The properties of the “transverse” spin-dependent structure function, $g_2(x, Q^2)$, measurable in deep inelastic scattering of polarized leptons from polarized nucleons are reviewed.

**Key Words:** spin-dependent structure function, deep inelastic scattering, operator product expansion

### I. INTRODUCTION

The “transverse” spin-dependent structure function, $g_2(x, Q^2)$, may soon be measured in experiments being prepared at CERN\(^1\) and HERA\(^2\). The existing literature on $g_2(x, Q^2)$ is meagre, and what exists is often contradictory and confusing to both theorists and experimentalists.\(^3\) One can find reputable “parton model” or “operator product expansion” predictions which state $g_2(x) = 0$, or $g_1(x) + g_2(x) = 0$, or $g_2(x) = -g_1(x) + \int_x^1 (dy/y)g_1(x)$, or $g_2(x) \propto m_{\text{quark}}$, or $g_2(x) \propto \langle k_{\perp} \rangle_{\text{quark}}$. Not only are these inconsistent with one another, none is correct in QCD. Much of the recent work has been confined to (or based on) the Russian literature.\(^4\)\(^5\)\(^6\)

The purpose of this paper is to gather together what I believe to be reliably known (and not known) about $g_2(x, Q^2)$ and to call attention to its unique sensitivity to higher twist, i.e., interaction-dependent effects in QCD. This paper is based on a review of
traditional sources,\textsuperscript{7} on unpublished work,\textsuperscript{8} and on an important paper by Shuryak and Vainshtein.\textsuperscript{4} The bag model calculation outlined in the Introduction is new.\textsuperscript{9}

In Section II I summarize the kinematics of spin-dependent terms in electron (or muon) scattering. It is surprisingly hard to find in the literature a simple, correct expression for the relevant cross section in terms of variables defined and easily measured in the laboratory. This is corrected, I hope, by Eq. (2.8). Also in this section I write down and discuss dispersion relations involving the two spin-dependent structure functions $g_1(x, Q^2)$ and $g_2(x, Q^2)$. These are uncontroversial but necessary for issues discussed subsequently.

Section III is concerned with the Burkhardt–Cottingham sum rule.\textsuperscript{10} $\int_0^1 dx \, g_2(x, Q^2) = 0$. Despite claims to the contrary,\textsuperscript{3,5,6} I do not believe this to be a current algebra, short distance or parton model sum rule. Instead, it is a super-convergence relation based on Regge asymptotics. It could be spoiled by Regge cuts, in which case the integral fails to converge, or by a "non-polynomial residue $J = 0$ fixed pole," in which case the integral converges but not to zero. There are arguments against such fixed poles\textsuperscript{11} but they need to be re-examined in the context of QCD.

In Section IV I summarize what is known about $g_2(x, Q^2)$ from the operator product expansion in QCD. A lot is known, but apparently not widely known. The initial work on the OPE analysis of $g_2$ can be found in Ref. 7. This section is largely a recapitulation of the work of Shuryak and Vainshtein\textsuperscript{4} based on the relation between higher-twist and interactions in QCD suggested by Politzer\textsuperscript{12} and subsequently developed by Shuryak and Vainshtein\textsuperscript{4,13} and by Soldate and myself.\textsuperscript{14} Ignoring quark mass effects ($O(m_q/\Lambda_{QCD})$), $g_2(x, Q^2)$ can be thought of as the sum of two terms

$$g_2(x, Q^2) = g_2^{ww}(x, Q^2) + \bar{g}_2(x, Q^2). \quad (1.1)$$

The first, $g_2^{ww}(x, Q^2)$, is a twist-2 contribution determined entirely by $g_1(x, Q^2),$

$$g_2^{ww}(x, Q^2) = -g_1(x, Q^2) + \int_x^1 \frac{dy}{y} g_1(y, Q^2). \quad (1.2)$$
This contribution results from the infelicitous but by now conventional definition of \( g_2(x, Q^2) \). The idea of separating it out was first proposed by Wandzura and Wilczek\(^1\) (who, incorrectly it seems, argued \( \bar{g}_2 = 0 \)). The second term, \( \bar{g}_2(x, Q^2) \), comes from twist-3 operators which are explicitly interaction dependent\(^5\): they are proportional to the QCD coupling \( g \) and the gluon field strength \( G_{\mu\nu}^a \); they measure a quark–gluon correlation function in the target nucleon.

This is the only case (at least in exclusive lepto-production) in which a higher twist effect can be observed without first precisely measuring and subtracting away a leading, twist-2 effect. This makes the experimental determination of \( g_2(x, Q^2) \) very interesting.

