Ref.TH.3241-CERN

SENSE AND NONSENSE IN THE RENORMALIZATION-SCHEME-DEPENDENCE PROBLEM

P. Stevenson
CERN - Geneva

A B S T R A C T

I discuss in detail the issues involved in the renormalization-scheme-dependence problem in perturbative QCD. The problem is not to find a good universal scheme; nor is a question of how to make coefficients small. The problem is that finite-order results depend on the choice of scheme, even though it is "arbitrary" (in that the exact result is scheme-independent). Only the "Principle of Minimal Sensitivity" approach gets to the heart of the problem, and attempts to reconcile this contradiction. I explain the motivation for this approach, defend it against recent criticisms, and explain to what extent it provides a solution to the problem.

Ref.TH.3241-CERN
9 February 1982

1000

1000 1000 1000 1000 1000 1000 1000 1000 1000 1000

1000 1000 1000 1000 1000 1000 1000 1000 1000 1000

1000 1000 1000 1000 1000 1000 1000 1000 1000 1000

1000 1000 1000 1000 1000 1000 1000 1000 1000 1000

1. - INTRODUCTION

One cannot make quantitative predictions from perturbative field theory without tackling the problem that finite-order results depend on the choice of renormalization scheme (RS) ¹⁾⁻¹⁰⁾. This is an old problem, but it is only recently - with the advent of next-to-leading order QCD calculations ²⁾, and through the efforts of Celmaster and others ^{3),4),10)} - that its importance has been clearly recognized. Unfortunately, there is still no consensus as to how to deal with the problem. The present paper aims to provide a thorough discussion of the issues involved, and to dispel some myths surrounding the subject. I shall also explain the motivation behind the "Principle of Minimal Sensitivity" or "optimization" approach ^{*) ,6),7)}, and the reasons for believing in it. Some recent criticisms ⁸⁾⁻¹⁰⁾ of this method will be answered.

This is a controversial subject, for it requires a painful re-appraisal of the fundamentals of perturbation theory. I shall be forthright in opposing many popular misconceptions and comfortable old prejudices. These are not specific to any one paper, but stem from the collective unconscious. Where I cite a specific article it is often because it provides a particularly clear expression of a common, and in my view mistaken, viewpoint. This is hardly fair to the authors involved, and I apologize at the outset. A balanced assessment of the literature must be left to future reviewers. It is more pressing to arrive at a resolution of the physics issues.

Perturbation theory involves an expansion in a renormalized coupling constant. The question is: - Which coupling constant? In QCD we are used to thinking of the coupling constant as a function of Q^2 . This idea is so familiar that people have almost forgotten how it comes about. The coupling constant depends on how it is defined - that is, it depends on the way the theory is renormalized. In particular, it depends on the "subtraction point" or "renormalization scale" μ . The idea of "renormalization group improved perturbation theory" ^{11),12)} is that it is advantageous to choose $\mu \approx Q$, where Q^2 is some kinematic invariant that governs the "energy scale" of the process in question. This gives a partial answer to the question of which coupling constant to use. However, (i) it only gives an order-of-magnitude for μ , and (ii) it does not say how the other details of the RS should be chosen. These are the issues to be addressed here.

*) It is not presupposed that the reader is familiar with Ref. 6). However, none of the mathematical details will be repeated here. Brief accounts of the method can be found in Refs. 7), 8).

2. - THE GREAT QED MYTH

Perhaps the first question to ask about the RS dependence problem is:
- Why did we not have this problem before, in QED? Most papers answer this as follows: "...there is a unique, physically motivated, choice for the QED coupling constant"; or "...we have a natural choice of renormalization in QED - the physical or on-shell scheme"; and they then conclude that: "Ideally, one would like to have a 'natural definition' of the QCD coupling constant comparable to the one we have for the fine-structure constant in QED". This attitude is wholly mistaken. In fact, the answer to the question: "Why did we not have this problem in QED?" is that we did: the problem has simply been ignored.

