Power Corrections and Higher Twist in QCD

H. David Politzer*

California Institute of Technology, Pasadena, California 91125, U.S.A.

People believe QCD to be the theory of strong interactions for a variety of reasons. Firstly, $SU(3)$ of color is a subgroup of $SU(5)$; second, it has instantons and a variety of other interesting topological properties; and, lastly, it agrees with high energy experiments. But on this last point, one must confess that there are a variety of obstacles that have thus far precluded any overwhelmingly convincing numerical confrontation with data. Theorists divide the sources of troubles into several classes. One of these is the power or higher twist corrections. I describe here a tentative approach of a systematic analysis of power corrections. The analysis is motivated by, and generalizes on, the operator product analysis of leptoproduction. In essence, I try to identify the appropriate set of process-independent, non-perturbative functions that characterize long-distance physics; these are to be convoluted with calculable short-distance amplitudes via an algorithm that generalizes the long distance-short distance factorization for the scaling limit.

By way of introduction, I wish to emphasize why power corrections are phenomenologically important at present energies. Consider the examples of scaling violations, $p + p \rightarrow \pi + X$ at large $p_\perp$, and jet shapes:

Power corrections to the scaling behavior of structure functions are expected to be roughly of the form

$$f(x, \log Q^2/\Lambda^2) \left[1 + \frac{1}{1-x} \frac{M^2}{Q^2} \right] \tag{1}$$

The $(1-x)^{-1}$ is necessary as $x \rightarrow 1$ because it is precisely the behavior needed to make the power correction term in (1) of constant magnitude as $Q^2 \rightarrow \infty$ for x in the resonance region. Data can be fit by $M \simeq 1.2$ GeV and $\Lambda \simeq 0$ almost as well as by $M \simeq 0.4$ GeV and $\Lambda \simeq 0.5$ GeV. Clearly, anything inbetween would do as well. In principle from QCD, we know that $M \propto \Lambda$. But without a theory of the constant of proportionality, we must think seriously about non-negligible power corrections. If we were only interested in the scaling limit, it is plausible that the M^2/Q^2 term is small compared to one. But if we wish to ask about deviations from exact scaling, the real question is how does $(1-x)^{-1}M^2/Q^2$ compare to $\alpha_s(Q^2)$? The very fact that the highest-Q^2 data now available favors a Λ smaller than that needed to fit earlier data suggests that indeed power corrections are significant in the scaling violations for $Q^2 \lesssim 15$ GeV^2.

In processes like $p + p \rightarrow \pi + X$, it is now appreciated that a smearing in "primordial" $k_\perp$ of order 0.5 GeV can enhance the production of π's at $p_\perp \sim 2$ GeV by an order of magnitude. This is because the underlying hard quark–quark scattering falls so steeply with $p_\perp$. These smearing effects are also formally power

corrections, and they disappear for large enough $p_\perp, \ldots$ but not yet.

In e^+e^- annihilation, we study the shapes of jets and attempt to extract a value for α_s and the gluon spin. The experiments look very little like the direct quark–gluon predictions, which again must be smeared by a non-perturbative $k_\perp$ that occurs in hadronization. But the significance on shapes of a $k_\perp$ of order 0.4 GeV is relative to the average $k_\parallel$ (longitudinal) per particle, i.e., 0 (2 GeV), and not to the total center-of-mass energy. And 0.4/2 is comparable to α_s. Of course, $k_\perp$ is roughly fixed, while $k_\parallel$ grows almost linearly with the total energy. So again these considerations are negligible asymptotically, ... but not yet.

Next, let me set some ground rules: I will consider here only $m^2_{\text{quark}} \approx 0$ (i.e., $\ll \Lambda^2$). Also, to make the structure clearer, I will assume that the masses of the observed hadrons are also negligible. (One might imagine measuring pions in a world where $m^2_\pi \ll \Lambda^2$.) There are kinematic occurrences of observed hadron masses in the real world as given by the ξ-scaling analysis, and these m^2_{hadron}/Q^2 effects can be restored unambiguously. (They are often non-negligible.) However, it is all the remaining $1/Q^2$ dependences (i.e., Λ^2/Q^2) that are the real, dynamical problem.

