General bounds on form factors and propagators from analyticity and unitarity: Application to the nucleon renormalization constant*

Varouzhan Baluni

Department of Mathematics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

D. J. Broadhurst

Department of Physics, Open University, Milton Keynes, England
and Laboratory for Nuclear Science and Department of Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

(Received 8 July 1976)

The following extremal problem is solved: Let O be an operator whose lowest-mass coupling is to the two particles A and B and is given by a form factor F which is analytic in a complex plane cut along a section L of the real line. Given values of F, at points not on L, and the partial-wave amplitudes for AB scattering in the channels with the same quantum numbers as O, over some portion of L, what is the optimal lower bound for a given positively weighted integral over L of the spectral function of O, consistent with the elastic *and* inelastic unitarity relations on L? The solution involves a system of two inhomogeneous singular integral equations of the Muskhelishvili type, which can be reduced to a singular integral equation of the Fredholm type. The results are applied to establish an upper bound of 0.25 for the nucleon renormalization constant, using the strong-coupling constant and πN scattering data in the P_{11} and S_{11} channels up to a c.m. energy of 1.7 GeV. The bound indicates that the nucleon is at least 75% composite.

I. INTRODUCTION

In this paper we shall examine some of the constraints imposed upon form factors and propagators by analyticity and unitarity.

In the last few years it has become clear that the general constraints on *scattering amplitudes* which follow from the requirements of analyticity and unitarity are of considerable practical importance.[1-4] The investigation of the content of analyticity and unitarity for *form factors* has been stimulated by results from electron-positron colliding beams. Following the approach originated by Meiman[5] and by Nguyen Van Hieu,[6] several important results have been reported. Restrictive bounds have been derived for the pion electromagnetic form factor[7,8] and for the scalar K_{l3} form factor[9,10] as consequences of their analytic properties.

Several important features distinguish the analysis of form factors from that of scattering amplitudes. The most obvious distinction is that a form factor depends upon only one complex variable, so that its analytic properties are much simpler than those of a scattering amplitude. In general, however, only the *modulus* of the form factor is measurable on its cut, which corresponds to the annihilation channel. This limitation is compensated by the possibility of measuring the form factor *inside* its domain of analyticity, either from scattering experiments (for the pion form factor in the spacelike region) or from three-body decays (for the K_{l3} form factors in the timelike region). Most importantly, the measurement of partial-wave amplitudes enables one further to constrain the form factor, using unitarity. In the simplest situation the elastic partial-wave amplitude fixes the phase of the form factor on the elastic cut, by Watson's theorem.

The general problem we study here is that of establishing relations between a form factor in the annihilation channel (alternatively, information about a propagator) and its values at measureable or theoretically interesting points inside its domain of analyticity. We shall be concerned *fully* to exploit any partial-wave information available. A comprehensive list of problems of physical interest has been given by Okubo.[11] We shall exemplify the content of our new inequalities by giving results, for the coupling of an off-shell nucleon to a physical nucleon and pion, which determine an upper bound for the renormalization constant of the nucleon.

The first major problem in this program is that of finding an *optimal* lower bound, consistent with *elastic* unitarity, for any given quadratic functional of the form factor in the annihilation channel (specifically a two-particle contribution to the dispersion relation for a propagator). By optimal we mean the most restrictive bound consistent with analyticity, elastic unitarity of the form factor, and an arbitrary amount of discrete information about the form factor (or its derivatives) on the real line. This extremal problem has been tackled independently by Okubo[11] and ourselves,[12] using different techniques. The problem has been reduced to that of solving a singular integral equation of the Fredholm type. Recently, a quite different and elegant mathematical method has been developed by Auberson, Mahoux, and Simão.[13]

These general results, which use *elastic* unitarity, have been applied to the scalar K_{l3} form factor, assuming unsubtracted dispersion relations for the propagators of current divergences.[12,14] The resulting inconsistency between the experimental data and the general structure of spontaneously broken SU(3)×SU(3) symmetry strongly indicates that the subtraction assumption is invalid. One is led to the important physical conclusion that the scale dimension of the chiral-symmetry-breaking Hamiltonian is greater than or equal to 3.[14]

Encouraged by this result, we proceed further with our program. The general problem we solve in this paper is that of finding an *optimal* lower bound for any given positively weighted integral of a spectral function (for example that in the sum rule for the nucleon renormalization constant, Z_2), given values of the form factor and both *elastic* and *inelastic* partial-wave amplitudes over some region of energies. The solution involves a system of *two inhomogeneous singular integral equations* of the type investigated by Muskhelishvili,[15] which we solve completely for the πNN form factor, given P_{11} and S_{11} partial waves up to ~1.7 GeV, to obtain the bound

$$Z_2 \leq 0.25 \; . \tag{1.1}$$

Our new results complete the program of implementing the content of unitarity for form factors, which had previously been partially exploited for the pion electromagnetic form factor by Auberson and Li,[16] for the scalar K_{l3} form factor by Bourrely[17] and ourselves,[12] and for the πNN form factor by ourselves.[18] We believe that our new results represent optimal bounds respecting elastic *and* inelastic unitarity and that no numerical improvement can be achieved without new input. Our method should be readily applicable to the computation of more restrictive bounds on the hadronic contributions to $g-2$ (Ref. 16) and on the dimension of the scale-breaking part of the Hamiltonian density.[19] (See Ref. 11 for a list of other possible applications.)

The paper is organized as follows. In Sec. II we formulate a general extremal problem based on the unitarity equations for propagators, form factors, and partial-wave amplitudes. We discuss inelastic contributions to the unitarity sums, particularly those relevant to the Z_2 calculations. Section III gives the details of solving the extremal problem and the resulting coupled integral equations. Section IV gives bounds on Z_2, together with the numerical results. In Sec. V we summarize our results.

II. UNITARITY AND ANALYTICITY AS CONSTRAINTS IN AN EXTREMAL PROBLEM

First the general class of physical problems which are of interest to us will be described. Then these problems will be recast in the form of an extremal problem for a certain functional. The functional will be defined on a class of functions restricted by various constraints including both unitarity and analyticity for the form factors. The general considerations will be applied to the problem of the nucleon renormalization constant, for which the relevant kinematics will be given.

A. Formulation of the problem

Suppose there is some local operator O, whose propagator $\Delta(t)$ has its *lowest* branch point at $t=(m_A+m_B)^2$ owing to the coupling to particles A and B. Let $F(t)$ represent the form factor describing the coupling of O to particles A and B and let $T(t)$ represent the partial-wave amplitude for the elastic scattering $AB \to AB$ in a state with the same quantum numbers (angular momentum, parity, isospin, etc.) as O.

Both functions $\Delta(t)$ and $F(t)$ will be analytic in the complex t plane cut along the real axis $[(m_A+m_B)^2, \infty]$. On the upper lip of this cut we have the three unitarity relations

$$\mathrm{Im}\Delta(t)=w_1(t)\left[\,|F(t)|^2+\sum_N |A_N(t)|^2\right], \tag{2.1}$$

$$\mathrm{Im}F(t)=F(t)T^*(t)+\sum_N A_N(t)B_N^*(t) \; , \tag{2.2}$$

$$\mathrm{Im}\,T(t)=|T(t)|^2+\sum_N |B_N(t)|^2 \; . \tag{2.3}$$

Here $w_1(t)$ is some positive weight function determined by the spin and other quantum numbers of O, A, and B. The functions $A_N(t)$ and $B_N(t)$ are proportional to the couplings of inelastic states N to O and to AB, respectively. The sum over N includes an integration over the multiparticle momentum phase space. Equations (2.1)–(2.3) are represented diagramatically by Figs. 1(a)–1(c).

We are concerned with finding the optimal lower bound for

$$I=\frac{1}{\pi}\int_{(m_A+m_B)^2}^{\infty} dt\, w_2(t)\,\mathrm{Im}\Delta(t) \; , \tag{2.4}$$

where $w_2(t)$ is a given positive weight function, and we are given $F(t)$ or its derivatives at points on the real line $[-\infty,(m_A+m_B)^2]$ and $T(t)$ on some portion of the cut $[(m_A+m_B)^2, \infty]$.

