced at least a small distance past the first normal threshold. Therefore such a curve of singularities cannot exist unless the displacement of C is prevented at some earlier stage. Since a curve of negative slope never extends into the region where z_1 and z_2 have imaginary parts of the same sign the only type of curve that could give trouble must have positive slope and negative second derivative. Thus it would have to bound a convex region of the real s, t plane near which $F(z_1,z_2)$ is analytic. We will prove that such a curve cannot exist by showing that curves of singularities cannot leave the physical sheet through a normal threshold at a finite point, and by using continuity of curves of singularities.

We will first consider the general possibility of a curve of singularities leaving the physical sheet. When $F(z_1,z_2)$ is continued along the path indicated by Eqs. (3.3) and (3.4) the paths of the α integrations become distorted but still have the end points 0 and 1. On a curve of singularities all the variables have either end-point or coincident (pinching) singularities. When the curve is followed, it can leave the physical sheet only at a point where one or more coincident singularities falls off the end of the contour of integration. At this point these coincident singularities are also end-point singularities. Therefore there is also a curve of singularities for a reduced diagram at the same point and having the same critical values of the α variables. (The "critical values" are those that give the coincident and the end-point singularities.) Since the slope of the two curves of singularities, given by Eq. (2.9), is a function of these critical values, the curves must touch where the original curve left the physical sheet. We can therefore continue along the curve for the reduced diagram which is on the physical sheet. The situation in the α integration is illustrated in Fig. 5. This establishes that every curve of singularities is associated with a continuous curve whose slope is continuous and along which $F(z_1,z_2)$ is singular in the limit given by Eq. (3.14), after analytic continuation in the physical sheet.

It may be possible for this continuous curve of singularities to terminate along a straight line of singularities. The only such lines are the normal thresholds. We will now show that a curve cannot leave the physical sheet through a normal threshold except asymptotically.

The discriminant $D(\alpha,s,t)$ has the form,

$$D(\alpha,s,t)=sf(\alpha)+tg(\alpha)-m^2K(\alpha). \tag{3.16}$$

We will assume that this corresponds to a suitably reduced diagram so that on its curve of singularities all the critical α values correspond to coincident singularities. We will also assume that this curve leaves the physical sheet at a point of tangency to a normal threshold, $t=4m^2$ for example. If this diagram can be further reduced to give another curve of singularities touching the normal threshold and leaving the physical sheet at the same point we fix our attention on the most fully reduced diagram that has this property. At the point of tangency, for the critical α values,

$$dt/ds=-f(\alpha)/g(\alpha)=0, \tag{3.17}$$

and since $g(\alpha)$ is bounded,

$$f(\alpha)=0. \tag{3.18}$$

Since the curve leaves the physical sheet at this point at least one parameter, α_1 say, must have its coincident singularities at the end point 0. If this was the only such

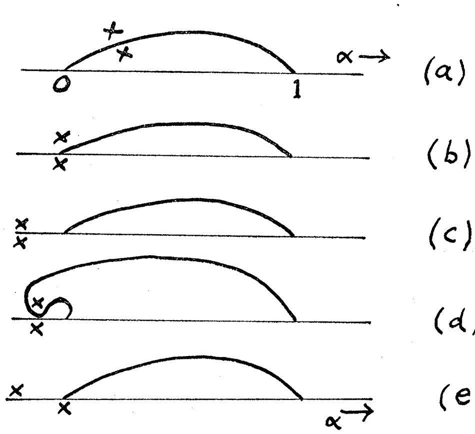

FIG. 5. An α integration; (a), (b), (c) show a coincident singularity following off the contour of integration as s, t vary along a curve of singularities through a point where it leaves the physical sheet. (c) keeps on the physical sheet but (d) follows the curve to an unphysical sheet. (e) shows the related singularity for a reduced diagram.

