CERN-TH/95-305
hep-ph/9511345

POLARIZED STRUCTURE FUNCTIONS: A THEORETICAL UPDATE

Stefano Forte*

Theory Division, CERN,
CH-1211 Genève 23, Switzerland.

Abstract

We review recent developments in the theory and phenomenology of polarized structure functions. We summarize recent experimental data on the proton and deuteron structure function g_1, and their impact on the understanding of polarized sum rules. Specifically, we discuss how accurate measurements of the singlet and nonsinglet first moment of g_1 test perturbative and nonperturbative QCD, and critically examine the way these measurements are arrived at. We then discuss how the extraction of structure functions from the data can be improved by means of a theoretical analysis of their x and Q^2 dependence, and how, conversely, experimental information on this dependence can be used to pin down the polarized parton content of the nucleon.

Invited talk at the VI Blois Workshop
Frontiers in Strong Interactions
Blois, France, June 1995

to be published in the proceedings

CERN-TH/95-305
November 1995

* On leave from INFN, Sezione di Torino, via P. Giuria 1, I-10125 Turin, Italy (address after December 1, 1995).

1. The state of the art

The state of the knowledge of polarized structure functions measured in deep-inelastic scattering has progressed during the last couple of years from the level of testing naive parton model ideas to the point of probing QCD at leading order (LO) and next-to-leading order (NLO). On the one hand, experiments at CERN and SLAC[1−4)] have accurately measured the structure function g_1, which determines the leading twist contribution to the longitudinal polarization asymmetry, for proton and deuteron targets, over a reasonably wide range in x and at a pair of values of Q^2. On the other hand, the perturbative behavior of g_1 in the (x, Q^2) plane can now be quantitatively studied at NLO[5)], thanks to the recent determination of the full set of two-loop polarized anomalous dimensions[6)]. This means that polarized structure function data can now be used to test perturbative QCD and obtain detailed information on the structure of polarized nucleons, while, conversely the data can be understood and analyzed using accurate NLO methods.

Here we will review these recent developments, emphasizing the impact of the new data on the theoretical understanding of polarized structure functions. As it often happens, quantities that are easier to understand theoretically are the hardest to measure, and vice versa. We will take a theorist's viewpoint: we will start from the experimental results which have the most direct theoretical interpretation, and discuss them without, at first, questioning the procedure through which they have been obtained; we will then proceed to quantities whose interpretation is more subtle, and eventually discuss the procedure which is used to extract information from the data.

In Sect. 2 we will review the determinations of the first moment of g_1 (as given by the experimental collaborations) and discuss in particular the behavior of the isotriplet first moment for which there is an absolute QCD prediction (the Bjorken sum rule). In Sect. 3 we will turn to the singlet component of the first moment, discuss its extraction from the data and its theoretical interpretation, which is complicated by the presence in this channel of the axial anomaly. In Sect. 4 we will then turn to the full x, Q^2 dependence of the structure function g_1; we will discuss how it is described in perturbative QCD, and how it may be used to obtain information on the general structure of polarized parton distributions. Finally, we will come full circle, and show how this information can be used to improve the precision of the extraction from the data of quantities of simple and direct theoretical relevance, such as the moments of parton distributions.

In this review we will concentrate on the physics of g_1[7)], the discussion of other polarized structure functions being outside the scope of our treatment. It is however worth mentioning that first measurements of g_2, the other structure function which contributes to the deep-inelastic polarized cross section have been performed recently[8)]. While the theoretical interpretation of this structure function is worth studying for its own sake[9)], its contribution to the longitudinally polarized cross-section, on which we will concentrate, vanishes asymptotically (as $\frac{1}{Q^2}$). In this contex, it is thus important mostly as a background: the recent experiments find that g_2 is rather small, and essentially compatible with zero within errors, thus supporting the view that g_1 may be accurately determined by simply neglecting g_2.

2. The nonsinglet first moment and the Bjorken sum rule

The main outcome of the recent precise experiments[1−4)] is a set of determinations of the first moment

$$\Gamma_1(Q^2) = \int_0^1 dx\, g_1(x, Q^2) \tag{2.1}$$

Ref.	Target	$\langle Q^2\rangle$	Γ_1	a_0
1	p	10	$0.136 \pm 0.011 \pm 0.011$	0.22 ± 0.14
2	p	3	$0.129 \pm 0.004 \pm 0.009$	0.29 ± 0.10
3	d	10	$0.034 \pm 0.009 \pm 0.006$	0.20 ± 0.11
4	d	3	$0.042 \pm 0.003 \pm 0.004$	0.30 ± 0.06

Table 1: Summary of recent experimental determinations of Γ_1 eq. (2.1) and a_0 eq. (2.4) (as presented by the experimental collaborations). The first error on Γ_1 is statistical and the second systematic. All results hold at the average scale $\langle Q^2\rangle$ (in GeV^2) of the respective experiments.

for proton and deuteron targets, displayed in table 1.* We will discuss later the theoretical input that goes into these determinations. If we take them at face value, however, they provide us directly with information that admits a simple theoretical interpretation, because Γ_1 measures the nucleon matrix element of the axial current:

$$M^t a_i^t s^\mu \equiv \langle t; p, s|\bar\psi_i \gamma_\mu \gamma_5 \psi_i|t; p, s\rangle, \tag{2.2}$$

$$\Gamma_1^t(Q^2) = \frac{1}{2}\sum_{i=1}^{n_f} e_i^2 C_i(Q^2) a_i^t = \frac{1}{2}\left[\langle e^2\rangle C_{\rm S}(Q^2) a_0^t(Q^2) + C_{\rm NS}(Q^2) a_{\rm NS}^t(Q^2)\right], \tag{2.3}$$

where t indicates the target hadron (proton or deuteron, in our case) with mass M^t, momentum p^μ and spin s^μ, and the sum runs over all activated quark flavors; in the sequel when dropping the label t we will tacitly refer to a proton target. In the last step we have introduced the average quark charge $\langle e^2\rangle = \frac{1}{n_f}\sum_{i=1}^{n} e_i^2$, and the singlet and nonsinglet axial charges

$$a_0 = \sum_{i=1}^{n_f} a_i \tag{2.4}$$

$$a^{\rm NS} = \sum_{i=1}^{n_f}\left(e_i^2 - \langle e^2\rangle\right) a_i, \tag{2.5}$$

exploiting the fact that the Wilson coefficient functions $C(Q^2)$ only distinguish between singlet and nonsinglet currents.

The nonsinglet axial charge is particularly simple because the corresponding current is conserved (neglecting quark masses) and can therefore be decomposed into a sum of scale–independent contributions:

$$\begin{aligned} a_{\rm NS}(Q^2) &= \sum_{k=2}^{n_f}\left[\left(\langle e^2\rangle_{k-1} - \langle e^2\rangle_k\right)\Theta(Q^2 - Q_k^2)\left(\left[\sum_{i=1}^{k-1} a_i\right] - (k-1)a_k\right)\right] \\ &= \frac{1}{6}(a_u - a_d) + \frac{1}{18}(a_u + a_d - 2a_s) \\ &\quad - \frac{1}{18}\Theta(Q^2 - Q_c^2)(a_u + a_d + a_s - 3a_c) + \dots \end{aligned} \tag{2.6}$$

* There exists also a direct determination of the neutron structure function and its first moment, obtained from scattering on a ^{3}He target[10)]. We will not include this result in our discussion, because of its large statistical and systematic uncertainties — partly due to the problems related to scattering on a nuclear target[7)]— and also because a recent reanalysis of the same raw data has lead to a rather different determination of the corresponding structure function[11)].

where $\langle e^2\rangle_k$ is the average charge computed with $n_f = k$ flavors, Q_c^2 is the threshold for the k-th flavor, and all the scale dependence comes from the function $\Theta(Q^2 - Q_k^2)$, which vanishes below threshold and is equal to 1 above threshold (having a nucleon target in mind in the last step we explicitly displayed thresholds for the charm and heavier quarks).

The simplest nonsinglet scale-independent currents are then those related to light quarks, i.e. the isotriplet axial charge g_A and the SU(3) octet charge a_8

$$g_A = a_u - a_d; \qquad a_8 = a_u + a_d - 2a_s. \tag{2.7}$$

Using isospin symmetry, knowledge of Γ_1 for proton and neutron targets determines the isotriplet combination:

$$\begin{aligned}\Gamma_1^{I=1} \equiv \Gamma_1^p - \Gamma_1^n =& C_{\rm NS}(Q^2)\frac{1}{6}\left(a_u^p - a_d^p\right)\\ =&\Bigg[1 - \left(\frac{\alpha_s}{\pi}\right) - \left(\tfrac{55}{12} - \tfrac{1}{3}n_f\right)\left(\frac{\alpha_s}{\pi}\right)^2\\ &- \left(41.4399 - 7.6072 n_f + \tfrac{115}{648}n_f^2\right)\left(\frac{\alpha_s}{\pi}\right)^3 + O(\alpha_s^4)\Bigg]\frac{1}{6}g_A.\end{aligned} \tag{2.8}$$

The only assumption that goes into the derivation of eq. (2.8) is that of exact isospin symmetry, which implies $a_u^p = a_d^n$, $a_d^p = a_u^n$ and $a_h^p = a_h^n$, where h indicates strange or heavy quark flavors. Furthermore, using isospin algebra again, g_A in eq. (2.7) can be related to the matrix element of the axial current which induces beta decay, and thus to the beta decay constant of the same name, which is known rather accurately: $g_A = 1.2573 \pm 0.0028$. The relation eq. (2.8) between the beta decay constant and the isotriplet first moment of g_1 is known as the Bjorken sum rule[12)]. In practice, the neutron structure function can be extracted from the deuteron one by using additivity of the deep-inelastic cross sections after correcting for nuclear effects[13)]:

$$\Gamma_1^d = (1 - 1.5\omega_D)\frac{\Gamma_1^p + \Gamma_1^n}{2}, \tag{2.9}$$

where $\omega_D \approx 0.05$ is the probability that the deuteron is in a D-wave state.

