

Quark-hadron duality in a relativistic, confining model

Sabine Jeschonnek

Physics Department, The Ohio State University, Lima, Ohio 45804

J. W. Van Orden

*Jefferson Lab, 12000 Jefferson Avenue, Newport News, Virginia 23606
and Department of Physics, Old Dominion University, Norfolk, Virginia 23529*

(Received 18 January 2002; published 17 May 2002)

Quark-hadron duality is an interesting and potentially very useful phenomenon, as it relates the properly averaged hadronic data to a perturbative QCD result in some kinematic regimes. While duality is well established experimentally, our current theoretical understanding is still incomplete. We employ a simple model to qualitatively reproduce all the features of the Bloom-Gilman duality as seen in electron scattering. In particular, we address the role of relativity, give an explicit analytic proof of the equality of the hadronic and partonic scaling curves, and show how the transition from coherent to incoherent scattering takes place.

DOI: 10.1103/PhysRevD.65.094038

PACS number(s): 12.40.Nn, 12.39.Ki, 13.60.Hb

I. INTRODUCTION

Quark-hadron duality has been well established experimentally [1] for over 30 years, but our theoretical understanding of the phenomenon is quite limited so far. In the past year, there has been renewed interest in duality, both on the experimental [2,3] and theoretical sides [4–8]. Duality is a major point in the planned 12 GeV upgrade of the Continuous Electron Beam Accelerator Facility (CEBAF) at Jefferson Lab [9]. Duality between partons and hadrons is also employed in QCD sum rules [10].

In a recent publication [4], we presented results obtained in a confining, relativistic model which qualitatively reproduced the features seen in Bloom-Gilman duality. In [4], we only discussed a reaction where all particles involved—electrons, photons, and quarks—were treated as scalars. In this paper, we present the model for physical electrons and photons, and only treat the quarks as scalars. We describe the model and its properties in more detail, and discuss the Coulomb sum rule, the transition from coherent to incoherent scattering, and duality in the form factors. We also put special emphasis on the role of relativity. A relativistic treatment was one of four basic conditions imposed in [4] to obtain duality, and here we show the consequences of relaxing this condition, and compare nonrelativistic and relativistic calculations. One main point of this paper is the explicit derivation of the scaling curve for scattering from quarks confined in their initial and final states and a comparison of this scaling curve to that obtained in a parton model calculation.

For the convenience of the reader, we define the concept of duality in the following and briefly discuss a few basic implications. In the literature, there exist many slightly varying “definitions” and usages of the term duality, and the phenomenon manifests itself experimentally in many different processes. We begin with a definition that covers all these cases. First, we need to make an obvious observation: any hadronic process can be correctly described in terms of quarks and gluons. In other words, quantum chromodynamics (QCD) is the correct theory for strong interactions. While this statement is obvious, it has little practical value, as in

most cases we cannot perform a full QCD calculation. For example, in order to calculate a resonance excitation form factor, one would need to include very many quarks and gluons, and they would all couple strongly. We will refer to the above statement that any hadronic process can be described by a full QCD calculation as “degrees of freedom duality.”

There exists a more practical and less obvious version of the first statement: in certain kinematic regions, the average of hadronic observables is described by a perturbative QCD (PQCD) result. This is the statement of duality, and we are going to explain the details in the following.

With the PQCD result, we indicate the result for the underlying quark process—for inclusive inelastic electron scattering from a proton, it is free electron-quark scattering; for semileptonic decays, e.g., $\bar{B} \rightarrow X_c l \bar{\nu}_l$, it is the underlying quark decay rate, in this case obtained from the process $b \rightarrow c l \bar{\nu}_l$ [11]; for $e^+ e^- \rightarrow$ hadrons, it is the underlying $e^+ e^- \rightarrow q \bar{q}$ process. Now, it is clear that we expect perturbative QCD to describe nature in a certain kinematic regime, i.e., for very large Q^2 . In this regime, as a result of the fact that full QCD is approximated by perturbative QCD, the statement of duality turns into the statement of the “degrees of freedom duality.” So we have identified one kinematic regime in which even the nonobvious version of duality must hold.

We also can identify a kinematic regime for which duality cannot hold: for $Q^2 \rightarrow 0$. While the underlying reason for the breakdown of duality at low four-momentum transfers is the nonperturbative, strong interaction of the hadrons, one can see the breakdown of duality most easily by considering the transition from incoherent to coherent scattering. For duality to hold for the nucleon structure functions in this case, we would need the following: the elastic proton and neutron form factors, which take the value of the nucleon charge for $Q^2 \rightarrow 0$, would have to be reproduced by electron scattering off the corresponding u and d quarks. Now, for the proton this can work, as the squares of the charges of two u quarks and one d quark add up to 1. However, for the neutron, the

squared quark charges cannot add up to 0, so it is clear that duality in inclusive inelastic electron scattering from a neutron must fail for $Q^2 \rightarrow 0$. In addition, we know from gauge invariance that for $Q^2 \rightarrow 0$, at fixed energy transfer ν , the function $\nu W_2(\nu, Q^2)$ must approach 0. It is clear that the scaling function $F_2(x)$ does not show that behavior, which gives us an additional reason to expect the breakdown of duality at low Q^2 .

So now we know that duality has to hold in one kinematic regime and that it has to break down in another kinematic regime. Obviously, a very interesting question is what happens in between these regimes, i.e., how exactly does duality break down, how far does it hold in the regime where it is nontrivial, i.e., for moderate values of Q^2 , and how accurately does it hold where it holds.

II. MODEL

Here, we present a model for the study of quark-hadron duality that uses only a few basic assumptions. Namely, we assume that it is necessary to include confinement and relativity in our model, that it is sufficient to base our model solely on valence quarks, and that these quarks can be treated as scalars. A model with these features will not give a realistic description of any data, but it should allow us to obtain duality and study the critical questions of when and how accurately duality holds.

Although it is our aim to study duality in electron scattering from the nucleon, i.e., from a three-quark-system, as a first step we simplify the problem at hand by substituting two quarks by an antiquark, as the representation $3 \otimes 3$ in SU(3) contains the representation $\bar{3}$. This means we have a two-body problem now, and we have to solve the Bethe-Salpeter equation. In the special case of the mass of the antiquark, M , going to infinity, the problem further simplifies to a one-body problem. In the case of scalar quarks considered here, we obtain a Klein-Gordon equation. In contrast to [4] where we assumed that all particles involved—electrons, photons, and quarks—are scalars, here we treat only the valence quarks as scalars. Note that the experiment which our model resembles most would be electron scattering from a B meson. Still, we expect to gain valuable insights from considering this case and would like to stress that none of the assumptions we made here prevent us from extending our model to describe more realistic circumstances.

We have chosen to implement confinement by using a linear potential, which leads to a relativistic harmonic oscillator solution. This has the advantage that analytic solutions can be readily obtained and that a comparison with the nonrelativistic case is easily feasible.

We have to solve the Klein-Gordon equation with a scalar potential:

$$\left(\frac{\partial^2}{\partial t^2} - \vec{\nabla}^2 + m^2 + V^2 \right) \Phi(x) = 0, \quad (1)$$

with the usual ansatz $\Phi^\pm(x) = \Phi(\vec{r}) \exp(\mp iEt)$ and the confining potential

$$V^2(\vec{r}) = b^2 r^2, \quad (2)$$

where b is the relativistic string tension and has dimension $[b] = [\text{mass}^2]$. The superscript of the wave function denotes positive and negative energy solutions to the Klein-Gordon equation. The mass of the quark is denoted by m and we use $m = 0.33 \text{ GeV}$ throughout this paper.

The Klein-Gordon equation in this form can be easily rearranged to have the form of a Schrödinger equation to give

$$\left(-\frac{\vec{\nabla}^2}{2m} + \frac{1}{2} \frac{b^2}{m} r^2 \right) \Phi(\vec{r}) = \frac{E^2 - m^2}{2m} \Phi(\vec{r}), \quad (3)$$

where the similarity to the Schrödinger equation for a nonrelativistic harmonic oscillator potential becomes apparent. The solutions to this equation are easily obtained by making the substitutions $\tilde{\kappa} \equiv b^2/m$ and $\tilde{E} \equiv (E^2 - m^2)/2m$ and using the well-known solutions of the nonrelativistic case.

The energy eigenvalues for the Klein-Gordon equation are $E = \pm E_N$ where

$$E_N = \sqrt{2\beta^2 \left(N + \frac{3}{2} \right) + m^2}. \quad (4)$$

N is the principal oscillator quantum number and $\beta \equiv b^{1/2}$. The corresponding wave functions are the usual nonrelativistic oscillator wave functions. For the present application, it is convenient to express the oscillator wave function in Cartesian form as

$$\Phi_N(\vec{r}) = \phi_{n_x}(x) \phi_{n_y}(y) \phi_{n_z}(z), \quad (5)$$

where

$$\phi_{n_x}(x) = \frac{\sqrt{\beta}}{\sqrt{2^n n_x!} \sqrt{\pi}} H_{n_x}(\beta x) \exp\left(-\frac{1}{2} \beta^2 x^2\right), \quad (6)$$

with similar expressions for y and z coordinates. The H_n are the Hermite polynomials. Unless noted otherwise, we use $\beta = 0.4 \text{ GeV}$, which was chosen to give reasonable values for the mass splitting and charge radius.

It should be noted that the negative energy solutions are just that since we are using a one-body wave equation and not a field theory. Therefore, these are an artifact of the model, but are necessary to provide a complete set of relativistic states.

As we choose to retain the nonrelativistic wave functions, we differ from the usual relativistic normalization: using the Klein-Gordon normalization condition for these wave functions gives

$$i \int d^3r \{ \Phi_N^{\pm*}(x) \partial^0 \Phi_N^\pm(x) - [\partial^0 \Phi_N^{\pm*}(x)] \Phi_N^\pm(x) \} = \pm 2E_N. \quad (7)$$

This leads to the factor $1/4E_0E_N$ in the response functions and to the energy factor in the current operator defined be-

low. Of course, we could have used explicitly relativistic normalized wave functions, which would have led to a different expression for the form factor, given below in Eq. (13).

The basic difference between the relativistic and nonrelativistic oscillator equations is the difference in the energy spectrum: while the nonrelativistic solutions are equally spaced, as $E_N^{nonrel} \propto N$, the relativistic spectrum goes as $E_N \propto \sqrt{N}$ for large N so the density of states increases with increasing N . We note in passing that with this relativistic spectrum our simple model gives rise to linear Regge trajectories [12] as seen in nature.

In the following, we will consider electron scattering from a meson with an infinitely heavy antiquark. In contrast to our previous publication [4], where we treated all particles as scalars, the electrons in this paper are spin-1/2 fermions, and the virtual photons have spin 1. Unless otherwise noted, we assume in this paper that only the light quark carries a charge and that the photon therefore couples only to the light quark, not to the heavy antiquark.

For a photon coupling to the quark in the positive energy harmonic oscillator ground state and transferring the four-momentum $q = (\nu, \vec{q})$, we have the following current matrix element:

$$j^\mu(q) = \frac{i}{\sqrt{4E_0E_N}} \int d^4x \exp(-iq \cdot x) \times \{ \Phi_{N_f}^*(x) \partial^\mu \Phi_0(x) - [\partial^\mu \Phi_{N_f}^*(x)] \Phi_0(x) \}, \quad (8)$$

where N_f can designate either a positive or negative energy state. Using this definition of the current along with Eq. (1), it can be easily shown that the current is conserved, $q_\mu j^\mu = 0$.

