Sum rules for heavy flavor transitions in the small velocity limit

I. Bigi,[1,*] M. Shifman,[2,†] N.G. Uraltsev,[2,3] and A. Vainshtein[2,4,‡]

[1] *Department of Physics, University of Notre Dame du Lac, Notre Dame, Indiana 46556*

[2] *Theoretical Physics Institute, University of Minnesota, Minneapolis, Minnesota 55455*

[3] *St. Petersburg Nuclear Physics Institute, Gatchina, St. Petersburg 188350, Russia*[§]

[4] *Budker Institute of Nuclear Physics, Novosibirsk 630090, Russia*

(Received 27 December 1994)

We show how sum rules for the weak decays of heavy flavor hadrons can be derived as the moments of spectral distributions in the small velocity (SV) limit. Our derivation of the sum rules is based on the operator product expansion. This systematic approach allows us to determine corrections to these sum rules, to obtain new sum rules, and it provides us with a transparent physical interpretation; it also opens a new perspective on the notion of the heavy quark mass. Applying these sum rules we derive a *lower* bound on the deviation of the exclusive form factor $F_{B\to D^*}$ from unity at zero recoil; likewise we give a field-theoretical derivation of a previously formulated inequality between the expectation value for the kinetic energy operator of the heavy quark and for the chromomagnetic operator. We analyze how the known results on nonperturbative corrections must be understood when one takes into account the normalization point dependence of the low scale parameters. The relation between the field-theoretic derivation of the sum rules and the quantum-mechanical approach is elucidated.

PACS number(s): 13.20.He, 11.55.Hx, 12.38.Lg, 12.39.Hg

I. INTRODUCTION

The theory of preasymptotic effects in inclusive weak decays of heavy flavor hadrons has been developing from the early 1980s on [1–8] and is entering now a rather mature stage [9–19]. In recent papers [15–17] it was shown in particular how the effects due to the motion of the heavy quarks inside the hadrons can be incorporated in a systematic way, namely, through distribution functions which crucially depend on the ratio $\gamma = m_q/m_Q$ where m_Q and m_q are the masses of the initial and final state quarks, respectively. Among other things, it was mentioned that the formalism developed in Ref. [15] automatically ensures the Bjorken sum rule [20–22] in the small velocity (SV) limit [23] (Sec. IV of Ref. [15]). In the present work we discuss in more detail the sum rules of Refs. [23, 20, 21], the so-called optical sum rules derived later by Voloshin [24] (see also [25, 26]), as well as other similar sum rules including those considered by Lipkin [27] within a quantum-mechanical approach.

These are the main observations of our paper.

We demonstrate that these apparently isolated sum rules represent merely different moments of observable spectral distributions. Their physical meaning becomes absolutely transparent within the formulation of the problem suggested in Ref. [15]. Actually many crucial elements are already included, implicitly and explicitly, in Ref. [15], and so our task is to combine them. This approach will allow us to get new sum rules and to obtain corrections to the old ones in a systematic and comprehensive way.

The sum rules are used to derive a lower bound on the deviation of the exclusive form factor $F_{B\to D^*}$ from unity at zero recoil. The lower bound is essentially determined by the average value of the chromomagnetic operator μ_G^2, familiar from previous studies. We also obtain a new field-theoretic derivation of the previously formulated inequality $\mu_\pi^2 > \mu_G^2$, where

$$\mu_G^2 = \frac{1}{2M_B}\left\langle B\left|\bar b\,\frac{i}{2}\sigma_{\mu\nu}G^{\mu\nu}\,b\right|B\right\rangle,$$
$$\mu_\pi^2 = \frac{1}{2M_B}\langle B|\bar b\,(i\vec D)^2\,b|B\rangle. \qquad (1)$$

Its validity has been questioned by some authors under the pretext that the original line of reasoning was purely quantum mechanical.

The analysis of the sum rules gives us an opportunity to discuss the notion of the heavy quark mass from a point of view complementary to recent investigations [28, 29]. It is shown that the key theoretical parameter $\bar\Lambda(\mu)$ is directly related to quantities measurable in inclusive semileptonic decays of B mesons in a certain kinematical regime. The relation obtained makes absolutely explicit the fact that $\bar\Lambda$ does *not* exist as a universal constant, as had previously been believed. Any consistent treatment must deal with a μ-dependent function $\bar\Lambda(\mu)$, where μ is a normalization point.

The assertions formulated above for $\bar\Lambda(\mu)$ are applicable in full to the expectation value of the kinetic energy operator μ_π^2.

The conventional derivation of the sum rules in the SV limit based on QCD and the heavy quark expansion is shown to be equivalent to Lipkin's approach [27] in all as-

*Electronic address: bigi@undhep.hep.nd.edu

†Electronic address: shifman@vx.cis.umn.edu

‡Electronic address: vainshte@vx.cis.umn.edu

§Permanent address.

pects where the ultraviolet domain (i.e., virtual momenta of order of the heavy quark masses) is unimportant. In Lipkin's formalism the problem is treated quantum mechanically from the very beginning; the $b \to c$ transition is treated as an instantaneous perturbation. These two formalisms are complementary with respect to each other.

Generically the sum rules considered express the moments of the energy of the hadronic final state in $b \to c$ transitions as an expansion in various parameters: a perturbative expansion in α_s, a nonperturbative one in $1/m_{b,c}$; it is often convenient to expand also in $\vec{q}/m_c$, with q being the four-momentum carried away by the current inducing the $b \to c$ transition under consideration. This latter expansion represents the SV limit and $\vec{q} = 0$ the zero recoil point; near it $\vec{q}/m_c$ can be treated as a small parameter, and all quantities can consistently be expanded in it. Each term in this expansion generates its own set of sum rules. We will mostly be dealing with sum rules at zero recoil and those arising to order $(\vec{q}/m_c)^2$.

Our investigation focuses on nonperturbative corrections expressed through terms proportional to powers of $1/m_{b,c}$. The Wilson operator product expansion (OPE) [30] provides the theoretical foundation for calculating these corrections. The version of the OPE used in the problem at hand is somewhat peculiar and is supplemented by a systematic expansion in inverse heavy quark masses which is close, both conceptually and technically, to what is currently known as heavy quark effective theory (HQET) [31]. In many papers on the subject it is not realized that HQET can be consistently formulated only in the context of the OPE where one has fully implemented Wilson's idea of separating large and short distance contributions. In particular, one *cannot* treat $1/m_{b,c}$ corrections properly without introducing a normalization point μ as the boundary between the large and short distance domains. All theoretical parameters appearing in the expansion are in general μ dependent; the predictions for observable quantities are of course independent of μ. The issue of the μ dependence can be kept implicit as long as one neglects perturbative corrections. Yet once one undertakes the required simultaneous treatment of both perturbative and nonperturbative contributions one has to introduce the normalization point and thus to make the μ dependence explicit. This aspect is discussed in a special section, and throughout the paper we repeatedly emphasize that all HQET-type expansions become well defined only when formulated within the OPE framework.

The organization of the paper is as follows. In Sec. II we outline the specifics of the general OPE-based approach applied to heavy quark decays and derive the expansion for the heavy flavor hadron masses in QCD. In Sec. III basic elements of our approach are demonstrated in a simplified model where the heavy quarks are deprived of their spins and the external "weak" current considered is scalar. Sections IV and V are devoted to real QCD; here we derive a set of sum rules describing the inclusive semileptonic decay $B \to X_c l\nu$. The sum rules are considered in detail at zero recoil in Sec. IV and in the general situation with the emphasis on the SV kinematics, $\vec{q} \ll m_c$, in Sec. V. We find a lower bound on $1 - F^2_{B\to D^*}$ at zero recoil and present expressions for $\overline{\Lambda}(\mu)$ in terms of the differential distributions observable in semileptonic B meson decays. In Sec. VI we establish a relation between field-theoretic and quantum-mechanical derivations of the sum rules in the SV limit. Section VII addresses practical implications of "running" of basic low-energy parameters of the heavy quark theory. Section VIII briefly summarizes the main results. In the Appendix we give a quantum-mechanical derivation of a sum rule relating $\overline{\Lambda}$ to observable quantities.

II. GENERAL OPE APPROACH AND MASS FORMULAS

Inclusive heavy flavor decays are closely related to deep inelastic scattering (DIS). While the latter was in the focus of theoretical investigations in the early days of QCD, heavy flavor decays received marginal attention. Recently it has been shown that many elements of the theory of DIS find their parallels in heavy flavor decays. In this paper we discuss one more aspect with an apparent analogy in the theoretical treatments, namely, the sum rules.

Let us recall that the standard analysis of DIS [32] proceeds as follows. One starts from the operator product expansion (OPE) for the T product

$$\hat{T}_{\mu\nu} = i \int d^4x e^{-iqx} T\{j^\dagger_\mu(x) j_\nu(0)\} = \sum_r c^{(r)}_{\mu\nu\gamma_1\cdots\gamma_n}(q) O^{(r)}_{\gamma_1\cdots\gamma_n}(0), \quad (2)$$

where j_μ is the electromagnetic or some other current of interest and $O^{(r)}(0)$ are local operators. The average of $\hat{T}_{\mu\nu}$ over the nucleon state with momentum p presents a forward scattering amplitude of the Compton type. This amplitude depends on two kinematic variables q^2 and $\nu = qp$. For large Euclidean q^2 OPE leads to a set of predictions for the coefficients of the Taylor expansion in ν of the Compton amplitude at $\nu = 0$. These predictions are formulated in terms of the expectation values of local operators $\langle N|O^{(r)}|N\rangle$ over the nucleon state. The coefficients of the expansion are related via dispersion relations to integrals over the imaginary part of the amplitude at hand (moments of the structure functions).

The strategy, as well as the results obtained, is quite general. At the same time, certain moments play a distinguished role due to the fact that they turn out to be proportional to operators whose matrix elements between the nucleon state are known on general grounds. Relations emerging in this way are called sum rules proper and possess particular names. In the case of unpolarized targets we deal with the Adler sum rule for neutrino scattering and the Gross-Lewellyn-Smith sum rule [32].

Conceptually a very similar description can be applied to the weak decays of heavy flavor hadrons H_Q. One relates the observable quantities to a nonlocal transition operator $\hat{T}$, expands the latter into a series of local operators, and determines their expectation values between

the state H_Q. There exist, of course, several technical differences relative to the case of DIS: One deals with heavy quark currents, uses the heavy quark mass as the expansion parameter instead of the square of the momentum transfer, and forms the expectation values for the heavy flavor hadron. Moreover, a much larger number of these expectation values is known (in the leading approximation in $1/m_Q$), as compared to light hadrons. It is due to the fact that the distribution of the heavy quark inside the hadron is trivial to leading order, in contrast with the case of light quarks. As a consequence one can predict, for example, absolute decay rates, not only their evolution as the scale changes.

Now we make a digression of a general nature concerning the heavy quark expansion for the hadron masses. The corresponding expressions will appear in many instances in the sum rules below.

For any hadron its mass can be written as a matrix element of the trace of the full energy-momentum tensor of the theory, in particular,

$$M_{H_Q} = \frac{1}{2M_{H_Q}} \langle H_Q | \theta_{\mu\mu} | H_Q \rangle, \tag{3}$$

where $\theta_{\mu\nu}$ denotes the energy-momentum tensor; we use the relativistic normalization of the states,

$$\langle H_Q | H_Q \rangle = 2M_{H_Q} V,$$

for the hadron at rest. Our goal is to expand M_{H_Q} in inverse powers of m_Q. In particular, we need to construct the $1/m_Q$ expansion for $\theta_{\mu\mu}$ which follows from the expansion of the Lagrangian.

The original QCD Lagrangian has the form

$$\mathcal{L}_{\rm QCD} = -\frac{1}{4g_s^2}(G^a_{\mu\nu})^2 + \sum_q \bar{q}(i\not{D} - m_q)q, \tag{4}$$

with the sum running over all existing quarks. In the field theory one has to specify exactly the normalization point μ where all operators are defined; thus all couplings, viz., g_s and m_q, are functions of μ. The standard equation (4) assumes that the normalization point is much higher than all masses in the theory, $\mu \gg m_q$. In particular, no terms $\sim (m/\mu)^n$ are kept.

Constructing the effective theory designated to describe the low-energy properties of heavy flavor hadrons we need to have the normalization point μ below the mass of the heavy quark m_Q. This changes the generic form of the Lagrangian and a series of operators of higher dimension can appear. Their coefficients contain inverse powers of m_Q. It is important that all these effective operators are Lorentz scalars.

In what follows we are not interested in terms beyond order $1/m_Q^2$. To this accuracy the general form of the Lagrangian is

$$\mathcal{L} = -\frac{1}{4g_s^2}(G^a_{\mu\nu})^2 + \bar{Q}(i\not{D} - m_Q)Q + \sum_q \bar{q}(i\not{D} - m_q)q + \frac{c_G}{4m_Q}\bar{Q}i\sigma_{\mu\nu}G_{\mu\nu}Q$$
$$+ \sum_q \frac{f_q}{m_Q^2} m_q \bar{q} i\sigma_{\mu\nu}G_{\mu\nu}q + \sum_{\Gamma,q} \frac{d^{(\Gamma)}_{Qq}}{m_Q^2}\bar{Q}\Gamma Q \bar{q}\Gamma q + \sum_{\Gamma,q,q'} \frac{d^{(\Gamma)}_{qq'}}{m_Q^2}\bar{q}\Gamma q \bar{q}'\Gamma q' + \frac{h}{m_Q^2}\,\mathrm{Tr}\, G_{\mu\nu}G_{\nu\rho}G_{\rho\mu} + O\left(\frac{1}{m_Q^3}\right), \tag{5}$$

where $G^a_{\mu\nu}$ is the gluon field strength tensor and $G_{\mu\nu} \equiv t^a G^a_{\mu\nu}$ is its matrix representation ($\mathrm{Tr}\, t^a t^b = \delta_{ab}/2$); below we will often use the shorthand notation $i\sigma G = i\sigma_{\mu\nu}G_{\mu\nu} = i\gamma_\mu\gamma_\nu G_{\mu\nu}$. The sum over light quark flavors is shown explicitly as well as the sum over possible structures Γ of the four fermion operators. Again, all masses, couplings, and coefficients of higher dimension operators depend on the renormalization point. For example, the coefficient c_G has the form

$$c_G(\mu) = \left(\frac{\alpha_s(\mu)}{\alpha_s(m_Q)}\right)^{\frac{3}{b}} - 1 + O(\alpha_s), \quad b = 11 - \frac{2}{3}n_f. \tag{6}$$

Expansion in the parameter $1/m_Q$ leads to the fact that the logarithmic evolution of operators generally mixes them with the ones of *higher* dimension, in contrast with the expansion in m_q for light quarks. The above relation is the effect of such mixing of the operator $\bar{Q}Q$ with the chromomagnetic one, $\bar{Q}i\sigma GQ$.

In evolving the parameters of the effective Lagrangian down to the normalization point μ one can use the values of c, d, f, h obtained at $\mu \sim m_Q$ as the initial conditions. In fact all operators with dimension $d \geq 5$ appear at this scale only as a result of quantum corrections and contain explicit powers of $\alpha_s(m_Q)/\pi$. For the sake of simplicity we will generally neglect them in our subsequent analysis, although it is easy to keep track of them if necessary. This neglect does not mean that we disregard all effects of order $1/m_Q^2$; they appear from the leading "tree level" operators as well and will be accounted for consistently. Note that neglecting higher order terms in Eq. (5) give us the standard QCD equations of motion for both the gluon and quark fields. If we were interested in quantum corrections at the order of $1/m_Q$ or $1/m_Q^2$ we would need to incorporate additional terms in the equation of motions or, alternatively, consistently consider these higher order terms in the Lagrangian as a perturbation and expand in them to the necessary order.

In the theory with Lagrangian (4) the trace of the energy-momentum tensor has the form

$$\theta_{\mu\mu} = -m_Q \frac{d}{dm_Q}\mathcal{L} + \left(\mu\frac{d}{d\mu} - m_q\frac{d}{dm_q}\right)\mathcal{L}, \tag{7}$$

$$-m_Q \frac{d}{dm_Q}\mathcal{L} = m_Q \bar{Q}Q, \tag{8}$$

$$\left(\mu\frac{d}{d\mu} - m_q\frac{d}{dm_q}\right)\mathcal{L} \equiv \mathcal{D} = \frac{\beta(\alpha_s)}{16\pi\alpha_s^2}(G^a_{\mu\nu})^2 + \sum_q m_q(1+\gamma_m)\,\bar{q}q - \mu\frac{dm_Q(\mu)}{d\mu}\,\bar{Q}Q, \tag{9}$$

where γ_m is the anomalous dimension of light quark mass, $\mu\, dm_q(\mu)/d\mu = -\gamma_m m_q$, and $\beta(\alpha_s)$ is the Gell-Mann–Low function, $\mu\, d\alpha_s(\mu)/d\mu = \beta(\alpha_s)$. Note that the scale dependence of mass enters $\theta_{\mu\mu}$ in the same way for light and heavy quarks; the explicit form of this dependence is, of course, different. The modification of $\theta_{\mu\mu}$ for the case when higher dimension operators are included is straightforward using the general expression Eq. (7). In Eq. (9), $\mathcal{D}$ has the meaning of the part associated with light degrees of freedom.

In the case we consider, thus, the hadron mass can be expressed in terms of two expectation values:

$$M_{H_Q} = \frac{1}{2M_{H_Q}}\langle H_Q|m_Q\bar{Q}Q|H_Q\rangle + \frac{1}{2M_{H_Q}}\langle H_Q|\mathcal{D}|H_Q\rangle. \tag{10}$$

This formula has an obvious advantage over the usual representation [31], giving the mass through the expectation value of the Hamiltonian: Mass is a Lorentz scalar quantity, and in Eq. (10) it is expressed through expectation values of Lorentz scalar operators. Of course, if one wants to develop the $1/m_Q$ expansion for M_{H_Q}, one eventually arrives at the standard expansion involving the same nonrelativistic operators. All relevant terms in the effective Hamiltonian appear to be directly related to the $1/m_Q$ expansion of the operator $\bar{Q}Q$. Let us elucidate this point in more detail dealing with the two operators in Eq. (10) in turn.

To calculate the matrix element of $\bar{Q}Q$ we, following Refs. [6, 7], start from the heavy quark current. In the rest frame of H_Q we have

$$\frac{1}{2M_{H_Q}}\langle H_Q|\bar{Q}\,\gamma_0 Q|H_Q\rangle = 1. \tag{11}$$

Using the decomposition

$$iD_\mu = m_Q v_\mu + \pi_\mu, \quad v_\mu = \left(p_{H_Q}\right)_\mu / M_{H_Q}, \tag{12}$$

and the equation of motion $i\not{D}\,Q = m_Q\,Q$, which imply

$$\frac{1-\gamma_0}{2}\,Q = \frac{\not{\pi}}{2m_Q}\,Q, \quad \pi_0\,Q = -\frac{\pi^2 + \frac{i}{2}\sigma G}{2m_Q}\,Q, \tag{13}$$

we arrive at the identity

$$\bar{Q}\,Q = \bar{Q}\gamma_0 Q + 2\bar{Q}\left(\frac{1-\gamma_0}{2}\right)^2 Q = \bar{Q}\gamma_0 Q - 2\bar{Q}\,\frac{\overleftarrow{\not{\pi}}}{2m_Q}\cdot\frac{\overrightarrow{\not{\pi}}}{2m_Q}\,Q = \bar{Q}\gamma_0 Q + \bar{Q}\frac{\not{\pi}^2}{2m_Q^2}Q + \text{ total derivative;} \tag{14}$$

in the first relation above the operators $\overleftarrow{\pi}_\mu$ act on the $\bar{Q}$ field. We are considering the forward matrix elements with zero momentum transfer and thus can drop all terms with total derivatives. Equation (14) supplemented with the equations of motion (13) then generates the complete $1/m_Q$ expansion for the scalar density:

$$\bar{Q}\,Q = \bar{Q}\,\gamma_0\,Q + \frac{1}{2m_Q^2}\bar{Q}\left(\pi^2 + \frac{i}{2}\sigma G\right)Q = \bar{Q}\,\gamma_0\,Q - \frac{1}{2m_Q^2}\bar{Q}\,(\vec{\pi}\cdot\vec{\sigma})^2 Q - \frac{1}{4m_Q^3}\,\bar{Q}\left[-(\vec{D}\cdot\vec{E}) + 2\vec{\sigma}\cdot\vec{E}\times\vec{\pi}\right]Q + O\left(\frac{1}{m_Q^4}\right). \tag{15}$$

Here $E_i = G_{i0}$ is the chromoelectric field, and its covariant derivative is defined as[1] $D_j E_k = -i[\pi_j, E_k]$; we have omitted the term $\bar{Q}([\pi_k,[\pi_0,\pi_i]] - [\pi_i,[\pi_0,\pi_k]])Q$ using the Jacobi identity. Moreover, $\vec{D}\cdot\vec{E} = g_s^2 t^a J_0^a$ by virtue of the QCD equation of motion, where $J_\mu^a = \sum_q \bar{q}\gamma_\mu t^a q$ is the color quark current; therefore the first of the $1/m_Q^3$ terms can be rewritten as the local four-fermion interaction [15].

We see that the $m_Q\bar{Q}Q$ part of $\theta_{\mu\mu}$ in Eq. (10) generates

$$m_Q - \frac{1}{2M_{H_Q}}\left\langle H_Q\middle|\bar{Q}\left[\frac{(\vec{\sigma}\cdot\vec{\pi})^2}{2m_Q} + \frac{1}{4m_Q^2}\left[-(\vec{D}\cdot\vec{E}) + 2\vec{\sigma}\cdot\vec{E}\times\vec{\pi}\right]\right]Q\middle|H_Q\right\rangle + O\left(\frac{1}{m_Q^3}\right). \tag{16}$$

Equation (16) shows the explicit dependence of $\theta_{\mu\mu}$ on the heavy quark mass. However, it is not the only source of m_Q dependence, since also the hadronic state H_Q depends on m_Q. To account for this dependence one needs

[1]Note that in our notation $\vec{D} = -\partial/\partial\vec{x} - i\vec{A}$; therefore, $(\vec{D}\cdot\vec{E}) = -\text{div}\,\mathbf{E}$ in the Abelian case.

to introduce the basis of asymptotic ($m_Q = \infty$) states, $|H_Q\rangle_{m_Q=\infty}$, and develop the perturbation theory in explicit $1/m_Q$ terms in the Lagrangian. Let us emphasize that the explicit $1/m_Q^n$ terms above come from short distances of order $1/m_Q$ whereas effects due to the m_Q dependence of states H_Q are large distance ($\sim \Lambda_{\rm QCD}^{-1}$) ones.

We are interested in terms $\sim \Lambda_{\rm QCD}$, $\Lambda_{\rm QCD}^2/m_Q$ and $\Lambda_{\rm QCD}^3/m_Q^2$ in mass. Therefore we must account for $1/m_Q$ effects in the matrix elements of $\bar{Q}(\vec{\sigma}\cdot\vec{\pi})^2 Q$ as well as terms m_Q^{-1} and m_Q^{-2} in the expectation value of $\mathcal{D}$.

The term m_Q in Eq. (16) is the only one in M_{H_Q} which is linear in mass. To zeroth order in m_Q the $m_Q\bar{Q}Q$ part gives no contribution and the only m_Q-independent term comes from the expectation value of $\mathcal{D}$. To leading order in m_Q this expectation value presents what is usually called $\overline{\Lambda}$:

$$\frac{1}{2M_{H_Q}}\langle H_Q|\mathcal{D}|H_Q\rangle = \frac{1}{2M_{H_Q}}\left\langle H_Q\left|\frac{\beta(\alpha_s)}{16\pi\alpha_s^2}G^2\right|H_Q\right\rangle + \frac{1}{2M_{H_Q}}\left\langle H_Q\left|\sum_q m_q(1+\gamma_m)\bar{q}\,q\right|H_Q\right\rangle - \mu\frac{dm_Q}{d\mu} = \overline{\Lambda} + O(m_Q^{-1}). \tag{17}$$

In a sense, the first term is an analogue of the gluon condensate in the QCD vacuum [33]. Instead of the QCD vacuum we deal now, however, with the ground state in the sector with the heavy quark charge equal to unity.

The analogy continues further and allows us to derive low-energy theorems very similar to those taking place [34] for the vacuum correlation functions involving G^2 and other operators. In the present context the low-energy theorems take the form

$$-i\int d^4x\,\frac{1}{2M_{H_Q}}\langle H_Q|T\left\{\mathcal{D}(0),\bar{Q}(x)\mathcal{O}Q(x)\right\}|H_Q\rangle = d\frac{1}{2M_{H_Q}}\langle H_Q|\bar{Q}\mathcal{O}Q|H_Q\rangle[1+O(m_Q^{-1})], \tag{18}$$

where the operator $\bar{Q}(x)\mathcal{O}Q(x)$ is bilinear in Q and $\bar{Q}$ and $\mathcal{O}$ contains derivatives and light fields A_μ, q; the factor d is equal to the dimension[2] of $\mathcal{O}$. Derivation of these low-energy relations merely parallels that given in Ref. [34].

Armed with these theorems we will be able to calculate subleading terms in Eq. (17). But first we need the expansion of the QCD Lagrangian to the necessary order in $1/m_Q$. We obtain, in the standard way,

$$\begin{aligned}\bar{Q}(i\not{D}-m_Q)Q = \bar{Q}\frac{1+\gamma_0}{2}\left(1+\frac{(\vec{\sigma}\cdot\vec{\pi})^2}{8m_Q^2}\right)\left[\pi_0-\frac{1}{2m_Q}(\vec{\pi}\cdot\vec{\sigma})^2\right.\\ \left.-\frac{1}{8m_Q^2}\left[-(\vec{D}\cdot\vec{E})+2\vec{\sigma}\cdot\vec{E}\times\vec{\pi}\right]\right]\left(1+\frac{(\vec{\sigma}\cdot\vec{\pi})^2}{8m_Q^2}\right)\frac{1+\gamma_0}{2}Q+O\left(\frac{1}{m_Q^3}\right)\\ = \varphi^+(\pi_0-\mathcal{H}_Q)\varphi+O\left(\frac{1}{m_Q^3}\right),\end{aligned} \tag{19}$$

where

$$\varphi = \left(1+\frac{(\vec{\sigma}\cdot\vec{\pi})^2}{8m_Q^2}\right)\frac{1+\gamma_0}{2}Q \tag{20}$$

is a nonrelativistic two-component spinor field. This substitution is nothing but the Foldy-Wouthuysen transformation which is necessary to keep the term linear in π_0 in its canonical form to ensure the mass independence of the field φ (for a recent discussion, see Ref. [35]). If it is not accomplished, there would be an implicit dependence of the heavy quark fields on mass and the low-energy theorems will be modified. In Eq. (19),

$$\mathcal{H}_Q = \frac{1}{2m_Q}(\vec{\sigma}\cdot\vec{\pi})^2 + \frac{1}{8m_Q^2}[-(\vec{D}\cdot\vec{E})+2\vec{\sigma}\cdot\vec{E}\times\vec{\pi}],\quad (\vec{\sigma}\cdot\vec{\pi})^2 = \vec{\pi}^2 + \vec{\sigma}\vec{B}, \tag{21}$$

is the nonrelativistic Hamiltonian through second order in $1/m_Q$. It coincides with the textbook expression [36] in the Abelian case. Notice that $1/m_Q^2$ and $1/m_Q^3$ terms in Eq. (15) are just $d\mathcal{H}_Q/dm_Q$.

---

[2]In fact it includes the anomalous dimension of $\bar{Q}\mathcal{O}Q$ which appears when radiative corrections are incorporated.

Now we can calculate power (in $1/m_Q$) corrections to matrix elements. Indeed, the next-to-leading term in $(2M_{H_Q})^{-1}\langle H_Q|\,\mathcal{D}\,|H_Q\rangle$ is given by

$$m_Q\left\{\frac{1}{2M_{H_Q}}\langle H_Q|\,\mathcal{D}\,|H_Q\rangle\right\}_{1/m_Q}=-i\int d^4x\,\frac{1}{2M_{H_Q}}\left\langle H_Q\middle|T\left\{\mathcal{D}(0),\bar{Q}(x)\frac{(\vec{\sigma}\cdot\vec{\pi})^2}{2}Q(x)\right\}\middle|H_Q\right\rangle$$
$$=2\left\{\frac{1}{2M_{H_Q}}\left\langle H_Q\middle|\bar{Q}\frac{(\vec{\sigma}\cdot\vec{\pi})^2}{2}Q\middle|H_Q\right\rangle\right\}_{m_Q=\infty}. \tag{22}$$

Combining now Eqs. (22) and (15) we arrive at a standard formula [31, 38] for the hadron mass:

$$M_{H_Q}=m_Q+\bar{\Lambda}+\frac{1}{m_Q}\left\{\frac{1}{2M_{H_Q}}\left\langle H_Q\middle|\bar{Q}\frac{(\vec{\sigma}\cdot\vec{\pi})^2}{2}Q\middle|H_Q\right\rangle\right\}_{m_Q=\infty}+O(m_Q^{-2}). \tag{23}$$

Using the same technology it is not difficult to develop the expansion one step further to include the $O(m_Q^{-2})$ term. To this end we must find and take into account the $1/m_Q$ contribution in $\langle H_Q|\bar{Q}(\vec{\sigma}\cdot\vec{\pi})^2Q|H_Q\rangle$ in Eq. (15) as well as the $1/m_Q^2$ contribution in $\langle H_Q|G^2|H_Q\rangle$. For the first matrix element we have

$$m_Q\left\{\frac{1}{2M_{H_Q}}\langle H_Q|\bar{Q}(\vec{\sigma}\cdot\vec{\pi})^2Q|H_Q\rangle\right\}_{1/m_Q}$$
$$=-i\int d^4x\,\frac{1}{4M_{H_Q}}\langle H_Q|T\{\bar{Q}(x)(\vec{\sigma}\cdot\vec{\pi})^2Q(x),\bar{Q}(0)(\vec{\sigma}\cdot\vec{\pi})^2Q(0)\}|H_Q\rangle'_{m_Q=\infty}\equiv-\rho^3, \tag{24}$$

where ρ^3 is a positive parameter of the order of $\Lambda^3_{\rm QCD}$ measuring the correlation function above. The prime in Eq. (24) indicates that the diagonal transitions have to be removed from the correlation function (analogously to elimination of the disconnected parts in the vacuum correlators).