It thus looks as though the Bjorken sum rule (2.8) is an ideal place to test QCD in a clean and simple way: in principle, with a pair of determinations of $\Gamma_1^{I=1}$ at different scales both its normalization and its scale dependence may be measured, and thus isospin symmetry (which determines the value of g_A) and perturbative QCD [which determines $C_{\rm NS}(Q^2)$] may be simultaneously tested. Now, isospin symmetry violation in QCD is expected to be suppressed by powers of the current quark mass on a typical strong interaction scale; in particular in the unpolarized quark distributions it is at most of a few per cent [14)]; the observation of larger isospin violation in this channel would presumably be the sign of nonperturbative physics and rather interesting per se [15)].

On the other hand, $C_{\rm NS}(Q^2)$ is known up to order α_s^3, i.e., at the four–loop order, thereby allowing very precise tests of perturbation theory. In fact, at this order it might already be possible to see[16)] manifestations of the fact that the perturbative expansion in powers of α_s of QCD observables diverges[17)]. A detailed treatment of this issue is beyond the scope of the present paper; its main implication for our purposes is that because of its divergent nature the perturbative series is ambiguous, in that different resummations of the series may differ by terms which are proportional to powers of $\frac{\Lambda^2_{\rm QCD}}{Q^2}$. Because the values of physical observables are unambiguous, these ambiguities must cancel against corresponding ones in other power-suppressed contributions, i.e. higher twist contributions.

Indeed, in general eq. (2.8) should read

$$\Gamma_1^{I=1} = \frac{1}{6}\left[C_{NS}(Q^2)g_A + \frac{C_{\rm HT}}{Q^2} + O\left(\frac{1}{Q^4}\right)\right]. \tag{2.10}$$

The coefficient $C_{\rm HT}$ can be computed once a prescription for the treatment of the perturbative expansion is specified; the result for Γ_1 is then unambiguous[18)]. This program has never been carried through explicitly (there exists a proposal[19)] based on an explicit cutoff scheme which however would require a determination of $C_{\rm HT}$ on the lattice). It is nevertheless possible to estimate the size of $C_{\rm HT}$ phenomenologically, for instance using QCD sum rules; this leads to[20)] $C_{\rm HT} \approx -0.1$ with an error of order 50% or more. Such a determination, however, is only meaningful if the ambiguity is substantially smaller than the value of $C_{\rm HT}$. Even though a qualitative argument[17)] suggests that this is the case, an estimate[16)] based on an analysis of the known terms in the expansion of $C_{\rm NS}$ with the method of Padé approximants indicates instead that the ambiguity is of the same size as $C_{\rm HT}$.

Be that as it may, these estimates give an indication of the level of accuracy at which perturbative QCD can be tested by comparing the Bjorken sum rule with the data. This comparison is shown in fig. 1, where $\Gamma_1^{I=1}$ is displayed as a function of Q^2 at various perturbative orders, with and without typical higher twist corrections, for different values of the strong coupling, and compared with an experimental value obtained[16)] by averaging all the data of table 1 (plus some less precise earlier data) evolved to a common scale (we will discuss this evolution in the next section):

$$\Gamma_1^{I=1}(3\ {\rm GeV}^2) = 0.164 \pm 0.011. \tag{2.11}$$

The comparison is striking: the result appears to be very sensitive to perturbative corrections — it would disagree with the theoretical prediction if this were computed at LO rather than NNLO — while being reasonably insensitive to higher twists and associated ambiguities. In fact, the value of α_s appears to be tested here to an accuracy comparable with the world average of all other determinations[21)]. In other words, if, rather than assuming isospin and using eq. (2.8) to measure α_s, we assign the value of α_s and test isospin, then the uncertainty due to the error on the value of α_s is already larger than the experimental uncertainty on the value of $\Gamma_1^{I=1}$ quoted in eq. (2.11).

This determination of $\Gamma_1^{I=1}$, however, while being a theorist's dream, is a phenomenologist's nightmare, and should be taken with extreme care. Firstly, the inclusion of nuclear effects in eq. (2.9) as a simple multiplicative correction factor may be an oversimplification: in a more detailed treatment of the deuteron wave function the correction is actually x dependent, and its effect on the first moment depends on the shape of the structure function $g_1(x)$; the effect has been argued to be non–negligible at the level of present-day accuracy, especially in the large–x region[23)]. Other nuclear effects, such as shadowing, which has a substantial effect[14,24)] on the unpolarized counterpart of $\Gamma_1^{I=1}$ (the isotriplet first moment of F_2), have been pointed out in this context[25)], but not studied systematically.

Furthermore, it is clear that the reason why this test of perturbative QCD appears to be so sensitive is that the measurement is performed at a low scale, which exponentially magnifies the sensitivity to perturbative evolution. At such a low scale, however, other sorts of higher twist corrections, besides those discussed above, could be relevant, in particular those

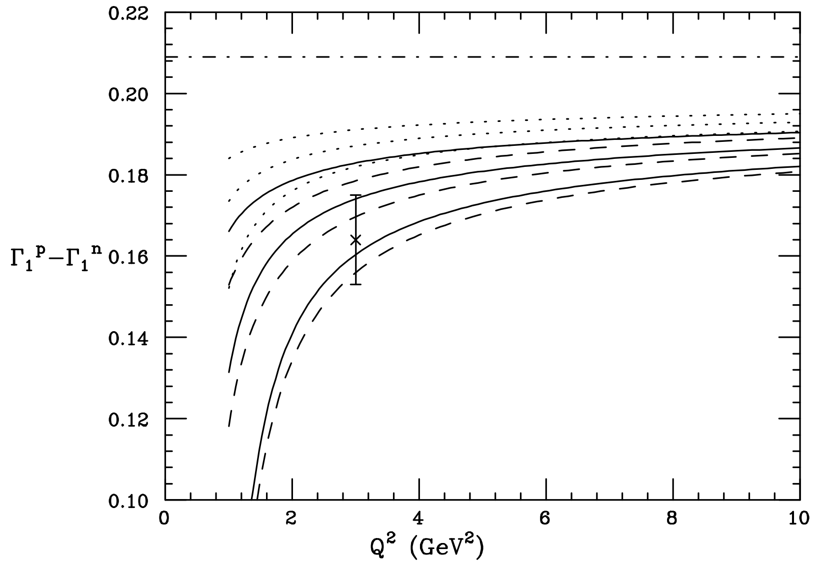

Figure 1: Scale dependence of $\Gamma_1^{I=1}$ eq. (2.10) [adapted from ref. [22]]. The dashed curve corresponds to eq. (2.8) and the solid curve to eq. (2.10) with $C_{\rm HT} = -0.1$; the dotted curve displays the $O(\alpha_s)$ contribution to eq. (2.8), and the dotdashed is the asymptotic ($\alpha_s = 0$) value. The three sets of curves correspond to the values $\alpha_s(M_z) = 0.118 \pm 0.007$. The data point is as in eq. (2.11).

related to the nucleon mass: even assuming that corrections due to transverse polarization (i.e. proportional to g_2) are negligible, there are still kinematic higher twist corrections such as target mass corrections[26]. Their effect on the first moment of $\Gamma_1^{I=1}$ is[27] comparable to the uncertainty in eq. (2.11) (which does not include it) at that scale. Moreover, such effects have never been studied systematically and could affect the extraction of all the first moments of table 1 from the data: even determinations with large $\langle Q^2 \rangle$ have several data points at low Q^2. Finally, the result (2.11) is obtained by evolving the full first moment Γ_1, which in turn is obtained from data taken at various values of Q^2. The corresponding uncertainty, related to perturbative evolution, could also be rather large at low scales: this will be the main subject of sect. 4.

In conclusion, the fact that $\Gamma_1^{I=1}$ is only measured at relatively low Q^2 allows testing of the Bjorken sum rule only to about 10% accuracy, due to the large sensitivity to the value of α_s. This could be turned around: assuming the validity of the sum rule (i.e. of exact isospin) the value of α_s can in principle be measured rather accurately. There is however a trade off, in that the determination of $\Gamma_1^{I=1}$ at a low scale is affected by substantial systematic uncertainties, and the errors in table 1 may turn out to be over–optimistic, as we will see explicitly in Sect. 4.

3. The singlet axial charge and the anomaly

The singlet component of Γ_1 can be extracted from the data in several distinct ways: by taking suitable linear combinations of different experiments, or by using SU(3) symmetry, or by exploiting the fact that the scale dependence of the singlet first moment differs from that of the nonsinglet. This scale dependence must also be taken into account when comparing different experiments, and turns out to pose some interesting theoretical questions. We will discuss the phenomenological and theoretical aspects of these issues in turn.

3.1. The first moment and its scale dependence

The simplest way of getting a value for the singlet first moment and the associated charge $a_0(Q^2)$ [eq. (2.4)] is to subtract the nonsinglet contribution off each experimental determination of Γ_1. The axial charge is then found from eq. (2.3) by dividing out the singlet coefficient function, which has been computed[28)] to NNLO:

$$C_{\rm S}(Q^2) = \left[1 - \left(\frac{\alpha_s}{\pi}\right) - \left(\tfrac{55}{12} - 1.16248 n_f\right)\left(\frac{\alpha_s}{\pi}\right)^2 + O(\alpha_s^3)\right]. \tag{3.1}$$

The values listed in table 1 were obtained in this way.

Neglecting all contributions from heavy quarks (we will come back to this assumption later) the nonsinglet is given by

$$\Gamma_1^{\rm nonsing.} = \frac{1}{6} C_{\rm NS}(Q^2)\left(g_A + \frac{1}{3}a_8\right). \tag{3.2}$$

Just as g_A can be obtained from the nucleon beta decay constants using isospin, a_8 can be obtained from hyperon beta decay constants using SU(3) symmetry, with the result $a_8 = 0.579 \pm 0.025$. This value is arrived at[29)] by a best fit based on the assumption of exact SU(3) symmetry; it could therefore by significantly affected by SU(3) symmetry breaking. This of course can only be introduced in a model-dependent way. One possibility is to introduce a parametrization of the hyperon beta decay constants which includes a term proportional to SU(3) breaking in the octet mass spectrum. A specific parametrization [30)] then leads to the value $a_8 = 0.40 \pm 0.19$. An alternative option is to allow for current mixing effects, i.e. let the singlet current have a nonvanishing SU(3) nonsinglet matrix element, whose size can then be estimated within a model for the quark content of the octet baryons[31)], leading to $a_8 \approx 0.4$ (and an uncertainty not smaller than the above) when SU(3) breaking is maximal.