Finally, I wish to outline a way to obtain at least a crude estimate of the \( x \)-dependence of \( \bar{g}_2(x, Q^2) \). The idea is that the most important gluons are those which are responsible for the confinement (and chiral symmetry breaking) of quarks in the nucleon. These gluons can be (fairly successfully) modeled by a bag boundary condition. This line of thought suggests repeating the analysis of Section IV, but with the QCD equation of motion, \( D\psi = 0 \), replaced by the bag equation of motion, \( i\partial\psi = \delta,\psi \), where \( \delta \) is a \( \delta \)-function on the bag’s surface. \( g_1(x, Q^2) \) can also be calculated in the bag model.\(^16\) By calculating \( g_1 \) and \( \bar{g}_2 \) in the same model we can explore the relation between the spreading of \( g_1 \) about \( x = 1/3 \) and the development of a non-zero \( \bar{g}_2 \): both are due to “Fermi motion” resulting from the confinement of quarks to a bag. The results of this calculation will be reported elsewhere.\(^9\)

II. KINEMATICS AND DISPERSION RELATIONS

A. Kinematics

The spin-dependent structure functions \( g_1(x, Q^2) \) and \( g_2(x, Q^2) \) are proportional to the imaginary part of forward virtual Compton scattering amplitudes. Consider the amplitude for forward scattering of a photon of four momentum \( q^\mu \) \( (q^2 = -Q^2 \leq 0) \) from

\*In Ref. 16, Hughes calculates \( g_1(x) \) and \( g_2(x) \) in the bag model. Unfortunately, his calculation of \( g_2(x) \) is not correct. It is based on an incorrect separation of the operators responsible for \( g_2 \). For a discussion of this problem, see Ref. 9.
a polarized (spin-1/2) target of four momentum $P^\mu$ ($P^2 = M^2$),
spin $S^\mu$ ($S^2 = -M^2$, $S \cdot P = 0$):

$$T_{\mu\nu}(q, P, S) = i \int d^4x e^{iq \cdot x} \langle PS \mid T(J_{\mu}(x)J_{\nu}(0)) \mid PS \rangle$$

$$= -g_{\mu\nu}T_1(q^2, \nu) + \frac{P_{\mu}P_{\nu}}{M^2} T_2(q^2, \nu)$$

$$+ i \frac{\epsilon_{\mu\nu\lambda\sigma}q^\lambda S^\sigma}{M^2} A_1(q^2, \nu)$$

$$+ \frac{i\epsilon_{\mu\nu\lambda\sigma}q^\lambda}{M^4} (\nu s^\sigma - q \cdot s P^\sigma) A_2(q^2, \nu)$$

$$+ q^\mu \text{ or } q^\nu \text{ terms}$$

where $\nu = P \cdot q$ and $\epsilon_{0123} = +1$. The imaginary part of $T_{\mu\nu}$ has
a similar decomposition defining $g_1$ and $g_2$,

$$W_{\mu\nu} = \frac{1}{2\pi} \text{Im} T_{\mu\nu}$$

$$= \frac{1}{4\pi} \int d^4x e^{iq \cdot x} \langle PS \mid [J_{\mu}(x), J_{\nu}(0)] \mid PS \rangle$$

$$= -g_{\mu\nu}F_1(x, Q^2) + \frac{P_{\mu}P_{\nu}}{\nu} F_2(x, Q^2)$$

$$+ q^\mu \text{ or } q^\nu \text{ terms}$$

$$+ \frac{i\epsilon_{\mu\nu\lambda\sigma}q^\lambda S^\sigma}{\nu} g_1(x, Q^2)$$

$$+ \frac{i\epsilon_{\mu\nu\lambda\sigma}q^\lambda (\nu s^\sigma - q \cdot s P^\sigma)}{\nu^2} g_2(x, Q^2)$$

242
where

\[ g_1(x, Q^2) = \frac{\nu}{2\pi M^2} \text{Im} A_1(q^2, \nu), \]  

\[ g_2(x, Q^2) = \frac{\nu^2}{2\pi M^4} \text{Im} A_2(q^2, \nu). \]  

For convenience we also define

\[ \alpha_1(x, Q^2) = \frac{\nu}{M^2} A_1(q^2, \nu), \]  

\[ \alpha_2(x, Q^2) = \frac{\nu^2}{M^4} A_2(q^2, \nu). \]

The cross section for inelastic lepton scattering is obtained by contracting \( W_{\mu\nu} \) with a similar tensor for the lepton currents (calculated to lowest non-trivial order in \( e \)),

\[ l_{\mu\nu} = \tilde{l}_{\mu\nu} \pm \Delta l_{\mu\nu}, \]  

where \( \tilde{l}_{\mu\nu} \) is symmetric and lepton spin-independent and \( \Delta l_{\mu\nu} \) is antisymmetric and lepton spin-dependent,

\[ \tilde{l}_{\mu\nu} = 4k_{\mu}k_{\nu} - 2k_{\mu}q_{\nu} - 2k_{\nu}q_{\mu} + 2k \cdot q g_{\mu\nu}, \]  

\[ \Delta l_{\mu\nu} = 2i\epsilon_{\mu\nu\alpha\beta}k^{\alpha}q^{\beta}. \]

Here \( k^{\alpha} (k'^{\alpha}) \) is the initial (final) lepton four momentum. The + or - sign in Eq. (2.5) refers to right- or left-handed initial leptons, respectively. The definitions of kinematic variables are given in Fig. 1.
FIGURE 1 Kinematic variables in inelastic lepton scattering (a) momentum space, (b) coordinate space.