In the above analysis it was tacitly assumed that O is an operator of definite *integer* spin. For

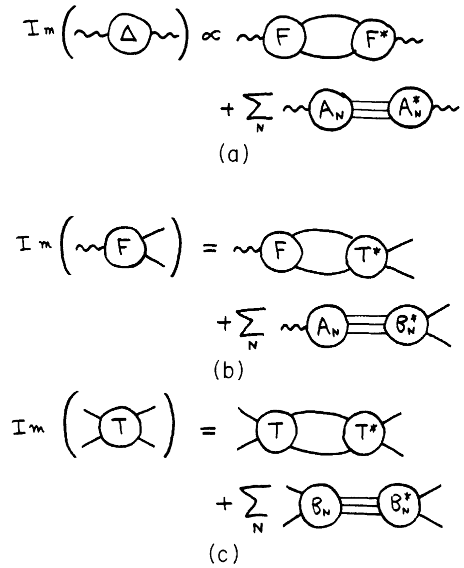

FIG. 1. The unitarity relations of Eqs. (2.1)–(2.3).

spinor operators of definite *half-integer* spin (for example an interpolating nucleon field) $\Delta(t)$ and $F(t)$ are analytic in the $w=\sqrt{t}$ plane with left-hand and right-hand cuts $[-\infty, -(m_A+m_B)]$ and $[(m_A+m_B), +\infty]$, the relevant partial-wave amplitudes in each case having the same spin and isospin as O but opposite parities by McDowell symmetry.

The relevant weight functions $w_1(t)$ and $w_2(t)$ for the K_{l3} problem are given by Li and Pagels[9] and for the Z_2 problem by Drell, Finn, and Hearn,[20] with subsequent amplification of the cut structure by Okubo[21] and of the relevant partial-wave amplitudes by ourselves.[18] We show in the next subsection that the inelastic couplings A_N and B_N for the Z_2 problem, as normalized by Eqs. (2.1) and (2.3), contribute to ImF in the symmetric fashion given in Eq. (2.2).

Incorporating half-integer-spin as well as integer-spin operators, O, we will establish the optimal lower bound for the functional given by Eqs. (2.1) and (2.4) as

$$I=\frac{1}{\pi}\int_{L(a,b)} dx\, h^2(x)\left[\,|F(x)|^2+\sum_N |A_N(x)|^2\right], \tag{2.5}$$

with

$$L(a,b)=(-\infty,-b)\,U(a,+\infty),\quad -b<a\ , \tag{2.6}$$

given the unitarity equations on the upper lip of the cut(s)

$$\mathrm{Im}F(x)=F(x)T^*(x)+\sum_N A_N(x)\,B_N^*(x),\quad \forall\, x\in L \tag{2.7}$$

$$0\le 1-\eta^2(x)=4\sum_N |B_N(x)|^2,\quad \forall\, x\in L \tag{2.8}$$

$$T(x)=\frac{1}{2i}\left[\eta(x)\,e^{2i\,\delta(x)}-1\right],\quad \forall\, x\in L\ , \tag{2.9}$$

where $F(x)$ is an analytic function in the x-plane cut along $L(a,b)$ and is known at some points in the interval $(-b,a)$. On a portion L_1 of the cut $L(a,b)$ both the phase shift $\delta(x)$ and the inelasticity $\eta(x)$ are known, whereas on the remaining portion L_2 only the latter is known within some limits

$$0\le\eta_-(x)\le\eta(x)\le\eta_+(x)\le 1,\quad \forall\, x\in L_2\ , \tag{2.10}$$

determined by experimental uncertaintites and unitarity bounds.

It is convenient to absorb the weight function $h(x)$ of Eq. (2.5) into the form factors $F(x)$ and $A_N(x)$. To preserve the analyticity of $F(x)$ one needs to construct the function $H(x)$ which is analytic in the x-plane cut along $L(a,b)$, satisfies

$$|H(x)|=h(x),\quad \forall\, x\in L(a,b)\ , \tag{2.11}$$

and has no zeros inside its domain of analyticity. This can be achieved as follows: Transform the complex x-plane onto the unit circle $|z|<1$, using the mapping

$$\frac{1-z}{1+z}=\left[\frac{(a-x)b}{(b+x)a}\right]^{1/2}\ , \tag{2.12}$$

which takes the points $x=-b,\ 0,\ a$ to $z=-1,0,1$, respectively, and maps the upper lips of the cuts $(-\infty,-b)$ and $(a,+\infty)$ to the upper half of the unit circle $z=e^{i\theta}$, $\pi\ge\theta\ge 0$, and the lower lips to the lower half. Then the function $H(x)$ is simply given by

$$H(x)=\exp\left\{\frac{1}{2\pi}\int_{-\pi}^{\pi} d\theta\,\ln|h(x(e^{i\theta}))|\left[\frac{e^{i\theta}+z(x)}{e^{i\theta}-z(x)}\right]\right\}. \tag{2.13}$$

With the notations

$$\overline{F}(x)=H(x)\,F(x)\ , \tag{2.14}$$

$$\overline{A}_N(x)=H(x)\,A_N(x)\ , \tag{2.15}$$

$$\overline{B}_N(x)=2i\,B(x)\exp[2i\arg H(x)]\ , \tag{2.16}$$

$$\epsilon(x)=\delta(x)+\arg H(x)\ , \tag{2.17}$$

the problem formulated in Eqs. (2.5)–(2.10) can be recast in the following general form.

Suppose that $\overline{F}(x)$ belongs to the class H_2 of functions which are analytic in the complex x plane, cut along L, and are quadratically integrable on L. Find the optimal lower bound for the functional

$$I=\frac{1}{\pi}\int_L dx\left[\,|\overline{F}(x)|^2+\sum_N|\overline{A}_N(x)|^2\right] \tag{2.18}$$

subject to the constraints

$$\mathrm{Re}\bar{F}(x) = \frac{1}{\pi} \mathrm{P} \int_L \frac{\mathrm{Im}\bar{F}(x')\,dx'}{x'-x}, \quad \forall x \in L \tag{2.19a}$$

$$\bar{F}(x) - \bar{F}^*(x)s(x) = \sum_N \bar{A}_N^*(x)\bar{B}_N(x), \quad \forall x \in L \tag{2.19b}$$

$$0 \leq 1 - |s(x)|^2 = \sum_N |\bar{B}_N(x)|^2, \quad \forall x \in L \tag{2.19c}$$

$$F_i \equiv \bar{F}(x_i) = \frac{1}{\pi} \int_L \frac{\mathrm{Im}\bar{F}(x)\,dx}{x - x_i}, \qquad i = 1, k;\ x_i \in (-b, a) \tag{2.20a}$$

$$s(x) = \eta(x)\exp[2i\epsilon(x)], \quad \forall x \in L_1 \tag{2.20b}$$

$$\eta_-(x) \leq |s(x)| \leq \eta_+(x), \quad \forall x \in L_2, \tag{2.20c}$$

where $L = L_1 \cup L_2$, $L_1 \cap L_2 = \emptyset$, and the real functions $\epsilon(x)$, $\eta(x)$, $\eta_\pm(x)$, and the numbers $\{F_i\}$ are given.

It is important to note than, besides the local constraints (2.19b) and (2.19c), the functions $\bar{A}_N(x)$ and $\bar{B}_N(x)$ are restricted by the integral constraint (2.19a) due to the analyticity of $\bar{F}(x)$. Unlike previous approaches, analyticity is implemented in the form of an integral relation between real and imaginary parts of $\bar{F}(x)$.

(We note that the extension to the case where derivatives of F are known is quite straightforward: there are constraints obtained by differentiating Eq. (2.20a) with respect to x_i.)

B. Nucleon renormalization constant

We assume that the nucleon can be represented by a renormalized interpolating local Dirac field $\psi(x)$. The nucleon propagator function $\Delta(w)$ is defined in terms of this field by

$$i\int d^4x\, e^{iq\cdot x}\langle 0|T(\psi(x)\bar{\psi}(0))|0\rangle = \frac{1}{2w}[(w + \not{q})\Delta(x) + (w - \not{q})\Delta(-w)], \tag{2.21}$$

where $q^2 = w^2$. The nucleon wave-renormalization constant Z_2 is then given by the sum rule[22]

$$Z_2^{-1} - 1 = \frac{1}{\pi}\left[\int_{-\infty}^{-(m+\mu)} + \int_{(m+\mu)}^{\infty}\right] dw\, \mathrm{Im}\Delta(w + i\epsilon), \tag{2.22}$$

where m and μ are the nucleon and pion masses, respectively. There is, of course, an infinity of possible interpolating fields. Our upper bound for Z_2, which involves only physical πN scattering data, applies for *any* interpolating field.