Even though the precise value of a_8 is thus affected by a large uncertainty, this has a rather small effect on the extraction of the singlet component, because its numerical coefficient is one third of that of the triplet and one fourth of the singlet. In fact, using these values of g_A and a_8 in the decomposition (2.3) of the first moment into its singlet and nonsinglet components, it is easy to see that about 90% of the proton first moment comes from the isotriplet component, and of the remaining 10% about two-thirds are singlet, the rest being octet. For the deuteron the isotriplet contribution vanishes and the singlet and octet are partitioned as in the proton. This trivial numerology has some important consequences for the extraction of a_0 from the data: a) if the singlet is extracted from the proton data the uncertainty on a_0 will be about ten times larger than that on Γ_1; b) an uncertainty of about 50% in the knowledge of a_8 has the same effect as an uncertainty of a few per cent on the knowledge of the proton's Γ_1, i.e. it generates an uncertainty of about 10% in the proton or deuteron a_0; c) the deuteron is in principle a better probe of the singlet component. This is reflected by the values of a_0 in table 1, derived neglecting SU(3) breaking.

In order to meaningfully compare values of a_0 obtained in different experiments, they must be evolved to the same scale. Indeed, the singlet axial current is not conserved because of the axial anomaly, as we will discuss more extensively in the next section, so its matrix elements can acquire an anomalous dimension, which actually vanishes at LO, and has been determined up to NNLO[28)]:

$$a_0(Q^2) = \left[1 + \tfrac{6 n_f}{33 - 2 n_f}\left(\frac{\alpha_s}{\pi}\right) + \tfrac{3087 n_f + 138 n_f^2 + 4 n_f^3}{12(33 - 2n_f)^2}\left(\frac{\alpha_s}{\pi}\right)^2 + O(\alpha_s^3)\right] a_0(\infty). \tag{3.3}$$

The full scale dependence of the singlet contribution to Γ_1 [eq. (2.3)] is obtained by combining this with the singlet coefficient function eq. (3.1).

It is then possible to determine a_0 from a global fit to the data of table 1 (rather than by simply averaging the determinations coming from each data set): if g_A and a_8 are fixed using SU(3) symmetry as discussed above, $a_0(\infty)$ is then the only free parameter. A global fit (including also as older data which however carry little weight due to their poor accuracy) with $\alpha_s(M_z) = 0.118 \pm 0.007$ leads to[32)]

$$a_0(\infty) = 0.29 \pm 0.04, \tag{3.4}$$

where the error is almost entirely statistical, because of the magnification of the statistical error on Γ_1 when a_0 is extracted.

Given the availability of independent determinations of Γ_1 at different scales and for different targets, it is actually possible to relax the assumptions that went into the determination eq. (3.4). First, it is possible to relax the assumption of SU(2) symmetry, and extract simultaneously a_0 and g_A[32)]:

$$a_0(\infty) = 0.30 \pm 0.04; \quad g_A = 1.12 \pm 0.10\,^{+0.10}_{-0.04}. \tag{3.5}$$

The error on g_A has been decomposed in statistical and systematic; the latter is entirely due to the uncertainty on α_s, the uncertainty on a_8 being essentially negligible. This shows, consistent with the discussion in the previous section and fig. 1, that the Bjorken sum rule can be tested at best to about 10% accuracy, due to the uncertainty on the value of α_s. Finally, it should in principle be possible to determine simultaneously g_A, a_8 and a_0 from the data: the proton–deuteron comparison fixes separately the isotriplet (i.e. g_A) while the contributions of a_0 and a_8 to the isosinglet could then be separated by their different scale dependence. This is however not yet feasible in practice because this scale dependence is only rather slight, while the sensitivity of the value of Γ_1 to a_8 is weak; the latter fact, however, implies that even if we were to assume a 30% variation of a_8 due to SU(3) violation the value of a_0 (3.4)-(3.5) would hardly be affected.

We now turn to polarized heavy quark contributions, which can be generated dynamically by assuming that the corresponding current matrix elements vanish on threshold. This then determines the various scale-independent nonsinglet contribution to eq. (2.6) in terms of the singlet at each threshold: for instance, the charm contribution is

$$a_u + a_d + a_s - 3a_c = [a_u + a_d + a_s](Q_c^2) = a_0(Q_c^2). \tag{3.6}$$

Thus, because of the small numerical value of a_0, the error made neglecting heavy quark contributions to the nucleon Γ_1 is indeed rather small.

We can finally compare all the results of table 1 to one another and to the Bjorken sum rule prediction of eq. (2.8) by evolving to a common scale. The result, shown in fig. 2, demonstrates the mutual consistency of various experiments and the agreement with the prediction of exact isospin as tested by the Bjorken sum rule.

3.2. The "spin crisis" and the anomaly

The value of the singlet axial charge thus determined has attracted a good deal of attention because of its smallness. Small here refers to the expectation, based on the Zweig rule, that $a_0 \approx a_8$, which the value in eq. (3.4) violates by several standard deviations.

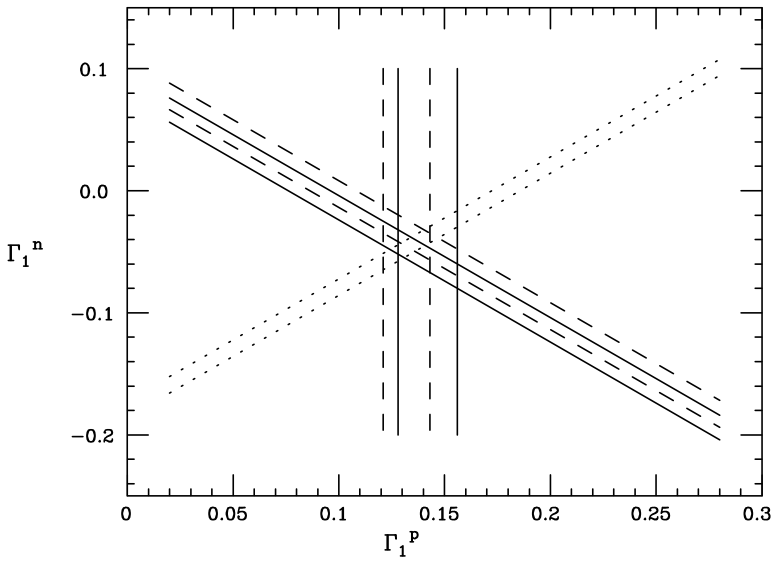

Figure 2: Values of the first moment Γ_1 eq. (2.1) from table 1 evolved to the common scale of $Q^2 = 5$ GeV^2. The solid lines correspond to the SMC experiments and the dashed lines to E143 experiments. The dotted line is the prediction of the Bjorken sum rule eq. (2.8)(as the solid lines in fig. 1).

However, a_8 is scale-independent, while a_0 depends on Q^2 according to eq. (3.3): this raises immediately the question of the scale at which these two quantities should be compared. Indeed, the Zweig rule is a phenomenological rule of the naive quark model. Matrix elements of operators measured in hard processes should agree with the predictions of this model at some typical low hadronic scale: it is however clear that, at a sufficiently low scale, $a_0(Q^2)$ [eq. (3.3)] grows as large as desired; in particular, it is of the same size as a_8 around 0.5 GeV^2. Whereas clearly perturbative evolution at such low scales cannot be trusted, nonperturbative estimates[33)] suggest that this scale dependence may actually be even stronger than perturbatively expected.

More specifically, one may take the point of view that naive constituent quark observables are related to parton distributions measured in hard processes by a (generally x– and Q^2–dependent) renormalization. A simple way of modelling this is to assume that the nucleon structure functions are given by combinations of constituent quark structure functions, each of which in turn is the convolution of a parton density inside the constituent quarks times the constituent quark density in the nucleon[32,34)]. The constituent quark densities are then scale independent but target-dependent, while the parton densities satisfy the Altarelli-Parisi equations and should be universal. The model is tested by verifying this universality (for instance, it allows predicting the second moment of the pion parton distributions in terms of the proton ones) and seems to be in good agreement with experimental data[35)]. The violation of the Zweig rule, which is then by construction satisfied by constituent quarks, is attributed to their partonic substructure: the apparent contradiction with the quark model is thus removed, but the observed effect is not explained.

A deeper understanding of the meaning of observed violation of the Zweig rule, and more in general of the first moment of g_1, can be obtained in the QCD parton model. Matrix elements of operators are endowed with a partonic interpretation by identifying them with moments of parton distributions. The definition of the operators is however ambiguous since a renormalized operator can always be redefined by a finite renormalization: given a specific definition of, say, the isotriplet axial current $j_{5\,I=1}^{\mu}$, we can define a new current

$$j_{5\,I=1}^{\mu\prime} = Z(\alpha_s) j_{5\,I=1}^{\mu}, \tag{3.7}$$

where $Z[\alpha_s(Q^2)] = 1 + Z^{(1)}\alpha_s(Q^2) + \ldots$. The relation of the current to physical observables will be modified accordingly: for example if $\Gamma_1^{I=1} = C_{\rm NS}(Q^2)\langle j_{5\,I=1}^{\mu}\rangle$, then $\Gamma_1^{I=1} = C_{\rm NS}(Q^2)Z^{-1}(Q^2)\langle j_{5\,I=1}^{\mu\prime}\rangle$. This is, a priori, as good a definition of the axial current as the original one, and the ensuing ambiguity (factorization scheme ambiguity) only goes away at $Q^2 \to \infty$, or to all perturbative orders. It may, however, be fixed by physical requirements, such as the preservation of physical symmetries. In particular, if a current is conserved (as the nonsinglet axial current) it is protected against scale dependence: because $\partial_\mu j_5^{\mu\,I=1} = 0$ then necessarily $Q^2 \frac{d}{dQ^2}\partial_\mu j_5^{\mu\,I=1} = 0$, hence j_5^{μ} (as a renormalized composite operator) does not depend on Q^2. It follows that the redefinition of eq. (3.7) is not allowed if one imposes that chiral symmetry be preserved, because it would spoil this scale independence, and there exists a "natural" normalization of the current, namely, that which preserves chiral symmetry.