The calculation of the double-differential cross section is straightforward and leads to the Rosenbluth equation

$$\frac{d\sigma}{dE_f d\Omega_f} = \sigma_{Mott} [v_L R_L(\vec{q}, \nu) + v_T R_T(\vec{q}, \nu)], \quad (9)$$

where σ_{Mott} is the Mott cross section, v_L and v_T are the usual leptonic coefficients,

$$v_L = \frac{Q^4}{q^4}, \quad v_T = \frac{1}{2} \frac{Q^2}{q^2} + \tan^2 \frac{\vartheta_e}{2},$$

$Q^2 \equiv -q^2 = q^2 - \nu^2$, and the longitudinal and transverse response functions are

$$R_L(\vec{q}, \nu) = \sum_{N=0}^{\infty} \frac{1}{4E_0E_N} |F_{0N}(\vec{q})|^2 [(E_0 + E_N)^2 \delta(\nu + E_0 - E_N) - (E_0 - E_N)^2 \delta(\nu + E_0 + E_N)] \quad (10)$$

and

$$R_T(\vec{q}, \nu) = 8 \frac{\beta^4}{q^2} \sum_{N=0}^{\infty} \frac{1}{4E_0E_N} N |F_{0N}(\vec{q})|^2 [\delta(\nu + E_0 - E_N) - \delta(\nu + E_0 + E_N)]. \quad (11)$$

In these expressions, $F_{0,N}$ stands for the excitation form factor

$$F_{0,N}(\vec{q}^2) = \int d^3\vec{r} \exp(i\vec{q} \cdot \vec{r}) \Phi_{N_f}^*(\vec{r}) \Phi_0(\vec{r}) = \int d^3\vec{p} \Psi_{N_f}^*(\vec{p}) \Psi_0(\vec{p} - \vec{q}), \quad (12)$$

where Ψ indicates a momentum space wave function. Making use of the recurrence relations of the Hermite polynomials, we find an explicit expression for $F_{0,N}$:

$$F_{0,N}(\vec{q}^2) = \frac{1}{\sqrt{N!}} i^N \left(\frac{|\vec{q}|}{\sqrt{2}\beta} \right)^N \exp\left(-\frac{\vec{q}^2}{4\beta^2}\right). \quad (13)$$

Note that some care is necessary in writing the expressions for the responses to properly include the negative energy states. The relative sign between the positive and negative energy contributions is associated with the negative norm of the negative energy states.

These expressions for the response functions have been derived assuming that the quark is excited from the ground state into a resonance state, N , and remains there without decaying. This is just the first step on the way to meson production in this picture. The δ function in the energies is an artifact of this assumption.

Note that because we assume scalar quarks, there is no magnetization current present. The only contribution to the transverse part of the cross section comes from the convection current. As a result, the transverse response falls faster than the longitudinal response with increasing momentum transfer, as will be shown explicitly below. This is in contrast to the case of spin-1/2 quarks where the magnetization current dominates. In turn, it causes the transverse response to dominate at large momentum transfer, giving rise to the Callan-Gross equation [13] in the scaling region.

The inclusive, inelastic electron scattering cross section can be reexpressed in terms of two structure functions W_1 and W_2 , which depend only on ν and Q^2 :

$$\frac{d\sigma}{dE_f d\Omega_f} = \sigma_{Mott} \left(W_2(\nu, Q^2) + 2W_1(\nu, Q^2) \tan^2 \frac{\vartheta_e}{2} \right). \quad (14)$$

where

$$\begin{aligned}
W_1(\nu, Q^2) &= \frac{1}{2} R_T(\sqrt{Q^2 + \nu^2}, \nu) \\
W_2(\nu, Q^2) &= \frac{Q^4}{(Q^2 + \nu^2)^2} R_L(\sqrt{Q^2 + \nu^2}, \nu) \\
&\quad + \frac{Q^2}{2(Q^2 + \nu^2)} R_T(\sqrt{Q^2 + \nu^2}, \nu). \quad (15)
\end{aligned}$$

III. COULOMB SUM RULE

For the moment, we will consider a wider class of models for hadrons made up from confined quarks: namely, the more general case of models where all quarks carry an electric charge. It is interesting to consider an apparent contradiction between a model such as the one discussed here and the parton model. In our model, since all states are bound states, all of the transition form factors are coherent in that they are the result of scattering from the total charge. The parton model, however, assumes that the cross sections are composed of incoherent scattering from the individual constituents, resulting in cross sections proportional to the sum of squares of individual charges. One method of examining the transition from coherent to incoherent scattering is the Coulomb sum rule [14]. Consider the longitudinal response function

$$\begin{aligned}
R_L(\vec{q}, \nu) &= \sum_f \langle \psi_0 | \rho^\dagger(\vec{q}) | \psi_f \rangle \langle \psi_f | \rho(\vec{q}) | \psi_0 \rangle \\
&\quad \times \delta(\nu + E_0 - E_f), \quad (16)
\end{aligned}$$

where the sum represents a generalized sum over all final states, bound or free, and $\rho(\vec{q})$ is the Fourier transform of the charge operator. Now define the longitudinal sum as

$$S(\vec{q}) = \int_{-\infty}^{\infty} d\nu R_L(\vec{q}, \nu). \quad (17)$$

Using the above definition of the longitudinal response and the completeness of the final states, this becomes

$$\begin{aligned}
S(\vec{q}) &= \int_{-\infty}^{\infty} d\nu \sum_f \langle \psi_0 | \rho^\dagger(\vec{q}) | \psi_f \rangle \langle \psi_f | \rho(\vec{q}) | \psi_0 \rangle \\
&\quad \times \delta(\nu + E_0 - E_f) \\
&= \sum_f \langle \psi_0 | \rho^\dagger(\vec{q}) | \psi_f \rangle \langle \psi_f | \rho(\vec{q}) | \psi_0 \rangle \\
&= \langle \psi_0 | \rho^\dagger(\vec{q}) \rho(\vec{q}) | \psi_0 \rangle. \quad (18)
\end{aligned}$$

This is a general result. To see how this relates to the problem of the transition between coherent and incoherent scattering, consider the simple case of a nonrelativistic system of two constituents with charges e_1 and e_2 . In this case,

$$\begin{aligned}
S(\vec{q}) &= \int d^3r \psi_0^\dagger(\vec{r}) (e_1 e^{-i\vec{q}\cdot\vec{r}_1} + e_2 e^{-i\vec{q}\cdot\vec{r}_2}) \\
&\quad \times (e_1 e^{i\vec{q}\cdot\vec{r}_1} + e_2 e^{i\vec{q}\cdot\vec{r}_2}) \psi_0(\vec{r}) \\
&= \int d^3r \psi_0^\dagger(\vec{r}) (e_1^2 + e_2^2 + e_1 e_2 e^{-i\vec{q}\cdot(\vec{r}_1 - \vec{r}_2)} \\
&\quad + e_1 e_2 e^{i\vec{q}\cdot(\vec{r}_1 - \vec{r}_2)}) \psi_0(\vec{r}) \\
&= e_1^2 + e_2^2 + 2e_1 e_2 \mathcal{F}(\vec{q}), \quad (19)
\end{aligned}$$

where $\vec{r} = \vec{r}_1 - \vec{r}_2$ is the relative coordinate of the two particles and

$$\mathcal{F}(\vec{q}) = \Re \int d^3r \psi_0^\dagger(\vec{r}) e^{i\vec{q}\cdot\vec{r}} \psi_0(\vec{r}) \quad (20)$$

is the real part of the Fourier transform of the ground-state probability density.

In order to understand the physical significance of the quantity determining the rate of falloff of the mixed term containing the product of e_1 and e_2 , it is necessary to write down the most general form of the charge form factor for two quarks with charges e_1, e_2 and masses m_1, m_2 ($M = m_1 + m_2$). Here, we have dropped the δ function obtained from integrating over the c.m. motion in the second step:

$$\begin{aligned}
F_{0, N_f}(\vec{q}) &= \int d^3r \vec{\Psi}_{N_f}^*(\vec{r}) (e_1 e^{i\vec{q}\cdot\vec{r}_1} + e_2 e^{i\vec{q}\cdot\vec{r}_2}) \Psi_0(\vec{r}) \\
&\rightarrow \int d^3r \vec{\Psi}_{N_f}^*(\vec{r}) (e_1 e^{i\vec{q}\cdot\vec{r}(m_2/M)} \\
&\quad + e_2 e^{-i\vec{q}\cdot\vec{r}(m_1/M)}) \Psi_0(\vec{r}). \quad (21)
\end{aligned}$$

From this expression one sees that in the most general case, $\mathcal{F}(\vec{q})$ cannot be interpreted in terms of the ground-state charge form factor. However, in the special case of $m_1 = m_2$,

$$F_{0,0}(\vec{q}) = (e_1 + e_2) \int d^3r e^{i\vec{q}\cdot\vec{r}/2} |\Psi_0(\vec{r})|^2 = (e_1 + e_2) f_{0,0}(\vec{q}). \quad (22)$$

In this case [15],

$$\mathcal{F}(\vec{q}) = f_{0,0}(2\vec{q}). \quad (23)$$

A special case of our general result (19) was discussed in [5], where the case of two scalar, equal mass quarks in a nonrelativistic harmonic oscillator potential was considered.

In the model we present in this paper, all of the charge is carried by one of the constituents: that is, $e_1 = 1$ and $e_2 = 0$. Since there is only a single charge, there is no difference between coherent and incoherent scattering, so we expect that

$$S(\vec{q}) = 1. \quad (24)$$

This then provides a useful test of the model.

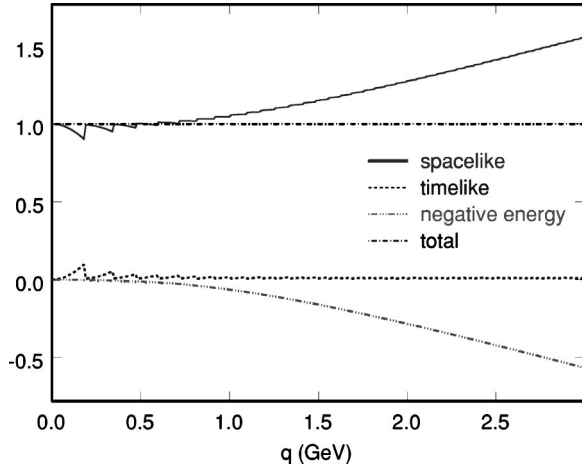


FIG. 1. Positive energy spacelike, timelike, and negative energy contributions to the Coulomb sum rule from our model.

Using Eq. (10),

$$\begin{aligned}
 S(\vec{q}) &= \int_{-\infty}^{\infty} d\nu \sum_{N=0}^{\infty} \frac{1}{4E_0E_N} |F_{0N}(\vec{q})|^2 [(E_0 + E_N)^2 \delta(\nu + E_0 \\
 &\quad - E_N) - (E_0 - E_N)^2 \delta(\nu + E_0 + E_N)] \\
 &= \sum_{N=0}^{\infty} \frac{(E_0 + E_N)^2 - (E_0 - E_N)^2}{4E_0E_N} |F_{0N}(\vec{q})|^2 \\
 &= \sum_{N=0}^{\infty} |F_{0N}(\vec{q})|^2. \tag{25}
 \end{aligned}$$

Using Eq. (13) it is straightforward to demonstrate that $S(\vec{q})=1$ in this case so the Coulomb sum rule is satisfied. Indeed, this will be true regardless of the form of the confining potential, as long as one considers a complete set of solutions. Note that for this model it is necessary that the integral in Eq. (17) be over both positive and negative energy transfers for the sum rule to be satisfied.

In electron scattering, only the spacelike region is accessible and the negative energy states are an artifact of the use of the Klein-Gordon equation as a wave equation. It is useful, therefore, to examine the contributions to the Coulomb sum from spacelike, timelike, and negative energy states. When referring to the spacelike and timelike contributions, only positive energy states are included. The different contributions are shown in Fig. 1. At $|\vec{q}|=0$, only the elastic form factor can contribute and therefore must saturate the sum rule. As the momentum transfer increases the elastic form factor falls off, resulting in a decrease in the spacelike contribution. It decreases until the momentum transfer increases to a point where the first excited state enters the spacelike region. The spacelike contribution then saturates the sum rule again. This process continues as new form factors become accessible in the spacelike region. The result is a sawtoothed behavior of the spacelike contribution. Because the density of states for this oscillator model increases with increasing energy, the magnitude of the “teeth” becomes

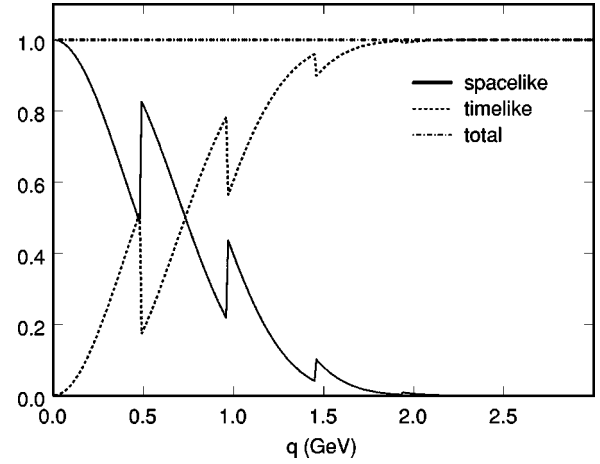


FIG. 2. Spacelike and timelike contributions to the Coulomb sum rule for a nonrelativistic oscillator.

smaller with increasing momentum transfer until the spacelike contribution is essentially smooth. Note that the spacelike contribution oversaturates the sum rule around momentum transfers of 0.5 GeV and above. Since the jaggedness in the spacelike region is associated with the migration of contributions from the timelike region, it is not surprising to see that the complement of this behavior does indeed show up in the timelike contribution. As the momentum transfer increases, the size of the timelike contribution becomes small. The negative energy contribution is smooth and compensates for the oversaturation of the spacelike contribution. This is clearly an artifact of using a one-body wave equation with negative energy contributions.