parameter there would be a reduced diagram formed from D, Eq. (3.16), by taking $\alpha_1=0$, that has a singularity on the physical sheet at the point of tangency. But we have assumed that D describes the most fully reduced diagram having a curve of singularities with this point as tangent and leaving the physical sheet. Therefore the reduced diagram having $\alpha_1=0$ must give a straight line of singularities. But the only lines of singularities are normal thresholds for which all the α variables are zero except those in generalized self energy parts.[2] It is not possible to go from a diagram giving a curve of singularities to a generalized self energy part by putting only $\alpha_1=0$. The simplest possibility is that also $\alpha_2=0$, that is to say another variable has its coincident singularities at an end point at the point of tangency. Further, both α_1 and α_2 must be factors of $f(\alpha)$ in this simplest case. Hence near the point of tangency to the normal threshold

$$\partial f(\alpha)/\partial\alpha_1\rightarrow 0. \tag{3.19}$$

We also have on the curve near this point,

$$\frac{\partial D}{\partial\alpha_1}=s\frac{\partial f}{\partial\alpha_1}+t\frac{\partial g}{\partial\alpha_1}-m^2\frac{\partial K}{\partial\alpha_1}=0, \tag{3.20}$$

and

$$t\frac{\partial g}{\partial\alpha_1}-m^2\frac{\partial K}{\partial\alpha_1}\neq 0. \tag{3.21}$$

If the expression in Eq. (3.21) were zero, then the reduced diagram with $\alpha_2=0$ would have a singularity for which α_1 gives a coincident singularity. Thus the line 1 would be on the mass shell although it is not one of the lines in the generalized self-energy part that gives the normal threshold. This is not possible, so the statement (3.21) is valid. Equation (3.19), (3.20) and (3.21) are consistent only if as the point of tangency is approached,

$$s\rightarrow\infty. \tag{3.23}$$

More generally, several α variables may become zero at the point of tangency to a normal threshold. The structure of diagrams giving rise to curves of singularities means that every term in $f(\alpha)$ must contain two factors α_1', α_2' from two sets α_1, α_1', α_1'', $\cdots$ and α_2, α_2', α_2'', $\cdots$ all of whose critical values tend to zero at the point of tangency.[2] The argument then proceeds as in the simple case considered above.

We consider next the possibility that a curve of singularities has a minimum, at which (dt/ds) is zero. If it did not leave the boundary of the physical sheet at the minimum, it would lead to a horn of singularities extending downwards as shown in Fig. 4. The method of analytic completion cannot be obstructed by such a downward pointing horn, as we noted above when discussing curves with positive slope and positive second derivative. If a curve had a maximum at some point but did not leave the physical sheet, it would either have a minimum at some other point or it would tend to infinity at some normal threshold in t. In order to stop the curve from going below the leading normal threshold in s, it is necessary for s to tend to plus infinity along both ends of the curve. This means that the curve would have a minimum in the variable s at which (ds/dt) is zero. As before, the method of analytic completion must not be obstructed by such a minimum. In this instance the analytic completion would begin from the analyticity in the upper half z_2 plane and proceed by varying z_1. If it were only necessary to consider one branch of a curve at a time, this would be sufficient to prove that all curves of singularities on the boundary of the physical sheet, where s and t are positive, must have negative slope. This is because only this type of curve satisfies the conditions we have proved, namely, that it leaves the physical sheet only asymptotically through a normal threshold and that it does not have a minimum. For a single branch of such a curve, a minimum would always lie below a maximum and therefore would obstruct the analytic completion in a manner that is not allowed. Also the curve cannot have a maximum without also having a minimum. Hence it must have negative slope everywhere in this region, as indicated by curve (a) in Fig. 1. These curves are allowed singularities of the function $F(z_1,z_2)$ analytically continued in the upper half z_1, z_2 planes and in the limit given by Eq. (3.14). Curves of singularities with negative slope do not lead to surfaces of singularities in the physical sheet in this region. Thus there would be no singularities of $F(z_1,z_2)$ for $z_1=s+i\epsilon$ and $z_2=t+i\epsilon'$ for any positive value of t if the curves of singularities all have negative slope.

However, it is necessary also to consider curves of singularities having more than one branch. A possible configuration is indicated in Fig. 6. In this case a maximum of a curve of singularities comes below a minimum.[9] It would obstruct the process of analytic completion since from either the s or from the t direction the maximum would be encountered first. A complex surface of singularities would extend into the physical sheet, (s_1+is_2, t_1+it_2) between the curves AB and CD.

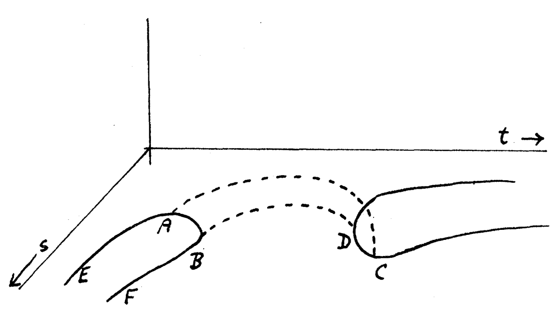

FIG. 6. Curves of singularities that would block the procedure of analytic completion. They are excluded because a single branch never has both ends tend to infinity in parallel directions.