The current matrix element can then be identified with the first moment of the polarized quark distribution: for instance the isotriplet charge is

$$a_{I=1} = g_A = \int_0^1 dx\, \Delta q_{I=1}(x) = \Delta u(1,Q^2) - \Delta d(1,Q^2), \tag{3.8}$$

where $\Delta u(x,Q^2)$, $\Delta d(x,Q^2)$ are the polarized up and down distributions, whose N-th moment is given by

$$\Delta q_i(N,Q^2) = \int_0^1 dx\, x^{N-1}\Delta q_i(x,Q^2); \tag{3.9}$$

and similarly for other nonsinglet charges. The conserved charge associated to $j_{5\,I=1}^{\mu}$ is the total (isotriplet) quark helicity, so $a_{I=1}$ can be interpreted as the total (isotriplet) quark polarization, in agreement with naive partonic expectation.

Things are more complicated in the singlet case; first, because there now are two parton distributions rather than one, i.e. the gluon distribution $\Delta g(x,Q^2)$ and the singlet quark distribution

$$\Delta\Sigma(x,Q^2) = \sum_{i=1}^{n_f}\Delta q_i(x,Q^2), \tag{3.10}$$

where $\Delta q_i(x,Q^2)$ is the quark distribution of flavor i and n_f is the number of flavors activated at the scale Q^2; and furthermore, because the current is no longer conserved, thus as a renormalized operator it depends on scale, and there is no unique "natural" way to normalize the quark singlet and gluon first moments, which will mix upon evolution.

Indeed, the classical conservation of the singlet axial current is spoiled at the quantum level by the axial anomaly[36]:

$$\partial_\mu j_5^{\mu} = n_f \frac{\alpha_s}{2\pi}{\rm tr}\epsilon^{\mu\nu\rho\sigma}F_{\mu\nu}F_{\rho\sigma}. \tag{3.11}$$

The singlet current is thus unprotected and requires an infinite renormalization, which induces a multiplicative scale dependence of the renormalized current of the form[37]

$$\frac{d}{dt}j_5^{\mu} = \gamma_5(\alpha_s) j_5^{\mu}, \tag{3.12}$$

where $t \equiv \ln(Q^2/\Lambda^2)$. The anomalous dimension $\gamma_5(\alpha_s) = \gamma_5^{(2)}\alpha_s^2 + \dots$ starts at two loops (because the nonconservation of the current is a one–loop effect) and depends on the specific choice of operator normalization; the scale dependence of a_0 induced by it was given in the $\overline{\rm MS}$ scheme in eq. (3.3).

Once this choice is made, however, the remaining normalization ambiguities are fixed by the anomaly equation. Indeed, eq. (3.11) holds as an equation between renormalized composite operators[38)], and it therefore implies that the operator $\mathrm{tr}\epsilon^{\mu\nu\rho\sigma}F_{\mu\nu}F_{\rho\sigma}$ must mix with the current in such a way as to compensate (to all orders in perturbation theory) its scale dependence:

$$\frac{d}{dt}n_f\frac{\alpha_s}{2\pi}\mathrm{tr}\epsilon^{\mu\nu\rho\sigma}F_{\mu\nu}F_{\rho\sigma} = \gamma_5(\alpha_s)\partial_\mu j_5^\mu. \tag{3.13}$$

This condition fixes the normalization of the singlet gluon operator that mixes with the axial current, as well as the coefficient of this mixing — that there exists a scheme where eq. (3.13) is satisfied is the content of the Adler–Bardeen theorem[37)] and modern versions thereof[38)].

However, there is still an obstacle in the identification of matrix elements of operators with partonic quantities: namely, we cannot simply identify the first moments $\Delta\Sigma(1,Q^2)$ and $\Delta g(1,Q^2)$ of the quark and gluon distribution [defined in analogy to eq. (3.9)] with matrix elements of two leading twist operators, because there is only one local leading twist (i.e. twist 2) operator with spin corresponding to the first moment of g_1 (i.e. spin 1). It follows that while it is natural to identify $\Delta g(1,Q^2)$ (up to an overall constant, still arbitrary at this stage), with the quantity that mixes with the matrix element a_0 of j_5^μ according to eq. (3.13), i.e.

$$-\frac{d}{dt}n_f\frac{\alpha_s}{2\pi}\Delta g(1,Q^2) = \gamma_5(\alpha_s)a_0(Q^2), \tag{3.14}$$

there is no unique way to identify $a_0(Q^2)$ itself. Indeed, if the quark singlet and gluon anomalous dimensions are computed for higher moments, where a tower of quark and gluon operators (for all odd $n \geq 3$) are naturally identified with the corresponding moments of quark and gluon distributions[39)], and then analytically continued to the first moment, the result depends on the adopted regularization: if dimensional regularization is used throughout[6)], then the anomalous dimension of $\Delta\Sigma(1,Q^2)$ coincides with $\gamma_5(\alpha_s)$; if instead infrared collinear singularities are regulated by means of an explicit regulator[39)] (such as putting the incoming particle off-shell) then the anomalous dimension of $\Delta\Sigma(1,Q^2) = \Delta\Sigma(1)$ vanishes. Note that in either case eq. (3.14) is satisfied, thereby fixing the constant of proportionality between $\Delta g(1,Q^2)$ and the operator which mixes with a_0 . These two results correspond, respectively, to identifying the first moment of the quark distribution with the two eigenvectors of perturbative evolution: in the former case the singlet quark distribution is identified as $\Delta\Sigma(1,Q^2) = a_0(Q^2)$, and in the latter case it is identified with the scale-invariant combination $\Delta\Sigma(1) = a_0(Q^2) + n_f\frac{\alpha_s}{2\pi}\Delta g(1,Q^2)$.

This ambiguity can be pinned down only on the basis of physical requirements on parton distributions. Now, it may be shown[40)] that, in schemes where $a_0 = \Sigma(1,Q^2)$, soft contributions are partly included in the hard coefficient function, rather than being properly factorized in the parton distributions: this explains why these contributions are removed by an infrared regulator, which instead yields schemes where the quark is identified with the conserved eigenvector. Furthermore, in the latter schemes $\Delta\Sigma(1)$ can be directly shown[41)] to coincide with the nucleon matrix element of the canonical, conserved quark helicity operator, the scale-dependence of a_0 then being due to a contribution to it from particle creation induced by the axial anomaly. In these schemes the Zweig rule expectation acquires a precise meaning[42], since it predicts the approximate equality of two scale–independent quantities: $a_8 \approx \Delta\Sigma(1)$. The (scale–dependent) deviation of a_0 from this value would then be explained by the gluon contribution to it[42]. Of course, whether this is actually the case rests with experiment: we will discuss this in sect. 4.2.

The choice of factorization scheme in the definition of polarized quark and gluon distribution, besides affecting the physical interpretation of the singlet first moment of g_1, has a significant effect in a finite-order computation of the scale dependence of Γ_1, and may thus substantially affect the extraction of this quantity from the data. Indeed, in the class of schemes where $\Delta\Sigma(1)$ is scale-independent, the singlet Γ_1 is given by

$$\Gamma_1^{\rm sing.} = \frac{1}{2}\langle e^2\rangle C_{\rm S}(Q^2)\left(\Delta\Sigma(1) - \frac{\alpha_s}{2\pi} n_f \Delta g(1,Q^2)\right) \tag{3.15}$$

$$= \frac{\langle e^2\rangle}{2}\left[C_q(Q^2)\Delta\Sigma(1) + 2n_f C_g(Q^2)\Delta g(1,Q^2)\right], \tag{3.16}$$

where the quark and gluon coefficient functions are given by

$$C_g(Q^2) = -\frac{\alpha_s}{4\pi} C_q(Q^2) = -\frac{\alpha_s}{4\pi} C_{\rm S}(Q^2) \tag{3.17}$$

to all orders in perturbation theory. This is to be contrasted to schemes where $\Delta\Sigma(1,Q^2) = a_0(Q^2)$, so $C_g(Q^2) = 0$.

If the coefficient functions are determined at any finite perturbative order, eq. (3.15) implies that $a_0(Q^2)$ is not simply found, as in sect. 3.1, dividing $\Gamma_1^{\rm sing.}$ by the coefficient function $C_{\rm S}(Q^2)$: indeed, using the expression of C_q and C_g to order α_s^k in eq. (3.16) one gets

$$\tfrac{2}{C_{\rm S}(Q^2)\langle e^2\rangle}\Gamma_1^{\rm sing.} - a_0(Q^2) = -n_f\left(\tfrac{C_{\rm S}^{(k)}}{2\pi}\right)\alpha_s^{k+1}\Delta g(1,Q^2) + O(\alpha_s^{k+2}), \tag{3.18}$$

where $C_{\rm S}^{(k)}$ is the coefficient of α_s^k in the perturbative expansion of $C_{\rm S}(Q^2)$. Thus all determinations of a_0 discussed so far are affected by an error of this size, and could only be improved by extracting $\Delta\Sigma(1)$ and $\Delta g(1,Q^2)$ directly from the data (for example fitting eq. (3.16) to the data) and then computing $a_0(Q^2) = \left(\Delta\Sigma(1) - \frac{\alpha}{2\pi}2n_f\Delta g(1,Q^2)\right)$. A more conservative point of view is that this is an intrinsic and unavoidable ambiguity of the computation. The size of the ambiguity of course depends crucially on the size of the gluon and decreases rapidly as the scale and the perturbative order increase. At order α_s, assuming that the quark respects the Zweig rule and the gluon makes up for the difference, and taking the value of a_0 from table 1, the correction at 3 GeV^2 is around 15% of the value if $a_0 \approx 0.5$, and around 50% if $a_0 \approx 0.15$. Note that this uncertainty is not included in the errors given in table 1.