The cross section obtained by contracting \( l_{\mu\nu} \) with \( W_{\mu\nu} \) contains both lepton spin-dependent and -independent parts. We define the cross section for right- or left-handed incident leptons by \( \sigma_{(R,L)} = \bar{\sigma} \pm \Delta \sigma \), with

\[
\frac{d\Delta \sigma(\alpha)}{dx \, dy \, d\phi} = \frac{e^4}{4\pi^2 Q^2} \left\{ \cos \alpha \left[ \left[ 1 - \frac{y}{2} - \frac{y^2}{4} (\kappa - 1) \right] g_1(x, Q^2) \right. \right. 
\]
\[-\frac{y}{2}(\kappa - 1)g_2(x, Q^2)\] \tag{2.7}

\[-\sin \alpha \cos \phi \sqrt{(\kappa - 1) \left( 1 - y - \frac{y^2}{4}(\kappa - 1) \right)} \]
\times \left( \frac{y}{2} g_1(x, Q^2) + q_2(x, Q^2) \right) \}

and
\[
\frac{d\sigma}{dx\, dy\, d\phi} = \frac{e^4}{4\pi^2 Q^2} \left\{ \frac{y}{2} F_1(x, Q^2) + \frac{1}{2xy} \right. \\
\times \left. \left( 1 - y - \frac{y^2}{4}(\kappa - 1) \right) F_2(x, Q^2) \right\}. \tag{2.8}
\]

The new variables introduced in Eqs. (2.7) and (2.8) simplify the expression for the cross section considerably. They are summarized in Table I. Note the following:

1. \( \alpha \) is the angle between the spin vector of the target \((\hat{s})\) and the incident electron beam \((\hat{k})\), not the virtual photon direction \((\hat{q})\).

### TABLE I

<table>
<thead>
<tr>
<th>Definitions and kinematic relations (see Fig. 1 also)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( q^2 = -Q^2 &lt; 0 ), ( q \cdot P = v \geq \frac{Q^2}{2} );</td>
</tr>
<tr>
<td>( x = \frac{Q^2}{2\nu}, \ y = \frac{v}{ME}; )</td>
</tr>
<tr>
<td>( \kappa = 1 + \frac{M^2Q^2}{\nu^2} = 1 + \frac{4M^2\nu^2}{Q^2} )</td>
</tr>
<tr>
<td>( \frac{2M\nu}{Q^2} = 1 )</td>
</tr>
<tr>
<td>( \cos \theta = \frac{1}{\sqrt{\kappa}} \left( 1 + \frac{y}{2}(\kappa - 1) \right) )</td>
</tr>
<tr>
<td>( \sin \theta = \sqrt{\frac{\kappa - 1}{\kappa}} \left( 1 - y - \frac{y^2}{4}(\kappa - 1) \right) )</td>
</tr>
</tbody>
</table>

245
2. $\phi$ is the azimuthal angle between the plane defined by $k$ and $k'$ and the plane defined by $k$ and $\hat{s}$.

3. Equations (2.7) and (2.8) are exact (except that lepton masses have been ignored): no scaling limit has been taken. $\kappa - 1 = M^2 Q^2 / v^2 = 4 M^2 x^2 / Q^2$ is a measure of the approach to the scaling limit, $Q^2 \rightarrow \infty$.

4. To eliminate spin-independent effects one may either (i) subtract cross sections for different values of $\alpha$; (ii) subtract cross sections for right- and left-handed leptons; or (iii) measure $\phi$-dependence.

Notice that effects associated with $g_2(x, Q^2)$ are suppressed by a factor $\sqrt{\kappa - 1} = 2 M_x / \sqrt{Q^2}$ with respect to the dominant structure function $g_1(x, Q^2)$. However, at $90^\circ$ the coefficient of the dominant term vanishes identically and allows the combination $(y/2) g_1 + g_2$ to be extracted cleanly at large $Q^2$. This is a unique feature of the spin-dependent scattering, and, we shall see, enables one to probe the higher twist structure of the target at leading asymptotic order in $Q^2$.