The form factor $F(w)$ is defined[23] in terms of the coupling of an off-shell nucleon to a physical pion (momentum k, isospin ρ) and physical nucleon (momentum p, isospin α, helicity λ) by

$$\langle 0|\eta(0)|p, \alpha, \lambda; k, \rho\rangle_{\mathrm{in}} = \frac{1}{2w}[(w + \not{q})F(w) + (w - \not{q})F(-w)] \times ig\gamma^5\tau_\rho u(p, \alpha, \lambda), \tag{2.23}$$

where $q = p + k$, $q^2 = w^2$, $\eta(x) = (i\not{\partial} - m)$, $\psi(x)$ is the nucleon source operator, and $\tau_\rho, \rho = 1, 2, 3$, are Pauli isospin matrices. The form factor is normalized by $F(m) = 1$, so that g is the πNN coupling constant obtained from analyses of πN dispersion relations. The analyticity of $F(w)$ has been proved by Bincer.[23]

To obtain the unitarity relations of Eqs. (2.1) and (2.2) we invert Eqs. (2.21) and (2.23) and take absorptive parts. For the propagator we find

$$\mathrm{Im}\Delta(\pm w + i\epsilon) = h^2(\pm w)|F(\pm w)|^2 + \frac{1}{2}\sum_{n \neq N\pi}(2\pi)^4\delta^4(q - n)|\langle 0|\psi_{1,3}(0)|n\rangle_{\mathrm{out}}|^2, \tag{2.24}$$

where $\psi_{1,3}$ are the first and third spinorial components of the Dirac field and the weight function h of Eq. (2.5) is given by[20]

$$h(\pm w) = \left\{\frac{3g^2[w \mp m)^2 - \mu^2]^{3/2}[(w \pm m)^2 - \mu^2]^{1/2}}{32\pi w^3(w \mp m)^2}\right\}^{1/2}. \tag{2.25}$$

For the form factor we find

$$\mathrm{Im}F(\pm(w + i\epsilon)) = F(\pm(w + i\epsilon))\,T^*(\pm w) - \frac{1}{2\sqrt{2}\,h(\pm w)}\sum_{n \neq N\pi}\langle 0|\psi_{1,3}(0)|n\rangle_{\mathrm{out}}\ {}_{\mathrm{out}}\langle n|N\pi, \pm\rangle, \tag{2.26}$$

where the state vector $|N\pi, \pm\rangle$ represents a positive-helicity πN system with c.m. energy w, total angular momentum $J = \frac{1}{2}$, isospin $I = I_3 = \frac{1}{2}$, and positive or negative parity; and is normalized as by Jacob and Wick.[24] The partial-wave amplitude T is defined by

$$T(+w) = \frac{1}{2i}[\eta_P(w)\,e^{2i\delta_P(w)} - 1]\ ,$$
$$T(-w) = \frac{1}{2i}(\eta_S(w)\,e^{2i\delta_S(w)} - 1)\ , \tag{2.27}$$

where $\delta_{S,P}(w)$ and $\eta_{S,P}(w)$ are the phase shifts and inelasticities in the P_{11} and S_{11} πN channels, respectively.[18]

Comparing Eqs. (2.22), (2.24), and (2.26) with Eqs. (2.5) and (2.7), we see that the problem discussed here is of the general form treated in the previous subsection, with the identifications

$$I \equiv Z_2^{-1} - 1, \tag{2.28}$$

$$a = b \equiv m + \mu, \tag{2.29}$$

$$A_n(\pm w) \equiv -\frac{1}{\sqrt{2}\,h(\pm w)}\,[\langle 0|\,\psi_{1,3}(0)|\,n\rangle_{\text{out}}], \tag{2.30a}$$

$$(2\pi)^4\delta^4(q-n)B_n(\pm w) \equiv \tfrac{1}{2}[\,_{\text{out}}\langle n|\,N\pi, \pm\rangle\,]. \tag{2.30b}$$

It remains to give the partial-wave unitary relations

$$\operatorname{Im} T(\pm w) = |\,T(\pm w)|^2 + \sum_{n\neq N\pi}(2\pi)^4\delta^4(q-n)|\,B_n(\pm w)|^2, \tag{2.31}$$

which reproduce Eq. (2.8).

The analytic function H, constructed from the weight function h of Eq. (2.25) according to the prescription of Eq. (2.13), is

$$H(w) = \left[\frac{3g^2}{32\pi}\right]^{1/2}\frac{[C(w, m+\mu)C(w, m-\mu)]^{3/2}[C(w, -m-\mu)C(w, -m+\mu)]^{1/2}}{[C(w,0)]^3[C(w,m)]^2}\ , \tag{2.32}$$

where

$$C(w,c) = \left[\left(\frac{m+\mu-w}{2}\right)\left(1+\frac{c}{m+\mu}\right)\right]^{1/2} + \left[\left(\frac{m+\mu+w}{2}\right)\left(1-\frac{c}{m+\mu}\right)\right]^{1/2}, \tag{2.33}$$

for $c^2 \leqslant (m+\mu)^2$. The function $H(w)$ is analytic in the w-plane cut along $L = (-\infty, -m-\mu) \cup (m+\mu, \infty)$ and satisfies Eq. (2.11) since

$$|\,C(w,c)|^2 = |\,w-c|\ ,\quad \forall w \in L. \tag{2.34}$$

We conclude this subsection by remarking that it has been assumed that no negative-metric states contribute to the unitarity sums of Eqs. (2.24) and (2.31). (Such states occur in gauge theories like QED and for these theories renormalization constants are gauge dependent.)

III. SOLUTION OF THE EXTREMAL PROBLEM

In this section the general extremal problem of Eqs. (2.18)–(2.20) will be exactly solved by the Lagrange-multiplier method. First it will be shown that the lower bound is determined by the solution of a system of two inhomogeneous singular integral equations of the Muskhelishvili type. The system will then be reduced to a Fredholm-type singular integral equation.

A. Lagrange equations for the extrema

Let us introduce Lagrange multipliers $\chi_I(x)$, $g(x)$, $h(x)$, $\{\lambda_i\}$ and $\mu_-(x)$, $\mu_+(x)$ corresponding to the equality constraints (2.19a), (2.19b), (2.19c), (2.20a) and inequality constraints (2.20c), respectively.[25] The meaning of the subscript I in $\chi_I(x)$ will become apparent. Unlike $\chi_I(x)$ and $h(x)$, the function $g(x)$ will be assumed to be complex, since the constraint (2.19b) is a relation for complex quantities and must be considered as a system of two real constraints. We record the important properties of the inequality multipliers μ_- and μ_+

$$\mu_-(x) = 0,\quad \text{if } \eta(x) > \eta_-(x),$$
$$\mu_-(x) > 0,\quad \text{if } \eta(x) = \eta_-(x); \tag{3.1a}$$

$$\mu_+(x) = 0,\quad \text{if } \eta(x) < \eta_+(x),$$
$$\mu_+(x) < 0,\quad \text{if } \eta(x) = \eta_+(x). \tag{3.1b}$$

The restrictions (3.1) merely reflect the fact that an increase (decrease) of η_+ (η_-) cannot result in an increase of the minimum value of the functional of Eq. (2.18).[26]

The Lagrange functional of the extremal problem is given by

$$\mathcal{L} = \frac{1}{\pi}\int_L dx\Big(|\bar F(x)|^2 + \sum_N |\bar A_N(x)|^2 - 2[\chi_I(x)F_R(x) + \chi_R(x)F_I(x)] - \operatorname{Im}\Big\{g(x)\Big[\bar F(x) - \bar F^*(x)s(x) - \sum_N \bar A_N^*(x)\bar B_N(x)\Big]\Big\} - h(x)\Big[1 - |s(x)|^2 - \sum_N |\bar B_N(x)|^2\Big] - 2f(x)F_I(x)\Big) + 2\sum_{i=1}^{k}\lambda_i F_i + \frac{2}{\pi}\int_{L_2} dx\{\mu_-(x)[\eta_-(x) - \eta(x)] + \mu_+(x)[\eta_+(x) - \eta(x)]\}, \tag{3.2}$$

where

$$F_R(x) \equiv \operatorname{Re}\bar F(x), \quad F_I(x) \equiv \operatorname{Im}\bar F(x), \tag{3.3a}$$

$$\chi_R(x) \equiv \frac{1}{\pi} P\int_L \frac{\chi_I(x')}{x' - x}\,dx', \tag{3.3b}$$

$$f(x) \equiv \sum_{i=1}^{k} \frac{\lambda_i}{x - x_i}. \tag{3.3c}$$

Observe that Eq. (3.3b) defines the Hilbert transform of the Lagrange multiplier $\chi_I(x)$. Hence the function

$$\chi(z) = \frac{1}{\pi}\int_L \frac{\chi_I(x)dx}{x - z} \tag{3.4}$$

is analytic in a complex z plane with boundary values

$$\chi(x + i\epsilon) = \chi_R(x) + i\chi_I(x), \quad \forall x \in L. \tag{3.5}$$

This fact will be essential in finding the lower bound (see next subsection).