The scale dependence of a_0 induced by the axial anomaly is determined perturbatively by the anomalous dimension $\gamma_5(\alpha_s)$, which is a universal (i.e. target–independent) property of the axial current. It has been shown recently[43] that this universality actually persists beyond perturbation theory: the singlet axial charge satisfies a Goldberger-Treiman equation which relates it to the decay constant F_Φ and irreducible coupling to the nucleon $g_{\Phi\bar{B}B}$ of a singlet pseudoscalar state Φ (defined in analogy to the π coupling and decay constant in the usual Goldberger–Treiman relation): $a_0(Q^2) = g_{\Phi\bar{B}B}F_\Phi(Q^2)$. This pseudoscalar meson is however not a physical state of the theory due to the presence of the anomaly in this channel, so these couplings could only be measured on the lattice. However, the decay constant can be expressed in terms of the topological susceptibility, which is a property of the QCD vacuum.

Furthermore, all the scale dependence of a_0 comes from this quantity, and is thus contained in a universal, target-independent factor. It follows that if the explanation of the observed Zweig rule violation is in this scale dependence, i.e. in the smallness of the scale-dependent term F_Φ, the effect will be universal, and not specific of the nucleon. It is interesting to compare this possibility with an alternative proposal[44] where the effect is instead related to a small value of the coupling $g_{\Phi\bar{B}B}$ due to instantons. The effect would then be strongly target dependent, and have a different scale dependence. More data could thus shed light on our understanding of the QCD vacuum beyond perturbation theory.

4. The structure function in the (x, Q^2) plane

We have seen in the previous sections that the scale dependence of Γ_1 [eq. (2.3)] driven by the Wilson coefficient functions as well as by the anomalous dimension of the singlet axial current is quite large, that it substantially affects the extraction of current matrix elements from the data, and that it is actually responsible for a large part of the uncertainty on their determination. This suggests that a detailed understanding of the evolution of the full structure function g_1 in the (x, Q^2) plane is necessary in order to accurately assess this uncertainty, and pin it down as much as possible. Indeed, experimental data are taken within a limited range in x ($0.003 \leq x \leq 0.7$ for SMC and $0.03 \leq x \leq 0.75$ for E143) and at different scales in each x bin, with the lowest x points being taken at low Q^2 and conversely ($1.3 \leq Q^2 \leq 48.7$ GeV^2 for SMC and $1.3 \leq Q^2 \leq 9.2$ GeV^2 for E143). The determination of moments of g_1 requires thus both extrapolation in x and evolution to a common scale. The two problems are closely related because $g_1(x, Q^2)$ is determined in terms of polarized parton distributions

$$g_1(x, Q^2) = \tfrac{1}{2}\left[C_{\rm NS} \otimes \Delta q_{\rm NS} + \langle e^2 \rangle \left(C_{\rm S} \otimes \Delta\Sigma + 2n_f C_g \otimes \Delta g\right)\right], \tag{4.1}$$

(where $\otimes$ denotes the usual convolution with respect to x). These evolve perturbatively according to the Altarelli–Parisi equations, so parton distributions at (x_0, Q_0) are causally determined from their values at $x > x_0$, $Q < Q_0$.

The values of the first moments in table 1 are arrived at by assuming the scattering asymmetry A_1 to be scale–independent, i.e. by approximately evolving the data to a common scale on the assumption that the scale dependence of the polarized structure function g_1 is the same as that of the unpolarized one F_2, and extrapolating to small (and large) x by fitting a phenomenological shape to the last few data points. In order to critically examine these assumptions, we must study the small x behavior and scale dependence of g_1 in perturbative QCD.

4.1. The small x behavior of polarized parton distributions

Whereas the extrapolation of the structure function g_1 from the measured range to $x = 1$ is under theoretical control and amounts to a small uncertainty, the extrapolation to small x is more subtle. This is due to the fact that g_1 must vanish identically for kinematic reasons at $x = 1$; moreover, the form of its drop is already constrained by available data and agrees with expectations based on QCD counting rules [45]. On the contrary, the observed small x behavior has thwarted several times theoretical prejudice. In fact, at least in principle, the associated uncertainty is infinite, since the $x \to 0$ limit can never be attained (it corresponds to infinite energy) and, because perturbative evolution proceeds from larger to smaller x, it is a priori impossible to exclude a growth in the unexplored small x region.

The traditional expectation for the small x behavior of structure functions is embodied in Regge theory, which at some low scale should provide the small x behavior of parton distributions that are input to perturbative evolution . Dominance of Regge poles would predict[46)] that both the singlet and nonsinglet contributions to g_1 decrease or are at most constant as $x \to 0$. Notice that in the unpolarized case Regge theory would have the singlet contribution to F_2 behave like a constant (or perhaps grow slightly, for a supercritical Pomeron) and the nonsinglet drop as $\sqrt{x}$; both the nonsinglet[47)] and singlet[48)] predictions are in excellent agreement with available experimental data. Because in LO $F_2 \propto xq(x)$, while $g_1 \propto \Delta q(x)$, the Regge expectations for the singlet can be made consistent with the expectation[45)] (partly based on QCD) that $\frac{\Delta q}{q} \underset{x\to 0}{\sim} x$, if one assumes that both F_2 and g_1 are constant or almost constant at small x at a low scale. The SMC data on g_1^p, however, seem to indicate a rapid growth of g_1 as $x \to 0$ (see fig. 3a). One might have thought this to be due to a stronger increase of the singlet at small x, perhaps due to Regge cuts, but this is belied by the SMC data on g_1^d, which should then look like the proton in the small x region, whereas in actual fact it is negative there (fig. 3b).

Now, even if at some low scale parton distributions behave according to Regge expectations, they will be modified by perturbative evolution. In particular, it has been known for long that LO perturbative evolution eventually leads to a rise of g_1 at small x[49,50)] of the same form as the rise of the unpolarized structure function F_2[51)]; the main difference is that whereas F_2 is positive–definite g_1 is not, so the rise may correspond to g_1 growing either large and positive or large and negative. A closer look reveals both analogies and differences between the polarized and unpolarized cases[50)]. The small x behavior of parton distributions is dominated by the rightmost singularity of perturbative anomalous dimensions in the space of moments. In the polarized case, this singularity is located at $N = 0$ for the nonsinglet as well as for all entries in the matrix of singlet anomalous dimensions, and it is already present at LO. In the unpolarized case it is located at $N = 1$ for the nonsinglet and at $N = 0$ for the singlet; in the latter case the singularity is present at LO in the gluon anomalous dimensions, but only starts at NLO in the quark ones [defining the moment variable N according to eq. (3.9)].

As a consequence, both the singlet and nonsinglet polarized distributions grow according to[50)]

$$\Delta f(x, Q^2) \sim \frac{1}{\sqrt{\sigma}} e^{2\gamma_f \sigma}, \tag{4.2}$$

where $\sigma \equiv \sqrt{\xi\zeta}$, $\rho \equiv \sqrt{\xi/\zeta}$, $\xi \equiv \ln \frac{x_0}{x}$, $\zeta \equiv \ln \frac{\alpha_s(Q_0^2)}{\alpha_s(Q^2)}$, and x_0 is a reference value of x such that the approximate small x form of the anomalous dimensions is applicable for $x \lesssim x_0$ and $Q^2 \gtrsim Q_0^2$. In the nonsinglet case, $\Delta f = \Delta q_{\rm NS}$ and

$$\gamma_{\rm NS}^2 = \frac{8}{33 - 2n_f}. \tag{4.3}$$

In the singlet case the quantities that display this growth are the linear combinations of quark and gluon $v^{\pm} = \Delta\Sigma + C^{\pm}\Delta g$, with

$$C^{\pm} = 2\left(1 \pm \sqrt{1 - \frac{3n_f}{32}}\right); \qquad \gamma_{\pm}^2 = \gamma_{\rm NS}^2 \left(5 \pm 4\sqrt{1 - \frac{3n_f}{32}}\right). \tag{4.4}$$

Notice that this is a scheme independent result. The behavior of the unpolarized nonsinglet is the same as that of the polarized singlet (but with different γ), while the behavior of the unpolarized singlet differs, first because the quantity which grows according to eq. (4.2) is $xg(x,Q^2)$ rather than $\Delta g(x,Q^2)$, and furthermore because the growth is driven by the gluon, i.e. the unpolarized quark is proportional to the unpolarized gluon (and of order α_s with respect to it).

This implies that eventually the singlet distributions indeed dominate g_1 at small x. In both singlet eigenvectors the quark and gluon have opposite sign, and asymptotically $\Delta\Sigma > 0$ if $\Delta\Sigma(x_0,Q_0) > -C^+\Delta g(x_0,Q_0)$: hence, for most plausible starting quark and gluon $\Delta\Sigma$ grows large and negative and Δg large and positive. As a consequence, asymptotically g_1 at small x must be large and negative. The nonsinglet distribution however also grows asymptotically; the growth has the same form as for the singlet distributions, only with a smaller value of γ_f, its sign is the same as that of the starting distribution. Therefore, the value of x where the singlet actually does dominate could in practice be extremely small. This is to be contrasted to the unpolarized case, where the growth of eq. (4.2) is down by a power of x in the nonsinglet compared to the singlet, so that the singlet already dominates around $x \sim 10^{-3}$.