As a further introduction to the analysis I propose, I raise certain criticisms of the method of smearing commonly employed. One typically thinks of an incoming hadron of momentum P as giving rise to a parton of momentum p such that

$$p \cong xP + p_\perp \tag{2}$$

or an outgoing parton gives rise to a spray of hadrons with momenta

$$P \cong zp + P_\perp \tag{3}$$

The $p_\perp$'s and $P_\perp$'s are generated according to universal probability distributions (e.g., Gaussians with a mean of several hundred MeV) and are convoluted à là Monte Carlo with the (on-shell) parton scattering cross sections. The trouble is that for relativistic bound states, the off-shellness, p^2, must be of the same order as $p_\perp^2$. Hence, if one is interested in $p_\perp$'s at all, it is inconsistent to do a convolution of probabilities with on-shell cross sections. The $p_\perp$ phenomena must be tied up with quantum interference effects. If one continues to use the very vivid picture of $p_\perp$ smearing, one must expect that the effective $p_\perp$ probability distributions are not process independent, but they in fact reflect interference within the specific process. — I suspect that $p_\perp$ smearing will continue to be a practical method for estimating power corrections.

Inclusive leptoproduction is a process for which a theory of power corrections already exists — the operator product expansion (OPE). (One may ask even here to what extent non-perturbative effects are accounted for, beyond their occurrence in operator matrix elements. One really doesn't know, except
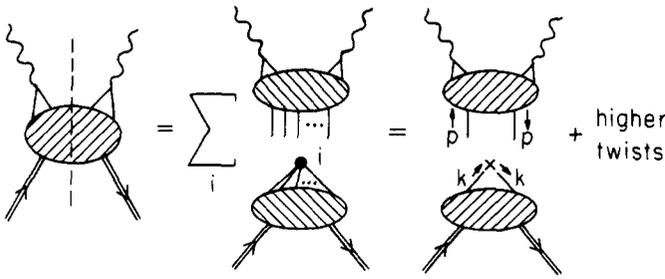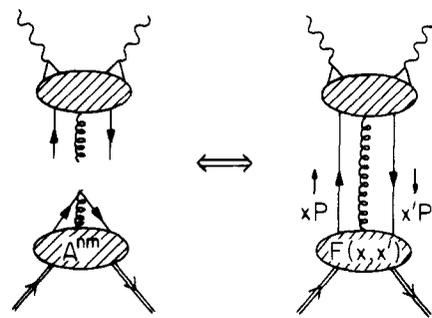
Fig. 1.

Fig. 2.

that the effects of small, dilute instantons appears only in very high powers of $1/Q^2$.) The structure of the OPE is indicated in Fig. 1. The leading or scaling contribution comes from bilinear operators, which have a direct single-active-parton interpretation. Note that one never assumes that the quark or gluon fields are on-shell. In the lower part of the figure, it is essential that the four-momentum k run over all values. That is why the matrix elements or parton structure functions are not trivial, free field δ-functions, but rather they reflect that binding. The constituent momentum p in the coefficient function or hard scattering part need not be on-shell. However, the scaling contributions are only sensitive to its longitudinal part, while effects of $p^2 \neq 0$ are down by p^2/Q^2.

Power corrections come from operators of higher twist, where twist is defined as "mass dimension minus spin". The extra twists come from either the presence of more fields (i.e., more active constituents in the hard scattering) or from contracted derivatives in the operators, which measure the off-shellness of the active constituents. For example, the ∂^2 in $\bar{\psi}\gamma^\mu \partial^\nu \partial^2 \psi$ measures the off-shellness of the ψ.

But the set of operators that are generated in the OPE are overcomplete, if we take hadron matrix elements, by virtue of the operator equations of motion. That is to say that we know the matrix elements must satisfy certain linear relationships, even if we know nothing about the states. For example, in a ϕ^4 theory

$$\langle P|\phi\partial^{\mu_1}\ldots\partial^{\mu_n}(\partial^2 + m^2 + \lambda\phi^2)\phi|P\rangle = 0 \qquad (4)$$

So the effects of off-shellness (the non-vanishing of $\partial^2 + m^2$) are always related to the possible presence of yet more fields (the extra ϕ^2). But the contribution of the $\phi\partial^{\mu_1}\ldots\partial^{\mu_n}\phi^2\phi$ operator in the hard scattering or coefficient function can be evaluated by treating all the ϕ's as on-shell, as in the parton picture of twist-2 operators.

In general, we can always choose an operator basis from which we have eliminated all contracted derivatives. In QCD we would eliminate all $D, D^2, D \cdot F$ type operators in favor of ones containing more fields. The physical effects of the former are simply re-expressed using the latter basis. In the basis thus chosen, all coefficient functions or hard scattering amplitudes can be evaluated by placing the active participants on-shell, and, again, only their longitudinal momenta are relevant. Hence in the massless limit they are collinear.