It is convenient to complete squares in Eq. (3.2) to obtain

$$I = \mathcal{L} = 2\sum_i \lambda_i F_i + \frac{1}{\pi}\int_L dx\Big\{|\bar F(x) - i\chi^*(x) - if(x) - \tfrac{1}{2}i[g^*(x) + g(x)s(x)]|^2 + \sum_N |\bar A_N(x) - \tfrac{1}{2}ig(x)\bar B_N(x)|^2 + [h(x) - \tfrac{1}{4}|g(x)|^2]\sum_N |\bar B_N(x)|^2 - |\chi^*(x) + f(x) + \tfrac{1}{2}[g^*(x) + g(x)s(x)]|^2 - h(x)[1 - |s(x)|^2]\Big\} + \frac{2}{\pi}\int_{L_2} dx\Big\{\mu_-(x)[\eta_-(x) - \eta(x)] + \mu_+(x)[\eta_+(x) - \eta(x)]\Big\}. \tag{3.6}$$

The functions $F_{R,I}(x), \bar A_N(x), \bar B_N(x)$ for $x \in L$ and $\eta(x)$, $\epsilon(x)$ for $x \in L_2$ will be treated as independent variables. The requirement of vanishing first-order variations of $\mathcal{L}$ with respect to these variables leads to the Lagrange equations

$$\bar F(x) - i\chi^*(x) - if(x) - \tfrac{1}{2}i[g^*(x) + g(x)s(x)] = 0, \quad \forall x \in L \tag{3.7a}$$

$$\bar A_N(x) - \tfrac{1}{2}ig(x)\bar B_N(x) = 0, \quad \forall x \in L \tag{3.7b}$$

$$h(x) - \tfrac{1}{4}|g(x)|^2 = 0, \quad \forall x \in L \tag{3.7c}$$

$$h(x)\eta(x) + \frac{1}{2\eta(x)}\operatorname{Im}[\bar F^*(x)g(x)s(x)] - \mu_-(x) - \mu_+(x) = 0, \quad \forall x \in L_2 \tag{3.7d}$$

$$\operatorname{Re}[\bar F^*(x)g(x)s(x)] = 0, \quad \forall x \in L_2, \tag{3.7e}$$

where, in deriving the last two relations, Eq. (3.7a) has been taken into account. Equations (3.7) can be greatly simplified using constraints (2.19). Indeed, combining Eq. (3.7b) with Eqs. (2.19b), (2.19c), one finds

$$\bar F(x) - \bar F^*(x)s(x) = -\tfrac{1}{2}ig^*(x)[1 - |s(x)|^2].$$

Replacing $\bar F(x)$ by its expression from Eq. (3.7a), we arrive at

$$[\chi^*(x) + f(x)] + [\chi(x) + f(x)]s(x) = -[g^*(x) + g(x)s(x)].$$

Hence

$$g(x) = -[\chi(x) + f(x)], \tag{3.8}$$

which gives

$$\bar F(x) = -\tfrac{1}{2}i[g^*(x) - g(x)s(x)]. \tag{3.9}$$

Returning to Eqs. (3.7c)–(3.7e) with this relation, we derive

$$\frac{1}{4\eta(x)}\operatorname{Re}[g^2(x)s(x)]-\mu_-(x)-\mu_+(x)=0,\quad \forall x\in L_2 \tag{3.10a}$$

$$\operatorname{Im}[g^2(x)s(x)]=0,\quad \forall x\in L_2. \tag{3.10b}$$

We show that Eqs. (3.10a) and (3.10b) cannot be satisfied simultaneously unless $\eta(x)=\eta_+(x)$ or $\eta_-(x)$. The solution $g(x)=0$ to Eq. (3.10b) is trivial due to Eq. (3.9). Hence

$$\arg[g^2(x)]=-\arg s(x)+\pi n,\quad \forall x\in L_2, \tag{3.11}$$

where n is an arbitrary integer. Now Eq. (3.10a) becomes

$$\tfrac{1}{4}(-1)^n|g(x)|^2-\mu_-(x)-\mu_+(x)=0,\quad \forall x\in L_2.$$

Recalling conditions (3.1), we find two solutions:

$$n \text{ is odd: } \eta(x)=\eta_+(x),\quad \forall x\in L_2, \tag{3.12a}$$

$$n \text{ is even: } \eta(x)=\eta_-(x),\quad \forall x\in L_2, \tag{3.12b}$$

which correspond to different extrema. The characters of these extrema are determined by the signs of the second-order variations of $\mathcal{L}$ with respect to the independent variables. It is easy to show that condition (3.12a) corresponds to a local minimum. The extremal character of condition (3.12b) is more difficult to determine. We shall content ourselves with proving that condition (3.12a) corresponds to the *optimal lower bound* we are seeking. Consider the following inequality which can be obtained from Eqs. (2.20c) and (3.6):

$$I\geqslant 2\sum_i \mu_i F_i-\frac{1}{2\pi}\int_{L_1}dx\{|G(x)|^2-\operatorname{Re}[G^2(x)s(x)]\}-\frac{1}{2\pi}\int_{L_2}dx|G(x)|^2[1+\eta_+(x)], \tag{3.13}$$

where $\{\mu_i\}$ are *any* real numbers and

$$G(x)=-\left[\frac{1}{\pi}\,\mathrm{P}\int_L\frac{I(x')dx'}{x'-x}+iI(x)+\sum_i\frac{\mu_i}{x-x_i}\right],$$

where $I(x)$, $x\in L$, is *any* real function. We compare the general lower bound of inequality (3.13) with the specific extremum corresponding to conditions (3.7), (3.11), and (3.12a). We find

$$\mathcal{L}_{\text{ext}}=2\sum_i\lambda_i F_i-\frac{1}{2\pi}\int_{L_1}dx\{|g(x)|^2-\operatorname{Re}[g^2(x)s(x)]\}-\frac{1}{2\pi}\int_{L_2}dx|g(x)|^2[1+\eta_+(x)], \tag{3.14}$$

where

$$g(x)=-\left[\chi_R(x)+i\chi_I(x)+\sum_i\frac{\lambda_i}{x-x_i}\right],$$

with $\{\lambda_i\}$ and $\chi_I(x)$, $x\in L$, the Lagrange multipliers obtained from the extremal conditions and the constraints of Eqs. (2.19) and (2.20). Comparison of Eqs. (3.13) and (3.14) reveals that $\mathcal{L}_{\text{ext}}$ is indeed a *lower bound* for I. It is the *optimal* lower bound because it can be saturated consistent with all the constraints of analyticity, unitarity, and data which we have imposed on the problem. Hence

$$I=\mathcal{L}\geqslant\mathcal{L}_{\text{ext}}\equiv I_{\min}.$$

We conclude this subsection by expressing the extremal conditions in a form suitable for their solution in the next subsection. The form factor is given by Eqs. (3.9), (3.11), and (3.12a) as

$$\bar{F}(x)=-\left\{\frac{1}{\gamma(x)}\operatorname{Im}g(x)+w(x)g(x)\right\},\quad \forall x\in L \tag{3.15}$$

with

$$w(x)=\frac{1}{2i}[\eta(x)e^{2i\epsilon(x)}-1],\quad \forall x\in L_1$$
$$=\tfrac{1}{2}i[1+\eta_+(x)],\quad \forall x\in L_2 \tag{3.16a}$$

$$\gamma(x)=1,\quad \forall x\in L_1$$
$$=[1+\eta_+(x)]^{-1},\quad \forall x\in L_2. \tag{3.16b}$$

The lower bound is given by Eq. (3.14) as

$$I_{\min}=2\sum_i\lambda_i F_i+\frac{1}{\pi}\int_L dx\operatorname{Im}[g(x)\bar{F}(x)]$$
$$=\sum_i\lambda_i F_i \tag{3.17}$$

by virtue of Eq. (3.8) and constraints (2.19a) and (2.20a).

B. Muskhelishvili equations for the lower bound

The analytic functions $\bar{F}(x)$ and $\chi(x)$ are related by Eqs. (3.8) and (3.15), which give

$$\bar{F}(x)=\frac{1}{\gamma(x)}\operatorname{Im}\chi(x)+w(x)\chi(x)+w(x)f(x),\quad \forall x\in L. \tag{3.18}$$

Eq. (3.18) can be decomposed into a system of two equations of the Muskhelishvili type[15]

$$\operatorname{Im}[w(x)\chi(x)]=F_I(x)-f(x)\operatorname{Im}w(x), \tag{3.19a}$$

$$\operatorname{Im}[w^*(x)\bar{F}(x)]=\chi_I(x)\left[|w(x)|^2-\frac{1}{\gamma(x)}\operatorname{Im}w(x)\right]. \tag{3.19b}$$

This system of equations is a generalized form of the Hilbert problem. It will be solved for $\bar{F}(x)\in H_2$, the class of analytic functions which are quadratically integrable on their cut L.