By the same token using low-energy theorems, the $1/m_Q^2$ piece in $(2M_{H_Q})^{-1}\langle H_Q|\,\mathcal{D}\,|H_Q\rangle$ can be written in terms of a similar correlation function:

$$m_Q^2\left\{\frac{1}{2M_{H_Q}}\langle H_Q|\,\mathcal{D}\,|H_Q\rangle\right\}_{1/m_Q^2}=\frac{1}{2M_{H_Q}}(-i)\int d^4x\left\langle H_Q\middle|T\left\{\mathcal{D}(0),\,\frac{1}{8}\bar{Q}(x)[-(\vec{D}\cdot\vec{E})+2\vec{\sigma}\cdot\vec{E}\times\vec{\pi}]Q(x)\right\}\middle|H_Q\right\rangle$$
$$-\frac{1}{8}\frac{1}{2M_{H_Q}}\int d^4x\,d^4y\,\langle H_Q|T\left\{\mathcal{D}(0),\bar{Q}(x)(\vec{\sigma}\cdot\vec{\pi})^2Q(x),\bar{Q}(y)(\vec{\sigma}\cdot\vec{\pi})^2Q(y)\right\}|H_Q\rangle$$
$$=\frac{3}{8}\frac{1}{2M_{H_Q}}\langle H_Q|\,\bar{Q}(x)[-(\vec{D}\cdot\vec{E})+2\vec{\sigma}\cdot\vec{E}\times\vec{\pi}]Q(x)|H_Q\rangle-\frac{3}{4}\rho^3, \tag{25}$$

where the proper analogue of Eq. (18) is used.

Now we have at our disposal everything needed to write down the mass formula including $1/m_Q^2$ corrections. Combining Eqs. (16), (24), and (25) we obtain

$$M_{H_Q}=m_Q+\bar{\Lambda}+\frac{1}{m_Q}\left\{\frac{1}{2M_{H_Q}}\left\langle H_Q\middle|\bar{Q}\frac{(\vec{\sigma}\cdot\vec{\pi})^2}{2}Q\middle|H_Q\right\rangle\right\}_{m_Q=\infty}$$
$$+\frac{1}{8m_Q^2}\left\{\frac{1}{2M_{H_Q}}\langle H_Q|\bar{Q}[-(\vec{D}\cdot\vec{E})+2\vec{\sigma}\cdot\vec{E}\times\vec{\pi}]Q|H_Q\rangle\right\}_{m_Q=\infty}-\frac{\rho^3}{4m_Q^2}+O(m_Q^{-3}). \tag{26}$$

It is clear that the m_Q dependence obtained has the usual quantum-mechanical interpretation: The third and fourth terms in Eq. (26) present the expectation values of the Hamiltonian $\mathcal{H}_Q$, Eq. (21), over the asymptotic state $|H_Q\rangle_{m_Q=\infty}$, and the term $-\rho^3/4m_Q^2$ presents the second-order iteration of the Hamiltonian.

For further usage we introduce here the following notations for the expectation values of the two terms in $1/m_Q^2$ Hamiltonian, which are the Darwin and the convection current (spin-orbital) interactions, respectively:

$$\rho^3_D=\frac{1}{2M_B}\langle B|\,\bar{b}(-\tfrac{1}{2}\vec{D}\cdot\vec{E})b\,|B\rangle=\frac{1}{2M_{B^*}}\langle B^*|\,\bar{b}(-\tfrac{1}{2}\vec{D}\cdot\vec{E})b\,|B^*\rangle$$
$$\rho^3_{LS}=\frac{1}{2M_B}\langle B|\,\bar{b}\vec{\sigma}\cdot\vec{E}\times\vec{\pi}b\,|B\rangle=-\frac{3}{2M_{B^*}}\langle B^*|\,\bar{b}\vec{\sigma}\cdot\vec{E}\times\vec{\pi}b\,|B^*\rangle. \tag{27}$$

ρ_D^3 is directly related to the third moment of the heavy quark distribution function and has been estimated in Ref. [15] (see also [37]). On the other hand, the value of ρ_{LS}^3 is expected to be suppressed for $L=0$ states such as B and B^*.

A similar decomposition can be obviously made for the nonlocal correlators whose expectation value has been generically denoted by ρ^3; the four unknown parameters appearing here can be related to the $1/m_Q$ part of the expectation values of $\bar{Q}\vec{\pi}^{\,2}Q$ and $\bar{Q}\vec{\sigma}\cdot\vec{B}Q$ in pseudoscalar and vector ground states:

$$\rho_{\pi\pi}^3 = i\int d^4x\,\frac{1}{4M_B}\langle B|T\{\bar{b}\vec{\pi}^{\,2}b(x),\bar{b}\vec{\pi}^{\,2}b(0)\}|B\rangle',$$
$$\rho_{\pi G}^3 = i\int d^4x\,\frac{1}{2M_B}\langle B|T\{\bar{b}\vec{\pi}^{\,2}b(x),\bar{b}\vec{\sigma}\cdot\vec{B}b(0)\}|B\rangle',$$
$$\frac{1}{3}\rho_S^3\delta_{ij}\delta_{kl}+\frac{1}{6}\rho_A^3(\delta_{ik}\delta_{jl}-\delta_{il}\delta_{jk}) = i\int d^4x\,\frac{1}{4M_B}\langle B|T\{\bar{b}\sigma_iB_kb(x),\bar{b}\sigma_jB_lb(0)\}|B\rangle'. \tag{28}$$

Then one has, for the parameters ρ^3 in B and B^*, respectively,

$$(\rho^3)_B = \rho_{\pi\pi}^3+\rho_{\pi G}^3+\rho_S^3+\rho_A^3,\ (\rho^3)_{B^*} = \rho_{\pi\pi}^3-\frac{1}{3}\rho_{\pi G}^3+\rho_S^3-\frac{1}{3}\rho_A^3. \tag{29}$$

The $1/m_Q$ corrections to the expectation values of $\vec{\pi}^{\,2}$ and $\vec{\sigma}\cdot\vec{B}$ are

$$\frac{1}{2M_B}\langle\vec{\pi}^{\,2}\rangle_B = \mu_\pi^2-\frac{2\rho_{\pi\pi}^3+\rho_{\pi G}^3}{2m_b}+O(m_b^{-2}),$$
$$\frac{1}{2M_{B^*}}\langle\vec{\pi}^{\,2}\rangle_{B^*} = \mu_\pi^2-\frac{2\rho_{\pi\pi}^3-\frac{1}{3}\rho_{\pi G}^3}{2m_b}+O(m_b^{-2}),$$
$$\frac{1}{2M_B}\langle\vec{\sigma}\cdot\vec{B}\rangle_B = -\mu_G^2-\frac{\rho_{\pi G}^3+2\rho_S^3+2\rho_A^3}{2m_b}+O(m_b^{-2}),$$
$$\frac{1}{2M_{B^*}}\langle\vec{\sigma}\cdot\vec{B}\rangle_{B^*} = \frac{1}{3}\mu_G^2-\frac{-\rho_{\pi G}^3+6\rho_S^3-2\rho_A^3}{6m_b}+O(m_b^{-2}). \tag{30}$$

Very similar parameters have been introduced in Ref. [37] where the mass formulas were discussed in the standard HQET approach.[3]

In the subsequent section we will consider in detail the toy model where heavy quarks are spinless and light quark masses vanish [15]. The absence of spin simplifies the analysis above. Let us briefly review the changes. In this toy model the trace of energy-momentum tensor is given by

$$\theta_{\mu\mu} = 2m_Q^2\bar{Q}Q+\frac{\beta(\alpha_s)}{16\pi\alpha_s^2}G_{\mu\nu}^aG_{\mu\nu}^a-2\mu\frac{dm_Q^2}{d\mu}\bar{Q}Q; \tag{31}$$

the heavy quark field is still denoted by Q but its dimension is m now, not $m^{3/2}$. Moreover, this expression, as well as the most of other ones, can be obtained from the spinor case by substitution:

$$Q_{\rm spinor}\to\sqrt{2m_Q}Q_{\rm scalar}.$$

Spin matrices are now absent, i.e., $(\vec{\sigma}\cdot\vec{\pi})^2_{\rm spinor}\to(\vec{\pi})^2_{\rm scalar}$. In particular, the equation of motion and the expansion of the expectation value for $\bar{Q}Q$ take the form

$$\pi_0Q = -\frac{\pi^2}{2m_Q}Q,$$
$$\frac{1}{2M_{H_Q}}\langle H_Q|2m_Q\bar{Q}Q|H_Q\rangle = 1-\frac{1}{2M_{H_Q}}\left\langle H_Q\left|\bar{Q}\frac{\vec{\pi}^2}{m_Q}Q\right|H_Q\right\rangle+O(m_Q^{-4}). \tag{32}$$

Notice that in the scalar case there is no explicit $1/m_Q^3$ terms; it corresponds to the absence of $1/m_Q^2$ terms in the nonrelativistic Hamiltonian for scalar particles:

$$\mathcal{H}_Q = \frac{\vec{\pi}^2}{2m_Q}+O(m_Q^{-3}).$$

---

[3]We do not agree with some numerical coefficients of Ref. [37]. In particular, Eqs. (50)–(55) of that paper are not quite consistent. Also, our result for the contribution of the Darwin term in mass is positive in the factorization approximation and smaller by a factor of 2.

Low-energy theorems have exactly the same form as before with

$$\mathcal{D} = \frac{\beta(\alpha_s)}{16\pi\alpha_s^2}G^2 - 2\mu\frac{dm_Q^2}{d\mu}\,\bar{Q}Q.$$

As a result the expansion for the hadron mass in the scalar case takes the form

$$M_{H_Q} = m_Q + \bar{\Lambda} + \frac{1}{2m_Q}\left\{\langle H_Q|\bar{Q}\vec{\pi}^2 Q|H_Q\rangle\right\}_{m_Q=\infty} - \frac{\rho^3}{4m_Q^2} + O(m_Q^{-3}), \tag{33}$$

with

$$\bar{\Lambda} = \left\{\frac{1}{2M_{H_Q}}\langle H_Q|\mathcal{D}|H_Q\rangle\right\}_{m_Q=\infty},$$
$$\rho^3 = 2m_Q^2 i\int d^4x\frac{1}{2M_{H_Q}}\langle H_Q|T\{\bar{Q}(x)\vec{\pi}^2Q(x),\bar{Q}(0)\vec{\pi}^2Q(0)\}|H_Q\rangle', \tag{34}$$

and

$$\langle H_Q|\bar{Q}\vec{\pi}^2Q|H_Q\rangle = \frac{M_{H_Q}}{m_Q}\left\{\langle H_Q|\bar{Q}\vec{\pi}^2Q|H_Q\rangle\right\}_{m_Q=\infty} - \frac{\rho^3}{m_Q} + O(m_Q^{-2}). \tag{35}$$

Let us return to ordinary QCD and briefly discuss the normalization point dependence of $\bar{\Lambda}$. In this context we can neglect terms of order $1/m_Q$ and higher; Eq. (17) gives the following definition of $\bar{\Lambda}(\mu)$,

$$\bar{\Lambda}(\mu) \equiv \left\{M_{H_Q} - m_Q(\mu)\right\}_{m_Q=\infty} = \frac{1}{2M_{H_Q}}\left\langle H_Q\left|\frac{\beta(\alpha_s)}{16\pi\alpha_s^2}G^2 + \sum_q m_q(1+\gamma_m)\bar{q}\,q\right|H_Q\right\rangle - \mu\frac{dm_Q}{d\mu}. \tag{36}$$

The last term is specific for the field theory and is absent in the naive quantum-mechanical approach.

In the perturbation theory the μ dependence of m_Q appears in the order α_s,

$$\frac{dm_Q}{d\mu} = -c_m\frac{\alpha_s}{\pi} + O(\alpha_s^2). \tag{37}$$

The matrix element in the right-hand side of Eq. (36), constituting the first part of $\bar{\Lambda}$, does not undergo perturbative renormalization[4] in order α_s. Nevertheless it is μ dependent in the theory with heavy quarks; its scale dependence appears in terms $\sim\alpha_s^2$. It might be tempting to drop the last term in Eq. (36) and consider

$$\bar{\lambda} = \frac{1}{2M_{H_Q}}\left\langle H_Q\left|\frac{\beta(\alpha_s)}{16\pi\alpha_s^2}G^2 + \sum_q m_q(1+\gamma_m)\bar{q}\,q\right|H_Q\right\rangle = M_{H_Q} - m_Q(\mu) + \mu\frac{dm_Q}{d\mu} \tag{38}$$

as an option for $\bar{\Lambda}$. However, it would not help defining a "purely nonperturbative" $\bar{\Lambda}$; in particular, an attempt to put $\mu\to 0$ would lead to the same infrared renormalon as in the pole mass [28, 29]. The transparent physical interpretation of this phenomenon will be discussed below in Secs. III D and V A. It is worth noting that the combination $m_Q(\mu) - \mu\,dm_Q/d\mu$ in Eq. (38) has the meaning of the *one-loop* pole mass of the heavy quark and therefore it naturally enters the "practical OPE" calculations performed at the one-loop level (for a detailed discussion see Sec. VII).

The $1/m_Q^2$ terms in mass M_{H_Q} obtained in this section will be used in derivation of the second sum rules at zero recoil. Equation (36) might be interesting by itself, though, as an alternative, compared to the standard HQET analysis, definition of one of its key parameters $\bar{\Lambda}$.

[4]The term with the anomalous dimension of the light quark mass *per se* is not renormalization invariant in the order α_s^2, but its μ dependence is compensated by the mixing with the G^2 term.

III. BJORKEN, "OPTICAL," AND OTHER SUM RULES: TOY MODEL

To introduce the reader to the range of questions to be considered below in the most straightforward way we first eliminate all inessential technicalities, such as the quark spins, and resort to a simplified model which has been previously discussed in Ref. [15]. It will be a rather simple exercise to return afterwards to standard QCD which will be done in Sec. IV.

A. Description of the model

We consider a toy example where all quarks are spinless; two, denoted by Q and q, couple to a massless real scalar field ϕ:

$$\mathcal{L}_\phi = h\bar{Q}\phi q + \text{H.c.}, \tag{39}$$

where h is the coupling constant and $\bar{Q} = Q^\dagger$. The masses of the quarks Q and q are both *large*. Later on we will analyze the SV limit where

$$\Lambda_{\rm QCD} \ll m_Q - m_q \ll m_Q \tag{40}$$

has to hold, but for the time being the ratio $\gamma = m_q/m_Q$ can be arbitrary provided that[5] $m_Q - m_q \gg \Lambda_{\rm QCD}$. The field ϕ carries color charge zero; the reaction $Q \to q + \phi$ is thus a toy model for the radiative decays of the type $B \to X_s\gamma$.

The total width for the free quark decay $Q \to q + \phi$ is given by the expression

$$\Gamma(Q \to q\phi) = \frac{h^2 E_0}{8\pi m_Q^2} \equiv \Gamma_0, \tag{41}$$

where

$$E_0 = \frac{m_Q^2 - m_q^2}{2m_Q}. \tag{42}$$

As explained in Refs. [1–15] the theory of preasymptotic effects in inclusive decays is based on introducing the transition operator

$$\hat{T} = i\int d^4x\, e^{-iqx} T\{\bar{Q}(x)q(x)\,,\ \bar{q}(0)Q(0)\}. \tag{43}$$

Then the energy spectrum of the ϕ particle in the inclusive decay is obtained from $\hat{T}$ in the following way:

$$\frac{d\Gamma}{dE} = \frac{h^2 E}{4\pi^2 M_{H_Q}} \text{Im}\,\langle H_Q|\hat{T}|H_Q\rangle, \tag{44}$$

where H_Q denotes a hadron built from the heavy quark Q and the light cloud (including the light antiquark). If not stated otherwise, H_Q will denote the ground state in a given channel. Moreover, one can (and must) apply the Wilson operator product expansion [30] (OPE) to express the nonlocal operator $\hat{T}$ through an infinite series of local operators with calculable coefficients.

B. OPE and predictions for observable quantities

In the Born approximation the transition operator has the form (Fig. 1)

$$\hat{T} = -\int d^4x \left(x\left|\bar{Q}\frac{1}{(P_0 - q + \pi)^2 - m_q^2}Q\right|0\right), \tag{45}$$

$$iD_\mu \equiv (P_0)_\mu + \pi_\mu \equiv m_Q v_\mu + \pi_\mu,$$

which is particularly suitable for constructing the OPE by expanding Eq. (45) in powers of π (see Ref. [15] where

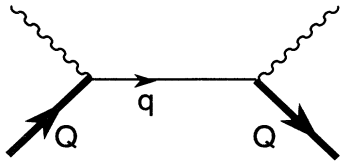

FIG. 1. The tree graph for the transition operator.

all notations have been introduced). The operators appearing as a result of this expansion are ordered according to their dimension. The leading operator $\bar{Q}Q$ has dimension 2 (let us recall that the scalar Q field has dimension 1 in contrast to the real quark fields of dimension 3/2, which leads in particular to different normalization factors; we still use relativistic normalization in the bulk of the paper, except in Sec. VI). Its expectation value for the state H_Q reduces to unity to leading order in $1/m_Q$; this contribution gives rise to the parton result (41). Beyond the leading approximation, according to Eq. (32), it takes the form

$$\frac{1}{2M_{H_Q}}\langle H_Q|\bar{Q}Q|H_Q\rangle$$

$$= \frac{1}{2m_Q}\left(1 - \frac{1}{2m_Q^2}\langle H_Q|\bar{Q}\vec{\pi}^2 Q|H_Q\rangle + \cdots\right); \tag{46}$$

thus, one gets a correction of order $1/m_Q^2$.

Corrections of the same order come from higher-dimensional operators in the expansion of the transition operator. There are no relevant operators of the next-to-leading dimension 3 (more exactly, as first noted in Ref. [4] in the framework of HQET, they vanish because of the equations of motion[6]). The only operator of dimension 4 in our toy model has the form $\bar{Q}\vec{\pi}^2 Q$, and it generates a $1/m_Q^2$ correction after taking the matrix element over H_Q.

After some simple algebra one finds

$$\begin{aligned}\frac{1}{\pi}\text{Im}\,\langle\hat{T}\rangle = &\left(\frac{\langle\bar{Q}Q\rangle}{2m_Q} - \frac{\langle\bar{Q}\vec{\pi}^2 Q\rangle}{12m_Q^3}\right)\delta(E - E_0)\\ &- \frac{E_0\langle\bar{Q}\vec{\pi}^2 Q\rangle}{12m_Q^3}\delta'(E - E_0)\\ &+ \frac{E_0^2\langle\bar{Q}\vec{\pi}^2 Q\rangle}{12m_Q^3}\delta''(E - E_0) + \cdots,\end{aligned} \tag{47}$$

[5]To reach the SV limit in the $B \to X_c e\nu$ transitions there is no need to assume that $m_b - m_c \ll m_c$. The small velocity regime for the c quark can be ensured by adjusting q^2 appropriately.

[6]This statement is sometimes erroneously interpreted as a proof of the absence of a term linear in $1/m_Q$ in the total width; see below. As a matter of fact the authors of Ref. [4] believed that the linear term *may* appear in the total width from the overall normalization of $\langle H_Q|\bar{h}_v h_v|H_Q\rangle$, as explicitly stated, e.g., on p. 404 of Ref. [4]. Also, "matching" to QCD, the necessary step in the HQET approach, was not considered. In fact, the question of the absence or presence of the $1/m_Q$ correction cannot be solved in HQET *per se* since the problem requires a full-QCD analysis of the total decay rate to perform such a matching; see the corresponding discussion in Ref. [28].

where operators of higher dimension have been ignored, and we have used a shorthand notation for the expectation value over H_Q: $\langle\cdots\rangle \equiv \langle H_Q|\cdots|H_Q\rangle$.

The expansion of $\mathrm{Im}\,\hat{T}$ into local operators generates more and more singular terms at the point where the ϕ spectrum would be concentrated in the free quark approximation. The physical spectrum, on the other hand, is a smooth function of E. In principle, one could derive a smooth spectrum by summing up an infinite set of operators to all orders (for more details see Refs. [15–17]). There is no need to carry out this summation here, however, since we are interested only in integral characteristics (we will discuss certain sum rules); as far as they are concerned, the expansion in Eq. (47) is perfectly legitimate.

At first, we calculate the total width by substituting Eq. (47) into Eq. (44) and integrating over E. As a matter of fact, the result has already been given in Ref. [15],

$$\Gamma = \int dE \frac{d\Gamma}{dE} = \Gamma_0 \left(1 - \frac{\mu_\pi^2}{2m_Q^2} + \cdots\right), \tag{48}$$

where the integration runs from 0 to the physical boundary $E_0^{\rm phys}$, expressed in the hadron masses,

$$E_0^{\rm phys} = \frac{M_Q^2 - M_q^2}{2M_Q}, M_Q \equiv M_{H_Q}, \tag{49}$$

and the same convention for M_q. We use here the standard notation for the expectation value of $\vec{\pi}^2$:

$$\mu_\pi^2 = \frac{1}{2M_{H_Q}} \langle H_Q|2m_Q \bar{Q}\vec{\pi}^2 Q|H_Q\rangle. \tag{50}$$

The expression in the parentheses in Eq. (48) is nothing else than the (corrected) matrix element of the operator $\bar{Q}Q$; the only possible effective operator $\bar{Q}\vec{\pi}^2 Q$ of dimension 4 is not a Lorentz scalar and thus cannot appear in Γ [15]. The meaning of this term in $\langle\bar{Q}Q\rangle$ is quite transparent: It reflects time dilation for the moving quark, and the coefficient $(-1/2)$ could, therefore, have been guessed from the very beginning. The absence of a correction of order $1/m_Q$ in the total width in the toy example under consideration is a manifestation of the general theorem established in Refs. [6, 7] and discussed in more detail recently in Ref. [28].

As we will see shortly, Eq. (48) treated in the SV limit is equivalent to two results simultaneously, that of Ref. [23] and the Bjorken sum rule [20, 21], with appropriate corrections due to terms which were not considered in Refs. [20, 21, 23]. All sum rules analogous to that of Eq. (48) below will generically be referred to as *the first sum rule.*

Equation (48) is the first example of the sum rules we will be dealing with throughout the paper. Its derivation (as well as that of all similar sum rules) intuitively is perfectly clear—the integrated spectrum of the decay obtained at the quark-gluon level using the OPE is equated to the integrated physical spectrum, saturated by the genuine hadronic states. We will elaborate on the justification for this procedure in Sec. IV A where, among other things, we discuss the accuracy one may expect from relations of this type; the general idea lying behind all such relations is widespread in QCD. For the moment we just adopt the heuristic attitude outlined above without submerging into further, less pragmatic, questions.

Next, we calculate the average energy of the ϕ particle. The corresponding sum rule in the SV limit is just a version of Voloshin's optical sum rule.

To be more exact, let us define moments I_n,

$$I_n = \int_0^{E_0^{\rm phys}} dE\,(E_0^{\rm phys} - E)^n \frac{1}{\Gamma_0}\frac{d\Gamma}{dE}, \tag{51}$$

and consider the first moment I_1. Notice that in the SV limit $E_0^{\rm phys} - E$ reduces to the excitation energy of the final hadron produced in the decay. The factor $E_0^{\rm phys} - E$ in the integrand eliminates the "elastic" peak, so that the integral is saturated only by the inelastic contributions. Let us note in passing (we will discuss this point later in more detail) that the optical sum rule [24] in its original formulation is actually divergent; a cutoff must be introduced by hand. In our formulation, of course, there is no place for divergences; the expectation value of $E_0^{\rm phys} - E$ is defined via a convolution with the physical spectrum, and since the ϕ energy in the decay is finite, the expectation value of $E_0^{\rm phys} - E$, as well as higher moments, is certainly finite. The decay kinematics provides us with a natural cutoff at the scale of the energy release.

Explicit calculation of I_1 using Eq. (47) yields

$$I_1 = \Delta - \frac{\mu_\pi^2 E_0^{\rm phys}}{2m_Q^2}, \tag{52}$$

where Δ is defined as

$$\Delta = E_0^{\rm phys} - E_0; \tag{53}$$

see Eqs. (42) and (49). The sum rule (52) contains Voloshin's result, again with corrections left aside previously [24, 25]. The sum rules analogous to that of Eq. (52) below will be generically referred to as *the second sum rule.*

It is quite evident that the series of such sum rules can readily be continued further. For the next moment, for instance, we get

$$I_2 = \int_0^{E_0^{\rm phys}} dE\,(E_0^{\rm phys} - E)^2 \frac{1}{\Gamma_0}\frac{d\Gamma}{dE} = \Delta^2 + \frac{\mu_\pi^2 E_0^2}{3m_Q^2}. \tag{54}$$

Relations of this type will be referred to as *the third sum rule.* Analyzing this sum rule in the SV limit one obtains, in principle, additional information, not included in the results of Refs. [20, 21, 23, 24]. It is worth emphasizing that in Eqs. (48), (52), and (54) we have collected all terms through order $\Lambda_{\rm QCD}^2$, whereas those of order $\Lambda_{\rm QCD}^3$ are systematically omitted. Predictions for higher moments would require calculating terms $O(\Lambda_{\rm QCD}^3)$ and higher.

C. SV limit and the sum rules

We proceed now to discussing the sum rules (48), (52), and (54) in the SV limit, i.e., in the limit of small velocity $\vec{v} = -\vec{q}/E$, where $-\vec{q}$ and E are the momentum and energy of the final hadronic system. The assumption Eq. (40) means that $|\vec{v}| \ll 1$. The availability of the extra expansion parameter, $|\vec{v}| \ll 1$, makes the SV limit a very interesting theoretical laboratory. We have already mentioned that in this limit the Bjorken sum rule relates the distortion of the "elastic" peak to the integral over the inelastic contributions. Here we will elaborate on this issue. First, we briefly recall what is already known and then present new results.

For technical reasons we will assume that

$$\Lambda_{\rm QCD} \ll m_Q - m_q \ll \sqrt{m_Q \Lambda_{\rm QCD}}. \tag{55}$$

The inequality on the right-hand side is not essential and can easily be lifted; it helps, however, to make all formulas more concise, and we will accept it for a while, thus replacing Eq. (40) by the stronger condition of Eq. (55). We now supplement the expansion in $\Lambda_{\rm QCD}$ carried out above by an expansion in $\vec{v}$. The natural hierarchy of parameters in the domain (55) is

$$\vec{v}^2, \ |\vec{v}|(\Lambda_{\rm QCD}/m_Q), \ \Lambda^2_{\rm QCD}/m_Q^2.$$

Terms of order $\vec{v}^4 \ll \Lambda^2_{\rm QCD}/m_Q^2$ will be omitted.

Expanding the hadron masses in terms of the heavy quark masses m_Q and m_q one finds, for scalar quarks,

$$M_Q = m_Q + \overline{\Lambda} + \frac{\mu_\pi^2}{2m_Q} + O(1/m_Q^2),$$

$$M_q = m_q + \overline{\Lambda} + \frac{\mu_\pi^2}{2m_q} + O(1/m_q^2). \tag{56}$$

Equation (56) implies, in turn, that

$$\Delta = E_0^{\rm phys} - E_0 = \frac{1}{2}v_0^2\left(\overline{\Lambda} + \frac{\overline{\Lambda}^2 + \frac{1}{2}\mu_\pi^2}{M_Q}\right) - v_0\frac{\mu_\pi^2}{2M_Q} + \cdots, \tag{57}$$

where for convenience we have introduced an auxiliary parameter v_0,

$$v_0 = (M_Q - M_q)/M_Q \simeq \frac{E_0}{M_Q}. \tag{58}$$

This parameter approximately coincides in the SV limit with the velocity of the heavy hadron produced in the transition $Q \to q\phi$. Indeed, if the mass of the produced excited hadron is $M_q + \epsilon$, then its velocity is

$$|\vec{v}| = v_0 - \frac{\epsilon}{M_Q} - \frac{v_0\epsilon}{M_Q}, \tag{59}$$

plus terms of higher order in ϵ and/or v_0.