It should be pointed out here that the small size of the timelike contributions is an essential consequence of using a relativistic wave equation. This can be seen by considering a similar situation where the nonrelativistic oscillator model is used to describe the system. The spacelike and timelike contributions of such a model are shown in Fig. 2. The Schrödinger equation of course has only positive energy solutions. Here the sawtoothed behavior in the spacelike contributions is similar to the relativistic case with the important difference that due to the linear character of the nonrelativistic spectrum (see Fig. 13) the contributions from states entering the spacelike region is not rapid enough to compensate for the falloff in the form factors. Therefore, with increasing momentum transfer, the size of the spacelike contribution approaches zero with all of the strength appearing in the timelike region.

Duality in the sum of the form factors in the spacelike region is related to the Coulomb sum rule. While the concept of duality in the form factors is not clearly related to an observable (in contrast to duality in the structure functions), it still has received attention in the literature [5]. We describe it within our model in Appendix A.

This model can be easily extended to the case where both quarks are charged by examining the behavior of the two-body Gross equation in the limit where the mass of one of the particles becomes infinite [16]. The contribution of the infinite mass particle to the structure function is simple and straightforward. As a result of its infinite mass, this particle remains stationary for any finite momentum transfer, and it is

pointlike. This particle therefore contributes only to elastic scattering and has a constant form factor. The structure function can then be written as

$$R_L(\vec{q}, \nu) = |e_1 F_{00}(\vec{q}) + e_2|^2 + e_1^2 \sum_{N=1}^{\infty} \frac{1}{4E_0 E_N} \\ \times |F_{0N}(\vec{q})|^2 [(E_0 + E_N)^2 \delta(\nu + E_0 - E_N) \\ - (E_0 - E_N)^2 \delta(\nu + E_0 + E_N)]. \quad (26)$$

The Coulomb sum can be easily calculated to be

$$S(\vec{q}) = e_1^2 \sum_{N=0}^{\infty} |F_{0N}(\vec{q})|^2 + e_2^2 + 2e_1 e_2 \Re F_{00}(\vec{q}) \\ = e_1^2 + e_2^2 + 2e_1 e_2 \Re F_{00}(\vec{q}). \quad (27)$$

So in this case,

$$\mathcal{F}(\vec{q}) = \Re F_{00}(\vec{q}) = F_{00}(\vec{q}). \quad (28)$$

After examining how the apparent contradiction between coherent and incoherent scattering is resolved in a more general framework, we now proceed to investigate duality in our model. The first condition for duality is that one obtains scaling in the structure function calculated solely with resonances and that the scaling curve thus obtained agrees with the scaling curve obtained in the parton model.

IV. PARTON MODEL

The usual assumption of the parton model is that at large momentum transfers the final state quarks can be treated as though they were free. Examination of the structure functions for large Q^2 and fixed Bjorken x leads to identification of the scaling functions.

For our simple model, the response functions for excitation of a bound (off-mass-shell) quark to a plane-wave final state can be calculated analytically as

$$R_L(\vec{q}, \nu) = \frac{(2E_0 + \nu)^2}{4\pi^{1/2}\beta E_0 |\vec{q}|} \left[\exp\left(-\frac{y^2}{\beta^2}\right) \right. \\ \left. - \exp\left(-\frac{(y+2|\vec{q}|)^2}{\beta^2}\right) \right] [\theta(\nu + E_0 - m) \\ - \theta(-\nu - E_0 - m)] \quad (29)$$

and

$$R_T(\vec{q}, \nu) = \frac{\beta}{2\pi^{1/2}E_0 |\vec{q}|^2} \left[\left(2(y+|\vec{q}|) - \frac{\beta^2}{|\vec{q}|} \right) \exp\left(-\frac{y^2}{\beta^2}\right) \right. \\ \left. - \left(2(y+|\vec{q}|) + \frac{\beta^2}{|\vec{q}|} \right) \exp\left(-\frac{(y+2|\vec{q}|)^2}{\beta^2}\right) \right] \\ \times [\theta(\nu + E_0 - m) - \theta(-\nu - E_0 - m)], \quad (30)$$

where $y = \sqrt{(\nu + E_0)^2 - m^2} - |\vec{q}|$.

For our model it is not possible to define the scaling variable in terms of the target mass since it is infinite. For this reason we define a new Bjorken variable

$$u_{Bj} = \frac{M}{m} x_{Bj} = \frac{Q^2}{2m\nu}, \quad (31)$$

which covers the interval $-\infty < u_{Bj} < \infty$. Using $\nu = Q^2/2mu_{Bj}$ and taking the limit $Q^2 \rightarrow \infty$, the structure functions become

$$R_L(Q^2, u_{Bj}) \rightarrow \frac{Q^2}{8\pi^{1/2}\beta m E_0 u_{Bj}} \exp\left(-\frac{(E_0 - mu_{Bj})^2}{\beta^2}\right) \\ \times \left[\theta\left(\frac{Q^2}{2mu_{Bj}}\right) - \theta\left(-\frac{Q^2}{2mu_{Bj}}\right) \right] \quad (32)$$

and

$$R_T(Q^2, u_{Bj}) \rightarrow \frac{2\beta mu_{Bj}}{\pi^{1/2}E_0 Q^2} \exp\left(-\frac{(E_0 - mu_{Bj})^2}{\beta^2}\right) \\ \times \left[\theta\left(\frac{Q^2}{2mu_{Bj}}\right) - \theta\left(-\frac{Q^2}{2mu_{Bj}}\right) \right]. \quad (33)$$

Since in this limit

$$|\nu| \frac{Q^4}{\vec{q}^4} \rightarrow \frac{8m^3 |u_{Bj}|^3}{Q^2} \quad \text{and} \quad |\nu| \frac{Q^2}{2\vec{q}^2} \rightarrow m |u_{Bj}|, \quad (34)$$

the structure functions have the limits

$$W_1(u_{Bj}, Q^2) \rightarrow F_1(u_{Bj}) = 0 \quad (35)$$

and

$$|\nu| W_2(\nu, Q^2) \rightarrow F_2(u_{Bj}) = \frac{m^2 u_{Bj}^2}{\pi^{1/2}\beta E_0} \exp\left(-\frac{(E_0 - mu_{Bj})^2}{\beta^2}\right). \quad (36)$$

Note that the choice of $|\nu|$ in defining W_2 is necessary to provide a properly normalized scaling function as will be seen below.

Although we have used the Bjorken scaling variable to obtain these results, this will be true for all such variables since all acceptable scaling variables must reduce to the

Bjorken scaling variable as $Q^2 \rightarrow \infty$. Therefore, a more general expression for F_2 for any scaling variable and any initial state can be written as

$$F_2(u) = \frac{m^2 u^2}{4\pi^2 E_0} \int_{|E_0 - mu|}^{\infty} dp p N(p), \quad (37)$$

where $N(p)$ is the ground-state momentum distribution normalized such that

$$\frac{1}{2\pi^2} \int_0^{\infty} dp p^2 N(p) = 1. \quad (38)$$

After obtaining the scaling curve in the parton model, i.e., the scaling curve for a quark initially bound and then free, we proceed to find an expression for the scaling curve in our model, where the quark makes the transition from the ground state to an excited bound state.

V. CONTINUUM LIMIT

An interesting feature of our relativistic oscillator model is that the scaling behavior of the model can be determined analytically by making a continuum approximation. The justification for this is that at increasing momentum transfer the contributions to the response functions are dominated by higher-energy states. Since the density of states increases with increasing energy, it is reasonable that a continuum approximation should provide a good description of the averaged response for large momentum transfers.

Using Eqs. (10) and (13) we can write

$$R_L(\vec{q}, \nu) = \sum_{N=0}^{\infty} \Delta N \frac{1}{4E_0 E_N} \frac{1}{N!} \left(\frac{\vec{q}^2}{2\beta^2} \right)^N \exp\left(-\frac{\vec{q}^2}{2\beta^2}\right) \times [(E_N + E_0)^2 \delta(E_N - E_0 - \nu) - (E_0 - E_N)^2 \delta(E_N + E_0 + \nu)], \quad (39)$$

where $\Delta N = 1$. It is convenient to write

$$E_N = \sqrt{p_N^2 + E_0^2}, \quad (40)$$

where

$$p_N^2 = 2\beta^2 N. \quad (41)$$

From this it can be determined that, for a variation in N ,

$$\Delta N = \frac{p_N \Delta p_N}{\beta^2} \quad (42)$$

and

$$\Delta E_N = \frac{p_N \Delta p_N}{E_N}. \quad (43)$$

The longitudinal response function can then be rewritten as

$$R_L(\vec{q}, \nu) = \frac{1}{4\beta^2 E_0} \sum_{N=0}^{\infty} \frac{\Delta E_N}{\Gamma\left(1 + \frac{E_N^2 - E_0^2}{2\beta^2}\right)} \times \left(\frac{\vec{q}^2}{2\beta^2}\right)^{(E_N^2 - E_0^2)/2\beta^2} \exp\left(-\frac{\vec{q}^2}{2\beta^2}\right) \times [(E_N + E_0)^2 \delta(E_N - E_0 - \nu) - (E_0 - E_N)^2 \delta(E_N + E_0 + \nu)]. \quad (44)$$

This sum can now be approximated by the integral

$$R_L(\vec{q}, \nu) = \frac{1}{4\beta^2 E_0} \int_{E_0}^{\infty} \frac{dE}{\Gamma\left(1 + \frac{E^2 - E_0^2}{2\beta^2}\right)} \left(\frac{\vec{q}^2}{2\beta^2}\right)^{(E^2 - E_0^2)/2\beta^2} \times \exp\left(-\frac{\vec{q}^2}{2\beta^2}\right) [(E + E_0)^2 \delta(E - E_0 - \nu) - (E - E_0)^2 \delta(E + E_0 + \nu)], \quad (45)$$

which can be trivially evaluated to give

$$R_L(\vec{q}, \nu) = \frac{(\nu + 2E_0)^2 [\theta(\nu) - \theta(-2E_0 - \nu)]}{4\beta^2 E_0 \Gamma\left(1 + \frac{(\nu + E_0)^2 - E_0^2}{2\beta^2}\right)} \times \left(\frac{\vec{q}^2}{2\beta^2}\right)^{[(\nu + E_0)^2 - E_0^2]/2\beta^2} \exp\left(-\frac{\vec{q}^2}{2\beta^2}\right). \quad (46)$$

Similarly, the transverse response becomes

$$R_T(\vec{q}, \nu) = \frac{(\nu^2 + 2E_0\nu) [\theta(\nu) - \theta(-2E_0 - \nu)]}{E_0 \vec{q}^2 \Gamma\left(1 + \frac{(\nu + E_0)^2 - E_0^2}{2\beta^2}\right)} \times \left(\frac{\vec{q}^2}{2\beta^2}\right)^{[(\nu + E_0)^2 - E_0^2]/2\beta^2} \exp\left(-\frac{\vec{q}^2}{2\beta^2}\right). \quad (47)$$

In the scaling limit the argument of the Γ function becomes large, so Stirling's formula can be used to write the longitudinal response function as

$$R_L(\sqrt{Q^2 + \nu^2}, \nu) = \frac{(\nu + 2E_0)^2}{4\pi^{1/2}\beta E_0} \frac{\exp\left[\frac{\nu^2 + 2E_0\nu}{2\beta^2} \ln\left(\frac{Q^2 + \nu^2}{\nu^2 + 2E_0\nu}\right) - \frac{Q^2 + \nu^2}{2\beta^2} + \frac{\nu^2 + 2E_0\nu}{2\beta^2}\right]}{\sqrt{\nu^2 + 2E_0\nu}} [\theta(\nu) - \theta(-2E_0 - \nu)]. \quad (48)$$

Using

$$\nu = \frac{Q^2}{2mu_{Bj}} \quad (49)$$

and taking the limit $Q^2 \rightarrow \infty$, the structure functions become

$$R_L(Q^2, u_{Bj}) \rightarrow \frac{Q^2}{8\pi^{1/2}\beta m E_0 u_{Bj}} \exp\left(-\frac{(E_0 - mu_{Bj})^2}{\beta^2}\right) \times \left[\theta\left(\frac{Q^2}{2mu_{Bj}}\right) - \theta\left(-\frac{Q^2}{2mu_{Bj}}\right)\right]. \quad (50)$$

Similarly,

$$R_T(Q^2, u_{Bj}) \rightarrow \frac{2\beta mu_{Bj}}{\pi^{1/2} E_0 Q^2} \exp\left(-\frac{(E_0 - mu_{Bj})^2}{\beta^2}\right) \times \left[\theta\left(\frac{Q^2}{2mu_{Bj}}\right) - \theta\left(-\frac{Q^2}{2mu_{Bj}}\right)\right]. \quad (51)$$

Since Eqs. (50) and (51) are identical to Eqs. (32) and (33), the model scales to the parton model result. So even though our model describes a bound quark being excited to resonance states, we do obtain a scaling curve in the Bjorken limit. This as well as the results presented in [17–19,6] shows that scaling does not necessarily imply scattering off free constituents, a belief which is encountered widely.