Consider electron-electron scattering at 300 GeV centre-of-mass energy. To use the canonical RS to predict the cross-section would not be sensible. A much better approximation is obtained in an RS in which the renormalization is performed at a scale of order 300 GeV. [Ref. 6] shows how to make this idea precise, but here I am appealing to entirely conventional ideas, dating back to Gell-Mann and Low^{11),12)}.] This means that the canonical RS is not always the best. The fine-structure constant is not the "natural" expansion parameter of QED under all circumstances. Therefore, the traditional QED practice must be wrong.

If this sounds outrageous, then it should not. One cannot espouse the notion of a running coupling constant in QCD, while simultaneously asserting that QED has a "natural", fixed coupling constant. Both theories - as indeed all renormalizable field theories - have a running coupling constant. (They "run" in opposite directions, but that is the only difference.) Fixed coupling constants belong to history.

There is no question that the canonical RS gives the definition of electric charge appropriate to schoolroom electrodynamics; i.e., the long-distance, classical limit. However, as soon as one turns to the quantum theory, and shorter distances, one must work with an effective charge, which can be different under different circumstances. In particular, the effective coupling will depend on the energy scale(s) of the process. It happens that all the celebrated, high-precision tests of QED (anomalous magnetic moments, etc.) involve a low energy scale, set by the electron or muon mass. Since the canonical RS involves the same kind of energy scale, it is not surprising that it turns out to be a good scheme, under these circumstances. (Allowing the coupling constant to run, and "optimizing"

gives only a marginal change, leaving the successful phenomenology intact ^{*)}.) However, the apparent success of the canonical scheme is merely a historical phenomenon, because QED was first tested with e, μ magnetic moments, and not with e^+e^- scattering at 300 GeV.

3. - FREE PARAMETERS AND EXPANSION PARAMETERS

Traditionally, the coupling constant has played a dual rôle as both a free parameter and an expansion parameter. It is very important to distinguish between these two functions.

Even in the exact theory, the predictions do not have the form of definite, arithmetical numbers. However, we can conveniently express the predictions algebraically, as known functions of an unknown, or "free", parameter. How we do this is purely a matter of convention. If we have a set of predictions $R_n(g)$ in terms of some free parameter g , then we can always re-express them in terms of some other free parameter $\lambda = f(g)$. The predictions $R'_n(\lambda)$, where the functions R'_n are given by $R_n \circ f^{-1}$, will look very different, but they express exactly the same physics. The convention adopted truly does not matter. (And it is quite irrelevant whether the free parameter has itself any physical significance.) A free parameter serves only as a book-keeping device which permits a consistent comparison of different physical quantities. Naturally, it must be fixed universally, and not varied from one process to another.

The status of an expansion parameter is quite different. When we approximate the predictions of the theory by truncating a series expansion, the results will obviously depend on the expansion parameter used. Thus, unlike the free parameter, the expansion parameter is a choice which matters: it deserves careful consideration, and should not be made arbitrarily. (Measurability or physical significance is, in itself, no indication that a quantity will make a good expansion parameter.) Moreover, it is entirely possible that a sensible choice of expansion parameter in one case may not be sensible in another. This is a vital consideration; particularly so when there is a large set of possible expansion parameters - the renormalized coupling constants of different RS's - which are all a priori on an equal footing.

*) See Section VIB of Ref. 6). A caveat should be made concerning mass renormalization, but this does not materially affect the discussion here.

I have explained the distinction between a free parameter and an expansion parameter in order to make the following point. It is not necessary to employ a universal definition of the strong coupling constant to make consistent comparisons of QCD effects in different processes [contrary to remarks in, for example, Ref. 13)]. One can use different expansion parameters as long as each can be related back to a universal free parameter. In fact, in the early QCD literature, e.g., Ref. 12), one had the running coupling constant as the expansion parameter, varying from case to case, while Λ served as the free parameter. This picture is not wholly satisfactory - because of the RS ambiguity - but it is basically correct. It is unfortunate that it has been obscured by loose talk of different Λ 's for different processes, when what is meant is different μ 's. (Λ is a boundary condition on the β function equation, and has no business varying from process to process.)