A Mellin transformation always allows us to translate from the OPE description, which uses spins as variables, to a parton picture which uses longitudinal momentum fractions as variables. For example, one can define $F(x, x')$ by

$$\int F(x, x')x^n x'^m dx\, dx' = \langle P|\bar{\psi}\overbrace{D \ldots D}^{n} F \overbrace{D \ldots D}^{m}\psi|P\rangle$$

$$= A^{nm} \qquad (5)$$

whose interpretation is suggested in Fig. 2.

To be honest, I do not yet have a simple algorithm for computing the hard amplitude directly from Feynman diagrams. The problem is particularly acute for gauge theories because the diagrams do not distinguish between a gauge field coming from a D^μ and one from an $F^{\mu\nu}$. Also, there is a rule that the collinear legs cannot coalesce but must each enter a vertex through which there is a hard flow.

Since I have chosen a collinear, on-shell basis (to make clear the separation of long and short distances), one may ask where in fact do transverse momenta arise. In leptoproduction, σ_L is sensitive to $p_\perp$'s. In the scaling limit, Fig. 3(a) gives the $a \cdot 1$ contribution to σ_T in eq. (6) below, while the calculable bremsstrahlung correction in Fig. 3(b) gives a σ_L of order g^2, coming from the region of integration where the virtual quark is off-shell and has $p_\perp^2$ of order Q^2. But the same quark distribution enters. Hence, schematically $\sigma_L = ag^2$. We can now go to some twist-4 M^2/Q^2 corrections to σ_T which come from diagrams such as in Fig. 3(c). Their contribution is schematically called $a'M^2/Q^2$ in eq. (6). But to get a σ_L from these operators or constituent sets, we need a hard gluon to be radiated as in Fig. 3. This is similarly a calculable process, given the non-perturbative information represented by a'. In summary,

$$\sigma_T = a \cdot 1 + \ldots \qquad + a'\frac{M^2}{Q^2} + \ldots$$

$$\sigma_L = a \cdot g^2(Q^2) + \ldots + a'g^2(Q^2)\frac{M^2}{Q^2} + \ldots \qquad (6)$$

The OPE is in essence a high Q^2 expansion, and we are using perturbation theory to compute the high $p_\perp$ wave function using radiative corrections to the low $p_\perp$ components (treated for simplicity as being at zero $p_\perp$). The low $p_\perp$ or longitudinal wave functions are non-perturbative.

A note of phenomenological interest: If M^2 is the scale that enters the power corrections in σ_T as in eq. (6), then the effective $\langle p_\perp^2 \rangle_{\text{primordial}}$ is of order $g^2(Q^2)M^2$. Hence if the $\langle p_\perp^2 \rangle_{\text{primordial}}$ is of order $(400\,\text{MeV})^2$, then the higher twist corrections to scaling may indeed be more important than originally imagined — with a net result of decreasing the best fit for Λ from scaling violations.

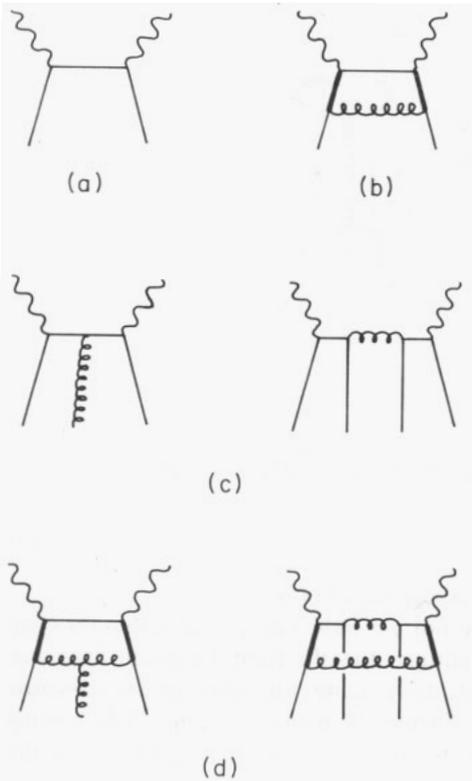I now propose a generalization to arbitrary inclusive hard scatterings, e.g., for $A + B \to C + X$:

Fig. 3.