Others [17–19] have studied the transition from ground state to excited bound states and found scaling in similar

models. However, duality is satisfied only when (i) the transition from ground state to excited bound states scales, (ii) the transition from ground state to a plane-wave final state scales, *and* (iii) both scaling curves coincide. In this section and the preceding section, we have shown that duality holds explicitly in our model. The numerical approach towards the scaling curve is shown in Fig. 3.

Note that as we explicitly made use of Stirling's formula in the derivation of the scaling function in the continuum limit, it is clear that for the lower-lying resonances, which correspond to lower N values, we will never quite see scaling in the subasymptotic regime. This is of no practical relevance, as these resonances are pushed out to very high values of u for larger Q^2 , and the structure functions practically vanish in this region. This is completely analogous to the fact that for electron scattering from a proton, one always picks up the elastic scattering at $x_{Bj} = 1$, independent of Q^2 —even though the elastic form factor will have fallen off to negligible values at high enough Q^2 .

VI. APPROACH TO SCALING IN THE STRUCTURE FUNCTIONS

After establishing analytically that one of the necessary conditions for duality is satisfied—namely, scaling to the scaling curve obtained from a free quark in the final state—we proceed to investigate the approach to scaling numerically.

We would like to remind the reader that our results should not be compared to the available nucleon data—our model

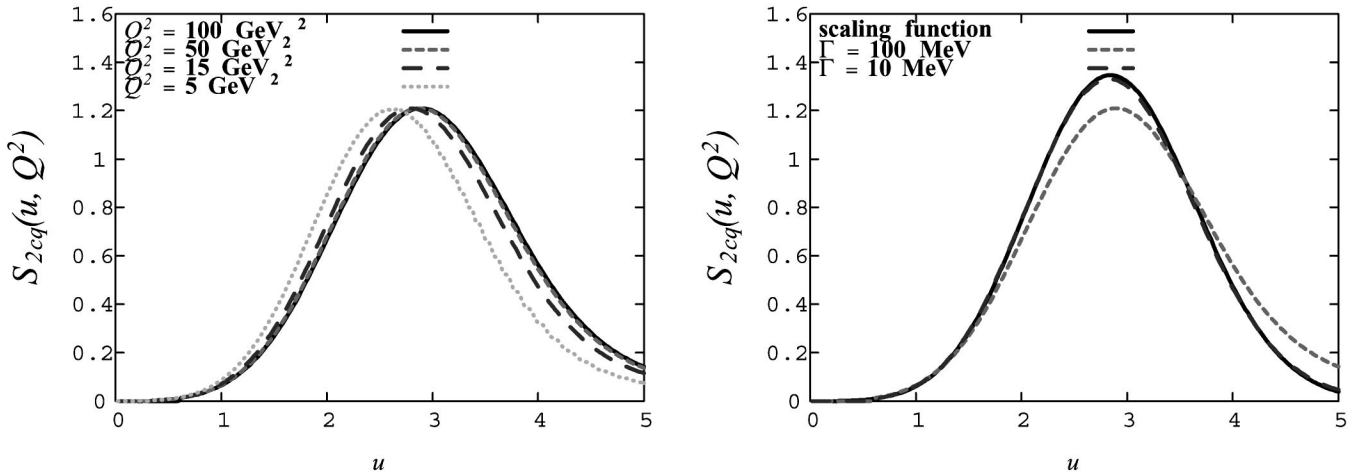


FIG. 3. The high-energy scaling behavior of S_{2cq} as a function of u for various values of Q^2 . In the left panel, we have used $\Gamma = 100$ MeV to give the impression of real resonances even though this large value distorts the scaling curve somewhat. For any width equal to or smaller than this, the distortion is rather innocuous, and for $\Gamma \rightarrow 0$, the structure function approaches the scaling function (solid line) in Eq. (36), as shown in the right-hand panel. The structure functions in the left panel are shown for $Q^2 = 100$ GeV².

calculations describe a process that might resemble electron scattering from a B meson, which has never been measured. In general, when we consider scattering from a meson target, scaling will set in later than for a baryon target: momentum sharing for higher momenta between fewer constituents is easier, which leads to a slower falloff of the individual form factors and to a later onset of scaling. In our situation, where the system is not allowed to decay, we have a somewhat extreme case.

To see duality clearly both experimentally and theoretically, one needs to go beyond the Bjorken scaling variable x_{Bj} and the scaling function $\mathcal{S}_{2Bj} = \nu W_2$ that goes with it. This is because in deriving Bjorken's variable and scaling function, one not only assumes Q^2 to be larger than any mass scale in the problem, but also that high Q^2 (PQCD) dynamics controls the interactions. However, duality has its onset in the region of low to moderate Q^2 , and there masses and violations of asymptotic freedom do play a role. Bloom and Gilman used a new, *ad hoc* scaling variable ω' [1] in an attempt to deal with this fact. In most contemporary data analyses, the Nachtmann variable [20,21] is used together with \mathcal{S}_{2Bj} . Nachtmann's variable contains the target mass as a scale, but neglects quark masses. For our model, the constituent quark mass (assumed to arise as a result of spontaneous chiral symmetry breaking) is vital at low energy, and a scaling variable that does not make any assumptions about the size of the quark and target masses compared to Q^2 is desirable. Such a variable was derived more than 20 years ago by Barbieri *et al.* [22] to take into account the masses of heavy quarks; we use it here given that after spontaneous chiral symmetry breaking the nearly massless light quarks have become massive constituent quarks, calling it x_{cq} :

$$x_{cq} = \begin{cases} \frac{1}{2M} (\sqrt{\nu^2 + Q^2} - \nu) \left(1 + \sqrt{1 + \frac{4m^2}{Q^2}} \right) & \text{for } \nu > 0, \\ -\frac{1}{2M} (\sqrt{\nu^2 + Q^2} + \nu) \left(1 + \sqrt{1 + \frac{4m^2}{Q^2}} \right) & \text{for } \nu < 0, \end{cases} \quad (52)$$

where the definition for negative energy is chosen such that it satisfies the kinematic constraints in this region and reproduces the behavior of x_{Bj} for large Q^2 . The scaling function associated with this variable is given by

$$\mathcal{S}_{2cq} \equiv |\vec{q}| W_2 = \sqrt{\nu^2 + Q^2} W_2. \quad (53)$$

This scaling function and variable were derived for scalar quarks which are free, but have a momentum distribution. The derivation of a new scaling variable and function for bound quarks will be published elsewhere. Numerically, this scaling variable does not differ very much from the one in Eq. (52). Of course all versions of the scaling variable must converge to x_{Bj} and all versions of the scaling function must converge towards \mathcal{S}_{Bj} for large enough Q^2 . One can also easily verify that in the limit $m \rightarrow 0$ one obtains from Eq.

(52) the Nachtmann scaling variable. In the following, we use the variable x_{cq} and the scaling function \mathcal{S}_{2cq} .

We are now ready to look at scaling and duality in our model. Since the target has mass $M \rightarrow \infty$, it is convenient to rescale the scaling variable x_{cq} by a factor M/m :

$$u \equiv \frac{M}{m} x_{cq}. \quad (54)$$

The variable u takes values from 0 to a maximal, Q^2 -dependent value, which can go to infinity. The high-energy scaling behavior of the appropriately rescaled structure function \mathcal{S}_{cq} is illustrated in Fig. 3. The structure function has been evaluated using the phenomenologically reasonable parameters $m = 0.33$ GeV and $\beta = 0.4$ GeV. To display it in a visually meaningful manner, the energy-dependent δ function has been smoothed out by introducing an unphysical Breit-Wigner shape with an arbitrary but small width Γ , chosen for purposes of illustration:

$$\delta(E_N - E_0 - \nu) \rightarrow \frac{\Gamma}{2\pi} \frac{f}{(E_N - E_0 - \nu)^2 + (\Gamma/2)^2}, \quad (55)$$

where the factor

$$f = \pi \left/ \left[\frac{\pi}{2} + \arctan \frac{2(E_N - E_0)}{\Gamma} \right] \right.$$

ensures that the integral over the δ function is identical to that over the Breit-Wigner shape. As for the all scalar case discussed in [4], the smearing out of the δ function in energy with the Breit-Wigner shape leads to a slight widening of the curve and flattening of the peak height. However, when choosing a smaller value for the Breit-Wigner width, these effects disappear, as seen in the right panel of Fig. 3.

Not unexpectedly, the scaling curve we find here, when using photons and electrons with their appropriate spins, differs from the one in the all scalar case both in its final shape and the approach to scaling. Now, we have two terms the longitudinal and transverse response functions—contributing to the structure function W_2 and therefore to the scaling curve. More importantly, the terms themselves are different and more complicated in the case considered here. The longitudinal part of the structure function contains an additional factor $(E_0 + E_N)^2$, which was not present in the all scalar case. As shown in Sec. V, the transverse response vanishes like $1/Q^2$ in the limit $Q^2 \rightarrow \infty$, while the longitudinal response grows like Q^2 . This leads to a vanishing of W_1 and to a Q^2 -independent value for $|\vec{q}| W_2$, even though, from Eqs. (15),(50),(51), it is clear that at lower Q^2 the transverse contribution to W_2 will not vanish immediately, therefore making the approach to scaling slower. The effect is rather significant, though, as the contributions of the convection current are very small. They delay scaling slightly for the low- u part of the curve. The effect would be more important for contributions of similar size within a certain kinematic range. In our case, the main effect of the transverse contribution is to slightly broaden the curve. For smaller Q^2 , this effect is more pronounced for low values of the scaling vari-

able u , as the higher u correspond to lower-lying resonances, which have only tiny contributions from the transverse part.

As already mentioned above, for a proton target, the dominant contribution to the transverse response and overall is the magnetization current, which does not contribute for our scalar “quarks.” Note that both the transverse and longitudinal contributions to W_2 are positive definite. If a dominant contribution in the transverse response is present, it should lead to a different scaling behavior in the structure function W_1 than in νW_2 . For νW_2 , the longitudinal term with different Q^2 behavior will most likely slow the approach to scaling down, as it is going to be of comparable size to the magnetization current contribution at low Q^2 . This is a completely general observation, and one would expect to see faster scaling in F_1 once the data are available. The same conclusion was reached on a different basis in [5].

The shape of the scaling curve is also different than for the all scalar case. The peak is higher, the curve extends to larger values of the scaling variable, and for $u \rightarrow 0$, the scaling function actually vanishes now, as expected from a valence quark distribution, even though we do not find a behavior $\propto \sqrt{u}$ as seen for proton targets. However, we cannot expect to reproduce the correct distribution function for quarks in our simple model with scalar “quarks.”

From the explicit expression for the scaling curve, it is clear that it depends both on the binding strength β of the harmonic oscillator and on the quark mass. It peaks at $u_{peak} = (E_0 + \sqrt{E_0^2 + \beta^2})/2m$, slightly above the value $u_{peak}^{scalar} = E_0/m$ which we found for the all scalar case. Naively, for a target of mass M made up of noninteracting quarks of mass m_q , one expects a spike at $x_{Bj} = m_q/M$. In our case, the role of the mass of the quark m_q is played by the ground-state energy E_0 , which appears everywhere (e.g., in flux factors, normalization) where one would have the mass in the free particle case. As our variable is rescaled with the factor M/m , we expect $u_{peak} \approx E_0/m$. It is interesting to note that this value receives a slight binding correction due to the conserved current employed here. We note that for weaker binding, the peak gets narrower and its position slides towards $u \approx 1$. As expected, in the limit of a free particle, $\beta \rightarrow 0$, we do obtain a spike of infinite height at $u = 1$ [in this limit, the scaling function becomes $\delta(u)/(mE_0)$].