B. Dispersion Relations

The amplitudes $A_j(q^2, \nu), j = 1,2$, are analytic functions of $\nu$ (at fixed $q^2$) with cuts in the physical regions $|\nu| \geq |q^2/2|$. $A_1(q^2, \nu)$ and $A_2(q^2, \nu)$ are, respectively, even and odd under crossing ($q^\mu \rightarrow -q^\mu, P^\mu \rightarrow P^\mu$):

$$A_1(q^2, -\nu) = A_1(q^2, \nu),$$

$$A_2(q^2, -\nu) = -A_2(q^2, \nu).$$

Combining Cauchy’s theorem with crossing one obtains dispersion relations*

$$A_1(q^2, \nu) = \frac{2}{\pi} \int_{-q^2/2}^{\infty} \frac{v' dv'}{\nu'^2 - \nu^2} \text{Im} A_1(q^2, \nu'),$$

$$A_2(q^2, \nu) = \frac{2}{\pi} \nu \int_{-q^2/2}^{\infty} \frac{dv'}{\nu'^2 - \nu^2} \text{Im} A_2(q^2, \nu').$$

*Actually $A_j(q^2, \nu)$ is only defined by Eq. (2.10) up to a polynomial in $\nu$ (with $q^2$-dependent coefficients). Such an ambiguity does not disrupt any of our subsequent analysis.

246
The convergence of these dispersion relations can be studied using Regge theory applied to virtual Compton amplitudes. A standard analysis yields, as \( \nu \to \infty \), at fixed \( q^2 \),

\[
A_1(q^2, \nu) \sim \sum_i \beta_i(q^2) \nu^{\alpha_i(0) - 1},
\]

\[
A_2(q^2, \nu) = \frac{\beta_c(q^2)}{(\ln \nu)^2} + \sum_i \frac{\beta_i(q^2)}{\ln \nu} \nu^{\alpha_i(0) - 1}.
\]

\[ (2.11) \]

\( A_1 \) receives contributions from Regge poles with intercept \( \alpha_i(0) \) and residue \( \beta_i(q^2) \). \( A_2 \) receives contributions from a possible multi-Pomeron (three or more) cut (\( \sim (\ln \nu)^{-5} \)) and Regge-Pomeron cuts (\( \sim \nu^{\alpha_i(0) - 1}/\ln \nu \)). For further discussion, see Ref. 6. On very general grounds we expect \( \alpha_i(0) < 1 \) for all trajectories, so the integrals in Eq. (2.10) converge.

For large \( q^2 \) it is more convenient to study scaling forms,

\[
\alpha_1(\omega, Q^2) = 4 \omega \int_1^{\infty} \frac{d\omega'}{\omega'^2 - \omega^2} g_1(\omega', Q^2),
\]

\[
\alpha_2(\omega, Q^2) = 4 \omega^3 \int_1^{\infty} \frac{d\omega'}{\omega'^2(\omega'^2 - \omega^2)} g_2(\omega', Q^2)
\]

\[ (2.12) \]

where \( \omega = 1/x \). Both \( \alpha_1 \) and \( \alpha_2 \) are analytic within the circle \(|\omega| \leq 1\), and Eq. (2.13) gives us Taylor series

\[
\alpha_1(x, Q^2) = \sum_{n=0,2,4,\ldots} \left( \frac{1}{x^n} \right) \int_0^1 dy y^n g_1(y, Q^2),
\]

\[
\alpha_2(x, Q^2) = \sum_{n=0,2,4,\ldots} \left( \frac{1}{x^n} \right) \int_0^1 dy y^{n+2} g_2(y, Q^2)
\]

\[ (2.13) \]

guaranteed to converge for \(|\omega| < 1 \) (\(|x| > 1\)). These equations are useful in the operator product expansion analysis of \( g_2(x, Q^2) \).

\[ \dagger \text{The Pomeron pole doesn't couple to either } A_1 \text{ or } A_2. \]
III. BURKHARDT–COTTINGHAM SUM RULE

Twenty years ago, Burkhardt and Cottingham\textsuperscript{10} proposed a sum rule

\[ \int_{-q^2/2}^{\infty} dv \text{Im} A_2(q^2, \nu) = 0 \]  

or

\[ \int_0^1 dx g_2(x, Q^2) = 0 \]  

in terms of scaling variables. Subsequently, claims have been made to have derived this sum rule using parton model or operator product expansion methods. These claims notwithstanding, the sum rule relies on assumptions of Regge theory which have never been tested and may not apply to the large \( q^2 \) limit of Compton scattering at all. Nevertheless, the sum rule is interesting; its deviation should be re-examined and it should be tested.

Burkhardt and Cottingham's derivation of Eq. (3.1) is deceptively simple. First, they suppose that the large \( \nu \) behavior of Compton amplitudes at fixed but non-zero \( q^2 \) is governed by Regge theory. Next they argue that all known Regge singularities contributing to \( A_2(q^2, \nu) \) have intercept less than zero. Then from Eq. (2.11) (ignoring the \( (\ln \nu)^{-5} \) term), as \( \nu \rightarrow \infty \),

\[ A_2(q^2, \nu) \sim \nu^{-1-\epsilon} \]  

for some \( \epsilon < 0 \) (modulo logarithms). In this case the large \( \nu \) behavior of \( A_2 \) can be examined by taking \( \nu \rightarrow \infty \) under the \( \nu' \)-integral in Eq. (2.10),

\[ A_2(q^2, \nu) \sim -\frac{2}{\pi \nu} \int_{-q^2/2}^{\infty} dv' \text{Im} A_2(Q^2, \nu') \]  

which contradicts the assumed behavior (Eq. (3.3)) unless the integral vanishes; hence the sum rule.