$$d\sigma^{\text{hadron}}(P_A, P_B, P_C; \{q's\})$$

$$= \sum_{ijk} \int \Gamma^{\text{in}}_{A,i}(P_A, x^i_1, x^i_2, \ldots; M^2) \, dx^i_1 \, dx^i_2 \ldots$$

$$\Gamma^{\text{in}}_{B,j}(P_B, x^j_1, \ldots) \, dx^j_1 \ldots \Gamma^{\text{out}}_{C,k}(P_C, x^k_1, \ldots) \, dx^k_1 \ldots$$

$$d\sigma^{\text{hard}}_{ijk}(x^i_1 P_A, x^i_2 P_A, \ldots x^j_1 P_B, \ldots \frac{1}{x^k_1} P_C, \ldots,$$

$$\times \{q's\}, M^2$$

$$+ 0 \tag{7}$$

Figure 4 suggests the origin of the structure of this equation. In each channel corresponding to a measured hadron there is a sum over on-shell, collinear constituent sets, labelled $i, j, k \ldots$. A scaling contribution comes from the simplest sets of one parton on each side of the cut. Sets with more active constituents contribute with more inverse powers of Q^2 in the hard scattering. There may be any number of hard currents with momenta q that attach directly into the hard scattering amplitude. The Γ's are the non-perturbative, process independent distribution functions which depend on the longitudinal momentum fractions, x, of the constituents. An arbitrary scale M^2 must be introduced to formally separate long distance from short distance. Its dependence in the Γ's is cancelled by its dependence in $d\sigma^{\text{hard}}$. The sets i have the quantum numbers of $A \otimes A$; thus they are color singlets, but can carry isospin. However, the constituents on one side of the cut need not be color singlets. Furthermore, they need not be identical to the constituents on the other side of the cut. Hence, each term in $d\sigma^{\text{hard}}$ is not really a proper cross section.

In the proof (in perturbation theory) of the scaling factorization theorem, one considered the scattering of slightly off-shell partons, i.e., $p^2 \neq 0$, and found violations of the factorization of order p^2/Q^2. The present point of view is to take $p^2 \rightarrow 0$.
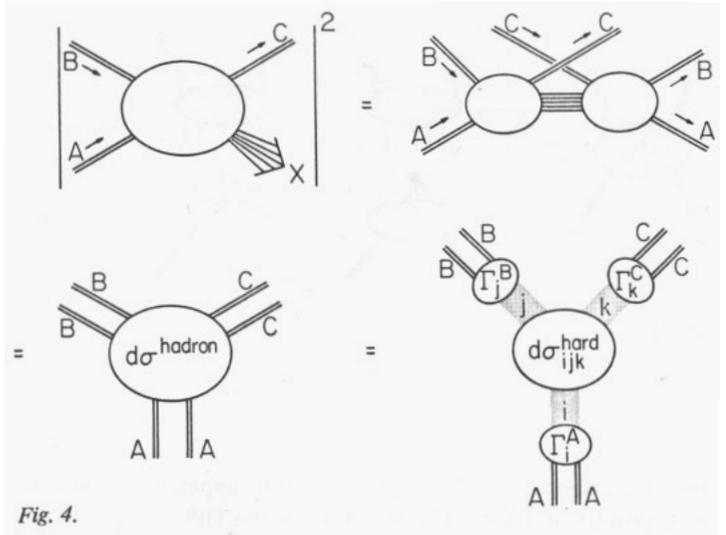
Fig. 4.

There is then perfect factorization for single parton scattering, but to give a complete description of hadron scattering one must consider also sets of partons.

Since no one has produced hadrons from QCD, eq. (7) cannot be derived from the theory. I am certain that an all-orders, all logarithms, perturbative proof of the appropriate analog theorem for quark and gluon scattering can be derived using the tools developed for the scaling factorization theorem. But I have no special wisdom on the very central question of to what extent are there non-perturbative effects that cannot be absorbed into the Γ's.

There is a very serious problem with actually using eq. (7) in practice. Put simply, the Γ's are virtually impossible to measure. This is because the external hadron kinematics can be used to fix the total momentum flowing in a given channel up one side of the cut, but one necessarily integrates over how that momentum is shared between the constituents. In different processes one integrates with different weight functions. So it is impossible to extract from experiment exactly how the Γ's depend on all their x's.