[In "optimized perturbation theory" ⁶⁾ the situation is clear. The "optimum" expansion parameter is different in different cases, and there is a definite algorithm for identifying it. The resulting predictions are expressible in terms of a single free parameter $\tilde{\Lambda}_{\text{ref}}$, the $\tilde{\Lambda}$ parameter of some fixed reference scheme; $\overline{\text{MS}}$, say. One is, of course, at liberty to re-express the results in terms of some other free parameter. Provided no further approximations are made, this has no effect on the predictions.]

In QED, the fine-structure constant α_{fs} is an obvious choice for the rôle of the free parameter. However, the question of whether it is a good expansion parameter is quite another matter. At low energies α_{fs} does well in this rôle: at high energies it fails badly. Thus, the existence of a "natural" RS is no miracle cure for the RS-ambiguity in QED. The observability of α_{fs} , its classical significance, and "low-energy theorems" are all entirely beside the point. Likewise, the RS-dependence problem in QCD is not due to infra-red difficulties, or to the inability to renormalize on-shell. The problem is common to all renormalizable field theories, irrespective of their dynamical properties.

4. - THE MYTH OF A UNIVERSAL CONVENTION

Now, of course, no one would dream of using a fixed coupling constant in QCD: the idea of a Q^2 dependent coupling constant is far too popular. However, what is often proposed (or tacitly assumed) is something nearly as reactionary - namely that only the renormalization scale μ can be process dependent, while all other aspects of the RS should be kept fixed. This common idea is most explicitly stated in the "Good Universal Convention" (GUC) conjecture of Ref. 10). I believe that this is a reflection of a popular, but too narrow, view of the renormalization group.

Consider, for example, an RS based on subtracting the three-gluon vertex $\Gamma_{3g}(p,q,r)$ at the external momentum values $p^2 = -3\mu^2$, $q^2 = -5\mu^2$, $r^2 = -17\mu^2$. Finite order results obtained with this scheme will depend on (i) the choice of the value of μ ; (ii) the choice of the numbers 3, 5, 17; (iii) the choice of this particular vertex function; (iv) the gauge choice; (v) the tensor decomposition of the vertex used; (vi) etc. However, the renormalization group (RG)¹⁴⁾ tells us that the exact all-orders results for physical quantities will be independent, not only of the choice of μ , but of all the choices (i)-(vi) above. The RG is not just about scale dependence: there is much more to it than that. The greater generality of the RG has tended to be overlooked, and has been kept alive largely through the efforts of Peterman^{15),*)}.

The problem has been that the way of describing RS-dependence used above is impossibly cumbersome to handle. However, the technical problem of how to label RS's in a convenient and complete manner has now been solved⁶⁾. One can label an RS by the scale μ (or, more correctly, its ratio with the $\tilde{\Lambda}$ parameter of the scheme), and by the coefficients of the β function. Of course, there are other equivalent choices for an RS "co-ordinate system". However, this parametrization is convenient because only a finite number of the parameters are relevant in finite order, and because no one RS is singled out as special^{**)}. Using this parametrization one can write down an infinite set of RG equations, the first of these being the familiar RG equation, expressing the μ independence of physical

*) An attempt to utilize the full RG by studying more general definitions of the three-point function was made in 1976 by R.G. Moorhouse, M.R. Pennington, and G.G. Ross (unpublished), and is referred to in 16).