With all these definitions the sum rules (48), (52), and (54) take the form

$$I_0 = 1 - \frac{\mu_\pi^2}{2m_Q^2} + \cdots, \tag{60}$$

$$I_1 = \frac{1}{2}v_0^2\overline{\Lambda} - v_0\frac{\mu_\pi^2}{M_Q} + \cdots, \tag{61}$$

$$I_2 = \frac{1}{3}v_0^2\mu_\pi^2 + \cdots, \tag{62}$$

where the ellipses denote the systematically omitted terms of order $\Lambda^3_{\rm QCD}$, as well as $O(v_0^2\Lambda^2_{\rm QCD})$ terms for I_1 and $O(v_0^4\Lambda^2_{\rm QCD})$ ones for I_2. The first term on the right-hand side of Eq. (61) corresponds to Voloshin's relation; the second term is a correction whose physical meaning will soon become clear.

To interpret the sum rules derived it will be instructive to consider the transition operator off the physical cut. Expressing Eq. (47) in terms of $E - E_0^{\rm phys}$ we then get

$$M_Q^{-1}\langle H_Q|\hat{T}|H_Q\rangle = \frac{1}{2m_Q^2}\left\{\left(1 - \frac{2\mu_\pi^2}{3m_Q^2}\right)\frac{1}{E - E_0^{\rm phys}} + \left(-\Delta + \frac{E_0\mu_\pi^2}{6m_Q^2}\right)\frac{1}{(E - E_0^{\rm phys})^2} + \frac{E_0^2\mu_\pi^2}{3m_Q^2}\frac{1}{(E - E_0^{\rm phys})^3} + \cdots\right\}. \tag{63}$$

To zeroth order in $\vec{v}$ or $\Lambda_{\rm QCD}$ there remains only the first term on the right-hand side of Eq. (63). The inclusive decay rate is then totally saturated by a single "elastic" channel, the production of the ground state meson containing q. This is the perfect inclusive-exclusive duality noted in Ref. [23]. The peak in the ϕ spectrum obtained in the quark transition $Q \to \phi q$ survives hadronization in this approximation; at the hadronic level $H_Q \to H_q\phi$ we still have the same peak at the same energy. Equations (61) and (62) are, of course, trivially satisfied in this case because in the absence of inelastic contributions both sum rules yield vanishing numbers.

Furthermore, the terms with $\Lambda_{\rm QCD}$ come with v. To order $\overline{\Lambda}$ there is only one such term, appearing in Eq. (61) for I_1. This term shows that the inelastic production *must* already be present at this level. It corresponds to the production of a meson H_q^* with excitation energy $\sim \Lambda_{\rm QCD}$ and a residue $\propto \vec{v}_0^{\,2}$. Then Eq. (60) implies that the height of the elastic peak is reduced by $\vec{v}_0^{\,2}$. A rough model exemplifying this picture can be obtained from Eq. (63). In this approximation, one can rewrite it as follows, omitting all numerical factors:

$$M_Q^{-1}\langle H_Q|\hat{T}|H_Q\rangle = \frac{1}{2m_Q^2}\left\{(1 - v_0^2)\frac{1}{E - E_0^{\rm phys}} + v_0^2\frac{1}{(E - E_0^{\rm phys} + \overline{\Lambda})} + \cdots\right\}. \tag{64}$$

The first term is the elastic peak while the second is an inelastic contribution. Of course, Eq. (64) is an illustration and not a unique solution.

Technically, the term $\vec{v}^{\,2}\bar{\Lambda}$ in Eq. (61) arises because the "elastic" peak of the quark transition situated at $E = E_0$ is slightly shifted when we pass to the hadronic transition; the genuine hadronic elastic peak is situated at $E = E_0^{\rm phys}$, to the right of the quark peak (see Fig. 2).

Including $O(\bar{\Lambda}^2)$ contributions we get correction terms in Eqs. (60) and (61), and Eq. (62) for I_2 becomes nontrivial. The new term in Eq. (61) has a simple meaning. In the toy example at hand the excited mesons produced in the decay at order $\vec{v}^{\,2}$ are spin-1 mesons, with the vertex proportional to $(\vec{v}\cdot\vec{\epsilon})$ where $\vec{\epsilon}$ is the polarization vector and $\vec{v}$ is the velocity of the given meson. This velocity differs, however, from v_0 by terms of order $\Lambda_{\rm QCD}/m_Q$; see Eq. (59). This rather trivial shift in velocity nicely explains all the sum rules above. Indeed, let us take into account that the physical $\vec{v}^{\,2}$ is reduced by the amount $\sim v_0(\bar{\Lambda}/M_Q)$. The height of the inelastic contribution is proportional to the square of the physical velocity, while the size of the inelastic domain is $\sim \bar{\Lambda}$. Hence, with our accuracy the right-hand side of Eq. (61) is expected to be $\sim v_0^2\bar{\Lambda} - v_0(\bar{\Lambda}^2/m_Q)$, while the right-hand side of Eq. (62) is expected to be $\sim \bar{\Lambda}^2 v_0^2$, in full accordance with what we actually have. A qualitative model of saturation now takes the form

$$M_Q^{-1}\langle H_Q|\hat{T}|H_Q\rangle = \frac{1}{2m_Q^2}\left\{\left[1 - \frac{2\mu_\pi^2}{3m_Q^2} - \left(\frac{\Delta}{\lambda} - \frac{E_0\mu_\pi^2}{6\lambda m_Q^2}\right)\right]\frac{1}{E - E_0^{\rm phys}} + \left(\frac{\Delta}{\lambda} - \frac{E_0\mu_\pi^2}{6\lambda m_Q^2}\right)\frac{1}{E - E_0^{\rm phys} + \lambda}\right\}, \quad (65)$$

where

$$\lambda \simeq \frac{2}{3}\frac{\mu_\pi^2}{\bar{\Lambda}} \sim \Lambda_{\rm QCD}.$$

D. Perturbative gluon corrections

So far we have assigned the gluon field to play the role of a soft medium to incorporate the effects of long distance dynamics and have completely ignored perturbative gluon corrections. Yet those have to be included; among other things the emission of hard gluons generates the spectral density outside the end-point region which is very relevant for our analysis.

In calculating radiative gluon correction we can disregard, in the leading approximation, nonperturbative effects, such as the difference between m_Q and M_Q or the "Fermi" motion of the initial quark. Thus we deal with the decay of the free quark Q at rest into $q + \phi +$ gluon. The virtual gluon contribution merely renormalizes the constant h in the analysis presented above. The analogous renormalization for the spinor case has been calculated in Ref. [23].

The effects from real gluon emission are most simply calculated in the Coulomb gauge, where only the graph shown in Fig. 3 contributes. A straightforward computation yields, to leading order in $\vec{v}^{\,2}$,

$$\frac{d\Gamma^{(1)}}{dE} = \Gamma_0\frac{8\alpha_s}{9\pi}\frac{E^3}{E_0 m_Q^2}\frac{1}{E_0 - E}. \quad (66)$$

It is well known that the logarithmic singularity in the integral over E for $\Gamma_{\rm tot}$ is canceled by a contribution of soft virtual gluons to the renormalization of h. For the second sum rule we are going to discuss here, this infrared range is not singular.

Gluon emission obviously contributes to the spectrum in its entire domain $0 < E < E_0$. In this order α_s does not run, of course. To demonstrate its scale dependence one has to carry out a two-loop calculation; it is quite evident, however, that it is $\alpha_s(E - E_0)$ that enters. Therefore, strictly speaking, one cannot apply Eq. (66) too close to E_0. Even leaving aside the blowing up of $\alpha_s(E - E_0)$, there exists another reason not to use Eq. (66) in the vicinity of E_0: If E is close to E_0, the emitted gluon is soft; such gluons are to be treated as be-

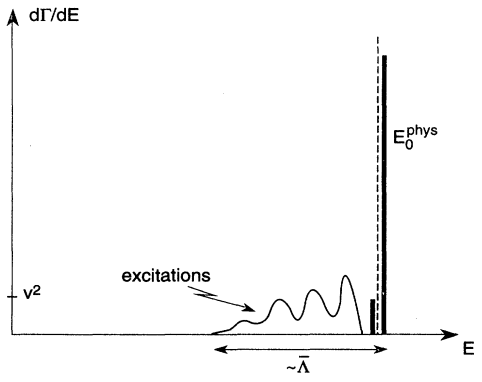

FIG. 2. A qualitative picture of the spectrum $d\Gamma/dE_\phi$ in the $H_Q \to \phi X_q$ with $O(\vec{v}^{\,2})$ terms included. The monochromatic line of the quark transition $Q \to \phi q$ (the dashed line at $E = E_0$) is dual to the physical line corresponding to the elastic decay $H_Q \to H_q\phi$ at $E = E_0^{\rm phys}$ plus a shoulder due to the transitions to the excited states $H_Q \to \phi H_q^*$. The height of the shoulder is $\sim \vec{v}^{\,2}$. Hard gluons are neglected.

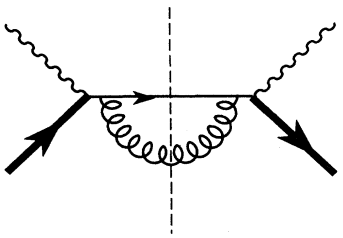

FIG. 3. The diagram responsible for the one-gluon correction in the energy distribution $d\Gamma(Q \to \phi q +$ gluon$)/dE_\phi$.

longing to the soft gluon medium in order to avoid double counting. The separation between soft and hard gluons is achieved by explicitly introducing a normalization point μ. The value of μ should be large enough to justify a small value for $\alpha_s(\mu)$. On the other hand, we would like to choose μ as small as possible. The possible choice is to have μ proportional to $\Lambda_{\rm QCD}$, but with a constant of proportionality that is much larger than unity:

$$\mu = C\Lambda_{\rm QCD}, \quad C \gg 1.$$

Then we draw a line: To the left of $E_0 - \mu$ the gluon is considered to be hard, to the right soft. Of course, the consistent introduction of the infrared renormalization point μ requires that the purely perturbative corrections to the weak vertex, even if they happen to be infrared convergent, have to be calculated using this explicit cutoff as well (see, e.g., the discussion in Ref. [28]), which is almost never done in practice. The corresponding modifications will be discussed below in Sec. VII.

Let us discuss now the sum rule corresponding to the first moment of $E_0^{\rm phys} - E$, i.e., an analogue of Eq. (61) with radiative corrections now included. Since our main purpose in this section is methodical, we will limit ourselves to the first order in $\overline{\Lambda}$ and the second order in v. A qualitative sketch of how $d\Gamma/dE$ looks like is presented in Fig. 4. Then the prediction for I_1 can obviously be rewritten as

$$I_1 = \int_0^{E_0^{\rm phys}} (E_0^{\rm phys} - E)\frac{d\Gamma}{dE}dE = \Gamma_0 \frac{1}{2}v_0^2 \left[\overline{\Lambda}(\mu) + v_0^{-2}\int_0^{E_0-\mu} \frac{16\alpha_s}{9\pi}\frac{E^3}{E_0 m_Q^2}dE\right], \quad (67)$$

where by definition

$$\overline{\Lambda}(\mu) = \int_{E_0^{\rm phys}-\mu}^{E_0^{\rm phys}} \frac{2}{v_0^2}\frac{1}{\Gamma_0}\frac{d\Gamma}{dE}(E_0^{\rm phys} - E)\,dE. \quad (68)$$

Without the radiative tail the prediction for I_1 could be obtained by integrating the theoretical expression (47) over a very narrow domain near E_0. Clearly, there is no way to switch off the radiative corrections in QCD: One has to deal with the perturbative and nonperturbative contributions simultaneously. The introduction of the parameter μ thus becomes mandatory. Equation (68) then can provide us with one possible physical definition of $\overline{\Lambda}(\mu)$ (among others) relating this quantity to an integral over a physically measurable spectral density. One may rephrase this statement as follows. Since quarks are permanently confined, the notion of the heavy quark mass becomes ambiguous. To eliminate this ambiguity one must explicitly specify the procedure of measuring "the heavy quark mass." Any conceivable procedure will necessarily involve a cutoff parameter μ much in the same way as the procedure defined above, and then $\overline{\Lambda}(\mu) = M_Q - m_Q(\mu)$. In the "most inclusive" procedure when one does not try to separate out any kind of effects, one integrates the tail to the kinematical bound $E \simeq m_Q - m_q$ and, therefore, obtains $\overline{\Lambda}$ normalized at the scale of energy release.

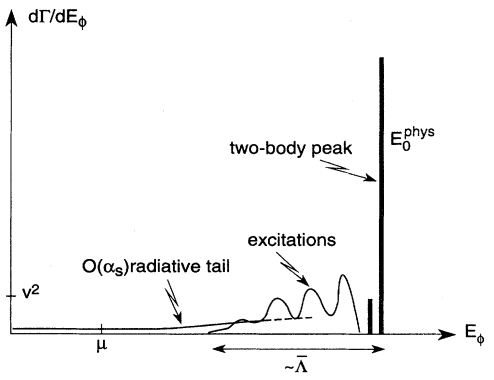

FIG. 4. A sketch of the energy spectrum $d\Gamma/dE_\phi$ with the $O(\alpha_s)$ radiative tail included.

Since this question is very important let us look at it from a slightly different angle. It had widely been believed that $\overline{\Lambda}$ can be defined as a universal constant. The standard definition, being applied to our example, would involve three steps: (i) Take the radiative perturbative tail to the left of the shoulder and extrapolate it all the way to the point $E = E_0$; (ii) subtract the result from the measured spectrum; (iii) integrate the difference over dE with the weight function $(E - E_0^{\rm phys})$. The elastic peak drops out and the remaining integral is equal to $\Gamma_0(\vec{v}^{\,2}/2)\overline{\Lambda}$. It is quite clear that this procedure cannot be carried out consistently—there exists no unambiguous way to extrapolate the perturbative tail too close to $E_0^{\rm phys}$, the end point of the spectrum.[7] Our procedure, with the normalization point μ introduced explicitly, is free from this ambiguity. We will further comment on this issue in Sec. V A where we discuss the possibility of measuring $\overline{\Lambda}(\mu)$ in the inclusive $B \to X_c l\nu$ decays.

In practice, the μ dependence of $\overline{\Lambda}(\mu)$ may turn out to be rather weak. This is the case if the spectral density is such as shown in Fig. 4, where the contribution of the first excitations (lying within $\sim \Lambda_{\rm QCD}$ from $E_0^{\rm phys}$) is numerically much larger than the radiative tail representing (at least in the sense of duality) high excitations. It is quite clear that if the physical spectral density resembles that of Fig. 4 and μ = several units $\times \Lambda_{\rm QCD}$, the running $\overline{\Lambda}(\mu)$ is rather insensitive to the particular choice of μ. As known from QCD sum rules it is just this situation which occurs for the standard quark and gluon condensates (the so-called practical version of Wilson's OPE).

Still, even if the μ dependence of $\overline{\Lambda}$ is numerically weak, conceptually it is impossible to define $\overline{\Lambda}$ in the limit $\mu \to 0$. Physically it is quite clear from the discussion presented above. This consideration can be thought

[7]The standard routine corresponds to using the literal perturbative expression for this tail with the nonrunning strong coupling (for one-loop calculations) or accounting for the first term in the expansion of $\alpha_s(k)$ in terms of $\alpha_s(m)$ (in the two-loop ones).

of as an illustration to a more formal argument presented recently [28, 29].

If one replaces E in the numerator of Eq. (66) by E_0 (the nonrelativistic approximation), one arrives at the formula obtained previously in Ref. [24]. Notice that the total prediction for I_1 is, of course, μ independent. Equation (67) shows how $\overline{\Lambda}(\mu)$ changes under the variation of the normalization point:

$$\delta\overline{\Lambda} = \delta\mu \frac{16}{9}\frac{\alpha_s(\mu)}{\pi}. \tag{69}$$

The numerical coefficient in front of $\delta\mu$ is slightly different from that found in Ref. [28] (16/9 versus $2\pi/3$). The reason is obvious—we use here a different procedure for defining $\overline{\Lambda}(\mu)$ compared to that suggested in [28]: Introducing a gluon mass λ "switches off" the perturbative tail in a "soft" rather than "hard" way at $\mu = \lambda$. The fact itself of the presence of a linear (in μ) renormalization is obviously common for all proper procedures. In other words, running of $\overline{\Lambda}$ is not logarithmic, but power like. This fact can be traced back to the mixing between the operators $\bar{Q}(ivD)Q$ and $\bar{Q}Q$ established in [28]. Numerically $\delta\overline{\Lambda} \sim 0.1\,\mathrm{GeV}$ if μ changes from 1 to 1.5 GeV, i.e., $\delta\overline{\Lambda}(\mu)/\overline{\Lambda}(\mu) \sim 0.2$.

By the same token, using the third sum rule, Eq. (62), one can give a physical definition of $\mu_\pi^2(\mu)$,

$$I_2 = \int_0^{E_0^{\rm phys}} (E_0^{\rm phys} - E)^2 \frac{d\Gamma}{dE} dE$$
$$= \Gamma_0 \frac{1}{3} v_0^2 \left[\mu_\pi^2(\mu) + v_0^{-2} \int_0^{E_0-\mu} \frac{8\alpha_s}{3\pi} \frac{E^3(E_0 - E)}{E_0 m_Q^2} dE \right], \tag{70}$$

where, by definition,[8]

$$\mu_\pi^2(\mu) = 3\Gamma_0^{-1} v_0^{-2} \int_{E_0^{\rm phys}-\mu}^{E_0^{\rm phys}} \frac{d\Gamma}{dE} (E_0^{\rm phys} - E)^2\, dE. \tag{71}$$

The μ dependence is then obtained as

$$\delta\mu_\pi^2 \sim \frac{4\alpha_s}{3\pi}\delta\mu^2 \sim 0.1\,\mathrm{GeV}^2 \text{ if } \delta\mu^2 \sim 1\,\mathrm{GeV}^2.$$

Because the integral of the perturbative tail does not depend on the particular heavy flavor hadron in the initial state, this equation is equivalent to the corresponding power mixing of the kinetic energy operator with the leading one:

$$\frac{d}{d\mu^2}\left[\bar{Q}\,(\vec{\pi})^2\,Q\right]_\mu = \frac{4\alpha_s(\mu)}{3\pi}\left[\bar{Q}Q\right]_\mu, \tag{72}$$

where the numerical value of the coefficient quoted above $(4/3\pi)$ refers to this particular way of introducing the renormalization point.

The fact that the operators $\left[\bar{Q}\,(\vec{\pi})^2\,Q\right]_\mu$ and $\left[\bar{Q}Q\right]_\mu$ mix with each other, and the mixing is quadratic in μ, has been noted some time ago in lattice calculations (G. Martinelli raised this question in a private discussion with M.S. in 1991; see [39]).

[8] We hope that this rather clumsy notation will not cause confusion: μ_π^2 is the matrix element of the kinetic energy operator while μ (with no subscript) is a normalization point.

E. Sum rules for the form factors at zero recoil

We continue investigating our toy model with the aim of establishing sum rules for the form factors at zero recoil. These sum rules will allow us to find corrections to the elastic form factor at *zero recoil* in terms of inelastic contributions.

The starting idea is to consider the kinematical point $\vec{q} = 0$ and q_0 close to $\Delta M = M_B - M_D$. In other words we abandon the case $q^2 = 0$ and turn to $q^2 \neq 0$ where q is now the momentum transfer carried away by our hypothetical ϕ quantum. Since $\vec{q} = 0$, the ground state B meson cannot decay into a P-wave state. Unlike the case of the Bjorken sum rule there is no "external" vector $\vec{v}$. The fact that we are at zero recoil implies that inelastic contributions are suppressed as Λ^2/m^2, i.e., produce effect of the same order of magnitude as the corrections to the elastic form factor.

For $\vec{q} \neq 0$ the strength of the contribution of the excited resonance is proportional to $\vec{q}^{\,2}/M_D^2$ [21]. In our kinematics $\vec{q} = 0$ and the only relevant velocity is provided by the primordial motion of the c quark inside D (more exactly, it is the "difference" between the quark motions in B and D that counts). The limit $\vec{q} \to 0$ is obtained, qualitatively, if instead of the "external" velocity we use that of the primordial motion:

$$\vec{q}^{\,2}/M_D^2 \to \Lambda_{\rm QCD}^2/M_D^2.$$

This analogy nicely illustrates why the residues of the excited resonances are proportional to $\Lambda_{\rm QCD}^2/M_D^2$ and match the picture of Fig. 4 where the contribution from the excited resonances is proportional to $\vec{q}^{\,2}/M_D^2$.

Let us sketch how the sum rules at zero recoil emerge technically. To this end we again consider the T product (43) sandwiched between the B meson state. The contribution of a given hadronic state in this amplitude is proportional to

$$\frac{1}{(M_Q - q_0)^2 - M_q^2} = \frac{1}{(\Delta M - q_0)(\Delta M - q_0 + 2M_q)}. \tag{73}$$

It is convenient to introduce the variable

$$\epsilon = \Delta M - q_0. \tag{74}$$

We are interested in singularities of the amplitude at $\epsilon \ll M_q$. The second pole in Eq. (73) at $\epsilon = -2M_q$ is a reflection of a distant singularity corresponding to the transition $B + \bar{D} \to \phi^*$ where the asterisk marks the virtual ϕ quantum. We assume, as usual, the hierarchy $\Lambda_{\rm QCD} \ll \epsilon \ll m_{Q,q}$. Correspondingly, we will expand in $\Lambda_{\rm QCD}/\epsilon$ and in ϵ/m. In particular, the factor responsible for the second pole in Eq. (73) becomes

$$\frac{1}{\Delta M - q_0 + 2M_q} = \frac{1}{2M_q}\left(1 - \frac{\epsilon}{2M_q} + \cdots\right).$$

The second- and higher-order terms in the brackets can be omitted since they lead to a nonsingular (at $\epsilon = 0$) expression; the nonsingular expressions have no imaginary part and are irrelevant for our purposes.

Next one considers the theoretical expression for the quark level transition operator:

$$-\hat{T} = \bar{Q}\frac{1}{(m_Q - m_q - q_0)(m_Q + m_q - q_0) + (\pi^2 + 2m_Q\pi_0 - 2q_0\pi_0)}Q. \tag{75}$$

Our task is to expand the transition operator in $\Lambda_{\rm QCD}/\epsilon$ and in $\epsilon/m_{Q,q}$ and then to compare the terms singular in $1/\epsilon$ in this theoretical expansion with the phenomenological expression obtained in the language of the resonance saturation. A technical point which deserves mentioning right at the beginning is a mismatch in the definitions of ϵ. The theoretical expression (75) is phrased in terms of the quark mass differences without reference to mass differences of mesons. Since we would like to get the sum rules written in terms of the physical excitation energies (measured from the mass of the lowest-lying meson state), we have to express Eq. (75) in terms of ΔM rather than Δm before expanding it. As we will see shortly, for our purposes it is necessary to keep all effects through order Λ^3/m^3; those of order Λ^4/m^4 and higher can be neglected.

The expansion for the meson mass in inverse powers of the heavy quark mass for spinless quarks considered here is given in Eq. (33). It is worth remembering that ρ^3 defined in Eq. (34) is the second-order perturbation in $1/m_Q$, and as such is positive definite. Using this mass formula one gets

$$\Delta M = \Delta m\left(1 - \frac{\langle\vec{\pi}^2\rangle_0}{2m_Qm_q} + \rho^3\frac{m_Q + m_q}{4m_Q^2m_q^2}\right), \tag{76}$$

where the following shorthand notation is used:

$$\langle\vec{\pi}^2\rangle_0 \equiv \langle H_Q|\bar{Q}\vec{\pi}^2Q|H_Q\rangle_{m_Q=\infty}.$$

Substituting Eq. (76) into Eq. (75) and expanding in Λ/ϵ and $\epsilon/m_{Q,q}$ using Eq. (35) we arrive at

$$-\langle H_Q|\hat{T}|H_Q\rangle = \frac{1}{\epsilon}\frac{1}{2m_q}\left(\langle\bar{Q}Q\rangle - \frac{\langle\bar{Q}\vec{\pi}^2Q\rangle}{2m_q^2}\right) + \frac{1}{\epsilon^2}\rho^3\frac{(\Delta m)^2}{8m_Q^2m_q^3}, \tag{77}$$

plus terms of higher order in $1/\epsilon$ and in $\Lambda_{\rm QCD}$.

This theoretical expression is to be confronted now with what one obtains from saturating the amplitude at hand with meson poles. From Eq. (73) it is not difficult to see that

$$-\langle H_Q|\hat{T}|H_Q\rangle = \frac{M_Q}{2m_qm_Q}\sum_{i=0,1,\ldots}\left(\frac{1}{\epsilon}F_i^2 + \frac{1}{\epsilon^2}\epsilon_iF_i^2 + \cdots\right), \tag{78}$$

where F_i is the form factor for the transition $H_Q \to H_q^i$ induced by the vertex (39),

$$\langle H_q^{(i)}|\bar{q}Q|H_Q\rangle = \left(\frac{M_QM_q^{(i)}}{m_Qm_q}\right)^{\frac{1}{2}}F_i, \tag{79}$$

$M_q^{(i)}$ is the mass of the ith state ($M_q^{(0)} = M_q$ is the mass of ground state), and

$$\epsilon_i = M_q^{(i)} - M_q^{(0)} \tag{80}$$

is the excitation energy of the ith state; all form factors are taken at the zero recoil point, where the meson produced in the transition $Q \to q$ is at rest in the rest frame of H_Q.

Comparing Eqs. (78) and (77) and using Eq. (32) we find that

$$\sum_{i=0,1,\ldots}F_i^2 = 1 - \frac{1}{2M_Q}\langle H_Q|2m_Q\bar{Q}\vec{\pi}^{\,2}Q|H_Q\rangle\left(\frac{1}{2m_Q^2} + \frac{1}{2m_q^2}\right) + O\left(\frac{1}{m^4}\right) \tag{81}$$

and

$$\sum_{i=1,\ldots}\epsilon_iF_i^2 = \rho^3\frac{(\Delta m)^2}{4m_Q^2m_q^2} + O\left(\frac{1}{m^3}\right). \tag{82}$$

Note that the elastic pole gives no contribution in Eq. (82). Hence we conclude that the residues of the excited states are proportional to Λ^2. Transferring then all excited states in Eq. (81) to the right-hand side we observe that the square of the elastic form factor receives corrections of order Λ^2, of local nature as well as due to excited states. Since all excitation energies and all residues are positive an obvious inequality holds:

$$\sum_{i=1,2,\ldots}F_i^2 < \frac{1}{\epsilon_1}\rho^3\frac{(\Delta m)^2}{4m_Q^2m_q^2}, \tag{83}$$

where ϵ_1 is the excitation energy of the first excited state. Therefore

$$1-\frac{1}{2M_{H_Q}}\langle H_Q|2m_Q\bar{Q}\vec{\pi}^{\,2}Q|H_Q\rangle\left(\frac{1}{2m_Q^2}+\frac{1}{2m_q^2}\right)-\frac{\rho^3}{4\epsilon_1}\frac{(\Delta m)^2}{m_Q^2m_q^2}$$

$$<F_0^2<1-\frac{1}{2M_{H_Q}}\langle H_Q|2m_Q\bar{Q}\vec{\pi}^{\,2}Q|H_Q\rangle\left(\frac{1}{2m_Q^2}+\frac{1}{2m_q^2}\right), \quad (84)$$

where F_0 is the form factor of the "elastic" transition $H_Q \to H_q$.

It is worth noting that the explicit form of the corrections, in particular to the first sum rule (81) and, therefore, the "local" terms $\propto \langle\vec{\pi}^2\rangle$ in Eq. (84), depends on the structure of the "weak" current considered and refers to the case of the scalar vertex. Should we use the vector current, coefficients in the sum rules would take a form leading to $F_0 = 1$ at $m_q = m_Q$ in accordance with the exact conservation of the vector current for equal masses.

Concluding this section let us mention a convenient computational device. It is helpful to let the initial quark mass go to infinity and retain corrections only in $1/m_q$. In this way one removes nonperturbative corrections originating in the initial state. The results referring to finite m_Q can be simply reconstructed at the very end. On the other hand, in the opposite limit, $m_q \gg m_Q$ one suppresses nonperturbative effects in the final state; in this way it is convenient to obtain relations for static hadronic quantities.

Similar sum rules at zero recoil in real QCD will be discussed in Sec. IV B.