The integration over sharing fractions is characteristic of a truly short distance scattering, of which Fig. 5(a) is an example from inclusive leptoproduction. There are processes in which one might imagine measuring the longitudinal momenta separately. For example in $p + \bar{p} \rightarrow \mu^+\mu^- e^+ e^- + X$, Fig. 5(b) contributes, and the lepton pair masses can be used to deduce all of the quark longitudinal momenta. However, the separation of the two annihilations, transverse to the collision axis, is restricted only by the proton's radius. In next order, Fig. 5(b) will receive a Coulombic correction, not absorbable into the
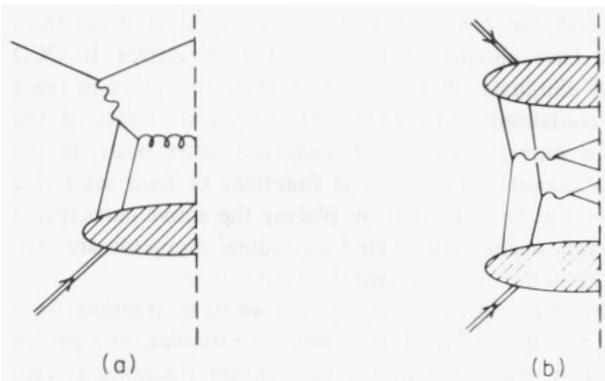

Fig. 5.

blobs, which is cut off only by the color neutrality on the scale of 1 fm. So the diquark correlation measured in principle by Fig. 5(b) is convoluted with the transverse wave function and is not precisely the one needed in Fig. 5(a). Furthermore, the contribution of Fig. 5(b) is down by $1/Q^2$ relative to a double radiation of pairs from a single quark–antiquark collision.

There are a few lessons that may already be drawn from these considerations. First, if one does think in terms of $p_\perp$ smearing, then the effective $p_\perp$ distributions depend on the specific hard scattering involved. The smearing algorithms may remain the simplest method of estimating the magnitude of power corrections on a given observable, but I conclude that they should not be taken literally.

Second, one might model, estimate, and/or parametrize the Γ's according to some theoretical ideas and thereby reduce the enormous task of extracting the Γ's completely from experiment. Bag-like or exclusive QCD-like arguments could be used to estimate the overall magnitudes and/or shapes of the Γ's. One might also choose to look in precisely those kinematic regions where power corrections are expected to be really large, e.g., as various k's $\to 1$. And one might identify qualitative features of the effects expected that are relatively insensitive to the precise values of the Γ's. For example, the effective $\langle p_\perp^2(x_{total})\rangle_{primordial}$, since it is generated by bremsstrahlung, probably vanishes as the x_{total} goes to zero or one for lack of available energy to do otherwise.

And third, by looking at certain moments, sum rules, or energy flows, one can extract predictions that can be tested more directly. If one goes back to the spin-space description natural to the OPE and remembers that a given Nachtmann moment projects out a fixed total spin in a given channel, the situation simplifies: for a fixed total spin, there are only a restricted, finite number of independent operators corresponding to dividing that spin among the constituents. For spin-2 operators, which govern the energy-momentum flow in a channel, I suspect there is a unique operator for each constituent set, leaving no ambiguity as to how to parse out its spin among the constituents.

I do not yet have a proper derivation, but I further conjecture that this last supposition leads to the following enormous simplification when one sums over observed hadrons, integrates over how momentum is shared among collinear constituents in a set, and looks only at energy flow. Then, the energy flow pattern should be something like

$$\frac{d\Sigma^{hadron}}{d\Omega} = \sum_{\{i,j,k\ldots\}} [\gamma_i\gamma_j\gamma_k\ldots] \frac{d\Sigma^{hard}_{ijk\ldots}}{d\Omega}. \tag{8}$$

Here $d\Sigma^{hard}_{ijk\ldots}$ is a calculable quantity telling how energy is distributed in space when the constituent sets $i, j, k, \ldots$ are produced. The γ_i's are dimensionful parameters, depending in principle on the scale M^2 (as does $d\Sigma^{hard}$), which are measures of the efficacy of the sets i that produce hadrons. For the simplest sets, with a single active parton, $\gamma_{\bar{q}q}$ or γ_{GG} are identically equal to 1, and eq. (8) reduces to the scaling prediction. The power corrections come when we treat as "physical" objects sets containing more constituents.

I have no real conclusion except that this subject requires more thought.

References

1. Politzer, H. D., Power Corrections at Short Distances, to be published in Nuclear Physics, and references therein.

Noted added in proof: "Factorization" for collinear, on-shell parton sets is necessary but not sufficient for the corresponding statement for hadron cross sections. I now believe that the latter is indeed more complicated than suggested in this paper.