[9] I am indebted to Dr. J. C. Polkinghorne and P. Landshoff for discussions on this point.

However, we will now show that the type of curve shown as $EABF$ is not allowed. At E, near the normal threshold in t, the coefficient g of t in the discriminant D, Eq. (3.16) satisfies

$$g(\alpha)>0, \tag{3.24}$$

for the critical values of the α variables. Since the slope is negative, we also have at E, from Eq. (3.17),

$$f(\alpha)>0. \tag{3.25}$$

At A the slope becomes infinite and g changes sign; hence, on AB, g is negative and f is positive. Similarly at B, the slope is zero so f changes sign. Hence at F both $f(\alpha)$ and $g(\alpha)$ must be negative. This contradicts the fact that near a normal threshold in t, it is necessary for g to be positive, as in Eq. (3.24). Hence the type of curve $EABF$ in Fig. 6 is not allowed. A similar argument shows that no curve can have an odd number of turning points. Since this discussion depends only on one branch of the curve, it is sufficient to show not only that the situation in Fig. 6 is not allowed, but also that we need not consider more complicated topologies. Each single branch of the curve must have an even number of turning points (if any) in each variable in order to have f and g with the known positive signs near the normal thresholds. This ensures that it starts from a normal threshold in t and ends at a normal threshold in s, in each case approaching the threshold asymptotically. For this type of curve a minimum always occurs below a maximum, that is for smaller values of t. Then the method of analytic completion applies and shows that no minimum is allowed.

This leads to the conclusion that the only curves of singularities of $F(z_1,z_2)$, in the limit shown in Eq. (3.14) in the region where s and t are positive must have negative slope. Their characteristic form is shown by the curve (a) in Fig. 1. In the limit shown in Eq. (3.14), there can be no curves of singularities in the region where t is positive and s is negative. We deduce that for

$$z_1=s+i\epsilon, \quad z_2=t+i\epsilon', \tag{3.25}$$

there are no singularities of $F(z_1,z_2)$ for any positive value of t. There is no special significance attached to our earlier use of the threshold $t=4m^2$ as an illustration. Similar arguments apply to all thresholds.

This completes the proof that the function

$$F(z_1, t+i\epsilon'), \tag{3.26}$$

has no singularities in the upper half z_1 plane for all positive values of t.

The analytic properties of the function

$$F(z_1, t-i\epsilon') \tag{3.27}$$

can be similarly investigated. By making a change of variable to z_3, the argument proceeds exactly as above. It shows that $F(z_1,z_2)$ is singular in the limit (with

$\epsilon > \epsilon'$),

$$z_1 = s+i\epsilon \rightarrow s, \quad z_2 = t-i\epsilon' \rightarrow t, \tag{3.28}$$

only in the region $t>4m^2$, $u>4m^2$. There it has curves of singularities that touch the normal thresholds in t and u. They are illustrated by curve (b) in Fig. 1. The restriction in Eq. (3.26) ensures that z_3 has a negative imaginary part. This is required to avoid the normal thresholds in u. This proves that the function (3.27) is analytic in the upper half z_1 plane above the line $z_1 = s+i\epsilon'$.

4. DOUBLE DISPERSION RELATION

The results of Sec. 3 permit us to transform the first integral on the right-hand side of Eq. (2.15). From the analyticity properties obtained for $F(z_1,z_2)$ the contour C in Eq. (3.1) can be displaced to give, for $t>4m^2$,

$$F(z_1, t+i\epsilon') = \frac{1}{2\pi i}\int_{-\infty}^{\infty}\frac{F(s+i\epsilon, t+i\epsilon')ds}{s+i\epsilon-z_1}, \tag{4.1}$$

$$F(z_1, t-i\epsilon') = \frac{1}{2\pi i}\int_{-\infty}^{\infty}\frac{F(s+i\epsilon, t-i\epsilon')ds}{s+i\epsilon-z_1}. \tag{4.2}$$