We first construct the function

$$X(z)=\exp\left[\frac{z}{\pi}\int_L \frac{dx\,\phi(x)}{x(x-z)}\right], \tag{3.20}$$

where

$$\phi(x)=\arg w(x)\pmod{\pi}. \tag{3.21a}$$

To specify $X(z)$ completely we need a prescription for the choice of $\phi(x)$ in Eq. (3.21a). For the case with both a left-hand and right-hand cut, we require

$$|\phi(x_1)-\phi(x_2)|<\tfrac{1}{2}\pi+K|x_1-x_2|\,,\quad \forall x_1,x_2\in L, \tag{3.21b}$$

where K is a positive number, and

$$0<\lim_{z\to a,-b}\left|\frac{X(z)}{H(z)}\right|<\infty. \tag{3.21c}$$

Equations (3.20)–(3.21c) specify $X(z)$ completely. Equation (3.20) ensures that the functions $X(z)$ and $X^{-1}(z)$ are analytic in the complex-z-plane cut along L. Equation (3.21b) ensures that they are quadratically integrable on L, except possibly at the end points $z=a,-b,\pm\infty$. The behavior at the branch points $z=a,-b$ is fixed by Eq. (3.21c). The behavior at infinity is given by

$$0<\lim_{|z|\to\infty}|z^\lambda X(z)|<\infty, \tag{3.21d}$$

where

$$\lambda=[\phi(\infty)-\phi(-\infty)]/\pi$$

is determined by the input. From Eqs. (3.16b) and (3.21a) we see that

$$\phi(x)=\tfrac{1}{2}\pi\pmod{\pi},\quad \forall x\in L_2,$$

and hence λ is an integer (assuming that partial-wave amplitudes are known only up to a finite energy). For the Z_2 calculation of the next section we find $\lambda=0$, i.e., $X(z)$ is constant as $|z|\to\infty$.

From Eq. (3.21a)

$$w(x)=m(x)e^{i\phi(x)}, \tag{3.22}$$

where the real function $m(x)$ is not necessarily positive. We use $X(x)\equiv X(x+i\epsilon)$ to recast Eq. (3.19) in the form

$$\operatorname{Im}[X(x)\chi(x)]=\frac{a(x)}{c(x)}F_I(x)-f(x)\operatorname{Im}X(x), \tag{3.23a}$$

$$\operatorname{Im}[X^{-1}(x)\bar{F}(x)]=-\frac{b(x)}{a(x)}\chi_I(x)\equiv-\beta(x)\chi_I(x), \tag{3.23b}$$

where for convenience the following notations have been introduced:

$$a(x)\equiv\gamma(x)|X(x)|^2, \tag{3.24a}$$

$$c(x)\equiv\gamma(x)m(x)|X(x)|\,, \tag{3.24b}$$

$$b(x)\equiv\operatorname{Im}X(x)-c(x), \tag{3.24c}$$

$$\beta(x)\equiv\frac{b(x)}{a(x)}. \tag{3.24d}$$

It may be verified from Eqs. (3.16) and (3.22) that $\beta(x)=0$ in the elastic region $[\eta(x)=1]$.

We consider Eq. (3.23b) as an inhomogeneous equation for $\bar{F}(x)$. The most general solution is

$$\bar{F}(x)=-X(x)\left[P(x)+\frac{1}{\pi}\int_L\frac{\beta(x')\chi_I(x')dx'}{x'-x}\right], \tag{3.25}$$

where $P(x)$ is a polynomial. We require a solution $\bar{F}(x)\in H_2$ and thus from Eq. (3.21d) we have that $P(x)$ is a polynomial of degree $\lambda-1$, if $\lambda>0$. For $\lambda=0$, $P(x)=0$. For $\lambda<0$, $P(x)=0$ and $\chi_I(x)$ must satisfy $|\lambda|$ superconvergence relations on L. We have studied the question of subtraction constants and superconvergence relations for $\lambda\neq0$ and find that they can be accommodated in a generalization of the method we give below. In the rest of this section we assume $\lambda=0$, since this is the case for the calculation of the next section and simplifies the general analysis.

Dropping the polynomial from Eq. (3.25) we obtain

$$\bar{F}(x)=-X(x)[\beta(x)\chi(x)+\mathcal{K}\chi_I(x)], \tag{3.26}$$

where the integral kernel $\mathcal{K}$ is defined by

$$\mathcal{K}\chi_I(x)\equiv\frac{1}{\pi}\int_L\left(\frac{\beta(x')-\beta(x)}{x'-x}\right)\chi_I(x')dx'.$$

Note that the threshold behavior of the extremal function $\bar{F}(x)$ is given by Eqs. (3.21c) and (3.25) which ensure that the extremal form factor $F(x)=\bar{F}(x)/H(x)$ is finite for $x=a,-b$.

We now eliminate $F_I(x)$ from Eq. (3.23a), using Eqs. (3.24) and (3.26), to obtain

$$\operatorname{Im}[X(x)\chi(x)]=-a(x)\mathcal{K}\chi_I(x)-c(x)f(x). \tag{3.27}$$

Our goal is to reduce Eq. (3.27) to an equation for $\chi_I(x)$. Application of Cauchy's theorem to the product of analytic functions $X(z)$ and $\chi(z)$ gives

$$\chi(x)=-X^{-1}(x)\frac{1}{\pi}\int_L[a(x')\mathcal{K}\chi_I(x')+c(x')f(x')]\times\frac{dx'}{x'-x}. \tag{3.28}$$

It is straightforward to recast the imaginary part of Eq. (3.28) as the Fredholm-type integral equation

$$|X(x)|\,\chi_I(x)+\frac{1}{\pi}\int_L dx'[\beta(x')K_a(x,x')-K_b(x,x')]\chi_I(x')=\sum_i \lambda_i K_c(x,x_i), \tag{3.29}$$

where

$$K_0(x,x')\equiv\frac{[\Theta_H(x)-\Theta_H(x')]\sin\phi(x)-[\Theta(x)-\Theta(x')]\cos\phi(x)}{x-x'},$$
$$\Theta_H(x)\equiv\frac{1}{\pi}\,\mathrm{P}\int_L\frac{\Theta(x')}{x'-x}\,dx' \tag{3.30}$$

for $\Theta(x)=a(x)$, $b(x)$, and $c(x)$ with a prescribed convention that $c(x_i)=0$, $x_i\in L$.

This completes the formal derivation of the solution. Equations (3.29) and (3.26) determine $\chi_I(x)$ and $\overline{F}(x)$, respectively, in terms of the Lagrange multipliers $\{\lambda_i\}$ which are in turn fixed by the constraint Eq. (2.20a). The evaluation of the extremum according to Eqs. (3.14) or (3.17) is a straightforward but arduous task.

IV. BOUNDS ON Z_2

We apply the general results of the previous section to establish an optimal bound for the nucleon renormalization constant. We also give a conservative bound which is valid even if the solution to the integral equation (3.29) is known only approximately.

A. Optimal bound

It is convenient to map the cut $L=(-\infty,-m-\mu)\cup(m+\mu,\infty)$ in the complex x plane onto the segment $(-1,1)$ of the real line in a complex y plane defined by the change of variables

$$y=(m+\mu)/x. \tag{4.1}$$

The solution $\chi_I(x)$ of the integral equation (3.29) may then be written as

$$\chi_I(x)=-\frac{1}{x}\sum_j\lambda_j S_j\left(\frac{m+\mu}{x}\right), \tag{4.2}$$

where the functions $S_j(y)$, $y\in(-1,1)$, are the solutions of the integral equations

$$\exp[\tilde{\phi}_H(y)]S_j(y)+\frac{1}{\pi}\int_{-1}^{+1}dy'[\tilde{\beta}(y')\tilde{K}_a(y,y')-\tilde{K}_b(y,y')]S_j(y')+y_j\tilde{K}_c(y,y_j)=0,\quad y_j\in(-1,1), \tag{4.3}$$

where

$$\tilde{K}_0(y,y')\equiv\cos[\tilde{\phi}(y)]\left[\frac{\tilde{\Theta}(y)-\tilde{\Theta}(y')}{y-y'}\right]-\sin[\tilde{\phi}(y)]\left[\frac{\tilde{\Theta}_H(y)-\tilde{\Theta}_H(y')}{y-y'}\right], \tag{4.4}$$

$$\tilde{\Theta}_H(y)\equiv\frac{1}{\pi}\,\mathrm{P}\int_{-1}^{+1}\frac{dy'\,\tilde{\Theta}(y')}{y-y'}, \tag{4.5}$$

$$y_j\equiv(m+\mu)/x_j, \tag{4.6}$$

and we use the tilde to denote a function of y obtained from the corresponding function of x by the mapping of Eq. (4.1), i.e.,

$$\tilde{\Theta}(y)\equiv\Theta\left(\frac{m+\mu}{y}\right),\quad\Theta=a,\ b,\ c,\ \phi,\ \text{etc.} \tag{4.7}$$