IV. REAL QCD

We proceed now to discuss the sum rules emerging in QCD for processes of the type $B \to X_c l\nu$. It is clear that the approximation $m_b - m_c \ll m_{b,c}$ does not apply in this case; one can still reach the SV limit, however, by using the fact that q^2 is not necessarily zero in this transition (from now on q is the momentum of the lepton pair). Indeed, if q^2 is close to its maximal value,

$$q^2_{\max} = (M_B - M_D)^2,$$

the D meson velocity is small. At the maximal value of q^2 the velocity vanishes. The velocity $\vec{v} = -\vec{q}/E_D$ is related to q^2 as

$$\frac{E_D - M_D}{M_D} = \frac{1}{\sqrt{1-\vec{v}^{\,2}}} - 1 = \frac{(M_B - M_D)^2 - q^2}{2M_BM_D}, \quad (85)$$

whereas, at the quark level,

$$\frac{1}{\sqrt{1-\vec{v}^{\,2}}} - 1 = \frac{(m_b - m_c)^2 - q^2}{2m_bm_c}. \quad (86)$$

For a sizable fraction of events measured in the semileptonic B decays the values of q^2 are such that these events actually do belong to the SV limit (i.e., v is small). An indirect proof of the relevance of the SV limit to the inclusive semileptonic B decays comes from the fact that about 65% of the total semileptonic rate is given by the "elastic" transitions to D and D^* [40]. The analysis below is carried out under the assumption that the right-hand side in Eq. (85) is small. Even though we do not assume that $m_b - m_c \ll m_{b,c}$ both quarks, b and c, will be treated as heavy, $m_{b,c} \gg \Lambda_{\rm QCD}$.

Needless to say, the proximity of q^2 to $q^2_{\max}$ can be realized in different ways; for instance, one can put $\vec{q} = 0$—this is especially convenient if we are interested in the zero recoil point—or one can keep $\vec{q} \neq 0$, but small and study the terms proportional to $\vec{q}^{\,2}$. This yields a practically realizable method of measuring $\overline{\Lambda}(\mu)$ in the semileptonic decays $B \to X_c e\nu$. We will consider first the simplest sum rule for the total decay width analogous to Eq. (81) in Sec. III E. Surprisingly, this analysis produces a lower bound on the deviation from unity in the $B \to D^*$ elastic form factor at zero recoil which does not quite agree with previous estimates obtained by a different method [41] (see also Ref. [37]). The result is of paramount importance for the experimental determination of $|V_{cb}|$, the Cabibbo-Kobayashi-Maskawa (CKM) matrix element, from the exclusive decay $B \to D^* e\nu$. Then we turn to an analogue of Voloshin's sum rule which appears to be a promising tool for extracting $\overline{\Lambda}(\mu)$.

The relation for $\overline{\Lambda}$ as the quantity measuring the mass difference between the heavy flavor hadron and the heavy quark has been obtained in Ref. [15] by analyzing the heavy quark distribution function appearing in the SV limit for the final state quark. It is convenient to rewrite the expressions obtained in Ref. [15] as follows.

Consider, first, Eq. (75) of Ref. [15]:

$$\frac{1}{2M_{H_Q}}\mathrm{Im}\,\langle H_Q|\hat{T}|H_Q\rangle = \frac{\pi}{4m_Q^2\overline{\Lambda}}\left[\delta(x)\left(1-\frac{1}{3}\frac{\vec{q}^{\,2}}{m_Q^2}\int dy y^{-2}G(y)\right)+\delta'(x)\left(\frac{q_0\langle\vec{\pi}^2\rangle}{2m_Q^2\overline{\Lambda}}+\frac{1}{3}\frac{\vec{q}^{\,2}}{m_Q^2}\int dy y^{-1}G(y)\right)\right.$$

$$\left.+\frac{1}{3}\frac{\vec{q}^{\,2}}{m_Q^2}x^{-2}G(x)\right], \quad (87)$$

where G is the temporal distribution function defined there. The parameter x in this expression measures the energy with respect to the quark boundary E_0, rather than to $E_0^{\rm phys}$:

$$x = \frac{E - E_0}{\bar{\Lambda}}, E_0 = \frac{m_b^2 - m_c^2 + q^2}{2m_b}.$$

It is clear then that the δ' term in Eq. (87) merely shifts the argument of the δ term to its physical value, $x^{\rm phys} = (E - E_0^{\rm phys})/\bar{\Lambda}$.

Limiting ourselves to the effects linear in $\bar{\Lambda}$ we get, as a consequence of this requirement,

$$\int_{-\infty}^{1} \frac{dx}{x} G(x) = -\frac{3}{2}. \tag{88}$$

This sum rule constraining the temporal distribution function is interesting in itself. We can transform it further using its integral representation, Eq. (76) of Ref. [15]:

$$G(y) = \frac{1}{2\pi\bar{\Lambda}} \int dt e^{-iyt\bar{\Lambda}} e^{im_b t} \langle B|\bar{b}(t=0,\vec{x}=0)\,\pi_i T\, e^{-i\int_0^t A_0(t',0)dt'} \pi_i\, b(t,\vec{x}=0)\,|B\rangle \tag{89}$$

[the factor $\exp(im_b t)$ accounts for the explicit time dependence associated with the rest energy m_b, which has not been factored out here]. Multiplying this relation by y^{-1} and integrating over y we arrive at

$$\bar{\Lambda} = \frac{2}{3}\int_0^\infty \frac{d\xi}{\xi} \int \frac{dt}{2\pi} e^{-it\xi} \frac{e^{im_b t}}{2M_B} \langle B|\bar{b}(t=0,\vec{x}=0)\pi_i T\, e^{-i\int_0^t A_0(t')dt'} \pi_i\, b(t,\vec{x}=0)|B\rangle. \tag{90}$$

Thus $\bar{\Lambda}$ is expressed in terms of a *nonlocal* correlator of heavy quark currents (see also Ref. [28]). Let us remember that the definition (36) involves local operators made of light fields. Its normalization point dependence can easily be traced formally through the properties of the path-ordered exponent which, being a field operator, requires specification of the normalization point for the gauge fields. In Sec. III D we have illustrated the renormalization point dependence using the saturation of the correlator by intermediate states, which is equivalent to calculating the correlator via dispersion relations.

It will be demonstrated in Sec. VI that matrix elements of the type entering Eq. (90) have a simple interpretation in ordinary quantum mechanics which uses the first-quantized language. Here we illustrate it for the case under consideration. Equation (90) is written in the heavy quark limit where all corrections $1/m$ are neglected. To this accuracy the time-ordered exponent is nothing but the (time) correlation function of a nonrelativistic heavy quark in the external gluon field:

$$\langle Q(0)\bar{Q}(x)\rangle_{A_\mu} = T\, e^{-i\int_0^t A_0(t')dt'} \delta^3(\vec{x})\, e^{im_Q t}. \tag{91}$$

Then the matrix element in Eq. (90) can be written as

$$\frac{e^{im_b t}}{2M_B} \langle B|\bar{b}(t=0,\vec{x}=0)\pi_i T\, e^{-i\int_0^t A_0(t',0)dt'} \pi_i\, b(t,\vec{x}=0)|B\rangle$$

$$= \int d^3x \frac{1}{2M_B} \langle B|\bar{b}\,\pi_i\, Q(t=0,\vec{x})\bar{Q}\,\pi_i b\,(t,\vec{x}=0)|B\rangle e^{-i(m_Q - m_b)t}, \tag{92}$$

where Q is a heavy quark, b or c. This form is convenient for the transition to quantum-mechanical notations [see Eq. (157) in Sec. VI] in which this matrix element becomes

$$\langle B|\pi_i(t=0)\pi_i(t)\,|B\rangle_{\rm QM} = \sum_n |\langle B|\pi_i|n\rangle_{QM}|^2 e^{i\epsilon_n t}. \tag{93}$$

Here π is the momentum operator of the heavy quark; matrix elements with the subscript QM are to be understood in the quantum-mechanical sense; namely, only the states with zero total momentum are considered and the nonrelativistic normalization is assumed; ϵ_n are the excitation energies of the heavy meson, $\epsilon_n \simeq M_{B_n} - M_B \simeq M_{D_n} - M_D$. Integrating the correlator over t and ξ according to Eq. (90) we thus get

$$\bar{\Lambda} = \frac{2}{3}\sum_n \frac{|\langle B|\,\pi_i\,|n\rangle_{\rm QM}|^2}{E_n - M_B}. \tag{94}$$

The expressions (90) and (94) are relations that give an alternative to Eq. (36), a phenomenological definition of $\bar{\Lambda}$ in terms of measurable correlators. In Eq. (94) the sum over n extends up to excitation energy $\epsilon_n = \mu$ and in this way yields $\bar{\Lambda}(\mu)$.

The quantum-mechanical derivation of Eq. (94) as a relation between the mass of the heavy quark and the total mass of the quantum-mechanical system containing the heavy quark will be given in the Appendix.

A. Sum rules at zero recoil: Generalities

Our analysis of the $B \to X_c e\nu$ problem at zero recoil will parallel the corresponding consideration carried out in the toy model of Sec. III E. The presence of spin is a technicality which can easily be incorporated. As a matter of fact, all formulas necessary for derivation of the first and second sum rules exist in the literature; we

will borrow them from Ref. [10] as well as all relevant notations.

The point $\vec{q}=0$ represents zero recoil. Then the transition operator

$$\hat{T}_{\mu\nu}=i\int e^{-iqx}dxT\{j_\mu^\dagger(x)j_\nu(0)\}, \tag{95}$$

for the $b\to c$ transitions, with $j_\mu=\bar{c}\Gamma_\mu b$, $\Gamma_\mu\equiv\gamma_\mu(1-\gamma_5)$, can be presented (in the tree approximation) in the form of the expansion

$$\hat{T}_{\mu\nu}=\bar{b}\Gamma_\mu(k_0\gamma_0+m_c+\not{\pi})\frac{1}{(m_c^2-k_0^2)}\times\sum_{n=0}^{\infty}\left(\frac{2k_0\pi_0+\pi^2+(i/2)\sigma G}{m_c^2-k_0^2}\right)^n\Gamma_\nu b, \tag{96}$$

with

$$k_0=m_b-q_0.$$

The operator product expansion (96) is justified provided that

$$\Lambda_{\rm QCD}\ll|m_c-k_0|.$$

In other words, the expansion (96) is a series in $\Lambda_{\rm QCD}/(m_c-k_0)$. At the same time, apart from the poles $1/(m_c-k_0)$ it obviously contains powers of $1/(m_c+k_0)$ which develop "distant" singularities at $k_0=-m_c$. We want these singularities corresponding to the propagation of the antiquark $\bar{c}$ to be indeed distant so that the dispersion integrals we will be dealing with do not stretch up to these $\bar{c}$ containing states. To this end we must impose a second condition on $|m_c-k_0|$, namely,

$$|m_c-k_0|\ll m_c\,.$$

Once this condition is imposed we expand $\hat{T}_{\mu\nu}$ in powers of $(m_c-k_0)/m_c$ and $\Lambda_{\rm QCD}/(m_c-k_0)$. The result is then ordered with respect to the powers of $1/(m_c-k_0)$. The terms nonsingular in (m_c-k_0) are irrelevant and can be discarded. Each particular power $1/(m_c-k_0)^{n+1}$ in the expansion leads to a sum rule with the weight function $\propto(m_c-k_0)^n$.

Let us sketch the basic elements of the procedure in some detail. We start from a series in $1/(m_c-k_0)$. The next step is averaging $\hat{T}_{\mu\nu}$ over the B meson state:

$$h_{\mu\nu}=\frac{1}{2M_B}\langle B|\hat{T}_{\mu\nu}|B\rangle. \tag{97}$$

The hadronic tensor $h_{\mu\nu}$ consists of five different covariants [4, 10]:

$$h_{\mu\nu}=-h_1g_{\mu\nu}+h_2v_\mu v_\nu-ih_3\epsilon_{\mu\nu\alpha\beta}v_\alpha q_\beta+h_4q_\mu q_\nu+h_5(q_\mu v_\nu+q_\nu v_\mu). \tag{98}$$

Moreover, the invariant hadronic functions h_1–h_5 depend on two variables q_0 and q^2 or q_0 and $|\vec{q}|$. For $\vec{q}=0$ only one variable survives, and only two of five tensor structures in $h_{\mu\nu}$ are independent.

Each of these hadronic invariant functions satisfies a dispersion relation in q_0,

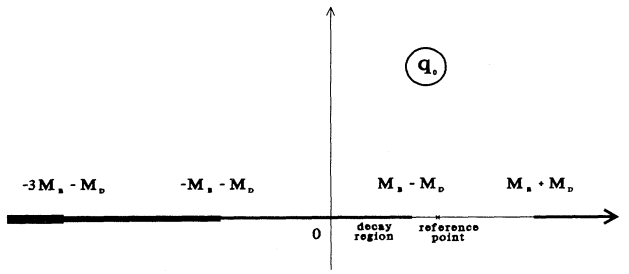

FIG. 5. The cuts of $h_i(q_0)$ in the complex q_0 plane at $\vec{q}=0$. The point $q_0\simeq M_B+M_D$ marks the beginning of three cuts: Two of them originate at the mass of B_c meson and the third starts at M_B+M_D. The three cuts can be differentiated considering $\vec{q}\neq 0$. Only a part of one cut at $0<q_0<M_B-M_D$ is relevant to the decay process under consideration. The masses are not shown in a real scale.

$$h_i(q_0)=\frac{1}{2\pi}\int\frac{w_i(\tilde{q}_0)d\tilde{q}_0}{\tilde{q}_0-q_0}+\text{polynomial}, \tag{99}$$

where w_i are observable structure functions:

$$w_i=2\,\mathrm{Im}\,h_i.$$

This dispersion representation for h_i assumes, as usual, that the integral on the right-hand side runs over all cuts that the transition operator may have. The general structure of the cuts in the complex q_0 plane is rather sophisticated; the issue deserves a special discussion since it is not always properly understood.

The structure of the cuts of the functions $h_i(q_0)$ is shown in Fig. 5. The part accessible in the decay channel of the B mesons covers the interval $[0,M_B-M_D]$. The dispersion integral (99) can be written as a sum of two integrals:

$$h_i(q_0)=\frac{1}{2\pi}\int_0^{M_B-M_D}\frac{w_i(\tilde{q}_0)d\tilde{q}_0}{\tilde{q}_0-q_0}+\frac{1}{2\pi}\int_A\frac{w_i(\tilde{q}_0)d\tilde{q}_0}{\tilde{q}_0-q_0}, \tag{100}$$

where the domain A consists of two subdomains $q_0<0$ (A_1) and $q_0>M_B+M_D$ (A_2). For real decays we are interested only in the first integral since the second one, rather than describing the $B\to X_ce\nu$ decay, refers to other physical processes. The subdomain A_1 actually describes a similar $b\to c$ amplitude, yet with negative q_0, and can be called the lower cut (at $q_0<-M_B-M_D$ it contains also the q^2 cut and for $q_0<-3M_B-M_D$ the u-channel contribution is present as well). The integral over A_2, on the other hand, will be referred to as the integral over the distant cuts. Two kinds of problems are encountered in evaluating the total dispersion integral. The first one emerges due to the fact that for real decay kinematics one has only $q_0>0$; therefore, say, for calculating the total width one does not have the integral over the whole physical cut, but needs to consider the smaller interval without the subdomain A_1. The corresponding problems of separating the contribution of the same type

of intermediate states, but at different values of q_0, are usually referred to as "local" duality. In the context of the present paper this is, however, not very important. For in our sum rules, from a purely theoretical point of view, it does not matter whether a particular transition can be measured in a real experiment or not; e.g., the lower bounds we will discuss rely only on the positivity of the corresponding transition probabilities.

There is generally another complication associated with the integral over the distant cuts (in particular, subdomain A_2) corresponding to quite different intermediate states. The problem of isolating these contributions can be generically referred to as "global" duality.

Both contributions thus represent a contamination for real decays. Fortunately, this contamination is irrelevant for our analysis. Indeed, to address the contamination due to the "lower" cut, let us choose the "reference" point of q_0 between the cuts (see Fig. 5), close to $M_B - M_D$,

$$q_0 = M_B - M_D - \epsilon, \tag{101}$$

where ϵ is a negative number,

$$m_D \gg -\epsilon \gg \Lambda_{\rm QCD}.$$

When calculating the functions h_i at the quark level from the operator product expansion we get a similar dispersion relation for the OPE coefficients (with the meson masses replaced by the quark ones). We then use duality concepts in identifying the physical relevance of these cuts. Local duality of QCD means that there is a one-to-one correspondence between the part of the OPE coefficients originating from the lower cut and the corresponding hadronic contribution in the phenomenological (hadronic) representation of h_i. The validity of local duality can be verified by itself by choosing q_0 in the complex plane close to the particular remote cut, the lower one for the case at hand. Therefore, we can systematically discard the contribution of that cut simultaneously, in the theoretical expression for h_i and in the "phenomenological" saturation. In this way we arrive at the relations

$$\frac{1}{2\pi}\int_0^\mu w_i^{\rm quark}\epsilon^n d\epsilon = \frac{1}{2\pi}\int_0^\mu w_i\epsilon^n d\epsilon, \tag{102}$$

with $\Lambda_{\rm QCD} \ll \mu \ll M_B$, where ϵ is the same variable as defined in Eq. (101), but on the cut it is positive. It represents the excitation energy. If, additionally, $\mu < M_B - M_D$, the region of integration in Eq. (102) lies completely in the physical domain. The left-hand side of Eq. (102) includes the perturbative corrections as well as the powerlike nonperturbative terms. The local duality we have invoked to discard the contribution from the "lower" cut has an accuracy of the type $\exp\{-\text{const} \times \mu/\Lambda_{\rm QCD}\}$.

An analogous analysis can be repeated almost verbatim for the contribution of the distant cuts to address the question of "global duality." This duality is even more transparent physically and is explicit in all perturbative calculations and for calculations in "soft" external fields. On the other hand, in principle, its accuracy is generally determined by the similar factor depending on m_c, namely, $\exp\{-\text{const} \times m_c/\Lambda_{\rm QCD}\}$, because $2m_c$ determines the distance to the remote cut. It is important to remember that it is the ratio $m_c/\Lambda_{\rm QCD}$, not $m_b/\Lambda_{\rm QCD}$, that enters, at least at zero recoil. In the real world $m_c/\Lambda_{\rm QCD}$ numerically is not so large, and since the constant in the exponent is unknown, one may be afraid of an insufficient accuracy of the duality for D's. At present theory provides us with no clues as to the value of the constant in the exponential; the degree of possible violations of the duality should be established empirically. With this caveat in mind we still believe that heavy quark expansion must work well in $B \to X_c e\nu$.

B. Sample calculation

After these more general remarks we return to concrete calculations of the theoretical part of the sum rules. To find $w_i^{\rm quark}$ we take the discontinuity of the transition operator (96),

$$\begin{aligned}\frac{1}{i}{\rm disc}\,\hat{T}_{\mu\nu} = 2\pi\bar{b}\Gamma_\mu(k_0\gamma_0 + m_c + \not{k})\\ \times\sum_{n=0}^{\infty}\frac{1}{n!}\delta^{(n)}(k_0 - m_c)\\ \times\frac{[2k_0\pi_0 + \pi^2 + (i/2)\sigma G]^n}{(m_c + k_0)^{n+1}}\Gamma_\nu b,\end{aligned} \tag{103}$$

where $\delta^{(n)}(x) = (d/dx)^n\delta(x)$. Using this discontinuity in the left-hand side of Eq. (102) it is not difficult to calculate all moments of the structure functions in the leading approximation. For definiteness we will first consider w_1^{AA}, the first structure function [see Eq. (98)] in the transition induced by the axial current. (The corresponding "elastic" contribution is $B \to D^*$.) At zero recoil h_1^{AA} is singled out merely by considering the spatial components of the axial-vector current:

$$h_1^{AA} = \frac{1}{3}h_{kk}^{AA} \quad (i = 1, 2, 3).$$

All other structure functions and other currents can be treated in a similar manner.

Technically it is convenient to carry out the computation in two steps: First one obtains auxiliary moments convoluted with the quark value of ϵ,

$$\tilde{\epsilon} = m_b - m_c - q_0 = k_0 - m_c, \tag{104}$$

and then, at the second stage, these results are converted into the true ϵ moments. Notice that in the case at hand ϵ should be defined as

$$\epsilon = M_B - M_{D^*} - q_0,$$

since the lowest-lying state produced is D^*, not D. Then, according to Eqs. (26)–(29),

$$\delta_A \equiv \epsilon - \tilde{\epsilon} = (M_B - m_b) - (M_{D^*} - m_c) = -(\mu_\pi^2 - \mu_G^2)\left(\frac{1}{2m_c} - \frac{1}{2m_b}\right) - \frac{2}{3m_c}\mu_G^2 + \left(-\rho_D^3 + \rho_{\pi\pi}^3 + \rho_S^3\right)\left(\frac{1}{4m_c^2} - \frac{1}{4m_b^2}\right) + \left(\rho_{LS}^3 - \rho_{\pi G}^3 - \rho_A^3\right)\left(\frac{1}{12m_c^2} + \frac{1}{4m_b^2}\right) + O\left(\frac{\Lambda_{\rm QCD}^4}{m^3}\right), \tag{105}$$

where μ_π^2 and μ_G^2, defined in Eq. (1), are the asymptotic expectation values of the kinetic and chromomagnetic operators. In the estimates of the neglected terms m generically denotes both m_c and m_b. The mass shift δ_A determines the difference in the threshold energy between the real hadrons at zero recoil and the quark mass difference; i.e., it is a direct zero recoil analogue of Δ in Eq. (53).

Let us define moments of the structure functions w_i as

$$I_n^{(i)} = \frac{1}{2\pi}\int \epsilon^n w_i(\epsilon) d\epsilon. \tag{106}$$

In the leading nontrivial approximation we get, for the moments of w_1,

$$\begin{aligned}
I_0^{(1)AA} &= \left\langle \bar{b}\gamma_0 \left\{ 1 - \frac{\vec{\pi}^2 + \vec{\sigma}\cdot\vec{B}}{4}\left(\frac{1}{m_c^2} + \frac{1}{m_b^2} + \frac{2}{3m_c m_b}\right) + \frac{1}{3m_c^2}\vec{\sigma}\cdot\vec{B} \right\} b \right\rangle + O(\Lambda_{\rm QCD}^3/m^3), \\
I_1^{(1)AA} &= \left\langle \bar{b}\left\{ (\vec{\pi}^2 + \vec{\sigma}\cdot\vec{B})\left(\frac{1}{2m_c} - \frac{1}{2m_b}\right) - \frac{2}{3m_c}\vec{\sigma}\cdot\vec{B} \right.\right. \\
&\quad \left.\left. + [-(\vec{D}\cdot\vec{E}) + 2\vec{\sigma}\cdot\vec{E}\times\vec{\pi}]\left(\frac{1}{8m_c^2} - \frac{1}{8m_b^2}\right) - \frac{1}{3m_c^2}\vec{\sigma}\cdot\vec{E}\times\vec{\pi} \right\} b \right\rangle + \delta_A + O(\Lambda_{\rm QCD}^4/m^3), \\
I_2^{(1)AA} &= \frac{1}{3}\left\langle \bar{b}\left\{ \left[(\vec{\pi}^2 + \vec{\sigma}\cdot\vec{B})\left(\frac{1}{2m_c} - \frac{1}{2m_b}\right)\sigma_k + \frac{1}{2m_c}[\sigma_k, \vec{\sigma}\cdot\vec{B}]\right] \right.\right. \\
&\quad \left.\left. \times \left[\sigma_k(\vec{\pi}^2 + \vec{\sigma}\cdot\vec{B})\left(\frac{1}{2m_c} - \frac{1}{2m_b}\right) - \frac{1}{2m_c}[\vec{\sigma}\cdot\vec{B}, \sigma_k]\right] \right\} b \right\rangle - \delta_A^2 + O(\Lambda_{\rm QCD}^5/m^3), \\
I_n^{(1)AA} &= \frac{1}{3}\left\langle \bar{b}\left\{ \left[(\vec{\pi}^2 + \vec{\sigma}\cdot\vec{B})\left(\frac{1}{2m_c} - \frac{1}{2m_b}\right)\sigma_k + \frac{1}{2m_c}[\sigma_k, \vec{\sigma}\cdot\vec{B}]\right] \pi_0^{n-2} \right.\right. \\
&\quad \left.\left. \times \left[\sigma_k(\vec{\pi}^2 + \vec{\sigma}\cdot\vec{B})\left(\frac{1}{2m_c} - \frac{1}{2m_b}\right) - \frac{1}{2m_c}[\vec{\sigma}\cdot\vec{B}, \sigma_k]\right] \right\} b \right\rangle + O(\Lambda_{\rm QCD}^{n+3}/m^3),
\end{aligned} \tag{107}$$

where $\vec{B}$ denotes the chromomagnetic field, and the last equation refers to $n > 2$. It is worth noting that to this order in $1/m_Q$ only the chromomagnetic field appears explicitly for $n \geq 2$ although in the original expansion (103) the field tensor G contains both chromomagnetic and chromoelectric components; the chromoelectric field cancels out. The underlying reason for that will become apparent shortly (see Sec. VI). Averaging $\langle\cdots\rangle$ over the initial hadron state is understood as $(2M_B)^{-1}\langle B|\cdots|B\rangle$ in all expressions in the right-hand side. In derivation of the above predictions for the moments we used the QCD equations of motion (13).

The structure of the solution of Eqs. (107) for the excitation function $w_1(\epsilon)$ is quite transparent. The first equation tells us that the sum of all probabilities is equal to unity up to small corrections $= O(\Lambda_{\rm QCD}^2/m^2)$. On the other hand, all higher moments I_n explicitly start with terms of order $\Lambda_{\rm QCD}^n(\Lambda_{\rm QCD}/m)^2$. Since the scale for the excitation energies ϵ_i is given[9] by $\Lambda_{\rm QCD}$, one immediately concludes that the probabilities of transitions to the excited states all scale like $\Lambda_{\rm QCD}^2/m^2$. To saturate the first sum rule one then needs state(s) which do not contribute to the higher moments I_n and are produced with practically unit probability; the only way to satisfy this constraint is to saturate by the final states D and D^* with the masses

[9]Effectively the same refers even to the thresholds associated with the $D^{(*)}$ + pion(s) in the chiral limit which, strictly speaking, have no excitation gap. This is true due to the fact that the corresponding amplitudes are proportional to the pion momentum; they can produce only chiral logarithms and do not change powers of mass in the analysis; see Ref. [42].

$$M_B - (m_b - m_c),$$

up to corrections vanishing in the limit $m_b, m_c \to \infty$. Moreover, these "elastic" transition amplitudes must be equal to unity up to terms inversely proportional to the *square* of the heavy quark masses, a fact observed originally in Ref. [23] and now known as Luke's theorem [43]. Although the expressions above are derived for the axial-vector current, similar results are valid for the vector current as well, where the lowest-lying final state is D, not D^*. In this way one obtains also the statement of the heavy flavor symmetry in the spectrum of hadrons.

The second equation of Eqs. (107) corresponding to $n = 1$ has a special status. When written in terms of the excitation energies $\tilde{\epsilon}$, counted from the quark threshold, the right-hand side does contain terms of order $O(\Lambda^2_{\rm QCD}/m)$ which are expressed in terms of μ_π^2 and μ_G^2; on the other hand, the structure of the solution described above implies that the contribution of the excited states in the phenomenological part is only of the order of $\Lambda^3_{\rm QCD}/m^2$. Therefore, this leading term is to be completely saturated by the elastic peak that resides not at $\tilde{\epsilon} = 0$ but is rather shifted by the amount $\delta_A = \epsilon - \tilde{\epsilon}$. This condition unambiguously determines the leading, $O(\Lambda^2_{\rm QCD}/m)$, correction to the masses of the heavy flavor hadrons which has been anticipated in Eq. (105). Similarly, to order $O(\Lambda^3_{\rm QCD}/m^2)$ the sum rule for the first moment yields the "local" $1/m^2$ term in the effective Hamiltonian, Eq. (21), and expresses the nonlocal correlators ρ^3 as the sum over inelastic probabilities.

It is not difficult to obtain a general expression for the function w_1^{AA} as it emerges from its moments $I_n^{(1)AA}$. The inelastic part of $w(\epsilon)$ appears at the $1/m^2$ level and is given by the Fourier transform of the time-dependent correlation functions of the leading operators in the effective Lagrangian:

$$\epsilon^2 w_1^{AA}(\epsilon) = \frac{1}{3}\int \frac{dt}{2\pi} e^{-it\epsilon} \frac{1}{2M_B} \left\langle B|\bar{b}(t=0,\vec{x}=0)\left[(\vec{\pi}^2 + \vec{\sigma}\cdot\vec{B})\left(\frac{1}{2m_c} - \frac{1}{2m_b}\right)\sigma_k + \frac{1}{2m_c}[\sigma_k, \vec{\sigma}\cdot\vec{B}]\right]\right.$$
$$\left. \times T\, e^{-i\int_0^t A_0(t',0)dt'}\left[\sigma_k(\vec{\pi}^2 + \vec{\sigma}\cdot\vec{B})\left(\frac{1}{2m_c} - \frac{1}{2m_b}\right) - \frac{1}{2m_c}[\sigma_k, \vec{\sigma}\cdot\vec{B}]\right] b(t,\vec{x}=0)|B\right\rangle'. \quad (108)$$

The prime here means subtraction of the "elastic" contribution (see the expression for I_2 and the discussion below).