The difference between the expressions on the left of Eq. (4.1) and (4.2) is real for $z_1 = s$, $(-4m^2 < s < 4m^2)$ in the limit as ϵ, ϵ' tend to zero. From Eqs. (4.1) and (4.2) both functions are analytic in the upper half z_1 plane in this limit. Hence their difference satisfies a dispersion relation. The first integral on the right-hand side of Eq. (2.15) then leads to

$$\frac{1}{(2\pi i)^2}\int_{4m^2}^{\infty} dt\left[\int_{-\infty}^{-t} ds + \int_{4m^2}^{\infty} ds\right]\frac{\rho(s,t)}{(s-z_1)(t-z_2)}, \tag{4.3}$$

where we have replaced s in Eq. (2.15) by z_1, so as to obtain the relevant term in $F(z_1,z_2)$.

$$\rho(s,t) = \lim_{\epsilon,\epsilon'\rightarrow 0}\left[F(s+i\epsilon, t+i\epsilon') - F(s+i\epsilon, t-i\epsilon') - F(s-i\epsilon, t+i\epsilon') + F(s-i\epsilon, t-i\epsilon')\right]. \tag{4.4}$$

The second integral on the right of Eq. (2.15) can be considered in a similar manner by making a change of variable so as to keep u constant when transforming the integrand to give the second dispersion relation. This procedure was described in an earlier paper.[7] After taking account of the use of oblique axes, the integrals can then be combined to give the Mandelstam representation. This completes the proof of the Mandelstam representation with equal-mass particles.

5. GENERAL MASSES WITHOUT ANOMALOUS THRESHOLDS

In order to prove in a similar manner the validity of the Mandelstam representation for general masses we require (a) that there is a region in the real s, t plane where the amplitude is real, (b) that when t exceeds its least value in this real region the amplitude has no curves of singularities in the limit of $F(s+i\epsilon, t+i\epsilon')$ for which dt/ds is positive and decreasing, and (c) similar restrictions with respect to the other variables.

In the equal-mass case the condition (b) was proved from the fact that the curves of singularities tend asymptotically to the normal thresholds in s and t, and do not leave the physical sheet in any other manner. This same condition applies in the general-mass case. If there are no anomalous thresholds, then the only straight lines of singularities are normal thresholds. Curves of singularities leave the physical sheet only asymptotically through normal thresholds (taking into account the discussion on continuity in Sec. 3). The condition (a), that the amplitude is real, is sufficient to prove single-variable dispersion relations. These establish the starting region for the method of analytic completion with the contour C. The required displacement of the contour C will not be prevented by normal thresholds nor by the surfaces related to the curves that tend asymptotically to normal thresholds. No curve of singularities can have a minimum since this would contradict the proven displacement of the contour C to positive values of t. Since the curves are continuous there can be no maximum either. This then establishes the characteristic negative slope of curves that are singularities of $F(z_1,z_2)$ in the limit $z_1 = s+i\epsilon \rightarrow s$, and $z_2 = t+i\epsilon' \rightarrow t$. Similarly, given that there are no anomalous thresholds, the required analyticity can also be proved in the other limits.

We will now show that condition (a), that there is a region of the s, t plane where the amplitude is real, is ensured for all orders if its holds in fourth order. The discriminant $D(\alpha,s,t)$ for a general diagram has the form,[2]

$$D(\alpha,s,t) = sf(\alpha) + tg(\alpha) + \sum_1^4 M_j^2 K_j(\alpha) - \sum_1^n \alpha_i m_i^2 C(\alpha). \tag{5.1}$$

If the variable α_i corresponds to an internal line of the diagram, then

$$\partial D(\alpha,s,t)/\partial\alpha_i = D(\alpha,\alpha_i^{-1},s,t) - \alpha_i m_i^2 C(\alpha,\alpha_i^{-1}) - m_i^2 C(\alpha). \tag{5.2}$$

The notation α_i^{-1} indicates that the line labelled α_i is to be removed before evaluating the expression concerned. The coefficient $C(\alpha)$ in Eq. (5.1) is positive for α real and positive for all diagrams. Hence $C(\alpha,\alpha_i^{-1})$ is positive.