This "derivation" raises many questions. First, does Regge theory apply to virtual Compton scattering at all? Arguments based
on generalized vector meson dominance suggest it does, at least at fixed $q^2$.
Assuming the applicability of Regge theory, Burkhardt and Cottingham argue that known Regge poles contributing to $A_2$ have intercepts $\alpha(0) < 0$. More recently, Ioffe, Khoze and Lipatov argue that Regge poles do not contribute to $A_2$ at all, but that Regge cuts with branch points at $\alpha(0) = 0$ (including a multi-Pomeron cut) spoil the sum rule: in this case $\int dv' \Im A_2(q^2, v')$ would not exist; it would diverge at its lower limit. Ioffe et al. then go on to argue (on the basis of light-cone current algebra) that the residues of these cuts fall rapidly with $Q^2$ so that the asymptotic form of the sum rule, Eq. (3.2) (where it is understood that the limit $Q^2 \to \infty$ is taken under the integral), remains valid. It seems to me, however, that their argument is circular: they exclude a short distance singularity which would have no other physical effect than to invalidate the sum rule.*

A final, more insidious possibility is that all conventional Regge trajectories do have intercept below zero, so $\int dv' \Im A_2(q^2, v')$ converges, but $A_2(q^2, v)$ has a Regge singularity at $\alpha(0) = 0$, a so-called “$J = 0$ fixed pole.” Fixed (in $t$) poles in hadron–hadron scattering amplitudes are excluded by generalized non-linear unitarity. Fixed poles are not excluded from Compton amplitudes because these are calculated only to lowest non-trivial order in $e$ — the electric charge — and are not subject to the full set of non-linear constraints of unitarity. The Adler sum rule and other current algebra sum rules require certain fixed poles in a (generalized) Compton amplitude. Note than an $\alpha(0) = 0$ pole does not contribute to $\Im A_2(q^2, v)$ since $\Im 1/v = 0$ for $v$ real and non-zero, and therefore does not spoil the convergence of the sum rule. We now recognize

$$\beta_0(q^2) = -\frac{2}{\pi} \int_{-q^2/2}^{\infty} dv \Im A_2(q^2, v)$$

as the “residue of a $J = 0$ fixed pole” coupling to $A_2(q^2, v)$.

Some years ago Cheng and Tung incorporated scaling into this

---

*A similar argument is made by Heimann, Ref. 3.
†Such singularities and their effect on current algebra sum rule are discussed at length in Refs. 19 and 20 and by Heimann, Ref. 3. My discussion is based on these references.
analysis to argue that $\beta_0(q^2) = 0$ in theories, like QCD, which scale (modulo logarithms) in the Bjorken limit. They point out that scaling ($\lim_{q^2 \to \infty} (v^2/M^4) A_2(q^2, v) = \alpha_2(\omega, Q^2) \approx \alpha_2(\omega) \text{ mod logarithms}$) requires

$$\beta_0(q^2) \sim \frac{1}{q^2}$$  \hspace{1cm} (3.6)

(mod logarithms) as $q^2 \to \infty$. Thus $\beta_0(q^2)$ cannot be an entire function (a polynomial, since essential singularities as $q^2 \to \infty$ are excluded on general grounds) of $q^2$ and must have singularities at some finite $q^2$. These correspond to the virtual photon turning into some on-shell hadronic state (and therefore occur for $Q^2 > 0$). Cheng and Tung then argue that this would result in $J = 0$ fixed poles in hadron–hadron scattering amplitudes which are excluded by unitarity. Thus, they conclude $\beta_0(q^2) \sim 1/q^2$ is forbidden and the Burkhardt–Cottingham sum rule is valid. Cheng and Tung’s argument is consistent with the fixed poles in Compton amplitudes required by current algebra, which in all cases have polynomial (in $q^2$) residues.$^{11,18}$

This chain of argument is complicated and should be re-examined in QCD. We are left with the conclusion that the Burkhardt–Cottingham sum rule is valid provided:

1. It converges—i.e., there are no ordinary (“moving”) Regge trajectories with $\alpha(0) \geq 0$ coupling to $A_2(q^2, v)$.
2. Non-polynomial residue $J = 0$ fixed poles are forbidden in Compton amplitudes.