Equation (108) has a very transparent quantum-mechanical meaning. It corresponds to the following picture. At time $t = 0$ the Hamiltonian of the system under consideration is suddenly changed by adding a perturbation

$$\left(\frac{1}{2m_c} - \frac{1}{2m_b}\right)(\vec{\pi}^2 + \vec{\sigma}\cdot\vec{B}). \quad (109)$$

Simultaneously the original wave function is changed due to the spin flip. This effect is represented by the term $(-1/2m_c)[\sigma_k, \vec{\sigma}\cdot\vec{B}]$ where k marks the spatial component of the axial-vector current. Had we considered the time component of the vector currents, the latter term would be absent, and the entire perturbation would reduce to Eq. (109). At time t the perturbation is switched off. The second-order term in this combined perturbation yields us the excitation probabilities. This interpretation brings us very close to Lipkin's quantum-mechanical formalism which will be discussed in some detail in Sec. VI.

Transitions induced by the time component of the vector current are given by the following combination of the structure functions:

$$w_0^{VV} \equiv -w_1^{VV} + w_2^{VV} + q_0^2 w_4^{VV} + q_0 w_5^{VV}. \quad (110)$$

The moments $I_n^{(0)VV}$ of w_0^{VV} defined in analogy to Eq. (106) look even simpler:

$$I_0^{(0)VV} = \left\langle \bar{b}\gamma_0\left\{1 - \frac{\vec{\pi}^2 + \vec{\sigma}\cdot\vec{B}}{4}\left(\frac{1}{m_c} - \frac{1}{m_b}\right)^2\right\} b\right\rangle + O(\Lambda^3_{\rm QCD}/m^3),$$
$$I_1^{(0)VV} = \left\langle \bar{b}\left\{(\vec{\pi}^2 + \vec{\sigma}\cdot\vec{B})\left(\frac{1}{2m_c} - \frac{1}{2m_b}\right) + [-(\vec{D}\cdot\vec{E}) + 2\vec{\sigma}\cdot\vec{E}\times\vec{\pi})]\left(\frac{1}{8m_c^2} - \frac{1}{8m_b^2}\right)\right\} b\right\rangle + \delta_V + O(\Lambda^4_{\rm QCD}/m^3),$$
$$I_2^{(0)VV} = \left\langle \bar{b}\left\{\left[(\vec{\pi}^2 + \vec{\sigma}\cdot\vec{B})\left(\frac{1}{2m_c} - \frac{1}{2m_b}\right)\right] \times \left[(\vec{\pi}^2 + \vec{\sigma}\cdot\vec{B})\left(\frac{1}{2m_c} - \frac{1}{2m_b}\right)\right]\right\} b\right\rangle - \delta_V^2 + O(\Lambda^5_{\rm QCD}/m^3),$$
$$I_n^{(0)VV} = \left\langle \bar{b}\left\{\left[(\vec{\pi}^2 + \vec{\sigma}\cdot\vec{B})\left(\frac{1}{2m_c} - \frac{1}{2m_b}\right)\right]\pi_0^{n-2}\left[(\vec{\pi}^2 + \vec{\sigma}\cdot\vec{B})\left(\frac{1}{2m_c} - \frac{1}{2m_b}\right)\right]\right\} b\right\rangle + O(\Lambda^{n+3}_{\rm QCD}/m^3). \quad (111)$$

Note that for the vector current the lowest charmed state is the D meson and, therefore, ϵ is to be defined now as $\epsilon = M_B - M_D - q_0$; the value of δ_V then differs from δ_A, and the second sum rule (the expression for the first moment) yields now

$$\delta_V \equiv \epsilon - \tilde{\epsilon} = (M_B - m_b) - (M_D - m_c) = -(\mu_\pi^2 - \mu_G^2)\left(\frac{1}{2m_c} - \frac{1}{2m_b}\right) - (\rho_D^3 + \rho_{LS}^3 - \rho_{\pi\pi}^3 - \rho_{\pi G}^3 - \rho_S^3 - \rho_A^3)\left(\frac{1}{4m_c^2} - \frac{1}{4m_b^2}\right) + O(\Lambda_{\rm QCD}^4/m^3), \tag{112}$$

instead of Eq. (105).

It is not difficult to trace a straightforward parallel between the derivation given above, Eq. (108), and the analysis of the effects of the heavy quark motion in the SV limit carried out in Ref. [15] that led us to the temporal distribution function. In that case, however, inelastic excitations appeared proportional to the velocity squared of the final quark, and it was the correlator of the spatial momentum operators $\vec{\pi}$ that emerged.

A remark is in order here. Obviously, one can, in principle, calculate higher-order power corrections to the sum rules (107) and (111). It is possible to see that, in what concerns the inelastic contributions to the structure functions, $w(\epsilon)$ will order by order reproduce the successive terms of the ordinary quantum-mechanical perturbation theory corresponding to both next-order iterations of the leading terms as well as to perturbations representing subleading power operators in the effective heavy quark Lagrangian; the local relativistic corrections in currents appear as well. Similarly, one can consider the sum rules in the case of nonzero recoil. This correspondence to the perturbation theory in quantum mechanics will be discussed in Sec. VI.

To conclude this section, let us say a few words elucidating why this general analysis based on the sum rules is important although it might seem to lead to no new results beyond the picture of the standard perturbation theory in quantum mechanics.

First, we shall see in the subsequent sections that this approach allows one to obtain useful bounds on the transition amplitudes and on the "static" matrix elements governing nonperturbative corrections.

Second, it demonstrates in a transparent and unambiguous way the necessity of introducing the infrared normalization point in addressing nonperturbative effects and clarifies both the physical meaning and the qualitative trend of the μ dependence. This is important in view of the existing misinterpretation of this problem in applications of HQET.

The last, but not the least, aspect is that we clearly see here that such assumptions about QCD as the validity of global duality discussed above, which sometimes are naively believed to be specific only for the OPE-based approach to inclusive transitions, are actually the most general requirement inherent to any consistent consideration. If for some particular physical reason the quark-hadron correspondence used above were modified—for example, there were a sizable "leak" in dispersion integrals from distant cuts—it would immediately lead to new contributions on the right-hand side of the sum rules; this would result in physically observable corrections not contained in the expansions that can be obtained in HQET.

C. Bound on the form factor at zero recoil from the sum rules

The general theory developed above is applied in this section to derive a lower bound on the deviation of the "elastic" form factor at zero recoil from unity. To this end we analyze the first sum rule.

We consider, for definiteness, transitions generated by the axial-vector current,

$$A_\mu = \bar{c}\gamma_\mu\gamma_5 b,$$

i.e., the transitions of the type $B \to D^*$ and $B \to$ excitations of the vector mesons. Practically they are most important in the exclusive approach to a determination of $|V_{cb}|$. These transitions are induced by the axial-vector current A_μ and as was mentioned above it is most convenient to focus on the spatial component of this current. For the spatial components of the current only h_1 survives.

To get the first sum rule at zero recoil we use the first equation of Eqs. (107), which to order $1/m^2$ reads as

$$\frac{1}{2\pi}\int d\epsilon\, w_1^{AA}(\epsilon) = 1 - \frac{1}{3}\frac{\mu_G^2}{m_c^2} - \frac{\mu_\pi^2 - \mu_G^2}{4}\left(\frac{1}{m_c^2} + \frac{1}{m_b^2} + \frac{2}{3m_c m_b}\right). \tag{113}$$

The chromomagnetic parameter is known experimentally:

$$\mu_G^2 \simeq \frac{3}{4}(M_{B^*}^2 - M_B^2) = 0.37\,\text{GeV}^2.$$

The sum rule stemming from Eq. (113) obviously takes the form

$$F_{B\to D^*}^2 + \sum_{\rm excit} F_{B\to \text{excitations}}^2 = 1 - \frac{1}{3}\frac{\mu_G^2}{m_c^2} - \frac{\mu_\pi^2 - \mu_G^2}{4}\left(\frac{1}{m_c^2} + \frac{1}{m_b^2} + \frac{2}{3m_c m_b}\right), \tag{114}$$

where the sum on the left-hand side runs over all excited states and all form factors are taken at the zero recoil point. This is a perfect analogue of Eq. (81). The form factor $B \to D^*$ at zero recoil is defined as

$$\langle B|A_k|D^*\rangle = \sqrt{4M_B M_{D^*}}\,F_{B\to D^*} e_k^*,$$

where e_k is the polarization of D^* meson.

If all terms $O(\Lambda_{\rm QCD}^2)$ are switched off, higher states

cannot be excited at zero recoil—only the elastic $B \to D^*$ transition survives—and we arrive at the well-known result that

$$F_{B\to D^*} = 1 \quad \text{(zero recoil)} ,$$

the statement of the heavy quark (or Isgur-Wise [44]) symmetry first noted in the SV limit in Ref. [23] (see also [45]). Including $O(\Lambda^2_{\rm QCD})$ terms we start exciting higher states; all transition form factors squared are proportional to $\Lambda^2_{\rm QCD}/m^2$. Simultaneously the form factor of the elastic transition shifts from unity.

Both power corrections on the right-hand side are negative. What is crucial is the fact that the contribution of the excited states is strictly positive. Transferring them to the right-hand side we arrive at the lower bound

$$1 - F^2_{B\to D^*} > \frac{1}{3}\frac{\mu_G^2}{m_c^2} + \frac{\mu_\pi^2 - \mu_G^2}{4}\left(\frac{1}{m_c^2} + \frac{1}{m_b^2} + \frac{2}{3m_c m_b}\right). \tag{115}$$

Since $\mu_\pi^2 > \mu_G^2$ (see below), we find that the (absolute value of) the deviation of the elastic form factor $F_{B\to D^*}$ from unity at zero recoil is definitely larger than

$$\frac{M^2_{B^*} - M^2_B}{8m_c^2} \sim 0.035.$$

As was mentioned, the second term on the right-hand side of Eq. (115) is also positive, $\mu_\pi^2 > \mu_G^2$. We will rederive this inequality within the framework of sum rules themselves in the next section. Previously it was obtained in this form in [46] where the quantum-mechanical argument of Ref. [15] was extended. The quantum-mechanical line of reasoning is applicable at a low normalization point.

The inequality $\mu_\pi^2 > \mu_G^2$ is in perfect agreement with the most refined QCD sum rule calculation of μ_π^2 [47] which leads to $\mu_\pi^2 = 0.6 \pm 0.1\ \text{GeV}^2$; the recent updated analysis along the same line of calculations yielded (see Ref. [48])

$$\mu_\pi^2 = 0.5 \pm 0.1\ \text{GeV}^2.$$

If this estimate is accepted, then the second term in Eq. (115) amounts to $\sim 1/2$ of the first one, and the lower bound for the deviation becomes $1 - F^2_{B\to D^*} > 0.1$. The actual deviation is probably twice as large. First, the sum rule derived above neglects perturbative α_s corrections. The first-order correction to the elastic form factor was calculated in Ref. [23]. If, in zeroth order in α_s, the $\bar b \gamma_\mu \gamma_5 c$ axial-vector vertex at zero recoil is unity, the first-order correction renormalizes it to

$$\eta_A = 1 + \frac{\alpha_s}{\pi}\left(\frac{m_b + m_c}{m_b - m_c}\ln\frac{m_b}{m_c} - \frac{8}{3}\right).$$

Numerically $\eta_A \simeq 0.97$ if one uses α_s normalized at the point $\mu = \sqrt{m_c m_b}$ (for a recent discussion see Ref. [49]). For the axial-vector current the perturbative correction is negative, so that unity in Eq. (113) is replaced by approximately[10] 0.95. Then, the contribution of the excited states in Eq. (113) is strictly positive, and this also reduces $F^2_{B\to D^*}$. This contribution may be as large as, roughly, the power correction on the right-hand side. An estimate of the excited state contribution supporting this statement is given in Ref. [42] where a more detailed numerical discussion of all corrections is given. Notice that in our approach the excited state contribution replaces a nonlocal contribution of Ref. [37].

We conclude that $1 - F^2_{B\to D^*}$ is definitely larger than 0.1, somewhat beyond the window obtained in Ref. [41]. The phenomenological impact of this observation is discussed in Ref. [42].

Let us note that in order to use the second sum rule similar to Eq. (84) we would need to know $O(\Lambda^3_{\rm QCD})$ terms both in the transition operator and in the relation between ΔM and Δm. The corresponding expressions are provided by Eq. (105) and the second equation of Eqs. (107); however, the relevant hadronic parameters ρ^3_D, ρ^3_{LS} and nonlocal correlators $\rho^3_{\pi\pi,\pi G,S,A}$ are not known yet.

In a very similar way one can obtain a bound and an estimate for the vector form factor of the $B \to D$ transition $F_{B\to D}$ at zero recoil. Here only the timelike component of the current contributes, and for this reason the full semileptonic decay amplitude is proportional to the lepton masses. Therefore this mode is not advantageous; the corresponding form factor is measurable (in principle) at zero recoil in the $B \to D + \tau\nu_\tau$ decays. Taking $\Gamma_\mu = \Gamma_\nu = \gamma_0$ we obtain for this case the sum rule

$$F^2_{B\to D} + \sum_{\rm excit} F^2_{B\to \rm excit} = 1 - \frac{\mu_\pi^2 - \mu_G^2}{4}\left(\frac{1}{m_c} - \frac{1}{m_b}\right)^2. \tag{116}$$

Perturbative corrections also differ and now look like [23]

$$1 + \frac{\alpha_s}{\pi}\left(\frac{m_b + m_c}{m_b - m_c}\ln\frac{m_b}{m_c} - 2\right). \tag{117}$$

The corrections in the case of the vector form factor obviously vanish at $m_b = m_c$, as they have to in view of the exact conservation of the current in this limit. Numerically therefore they are expected to be smaller for vector transitions than for axial-vector ones.

It is worth mentioning that the excitation probabilities entering the sum rules (113) and (116) are generated separately by the axial-vector or the vector current, respectively, but not by the V-A current that directly produces the experimental widths. Actually at zero velocity transfer the axial-vector and vector currents cannot interfere. Therefore for the V-A semileptonic transitions into massless leptons one has just to add to Eq. (113) (assuming that no final state identification is attempted) the contribution of the $\bar c \gamma_k b$ current. The sum rule for this current

[10] For our numerical estimate we use somewhat larger number than the literal value of η_A^2 above; see discussion in Sec. VII and Eq. (189).

is obtained in the next subsection. Combining the two sum rules one gets

$$F^2_{B\to D^*} = 1 - \frac{\mu_\pi^2 - \mu_G^2}{3m_b m_c} - \int_{\epsilon > M_{D^*} - M_D} d\epsilon \frac{w_1^{V\text{-}A}(\epsilon)}{2\pi}, \tag{118}$$

where the last term representing the inelastic contribution is expressed via the differential semileptonic width at zero recoil:

$$\frac{w_1^{V\text{-}A}(\epsilon)}{2\pi} = \frac{8\pi^3}{G_F^2 |V_{cb}|^2 q_0^2} \frac{1}{|\vec{q}|} \frac{d^2\Gamma_{\rm SL}}{d\vec{q}^{\,2}\, dq_0}\bigg|_{\vec{q}=0, q_0 = M_B - M_D - \epsilon} \tag{119}$$

(q is the momentum of the lepton pair in the process). Equation (118) is much less useful as an upper bound because the main part of the correction to the elastic form factor must now come from the presently unknown contribution of excitations.

D. Lower bound on μ_π^2

Reversing the line of reasoning that led to the upper bound on the form factors (the lower bound on $1-F$), we can exploit the very same idea to get a constraint on the matrix elements μ_G^2 and μ_π^2. At zero recoil there are only two independent structure functions for the correlator of the V-A currents; similar functions can be introduced for other weak vertices as well. Choosing a particular current one projects out certain combinations of the structure functions. It is important that for the Hermitian-conjugated currents in the two weak vertices one gets a definitely positive structure function expressed as a sum over certain transition probabilities. If an appropriate channel is found where the elastic peak is kinematically absent, the theoretical side of the corresponding sum rule will contain no unity term, and start with the leading $1/m^2$ corrections given just by a linear combination of μ_G^2 and μ_π^2. The "phenomenological" side, a sum over the excited states, is positive definite. In this way we arrive at a constraint on the linear combination of μ_G^2 and μ_π^2 at hand.

It is not difficult to find a specific example. Indeed, let us consider the vector current $\bar{q}\gamma_\mu Q$ and in particular its *spacelike* components. It is obvious that $-(1/3)h_{kk}^{VV} = -h_1^{VV}$ is populated only by the excited states: If the initial state H_Q is the ground state pseudoscalar, the final one H_q^* must be the axial-vector meson, nondegenerate with H_Q in the symmetry limit. (In the quark model language one would say that H_q^* is a P-wave state.) Using the results of Ref. [10] [Eq. (A1)] we find

$$-h_1^{VV}(q_0) = \frac{1}{\epsilon}\left[\frac{\mu_\pi^2 - \mu_G^2}{4}\left(\frac{1}{m_Q^2} + \frac{1}{m_q^2} - \frac{2}{3m_Q m_q}\right) + \frac{\mu_G^2}{3}\frac{1}{m_q^2}\right]. \tag{120}$$

The expression in the square brackets is equal to the sum over excitations:

$$\frac{\mu_\pi^2 - \mu_G^2}{4}\left(\frac{1}{m_Q^2} + \frac{1}{m_q^2} - \frac{2}{3m_Q m_q}\right) + \frac{\mu_G^2}{3}\frac{1}{m_q^2} = \sum_i F^2_{H_Q\to H_q^*}, \tag{121}$$

and, hence, is always positive, for any values of m_Q and m_q. Being interested in the "static" properties of the initial state only, it is convenient, according to the comment in the end of Sec. III E, to consider the theoretical limit $m_q \gg m_Q$ where only initial state effects survive.[11] Requiring positivity of Eq. (120) at $m_q \gg m_Q$ we conclude that

$$\mu_\pi^2 - \mu_G^2 = 4m_Q^2 \left\{\sum_i |F_{H_Q\to H_q^*}|^2\right\}_{m_q=\infty} \tag{122}$$

and, therefore,

$$\mu_\pi^2 > \mu_G^2. \tag{123}$$

This is literally the same inequality that has been obtained previously [46] in the quantum-mechanical language along the lines suggested in Ref. [15]. The argument presented above can be viewed as a consistent and transparent field-theoretic reincarnation. To be fully consistent, according to the general discussion of Sec. III D, one must introduce the cutoff in the "phenomenological" integral over the decay probabilities from the upper side of excitation energies ϵ. It is most important that the integral in the right-hand side of Eq. (122),

$$\sum_{\epsilon_i \le \mu} |F_{H_Q\to H_q^*}|^2,$$

is positive for *any* normalization point μ, and therefore ensures the validity of the inequality (123) for operators normalized at arbitrary values of μ, provided that the normalization point is consistently introduced in this particular way. For high enough μ the contribution of the excited states is given by perturbative expressions and the μ dependence is explicitly calculable; of course, for very large μ the inequality becomes trivial. We will discuss the issue in more detail in Sec. VII.

It is instructive to note that the zero recoil sum rule for the $V_k \times V_k$ transitions (or similar ones where there is no elastic peak) provides us with the direct way of determining the evolution of the kinetic energy operator. One proceeds here along the same line of reasoning as has been outlined in Sec. III D where we considered the small velocity kinematics with $\vec{q} \neq 0$. In this case the elastic peak identically vanishes for purely kinematical reasons. Therefore the only possible impact of introducing the infrared normalization point (say, via the gluon

[11] The very same bound can be obtained by considering, say, the correlator of two $i\gamma_5$ currents. In this case one gets directly the difference $\mu_\pi^2 - \mu_G^2$ with the coefficient $(1/m_c + 1/m_b)^2$.

"mass") can emerge from the gluon emission probability.

Technically, we can consider the relation (121) in perturbation theory. Then the hadron state H_Q must be replaced by the heavy quark Q, and the excited states are q + gluon states. For such an initial state the expectation value of operator σG vanishes and the sum rule (121) converts into the relation for the perturbative part of μ_π^2:

$$\frac{(\mu_\pi^2)_{\rm pert}}{4}\left(\frac{1}{m_Q^2}+\frac{1}{m_q^2}-\frac{2}{3m_qm_Q}\right)=\frac{1}{3}\sum_k\int_{\omega<\mu}\frac{d^3k}{2\omega(2\pi)^3}\frac{1}{4m_qm_Q}|\langle qg|\bar{q}\gamma_kQ|Q\rangle|^2; \tag{124}$$

the sum over gluon polarizations is implied. Following the procedure used throughout this paper, we have introduced the renormalization point μ as the upper limit for excitation energy. The calculation of the amplitude in the right-hand side is very simple:

$$\langle qg|\bar{q}\vec{\gamma}Q|Q\rangle = g_s\,\bar{q}\,t^a\left[\vec{\epsilon}^{\,a}\left(\frac{1}{2m_q}+\frac{1}{2m_Q}\right)-i\,(\vec{\epsilon}^{\,a}\times\vec{\sigma})\left(\frac{1}{2m_q}-\frac{1}{2m_Q}\right)\right]Q. \tag{125}$$

Taking the square of this amplitude and summing over the gluon polarizations $\vec{\epsilon}^{\,a}$ we get the same dependence on the masses m_q and m_Q as in the left-hand side of Eq. (121), which is expected on general grounds. In this way we obtain

$$(\mu_\pi^2)_{\rm pert}=\frac{4\alpha_s}{3\pi}\,\mu^2. \tag{126}$$

This result can be rewritten as

$$\frac{d}{d\mu^2}\bar{Q}\,(i\vec{D})^2\,Q=\frac{4\alpha_s(\mu)}{3\pi}\bar{Q}Q, \tag{127}$$

which coincides with Eq. (72). It is easy to check that the same evolution law is obtained for any suitable current and even for a case when heavy quarks were spinless.

Throughout this paper we have phrased our discussion of real QCD in terms of transitions where initial states were heavy flavor *mesons*, namely, B. It is clear that exactly the same reasoning can be applied for the transitions of heavy baryons, for example, when the initial hadron is Λ_b. The matrix elements are different, of course. In particular, the expectation value of the chromomagnetic operator vanishes for Λ_b. Moreover, contrary to the meson case no nontrivial lower bound on the kinetic term emerges. Therefore it is natural to expect smaller deviations from the symmetry limit for both vector and axial-vector form factors in baryons than for mesons.

An analysis of corrections to $F_{B\to D^*}$ resembling ours in spirit, but not technically, has been carried out recently in Ref. [37]. There $1-F_{B\to D^*}$ is expressed in terms of some local and nonlocal expectations values; the latter are unknown. In our analysis the role of the nonlocal expectation values is played by the contribution of the excited states. What is crucial is that this contribution is always positive. In [37] $1-F_{B\to D^*}$ is found to be positive (good), and a numerical estimate is presented relating $1-F_{B\to D^*}$ to μ_π^2, a parameter that is somewhat more uncertain than μ_G^2. As a matter of fact, the numerical values of μ_π^2 accepted in [37] probably under the influence of some recent claims [50] are in contradiction with the inequality (123).

V. SUM RULES AT $\vec{q}\neq 0$

We leave now the point of zero recoil and discuss the sum rules in the general situation. The first obvious complication is that there are more independent structure functions that enter separately the decay rate for any particular current, and each depends on two variables for which we choose as q_0 and $\vec{q}^{\,2}$. Inelastic processes corresponding to the transition to the states other than D and D^* are now not necessarily suppressed by powers of $\bar{\Lambda}/m_Q$. The nth moments of the structure functions are therefore proportional to $\bar{\Lambda}^n$. If nonperturbative effects are calculated explicitly through $O(\Lambda^2_{\rm QCD})$ terms, we obtain nontrivial corrections for the first three moments with $n=0,1$, and 2.

At $\vec{q}\neq 0$ the variable ϵ is defined as[12]

$$\epsilon = M_B-\sqrt{M_{D^*}^2+\vec{q}^{\,2}}-q_0 \tag{128}$$

and moments of the structure functions are

$$I_n^{(i)}(\vec{q}^{\,2})=\frac{1}{2\pi}\int d\epsilon\,\epsilon^n\,w_i(\epsilon,\,\vec{q}^{\,2}), \tag{129}$$

where i labels the structure function; they can be considered separately for axial-vector current (AA), vector current (VV), and the interference of the two (AV).

Using Eqs. (A1)–(A6) obtained in Ref. [10] and expanding the corresponding hadronic invariant functions h_i in powers of $1/\epsilon$ through terms $1/\epsilon^3$ one arrives at the set of relations which are valid up to terms $O(\Lambda^3_{\rm QCD})$. We present here these relations only for the structure functions $w_{1,2,3}$ which contribute to semileptonic decays with massless leptons (the most general case will be considered in Ref. [51]). We have the following sum rules for the zeroth moments:

[12]When vector current is considered, the lowest state appearing is D; we still use the D^* energy as a reference point, keeping in mind that explicit account for the D contribution in the phenomenological part of the first moments I_1^{VV} is necessary. The fact that the D contribution to the semileptonic width vanishes at $\vec{q}=0$ for massless leptons justifies such a choice.

$$I_0^{(1)AA} = \frac{E_c + m_c}{2E_c} - \frac{\mu_\pi^2 - \mu_G^2}{4E_c^2}\frac{m_c}{E_c}\left[\frac{m_c^2}{E_c^2} + \frac{E_c^2}{m_b^2} + \frac{2}{3}\frac{m_c}{m_b}\right] - \frac{\mu_G^2}{3E_c^2}\frac{m_c}{E_c}\frac{E_c^2 + 3m_c^2}{4E_c^2}, \tag{130}$$

$$I_0^{(2)AA} = \frac{m_b}{E_c}\left\{1 - \frac{\mu_\pi^2 - \mu_G^2}{3E_c^2}\left[2 - \frac{5}{2}\frac{E_c^2}{m_b^2} + \frac{3}{2}\frac{m_c^2}{E_c^2}\right] - \frac{\mu_G^2}{3E_c^2}\left[\frac{1}{2} + \frac{m_c}{m_b} + \frac{3}{2}\frac{m_c^2}{E_c^2}\right]\right\}, \tag{131}$$

$$I_0^{(3)AV} = -\frac{1}{2E_c}\left\{1 - \frac{\mu_\pi^2 - \mu_G^2}{3E_c^2}\left[1 + \frac{3}{2}\frac{m_c^2}{E_c^2}\right] - \frac{\mu_G^2}{2E_c^2}\left[1 + \frac{m_c^2}{E_c^2}\right]\right\}. \tag{132}$$

Expressions for the VV functions are obtained from the axial-vector ones by replacing $m_c \to -m_c$; the structure functions $w^{(1,2)AV}$ and $w^{(3)AA,VV}$ vanish.

The above equations are analogues of the Bjorken sum rule; however, they incorporate nonperturbative effects which appear at the $1/m_Q^2$ level; the corrections are not universal and differ explicitly for different currents and structure functions.

The first moments look like

$$I_1^{(1)AA} = \frac{E_c + m_c}{2E_c}\left\{\frac{\mu_\pi^2 - \mu_G^2}{2E_c}\left[1 - \frac{E_c}{m_b} - \frac{1}{3}\left(1 - \frac{m_c}{E_c}\right)\left(1 + 3\frac{m_c}{E_c} + 2\frac{E_c}{m_b}\right)\right]\right.$$
$$\left. + \frac{\mu_G^2}{2E_c}\left[1 - \frac{2}{3}\frac{m_c}{E_c} + \frac{m_c^2}{E_c^2}\right] + [(M_B - m_b) - (E_{D^*} - E_c)]\right\}, \tag{133}$$

$$I_1^{(2)AA} = \frac{m_b}{E_c}\left\{\frac{\mu_\pi^2 - \mu_G^2}{3E_c}\left[2 - \frac{7}{2}\frac{E_c}{m_b} + \frac{3}{2}\frac{m_c^2}{E_c^2}\right] + \frac{\mu_G^2}{3E_c}\left[\frac{1}{2} - \frac{E_c - m_c}{m_b} + \frac{3}{2}\frac{m_c^2}{E_c^2}\right] + [(M_B - m_b) - (E_{D^*} - E_c)]\right\}, \tag{134}$$

$$I_1^{(3)AV} = -\frac{1}{2E_c}\left\{\frac{\mu_\pi^2 - \mu_G^2}{3E_c}\left[1 - \frac{5}{2}\frac{E_c}{m_b} + \frac{3}{2}\frac{m_c^2}{E_c^2}\right] + \frac{\mu_G^2}{2E_c}\left[1 + \frac{m_c^2}{E_c^2}\right] + [(M_B - m_b) - (E_{D^*} - E_c)]\right\}. \tag{135}$$

At $\vec{q} = 0$ these relations determine $1/m$ terms in the masses of heavy mesons. Their derivatives with respect to $\vec{q}^{\,2}$ near zero recoil give the Voloshin's "optical" sum rule. Here they are obtained with better accuracy for arbitrary, not necessarily small, velocity and incorporate $1/m_Q$ relative corrections. The latter appear to be quite sizable when $\vec{q}$ is not particularly large.