The optimal upper bound for Z_2 is then given by Eq. (3.17) as

$$Z_2^{-1}-1\geq\sum_i\lambda_i F_i, \tag{4.8}$$

where

$$F_i=\overline{F}(x_i)=F(x_i)H(x_i), \tag{4.9}$$

and the extremal Lagrange multipliers $\{\lambda_i\}$ are determined by Eqs. (2.20a), (3.26), and (4.2), which lead to

$$F_i=\sum_j M_{ij}\lambda_j, \tag{4.10}$$

where

$$M_{ij}=\frac{y_i\exp[(\tilde{\phi}_H(y_i)]}{\pi(m+\mu)}\int_{-1}^{+1}\frac{dy\,\tilde{\beta}(y)S_j(y)}{y_i-y}. \tag{4.11}$$

Eliminating $\{\lambda_i\}$ from Eqs. (4.8) and (4.10), we obtain the upper bound

$$Z_2\leq\left[1+\sum_{i,j}F_i(M^{-1})_{ij}F_j\right]^{-1}, \tag{4.12}$$

where M^{-1} is the inverse of the square matrix M, whose elements are given by Eqs. (4.11).

Our program for calculating the upper bound of Eq. (4.12) is as follows:

(1) Use experimental πN partial-wave data to determine the input $\eta(x)$ and $\delta(x)$, $x\in L_1$, and set $\eta_+(x)=1$, $\forall x\in L_2$, to obtain the functions $w(x)$ and $\gamma(x)$ from Eqs. (3.16), (2.17), (2.32), and (2.33).

(2) Use $w(x)$, $x\in L$, to determine the phase $\tilde{\phi}(y)$, $y\in(-1,1)$, from Eqs. (3.21a)–(3.21c), which give

$$\tilde{\phi}(-1)=\pi/4, \tag{4.13a}$$

$$\tilde{\phi}(+1)=-3\pi/4 \tag{4.13b}$$

and ensure that any discontinuities of $\tilde{\phi}(y)$ are less than $\pi/2$ in magnitude for $0<|y|<1$.

(3) Calculate the Hilbert transform $\tilde{\phi}_H(y)$, using principal-value integration, and hence, from Eq. (3.20), determine

$$\tilde{X}(y)=\exp[\tilde{\phi}_H(y)+i\tilde{\phi}(y)], \quad \forall y\in(-1,1). \tag{4.14}$$

(4) Calculate $\tilde{a}(y), \tilde{b}(y), \tilde{c}(y)$ using Eq. (3.24), and hence, by principal-value integration, their Hilbert transforms, to obtain the kernels of Eqs. (4.3) and (4.4).

(5) Find approximate solutions $S_j(y)$, $y\in(-1,1)$, to the integral equations (4.3) by approximating them by finite-dimensional matrix equations.

(6) With these approximate solutions, calculate an approximate optimal bound from Eqs. (4.11) and (4.12) using as input the given values of the form factor $F(x_i)$ in Eq. (4.9).

(7) Verify the numerical convergence of the result as the mesh size in step (5) is decreased and the stability of the result against small variations of the parametrization of the data in step (1).

It will be appreciated that this is a long and demanding numerical calculation. The bound involves integrals of approximate matrix solutions to a singular integral equation whose kernel is obtained from parametrizations of the data via two successive principal-value integrations. We have therefore devised the following stringent test of the accuracy of our computation.

B. Conservative bound

Inequality (3.13) defines a class of rigorous upper bounds for Z_2 in terms of the arbitrary real parameters $\{\mu_i\}$ and the arbitrary real function $I(x)$, $x\in L$. In analogy with Eq. (4.2) we set

$$I(x)=-\frac{1}{x}\sum_j \mu_j T_j\left(\frac{m+\mu}{x}\right), \tag{4.15}$$

where the functions $T_j(y)$, $y\in(-1,1)$, are arbitrary. We now optimize the bound (3.13) with respect to $\{\mu_i\}$ by requiring that the right-hand side is an extremum with respect to variations of these parameters. (It is easy to show that the extremum is an absolute maximum.) The result is the following class of upper bounds:

$$Z_2\leq\left[1+\sum_{i,j}F_i(N^{-1})_{ij}F_j\right]^{-1}, \tag{4.16}$$

where

$$N_{ij}=\frac{1}{\pi(m+\mu)}\int_{-1}^{+1}dy\left\{\frac{1}{\tilde{\gamma}(y)}\operatorname{Im}h_i(y)\operatorname{Im}h_j(y)+\operatorname{Im}[\tilde{w}(y)h_i(y)h_j(y)]\right\}, \tag{4.17}$$

$$h_j(y)=\frac{1}{\pi}\,\mathrm{P}\int_{-1}^{+1}\frac{dy'T_j(y')}{y-y'}+iT_j(y)+\frac{y_j}{y-y_j}. \tag{4.18}$$

Equations (4.16)–(4.18) define a rigorous upper bound for Z_2 for *any* functions $T_j(y)$, $y\in(-1,1)$, such that the integrals of Eq. (4.17) exist. The optimal upper bound (4.12) is the special case of the general upper bound (4.16) for which $T_j(y)=S_j(y)$, the exact solutions to the integral equations (4.3). For any other choice of $T_j(y)$ the bound (4.16) is certainly valid, but it is not optimal. Indeed it can be shown that one obtains the integral equations (4.3) by requiring the right-hand side of inequality (4.16) to be an extremum with respect to variations of $T_j(y)$, $y\in(-1,1)$, and that this extremum be an absolute minimum. We are therefore able to take account of any numerical inaccuracies which may accrue in the computation of $S_j(y)$. We take the approximate values of $S_j(y)$ at the mesh points, obtained at step 5 of our computation of the optimal bound, and interpolate them by smooth functions $T_j(y)$ which we use in Eqs. (4.16)–(4.18) to obtain a bound which is certainly valid, but may be only approximately optimal. We call this bound a conservative bound. If in calculating this conservative bound we err from the optimal bound, we do so always on the safe side.

C. Input

Table I gives details of the partial-wave input data on the cut $(-1,1)$. The inelastic thresholds in the P_{11} and S_{11} channels are taken as $y=y_2$ and $-y_3$, respectively, where

$$y_n\equiv(m+\mu)/(m+n\mu).$$

Note that we take the S_{11} partial wave as elastic ($\eta=1$) below three-pion threshold. This is consistent with the data[27,28] and is to be expected, since for the S_{11} channel the two-pion final state is inhibited by angular momentum barriers which are not present in the P_{11} channel.

In the elastic regions, $I_{el}^{P,S}$, we use the same phase-shift parametrizations as were used in Ref.

TABLE I. Partial-wave input on the cut (−1,1).

Region	δ	η	Ref.
$I_{el}^S=(-1,-y_3]$	$-\delta_S$	1	27
$I_{in}^S=(-y_3,-y_S]$	$-\delta_S$	η_S	28
$I=(-y_S,y_P)$	unknown	≤ 1	$\cdots$
$I_{in}^P=[y_P,y_2)$	δ_P	η_P	28
$I_{el}^S=[y_2,1)$	δ_P	1	27

18, taken from the analysis of Roper *et al.*[27] In the inelastic regions, $I_{\rm in}^{P,S}$, we fit the single-energy partial-wave determinations of Almehed and Lovelace[28] with smooth functions $\delta_{P,S}$ and $\eta_{P,S}$. For each function we perform a least-squares fit of ~35 data points with ~10 parameters, demanding continuity with the elastic regions. The resultant phase $\tilde{\phi}(y)$, $y\in(-1,1)$, is sketched in Fig. 2.

We have found it convenient to choose the maximum energies up to which data is used such that $\tilde{\phi}(y)$ is continuous for $y=y_P, -y_S$, hence ensuring that

$$\infty > |\tilde{X}(y)| > 0, \quad |y| < 1 .$$

This is achieved with $y_P\approx 0.70$ and $y_S\approx 0.64$, corresponding to c.m. energies of 1.53 GeV and 1.68 GeV in the P_{11} and S_{11} channels, respectively. Below these energies the CERN and Saclay phase-shift analyses agree well.[29] At higher energies we use no data, but merely impose the unitarity bound $\eta\leq 1$.