IV. OPERATOR PRODUCT EXPANSION ANALYSIS OF $g_2(x, Q^2)$

The most reliable method for exploring the properties of structure functions in the deep inelastic limit of QCD is the operator product expansion (OPE). Where the parton model is trustworthy it can be mapped onto an OPE analysis.$^{21}$ In the case of $g_2(x, Q^2)$ use of parton model methods without sufficient attention to the underlying operator structure has resulted in a lot of confusion. As far as I know, the first completely correct discussion of the properties of $g_2(x, Q^2)$ at asymptotic $Q^2$ is that of Shuryak and Vainshtein.$^4$ I follow their analysis here, though much of the preliminary work
is drawn from Ref. 7. Further technical work along these lines is presented in Ref. 5. I consider only the case \( m_{\text{quark}} = 0 \). Light quark mass effects are suppressed by \( m/\sqrt{Q^2} \) in \( F_1(x, Q^2) \) or \( g_1(x, Q^2) \). They are more important in \( g_2(x, Q^2) \) where they enter \( \propto m/\Lambda_{\text{QCD}} \).* Up and down quark mass effects are still likely to be small, but strange quark mass effects are not small. They have been included in the OPE analysis by Kodaira et al.\(^7\)

The leading order contributions to the antisymmetric part of the Compton amplitude \( T_{\mu\nu} \) comes from a tower of simple operators

\[
\mathcal{J}_{[\mu\nu]} = i\epsilon_{\mu\nu\lambda\sigma}q^\lambda \sum_{n=0,2,4} \left(-\frac{2}{q^2}\right)^{n+1} q_{\mu_1} \cdots q_{\mu_n} \Theta^{\sigma_{\mu_1} \cdots \mu_n} \tag{4.1}
\]

where

\[
T_{[\mu\nu]}(q, P, S) = \langle PS|\mathcal{J}_{[\mu\nu]}(q)|PS\rangle \tag{4.2}
\]

and

\[
T_{[\mu\nu]} = \frac{1}{2} (T_{\mu\nu} - T_{\nu\mu}). \tag{4.3}
\]

The operators \( \Theta^{\sigma_{\mu_1} \cdots \mu_n} \) are defined by

\[
\Theta^{\sigma_{\mu_1} \cdots \mu_n} = i^n \bar{\psi} \gamma^\sigma \gamma^5 D^{(\mu_1} D^{\mu_2} \cdots D^{\mu_n)} \psi - \text{traces}. \tag{4.4}
\]

The terms denoted “traces” are whatever is necessary to subtract from the displayed term in order to render the resulting operator traceless: \( g_{\sigma\mu k} \Theta^{\sigma_{\mu_1} \cdots \mu_n} = g_{\mu k \mu} \Theta^{\sigma_{\mu_1} \cdots \mu_n} = 0. \)† These operators are renormalization scale \((\mu^2)\)-dependent, but I suppress that label.

---

*Much of the confusion surrounding \( g_2(x, Q^2) \) is due to effects \( O(m/\Lambda_{\text{QCD}}) \) which vanish as \( m \to 0 \) in QCD but are important, even dominant, if confinement is ignored \((\Lambda_{\text{QCD}} \to 0)\).\(^9\)

†The operator provided by the OPE does not come traceless; terms subtracted to render it traceless must be added back in and generate twist \( \geq 4 \) corrections to structure functions.\(^12\)-\(^14\)
The symbol \{ab \ldots \} denotes the symmetrization:

\[ \{a_1 \ldots a_n\} = \frac{1}{n!} \sum_{\text{permutations}, \mathcal{P}} P(a_1 a_2 \ldots a_n). \]  

Equation (4.1) is accurate to twist-4; that is, it correctly gives the scaling (twist-2) and the \(O(1/\sqrt{Q^2})\) (twist-3) behavior of \(T_{\mu\nu}\) for \(m_{\text{quark}} = 0.4, 5, 7\).

The symbol \{ab \ldots \} denotes the symmetrization:

\[ \{a_1 \ldots a_n\} = \frac{1}{n!} \sum_{\text{permutations}, \mathcal{P}} P(a_1 a_2 \ldots a_n) \]

The symbol \{ab \ldots \} denotes the symmetrization:

\[ \{a_1 \ldots a_n\} = \frac{1}{n!} \sum_{\text{permutations}, \mathcal{P}} P(a_1 a_2 \ldots a_n) \]

For \(n = 0, 2, 4, \ldots\), and

\[ \Theta_{\{\sigma_1 \mu_1 \ldots \mu_n\}} = \frac{1}{n + 1} \left[ \Theta_{\{\sigma_1 \mu_1 \ldots \mu_n\}} + \Theta_{\{\sigma_1 \mu_1 \mu_2 \ldots \mu_n\}} \right. \]

\[ + \Theta_{\{\mu_1 \sigma \mu_1 \ldots \mu_n\}} + \ldots \]

for \(n = 2, 4, 6, \ldots\). The proton matrix elements of each of these operators is described by a single (dynamically determined, scale-dependent) number because it is only possible to form a single traceless tensor of rank-\(n\) from \(P_{\mu}, g_{\lambda\nu}\), linear in \(s_o\) and traceless. Thus,

\[ \langle Ps|\Theta_{\{\sigma_1 \mu_1 \ldots \mu_n\}}|Ps\rangle = \frac{a_n}{n + 1} [s_o P_{\mu_1} P_{\mu_2} \ldots P_{\mu_n}] \]
for \( n = 0, 2, 4 \ldots \) and