**) Neither of these properties holds if one labels RS's by relating them all to one fixed "base" RS. It is all too easy, with such a parametrization, to end up with results that depend on the wholly arbitrary choice of base scheme.
[see also footnote on p. 18]

quantities. The additional equations express the independence of physical quantities from all the other aspects of the RS. These equations are simply an explicit realization of the original equations of Stueckelberg and Peterman¹⁴⁾. These equations imply constraints, or "self-consistency conditions", on the perturbation series coefficients, whose RS-dependence is thereby determined. One can re-express this information by saying that certain specific combinations of series coefficients and RS-parameters are RS-invariants; there being one such invariant ρ_i for each series coefficient r_i ^{*)}. The above results from Ref. 6) are proved facts. They are relevant to any discussion of RS-dependence, whether or not one believes in the "Principle of Minimal Sensitivity".

The purpose of re-iterating these results here is to emphasize that "scale dependence" and "convention dependence" are inextricably bound up together in a unified RG framework. It is wholly artificial, and often misleading, to discuss them separately. They are just two aspects of the same problem: RS-dependence.

It may be helpful at this point to make the terminology more precise. A "renormalization scheme" (RS), as I use the term, means a complete algorithm for removing the ultra-violet divergences of Feynman diagrams, including a specification of the renormalization scale μ . [NB., in gauge theories a gauge choice forms part of the RS algorithm. One can consider gauge dependence as another aspect of RS dependence. See Appendix B of Ref. 6).] I shall adopt Celmaster's term "convention"¹⁰⁾ to refer to a class of RS's differing only in the choice of μ ; i.e., they share the same $\tilde{A}$ and the same β function coefficients. For example, "MS-with- $\mu=Q$ ", referring to $R_{e^+e^-}$ at a total energy Q , is a RS, while "MS" is only a "convention". [As usual, MS and $\overline{\text{MS}}$ refer to "minimal subtraction" and "modified minimal subtraction", respectively²⁾.] The artificiality of separating scale dependence from convention dependence is illustrated by the fact that $\overline{\text{MS}}$ -with- $\mu=Q$ and MS-with- $\mu=Q/2.66$ are the same RS at all orders [if $\overline{\text{MS}}$ is generalized to higher orders in the most sensible way^{18),10)}]. At next-to-leading order one can make any two conventions equivalent by a shift of μ .

Having said all this, I can now summarize the case against the "Good Universal Convention" (GUC) conjecture.

(i) It is not plausible. Since scale- and convention-dependence are on an equal footing with respect to the full RG, it is highly unnatural to assume that only μ can vary from process to process, while the other RS-parameters have fixed, universal

*) Since the publication of 6), I happened upon a 1952 paper by A. Thellung¹⁷⁾, in which this result was first obtained. Thellung's invariants are not defined in terms of β function coefficients (the β function being unheard of at that time), and are formal, divergent expressions, but they are essentially equivalent to the ρ_i of 6).

values. We have abandoned the notion of a fixed scheme in favour of a running coupling constant, on the basis of RG arguments, and this proved to be a very fruitful advance. The full RG urges us to go further: the whole of the RS can - and the implication is, should - depend on which physical quantity one is considering.

(ii) It is not desirable. The schemes which have been proposed as candidates for a GUC are remarkably cumbersome and ugly. All sorts of formidable technicalities are encountered ^{3),4),18)-20)}. It is hard to see this as our "best hope" ¹⁰⁾. If the theory were trying to tell us that this is the wrong approach, it could hardly make its feelings plainer.

(iii) It would not solve the problem, anyway. One would still be left with the problem of how to choose the scale μ for each individual process. It is totally insufficient to say only that μ should be chosen to reflect "the 'typical' energy scale of the process". The ambiguity in that phrase is a major part of the RS-dependence problem. Both formally and numerically, the difference between using $\mu = Q$ and $\mu = \frac{1}{2}Q$ is at least as important as the difference between $\overline{MS}$ and momentum subtraction. Not to give a precise means of choosing μ is to evade the real issues in the RS-dependence problem.

5. - CAN THE PROBLEM BE EVADED?