In Sec. III it was claimed that $\tilde{X}(y)$ tends to a finite constant as $y\to 0$, thus obviating the necessity of considering subtractions of superconvergence relations in Eq. (3.25). This is true independently of the choice of $y_{P,S}$. Conditions (3.21a)–(3.21c) determine the crucial integer λ as

$$\lambda = \left[\frac{\phi(m+2\mu)}{\pi}\right] - \left[\frac{\phi(-m-3\mu)}{\pi}\right] = 0 ,$$

where the square brackets denote integer parts and we have used only the experimental values, $\delta_P(m+2\mu)$ and $\delta_S(m+3\mu)$.

Our input for the form factor is

$$F(m) = 1 , \tag{4.19}$$

which fixes the strong coupling constant, g. We take[30]

$$g^2/4\pi = 14.73 \pm 0.29 . \tag{4.20}$$

As in Ref. 18, we consider whether the bound for

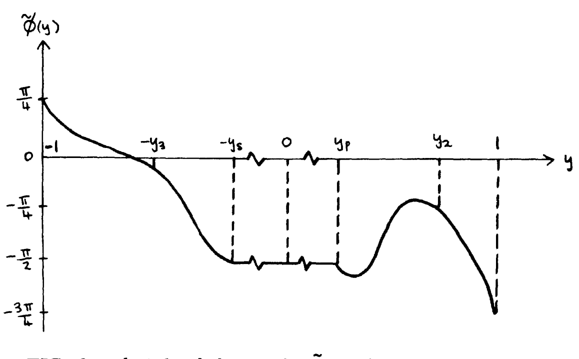

FIG. 2. Sketch of the angle $\tilde{\phi}$ as determined by πN data, using Eqs. (3.16) and (3.21).

Z_2 can be improved using the approximate current-algebra result[31]

$$1/F(-m) \approx g_A = 1.226 \pm 0.011 , \tag{4.21}$$

where g_A is the nucleon axial-vector coupling constant.

D. Results

For our optimal bound we use the strong coupling constant and πN scattering amplitudes to obtain

$$Z_2 \leq 0.25 \pm 0.02 ,$$

where the quoted error allows for the following uncertainties:

(i) the experimental error in the strong coupling constant;

(ii) mass uncertainties due to isospin violation;

(iii) variations of the parametrization of the elastic phase shifts, within the limits of Ref. 27;

(iv) variations of the parametrizations of the inelastic partial-wave amplitudes, within the small discrepancies between CERN and Saclay analyses[29];

(v) the approximations made in solving the integral equation.

The result converges quite rapidly as the matrix size is increased. For our most accurate calculation we evaluated $\tilde{\phi}_H(y)$ at 300 points in the interval $(-1,1)$. We then used cubic interpolation of the smooth, finite function

$$\frac{\exp[\tilde{\phi}_H(y)]}{(1-y)^{3/4}(1+y)^{1/4}}$$

to define $\tilde{a}(y)$, $\tilde{b}(y)$, and $\tilde{c}(y)$, whose Hilbert transforms were evaluated at 120 points in the interval $(-1,1)$. Thus the kernel of the integral equation was approximated by a 120×120 matrix, whose resolvent was found to a high precision. The result converged smoothly as the matrix size was increased from 60×60 to 120×120, changing by less than 4%.

No improvement of the optimal bound results from using the current-algebra result of Eq. (4.21). The extremal form factor corresponding to our upper bound satisfies this constraint, within the experimental errors in g_A. This is not surprising, since we found in Ref. 18 that the extremal form factor obtained using only *elastic* data is consistent with the current-algebra constraint.

To calculate conservative lower bounds we made smooth interpolations of the output of the matrix inversion, performed the Hilbert transformation of Eq. (4.18), and evaluated the matrix elements N_{ij} of Eq. (4.17). The resulting conservative bounds are *extremely* sensitive to the parametrization of $T_j(y) \approx S_j(y)$, unlike the optimal bound which is fairly insensitive to the parametrization of the solutions of the integral equation. It is easy to see how this arises. The matrix elements M_{ij} of Eq. (4.11), which define the optimal bound, do not involve $S_j(y)$ in the elastic region, $y \in I_{el}^{S,P}$, because $\beta = 0$ when $\eta = 1$. The matrix elements N_{ij} on the other hand are extremely sensitive to the parametrization of $T_j(y) \approx S_j(y)$ near the elastic thresholds at $y \to \pm 1$. As $y \to \pm 1$

$$S_j(y) \propto (1-y)^{-3/4}(1+y)^{-1/4},$$

and it appears as if the integrands of Eq. (4.17) possess singularities like $(1-y)^{-3/2}(1-y)^{-1/2}$ which would render the integrals divergent. In fact, several cancellations of this singularity occur. First of all, the leading singularity is canceled by virtue of the threshold angles of Eq. (4.13). A further cancellation occurs by virtue of the integral equation (4.3), so that the integrands actually *vanish* like $(1-y)^{3/2}(1+y)^{1/2}$ at the elastic thresholds. Clearly it is very difficult to achieve all these cancellations with an interpolation $T_j(y)$ of the approximate matrix solution. We used a parametrization of the threshold behavior and an algorithm for principal-value integration which guarantee that the integrands are no more singular than $(1-y)^{-1/2}$ as $y \to \pm 1$, thereby obtaining conservative bounds of around 0.32 and 0.36, with and without the current-algebra constraint. The extent to which these conservative bounds fall short of the optimal bound can be entirely attributed to the elastic regions. This may be demonstrated by replacing the integrand of Eq. (4.17), *in the elastic region only*, by an analytic function of the approximate solutions, $T_j(y)$, which is identical to the integrand in the limit that $T_j(y) = S_j(y)$, the exact solutions. The conservative bound then reproduces the optimal, within the quoted errors. This agreement gives us considerable confidence in our approximate solutions *outside* the elastic-threshold regions, which determine the approximate optimal upper bound $Z_2 \leq 0.25 \pm .02$. We believe that this bound accurately represents the maximal constraint on the renormalization of the nucleon's wave function, imposed by low-energy πN scattering data, through analyticity and unitarity.

E. Singularities of the kernel

It remains to comment on the use of finite-dimensional methods to obtain approximate solutions of the integral equation. This procedure is known to be legitimate for integral kernels which belong to the class of completely continuous compact operators.[32] In this case it can be proved that finite-dimensional solutions converge to the continuum solution as the dimensionality is increased. The kernel of our integral equation is not, however, completely continuous. It has singularities at $y = -y_S$ and $y = y_P$, which correspond to the energies beyond which πN data is ignored. At these points the functions $\tilde{a}(y)$ and $\tilde{b}(y)$ are discontinuous and hence their Hilbert transforms have logarithmic singularities. The resultant singularities of the kernel are of the same type as the singularities at $y = 0$ of the kernels

$$K_1(y, y') = \frac{\lambda_1}{\pi}\left[\frac{\theta(y) - \theta(y')}{y - y'}\right],$$

$$K_2(y, y') = \frac{\lambda_2}{\pi^2}\left[\frac{\ln|y/y'|}{y - y'}\right],$$

defined on some finite segment of the real line which includes $y = 0$. The kernels K_1 and K_2 are *bounded* operators,[33] with norms

$$\|K_1\| \leq |\lambda_1|,$$

$$\|K_2\| \leq |\lambda_2| \sum_{n=0}^{\infty} [\pi(n+\tfrac{1}{4})]^{-2} < 1.7425|\lambda_2|.$$

Convergence properties of special-matrix-inversion methods have been proved by Jones and Tiktopoulos[33] for kernels having a singular component with norm less than unity. This restriction on the norm is not essential; what is necessary is that the spectrum of eigenvalues of the operator does not contain unity. There exist powerful methods to deal with the more general case.[34] Our approach is a pragmatic one. Rather than undertake the task of calculating an upper limit on the norm of our integral kernel, we content ourselves with the observation that it is a bounded operator, that the results of the finite-dimensional calculations are well behaved as the dimensionality is increased from $n = 60$ to $n = 120$, and that we obtain conservative bounds which are less restrictive than the optimal bound only by virtue of approximation errors that accrue in the *elastic* regions, which are far from the singularities.

V. Conclusion

We have shown how to bound a spectral integral, given data on the appropriate form factor and scattering amplitudes. The combined constraints of analyticity, elastic unitarity, and inelastic are stringent and intricate. They lead to a system of two singular integral equations which determine the extremum. Using the methods developed by Muskhelishvili[15] we have succeeded in reducing the problem to the solution of a Fredholm integral equation, involving an integral operator with bounded norm. The numerical solution of this equation in the case of the nucleon spectral function yields a new, stringent bound

$$Z_2^{-1} - 1 > 3.0, \quad Z_2 < 0.25, \tag{5.1}$$

given reliable πN scattering data.