\[
(\langle s P | \Theta_{[\mu_1, \ldots, \mu_n]} | P s \rangle) = \frac{d_n}{n+1} [(s_{\mu_1} P_{\mu_2} - s_{\mu_2} P_{\mu_1}) P_{\mu_2} \ldots P_{\mu_n} \\
+ (s_{\mu_2} P_{\mu_3} - s_{\mu_3} P_{\mu_2}) P_{\mu_2} \ldots P_{\mu_n} + \ldots - \text{traces}]
\]  

for \( n = 2, 4, 6, \ldots \). It is now a simple matter to combine Eqs. (4.8) and (4.9) with Eqs. (4.1) and (2.1) to obtain expressions for \( \alpha_1(x, Q^2) \) and \( \alpha_2(x, Q^2) \),

\[
\alpha_1(x, Q^2) + \alpha_2(x, Q^2) = \sum_{n=0, 2, 4, \ldots} a_n + nd_n \frac{1}{n+1} x^{-n-1}, \tag{4.10}
\]

\[
\alpha_2(x, Q^2) = \sum_{n=2, 4, \ldots} \frac{n(d_n - a_n)}{n+1} \frac{1}{x^{-n-1}}. \tag{4.11}
\]

Equations (4.10) and (4.11) can be recognized as Taylor expansions of \( \alpha_1 \) and \( \alpha_2 \) about \( \omega = 1/x = 0 \), where they are analytic. If we compare these with the Taylor expansions derived from dispersion relations in Section II, we obtain moment sum rules for \( g_1 \) and \( g_2 \):

\[
\int_0^1 dx x^n g_1(x, Q^2) = \frac{1}{4} a_n \quad n = 0, 2, 4, \ldots, \tag{4.12}
\]

\[
\int_0^1 dx x^n g_2(x, Q^2) = \frac{1}{4n+1} (d_n - a_n) \quad n = 2, 4, \ldots. \tag{4.13}
\]

Note the absence of any information about \( \int_0^1 dx g_2(x, Q^2) \). Contrary to some assertions in the literature, the OPE tells us nothing about the Burkhardt–Cottingham sum rule: the absence of an \( n = 0 \) in Eq. (4.11) is matched by the absence of a \( 1/x \) term in Eq. (2.13). A knowledge of the matrix elements \( \{a_n, n = 0, 2, 4, \ldots \} \) together with the fact that \( g_1(x, Q^2) \) is even (\( g_1(x, Q^2) \))
= g_1(-x, Q^2)) completely determines g_1(x, Q^2) by analytic continuation. A knowledge of \{a_n\} and \{d_n, n = 2, 4, \ldots\} does not completely determine g_2(x, Q^2) because its lowest moment, the Burkhardt–Cottingham integral, is unknown.*

Notice that g_2(x, Q^2) receives contributions both from \Theta^{(\sigma, \mu_1, \ldots, \mu_n)} and \Theta^{[\sigma, [\mu_1] \mu_2, \ldots, \mu_n]}. The first operator has twist-2 (dimension \(n + 3\), spin \(n + 1\)), the second has twist-3 (dimension \(n + 3\), spin \(n\)). Suppose, for some reason, \(\{d_n\} = 0\); then comparing Eqs. (4.12) and (4.13) it follows that

\[
g_2^{WW}(x, Q^2) = -g_1(x, Q^2) + \int_0^1 dy \frac{y}{y} g_1(y, Q^2).
\]  

(4.14)

This result was first proposed by Wandzura and Wilczek\(^{15}\); hence the superscript "WW." In general then

\[
g_2(x, Q^2) = g_2^{WW}(x, Q^2) + g_2(x, Q^2)
\]  

(4.15)

where

\[
\int dx x^n \bar{g}_2(x, Q^2) = \frac{n}{4(n + 1)} d_n, \quad n = 2, 4, \ldots.
\]  

(4.16)

Note that the Wandzura–Wilczek piece of \(g_2\) obeys the Burkhardt–Cottingham sum rule—\(\int_0^1 dx g_2^{WW}(x, Q^2) = 0\)—provided \(g_1\) is sufficiently well-behaved as \(x \to 0\) to allow the exchange of the \(x\) and \(y\) integrations.

The arguments made by Wandzura and Wilczek\(^{15}\) and others\(^{22}\) to support the ansatz \(\{d_n\} = 0\) need not be repeated in detail here. They all amount to the observation that the matrix element of \(\Theta^{[\sigma, [\mu_1] \mu_2, \ldots, \mu_n]}\) vanishes in an on-shell massless quark state moving colinearly with the target in an infinite momentum frame. If this is defined to be the "parton model," then indeed \(\{d_n\} = 0\). However, in such a naive model, familiar structure functions like \(F_1\) and \(g_1\)

---

*An attempt to define the \(n = 0\) moment by analytic continuation in \(n\) would be frustrated by the Regge singularities with \(\alpha \geq 0\) discussed in Section III.
are $\delta$-functions at $x = 1/3$. Normally, parton model predictions are not altered by allowing quarks to have limited transverse momenta or virtuality. Such effects, however, completely eliminate the prediction $\{d_{11}\} = 0$.\(^8\)