From Eq. (5.2) we see that if, in some region of the real s, t plane, for real positive α, we have

$$D(\alpha,\alpha_i^{-1},s,t) < 0, \tag{5.3}$$

then in the same region

$$\partial D(\alpha,s,t)/\partial\alpha_i<0. \qquad (5.4)$$

This proves that in the region in which Eq. (5.3) is valid, the variable α_i cannot have coincident singularities (for the physical branch of the amplitude). Hence in considering the singularities corresponding to $D(\alpha,s,t)$, it is sufficient to consider those with $\alpha_i=0$ so that the corresponding line is reduced. In this manner we can obtain a lower bound on the leading singularities of any diagram either by reducing any internal line, or by removing any internal line. It is clear from Eq. (5.1) that an increase in the mass m_i of any internal line will not cause a negative D to become positive. If the line i is on the mass shell, D will decrease as m_i is increased. It is therefore sufficient to consider only diagrams in which the lightest "allowed" masses occur in internal lines. Allowed masses are those that are compatible with selection rules. In removing lines from a diagram, selection rules must not be violated if the best lower bound is to be obtained for the location of singularities. However, the statement in Eq. (5.3) and (5.4) does not itself depend on selection rules and it may be convenient in some instances to ignore selection rules in removing internal lines, provided it is remembered that this may not give the best lower bound.

The above method of reduction or removal of internal lines will lead in general to the simplest reduced diagram having the structure of a fourth order diagram. The simplest such diagram will have the lightest allowed masses. We conclude that the condition for the amplitude to be real in some region of the real s, t plane is that the amplitude given by this reduced "fourth order" diagram shall be real in some region of the s, t plane.

We consider next the possibility of anomalous thresholds. We assume that there are none for the reduced "fourth order" diagram. Then the region in which D is negative will extend up to the leading normal threshold in each variable. If the leading normal threshold bounds not only the real region but also bounds a physical scattering region in each variable then there can be no anomalous thresholds. This is because any straight line of singularities must intersect a physical region since it must not enter the region where the amplitude is real. In physical regions the only singularities are at normal thresholds.[2] Examples of this type of process are given by pion-pion scattering, and by pion-nucleon scattering. For these processes there are no anomalous thresholds.

However, the leading normal threshold may not bound a physical region for the scattering process we are considering. For example, the leading normal threshold for nucleon-antinucleon scattering is given by an intermediate state of two pions. In considering the possibility of anomalous thresholds in such a case it is sufficient to limit the discussion to vertex parts. Then D is a function of one variable only, and

$$D(\alpha,s)=sP_1(\alpha)+M_2^2P_2(\alpha)+M_3^2P_3(\alpha)-\sum\alpha_i m_i^2C(\alpha), \qquad (5.5)$$

where $P_i(\alpha)$ is positive when α is real and positive. For ease of description, the nucleon-antinucleon case will be discussed. Below the two-pion threshold, D is negative for all vertex parts since the fourth order term leads to a vertex part that has D negative in this region. Above the two-pion threshold, if two pion lines for any diagram contain four-momenta satisfying

$$s=(q_1+q_2)^2, \qquad (5.6)$$

then the diagram cannot have a singularity with

$$q_1^2=m_\pi^2, \quad q_2^2=m_\pi^2. \qquad (5.7)$$

This result follows from the fact that only at the threshold

$$s=(2m_\pi)^2, \qquad (5.8)$$

will the Feynman diagram have an end-point singularity in the relative momentum of the two pions. It follows that one of the two pion lines can be reduced when s exceeds the threshold value. In the reduced diagram s must be carried by at least three pion lines. It can be seen by inspection that the simplest of these vertex diagrams will have D negative below the three-pion threshold. This is done by noting that the derivative of D with respect to a parameter for one of the nucleon lines is always negative. Hence this line cannot be on the mass shell and the diagram can be further reduced. The reduced diagram is known to have no anomalous thresholds. An argument similar to that used below the two-pion threshold can now be applied to show that all vertex diagrams having at least three pion lines carrying the energy squared must have negative D below the three-pion threshold. This argument is a generalization of the "majorization" procedure developed by Symanzik.[10] It shows that below each normal threshold only the simplest vertex diagrams corresponding to that threshold need be considered. The argument can be fairly readily applied to any individual case, and it shows for example that multiplarticle states do not lead to anomalous thresholds in nucleon-antinucleon scattering.

We conclude that (1) when the leading normal threshold in each variable bounds a physical region for the process under consideration, the absence of anomalous thresholds ensures their absence to all orders, and (2) when the leading normal threshold lies below the relevant physical region, the absence of anomalous thresholds can be verified by considering the simplest multiparticle reduced diagrams whose normal thresholds lie in the gap between the region where the amplitude is real and the region where it is physical.

---

[10] K. Symanzik, Progr. Theoret. Phys. (Kyoto) **20**, 690 (1947).