Of course, the perturbatively generated growth of eq. (4.2) is only visible provided it dominates over a possible growth of the starting parton distributions: even though Regge theory does not favor such a growth, it cannot be firmly ruled out. For instance, if parton distributions rise as $\Delta f(x,Q_0) \sim x^{-\lambda}$, then the evolved distributions will behave according to eq. (4.2) only for $\rho \lesssim \frac{\gamma_f}{\lambda}$, but for smaller values of x they will behave as[50]

$$\Delta f \sim x^{-\lambda} \left(\frac{\alpha_s(Q_0^2)}{\alpha_s(Q^2)} \right)^{\frac{\gamma_f^2}{\lambda}}, \tag{4.5}$$

i.e. they reproduce the boundary condition up to an x–independent correction; in the singlet case if the starting quark and gluon have different small x–behavior the most singular one dominates asymptotically. In such case, the onset of the asymptotic small x behavior of eq. (4.5) will be yet slower[52]: firstly, because the small x solution to the evolution equations is then dominated by the singularity of the boundary condition rather than the singularity at $N = 0$ in the anomalous dimension, therefore, approximating the anomalous dimension with this singularity is less accurate; also, the two singlet eigenvectors only mix due to the x-independent correction in eq. (4.5), thus the dominance of the most singular behavior only sets in rather slowly.

It is interesting to see how these LO predictions are modified at higher orders. At NLO all polarized anomalous dimensions have a $\frac{1}{N^3}$ singularity. This leads to a correction to the LO of the form[5]

$$\Delta f^{\rm NLO}(x,Q^2) = \left[1 + \epsilon_f \left(\tfrac{\rho}{\gamma_f}\right)^3 \left(\alpha_s(Q_0^2) - \alpha_s(Q^2)\right)\right] \Delta f^{\rm LO}(x,Q^2), \tag{4.6}$$

where $\Delta f^{\rm LO}(x,Q^2)$ is given by eq. (4.2), and the coefficients ϵ are explicitly

$$\epsilon_{\rm NS} = \tfrac{8}{3\pi\beta_0}; \qquad \epsilon_\pm = \tfrac{112}{3\pi\beta_0}\left[(1 - \tfrac{n_f}{14}) \pm \tfrac{13}{14}(1 - \tfrac{11 n_f}{104}) \Big/ \sqrt{1 - \tfrac{3 n_f}{32}}\right], \tag{4.7}$$

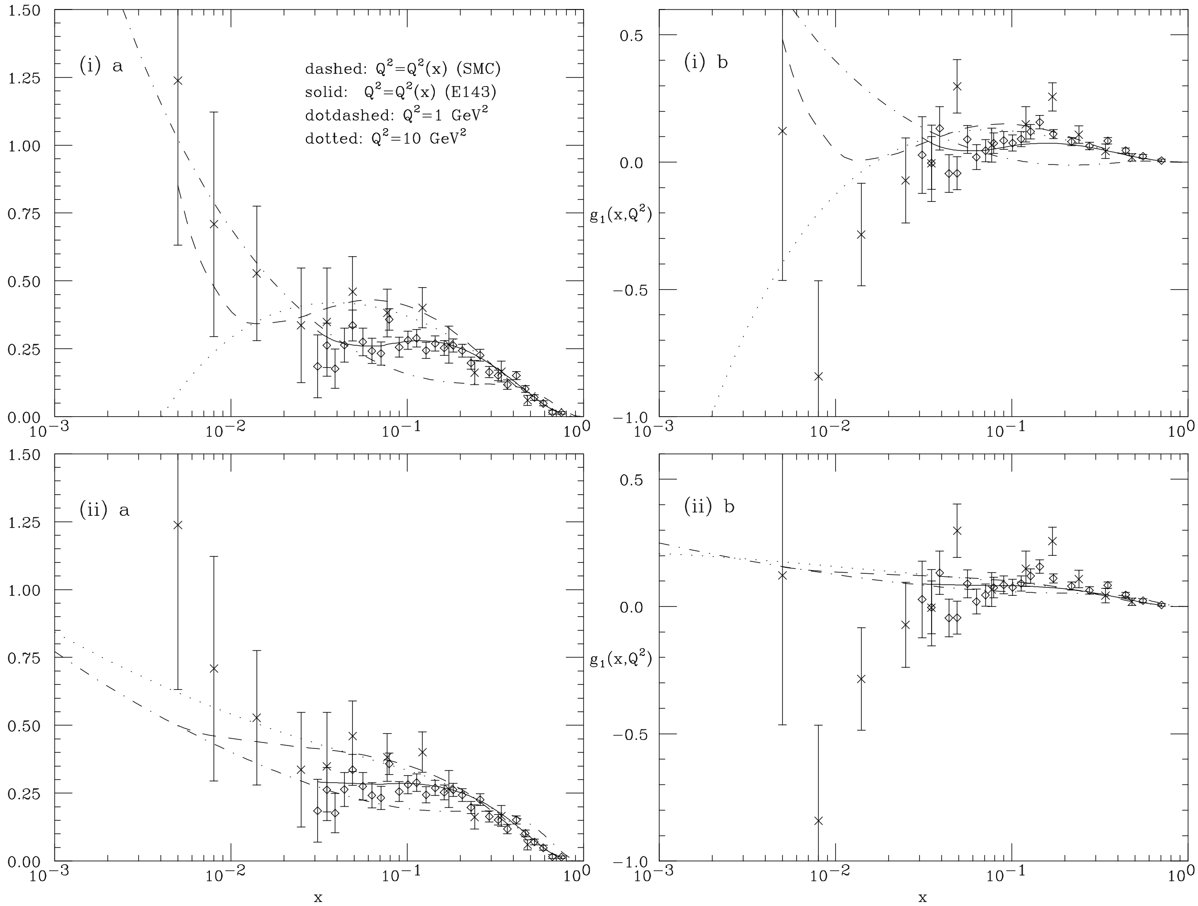

Figure 3: Plot of $g_1(x)$ for (a) proton and (b) deuteron. The crosses are SMC data and the diamonds E143 data. The curves correspond to a fit[50] to the proton data only with valence-like nonsinglet and (i) "maximal" gluon and steep input quark singlet or (ii) "minimal" gluon and flat input quark singlet.

corresponding to a further rise proportional to $\xi^{3/2}$ of the parton distributions at small x, with coefficient $\epsilon_{\rm NS}/\gamma^3_{\rm NS} \approx \epsilon_+/\gamma^3_+ \approx \frac{1}{2}$. Notice that the form of the small x eigenvectors is unaltered, and that these results are scheme–independent.

At NLO the Mellin transforms of the quark singlet and nonsinglet and gluon coefficient functions also acquire a singularity at $N = 0$, which is however only of order $\frac{1}{N^2}$ and therefore does not contribute to the asymptotic small x behavior eq. (4.2). This singularity corresponds to a $\ln\frac{1}{x}$ rise (with positive sign) of the functions $C(x, Q^2)$ in eq. (4.1). At small x, this leads to a rise of g_1 even when the various parton distributions are flat. Even though this rise is weaker than that induced by perturbative evolution, it could dominate at low scale where evolution effects are small; its precise form is scheme–dependent, however.

At yet higher orders, the expected generic behavior of anomalous dimensions is $\gamma \underset{N\to 0}{\sim} (\alpha_s^n/N^{2n-1})$. The small N behavior of the NLO coefficient functions suggests that these might instead behave as $C \underset{N\to 0}{\sim} \left(\alpha_s/N^2\right)^{n-1}$, which would guarantee scheme independence of the coefficients of the singularities in the anomalous dimensions at least in the nonsinglet case. Indeed a resummation of these singularities in the nonsinglet anomalous dimensions has been proposed long ago[53]. In the absence of appropriate factorization theorems, it is unclear whether this resummation, which reproduces [54] the known NLO leading singularity, correctly reproduces the leading singularities of the nonsinglet anomalous dimensions to all orders. If it does, it implies that the polarized and unpolarized nonsinglet quark distributions will display the same small x–behavior, in contradiction to Regge theory. More generally, one would expect[55] the summation of logarithmic effects in $\ln \frac{1}{x}$ induced by these singularities to lead to a rise of parton distributions, which should be asymptotically power-like; this rise should set in significantly faster than in the unpolarized case, due to the stronger nature of the singularity, and thus, contrary to the unpolarized case[55], it could set in before perturbation theory breaks down.

The combination of all these effects makes it rather hard to infer the asymptotic small x behavior of polarized parton distributions from present-day data: a HERA-like kinematic coverage would be necessary to disentangle the various effects. The only firm prediction at this stage is that parton distributions at small x will certainly grow (positive or negative) at least as fast as eq. (4.2), including the nonsinglet distributions, and that therefore a simple phenomenological extrapolation based on Regge expectations is not adequate at least in principle. More detailed statements can be made by studying the structure of the data in the (x, Q^2) plane.

4.2. The evolution of $g_1(x, Q^2)$

The only way of arriving at a precise and reliable determination of the moments of g_1 is to describe its evolution in the (x, Q^2) plane in terms of the evolution of polarized parton distribution. The fact that these evolution effects are large not only implies that they must be included in the data analysis, but also that they might provide useful information on the polarized parton content of the nucleon, and specifically on its gluon content. In fact, one would expect g_1 to be a very sensitive probe of Δg — much more than the unpolarized structure function F_2 is a probe of g — since, even though the coupling of the gluon to g_1 is formally NLO, thanks to the anomaly it does not decouple asymptotically[42]. In comparison to the unpolarized case, however, the analysis is complicated by the fact that the singlet and nonsinglet display a similar perturbative behavior in the small x region: this implies that only combining proton and neutron (or deuteron) data is it possible to disentangle the contributions of the various parton distributions to g_1, at least using data with the current relatively restricted kinematic coverage.

Thanks to the recent determination[6] of the full matrix of two-loop anomalous dimensions, it is possible to study the evolution of parton distributions at NLO. The computations of ref. [6] are performed in the $\overline{\text{MS}}$ scheme, where the gluon decouples from the first moment of g_1 and thus is not properly factorized, as we discussed in sect. 3.2. However, since the quark and gluon coefficient functions are known[39,42] in schemes where the gluon does contribute to g_1, the corresponding anomalous dimensions may be determined[5] from those of ref. [6]. A determination of parton distributions can then be made by fitting to the data, provided one makes some simplifying assumptions on the form of the starting parton distributions.* Specifically, a simple polynomial parametrization of parton distributions may be adopted, which entails the assumption that the asymptotic small x behavior of parton distributions starts setting in at the values of x explored at present:

$$\Delta f(x, Q_0^2) = N\left(\alpha_f, \beta_f, a_f\right) \eta_f x^{\alpha_f} (1-x)^{\beta_f} (1 + a_f x) \tag{4.8}$$

* For a review of LO polarized parametrizations see ref. [56]. A NLO fit has been presented in ref. [57], but in a scheme where the gluon does not couple to g_1. Approximate NLO fits (prior to the full determination of NLO anomalous dimensions) in properly factorized schemes are in refs. [58,50]. A full NLO computation in a variety of physical schemes is in ref. [5].