It has been suggested that one can fix the RS ambiguity by appealing to the data ²¹⁾. This would indeed be a sensible course if it were true that the scheme had to be universal, and could not be process-dependent. However, as I have been arguing, there is no reason to assume this. And that, unfortunately, is that. Obviously, if the RS can be different for each physical quantity, and is to be fitted to the data, one has more free parameters than measurements. One would then have little - if any - predictive power. [I contend that the analysis of Ref. 22) is faulty, and that in fact one could fit any data whatsoever. However, I do not think it is worthwhile to discuss this point further.]

What about the idea of "direct comparison of physical quantities", whereby one extracts a "physical" coupling constant from process B and uses it as the expansion parameter in process A? [e.g., Ref. 23)]. It has been pointed out elsewhere ¹⁾ that this does not work because it fixes only the "convention", leaving the scale ambiguity unresolved: there is no automatic identification between the " Q^2 " of one process and the " Q^2 " of another. Moreover, - even ignoring the distinction between an RS and a convention - there is a basic problem. Why should

one take the "physical" coupling constant from process B, and not from processes C, D, E,...? That is, one is still faced with an infinite number of different predictions for process A. This is just the RS-dependence problem all over again: nothing has been resolved. In defining a "physical" coupling constant from process B - such that there are no radiative corrections to process B - one is merely defining another RS. There is no a priori reason to suppose that this will be a good RS for process A.

There is a subtle point here. A coupling constant defined in terms of a physical quantity B is, in that sense, RS-invariant. However, that is quite irrelevant in any other context. As far as process A is concerned, the "physical" coupling constants of processes B, C, D,... are merely an infinite subset of possible RS's, and have no special merit. To define a coupling constant is to define an RS, and vice versa. Thus, a "scheme-independent coupling constant" ²⁰⁾ is a contradiction in terms. It is like speaking of a "gauge-independent gauge parameter".

6. - A RADICAL RE-THINK

At this point we have arrived at a rather radical conclusion: the assumption of a universal scheme or convention is unjustified, and probably false. Moreover, all the methods based on this assumption fail to meet the real issues of the RS-dependence problem.

If there is no over-all "best" RS, we must seek instead a sensible criterion for identifying the "best" RS in some given context. This radical viewpoint comes together with a radical simplification of the problem: one can do away with all the field-theoretic complications. In the words of Ref. 4): "The salient features of prescription dependence have nothing to do with the complexities of QCD". After all, QCD itself does not have the problem, since its exact results are RS-independent. The problem arises because our approximation scheme mangles this beautiful property. What we have is an example of a "non-invariant approximation". The exact result is independent of certain variables, but the approximate results do depend on these "unphysical" variables.

This is the core of the problem. What we have is a conflict between two pieces of information: the RG tells us that the answer does not depend on the RS - while our calculated result does depend on the RS. It would be foolish to make predictions without attempting to reconcile this conflict. It cannot make

sense for our prediction to be highly sensitive to variables on which it is not supposed to depend at all. This suggests that - in the absence of any further information to guide us - the most sensible course is to make our approximant as insensitive as possible to small variations in the unphysical variables. It is this piece of commonsense that the "Principle of Minimal Sensitivity" (PMS) ⁶⁾ embodies.

The point is that, although our calculated result is not invariant with respect to changes in RS, we can at least seek a region in which it is locally constant. This is where the contradiction between our two sources of information will be least severe. It is the region in which the calculation is most in agreement with itself. It is also the region in which the approximate result is behaving most like the exact result, which is known to be a constant in RS-space.

The PMS criterion goes to the heart of the problem, in a way that no other approach does. No other method makes use of the fact that the exact result is known to be RS-independent. Without this fact there would be no problem - or, at least, the problem would be totally different. It seems logical, therefore, to expect this fact to play a key rôle in the solution. That is indeed the case if we say that there is an "optimum" RS such that the result is minimally sensitive to small changes in RS.