6. CONCLUDING REMARKS

The proof of the Mandelstam representation given in this paper applies when there are no anomalous thresholds. It has also been shown that the absence of anomalous thresholds is ensured when they are absent from a limited number of lower order diagrams. In some instances the fourth order diagram itself gives this information, as for example with pion-pion scattering, or pion-nucleon scattering. It may also be true in general that the fourth order diagram determines the absence of anomalous thresholds to all orders, but this has not been proved yet. It can fairly easily be checked for special cases, and for example it is true for nucleon-nucleon scattering.

The method of proof used in this paper can in principle also be applied when there are anomalous thresholds but not when there are super-anomalous thresholds. The characteristic of the latter is the existence of two branches of a curve on which the amplitude is singular in one spectral region, the two branches being connected by a singular surface in the physical sheet. The difficulty in practice of applying the method when there are anomalous thresholds is that curves of singularities can leave the boundary of the physical sheet through these thresholds at a finite point. There is therefore no simple continuity argument that prevents such a curve from having a maximum at some positive value of one of the energy variables. The aim of further investigation of anomalous thresholds must be to find out whether certain low-order diagrams, possibly the fourth order alone, determine the validity of the Mandelstam representation to all orders when there are anomalous thresholds.

There are several points in the proof where the perturbation series has been explicitly used. The method of analytic completion beginning with knowledge obtained from a single-variable dispersion relation is not specific to a perturbation treatment of the problem. However, the perturbation series, in particular the parametric representation of a Feynman diagram, has been used in showing that the required analytic completion is possible without distortion of the contour surrounding the tube of analyticity. This representation was required (1) to extend the single-variable dispersion relation up to the leading normal threshold, (2) to establish continuity of curves of singularities on the boundary of the physical sheet, and (3) to prove that a curve of singularities that leaves the boundary of the physical sheet through a normal threshold does so asymptotically.

In a previous report on this proof,[11] an alternative method was mentioned whereby the nonexistence of spurious turning points (i.e., minima of curves of singularities not at anomalous thresholds) and disconnected complex singularities can be established for the physical sheet. This method consists of tracing curves of singularities through the complex part of the physical sheet beginning from either a minimum of a curve of singularities or a disconnected complex singularity if either of these exist in the physical sheet. It can be shown that this complex curve must lead to a singularity below the leading normal threshold, for example in the variable t, with t real. This contradicts the single-variable dispersion relation and proves that the starting assumption of a spurious turning point or a disconnected complex singularity was wrong. This alternative method will not be described in detail since the manner in which it traces complex singularities is a special case of another alternative method that has been developed independently.

This alternative method for proving the Mandelstam representation has been developed by Landshoff, Polkinghorne, and Taylor.[12] Their method of approach starts from the opposite side of the problem to that used in this paper. Here we have started from a region in z_1, z_2 space know to be free from singularities and by standard methods have extended this region to show that there are no singularities inside a tube in this space. Several applications of this procedure were sufficient to prove the analyticity required for the Mandelstam representation. In contrast to this method, Landshoff, Polkinghorne, and Taylor assume the existence of singularities in the physical sheet, and then study the properties of the associated surfaces of singularities. By tracing curves of singularities in the complex space, they show that if one part of a branch of a surface of singularities is in the physical sheet, then all parts of this branch of the surface must be singular in the physical sheet. Since any such branch must intersect, for example, a forward scattering region ($t=0$, s complex) where there are no complex singularities, they can deduce that their original assumption of singularities in the physical sheet was incorrect, and there can be no complex singularities in the physical sheet. Their method is dependent also on a continuity argument that is related to the one used in this paper, but it is expressed as an induction procedure which is the converse of the way in which it is used here.

ACKNOWLEDGMENTS

This work was carried out under the auspices of the U. S. National Science Foundation at the Institute for Advanced Study, Princeton, New Jersey, and the U. S. Atomic Energy Commission at the Lawrence Radiation Laboratory of the University of California, Berkeley. The author is indebted to many colleagues in Princeton and Berkeley for valuable discussions, and in particular he wishes to thank Professor J. R. Oppenheimer, Professor G. F. Chew, and Professor S. Mandelstam.

[11] R. J. Eden, Phys. Rev. Letters 5, 213 (1960).

[12] P. V. Landshoff, J. C. Polkinghorne, and J. C. Taylor Nuovo cimento (to be published). This work was reported by Dr. Polkinghorne at the Proceedings of the Tenth Annual International Rochester Conference on High-Energy Physics, August 1960 (to be published).