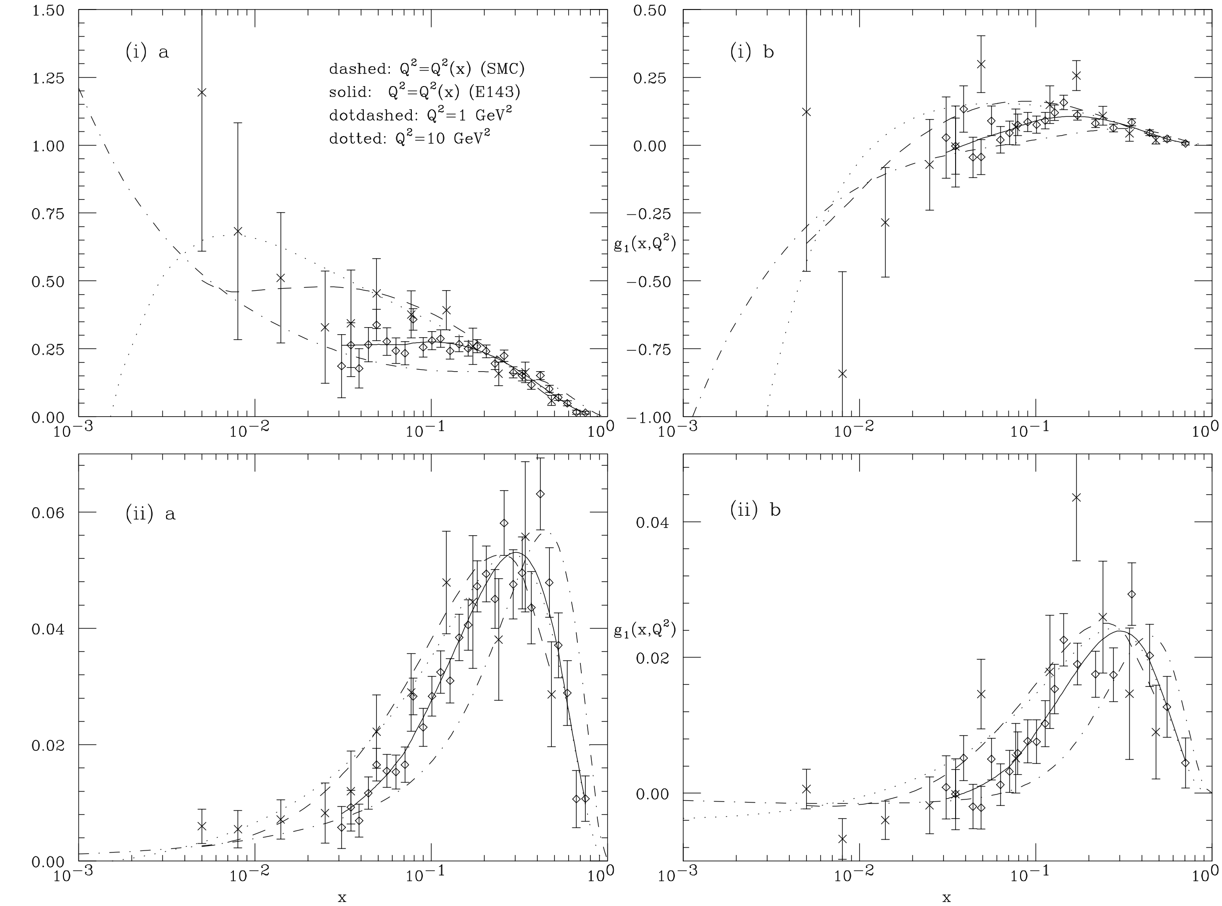

Figure 4: Plot of (i) $g_1(x)$ and (ii) $xg_1(x)$ for (a) proton and (b) deuteron. The crosses are SMC data and the diamonds E143 data. The curves correspond to a NLO fit[5] to all data.

(where $N(\alpha, \beta, a)$ is a normalization factor, fixed e.g. as $N(\alpha, \beta, a) \int_0^1 dx\, x^\alpha (1-x)^\beta (1+ax) = 1$).

Two fits[50] to the proton data only of parton distributions of this form are displayed in fig. 3, together with the corresponding predicted deuteron structure function. The nonsinglet is assumed to be valence-like ($\alpha_{\rm NS} = +0.2$); the gluon is assumed to be flat ($\alpha_g = 0$) and the singlet quark is either steep ($\alpha_q = -0.5$, fig. 3i) or flat ($\alpha_q = 0$, fig. 3ii). The starting gluon is fixed either by requiring $\Delta\Sigma(1)$ to satisfy the Zweig rule, so that the fit then forces a large gluon component to compensate ("maximal gluon", fig. 3i), or by symply taking the gluon distribution to vanish at the starting scale ("minimal gluon", fig 3ii). The quality of these two fits does not differ in a statistically significant way: while a larger gluon corresponds to stronger evolution effects, the proton data alone do not allow us to fix its size, or the small x behavior of the various parton distributions. Notice, however, that in order to obtain a satisfactory description of the data it is necessary to introduce a gluon coupling to g_1. This is due to the fact that the SMC and E143 experiments measure g_1 at different values of Q^2 for equal values of x. The two data sets can then be made consistent with each other only if there is a sufficient amount of perturbative evolution, and this can hardly be obtained in a LO computation[59], or if schemes where the gluon decouples from g_1 are adopted.

Inspection of the deuteron predictions (fig. 3b), however, clearly shows that more definite conclusions may be reached by combining the two sets of data: for instance the "minimal gluon" case seems to be excluded. This is a consequence of the fact that the deuteron is a more sensitive probe of the singlet and thus of the gluon, as explained in sect. 3.1. The results

Q^2	$\Delta\Sigma(N=1)$	$\Delta g(N=1)$	Γ_1^p	Γ_1^d	a_0
3	0.5 ± 0.1	2.3 ± 1.2	$0.118\,^{+0.016}_{-0.014}$	$0.022\,^{+0.015}_{-0.013}$	$0.15\,^{+0.17}_{-0.12}$
5	0.5 ± 0.1	2.6 ± 1.4	$0.120\,^{+0.016}_{-0.014}$	$0.023\,^{+0.015}_{-0.013}$	$0.14\,^{+0.16}_{-0.11}$
10	0.5 ± 0.1	3.0 ± 1.6	$0.122\,^{+0.018}_{-0.014}$	$0.023\,^{+0.016}_{-0.013}$	$0.14\,^{+0.16}_{-0.11}$

Table 2: Next-to-leading order determination[5)] of polarized first moments.

of a full NLO determination[5)] of g_1 (displayed in fig. 4) indeed confirm that all the parameters in eq. (4.8) may then be determined with reasonable accuracy. In particular, it is possible to disentangle the singlet and nonsinglet contributions, and thus determine independently the small x behavior of the various parton distributions. This turns out to be steep for the nonsinglet ($\Delta q_{\rm NS} \underset{x\to 0}{\sim} x^{\alpha_{\rm NS}}$ with $\alpha_{\rm NS} = -0.7 \pm 0.2$), and valence-like for the singlet (with large error). The result is quite suggestive in view of the discussion in the previous section: it seems consistent with the expectation, based on the resummations of ref. [53)], that the polarized and unpolarized nonsinglets should behave in the same way, despite the prediction of Regge theory. However, the asymptotic small x behavior is only starting to set in at the smallest values of x covered by present–day data, which, in this respect, are to be compared with the NMC unpolarized data [47)]: a kinematic coverage comparable with that available at HERA for determining [48)] the small x behavior of F_2 would be required to determine precisely the small x behaviour of polarized parton distributions.

Present–day data are however sufficient to determine the singlet first moments $\Delta\Sigma(1,Q^2)$ and $\Delta g(1,Q^2)$ with reasonable accuracy (see table 2), assuming isospin [and SU(3)] to be exact. Even though the first moment of Γ_1 at each Q^2 determines only the linear combination eq. (3.15) of $\Delta\Sigma(1,Q^2)$ and $\Delta g(1,Q^2)$, their different scale dependence is sufficient to determine them independently: the large evolution effect already observed requires a large gluon component. Equation (3.15) then forces a large value of $\Delta\Sigma(1,Q^2)$, consistent with the Zweig rule. Notice that the direct gluon contribution to g_1 is essential in order to get good agreement with the data. Thus current data support the proposal of ref. [42)] that the smallness of a_0 is due to a large gluon contribution to it.

The first moment of the quark and gluon distribution then determine both the axial charge and the first moment of Γ_1. These determinations turn out to be significantly smaller, and with significantly larger error, than the corresponding determinations from the experimental collaborations (table 1): as discussed in sect. 3.1, a difference of the order of 10% in the value of Γ_1^p results in a variation of a factor of about 2 in the value of a_0. We may actually understand in detail the origin of this discrepancy[50)], which is due to the approximate treatment of both scale dependence and extrapolation at small x used in refs.[1−4)]. Firstly, the data points which provide the bulk of Γ_1 are taken at a scale Q^2 larger than the nominal average scale of either experiment: for example in refs.[1,3)] $Q^2_{\rm exp}(x)$ is always larger than 20 GeV^2 for x above 0.1, while $\langle Q^2\rangle = 10$ GeV^2 for this experiment. Since in this region g_1 increases as the scale increases the substantial underestimate of evolution effects due to their approximate treatment in refs.[1−4)] leads to an overestimate of Γ_1. Furthermore, the flat small x extrapolation based on Regge theory overestimates the corresponding contribution, because it is obtained by extrapolating the smallest–x data points, which are taken at very low Q^2 (~ 1 GeV^2) but assumed to apply at larger values of Q^2 where perturbative evolution, as explained in sect. 4.1, leads rapidly to negative values of g_1 and thus to a small (or even negative) contribution to Γ_1 from the small x region.