The physical origin of the twist-3 contribution to $g_2(x, Q^2)$ was made apparent by Shuryak and Vainshtein,\(^4\) who pointed out that the equations of motion of QCD could be used to trade the antisymmetry of $\Theta^{[\sigma, \{\mu_1, \mu_2 \ldots \mu_n\}}$ for factors of the gluon field strength $G_{\mu \nu}$ and the QCD coupling $g$. Rather than reproduce their calculation which is tedious, here is an example applied to a simpler twist-3 operator $X_{\mu \nu} = \bar{\psi} D_\mu D_\nu \psi$. Using $D_\mu = \frac{i}{2} \{\gamma_\mu, D\},$ $\bar{D} \psi = 0$, $\bar{\psi} D = 0$ and $[D_\mu, D_\nu] = g G_{\mu \nu}$,

$$X_{\mu \nu} = \frac{1}{2} \bar{\psi} D_\mu D_\nu \gamma_\nu \psi$$

(4.17)

$$= \frac{1}{2} g \bar{\psi} G_{\mu \lambda} \gamma^\lambda \gamma_\nu \psi.$$

This is a specific case of a general result that operators of twist $\geq 3$ can always be written in a form in which they are manifestly interaction dependent.\(^{12-14}\) Shuryak and Vainshtein’s general result is

$$\Theta^{[\sigma, \{\mu_1, \mu_2 \ldots \mu_n\}}$$

$$= S_n \left[ \frac{g}{8} \sum_{i=0}^{n-2} i^{n-2} \bar{\psi} D_{\mu_1} \ldots D_{\mu_i} G_{\sigma \mu_{i+1}} D_{\mu_{i+2}} \ldots D_{\mu_{n-1}} \gamma_{\mu_n} \psi ight]$$

$$+ \frac{g}{16} \sum_{i=0}^{n-3} i^{n-3} \bar{\psi} D_{\mu_1} \ldots D_{\mu_i} (D_{\mu_{i+1}} G_{\sigma \mu_{i+2}}) D_{\mu_{i+3}} \ldots D_{\mu_{n-1}} \gamma_{\mu_n} \gamma_5 \psi$$

$$- \text{traces} \right]$$

(4.18)
where $\tilde{G}_{\alpha\beta} = \frac{1}{2} \epsilon_{\alpha \beta \lambda \sigma} G^{\lambda \sigma}$ and $S_n$ symmetrizes the indices $\mu_1 \ldots \mu_n$.

The simplest operator

$$\Theta_{[\sigma, (\mu_1 \mu_2)]} = \frac{g}{8} \bar{\Psi} (\tilde{G}_{\sigma \mu_1} \gamma_{\mu_2} + \tilde{G}_{\sigma \mu_2} \gamma_{\mu_1}) \psi$$  (4.19)

determines the second moment of $g_2(x, Q^2)$.

The observation of Shuryak and Vainshtein has both good and bad news for students of the spin structure of the nucleon. On the one hand, there is no a priori reason for $g_2(x, Q^2)$ to be small in an interacting theory. Furthermore, $g_2(x, Q^2)$ cannot be profitably studied in a parton model: $g_2(x, Q^2)$ is determined by the $x$-dependence of quarks' transverse momenta and off-shellness which are unknown in the parton models. On the other hand, $g_2(x, Q^2)$ is easily accessible to experiment, and measuring it will provide well-defined information about the nucleon matrix elements of specific interaction-dependent quark–gluon operators. If carried out, these would be the first measurements of higher twist operator matrix elements, eventually to be followed by the extraction of twist-4 effects from the approach to Bjorken scaling in spin-independent electron scattering.

Acknowledgments

The need for such a reconsideration of $g_2$ was impressed upon me by Richard Milner, Klaus Rith and Dietrich von Harrach. Much of this work evolved from conversations with Anees Manohar, Graham Ross and Arkady Vainshtein, many of which took place at a QCD Study Group in Santa Fe in August 1989. I am grateful to Gerry Garvey of LAMPF for his sponsorship and hospitality in Santa Fe. I am grateful to Xiangdong Ji for many conversations, for bringing several references to my attention and for a careful reading of the manuscript. I thank Lech Mankiewicz and Andreas Schäfer for checking Eq. (2.8). Finally, I wish to thank Allen Olsen for collaboration on issues surrounding $g_2$. This work is supported in part by funds provided by the U.S. Department of Energy (D.O.E.) under contract #DE-AC02-76ER03069.

R. L. JAFFE
Center for Theoretical Physics,
Laboratory for Nuclear Science
and Department of Physics,
Massachusetts Institute of Technology,
Cambridge, Massachusetts 02139
References

1. SMC proposal, CERN/SPSC 88-47.
2. Private communication, the HERMES Collaboration.
9. R. L. Jaffe and X. Ji, to be published.