VII. MOMENTS AND FURTHER SUM RULES

Now, we will discuss global duality, where the term global implies that we consider an average and integral over many resonances, which is compared with the corresponding integral over the scaling curve.

Local duality implies that we compare the contribution of one single resonance or just a few resonances with the scaling result, i.e., with the free quark result. This will be discussed in Sec. VIII. The concept of local duality is taken to its extreme when one focuses not just on one single resonance, but on one point only of the contribution of the single resonance, as was done first by Bloom and Gilman [1], when they compared the peak value of a single resonance with the value of the scaling curve. This version of duality has been investigated in Ref. [23]. As this ratio would depend strongly

on the Breit-Wigner width we use to smooth out the δ functions, it is not appropriate to consider it in this paper.

Global duality was first quantified by Bloom and Gilman [1] in the form of finite energy sum rules, where the integral over the scaling curve was compared to the integral over the resonance contribution. The integration range in both cases comprises the region of the scaling variable ω' or ν , respectively, which corresponds to the resonance region, defined as having an invariant mass $W < 2$ GeV:

$$\frac{2M_N}{Q^2} \int_0^{\nu_m} d\nu \nu W_2(\nu, Q^2) = \int_1^{1+W_m^2/Q^2} d\omega' \nu W_2(\omega'), \quad (56)$$

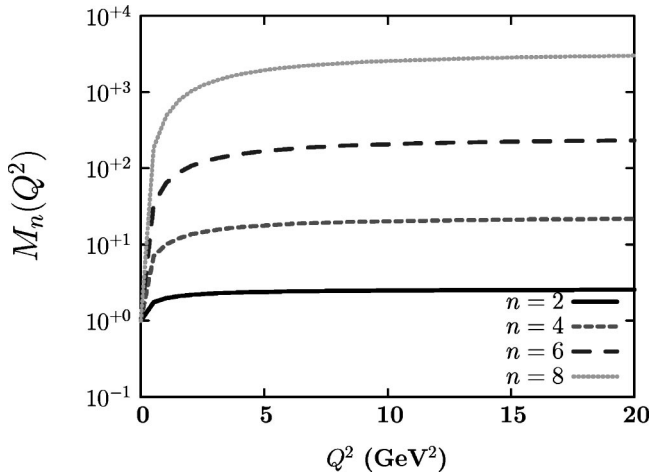
where $W_m \approx 2$ GeV and $\nu_m = (W_m^2 - M_N^2 + Q^2)/2M_N$. Here, M_N denotes the mass of the nucleon target. The agreement between the left- and right-hand sides of this equation is better than 10%; for the larger values of Q^2 , starting around $Q^2 \approx 2$ GeV², the agreement is quite impressive: 2% or better.

While it certainly would be desirable to calculate the same finite energy sum rule in our model, there is a practical problem and a philosophical problem. First, in our model, we deal with an infinitely heavy system, so that, in principle, the invariant mass W of the final state is always infinity. Even if we could define a reasonable substitute for the invariant mass of the final state, picking an integration limit is a problem in principle: for our model, the scaling curve consists solely of resonance contributions, even though they cannot be resolved and form a smooth curve. So any distinction between “resonance region” and “continuum” or “multiparticle final states” is artificial. The conventional definition of “resonance region” as the region where $W < 2$ GeV means the region where the resonances are prominent and dominant. However, it does not mean that for $W > 2$ GeV there are no resonances present, and it also does not mean that for $W < 2$ GeV, i.e., in the resonance region, there are no nonresonant contributions at all. Experimentally, there is background for $W < 2$ GeV, and there are resonances for $W > 2$ GeV, e.g., several N^* and Δ resonances. Also, from a theoretical point of view, it is obvious that resonances which decay by creation of a quark-antiquark pair in the final state must be accompanied by a corresponding nonresonant production mechanism, where pair creation takes place in the initial state and the photon interacts with the preformed meson.

Since the distinction between a “resonance region” and a “continuum region” has its problems, we utilize the moments of the scaling function S_{cq} , where the integration range comprises the whole interval of the scaling variable. The physical information contained in the finite energy sum rule and the moments M_n is the same. The moments of a scaling function $\mathcal{S}(Q^2, x)$ with a scaling variable x are defined as

$$M_n(Q^2) = \int_0^{x_{max}} dx x^{n-2} \mathcal{S}_2(Q^2, x). \quad (57)$$

Here, in contrast to [24], we do not include the unphysical region $[x_{max}, 1]$ in the integration interval. It is obvious that


 FIG. 4. The lowest moments M_n as a function of Q^2 .

higher moments, i.e., $n=4,6,\dots$, tend to emphasize the resonance region; as for fixed Q^2 , the resonances are found at large x . The values of the moments decrease with increasing n . In our case, we change to u -type scaling variables [see Eq. (54)] so that

$$M_n^{u\text{-based}}(Q^2) = \int_0^{u_{\max}} du u^{n-2} S_2(Q^2, u), \quad (58)$$

where u_{\max} corresponds to the maximum value of u which is kinematically accessible at a given Q^2 . By changing from x to u scaling variables, we change the upper integration limit from a value equal to or lower than 1 to a value considerably larger than 1 for $Q^2 > 1-2 \text{ GeV}^2$. This means that our higher moments will emphasize the low-lying resonances even more than the conventional, x -based moments. Also, the higher moments will be larger than the moments with small n .

Evaluating the moments of the structure function (53) explicitly one has

$$\begin{aligned} M_n(Q^2) &= \left(\frac{r}{2m}\right)^{n-1} \sum_{N=0}^{\infty} (\sqrt{\nu_N^2 + Q^2} - \nu_N)^{n-1} \\ &\quad \times \frac{E_0}{E_N} |F_{0N}(\sqrt{\nu_N^2 + Q^2})|^2 \\ &\quad \times \left[\frac{Q^4}{q_N^4} (E_0 + E_N)^2 + 4N\alpha \frac{Q^2}{q_N} \right], \end{aligned} \quad (59)$$

with $r \equiv 1 + \sqrt{1 + 4m^2/Q^2}$, and $\nu_N = E_N - E_0$ and $q_N = \sqrt{Q^2 + \nu_N^2}$.

The elastic contribution is given by

$$M_n^{\text{elastic}}(Q^2) = \left(\frac{r}{2m}\right)^{n-1} |Q|^{n-1} \exp\left(-\frac{Q^2}{2\beta^2}\right). \quad (60)$$

Note that for vanishing four-momentum transfer Q^2 , all moments take the value 1, independent of n .

In Fig. 4, we show the moments for $n=2,4,6,8$, which

were obtained by integrating over the positive energy states only. One can see clearly that all moments flatten out, even though they did not quite reach their asymptotic value at the highest Q^2 value shown. The lowest moment M_n , is within 9% of its asymptotic value at $Q^2=5 \text{ GeV}^2$ and within 2% of its asymptotic value at $Q^2=20 \text{ GeV}^2$. As expected, the higher moments, which by construction get more contributions from the lower-lying resonances, need higher Q^2 values in order to reach their asymptotic values. For M_6 , we find that it has reached 64% of its asymptotic value at $Q^2=5 \text{ GeV}^2$ and 88% of its asymptotic value at $Q^2=20 \text{ GeV}^2$. From these numbers, we can see that even though scaling does not set in for $Q^2 < 50 \text{ GeV}^2$, the asymptotic values at least of the lower moments are reached much earlier. This reflects the fact that $S_{c_q}(u, Q^2)$ approaches the scaling curve by shifting towards higher u , not by approaching it from below or above.

Since the continuum approximation provides a relatively simple analytic expression for the structure functions, it is possible to use this to study certain properties of the moments. First, however, it is necessary to determine the validity of this approximation for the calculation of moments. Figure 5 shows calculations of the first three moments M_0 , M_1 , and M_2 . In each panel the solid line represents the exact calculation according to Eq. (59). The dash-dotted curve is a calculation of the continuum approximation with both longitudinal and transverse contributions, while the dashed curve includes only the longitudinal contribution. Note that the continuum approximation works very well down to a couple of GeV^2 . Note also that while the inclusion of the transverse contribution slows convergence to the asymptotic value for M_0 it improves convergence for the higher moments.

The continuum approximation can then be used to obtain an expansion of the moments in powers of $1/Q^2$ reminiscent of the operator product expansion (OPE) series:

$$M_n^{2k}(Q^2) = \sum_{i=0}^k \frac{c_n^{2i}}{Q^{2i}}. \quad (61)$$

Note that we do not have any gluons in our model and, thus, no radiative corrections. The expansion coefficients c_n^{2i} correspond to the nonperturbative matrix elements of higher-twist operators in the OPE. Since this is an asymptotic series, the expansion will fail at low Q^2 with the point at which the series diverges being dependent upon the order of the series. The expansion coefficients for the five lowest moments are shown in Table I. Contributions to the coefficients from transverse and longitudinal responses are shown along with the total. The obvious feature of these coefficients is that they are not in general small, nor do they show any obvious convergence. The reason for this can be seen from Fig. 6. Here the exact result is shown as a solid line and is compared to the expansion with from one to four terms for both M_0 and M_4 . Since the moments all have a finite value at $Q^2=0$, the function cannot be analytic in $1/Q^2$. Any expansion in this variable to a finite number of terms will at some point diverge, either above or below the correct result. Using an

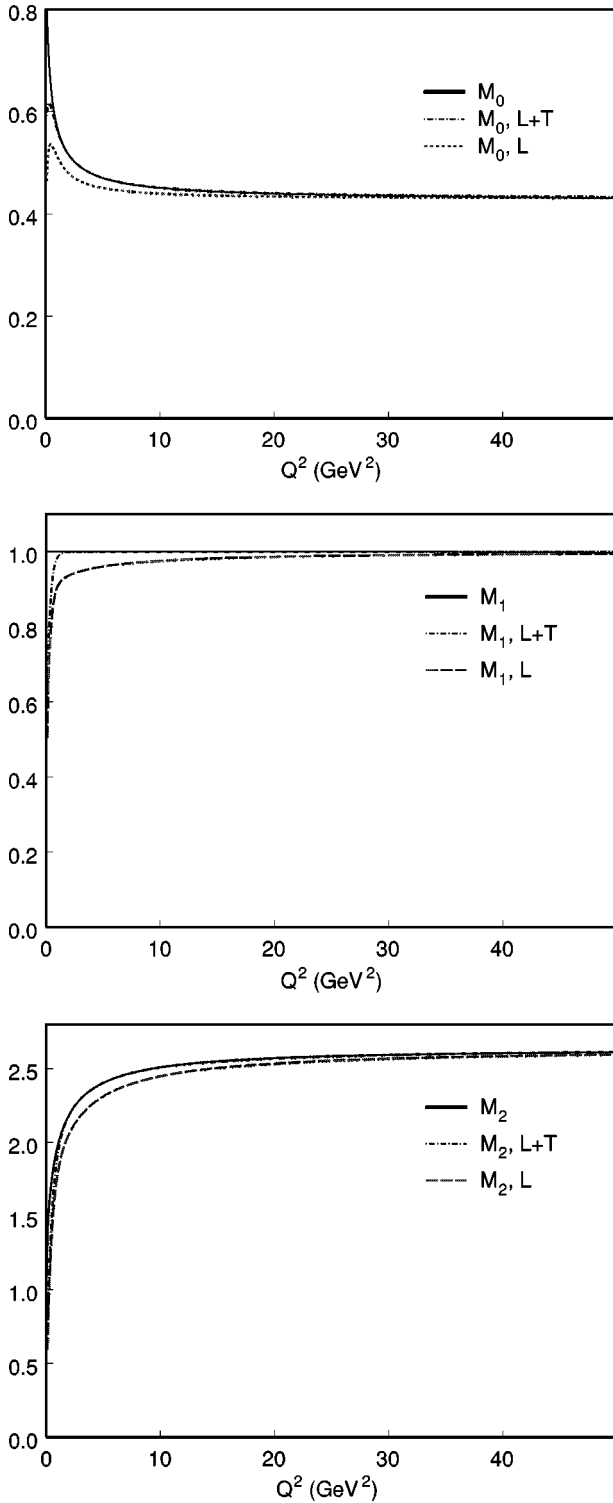


FIG. 5. The first three moments, M_0 , M_1 , and M_2 . The solid line shows the exact result calculated according to Eq. (59). The dash-dotted lines show the moment calculated in the continuum approximation; the dotted curves show the purely longitudinal contribution to the moment in the continuum approximation. Inclusion of the transverse contributions improves convergence of the moments to their asymptotic values.