The third sum rules, which are relations for the second moments of the structure functions, are calculated only in the leading nontrivial approximation. They look rather simple and manifestly satisfy the heavy quark symmetry relation [52]

$$\frac{2E_c}{E_c + m_c} I_2^{(1)AA} = \frac{2E_c}{E_c - m_c} I_2^{(1)VV} = \frac{E_c}{m_b} I_2^{(2)AA} = \frac{E_c}{m_b} I_2^{(2)VV} = -2E_c I_2^{(3)AV} = \frac{\mu_\pi^2}{3}\frac{E_c^2 - m_c^2}{E_c^2} + \bar{\Lambda}^2\left(1 - \frac{m_c}{E_c}\right)^2. \tag{136}$$

All higher moments vanish in our approximation. We used above the notation E_{D^*} for the energy of D^* and E_c for the energy of the c quark in the free quark decay:

$$E_{D^*} = \sqrt{M_{D^*}^2 + \vec{q}^{\,2}}, \quad E_c = \sqrt{m_c^2 + \vec{q}^{\,2}}. \tag{137}$$

The quantity

$$(M_B - m_b) - (E_{D^*} - E_c) \simeq \bar{\Lambda}\left(1 - \frac{m_c}{E_c}\right) - (\mu_\pi^2 - \mu_G^2)\left(\frac{1}{2E_c} - \frac{1}{2m_b}\right) - \frac{2\mu_G^2}{3E_c} - \frac{\bar{\Lambda}^2}{2E_c}\left(1 - \frac{m_c^2}{E_c^2}\right) + O\left(\frac{1}{m^2}\right), \tag{138}$$

which enters the first and second moments, is similar to the δ_A we have discussed in the case of zero recoil; at $\vec{q} \neq 0$ however, it is of the order of $\Lambda_{\rm QCD}$.

It is important for the analysis of the inclusive widths that the nonperturbative expansion of the invariant hadronic functions h_i and, therefore, of the moments of the structure functions always run in inverse powers of E_c rather than m_c; it is correlated with the fact that nonphysical singularities in q_0 representing distant cuts are also separated from the physical cut by $\sim E_c$. It implies that one can use the same expansion literally even if $m_c \to 0$ as long as $|\vec{q}| \gg \Lambda_{\rm QCD}$. In turn, this means that at $E_c \sim m_b$, dominating the inclusive width for $m_b \gg m_c$, nonperturbative corrections scale like the inverse square of the mass of the heavier decaying quark [6, 7].

Using the same fact one may hope to decrease $1/m_c$ corrections to the determination of $\bar{\Lambda}$ and μ_π^2 from the second and third sum rules, which are rather significant numerically in the SV limit, considering decay processes

with *large* recoil. A more detailed analysis of the sum rules at nonzero recoil is given in Ref. [51].

A. Second sum rule at $\vec{q} \neq 0$: Measuring $\overline{\Lambda}(\mu)$

We discuss now a few useful applications of the sum rules presented above in the the simpler SV kinematics when we can expand moments in $\vec{q}^{\,2}/m_c^2$.

If at $\vec{q} = 0$ the second sum rule requires knowledge of $O(\Lambda^3_{\rm QCD})$ terms, at $\vec{q} \neq 0$ (i.e., $\Lambda_{\rm QCD} \ll |\vec{q}| \ll M_D$) a nontrivial prediction, an analogue of Voloshin's sum rule [24], arises in order $\Lambda_{\rm QCD}$. Higher-order corrections are written in Eqs. (133)–(135). We need to consider the average value of $q_{0\max} - q_0$. If the value of q^2, rather than $\vec{q}^{\,2}$, is fixed, this quantity is related to $M^2_{X_c}$, the average invariant mass squared of the hadronic system produced.[13] Indeed, then

$$M^2_{X_c} = M_D^2 + 2M_B(q_{0\max} - q_0),$$

$$q_{0\max} = \frac{M_B^2 - M_D^2 + q^2}{2M_B}. \tag{139}$$

If $\vec{q}^{\,2}$ is fixed, then

$$M^2_{X_c} = M_D^2 + 2E_c(q_{0\max} - q_0) + (q_{0\max} - q_0)^2,$$

$$q_{0\max} = M_B - \sqrt{M_D^2 + \vec{q}^{\,2}}; \tag{140}$$

in this case the second moment also contributes, but only at the level of $O(\Lambda^2_{\rm QCD})$. We will use the energy of D^* in what follows, and therefore need to define the corresponding threshold energy $q^*_{0\max}$:

$$q^*_{0\max} = M_B - \sqrt{M_{D^*}^2 + \vec{q}^{\,2}}. \tag{141}$$

The basic idea is the same as that demonstrated in Sec. III D in the toy model: To order $\Lambda_{\rm QCD}$ in the SV limit the weighted integral over the excited states is proportional to $\overline{\Lambda}(\mu)$. In the previous section we considered the point of zero recoil; now we have to shift from this point and consider terms proportional to the square of the c quark velocity. The SV limit will be ensured by choosing $|\vec{q}| \ll M_D$.

Let us assume first that all structure functions in Eq. (98) are known separately. Then it is most convenient to consider the function h_1 for axial-vector current transitions. If we are aiming at effects linear in $\Lambda_{\rm QCD}$, all μ_π^2 and μ_G^2 corrections can be neglected; we can include them, however. Using Eq. (133) and the similar one for the vector current as well as the expansion (138) we get

$$\frac{1}{2\pi}\int_{q_{0\max}-\mu} dq_0 (q^*_{0\max} - q_0) w_1^{V-A} = \frac{\vec{q}^{\,2}}{2M_D^2}\left\{\overline{\Lambda} + \frac{\overline{\Lambda}^2}{m_c} - \frac{\mu_\pi^2}{3m_c} - 2\frac{\mu_\pi^2 - \mu_G^2}{3m_b}\right\} + O\left(\frac{\vec{q}^{\,4}}{M_D^4}, \Lambda^3_{\rm QCD}\right), \tag{142}$$

where the hadronic tensor considered is that induced by the V-A current. In this equation $\vec{q}$ is supposed to be fixed and to be small as compared to m_c^2.

Let us note the explicit presence of the normalization point μ in the lower limit of integration. Contrary to the naive quantum-mechanical description, in real QCD the structure functions do not vanish when the excitation energy becomes larger than a typical hadronic mass scale, but rather contain a long tail associated with the gluon emission. To reiterate the conclusion of Sec. III E: We introduce a normalization point μ in such a way that all frequencies smaller than μ can be considered as "soft" or inherent to the bound state wave function; at the same time $\alpha_s(\mu)$ has to be sufficiently small for the perturbative expansion in $\alpha_s(\mu)/\pi$ to make sense. We then draw a line at $q_0 = q^*_{0\max} - \mu$ (the picture is similar to that of Fig. 4, with $d\Gamma/dE_\phi$ replaced by w_1^{V-A} and E_ϕ by q_0). The integral (142) taken over the range $q^*_{0\max} - \mu$ to $q^*_{0\max}$ represents $(\vec{v}^{\,2}/2)\overline{\Lambda}(\mu)$ modulo corrections of higher order in v and in $\Lambda_{\rm QCD}$. The running mass is then defined as $m_b(\mu) = M_B - \overline{\Lambda}(\mu)$.

Practically it may be not so easy to separate different structure functions from each other, which would require a triple-decay distribution over q_0, q^2, and E_ℓ. A similar prediction can be given for the double-differential distribution in the semileptonic decay when the integral over the energy of the lepton is considered. The explicit form of the sum rule depends again on whether one fixes q^2 or $\vec{q}^{\,2}$ in the process. Below we assume that $\vec{q}^{\,2}$ is kept fixed.

Using the relation

$$\frac{d^2\Gamma}{dq_0 d\vec{q}^{\,2}} = |V_{cb}|^2 \frac{G_F^2}{16\pi^4}|\vec{q}|\left[(q_0^2 - \vec{q}^{\,2})w_1 - \frac{\vec{q}^{\,2}}{3}w_2\right] \tag{143}$$

and Eqs. (133), (134), (136), and (138) one arrives at

$$\int dq_0 (q^*_{0\max} - q_0)\frac{d^2\Gamma}{dq_0 d\vec{q}^{\,2}} = \frac{G_F^2}{8\pi^3}|V_{cb}|^2 (m_b - m_c)^2 \frac{|\vec{q}|^3}{2M_D^2}\left\{\overline{\Lambda} + \frac{\overline{\Lambda}^2}{m_c} - \frac{\mu_\pi^2 - \mu_G^2}{3}\left[\frac{1}{m_c} + \frac{20}{3(m_b - m_c)} + \frac{2}{m_b}\right]\right.$$
$$\left. - \frac{\mu_G^2}{3}\left[\frac{1}{m_c} + \frac{4}{m_b - m_c} - \frac{4m_c}{3(m_b - m_c)^2}\right]\right\} + O\left(|\vec{q}|^5, \Lambda^3_{\rm QCD}\right). \tag{144}$$

[13] The corresponding analysis of the average invariant hadronic mass presented in Ref. [4] was incorrect; see Ref. [11]. The average invariant mass of the final state hadrons is *not* given directly by matrix elements of local heavy quark operators at the level of nonperturbative corrections, contrary to claims in Ref. [4].

Again, one can consider radiative corrections. For simplicity we will neglect all terms that are $\sim \Lambda^2_{\rm QCD}$ and higher (then the difference between $q^*_{0\max}$ and $q_{0\max}$ becomes unimportant), and limit ourselves only to terms $\sim \vec{q}^{\,2}$. Including the radiative tail we have

$$\frac{8\pi^3}{G_F^2|V_{cb}|^2}\int_{|\vec{q}|}^{q_{0\max}} dq_0(q_{0\max}-q_0)\frac{1}{|\vec{q}|}\frac{d^2\Gamma}{dq_0 d\vec{q}^{\,2}}$$

$$=\frac{8\pi^3}{G_F^2|V_{cb}|^2}\left\{\int_{q_{0\max}-\mu}^{q_{0\max}} dq_0(q_{0\max}-q_0)\frac{1}{|\vec{q}|}\frac{d^2\Gamma}{dq_0 d\vec{q}^{\,2}}+\int_{|\vec{q}|}^{q_{0\max}-\mu} dq_0(q_{0\max}-q_0)\frac{1}{|\vec{q}|}\frac{d^2\Gamma}{dq_0 d\vec{q}^{\,2}}\right\}. \quad (145)$$

The first term on the right-hand side, up to the known factor, has the meaning of the running value of $\overline{\Lambda}$ if μ is much smaller than $|\vec{q}|$:

$$\frac{8\pi^3}{G_F^2|V_{cb}|^2}\int_{q_{0\max}-\mu}^{q_{0\max}} dq_0(q_{0\max}-q_0)\frac{1}{|\vec{q}|}\frac{d^2\Gamma}{dq_0 d\vec{q}^{\,2}}=(m_b-m_c)^2\frac{|\vec{q}|^2}{2m_c^2}\,\overline{\Lambda}(\mu). \quad (146)$$

If $q_{0\max}-q_0$ is still sufficiently large, the integrand coincides with the result obtained in the perturbative calculation. To first order in α_s it is given by the probability to emit a gluon with energy $\omega \approx q_{0\max}-q_0$. It looks particularly simple when $q_{0\max}-q_0 \ll |\vec{q}|$. As compared to the case of zero recoil we have discussed previously, now a dipole gluon emission appears which is proportional to $\vec{v}^{\,2}$, and for small ω one can neglect the $1/m$ suppressed amplitude considered before. In the SV limit, when $\vec{q}^{\,2} \ll m_c^2$, the results for the radiative corrections obtained in Sec. III D for the toy model are directly applicable to semileptonic decays. Indeed, the gluon can be emitted (absorbed) either by the color charge of the c quark or by its magnetic moment. Moreover, there is no interference—the gluon emitted by the magnetic moment has to be absorbed by the magnetic moment. As long as we consider gluons with momenta much less than $\vec{q}$, we can disregard the gluon interaction with the magnetic moment. Then we are left with a charge interaction only which is the same for spin-0 bosons and spin-1/2 fermions. As a result, the expression for the radiative correction obtained previously in the toy model is modified in a minimal way, only due to a slightly different kinematics, and for $q_{0\max}-q_0 \ll |\vec{q}|$ we get[14]

$$\frac{8\pi^3}{G_F^2|V_{cb}|^2}\,\frac{1}{|\vec{q}|}\frac{d^2\Gamma_{\rm pert}}{dq_0 d\vec{q}^{\,2}}$$

$$=(m_b-m_c)^2\frac{|\vec{q}|^2}{2m_c^2}\frac{16\alpha_s}{9\pi(q_{0\max}-q_0)}. \quad (147)$$

The above equation shows how the the value of $\overline{\Lambda}$ obtained from the sum rules depends on the parameter μ; it coincides with Eq. (69). If one does not introduce this explicit cutoff, the value of $\overline{\Lambda}$ would generally scale like $\alpha_s m_Q$.

[14]Virtual corrections lead to the term $\sim \delta(q_{0\max}-q_0)$ which vanishes in the moment we consider, but for the zeroth moment in the first sum rule it cancels the logarithmic singularity following from expression (147).

Thus, the sum rules (142) and (146),(147) can be used to elucidate what is actually meant by the heavy quark mass. This question is rather subtle since the heavy quark mass is a purely theoretical parameter which is not directly measurable. On the other hand, it is a very important parameter, crucial in a wide range of questions.

The sum rule (142) or (146) expresses $\overline{\Lambda}$ in terms of an integral over the physically observable quantities. Therefore, we have a suitable phenomenological definition of $\overline{\Lambda}$ and, through this quantity, the heavy quark mass. Of course, both of them depend explicitly on the normalization point.

Although the above equations yield $\overline{\Lambda}(\mu)$ in terms of the excitation distributions which are, in principle, experimentally measurable, this does not mean that it can be easily measured in practice. Needless to say, it has not been measured so far. Still, expression (142), combined with the Bjorken sum rule, implies [52] a lower bound on $\overline{\Lambda}(\mu)$ (see also [24]) which turns out to be quite restrictive:

$$\overline{\Lambda} > 2\Delta_1\,(\rho^2-\tfrac{1}{4}), \quad (148)$$

where Δ_1 is the mass difference between the first excitation of D and the D meson; ρ^2 is the slope of the Isgur-Wise function [20, 21]. Numerically the right-hand side of Eq. (148) is close to 0.5 GeV. Further details can be found in Ref. [52].

B. Third sum rule in the SV kinematics

In a similar manner one can use the third sum rule to relate the kinetic operator to the average value of $(q_{0\,\max}-q_0)^2$. The value of μ_π^2 can be extracted in a model-independent way from a sum rule similar to Eq. (71) if double-differential measurements are used; say, for small velocity events,

$$\mu_\pi^2(\mu)=3\tilde{\Gamma}^{-1}v^{-2}\int_{q_{0\max}-\mu}^{q_{0\max}}\frac{d^2\Gamma}{dq_0 d\vec{q}^{\,2}}(q_{0\max}-q_0)^2 dq_0, \quad (149)$$

where $v=|\vec{q}|/m_c$ or $|\vec{q}|/M_D$ (which particular mass is

used in the denominator does not matter in the approximation considered); the normalization $\tilde{\Gamma}$ is defined as

$$\tilde{\Gamma} = \int_{q_{0\,\max}-\mu}^{q_{0\,\max}} \frac{d^2\Gamma}{dq_0 d\vec{q}^{\,2}}\, dq_0. \tag{150}$$

This determines the value of μ_π^2 normalized at point μ.

Again, similarly to the situation with $\overline{\Lambda}$, one can get a lower bound on μ_π^2 without waiting for measurements of the double-differential distributions. To this end one combines the third sum rule (149) with the Bjorken sum rule and one gets [52]

$$\mu_\pi^2 > 3\Delta_1^2\left(\rho^2 - \frac{1}{4}\right), \tag{151}$$

where the quantities on the right-hand side are the same as in Eq. (148). Numerically the right-hand side is close to 0.5 GeV^2; see Ref. [52].

VI. QUANTUM-MECHANICAL INTERPRETATION

A quantum-mechanical approach to the derivation of the sum rules for the heavy quark transitions has been suggested by Lipkin [27]. The formalism he exploits is similar to that used in the theory of the Mössbauer line shape; it is discussed in detail in Lipkin's textbook [53].

Some of our results can be readily understood within the framework of Lipkin's approach, and, therefore, it seems instructive to provide a dictionary allowing one to translate (where possible) the field-theoretic consideration into the language of quantum mechanics.

As a matter of fact the expressions for the moments of the distribution functions (i.e., averages of powers of the excitation energy) obtained in Sec. IV B have a very transparent quantum-mechanical interpretation, which was already mentioned in brief. Namely, the $b \to c$ transition in the semileptonic decay is an instantaneous replacement of the b quark by c:

$$|X_c\rangle = \int d^3x\, e^{i\vec{q}\cdot\vec{x}}\bar{c}(0,\vec{x})\Gamma b(0,\vec{x})|B\rangle = j_{\vec{q}}|B\rangle, \tag{152}$$

where Γ is some Dirac matrix (γ_μ for the vector current and $\gamma_\mu\gamma_5$ for the axial-vector one). Using the fact that both b and c quarks are heavy we can use a nonrelativistic expansion for the current $j_{\vec{q}}$.

For definiteness, we will consider here the zero recoil limit which was discussed above in great detail. Generalization to $\vec{q} \neq 0$ is straightforward. As an example, for the vector and axial-vector currents we get

$$j_0^V(\vec{q}=0) = \int d^3x \varphi_c^\dagger(\vec{x})\left[1 - \frac{1}{8}\left(\frac{1}{m_c} - \frac{1}{m_b}\right)^2(\vec{\sigma}\cdot\vec{\pi})^2 + O\left(\frac{1}{m^3}\right)\right]\varphi_b(\vec{x}), \tag{153}$$

$$j_k^A(\vec{q}=0) = \int d^3x \varphi_c^\dagger(\vec{x})\left[\sigma_k - \frac{1}{8m_c^2}(\vec{\sigma}\cdot\vec{\pi})^2\sigma_k - \frac{1}{8m_b^2}\sigma_k(\vec{\sigma}\cdot\vec{\pi})^2 + \frac{1}{4m_c m_b}(\vec{\sigma}\cdot\vec{\pi})\sigma_k(\vec{\sigma}\cdot\vec{\pi}) + O\left(\frac{1}{m^3}\right)\right]\varphi_b(\vec{x}), \tag{154}$$

where $\varphi_{c,b}$ are two-component nonrelativistic spinor fields defined in Eq. (20). Other components of the currents do not contain the leading $(1/m)^0$ terms. Equations (153) and (154) are still written in the second-quantized form where $\varphi_{c,b}$ are field operators. However, since the number of heavy quarks does not change in the process at hand, it is convenient to proceed to the first-quantized form. It corresponds to an ordinary nonrelativistic description by a two-component wave function

$$\Psi_\alpha(\vec{x}_Q, \{x_{\rm light}\}), \alpha = 1,2,$$

where $\vec{x}_Q$ is the heavy quark coordinate, α is the heavy quark spinor index, and $\{x_{\rm light}\}$ represents an infinite number of light degrees of freedom. Note that it does not imply any nonrelativistic approximation for the light cloud; $\{x_{\rm light}\}$ are still field-theoretical coordinates of QCD. From now on we will not write out explicitly these coordinates in the argument of the wave functions.

Using this notation, Eq. (152) takes the form (at $\vec{q}=0$)

$$\Psi^{(X_c)}(\vec{x}_Q, \{\cdots\}) = \left[1 - \frac{1}{8}\left(\frac{1}{m_c} - \frac{1}{m_b}\right)^2(\vec{\sigma}\cdot\vec{\pi})^2 + O\left(\frac{1}{m^3}\right)\right]\Psi^{(B)}(\vec{x}_Q, \{\cdots\}), \tag{155}$$

for $j = j_0^V$, and for axial-vector current, $j = j_k^A$,

$$\Psi_k^{(X_c)}(\vec{x}_Q, \{\cdots\}) = \left[\sigma_k - \frac{1}{8m_c^2}(\vec{\sigma}\cdot\vec{\pi})^2\sigma_k - \frac{1}{8m_b^2}\sigma_k(\vec{\sigma}\cdot\vec{\pi})^2 + \frac{1}{4m_c m_b}(\vec{\sigma}\cdot\vec{\pi})\sigma_k(\vec{\sigma}\cdot\vec{\pi}) + O\left(\frac{1}{m^3}\right)\right]\Psi^{(B)}(\vec{x}_Q, \{\cdots\}). \tag{156}$$

The derivative appearing in the momentum operators $\pi^i = -i(\partial/\partial x_Q^i) + A^i(x_Q)$ acts on the coordinate x_Q of the wave function Ψ_α while the σ matrices act on its spin index.

In relativistic and nonrelativistic theories the expectation values $\langle B|\cdots|B\rangle$ are normalized differently. In relativistic theory $\langle B|B\rangle = 2M_B V$ (in the B meson rest frame) where V is the volume of the "large box," while $\langle\Psi^{(B)}|\Psi^{(B)}\rangle = 1$; the degree of freedom associated with motion of a system (B meson) as a whole is usually not considered, being treated separately. Correspondingly, the transition rule is

$$\langle B(\vec{p})| \int d^3x \bar{b}\mathcal{O}b(x)|B(\vec{p})\rangle_{\rm secondquant} \to \langle \Psi^{(B)}|\mathcal{O}|\Psi^{(B)}\rangle_{\rm QM}, \tag{157}$$

where $\mathcal{O}$ is some local operator, e.g., γ_0, π_i, etc.

The characteristic feature of the currents considered above is that the states produced by them from $\Psi^{(B)}$ are close to $\Psi^{(D)}$ and $\Psi^{(D^*)} \approx \sigma_k\Psi^{(D)}$, respectively. Moreover, at $\vec{q}=0$ the currents do not contain terms linear in $1/m$; corrections start with operators that explicitly contain $1/m^2$. Therefore,

$$\Psi^{(X_c)}(\vec{x}_Q) = \left[1 - \frac{1}{8}\left(\frac{1}{m_c}-\frac{1}{m_b}\right)^2 \langle B|(\vec{\sigma}\cdot\vec{\pi})^2|B\rangle\right]\Psi^{(B)}(\vec{x}_Q) - \frac{1}{8}\left(\frac{1}{m_c}-\frac{1}{m_b}\right)^2 \sum_n \Psi^{(B_n)}(\vec{x}_Q)\,\langle B_n|(\vec{\sigma}\cdot\vec{\pi})^2|B\rangle + O\left(\frac{1}{m^3}\right), \tag{158}$$

$$\Psi_k^{(X_c)}(\vec{x}_Q) = \left[1 + \frac{1}{6m_c^2}\langle B|\vec{\sigma}\cdot\vec{B}|B\rangle - \frac{1}{8}\left(\frac{1}{m_c^2}+\frac{1}{m_b^2}+\frac{2}{3m_cm_b}\right)\langle B|(\vec{\sigma}\cdot\vec{\pi})^2|B\rangle\right]\sigma_k\Psi^{(B)}(\vec{x}_Q) + \sum_n \Psi^{(B_n)}(\vec{x}_Q)\,\langle B_n|\left(-\frac{1}{8m_c^2}(\vec{\sigma}\cdot\vec{\pi})^2\sigma_k - \frac{1}{8m_b^2}\sigma_k(\vec{\sigma}\cdot\vec{\pi})^2 + \frac{1}{4m_cm_b}(\vec{\sigma}\cdot\vec{\pi})\sigma_k(\vec{\sigma}\cdot\vec{\pi})\right)|B\rangle + O\left(\frac{1}{m^3}\right). \tag{159}$$

Here $|B_n\rangle$ are excitations of the B meson and $\Psi^{(B_n)}$ are the corresponding wave functions. (Here and below in this section all matrix elements are written in the nonrelativistic normalization; the factor $(2M_B)^{-1}$, characteristic of the relativistic normalization, is absent.)

Let us emphasize that $\Psi^{(B)}$, $\Psi^{(B_n)}$ are the eigenfunctions of the Hamiltonian $\mathcal{H}(m_Q)$ with $m_Q = m_b$, and *not* the one with $m_Q = m_c$. $\mathcal{H}(m_Q)$ is the total Hamiltonian including that of light degrees of freedom, and its dependence on m_Q (apart from the trivial term m_Q) is given by Eq. (21). In the final state the observable hadrons are described by the eigenstates of $\mathcal{H}(m_Q = m_c)$. Therefore the amplitude of an individual state $|D_n\rangle$ is obtained by projecting $|\Psi^{(X_c)}\rangle$, $|\Psi_k^{(X_c)}\rangle$ onto these eigenfunctions. The projection onto the ground state is equal to unity up to $1/m^2$ terms, whereas the excited states are produced with the amplitudes which scale like $1/m$. This reexpansion is the only source of $1/m$ terms in the amplitudes.

With this information at hand let us return to the sum rules. We start from the first sum rules [see the expressions for I_0^{VV}, $I_0^{(1)AA}$ in Eqs. (111), (107)]. These sum rules have the meaning of the total probability of hadronic production at zero recoil if it is assumed that arbitrary energy can be carried away by the lepton pair produced in the decay process [see Eqs. (116), (113)]. In the nonrelativistic language these total probabilities are nothing but the norms of the states $\Psi^{(X_c)}$ and $\Psi_k^{(X_c)}$. There is no need for reexpansion in order to find these norms. As a matter of fact, the result is given by the coefficient in front of $\Psi^{(B)}$ (or $\sigma_k\Psi^{(B)}$); other terms contribute only at the level of $1/m^4$. The concrete expressions, thus, are

$$\langle\Psi^{(X_c)}|\Psi^{(X_c)}\rangle = \langle B|\,(j_0^V)^\dagger\, j_0^V\,|B\rangle = 1 - \frac{1}{4}\left(\frac{1}{m_c}-\frac{1}{m_b}\right)^2\langle B|(\vec{\sigma}\cdot\vec{\pi})^2|B\rangle + O\left(\frac{1}{m^3}\right), \tag{160}$$

$$\frac{1}{3}\sum_{k=1}^{3}\langle\Psi_k^{(X_c)}|\Psi_k^{(X_c)}\rangle = \frac{1}{3}\langle B|\,(j_k^A)^\dagger\, j_k^A\,|B\rangle = 1 + \frac{1}{3m_c^2}\langle B|\vec{\sigma}\cdot\vec{B}|B\rangle - \frac{1}{4}\left(\frac{1}{m_c^2}+\frac{1}{m_b^2}+\frac{2}{3m_cm_b}\right)\langle B|(\vec{\sigma}\cdot\vec{\pi})^2|B\rangle + O\left(\frac{1}{m^3}\right), \tag{161}$$

which exactly coincides with the results presented in Eqs. (116) and (113).

It is natural to ask at this point how it could happen that the relativistic current $\bar{b}\gamma_\mu c$ or $\bar{b}\gamma_\mu\gamma_5 c$ that was seemingly normalized to unity (through equal time commutators) led us to a nonrelativistic current with the normalization different from unity. The answer is that in the nonrelativistic consideration one excludes the states whose excitation energy $\sim m_Q$. In fact, adding such states would restore the normalization back to unity. This explains, in particular, the negative sign of $1/m^2$ corrections to the nonrelativistic normalization. (Let us mention in passing that Lipkin in his analysis did not consider these $1/m^2$ relativistic corrections altogether.)

The fact that the states $\Psi^{(B)}(\vec{x}_Q)$ and $\sigma_k\Psi^{(B)}(\vec{x}_Q)$ are not the eigenstates of the Hamiltonian $\mathcal{H}(m_c)$ becomes important for higher moments I_n, Eq. (106), which in the quantum-mechanical language take the form

$$I_n^{VV} = \langle\Psi^{(B)}(\vec{x}_c)|\left\{\mathcal{H}(m_c) - \langle D|\mathcal{H}(m_c)|D\rangle\right\}^n|\Psi^{(B)}(\vec{x}_c)\rangle\left[1+O\left(\frac{1}{m}\right)\right] \tag{162}$$

and

$$I_n^{AA} = \frac{1}{3}\sum_{k=1}^{3}\langle\Psi^{(B)}(\vec{x}_c)|\sigma_k\left\{\mathcal{H}(m_c) - \langle D^*|\mathcal{H}(m_c)|D^*\rangle\right\}^n\sigma_k|\Psi^{(B)}(\vec{x}_c)\rangle\left[1+O\left(\frac{1}{m}\right)\right]. \tag{163}$$

The $1/m^2$ corrections to currents do not show up in the leading term presented above. From now on we will abbreviate $\langle D|\mathcal{H}(m_c)|D\rangle = \langle\mathcal{H}(m_c)\rangle_D$, and likewise for D^*.