The presence of large corrections due to perturbative evolution then inevitably implies sizable uncertainties related to insufficiently precise knowledge of the polarized gluon distribution, which drives this evolution at least at the level of the first moment, and to the unknown higher order corrections (scheme ambiguities). These are indeed the dominant sources of the error on Γ_1 and thus a_0 in table 2. It is interesting to notice that even though the values of table 2 are obtained by assuming isospin [i.e. the Bjorken sum rule eq. (2.8)] to be exact, the values of a_0, $\Delta\Sigma(1)$ and $\Delta g(1, Q^2)$ are essentially insensitive to SU(2) [or SU(3)] violations. This suggests that if the isotriplet first moment Γ_1 were also determined by a global NLO analysis, rather than being fixed using the Bjorken sum rule, its value might not differ much from that found by combining values of Γ_1 from table 1. Such an analysis would however be required in order to test the Bjorken sum rule in a reliable way.

5. Outlook

Polarized structure functions determined in deep-inelastic scattering are clearly the preferred source of information on the polarized structure of the nucleon at high energy, since the availability of renormalization group methods and factorization theorems that characterizes fully inclusive processes allows a determination of physical observables which is free from the ambiguities due to our ignorance of the low–energy dynamics. In particular, due to the axial anomaly, polarized deep-inelastic scattering provides an ideal probe of the polarized gluon distribution and allows us to get a handle on quantities that have direct relevance for our understanding of the chiral structure of the QCD vacuum.

The new generation of experiments[1−4] has allowed us to confirm these expectations, by providing us with surprisingly accurate information on polarized parton distributions. The analysis of these data, however, requires sophisticated theoretical and phenomenological tools. A future generation of experiments with a wider kinematic coverage in x and Q^2, such as those that would be possible with a polarized proton beam at HERA[60], could lead to a full understanding of the small x behavior and scale dependence of polarized parton distributions, thereby testing our understanding of QCD in perturbation theory and beyond.

Acknowledgements I thank for several discussions G. Altarelli, G. Ridolfi and R. Ball, whom I also thank for a critical reading of the manuscript.

[1] SMC Collaboration, D. Adams et al., *Phys. Lett.* **B329** (1994) 399.
[2] E143 Collaboration, K. Abe et al., *Phys. Rev. Lett.* **74** (1995) 346.
[3] SMC Collaboration, D. Adams et al., *Phys. Lett.* **B357** (1995) 248.
[4] E143 Collaboration, K. Abe et al., *Phys. Rev. Lett.* **75** (1995) 25.
[5] R. D. Ball, S. Forte and G. Ridolfi, preprint CERN-TH/95-266, `hep-ph/9510449`.
[6] R. Mertig and W. L. van Neerven, Leiden preprint INLO-PUB-6/95; Erratum, private communication (November 1995).
[7] For a general introduction to the phenomenology of g_1, see for instance S. Forte, `hep-ph/9409416`, in "Radiative Corrections: Status and Outlook", B. F. L. Ward, ed. (World Scientific, Singapore, 1995) and references therein. We will follow the notation and conventions of this paper.
[8] SMC Collaboration, D. Adams et al., *Phys.* Lett. **B336** (1994) 125; E143 Collaboration, K. Abe et al., preprint SLAC-PUB-95-6982.
[9] See e.g. R. L. Jaffe, *Comm. Nucl. Part. Phys.* **14** (1990) 239; G. Altarelli, B. Lampe, P. Nason and G. Ridolfi, *Phys. Lett.* **B334** (1994) 187 and references therein.
[10] E142 Collaboration, P. L. Anthony et al., *Phys. Rev. Lett.* **71** (1993) 959.
[11] Y. Roblin, Ph.D. Thesis, University of Clermont–Ferrand (1995).
[12] J. D. Bjorken, *Phys. Rev.* **148** (1966) 1467
[13] L. L. Frankfurt and M. Strikman, *Nucl. Phys.* **A405** (1983) 557.

[14] S. Forte, *Phys. Rev.* **D47** (1993) 1842.
[15] S. Forte, *Phys. Lett.* **B309** (1993) 174.
[16] J. Ellis et al., preprint CERN-TH/95-155, `hep-ph/9509312`.
[17] See e.g. V. Braun, `hep-ph/9505317` and references therein for a simple phenomenological discussion.
[18] F. David, *Nucl. Phys.* **B263** (1986) 637; A. H. Mueller, *Nucl. Phys.* **B250** (1985) 327.
[19] X. Ji, *Nucl. Phys.* **B448** (1995) 51.
[20] For a comprehensive summary of the determinations of higher twist corrections to the Bjorken sum rule see L. Mankiewicz, E. Stein and A. Schäfer, `hep-ph/9510418`.
[21] J. Ellis and M. Karliner, *Phys. Lett.* **B341** (1995) 397.
[22] G. Altarelli, P. Nason and G. Ridolfi, *Phys. Lett.* **B320** (1994) 152.
[23] W. Melnitchouk, G. Piller and A. W. Thomas, *Phys. Lett.* **B346** (1995) 165.
[24] B. Badełek and J. Kwieciǹski, *Phys. Rev.* **D80** (1994) 4.
[25] N. N. Nikolaev and V. I. Zakharov, *Phys. Lett.* **B55** (1975) 397; S. J. Brodsky and H. J. Lu, *Phys. Rev. Lett.* **64** (1990) 1342.
[26] O. Nachtmann, *Nucl. Phys.* **B63** (1973) 237; S. Wandzura, *Nucl. Phys.* **B122** (1977) 412; S. Matsuda and T. Uematsu, *Nucl. Phys.* **B168** (1980) 181.
[27] H. Kawamura and T. Uematsu, *Phys. Lett.* **B343** (1995) 346.
[28] S. A. Larin, *Phys. Lett.* **B334** (1994) 192.
[29] F. E. Close and R. G. Roberts, *Phys. Lett.* **B336** (1994) 257.
[30] B. Ehrnsperger and A. Schäfer, *Phys. Lett.* **B348** (1995) 619.
[31] H. J. Lipkin, *Phys. Lett.* **B337** (1994) 157; J. Lichtenstadt and H. J. Lipkin, *Phys. Lett.* **B353** (1995) 119.
[32] G. Altarelli and G. Ridolfi, *Nucl. Phys. B (Proc. Suppl.)* **39B** (1995) 106 and private communication.
[33] R. D. Ball. *Phys. Lett.* **B266** (1991) 473.
[34] G. Altarelli et al., *Nucl. Phys.* **B69** (1974) 531.
[35] G. Altarelli, S. Petrarca and F. Rapuano, preprint CERN-TH/95-273, `hep-ph/9510346`.
[36] See R. Jackiw, in S. Treiman, R. Jackiw, B. Zumino and E. Witten, *"Current Algebra and Anomalies"* (World Scientific, Singapore, 1985) for a review.
[37] S. Adler and W. Bardeen, *Phys. Rev.* **182** (1969) 1517.
[38] M. Bos, *Nucl. Phys.* **B404** (1993) 215.
[39] J. Kodaira et al., *Phys. Rev.* **D20** (1979) 627; *Nucl. Phys.* **B159** (1979) 99; J. Kodaira, *Nucl. Phys.* **B165** (1979) 129.
[40] R. D. Carlitz, J. C. Collins and A. H. Mueller, *Phys. Lett.* **B214** (1988) 229; G. Altarelli and B. Lampe, *Z. Phys.* **C47** (1990) 315; W. Vogelsang, *Z. Phys.* **C50** (1991) 275.
[41] S. Forte, *Nucl. Phys.* **B331** (1990) 1.
[42] G. Altarelli and G. G. Ross, *Phys. Lett.* **B212** (1988) 391.
[43] G.M. Shore and G. Veneziano, *Phys. Lett.* **B244** (1990) 75; *Nucl. Phys.* **B381** (1992) 23; S. Narison, G.M. Shore and G. Veneziano, *Nucl. Phys.* **B433** (1995) 209.
[44] S. Forte and E. V. Shuryak, *Nucl. Phys.* **B357** (1991) 153.
[45] See S. J. Brodsky, M. Burkardt and I. Schmidt, *Nucl. Phys.* **B441** (1995) 197 and references therein.
[46] R. L. Heimann, *Nucl. Phys.* **B64** (1973) 429.
[47] NMC Collaboration, M. Arneodo et al., *Phys. Rev.* **D50** (1994) 1.
[48] R. D. Ball and S. Forte, *Phys. Lett.* **B335** (1994) 77.
[49] M. A. Ahmed and G. G. Ross, *Phys. Lett.* **B56** (1975) 385; M. B. Einhorn and J. Soffer, *Nucl. Phys.* **B74** (1986) 714; A. Berera, *Phys. Lett.* **B293** (1992) 445.
[50] R. D. Ball, S. Forte and G. Ridolfi, *Nucl. Phys.* **B 444** (1995) 287
[51] A. De Rújula et al., *Phys. Rev.* **D10** (1974) 1649.
[52] T. Gehrmann and W. J. Stirling, Durham preprint DTP/95/62, `hep-ph/9507332`.
[53] R. Kirschner and L. Lipatov, *Nucl. Phys.* **B213** (1983) 122.
[54] J. Blümlein and A. Vogt, preprint DESY 95-175, `hep-ph/9510410`.
[55] R. D. Ball and S. Forte, *Phys. Lett.* **B351** (1995) 313.
[56] T. Gehrmann and W. J. Stirling, Durham preprint DTP/95/78, `hep-ph/9510243`.
[57] M. Glück et al., Dortmund preprint DO-TH 95-13, `hep-ph/9508347`.
[58] T. Gehrmann and W. J. Stirling *Z. Phys.* **C65** (1995) 461.
[59] G. Altarelli, P. Nason and G. Ridolfi, *Phys. Lett.* **B320** (1994) 152
[60] V. Hughes, talk at the Workshop on Deep-Inelastic Scattering and QCD, Paris, May 1995.