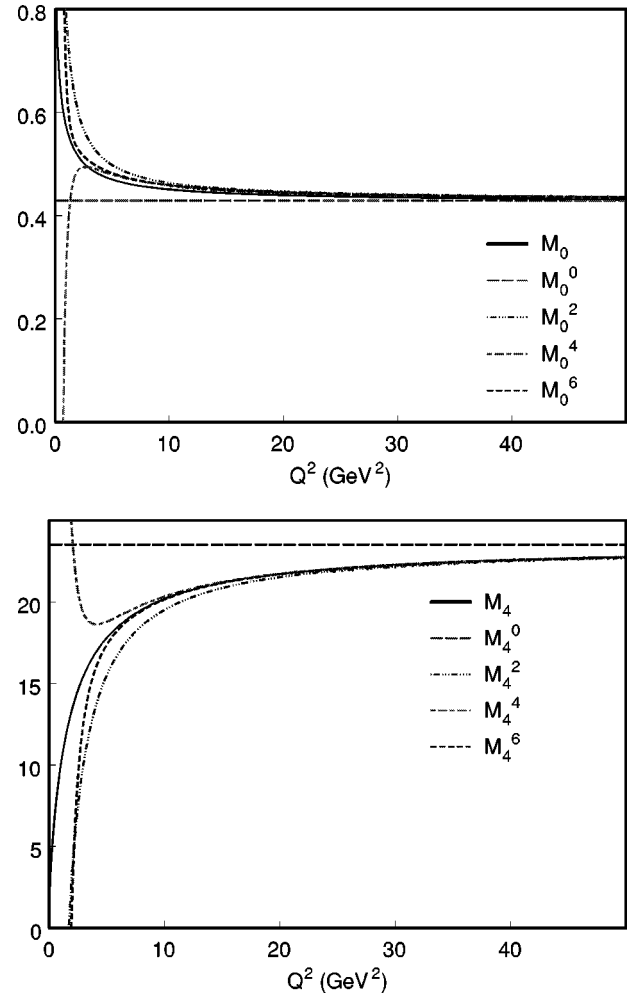


FIG. 6. Comparison of moments to the expansion of the continuum approximation in $1/Q^2$.

additional term to extend the approximation to lower Q^2 must require that the coefficient of this term be of opposite sign to the preceding term, leading to an alternating series. Higher moments have the curvature toward the finite result occurring at increasing values of Q^2 . This requires that the size of the coefficients for the higher terms in the series must also be increasing. This shows that the global duality observed in our model is the result of a delicate cancellation between many “higher twist” terms. However, while it is fascinating to speculate if duality in nature is realized by small higher-order expansion coefficients or by cancellations, our model is too simple to allow us to draw any conclusions about this. In fact, results presented in [25] indicate that the expansion coefficients have the same sign. One may hope that the building of more realistic models will allow us to gain a better insight into this question in the future.

The relation asymptotic behavior of the moments to sum rules can also be addressed in this model. Consider the moments of $F_2(u)$:

$$\mathcal{M}_n \equiv \int_{-\infty}^{\infty} du u^{n-2} F_2(u), \quad (62)$$

TABLE I. Leading coefficients of the expansion of the moments in $1/Q^2$.

		c_n^0	c_n^2	c_n^4	c_n^6
M_0	L	0.42859	0.20760	-0.26022	0.13530
	T	0.00000	0.13715	-0.19201	0.19056
	total	0.42859	0.34475	-0.45223	0.32586
M_1	L	1.00037	-0.32052	0.95887	-2.83879
	T	0.00000	0.32012	-0.95870	2.83881
	total	1.00037	-0.00039	0.00017	0.00002
M_2	L	2.64117	-2.37461	5.53199	-14.8231
	T	0.00000	0.84518	-3.22720	10.6710
	total	2.64117	-1.52944	2.30479	-4.1521
M_3	L	7.6117	-11.1960	26.8417	-71.9074
	T	0.0000	2.4357	-11.2072	40.7982
	total	7.6117	-8.7603	15.6345	-31.1092
M_4	L	23.5214	-47.6636	122.363	-336.920
	T	0.0000	7.5268	-40.233	159.245
	total	23.5214	-40.1368	82.130	-177.675

where the integral starts at $-\infty$ to include contributions from negative energy states as in the Coulomb sum rule. Using Eq. (37) this becomes

$$\begin{aligned}
 \mathcal{M}_n &= \frac{m^2}{4\pi^2 E_0} \int_{-\infty}^{\infty} du u^n \int_{|E_0-mu|}^{\infty} dp p N(p) \\
 &= \frac{m^2}{4\pi^2 E_0} \int_0^{\infty} dp p N(p) \int_{(E_0-p)/m}^{(E_0+p)/m} du u^n \\
 &= \frac{m^2}{4\pi^2 E_0} \int_0^{\infty} dp p N(p) \frac{1}{n+1} \\
 &\quad \times \left[\left(\frac{E_0+p}{m} \right)^{n+1} - \left(\frac{E_0-p}{m} \right)^{n+1} \right]. \quad (63)
 \end{aligned}$$

The two lowest moments

$$\mathcal{M}_0 = \frac{m}{2\pi^2 E_0} \int_0^{\infty} dp p^2 N(p) = \frac{m}{E_0} = 0.43002 \quad (64)$$

and

$$\mathcal{M}_1 = \frac{1}{2\pi^2} \int_0^{\infty} dp p^2 N(p) = 1 \quad (65)$$

are proportional to the normalization integral of the momentum distribution. Comparing these to the corresponding values of c_0^0 and c_1^0 in Table I shows that the contributions from negative energy states are small. Note also that for a spin-1/2 constituent the expression corresponding to Eq. (37) has a leading factor of u rather than u^2 as in this case. So the sum rules would be associated with M_1 and M_2 as expected.

VIII. LOW- Q^2 REGION

After studying the scaling behavior of our model at high Q^2 and the moments over a range of four-momentum transfers, we now discuss the behavior at low Q^2 . In this region, resonances are dominant for a wide range in the scaling variable.

Before discussing the numerical results, a remark on the kinematics is in order. For a fixed resonance in inclusive electron scattering from the nucleon, its position in terms of Bjorken's scaling variable is given by $x_{res} = Q^2 / (W_{res}^2 - M_N^2 + Q^2)$. This means that for higher Q^2 , the resonance position moves towards higher values of x_{Bj} , and for very large Q^2 , $x_{Bj} \rightarrow 1$. In our case, the maximal value of the scaling variable u is larger than 1, and for very large Q^2 , the resonances move out to very large values of u , where their contribution is extremely small.

If local duality holds, we expect the resonance curve to oscillate around the scaling curve and to average to it, once Q^2 is large enough. In Bloom-Gilman duality, the finite energy sum rule gets within 5% for $Q^2 \geq 1.75 \text{ GeV}^2$. For lower Q^2 , the resonances approach the scaling curve from below. In our case, we have the onset of scaling for larger values of Q^2 than observed in nature. This is not unexpected, as we consider an infinitely heavy meson as target and assume that this meson is made up of scalar quarks. For this reason, our cross section for photon exchange for large Q^2 is dominated by the longitudinal part, and the transverse part, comprising solely the convection current, is very small. In nature, for spin-1/2 quarks, we have the magnetization current, which is the dominant component of the cross section and which therefore determines the scaling behavior. So we cannot expect our model calculation to show the same behavior as experimental data for the same values of the four-momentum transfer.

In discussing local duality and resonances, the smoothing method used becomes important. The visual appearance of "resonances" depends on the chosen smoothing method: a

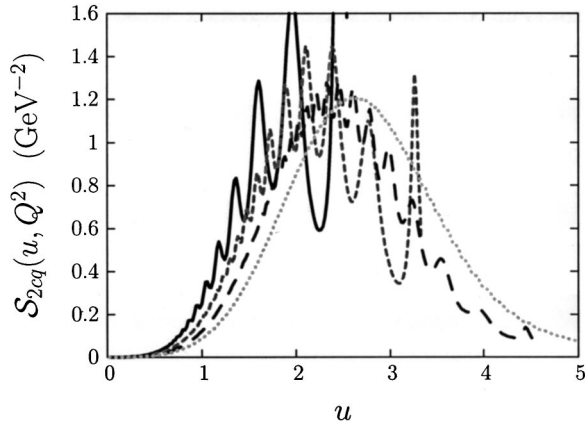


FIG. 7. The low Q^2 behavior of $S_{2cq}(u, Q^2)$ as a function of u for various values of Q^2 . The solid curve shows $Q^2=0.5$ GeV 2 , the short-dashed curve shows $Q^2=1$ GeV 2 , the long-dashed curve shows $Q^2=2$ GeV 2 , and the dotted curve shows $Q^2=50$ GeV 2 .

bumpy structure is seen only when a Breit-Wigner shape is inserted for the energy δ function. It also depends on the width chosen in the Breit-Wigner smoothing method. For a smaller width, the resonances are visible for higher Q^2 . Depending on the width of the Breit-Wigner, e.g., for $\Gamma=100$ MeV, we do not see any resonances for $Q^2=5$ GeV 2 , even though this value of Q^2 is below the scaling region. In this paper, the working definition of local duality which we use is “resonance curves oscillating around the scaling curve.” At some point, when considering more realistic models, it may be useful and necessary to introduce a sharper, more quantitative definition. However, at this stage, we are interested more in qualitative results and do not intend to quantify how well exactly local duality works for our simple model.

In Fig. 7, we show our results for the scaling function $S_{2cq}(u, Q^2)$ for various low values of Q^2 . The δ function in the energy has been smoothed out using the Breit-Wigner method, with a width of $\Gamma=100$ MeV. For visual purposes, we have assigned a small width to the elastic peak, too. One can see clearly from the figure that the resonances move out towards higher u with increasing Q^2 , as dictated by kinematics. While the elastic peak is rather prominent for $Q^2=0.5$ GeV 2 and $Q^2=1.0$ GeV 2 , it becomes negligible for $Q^2 \geq 2.0$ GeV 2 . This is the phenomenon we have observed already when studying the moments: the elastic contribution there vanishes rapidly with increasing Q^2 .

As already observed while studying the moments in the previous section, the approach to scaling when using a virtual photon is slower than for the all-scalar case discussed in [4]. It is clear that one needs to reach fairly large values of Q^2 before the “resonance curve” averages with good accuracy to the scaling function. Indeed, with our choice of Breit-Wigner width, this happens only when the bumps have already disappeared, i.e., for $Q^2 \geq 5$ GeV 2 .

In order to illustrate this point, we have included Fig. 8, where we used a value of $\Gamma=50$ MeV to smooth out the energy δ function. The curves are more jagged than for the larger width, and the $Q^2=5$ GeV 2 curve still shows plenty of resonance structure.

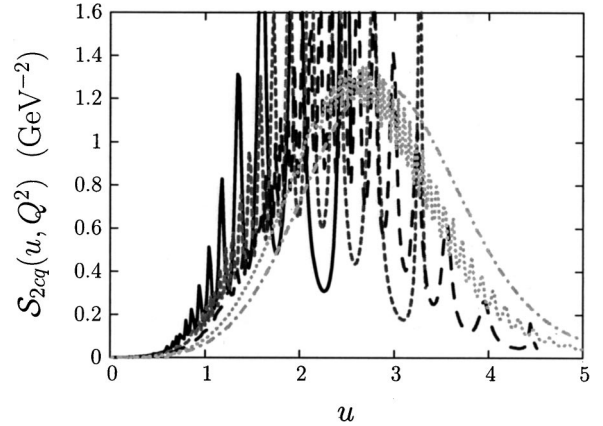


FIG. 8. The low Q^2 behavior of $S_{2cq}(u, Q^2)$ as a function of u for various values of Q^2 . The Breit-Wigner width chosen to smooth out the energy δ functions is $\Gamma=50$ MeV. The solid curve shows $Q^2=0.5$ GeV 2 , the short-dashed curve shows $Q^2=1$ GeV 2 , the long-dashed curve shows $Q^2=2$ GeV 2 , the dotted curve shows $Q^2=5$ GeV 2 , and the dash-dotted curve shows $Q^2=50$ GeV 2 .