As we have already seen, all excitation probabilities are proportional to $1/m^2$ and, therefore,

$$I_n \sim \frac{\Lambda_{\rm QCD}^{n+2}}{m_Q^2} \text{ for } n \geq 1.$$

The very same $1/m^2$ behavior is also readily seen from Eqs. (162) and (163). Indeed, $\mathcal{H}(m_c) - \langle\mathcal{H}(m_c)\rangle$ acting on the states $\Psi^{(B)}(\vec{x}_c)$ and $\sigma_k\Psi^{(B)}(\vec{x}_c)$ produces a nonvanishing result only at the level $1/m$:

$$\begin{aligned}\left[\mathcal{H}(m_c) - \langle\mathcal{H}(m_c)\rangle_D\right]|\Psi^{(B)}(\vec{x}_c)\rangle &= \left[\mathcal{H}(m_b) - \langle\mathcal{H}(m_c)\rangle_D + \mathcal{H}(m_c) - \mathcal{H}(m_b)\right]|\Psi^{(B)}(\vec{x}_c)\rangle \\ &= \left[(M_B - M_D) - (m_b - m_c) + (\vec{\sigma}\cdot\vec{\pi})^2\left(\frac{1}{2m_c} - \frac{1}{2m_b}\right)\right]|\Psi^{(B)}(\vec{x}_c)\rangle + O\left(\frac{1}{m^2}\right)\end{aligned} \tag{164}$$

and, by the same token,

$$\begin{aligned}\left[\mathcal{H}(m_c) - \langle\mathcal{H}(m_c)\rangle_{D^*}\right]\sigma_k|\Psi^{(B)}(\vec{x}_c)\rangle = \Bigg\{&\sigma_k\left[(M_B - M_{D^*}) - (m_b - m_c) + (\vec{\sigma}\cdot\vec{\pi})^2\left(\frac{1}{2m_c} - \frac{1}{2m_b}\right)\right] \\ &- \frac{1}{2m_c}[\sigma_k, \vec{\sigma}\cdot\vec{B}]\Bigg\}|\Psi^{(B)}(\vec{x}_c)\rangle + O\left(\frac{1}{m^2}\right).\end{aligned} \tag{165}$$

Using these equations it is easy to obtain the predictions for the moments with $n \geq 2$. To this end we act by the leftmost and rightmost operators $\mathcal{H}(m_c) - \langle\mathcal{H}(m_c)\rangle$ onto $\langle\Psi^{(B)}|$ and $|\Psi^{(B)}\rangle$, respectively. This produces an overall $1/m^2$ suppression. For vector currents we have

$$\begin{aligned}I_n^{VV} &= \langle\Psi^{(B)}|\left[\mathcal{H}(m_c) - \langle\mathcal{H}(m_c)\rangle_D\right]^n|\Psi^{(B)}\rangle \\ &= \left\langle\Psi^{(B)}\right|\left[(\vec{\sigma}\cdot\vec{\pi})^2\left(\frac{1}{2m_c} - \frac{1}{2m_b}\right) + \delta_V\right](\mathcal{H} - \langle\mathcal{H}\rangle)^{n-2}\left[(\vec{\sigma}\cdot\vec{\pi})^2\left(\frac{1}{2m_c} - \frac{1}{2m_b}\right) + \delta_V\right]\left|\Psi^{(B)}\right\rangle + O\left(\frac{1}{m^2}\right),\end{aligned} \tag{166}$$

while for the axial-vector currents

$$\begin{aligned}I_n^{AA} &= \frac{1}{3}\sum_{k=1}^{3}\langle\Psi^{(B)}|\sigma_k\left[\mathcal{H}(m_c) - \langle\mathcal{H}(m_c)\rangle_{D^*}\right]^n\sigma_k|\Psi^{(B)}\rangle \\ &= \frac{1}{3}\sum_{k=1}^{3}\left\langle\Psi^{(B)}\right|\left\{\left[(\vec{\sigma}\cdot\vec{\pi})^2\left(\frac{1}{2m_c} - \frac{1}{2m_b}\right) + \delta_A\right]\sigma_k + \frac{1}{2m_c}[\sigma_k, \vec{\sigma}\cdot\vec{B}]\right\}(\mathcal{H} - \langle\mathcal{H}\rangle)^{n-2} \\ &\quad\times\left\{\sigma_k\left[(\vec{\sigma}\cdot\vec{\pi})^2\left(\frac{1}{2m_c} - \frac{1}{2m_b}\right) + \delta_A\right] - \frac{1}{2m_c}[\sigma_k, \vec{\sigma}\cdot\vec{B}]\right\}\left|\Psi^{(B)}\right\rangle + O\left(\frac{1}{m^2}\right),\end{aligned} \tag{167}$$

where

$$\delta_V = (M_B - M_D) - (m_b - m_c),$$

$$\delta_A = (M_B - M_{D^*}) - (m_b - m_c)$$

are given in Eqs. (112) and (105). Note that in the remaining operators $(\mathcal{H} - \langle\mathcal{H}\rangle)^{n-2}$ one can already neglect the explicit heavy quark mass dependence. For $n > 2$ the terms with δ lead only to corrections higher than $1/m^2$ and can be omitted. Indeed, picking up δ rather than $(\vec{\sigma}\cdot\vec{\pi})^2$ allows one to act by $(\mathcal{H} - \langle\mathcal{H}\rangle)$ directly on Ψ or $\sigma_k\Psi$ which results, in turn, in an additional $1/m$ suppression. The case $n = 1$ is somewhat special and will be discussed shortly.

One can easily see that Eqs. (166) and (167) coincide with our previous results for the moments, Eqs. (111) and (107), provided that one identifies $\pi_0^{(n-2)}$ in the latter with $(\mathcal{H} - \langle\mathcal{H}\rangle)^{(n-2)}$ in the quantum-mechanical relations.

This correspondence has a transparent physical meaning and can be readily proven. Indeed, both play the role of the generator of the time evolution, $(\mathcal{H} - \langle\mathcal{H}\rangle)$ in the first-quantized approach and $\pi_0 = i\partial_0 + A_0$ in the second-quantized formalism.

To see that this is indeed the case, let us consider the state

$$|\Psi_{\mathcal{O}}(t)\rangle = \bar{b}(\vec{x}=0,t)\,\mathcal{O}\,b(\vec{x}=0,t)\,|B\rangle, \tag{168}$$

where $\mathcal{O}$ is an arbitrary operator. Then

$$\langle\Psi_{\mathcal{O}}(0)|\Psi_{\mathcal{O}}(t)\rangle = \langle B|\,\bar{b}(\vec{x}=0,t=0)\,\mathcal{O}^\dagger\,T\,e^{-i\int_0^t A_0(\vec{x}=0,\tau)d\tau}\,\mathcal{O}\,b(\vec{x}=0,t)\,|B\rangle \tag{169}$$

in the leading in $1/m$ approximation when the heavy quark Green function reduces to the T exponent above. Expanding the right-hand side in t one gets

$$\langle\Psi_{\mathcal{O}}(0)|\Psi_{\mathcal{O}}(t)\rangle = \langle B|\bar{b}(\vec{x}=0,t=0)\,\mathcal{O}^\dagger\sum_{n=0}^{\infty}(i\partial_0 + A_0)^n\frac{(-it)^n}{n!}\,\mathcal{O}\,b(\vec{x}=0,t)|B\rangle, \tag{170}$$

which proves the assertion above.

Let us return now to the discussion of the special case of the $n=1$ moments (the second sum rule). This case is singled out because, on the one hand, I_1 must be proportional to $\Lambda^3_{\rm QCD}/m^2$ on general grounds and, on the other hand, this suppression does not immediately show up as happens for $n \geq 2$, since now we have only one power of $\mathcal{H} - \langle\mathcal{H}\rangle$ sandwiched between the bra and ket states. To calculate I_1^{VV} and I_1^{AA} we need to account for all $1/m^2$ terms that have been neglected in Eqs. (164) and (165). Including these terms we get

$$\begin{aligned} I_1^{VV} = {} & [(M_B - M_D) - (m_b - m_c)] + \langle\Psi^{(B)}(\vec{x}_c)|\left(\frac{1}{2m_c} - \frac{1}{2m_b}\right)(\vec{\sigma}\cdot\vec{\pi})^2 \\ & + \frac{1}{8}\left(\frac{1}{m_c^2} - \frac{1}{m_b^2}\right)[-(\vec{D}\cdot\vec{E}) + 2\vec{\sigma}\cdot\vec{E}\times\vec{\pi}]|\Psi^{(B)}(\vec{x}_c)\rangle + O\left(\frac{\Lambda^4_{\rm QCD}}{m^3}\right) \end{aligned} \tag{171}$$

and

$$\begin{aligned} I_1^{AA} = {} & [(M_B - M_{D^*}) - (m_b - m_c)] + \langle\Psi^{(B)}(\vec{x}_c)|\left(\frac{1}{2m_c} - \frac{1}{2m_b}\right)(\vec{\pi}\cdot\vec{\sigma})^2 - \frac{2}{3m_c}\vec{\sigma}\cdot\vec{B} \\ & + \frac{1}{8}\left(\frac{1}{m_c^2} - \frac{1}{m_b^2}\right)[-(\vec{D}\cdot\vec{E}) + \vec{\sigma}\cdot\vec{E}\times\vec{\pi}] - \frac{1}{3m_c^2}\vec{\sigma}\cdot\vec{E}\times\vec{\pi}|\Psi^{(B)}(\vec{x}_c)\rangle + O\left(\frac{\Lambda^4_{\rm QCD}}{m^3}\right), \end{aligned} \tag{172}$$

where we have used expression (21) for the $1/m^2$ part of the Hamiltonian.

If one neglects the $1/m^2$ terms, then I_1^{VV} and I_1^{AA} vanish, and the known relations Eqs. (112) and (105) for hadron masses at the level $1/m$ arise. Equations (171) and (172) are extensions of these relations to the order $1/m^2$. There are two sources of $1/m^2$ corrections in the right-hand side: The first one is explicit $1/m^2$ terms in these equations; the second source is an implicit $(1/m^1)$ dependence of the expectation values of operators $\vec{\pi}^2$ and $\vec{\sigma}\cdot\vec{B}$ over the state $\Psi^{(B)}(\vec{x}_c)$. This dependence is due to the fact that $\Psi^{(B)}(\vec{x}_c)$ is the eigenfunction of the Hamiltonian $\mathcal{H}(m_b)$ rather than of the asymptotic one $\mathcal{H}(m_Q=\infty)$. It is obvious, by the way, that these matrix elements do not produce $1/m_c$ dependence.

Let us discuss in brief how this quantum-mechanical formalism works in the case when $\vec{q} \neq 0$. The instantaneous $b \to c$ transition now not only replaces b by c but also boosts the c quark providing it with the spatial momentum $-\vec{q}$. As a result the wave function of the system produced takes the following form for the vector and axial-vector currents, respectively:

$$\Psi^{(X_c)}(\vec{x}_Q,\{\cdots\}) = e^{-i\vec{q}\cdot\vec{x}_Q}\left(1 - \frac{\vec{v}^{\,2}}{8}\right)\cdot\Psi^{(B)}(\vec{x}_Q,\{\cdots\}) + O\left(\vec{v}^{\,4}, \frac{\Lambda_{\rm QCD}|\vec{v}|}{m}, \frac{\Lambda^2_{\rm QCD}}{m^2}\right), \tag{173}$$

$$\Psi_k^{(X_c)}(\vec{x}_Q,\{\cdots\}) = e^{-i\vec{q}\cdot\vec{x}_Q}\left(1 - \frac{\vec{v}^{\,2}}{8}\right)\cdot\sigma_k\Psi^{(B)}(\vec{x}_Q,\{\cdots\}) + O\left(\vec{v}^{\,4}, \frac{\Lambda_{\rm QCD}|\vec{v}|}{m}, \frac{\Lambda^2_{\rm QCD}}{m^2}\right), \tag{174}$$

where $\vec{v} = -\vec{q}/m_c$ is the velocity of the final c quark in the SV limit. The factor $1 - \vec{v}^{\,2}/8$ reflects the overall normalization of the currents; it results in the factor $1 - \vec{v}^{\,2}/4$ in the $n=0$ moments (the Bjorken sum rule). In higher moments it is not essential at the level of accuracy we accept here.

The moments $I_n(\vec{q})$ are defined in Eq. (129), but here we define ϵ for the vector case as $\epsilon = M_B - q_0 - \sqrt{M_D^2 + \vec{q}^{\,2}}$. The generic formulas for the moments take the form

$$I_0^{VV} = I_0^{AA} = 1 - \frac{\vec{v}^{\,2}}{4}, \tag{175}$$

$$I_n^{VV} = \langle \Psi^{(B)} | \left[e^{i\vec{q}\cdot\vec{x}_Q} \mathcal{H}(m_c) e^{-i\vec{q}\cdot\vec{x}_Q} - \sqrt{M_D^2 + \vec{q}^{\,2}} \right]^n |\Psi^{(B)}(\vec{x}_c)\rangle + O\left(\vec{v}^{\,4}, \frac{\Lambda_{\rm QCD}^2}{m^2}\right), \tag{176}$$

$$I_n^{AA} = \frac{1}{3}\sum_{k=1}^{3} \langle \Psi^{(B)} | \sigma_k \left[e^{i\vec{q}\cdot\vec{x}_Q} \mathcal{H}(m_c) e^{-i\vec{q}\cdot\vec{x}_Q} - \sqrt{M_{D^*}^2 + \vec{q}^{\,2}} \right]^n \sigma_k |\Psi^{(B)}(\vec{x}_c)\rangle + O\left(\vec{v}^{\,4}, \frac{\Lambda_{\rm QCD}^2}{m^2}\right). \tag{177}$$

Let us notice that the only effect of the exponent is to replace the operator $\vec{\pi}$ in the Hamiltonian by $\vec{\pi} - \vec{q}$:

$$\begin{aligned} e^{i\vec{q}\cdot\vec{x}_Q} \mathcal{H}(m_c; \vec{\pi}) e^{-i\vec{q}\cdot\vec{x}_Q} &= \mathcal{H}(m_c; \vec{\pi} - \vec{q}) \\ &= \mathcal{H}(m_c; \vec{\pi}) - \frac{\vec{q}\cdot\vec{\pi}}{m_c} + \frac{\vec{q}^{\,2}}{2m_c} \\ &\quad - \frac{1}{4m_c^2} \vec{q}\cdot\vec{\sigma}\times\vec{E} + O\left(\frac{1}{m_c^3}\right). \end{aligned} \tag{178}$$

As an example, let us consider the $\vec{q}^{\,2}$ dependence of the first few moments. For $n = 0$ the result for $dI_0/d\vec{v}^{\,2}$ at $\vec{v}^{\,2} = 0$ coincides with the Bjorken sum rules; it has been derived in Sec. V with even better accuracy.

The next moment to consider is $n = 1$. Here we get, for $dI_1/d\vec{v}^{\,2}$ at $\vec{v}^{\,2} = 0$,

$$\left.\frac{dI_1^{VV}}{d\vec{v}^{\,2}}\right|_{\vec{v}=0} = \left.\frac{dI_1^{AA}}{d\vec{v}^{\,2}}\right|_{\vec{v}=0} = \frac{\bar{\Lambda}}{2},$$

$$\bar{\Lambda} \simeq M_D - m_c \simeq M_{D^*} - m_c. \tag{179}$$

This is Voloshin's "optical" sum rule [24].

Finally, let us mention $n = 2$ case, where, at $\vec{v}^{\,2} = 0$,

$$\left.\frac{dI_2^{VV}}{d\vec{v}^{\,2}}\right|_{\vec{v}=0} = \left.\frac{dI_2^{AA}}{d\vec{v}^{\,2}}\right|_{\vec{v}=0} = \frac{1}{3}\langle \vec{\pi}^{\,2}\rangle. \tag{180}$$

Concluding this section, let us make a few comments on those aspects of the quantum-mechanical approach which we modified compared to the original Lipkin's presentation. The central point of the framework suggested in Ref. [27] is the fact that the wave function of the charmed system, immediately after the instantaneous $b \to c$ transition, is known in terms of the B meson wave function; see Eqs. (158), (159) and (173), (174). These equations differ from their counterparts in Ref. [27] in the normalization of the currents; there it was effectively set equal to unity. In our expressions the charm wave function is *not* normalized to unity, with the corrections appearing at the level $O(\Lambda_{\rm QCD}^2/m^2)$ or $O(\vec{v}^{\,2})$. This correction to normalization affects only the first sum rules, viz., I_0^{VV}, I_0^{AA}.

Another point of difference is that Lipkin did not account for the fact that the spin part of the Hamiltonian depends on m_Q at the same level as the kinetic energy. As a result, our sum rules at zero recoil are, strictly speaking, different from those of Ref. [27]. (Our results coincide with those of Lipkin provided the spin interaction is switched off.) At the level of accuracy we accept in this section the spin terms omitted in Ref. [27] do not affect at all the sum rules at $\vec{q} \neq 0$. In particular, the linear in $\vec{q}^{\,2}$ part in the first moments [see Eq. (179) for I_1^{VV} and I_1^{AA}] derived in Ref. [27] identically coincides with Voloshin's sum rule [24], although surprisingly it was not recognized in Ref. [27].

Our final remark concerns $\bar{c}b$ currents which vanish in the nonrelativistic limit. For example, in Sec. IV D we considered the spatial components of the vector current $\bar{c}\gamma_k b$ at $\vec{q} = 0$. In the nonrelativistic limit this current takes the form

$$\bar{c}\vec{\gamma}b\,|_{\vec{q}=0} = \phi_c^\dagger \left[\left(\frac{1}{2m_c} + \frac{1}{2m_b} \right) \vec{\pi} + \left(\frac{1}{2m_c} - \frac{1}{2m_b} \right) i\vec{\sigma}\times\vec{\pi} \right] \phi_b + O\left(\frac{1}{m^2}\right). \tag{181}$$

This current produces the charmed state with the wave function

$$\tilde{\Psi}_k(\vec{x}_Q) = \left[\left(\frac{1}{2m_c} + \frac{1}{2m_b} \right) \pi_k + \left(\frac{1}{2m_c} - \frac{1}{2m_b} \right) i[\vec{\sigma}\times\vec{\pi}]_k \right] \Psi^{(B)}(\vec{x}_Q). \tag{182}$$

The normalization of this state is

$$\frac{1}{3}\sum_{k=1}^{3} \langle \tilde{\Psi}_k(\vec{x}_Q) | \tilde{\Psi}_k(\vec{x}_Q)\rangle \simeq \frac{1}{4}\left[\left(\frac{1}{m_c^2} + \frac{1}{m_b^2} - \frac{2}{3m_c m_b} \right) \langle \vec{\pi}^{\,2}\rangle + \left(\frac{1}{m_b^2} - \frac{1}{3m_c^2} - \frac{2}{3m_c m_b} \right) \langle \vec{\sigma}\cdot\vec{B}\rangle \right]. \tag{183}$$

This expression is identical to the left-hand side of the sum rule (121) which is the $n = 0$ moment for this current. For higher moments we get

$$I_n^{(1)VV}(\vec{q}=0)=\frac{1}{3}\sum_{k=1}^{3}\langle\tilde{\Psi}_k(\vec{x}_Q)|\left[\mathcal{H}(m_c)-\langle\mathcal{H}(m_c)\rangle\right]^n|\tilde{\Psi}_k(\vec{x}_Q)\rangle$$
$$\simeq\frac{1}{3}\sum_{k=1}^{3}\left\langle\Psi^{(B)}\left|\left[\left(\frac{1}{2m_c}+\frac{1}{2m_b}\right)\pi_k-\left(\frac{1}{2m_c}-\frac{1}{2m_b}\right)i[\vec{\sigma}\times\vec{\pi}]_k\right]\left[\mathcal{H}(m_c)-\langle\mathcal{H}(m_c)\rangle\right]^n\right.\right.$$
$$\left.\left.\times\left[\left(\frac{1}{2m_c}+\frac{1}{2m_b}\right)\pi_k+\left(\frac{1}{2m_c}-\frac{1}{2m_b}\right)i[\vec{\sigma}\times\vec{\pi}]_k\right]\right|\Psi^{(B)}\right\rangle. \tag{184}$$

Now the suppression $1/m^2$ is explicit—it comes from the current itself—and all $1/m$ terms in $\mathcal{H}(m_c)$ can be omitted.

VII. IMPACT OF THE PERTURBATIVE EVOLUTION OF OPERATORS IN THE OPE

In this section we shall briefly discuss practical modifications that arise in the calculation of nonperturbative corrections in the heavy flavor decays if one accounts for the normalization point dependence of the corresponding operators. Although this question is rather standard, we feel the need to dwell on it in view of the apparent confusion taking place in applications of the heavy quark expansions existing in the literature.

We have already mentioned above that HQET is nothing else than a version of the Wilson operator product expansion. It is well understood (see, for example, a recent discussion in Ref. [28]) that consistent incorporation of nonperturbative effects in the framework of the Wilson procedure requires separating momenta below and above a certain normalization point μ. The low-momentum physics is then attributed to the matrix elements of the operators whereas the high-momentum part enters the Wilson expansion coefficients; both, therefore, depend explicitly on μ in such a way that all observables are μ independent.

This procedure is always performed when the corresponding operators undergo logarithmic renormalization, for an obvious reason: The Feynman integrals determining the coefficient functions in this case logarithmically diverge in the infrared domain, and one merely cannot put μ, the infrared regularization, to zero. In calculating the power corrections to the heavy flavor decays another situation can typically arise, when the operators under consideration have vanishing anomalous dimensions and no logarithmic mixing. This is the case, for instance, with the leading operator $\bar{Q}Q$ and the kinetic energy operator $\bar{Q}\vec{\pi}^2Q$ (the second one still mixes with the first one nonlogarithmically, through a power of μ^2; see below). In this case the coefficient functions possess a safe infrared limit, and it is very tempting just to calculate them in perturbation theory *per se*, with no infrared cutoff (or, which is the same, putting $\mu=0$). This is what is routinely done with respect to the coefficient of the operator $\bar{Q}Q$.

From a purely theoretical point of view there is no way one can justify such a procedure. It gives rise to questions which have no consistent answers, e.g., how the sum of all perturbative terms is to be understood, etc. One of the manifestations of these inconsistencies is the infrared renormalon (see a recent discussion in Refs. [28, 29] and references therein). It should be very clearly realized that in the consistent operator product expansion (and, hence, in HQET) one does not discriminate perurbative contributions versus nonperturbative, but, rather, large distance ones versus short distances (all distances are measured in the scale of μ^{-1}).

However, in practice no separation of the infrared part from the coefficient of $\bar{Q}Q$ is carried out. The usual routine is as follows: One takes the known expressions for the one-loop (or two-loop) perturbative corrections for a particular quantity and merely adds to these perturbative terms nonperturbative corrections expressed via certain matrix elements. (We also follow this routine for numerical calculations.) For instance, the perturbative correction η_A to the axial-vector current $\bar{c}\gamma_\mu\gamma_5 b$ calculated from the standard Feynman graphs with no subtractions was simply added to the nonperturbative contribution of Eq. (115). It is clear, however, that the nonperturbative contribution *per se* takes care of all relevant gluon exchanges with momenta below μ; on top of it the one-loop Feynman integral for the radiative correction η_A has some (small) piece coming from the same low-momentum domain. The question which immediately comes to one's mind is the menace of double counting.

The answer to this question is that in actual OPE-based calculations, which are always truncated at some finite order in α_s and keep only a few higher-dimensional operators, one can stick to what is called the "practical version" of the OPE in QCD [54] and still avoid double counting.

Indeed, it is not possible to define the perturbative part in, say, $\langle\bar{Q}\vec{\pi}^2Q\rangle$ to all orders in α_s. On the other hand, it is quite possible to introduce a "one-loop perturbative contribution" to $\bar{Q}\vec{\pi}^2Q$ normalized at μ. To this end we, *by definition*, take two gluon lines in $\bar{Q}\vec{\pi}^2Q$, contract them to form the gluon loop, use the bare gluon propagator, and calculate the loop cutting the integration off from above, at $k_g=\mu$. We then get

$$\bar{Q}\vec{\pi}^2Q|_{\text{one loop}}=\frac{4\alpha_s}{3\pi}\mu^2\bar{Q}Q.$$

Now let us subtract and add this "one-loop" π^2 from the properly and scientifically defined $\bar{Q}\vec{\pi}^2Q$. Moreover, let us combine the added part with $C(\mu)\bar{Q}Q$; then we get $\bar{Q}Q$ times the coefficient coinciding, up to corrections of higher order in μ^2/m^2, with the coefficient obtained from the full perturbative one-loop calculation, with no

subtractions whatsoever. Simultaneously, the matrix element of the kinetic energy operator is replaced by that with the subtracted "one-loop" part. Strictly speaking, this quantity does not represent now a matrix element of any operator; hence, factorization inherent to the genuine OPE is lost. This is unimportant for numerical analysis due to the fact that the term which we added and subtracted is numerically small, much smaller than the actual value of $\langle\pi^2\rangle$. For this reason the replacement of the added term plus $C(\mu)\bar{Q}Q$ by merely $C_{\text{one loop}}\bar{Q}Q$ introduces a very small numerical error, much smaller than π^2 correction itself. This fact, the numerical smallness of the "one-loop" value of the condensate compared to its actual value, constitutes the conceptual foundation for the practical version of OPE. In this version μ does not appear explicitly in the coefficient function of the lowest-dimension operator; formally this corresponds to setting $\mu = 0$ there. The fact of the numerical smallness is not an obvious property of QCD and is to be cross-checked in any new situation. So far it turns out that it is always valid, for reasons which are not completely understood; see the reprint volume cited in [54]. We repeat, however, that in principle one should calculate the Wilson coefficients by evaluating the relevant Feynman integrals with an explicit infrared cutoff in the propagators.

To further facilitate understanding of this subtle aspect let us discuss this general strategy, which we usually stick to, in a particular example, namely, the first sum rule, Eq. (113). This sum rule, with the perturbative corrections added, takes the form

$$F^2_{B\to D^*} + \sum_{\text{excit}} F^2_{B\to\text{excit}} = \xi_A - \frac{1}{3}\frac{\mu_G^2}{m_c^2} - \frac{\mu_\pi^2 - \mu_G^2}{4}\left(\frac{1}{m_c^2} + \frac{1}{m_b^2} + \frac{2}{3m_c m_b}\right). \tag{185}$$

For a consistent calculation of the perturbative correction factor ξ_A we need to introduce an infrared renormalization point $\mu \ll m_c$ and use $\xi_A(\mu)$ in Eq. (185); the operators entering this sum rule are then normalized at the scale μ. It is worth noting that without introducing μ the sum rule written above, strictly speaking, has little sense because the sum over excitation will diverge in the ultraviolet (see below). If a renormalization point is set, the sum will run only over states with excitation energy below μ. Keeping in mind the explicit μ dependence of ξ_A, it is clear that it cannot be equal to η_A^2 which is calculated without an infrared cutoff and is thus μ independent. In reality, however, the major part of both η_A^2 and $\xi_A(\mu)$ comes in the heavy quark limit from the momenta $\sim m_Q$ and therefore they must be similar. Actually, one does not even need to calculate the perturbative factor ξ_A anew as long as η_A^2 is known.

To determine ξ_A one notes that to any particular order in perturbation theory the following relation holds:

$$\eta_A^2 = \xi_A(\mu)\,|_{\mu=0}; \tag{186}$$

in fact it is merely the definition of η_A. It can be formally obtained considering the sum rule (185) in perturbation theory in the particular order (we will assume the first order in α_s in what follows). In this approximation the spectrum of particles is given by quarks and gluons, and matrix elements of operators, in particular μ_π^2, are nothing but the "perturbative one-loop" contributions considered above. The elastic contribution is given by the exclusive $b \to c$ probability whereas the sum over the excitation is represented by the process $b \to c + g$ whose probability is proportional to α_s. Then, at $\mu = 0$ one gets the relation (186) because other terms vanish.

To determine the μ dependence of ξ_A we can use the same technique as in Sec. IV D, namely, consider the perturbative analogue of the sum rule (185):

$$\eta_A^2 + \frac{1}{3}\sum_k \int_{\omega<\mu} \frac{d^3k}{2\omega(2\pi)^3}\frac{1}{4m_c m_b}|\langle cg|\bar{c}\gamma_k\gamma_5 b|b\rangle|^2 = \xi_A(\mu) - \frac{(\mu_\pi^2)_{\text{pert}}}{4}\left(\frac{1}{m_c^2} + \frac{1}{m_b^2} + \frac{2}{3m_c m_b}\right). \tag{187}$$

Notice that $(\mu_G^2)_{\text{pert}} = 0$ because the spin-dependent operators do not mix with $\bar{Q}Q$. Therefore, the chromomagnetic operator is irrelevant for the aspect under discussion now. The gluon emission amplitude is

$$\langle cg|\bar{c}\vec{\gamma}\gamma_5 b|b\rangle = g_s\,\bar{c}\,t^a\left[(\vec{\epsilon}^{\,a}\times\vec{n})\times\vec{\sigma}\left(\frac{1}{2m_c} + \frac{1}{2m_b}\right) + i\,(\vec{\epsilon}^{\,a}\times\vec{n})\left(\frac{1}{2m_c} - \frac{1}{2m_b}\right)\right] b. \tag{188}$$

Here $\vec{n} = \vec{k}/|\vec{k}\,|$. Using the expression (126) for $(\mu_\pi^2)_{\text{pert}}$ we immediately find that

$$\xi_A(\mu) \simeq \eta_A^2 + \frac{2\alpha_s}{3\pi}\mu^2\left(\frac{1}{m_c^2} + \frac{1}{m_b^2} + \frac{2}{3m_c m_b}\right). \tag{189}$$

A similar situation occurs in the sum rule for the vector current (the time component) where both the coefficient in front of μ_π^2 and the gluon emission probability are proportional to $(1/2m_b - 1/2m_c)^2$.