The significance of Z_2 as a measure of the probability for finding a "bare" particle inside the dressed, physical nucleon has been widely discussed.[35] We believe that the phenomenological success of constituent quark models and the approximate scaling of deep-inelastic structure functions leave no doubt that hadrons are *predominantly* composite. It is therefore interesting to inquire how the observed interactions of hadrons may be translated into a *quantitative* measure of the compositeness of the nucleon. As this measure we take Z_2 as defined for *any* interpolating nucleon field by the sum rule of Eq. (2.22). Our result (5.1) then indicates that the nucleon is at least 75% composite. We find it remarkable that this result may be obtained from low-energy πN data on general principles, without invoking specific assumptions or approximations. The stringency of our bound is a good measure of just how strong the strong interaction is.

The utility of scattering data in determining our bound is apparent when one compares it with previous results. Using only the strong coupling constant, Drell *et al.*[20] established the first rigorous bound on Z_2:

$$Z_2^{-1} - 1 > 0.17, \quad Z_2 < 0.85, \tag{5.2}$$

indicating that the nucleon is at least 15% composite. Using the observed P_{11} and S_{11} elastic shifts we were able[18] to improve this bound to

$$Z_2^{-1} - 1 > 1.42, \quad Z_2 < 0.41, \tag{5.3}$$

indicating that the nucleon is at least 60% composite. Comparison of results (5.2) and (5.3) shows that knowledge of the elastic phase shifts improves the bound on the spectral integral by a factor of 8. Comparison of results (5.1) and (5.3) shows that knowledge of the partial-wave amplitudes in the inelastic region up to 1.7 GeV gives a further improvement of more than 100% in this lower bound. The very inelasticity which forbade the use of Watson's theorem at higher energies in previous bounds[11-14,17,18] is now a virtue, since the inelastic contributions to the scattering amplitude set a lower bound on the inelastic contributions to the spectral integral, thanks to the analyticity of the form factor.

Attempts to bound Z_2 in terms of high-energy data on deep-inelastic structure functions[36-39] or form factors[40] have not produced results of comparable rigor. One ends up either with a plausibility argument[39,40] for $Z_2 = 0$ or else with an upper bound which involves the poorly determined longitudinal structure function.[37,38]

We conclude by remarking that our limit on the elementarity of the proton is more stringent and more rigorous than Weinberg's limit[41] on the elementarity of the deuteron. For the renormalization constant of the deuteron, Z_D, Weinberg used a nonrelativistic analysis of the nucleon-nucleon interaction to obtain

$$Z_D < 0.2\alpha, \tag{5.4}$$

where α is an unknown parameter of order unity. It appears that the nucleon is known to be at least as composite as the deuteron. Of course, little interest attaches to improving the deuteron bound. We are sufficiently convinced of the compositeness of the deuteron, having detected its constituents. Until constituents of the nucleon are found, or made to satisfy the requirements of confinement, one may have to be content with an upper limit of 25% on the "elementarity" of the nucleon.

ACKNOWLEDGMENT

One of the authors (V. B.) takes this opportunity to thank Professor Ken Johnson and Professor Francis Low for many helpful conversations and for the warm hospitality at the Center for Theoretical Physics, M. I. T.

---

*Work supported in part by NSF under Grant No. GP-29463 and by ERDA under Contract No. EV-76-C-02-3069.*000.

[1]A. Martin and F. Cheung, *Lectures on Analyticity Properties and Bounds of Scattering Amplitudes* (Gordon and Breach, New York, 1969).

[2]M. B. Einhorn and R. Blankenbecler, Ann. Phys. (N.Y.) 67, 480 (1971).

[3]S. M. Roy, Phys. Rep. 5, 125 (1972).

[4]F. J. Yndurain, Rev. Mod. Phys. 44, 645 (1972).

[5]N. N. Meiman, Zh. Eksp. Teor. Fiz. 44, 1228 (1963) [Sov. Phys.–JETP 17, 830 (1963)].

[6]Nguyen Van Hieu, Yad. Fiz. 7, 1111 (1968) [Sov. J. Nucl. Phys. 7, 667 (1968)].

[7]V. Baluni, in *High Energy Physics*, in proceedings of the Fifteenth International Conference on High Energy Physics, Kiev, 1970, edited by V. Shelest (Naukova Dumka, Kiev, U.S.S.R., 1972); Phys. Lett. 38B, 535 (1972) and references cited therein.

[8]I. Raszillier, Lett. Nuovo Cimento 2, 347 (1971).

[9]L. F. Li and H. Pagels, Phys. Rev. D 3, 2191 (1971); and 4, 255 (1971).

[10]S. Okubo and I. F. Shih, Phys. Rev. D 4, 2020 (1971).
[11]S. Okubo, J. Math. Phys. 15, 963 (1974).
[12]V. Baluni and D. J. Broadhurst, Nucl. Phys. B84, 178 (1975).
[13]G. Auberson, G. Mahoux, and F. R. A. Simão, Nucl. Phys. B98, 204 (1975).
[14]D. J. Broadhurst, Nucl. Phys. B85, 189 (1975).
[15]N. I. Muskhelishvili, *Singular Integral Equations*, edited by J. Radok (Groningen, Holland, 1953).
[16]G. Auberson and L. F. Li, Phys. Rev. D 5, 2269 (1972).
[17]C. Bourelly, Nucl. Phys. B43, 434 (1972); B53, 289 (1973).
[18]V. Baluni and D. J. Broadhurst, Phys. Rev. D 7, 3738 (1973).
[19]D. N. Levin, S. Okubo, and D. R. Palmer, Phys. Rev. D 4, 1847 (1971).
[20]S. D. Drell, A. C. Finn, and A. C. Hearn, Phys. Rev. 136, B1439 (1964).
[21]S. Okubo, Prog. Theor. Phys. 48, 1986 (1972).
[22]J. D. Bjorken and S. D. Drell, *Relativistic Quantum Mechanics* (McGraw-Hill, New York, 1965), p. 151. (The spectral functions are unrenormalized in this reference.)
[23]A. M. Bincer, Phys. Rev. 118, 855 (1960).
[24]M. Jacob and G. C. Wick, Ann. Phys. (N.Y.) 7, 404 (1959).
[25]A lucid discussion of the Lagrange multiplier method for functions in an n-dimensional Euclidian space is given in M. R. Hestenes, *Calculus of Variations and Optimal Control Theory* (Wiley, New York, 1966), Chap. 1. We are using the extension of this method to functionals in the space of quadratically integrable analytic functions.
[26]The usefulness of these simple considerations was demonstrated in Ref. 2 by Einhorn and Blankenbecler, who first applied the Lagrange method to the analysis of scattering amplitudes with equality *and* inequality constraints.
[27]L. D. Roper, R. M. Wright, and B. T. Feld, Phys. Rev. 138, B190 (1965). We use solution 24 of this reference.
[28]S. Almehed and C. Lovelace, CERN Report No. CERN-TH-1408, 1971 (unpublished).
[29]Particle Data Group, Rev. Mod. Phys. 45, S106 (1973).
[30]G. Ebel, H. Pilkuhn, and F. Steiner, Nucl. Phys. B17, 1 (1970).
[31]H. Suura and L. M. Simmons, Jr., Phys. Rev. 148, 1579 (1966); H. S. Mani, G. Mohan, and L. K. Pande, Nuovo Cimento 44, 265 (1966).
[32]K. Maurin, *Methods of Hilbert Spaces* (PWN—Polish Scientific, Warsaw, 1967), ch. 7.
[33]C. E. Jones and G. Tiktopoulos, J. Math. Phys. 7, 311 (1966).
[34]Yu. V. Vorobyev, *Methods of Moments in Applied Mathematics* (Gordon and Breach, New York, 1965).
[35]See the references cited in Ref. 18, in particular E. Tirapegui, Phys. Rev. D 4, 3807 (1971); and K. Hayashi *et al.*, Fortschr. Phys. 15, 625 (1967).
[36]F. Cooper and H. Pagels, Phys. Rev. D 2, 228 (1970).
[37]G. B. West, Phys. Rev. Lett. 27, 762 (1971).
[38]D. J. Broadhurst, Phys. Rev. D 5, 2615 (1972).
[39]M. Hirayawa and J. Ishida, Prog. Theor. Phys. 48, 2377 (1972).
[40]M. Gell-Mann and F. Zachariasen, Phys. Rev. 123, 1065 (1961). See G. Barton, *Introduction to Dispersion Techniques in Field Theory* (Benjamin, New York, 1965), p. 140, for a discussion.
[41]S. Weinberg, Phys. Rev. 137, B672 (1965).