Overall, we find that the onset of local duality is definitely slower than for the all-scalar case, which is what one expects due to the additional structure in the more realistic case discussed in this paper.

IX. SUMMARY AND OUTLOOK

We have presented a simple, quantum-mechanical model which allows us to obtain the qualitative features of Bloom-Gilman duality. The model assumptions we made are very basic: we assumed relativistic, confined valence scalar quarks, and treated the hadrons in the infinitely narrow resonance approximation. To simplify the situation further, we did not consider a three quark “nucleon” target, but a target composed of an infinitely heavy antiquark and a light quark. In contrast to [4] where all particles involved in the reaction—electrons, photons, and quarks—were considered to be scalar, we only use scalar quarks in this paper. This makes the present model more realistic—in particular, we were able to use a conserved current here. However, there is still much work to be done in modeling. The goal of our model calculations is to gain qualitative insight into duality, its applicability, and accuracy in various kinematic regions, not to quantitatively describe any data. In the future, we plan to describe more realistic situations in our model. Note that our assumptions are very basic and general, so that we will be able to extend our model in a straightforward manner.

There are several conditions that must be satisfied in order to see duality. In this paper, we put a special emphasis on three of these conditions: we demonstrated how the transition from coherent scattering at low Q^2 to incoherent scattering at high Q^2 takes place, we highlighted the role of relativity by considering the contributions to the Coulomb sum rule in a relativistic framework and a nonrelativistic framework, and we gave an analytic proof for the equality of the scaling curve in our model and the parton model result.

Quark-hadron duality is not only very interesting in itself; it also opens the door to very useful applications: duality

relates the resonance region data to data from the deep inelastic region. If duality is understood well enough and if the correct procedures for the averaging of the resonance data and the attendant errors are established, we may exploit duality to gather information in previously unreachable regimes. The investigation of polarized structure functions in the high-Bjorken- x region, $x_{Bj} \rightarrow 1$, is a major part of the experimental program at Jefferson Lab [26,9]. Even without knowing details about the correct averaging procedures, it is clear from the experimental results and our investigation of duality that the conventional, sharp distinction between the “resonance region,” corresponding to an invariant mass $W < 2$ GeV, and the “deep inelastic region” where $W > 2$ GeV, is entirely artificial.

While quark-hadron duality has been investigated by theorists before, modeling duality is an important new step in our way to a thorough understanding of this phenomenon. In the literature, one often finds the phrase that duality has been explained in terms of QCD by DeRujula, Georgi, and Politzer [24]. What was stated in their paper is that at moderate Q^2 , the higher-twist corrections to the lower moments of the structure function are small. The higher-twist corrections arise due to initial and final state interactions of the quarks and gluons. Hence, the average value of the structure function at moderate Q^2 is not very different from its value in the scaling region. While all this is true, the statement is merely a rephrasing of the experimentally observed fact that the resonance curve averages to the scaling curve in terms of the language of the OPE. However, the operator product expansion does not explain why a certain correction is small or why there are cancellations—the expansion coefficients which determine this behavior are not predicted in the OPE. The ultimate answer to this question might come from a numerical solution of QCD on the lattice, but an understanding of the physical mechanism that leads to the small values of the expansion coefficient in the framework of a model is highly desirable. Also, the OPE will break down for very low Q^2 . Duality was experimentally observed [2] to hold for Q^2 as low as 0.5 GeV^2 —a region where the validity of the OPE is questionable. In our analysis of the moments and their expansion coefficients, it became clear that a rigid application of the OPE at very low Q^2 will inevitably lead to large, alternating expansion coefficients.

The constant resonance to background ratio aspect of duality was addressed in several papers by Carlson and Mukhopadhyay [23]. They used counting rules to find the Q^2 dependence of the form factors of the resonances in the Breit frame and compared them to the behavior of the scaling curve for large x_{Bj} and to the behavior of the background in the same region. From these considerations, they could explain the constant ratio, provided the Q^2 -independent coefficients of the helicity amplitudes were not anomalously small, as in the case of the Δ resonance, for which the ratio vanishes. Still, there is no explanation why the coefficient is small in one case and not in others, and there exist several models with contradictory predictions.

The preceding observations clearly show the need for modeling. Even though one may obtain expansion coefficients from calculations on the lattice, an understanding of

the underlying physical mechanisms will most likely be gained only by considering models like ours. One great advantage of a purely analytical model like the one presented here is that explicit derivations of key quantities like the scaling function are feasible. The proof that the scaling function obtained for the transition from a bound quark to an excited bound quark is the same as the scaling function for the transition from a bound quark to a free quark was given here for a linear potential, which is the relativistic analogue of a harmonic oscillator. It is desirable to extend the investigation to other types of potential and to find a proof that applies to a general class of potentials. In a recent publication [6], numerical methods were applied to study the responses of a massless quark, and a disagreement between results with and without final state interactions was observed. This is in contrast to our findings, and it is important to understand the reasons for these differences.

The experimental data at very low Q^2 still average to one single curve, independent of their Q^2 [2]. However, this is not duality in the sense defined in this paper because this curve differs from the scaling curve. To investigate this interesting observation, one must go beyond the model presented here, which contains valence quarks only, and therefore must produce a valencelike shape. However, introducing sea quarks and modeling the decay of the excited resonances, along with the corresponding nonresonant production mechanisms from sea quark pairs, might shed considerable light on this issue.

ACKNOWLEDGMENTS

The authors thank F. Close, R. Ent, R. J. Furnstahl, N. Isgur, C. Keppel, S. Liuti, I. Niculescu, and W. Melnitchouk for stimulating discussions. We thank E. Braaten and R. J. Furnstahl for useful comments on the manuscript. The authors would like to acknowledge the contribution of our late colleague Nathan Isgur to this work by providing the initial motivation. This work was in part supported by funds provided by the U.S. Department of Energy (DOE) under cooperative research agreement No. DE-AC05-84ER40150.

APPENDIX A: DUALITY IN THE EXCITATION FORM FACTORS

In this appendix, we proceed to study duality in the excitation form factor $F_{0,N}(\vec{q}^2)$. While this duality is not directly related to an observable like the structure functions or response functions, it exhibits duality very clearly. Duality in the form factors has recently received some attention in [5].

The duality prediction is that it should not matter if we describe the process in question in a perturbative QCD picture, involving only a free quark, or in a hadronic picture with resonances. The form factor for a hypothetical free quark is just 1, as it does not have any structure. In the hadronic picture, we have inclusive electron scattering where we can excite all resonances—as the final hadronic state is not observed, we have to sum over the resonances incoherently. So we have to compare $\sum_N |F_{0,N}(\vec{q}^2)|^2$ to 1. Figure 9 shows single form factors for the lowest resonances, the elas-

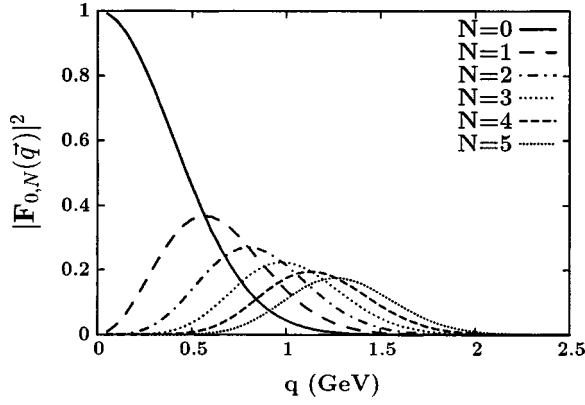


FIG. 9. The excitation form factor squared, $|F_{0N}(\vec{q})|^2$, for the lowest excitations, $N=0-5$, and $\beta=0.4$ GeV.

tic peak and inelastic excitations up to $N=5$. All form factors look qualitatively similar, except for the $N=0$ elastic form factor, which starts at 1 for $|\vec{q}|=0$. In general, the form factors increase in width and decrease in height with increasing N .

Using our previous expression for the form factor, Eq. (13), we find, for the sum up to a certain value N_{max} ,

$$\sum_{N=0}^{N_{max}} |F_{0N}(\vec{q})|^2 = \exp\left(-\frac{\vec{q}^2}{2\beta^2}\right) \sum_{N=0}^{N_{max}} \frac{1}{N!} \left(\frac{\vec{q}^2}{2\beta^2}\right)^N, \quad (\text{A1})$$

and it is obvious that

$$\sum_{N=0}^{N_{max}} |F_{0N}(\vec{q})|^2 = 1 \quad \text{if } N_{max} \rightarrow \infty, \quad (\text{A2})$$

as mentioned when discussing the Coulomb sum rule. However, we are limited in the maximal value of N_{max} not by technical problems, but by a physical constraint: we are considering electron scattering, i.e., spacelike kinematics, and therefore we must satisfy the condition

$$Q^2 > 0 \Leftrightarrow |\vec{q}| > \nu = E_N - E_0, \quad (\text{A3})$$

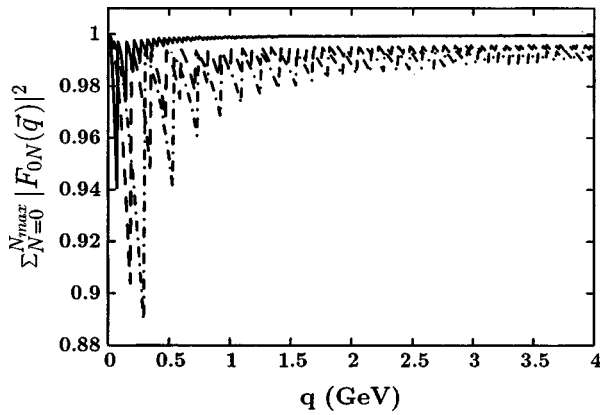


FIG. 10. The sum over the excitation form factors squared, for $\beta=0.2$ GeV (solid line), $\beta=0.4$ GeV (dashed line), and $\beta=0.6$ GeV (dash-dotted line).

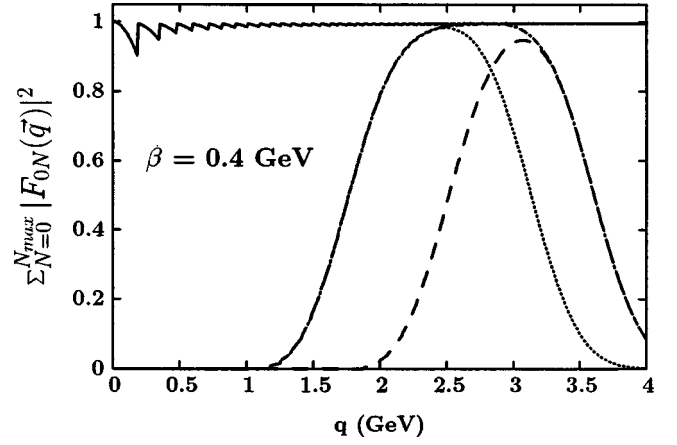


FIG. 11. The sum over the excitation form factors squared $|F_{0,N}(\vec{q}^2)|^2$ for $\beta=0.4$ GeV. The solid line shows the sum over all allowed N , the dash-dotted line shows the sum from $N=10$ to 40, the dotted line shows the sum from $N=10$ to 30, and the dashed line shows the sum from $N=20$ to 40.

so that for fixed three-momentum transfer $|\vec{q}|$, we find a limit on the value of N . The form factor sums are shown in Fig. 10 for various values of the oscillator parameter β . Larger values of β indicate a stronger binding.