The scale dependence of the strong coupling α_s in the above expressions is not essential in the one-loop calculations; if higher loops are accounted for, the sum of all Feynman integrals with the particular infrared cutoff μ will automatically have a suitable form to give logarithms

of the ratio of m/μ necessary to convert $\alpha_s(m)$ appearing in the perturbative calculations of ξ, into $\alpha_s(\mu)$ in terms responsible for μ^2 corrections.

In principle, one can now utilize Eq. (189) to use the exact perturbative value of $\xi_A(\mu)$ in Eq. (185), which differs from η_A^2 for $\mu \neq 0$. Obviously, $\xi_A(\mu) - \eta_A^2$ in Eq. (189) contains two pieces, one given by the perturbative one-loop sum over the excitation calculated with the upper cutoff μ and the second representing the "perturbative one-loop" contribution to $\mu_\pi^2(\mu)$. It is clear that one can formally subtract these two terms from the sum over excitations in the left-hand side of the sum rule (185) and from $\mu_\pi^2(\mu)$ term in the right-hand side, respectively, and then use as the perturbative factor ξ_A its one-loop value at $\mu = 0$, i.e., η_A^2. It is just what one does following the routine practice of the OPE where the perturbative coefficient is calculated without an explicit infrared cutoff. In such a case μ_π^2 entering the "practical" form of the sum rule can be merely understood as

$$\mu_\pi^2 \to \mu_\pi^2(\mu) - \mu^2 \frac{d\mu_\pi^2}{d\mu^2} \tag{190}$$

and, simultaneously,

$$\sum_{\epsilon \leq \mu} F_{B\to \text{excit}}^2 \to \sum_{\epsilon \leq \mu} F_{B\to \text{excit}}^2 - \frac{1}{3} \sum_k \int_{\omega<\mu} \frac{d^3k}{2\omega(2\pi)^3} \frac{1}{4m_c m_b} |\langle cg|\bar{c}\gamma_k\gamma_5 b|b\rangle|^2. \tag{191}$$

The normalization point μ is chosen high enough to ensure that the approximate duality of the parton computed probabilities to real hadronic ones above μ. Then one can remove the upper limit in the difference in the right-hand side of the above equation and assume that the sum runs over all energies: The region above μ cancels out automatically.

The expression in the right-hand side of Eq. (190) is nothing but a linear (in μ^2) extrapolation of the matrix element from the point μ to $\mu = 0$. Formally, it is independent of μ: The dependence appears in terms proportional to α_s^2. Therefore one can take any value of μ as long as $\Lambda_{\text{QCD}} \ll \mu \ll m_c$. At the same time, the sum over excitations now is to be understood, strictly speaking, as the one from which the small one-loop perturbative gluon probability is subtracted.

A similar analysis can be literally extended to the two-, three-, and, in principle, any finite-loop calculation of ξ_A. We would like to stress once again, though, that this procedure has no theoretical justification and is not even well defined in high orders. We dwell on it only for the reason that it literally corresponds to the routine procedure used in numerical calculations; using $\mu = 0$ not only does not allow one to use consistently the OPE, but would lead also to numerical problems in higher orders.

Now let us turn from this rather general theoretical discussion to more practical questions related to heavy flavor decays. Up to now nonperturbative effects have been discussed in detail through corrections of order $1/m_Q^3$. Some effects, such as the invariant mass of the final hadronic state in the decays considered above, have corrections starting at order $1/m_Q$ and are expressed via the parameter $\bar{\Lambda}$ (see also Ref. [11]). As pointed out in Ref. [28] and illustrated in the present paper, their effects are determined by Feynman integrals which have a linear behavior in the infrared region; the corresponding IR effects are not expressed in terms of matrix elements of any local operator.

The inclusive widths of heavy flavor particles *are* obtained from an expansion in local operators, and corrections start with terms scaling like $1/m_Q^2$. These are described by two universal operators: the chromomagnetic one $\bar{Q}\,\frac{i}{2}\sigma G\,Q$ and the kinetic operator $\bar{Q}\,(i\vec{D})^2\,Q$. The natural normalization scale for them is given by $\mu \simeq m_Q$, at least if one neglects the mass of the final state quark(s).[15] One cannot, however, use directly this high normalization point because then the matrix elements of high-dimension operators will scale like m_Q to the corresponding power due to purely perturbative contributions, and instead of an expansion in $1/m_Q$ one would obtain the suppression of the higher-order terms only as some powers of $\alpha_s(m_Q)$. To obtain the real power expansion one must evolve these operators down to a scale μ which is to be much smaller than m_Q but still much larger than Λ_{QCD}. The expansion one arrives at in this way runs, strictly speaking, in powers of μ/m_Q.

The present state of the art in this kind of calculation is limited only by one-loop corrections to Wilson coefficients which, apart from the chromomagnetic operator that has a logarithmic renormalization, are calculated (or even typically borrowed from old QED calculations) without an infrared cutoff. The same refers to the corrections to weak currents used in the present paper. The analysis above suggests therefore that the value of the kinetic term μ_π^2 can be understood in the corresponding expressions as a linear extrapolation to $\mu = 0$. This problem does not arise at all for the chromomagnetic operator $\mathcal{O}_G$ whose mixing with the leading one $\bar{Q}Q$ appears only in the next order in $1/m_Q$ due to the fact that it is not a spin singlet; no double counting occurs for this reason. From a practical viewpoint, because the value of μ_π^2 is basically unknown yet, it does not make a big difference at present to prefer this or an alternative definition. Let us note in passing that it is quite probable that the QCD sum rule estimates determine a similar quantity extrapolated to $\mu = 0$ because no explicit infrared cutoff in the integrals is introduced, though this question definitely de-

[15] It is important that for the hyperfine splitting in heavy mesons, which allows one to extract experimentally the chromomagnetic matrix elements, the same normalization point emerges.

serves a more careful analysis if one wants to go beyond the accuracy of the "practical" OPE.

The issue of an accurate understanding of the definition of matrix elements of operators becomes important when one turns to the real practical bounds on physical observables of the type discussed in this paper. For example, extrapolating the operator $\bar{Q}\,\vec{\pi}^2\,Q$ to the zero renormalization point implies a subtraction of a positive quantity which, in principle, might have even changed the sign of the matrix element. To state it differently, one may be concerned whether the inequality (123) survives the extrapolation of μ_π^2 to a low point, which is often assumed implicitly. We shall argue now that this effect is too small numerically and cannot upset the bound we used.

To see it, let us consider the reasonably high normalization point $\mu \simeq 1\,\mathrm{GeV}$. Using the explicit estimate of the renormalization point dependence of the operator $\bar{Q}\,\vec{\pi}^2\,Q$ in Eq. (72) and assuming $\alpha_s(\mu) \simeq 0.36$ one readily obtains that the amount one may need to subtract from $\mu_\pi^2(\mu)$ constitutes at most $0.15\,\mathrm{GeV}^2$, a value that does not exceed theoretical uncertainties in the existing estimates of μ_π^2. Most probably this number overestimates the real contribution to be subtracted, because approximate duality of the perturbative corrections is expected to start earlier; moreover, the perturbative coefficient functions (η_A, η_V, corrections to inclusive widths, etc.) are evaluated in one loop using a smaller value of $\alpha_s \simeq \alpha_s(m_Q)$. The second effect, though formally of higher order in α_s, is too transparent physically to raise doubts that a more realistic estimate of effects of potential double counting corresponds to using $\alpha_s(m_c, m_b)$ rather than $\alpha_s(1\,\mathrm{GeV})$ above.

At the same time, as emphasized in Sec. IV D, if the normalization point is introduced via the upper bound in the integral over the energy of the excited states, the inequality $\mu_\pi^2 > \mu_G^2$ holds for any normalization point (in other words, such regularization does not violate the positivity of the Pauli operator for the spinor quark). Therefore, it is legitimate to take the normalization point as low as 1 GeV. In this case, obviously, one deals with μ_G^2 normalized at this low point as well, and it is known [55] that the perturbative evolution *increases* its value toward lower μ. In our estimates we took μ_G^2 directly from the hyperfine splitting of B and B^*; therefore that value corresponded to $\mu \simeq 4.5\,\mathrm{GeV}$. Apparently its hybrid logarithmic enhancement would safely make up for the relatively insignificant subtraction of the "perturbative" contribution to μ_π^2. Based on these arguments we have stated in a previous paper [42] that the inequality

$$\mu_\pi^2 > \mu_G^2$$

must survive in QCD, in spite of recent claims [50] that it cannot hold true.

It is instructive to trace how this inequality works at different scales μ. Most trivially it is fulfilled when μ is taken parametrically large. Then μ_π^2 contains a large *positive* perturbative piece of the order of $\frac{\alpha_s}{\pi}\mu^2$ that grows faster than any possible change in μ_G^2 having no additive renormalization, even if the hybrid anomalous dimension of the latter were negative.

A more interesting consideration emerges when one wants to push μ toward lower values. Using the naive one-loop expression for the evolution of μ_G^2 corresponding to the hybrid anomalous dimension $\gamma_G = 3$ one would obtain an arbitrarily large value for μ_π^2 which, of course, makes little sense. The answer to this apparent paradox is rather obvious, especially if one looks at a hypothetical zero recoil excitation curve (for the external current $\bar{c}\gamma_i b$) similar to the one depicted in Fig. 4. The real evolution of the difference $\mu_\pi^2 - \mu_G^2$ at $m_q \gg m_Q$, according to the sum rule (121), is given by the decay probability occurring at energy $\epsilon = \mu$; obviously the latter in no way is given at low ϵ by the simple perturbative formulas based on the strong coupling α_s with the Landau pole, and rather stays finite at any μ. In other words, the evolution of μ_G^2 is to be smooth even when one approaches the strong coupling regime, and no formal contradiction emerges.

VIII. CONCLUSIONS

In the present paper we have addressed weak transitions between heavy quarks from the "inclusive" side most suitable theoretically for applying the technique of the Wilson OPE. This analysis naturally extend the consideration outlined in Ref. [15] which concerned the heavy quark distribution function relevant for the decays in the limit of small velocity for the final state hadron system. It has been demonstrated that a few sum rules discussed so far in the literature are in fact relations for the moments of a single distribution function, considered in different orders in $1/m_Q$ and in different kinematical regimes. We have shown how the expansions of HQET can consistently be obtained from QCD using this strategy, and in this way illustrated that such natural assumptions as "global duality," which usually are attributed only to the inclusive width calculations, are in fact necessary ingredients in *any* consistent model-independent treatment and, in particular, are implicitly used in HQET as well.

The analysis of the sum rules proved to be very instructive in elucidating the important fact that has usually been neglected in HQET, the necessity of introducing an explicit infrared normalization point μ ensuring true separation of low- and high-momentum physics, which *cannot* be set to zero. The consistent application of this approach leads to the fact that all nonperturbative parameters, including $\overline{\Lambda}$, cannot be sensibly defined as universal constants, but rather depend explicitly on the normalization point. This has been previously mentioned in our paper [15] and discussed in detail in Refs. [28, 29]. Here we gave a physical illustration of how it works, analyzing possible constructive phenomenological definitions of corresponding quantities in the presence of radiative corrections. In this way we have supplemented the previous calculations by estimates of the dependence of the kinetic energy operator $\bar{Q}\,\vec{\pi}^2\,Q$ on the renormalization point.

The fact that such "purely nonperturbative" objects such as the pole mass of the heavy quark, $\overline{\Lambda}$, a "purely nonperturbative" distribution function of heavy quarks

routinely used in HQET, are incompatible with a consistent OPE-based approach and are ill defined theoretically calls for the clarification of how the known results on nonperturbative corrections in HQET must be interpreted. This does not mean of course that the concrete calculations that have been done so far are irrelevant, and we formulated the way in which they are to be understood for a few typical examples.

As a practical application of our sum rules we have derived a model-independent lower bound on the deviation of the exclusive axial-vector form factor $F_{B\to D^*}$ of the $B \to D^* + l\nu$ decay at zero recoil, a process that for a long time has been believed to give the best theoretical accuracy to determine $|V_{cb}|$, and estimated a reasonable "central" value $F_{B\to D^*}(\vec{q}=0) \simeq 0.9$. The deviation appears to be essentially larger than the estimates that had been obtained before from model calculations based on standard HQET, and apparently better agree with quite general expectations about the size of corrections to the heavy quark symmetry for charmed particles. On the other hand, the theoretical clarification of the notion of the heavy quark mass made in recent papers [28, 29] suggests that the most accurate theoretical way to determine the CKM matrix elements for heavy quark decays is using the *inclusive* semileptonic widths. These results have been reported in Ref. [42].

It is worth clarifying in this respect that in our estimates of the exclusive form factors of the $b \to c$ transitions we consistently took into account terms through order $1/m_Q^2$ and discarded effects that scale like $1/m_Q^3$. The parameter $1/m_c$ is actually not very small and even the second-order corrections are as large as 10% here; therefore one can expect sizable relative corrections for real form factors due to higher-order terms. In particular, this applies to the model-independent upper bound for $F_{B\to D^*}$. Our result is strict in the sense that it holds for corrections through terms of order $1/m_Q^2$ that have been addressed in the literature so far.

One of the sum rules at zero recoil enabled us to rederive a model-independent lower bound on the value of the kinetic energy operator in B mesons,

$$\mu_\pi^2 = \frac{1}{2M_B}\langle B|\bar{b}\,\vec{\pi}^2\,b|B\rangle \gtrsim \frac{3}{4}\left(M_{B^*}^2 - M_B^2\right),$$

in a way that clearly showed its physical relevance even in real QCD, and not only in the approximate framework of quantum-mechanical considerations, as is sometimes stated.

ACKNOWLEDGMENTS

N.U. gratefully acknowledges the creative atmosphere and hospitality at Theoretical Physics Institute of University of Minnesota, and the exchange of ideas related to the subject of this paper with V. Braun. We would like to thank B. Blok, A. Falk, T. Mannel, M. Savage, and especially L. Koyrakh and M. Voloshin for useful discussions. This work was supported in part by the National Science Foundation under Grant No. PHY 92-13313 and by the DOE under Grant No. DE-FG02-94ER40823.

APPENDIX

Let us derive expression (94),

$$\frac{\bar{\Lambda}}{2} = \sum_n \frac{1}{3}\frac{|\langle B|\pi_i|n\rangle|^2}{E_n - M_B},$$

in ordinary quantum mechanics. We will use the nonrelativistic normalization of states below.

It is easy to see that in a simple potential description of the heavy hadron as a two-body system with the reduced mass $m_r \simeq m_{\rm sp}$ the sum on the right-hand side is given just by half of this mass (which in this case is identified with $\bar{\Lambda}$). Indeed,

$$-\sum_n \frac{|\langle B|\vec{\pi}\cdot\vec{v}|n\rangle|^2}{E_n - M_B}$$

represents an order $\vec{v}^{\,2}$ correction to the ground state energy, produced by a perturbation:

$$\delta H = \vec{\pi}\cdot\vec{v}.$$

On the other hand,

$$H + \delta H = H(\vec{\pi} + \vec{v}) - \frac{m_r\vec{v}^{\,2}}{2}.$$

The eigenvalues of the Hamiltonian given by the first term on the right-hand side are the same as for the unperturbed one, which leads to the relation

$$\sum_n \frac{|\langle B|\vec{\pi}\cdot\vec{v}|n\rangle|^2}{E_n - M_B} = \frac{m_r\vec{v}^{\,2}}{2}, \tag{A1}$$

which coincides with Eq. (94) if $\bar{\Lambda} = m_{\rm sp}$. In reality, of course, Eq. (94) is more general and accounts also for the binding energy. To show it we can apply the following general consideration.

First, let us note that a relation similar to Eq. (A1) can be obtained in a more general way. Namely, for any system whose Hamiltonian depends on the heavy quark momentum π in a nonrelativistic quadratic way,

$$\mathcal{H} = \frac{\vec{\pi}^2}{2m_Q} + \mathcal{H}_{\rm light}(\vec{x}_Q, \{x_{\rm light}\}), \tag{A2}$$

one has the exact commutation relation

$$\vec{\pi} = im_Q[\mathcal{H}, \vec{x}_Q], \tag{A3}$$

where $\vec{\pi}$ and $\vec{x}_Q$ are the operators of the heavy quark momentum and coordinate. Then one writes

$$\sum_n \frac{|\langle 0|\pi_k|n\rangle|^2}{E_n - E_0} = i\frac{m_Q}{2}\sum_n \left(\frac{\langle 0|\pi_k|n\rangle\langle n|[\mathcal{H}, x_k^Q]|0\rangle}{E_n - E_0} + \frac{\langle 0|[\mathcal{H}, x_k^Q]|n\rangle\langle n|\pi_k|0\rangle}{E_n - E_0} \right)$$
$$= i\frac{m_Q}{2}\sum_n \left(\langle 0|\pi_k|n\rangle\langle n|x_k^Q|0\rangle - \langle 0|x_k^Q|n\rangle\langle n|\pi_k|0\rangle \right)$$
$$= \frac{m_Q}{2}\sum_n \langle 0|i[\pi_k, x_k^Q]|0\rangle = \frac{m_Q}{2}. \tag{A4}$$

Note that we do not sum over spatial index k in this equation; it is valid for arbitrary k. Below we will assume that the summation over k is *not* performed. We emphasize that this relation is rigorous for any system as long as the heavy quark momentum $\vec{\pi}$ enters the Hamiltonian quadratically.

Now we apply Eq. (A4) to the B meson. We get

$$\frac{m_Q}{2} = \sum_n \frac{|\langle B|\pi_k|n\rangle|^2}{E_n - M_B}. \tag{A5}$$

The sum over intermediate states runs over all possible hadronic states which are marked, in particular, by their total momentum $\vec{\pi}$. The matrix elements, however, do not vanish only for zero momentum transfer:

$$\langle B(\vec{p}=0)|\pi_k|n(\vec{p})\rangle = \langle B|\pi_k|n\rangle_{\rm QM}\frac{(2\pi)^3\delta^3(\vec{p})}{\sqrt{V}}, \tag{A6}$$

where V is the volume. When squared only the equal momentum matrix elements are present, and the factor $(2\pi)^3\delta^3(0) = V$ cancels against $1/\sqrt{V}$ normalization of states associated with the continuous spectrum of total momentum. The situation requires more care when the state n is a B meson; the energy denominator $E_n - M_B = \vec{p}^{\,2}/2M_B$ has a pole in this case when integrated over $d^3\vec{p}$, and this singularity must be treated properly. We write therefore

$$\frac{m_Q}{2} = \sum_{n\neq B} \frac{|\langle B|\pi_k|n\rangle_{\rm QM}|^2}{E_n - M_B} + \int \frac{d^3\vec{p}}{(2\pi)^3}\,\frac{|\langle B(\vec{p}=0)|\pi_k|B(\vec{p})\rangle|^2}{\vec{p}^{\,2}/2M_B}. \tag{A7}$$

The matrix elements in the last term can be represented in the form

$$\langle B(\vec{p}=0)|\pi_k|B(\vec{p})\rangle = \frac{m_Q}{M_Q}\langle B(\vec{p}=0)|P_k|B(\vec{p})\rangle, \tag{A8}$$

where P is the total momentum operator. To calculate the integral over the momentum of the B meson we can again use Eq. (A4):

$$\int \frac{d^3\vec{p}}{(2\pi)^3}\,\frac{|\langle B(\vec{p}=0)|P_k|B(\vec{p})\rangle|^2}{\vec{p}^{\,2}/2M_B} = \frac{M_B}{2}; \tag{A9}$$

this is nothing but relation (A4) for "quantum mechanics" of free B mesons considered as elementary particles. Combining Eqs. (A7)–(A9) we finally get

$$\frac{m_Q}{2} = {\sum}' \frac{|\langle B|\pi_k|n\rangle_{\rm QM}|^2}{M_n - M_B} + \frac{m_Q^2}{2M_B}$$

or

$${\sum}' \frac{|\langle B|\pi_k|n\rangle_{\rm QM}|^2}{M_n - M_B} = \frac{m_Q}{M_B}\,\frac{M_B - m_Q}{2} \simeq \frac{\bar{\Lambda}}{2}. \tag{A10}$$

This equation is clearly the relation (94) we are looking for; if one sums over k, an additional factor of 3 emerges to match the exact coefficient in the latter.

[1] M. Shifman and M. Voloshin (unpublished); see V. Khoze and M. Shifman, Usp. Fiz. Nauk **140**, 3 (1983) [Sov. Phys. Usp. **26**, 387 (1983)].

[2] N. Bilić, B. Guberina, and J. Trampetić, Nucl. Phys. **B248**, 261 (1984); M. Voloshin and M. Shifman, Yad. Fiz. **41**, 187 (1985) [Sov. J. Nucl. Phys. **41**, 120 (1985)]; Zh. Eksp. Teor. Fiz. **91**, 1180 (1986) [Sov. Phys. JETP **64**, 698 (1986)].

[3] M. Voloshin, N. Uraltsev, V. Khoze, and M. Shifman, Yad. Fiz. **46**, 181 (1987) [Sov. J. Nucl. Phys. **46**, 112 (1987)].

[4] J. Chay, H. Georgi, and B. Grinstein, Phys. Lett. B **247**, 399 (1990).

[5] I.I. Bigi and N.G. Uraltsev, Phys. Lett. B **280**, 120 (1992).

[6] I. Bigi, N. Uraltsev, and A. Vainshtein, Phys. Lett. B **293**, 430 (1992); **297**, 477(E) (1993).

[7] B. Blok and M. Shifman, Nucl. Phys. **B399**, 441 (1993); **399**, 459 (1993).

[8] I. Bigi, B. Blok, M. Shifman, N. Uraltsev, and A. Vainshtein, in *The Fermilab Meeting*, Proceedings of the Meeting of the Division of Particles and Fields of the APS, Batavia, Illinois, 1992, edited by C.H. Albright *et al.* (World Scientific, Singapore, 1993), Vol. 1, p. 610.

[9] I. Bigi, M. Shifman, N. Uraltsev, and A. Vainshtein, Phys. Rev. Lett. **71**, 496 (1993).

[10] B. Blok, L. Koyrakh, M. Shifman, and A. Vainshtein, Phys. Rev. D **49**, 3356 (1994).

[11] I. Bigi and N. Uraltsev, Nucl. Phys. **B423**, 33 (1994).

[12] A. Manohar and M. Wise, Phys. Rev. D **49**, 1310 (1994).

[13] T. Mannel, Nucl. Phys. **B413**, 396 (1994).

[14] M. Neubert, Phys. Rev. D **49**, 3392 (1994).

[15] I. Bigi, M. Shifman, N. Uraltsev, and A. Vainshtein, Int. J. Mod. Phys. A **9**, 2467 (1994).

[16] A. Falk, E. Jenkins, A. Manohar, and M. Wise, Phys. Rev. D **49**, 4553 (1994).
[17] M. Neubert, Phys. Rev. D **49**, 4623 (1994).
[18] I. Bigi, M. Shifman, N. Uraltsev, and A. Vainshtein, Phys. Lett. B **328**, 431 (1994).
[19] T. Mannel and M. Neubert, Phys. Rev. D **50**, 2037 (1994).
[20] J. Bjorken, in *Results and Perspectives in Particle Physics*, Proceedings of the 4th Rencontres de Physique de la Vallèe d'Aoste, La Thuile, Italy, 1990, edited by M. Greco (Editions Frontières, Gif-Sur-Yvette, France, 1990).
[21] N. Isgur and M. Wise, Phys. Rev. D **43**, 819 (1991).
[22] J.D. Bjorken, I. Dunietz, and J. Taron, Nucl. Phys. **B371**, 111 (1992).
[23] M. Voloshin and M. Shifman, Yad. Fiz. **47**, 801 (1988) [Sov. J. Nucl. Phys. **47**, 511 (1988)].
[24] M. Voloshin, Phys. Rev. D **46**, 3062 (1992).
[25] A.G. Grozin, Report No. BudkerINP 92-97, 1992, Pt. 1 (unpublished).
[26] M. Burkardt, Phys. Rev. D **46**, R1924 (1992); **46**, R2751 (1992).
[27] H. Lipkin, Phys. Lett. B **308**, 105 (1993); Nucl. Phys. **A560**, 548 (1993).
[28] I. Bigi, M. Shifman, N. Uraltsev, and A. Vainshtein, Phys. Rev. D **50**, 2234 (1994).
[29] M. Beneke and V. Braun, Nucl. Phys. **B426**, 301 (1994); M. Beneke, V. Braun, and V.I. Zakharov, Phys. Rev. Lett. **73**, 3058 (1994).
[30] K. Wilson, Phys. Rev. **179**, 1499 (1969); K. Wilson and J. Kogut, Phys. Rep. **12**, 75 (1974).
[31] E. Eichten and B. Hill, Phys. Lett. B **234**, 511 (1990); H. Georgi, *ibid.* **240**, 447 (1990). For recent reviews see, e.g., N. Isgur and M.B. Wise, in *B Decays*, edited by S. Stone (World Scientific, Singapore, 1992); H. Georgi, in *Perspectives in the Standard Model*, Proceedings of the Theoretical Advanced Study Institute, Boulder, Colorado, 1991, edited by R.K. Ellis, C.T. Hill, and J.D. Lykken (World Scientific, Singapore, 1992); M. Neubert, Phys. Rep. **245**, 259 (1994).
[32] For a review see, e.g., B.L. Ioffe, V.A. Khoze, and L.N. Lipatov, *Hard Processes* (North-Holland, Amsterdam, 1984).
[33] M. Shifman, A. Vainshtein, and V. Zakharov, Nucl. Phys. **B147**, 385 (1979) [reprinted in *Vacuum Structure and QCD Sum Rules*, edited by M. Shifman (North-Holland, Amsterdam, 1992), p. 24].
[34] V. Novikov, M. Shifman, A. Vainshtein, and V. Zakharov, Nucl. Phys. **B191**, 301 (1981).
[35] S. Balk, J.G. Körner, and D. Pirjol, Nucl. Phys. **B428**, 499 (1994).
[36] J. Bjorken and S. Drell, *Relativistic Quantum Fields* (McGraw-Hill, New York, 1965).
[37] T. Mannel, Phys. Rev. D **50**, 428 (1994).
[38] Reference [6]; for more details see Report No. FERMILAB-PUB-92/158-T, 1992 (unpublished).
[39] L. Maiani, G. Martinelli, and C. Sachrajda, Nucl. Phys. **B368**, 281 (1992).
[40] ARGUS Collaboration, H. Albrecht *et al.*, Z. Phys. C **57**, 533 (1993); D.Z. Besson, in *Lepton and Photon Interactions*, Proceedings of the 16th International Symposium, Ithaca, New York, 1993, edited by P. Drell and D. Rubin, AIP Conf. Proc. No. 302 (AIP, New York, 1994); S. Stone, in *B Decays*, 2nd ed., edited by S. Stone (World Scientific, Singapore, 1994).
[41] A. Falk and M. Neubert, Phys. Rev. D **47**, 2965 (1993).
[42] M. Shifman, N. Uraltsev, and A. Vainshtein, Phys. Rev. D **51**, 2217 (1995).
[43] M.E. Luke, Phys. Lett. B **252**, 447 (1990).
[44] N. Isgur and M. Wise, Phys. Lett. B **232**, 113 (1989); **237**, 527 (1990).
[45] S. Nussinov and W. Wetzel, Phys. Rev. D **36**, 130 (1987).
[46] M. Voloshin, Report No. TPI-MINN-94/18-T (unpublished).
[47] P. Ball and V. Braun, Phys. Rev. D **49**, 2472 (1994); for earlier estimates see, M. Neubert, *ibid.* **46**, 1076 (1992); E. Bagan, P. Ball, V. Braun, and H. Dosch, Phys. Lett. B **278**, 457 (1992).
[48] E. Bagan, P. Ball, V. Braun, and P. Gosdzinsky, Phys. Lett. B **342**, 362 (1995); also V. Braun (private communication).
[49] N.G. Uraltsev, Report No. PNPI/T/B-9/94, hep-ph/9409382 (unpublished).
[50] M. Neubert (unpublished); see also, Phys. Lett. B **322**, 419 (1994).
[51] L. Koyrakh, N. Uraltsev, and A. Vainshtein (in preparation).
[52] I. Bigi, A. Grozin, M. Shifman, N. Uraltsev, and A. Vainshtein, Phys. Lett. B **339**, 160 (1994).
[53] H. Lipkin, *Quantum Mechanics* (North-Holland, Amsterdam, 1973).
[54] V. Novikov, M. Shifman, A. Vainshtein, and V. Zakharov, Nucl. Phys. **B249**, 445 (1985) [reprinted in *Vacuum Structure and QCD Sum Rules*, edited by M. Shifman (North-Holland, Amsterdam, 1992), p. 161].
[55] A. Falk, B. Grinstein, and M. Luke, Nucl. Phys. **B357**, 185 (1991).