The spiky character of the curves stems from the fact that the form factors for the N th state are only allowed to contribute to the sum if $|\vec{q}| > \nu_N$, where $\nu_N := E_N - E_0$. Therefore, the sum jumps up whenever another threshold is crossed. This effect can be observed best for the strongest binding, $\beta=0.6$ GeV, as the gaps between the energy levels are largest in this case. With increasing value of the three-momentum transfer, more and more resonance states can be excited and contribute to the sum: the spikes subside and the average value of the sum gets fairly close to 1. The curve for the weakest binding, $\beta=0.2$ GeV, becomes almost smooth and takes on a value of 0.9994; i.e., duality is violated by less than 0.06%. For $\beta=0.4$ GeV, the sum reaches 0.995, so duality is violated only by 0.5%. Even for the strongest binding, $\beta=0.6$ GeV, the duality prediction is satisfied within

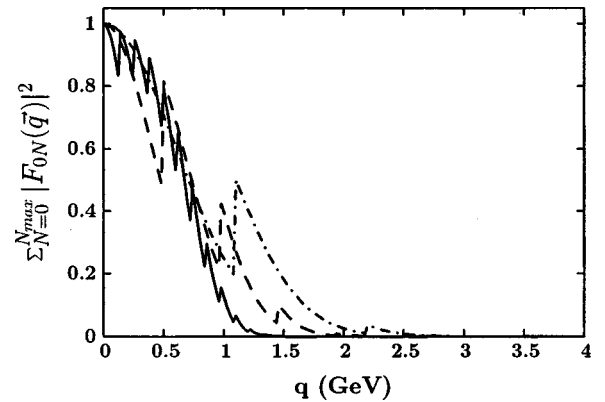


FIG. 12. The sum over the excitation form factors squared for the nonrelativistic case, for $\beta=0.2$ GeV (solid line), $\beta=0.4$ GeV (dashed line), and $\beta=0.6$ GeV (dash-dotted line). Note that the scale differs from the scale in Fig. 10.

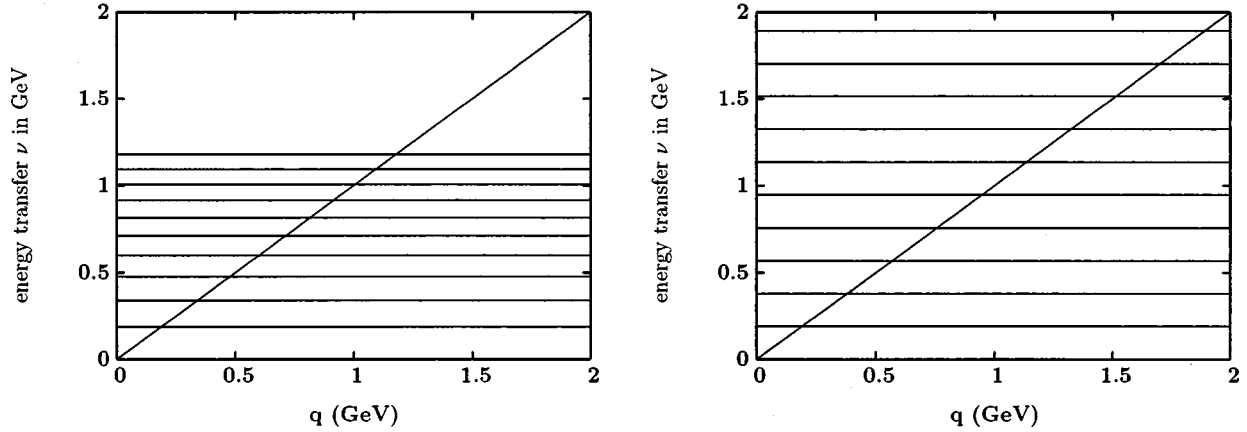


FIG. 13. The first eleven energy transfers ν for the relativistic case with $\beta=0.4$ GeV (left panel) and for the non-relativistic case with $\beta=0.25$ GeV (right panel). The diagonal line in both panels indicates the photopoint, i.e. $|\vec{q}|=\nu$. All energy transfers below this line are allowed in the space-like region.

1%. Here, we see a typical feature of duality: the need for many resonances to contribute in order to reproduce the behavior of a free quark. At low $|\vec{q}|$, where only a few resonances can contribute, the deviations from 1 are larger. The fact that duality is satisfied best for weak binding is what we expect: a quark that is bound very lightly and then receives a hard momentum transfer behaves essentially as if it were free. If the binding gets stronger, the situation gets more nonperturbative, and duality does not work as well.

The duality as seen in the form factors is reminiscent of duality in the decay rates of the semileptonic decay of heavy quarks [11,27]. There, in the limit of infinite masses of the b and c quarks, the loss of strength in the elastic channel is compensated for by the increase in the inelastic decay channels. Once one considers heavy, but not infinitely heavy, quark masses, one obtains a jagged structure, with peaks getting close to the free quark limit, quite similar to what we observe when considering the Coulomb sum rule and the excitation form factors.

Let us consider the number of resonances needed in more detail, so that we can draw further conclusions on the kind of duality we are observing here. In the calculations presented in Fig. 10, we summed up to the highest allowed N, N_{max} , which is quite large in general. In Fig. 11, we present the full curve, and three curves where we summed over a limited number of resonances only: namely, from $N=10$ to 40 (dash-dotted line), $N=10$ to 30 (dotted line) and $N=20$ to 40 (dashed line). One can see that for a small interval in three-momentum transfer of $\vec{q}=2.6-3.0$ GeV, it is sufficient to include only resonances from $N=10$ to 40 in order to reproduce the full, unrestricted curve. One also sees from the dotted line that for lower three-momentum transfer, the inclusion of just the 20 resonances from $N=10$ to 30 suffices to get close to the full curve $\vec{q}=2.6$ GeV, while the same number of resonances does not suffice to approximate the full curve at a slightly higher value of $|\vec{q}|$ (see the dashed line).

So while it is clear that the “degrees of freedom duality” holds very nicely over the whole kinematic range, we see that duality—the truncated version—does not hold as well: we do need a certain number of resonances to obtain duality

in a limited kinematic interval, and this number increases when we increase the three-momentum transfer.

APPENDIX B: THE ROLE OF RELATIVITY

We have stressed the importance of relativity before, and while it is quite obvious that one needs to include it for GeV momentum transfers to light quarks, it is instructive to see how relativity works for the form factors. In order to illustrate this point, we show the sum of the excitation form factors squared calculated for the nonrelativistic harmonic oscillator potential in Fig. 12. As mentioned in Sec. II, the wave functions, and therefore the form factors, are the same. The difference lies in the energy spectrum.

Obviously, duality in the nonrelativistic case does not work at all: the curves start out at 1, as the elastic form factor for $|\vec{q}|=0$ is 1, but then fall off immediately. Whenever a new threshold opens, the additional contribution is not sufficient to compensate the falloff of the other form factors: they do not contribute at their maximum value, but only with the “high- $|\vec{q}|$ ” side of the peak, where the form factor drops quickly; see Fig. 9. In the nonrelativistic case, the spacing between the energy levels is wider than in the relativistic case, where the levels shrink together. This means that considerably fewer resonances are allowed to contribute at the same, fixed $|\vec{q}|$, and therefore the resulting sum is much smaller and duality is violated. To clarify this point, we show a comparison of energy levels in Fig. 13.

In the left panel, we show the energy transfers ν for the first 11 energy levels (elastic and the first ten resonances) of a relativistic harmonic oscillator with $\beta=0.4$ GeV. The diagonal line marks the photopoint $|\vec{q}|=\nu$. This means that for a given $|\vec{q}|$, all the energy transfers below that line are in the spacelike region and therefore allowed. E.g., in the relativistic case for $|\vec{q}|=1$ GeV, nine resonances (counting the elastic) can be excited. The right panel of Fig. 13 shows the energy transfers ν for the energy levels of a nonrelativistic harmonic oscillator with $\beta=0.25$ GeV. We chose the β for the nonrelativistic oscillator in order to reproduce the energy

splitting between the ground state and the first excited state of the relativistic case. One clearly sees that fewer resonances can contribute here, e.g., only six compared to nine in the relativistic case at $|\vec{q}|=1$ GeV. The discrepancy grows larger for higher $|\vec{q}|$, as the relativistic energy levels move closer together, while the nonrelativistic ones are equally spaced.

In conclusion, we have seen that the relativistic description is necessary to ensure a correct treatment of the phase

space. Only with a proper relativistic phase space do we see duality in the excitation form factors. This was already clear from our discussion of the Coulomb sum rule in the main body of the paper.

Mathematically, degrees of freedom duality in the excitation form factor means that one can expand a plane wave (free quark) in a set of Hermite polynomials (bound quark), provided one uses a sufficiently large number of basis states. Any other set of orthonormal polynomials would also work.

-
- [1] E. D. Bloom and F. J. Gilman, Phys. Rev. Lett. **25**, 1140 (1970); Phys. Rev. D **4**, 2901 (1971).
- [2] I. Niculescu *et al.*, Phys. Rev. Lett. **85**, 1182 (2000); **85**, 1186 (2000); R. Ent, C. E. Keppel, and I. Niculescu, Phys. Rev. D **62**, 073008 (2000); **64**, 038302 (2001).
- [3] S. Liuti, R. Ent, C. E. Keppel, and I. Niculescu, hep-ph/0111063.
- [4] N. Isgur, S. Jeschonnek, W. Melnitchouk, and J. W. Van Orden, Phys. Rev. D **64**, 054005 (2001).
- [5] F. E. Close and N. Isgur, Phys. Lett. B **509**, 81 (2001).
- [6] M. W. Paris and V. R. Pandharipande, Phys. Lett. B **514**, 361 (2001); Phys. Rev. C **65**, 035203 (2002).
- [7] S. Simula, Phys. Lett. B **481**, 14 (2000); Phys. Rev. D **64**, 038301 (2001).
- [8] W. Melnitchouk, Phys. Rev. Lett. **86**, 35 (2001).
- [9] “The Science driving the 12 GeV upgrade,” edited by L. Cardman, R. Ent, N. Isgur, J.-M. Laget, C. Leemann, C. Meyer, and Z.-E. Meziani, 2001.
- [10] M. A. Shifman, A. I. Vainshtein, and V. I. Zakharov, Nucl. Phys. **B147**, 385 (1979); **B147**, 448 (1979); A. I. Vainshtein, V. I. Zakharov, V. A. Novikov, and M. A. Shifman, Sov. J. Nucl. Phys. **32**, 840 (1980); E. C. Poggio, H. R. Quinn, and S. Weinberg, Phys. Rev. D **13**, 1958 (1976); A. V. Radyushkin, in *Strong Interactions at Low and Intermediate Energies*, edited by J. L. Goity (World Scientific, Singapore, 2000); T. D. Cohen, R. J. Furnstahl, D. K. Griegel, and X. Jin, Prog. Part. Nucl. Phys. **35**, 221 (1995); B. Blok and M. Lublinsky, Phys. Rev. D **57**, 2676 (1998).
- [11] N. Isgur and M. B. Wise, Phys. Rev. D **43**, 819 (1991).
- [12] P. D. B. Collins, *An Introduction to Regge Theory and High Energy Physics* (Cambridge University Press, Cambridge, England, 1977).
- [13] C. G. Callan and D. J. Gross, Phys. Rev. Lett. **22**, 156 (1969).
- [14] K. W. McVoy and L. Van Hove, Phys. Rev. **125**, 1034 (1962).
- [15] F. E. Close (private communication); F. E. Close and Qiang Zhao (unpublished).
- [16] J. Zeng, J. W. Van Orden, and W. Roberts, Phys. Rev. D **52**, 5229 (1995).
- [17] O. W. Greenberg, Phys. Rev. D **47**, 331 (1993).
- [18] S. A. Gurvitz and A. S. Rinat, Phys. Rev. C **47**, 2901 (1993); S. A. Gurvitz, Phys. Rev. D **52**, 1433 (1995).
- [19] E. Pace, G. Salme, and F. M. Lev, Phys. Rev. C **57**, 2655 (1998).
- [20] O. Nachtmann, Nucl. Phys. **B63**, 237 (1973); **B78**, 455 (1974).
- [21] O. W. Greenberg and D. Bhaumik, Phys. Rev. D **4**, 2048 (1971).
- [22] R. Barbieri, J. Ellis, M. K. Gaillard, and G. G. Ross, Phys. Lett. **64B**, 171 (1976); Nucl. Phys. **B117**, 50 (1976).
- [23] C. E. Carlson and N. C. Mukhopadhyay, Phys. Rev. D **41**, 2343 (1990); **47**, R1737 (1993).
- [24] A. DeRujula, H. Georgi, and H. D. Politzer, Ann. Phys. (N.Y.) **103**, 315 (1977).
- [25] X. Ji and P. Unrau, Phys. Rev. D **52**, 72 (1995).
- [26] Jefferson Lab experiment E01-012, spokespersons J.-P. Chen, S. Choi, and N. Liyanage; Jefferson Lab Experiment E93-009, spokespersons G. Dodge, S. Kuhn, and M. Taiuti.
- [27] N. Isgur, Phys. Lett. B **448**, 111 (1999).