7. - IS THERE A THEOREM?

Of course, there is a tremendous difficulty here with language. If the PMS criterion is claimed to give the "optimum" result, then what exactly does "optimum" mean? Clearly, an ideal criterion would always minimize the error in the approximation. Is there a theorem that the PMS criterion is best in this incontrovertible sense? It is important to understand the answer to this question. The answer is: "No, - but that is the wrong question". An ideal criterion is simply not within our reach.

Definition If $\mathcal{R}^{(n)}(\text{RS})$ is an approximation to the exact result $\mathcal{R}$, then an ideal criterion fixes the RS such that the error $|\mathcal{R} - \mathcal{R}^{(n)}(\text{RS})|$ is a minimum.

Theorem In circumstances where the exact result $\mathcal{R}$ is unknown, an ideal criterion cannot, in general, exist.

Proof In general, there will be one or more RS's in which $\mathcal{R}^{(n)}(\text{RS})$ happens to take on the same numerical value as $\mathcal{R}$. [For example, in third order, $\mathcal{R}^{(3)}(\text{RS})$ takes on all values between $-\infty$ and $+\infty$, and therefore must coincide with $\mathcal{R}$ at some point(s).] The minimum of the error is then zero. Thus, an ideal criterion must, in general, produce the exact result $\mathcal{R}$. But this contradicts the condition that $\mathcal{R}$ is unknown.

[In special cases $\mathcal{R}^{(n)}(\text{RS})$ may have a finite upper (lower) bound, and $\mathcal{R}$ may lie above (below) this bound. In such cases an ideal criterion does exist: it picks out the maximum (minimum) of $\mathcal{R}^{(n)}(\text{RS})$. This is in accord with the PMS criterion. A classic example of this special case is the Rayleigh-Ritz theorem; see Section II of 6).]

The above theorem says that there are no theorems. There cannot be a rigorous solution of the RS-dependence problem - unless it be an exact solution of QCD. Asking for rigour, then, is asking the wrong question. We must ask instead for a solution that meets the highest standards short of rigour. The definition of "best" or "optimum" in this context must be something like "most sensible, in the absence of further information". The key point is our limited amount of information, and the difficulty in formulating a more satisfactory definition lies in expressing this absence-of-information in mathematical terms.

Ignorance of the exact result (and the non-existence of any rigorous bounds on the error) means that any approximation is a gamble. There can be no certainty. We can only ask whether our bet is sensible or not. It cannot be wise to gamble when there is an unresolved conflict in our available information. It cannot be sensible to bet on one particular RS if a small change in the scheme - which is not supposed to matter at all - makes a big difference to the result. The argument is that, in the absence of further information to guide us, the most sensible gamble is to believe the calculated result where it is least sensitive to small variations in RS. There is no theorem - artificial counter-examples can be constructed (though none have been exhibited) -; nor is there even a satisfactory definition of "most sensible"; but we are talking about gambling, not mathematics.

In using approximations, physicists are making gambles all the time, though they are not always so explicit about it. The argument for the minimal sensitivity criterion cannot be absolutely compelling - otherwise it would prove a theorem that cannot exist. However, the argument is very simple, general, and abstract. It applies to any kind of "non-invariant approximation", whatever its origin. Thus, there are several opportunities to put the idea to the test. (I shall have more to say about this in Section 10).

8. - EXTRA INFORMATION?

Any rational resolution of the RS ambiguity must be based on some information not already contained in the results of the diagram calculations. Therefore, it is very important to consider possible sources of extra information. In this section I shall argue two points: (i) that the PMS criterion does use such extra information, and (ii) that, with one exception, we have no other sources of information worth speaking of.

The PMS criterion relies on the fact that physical quantities - unlike, say, Green's functions - are independent of RS. That is, it uses the RG-invariance property of physical quantities, expressed by the RG-equations

$$\partial \mathcal{R} / \partial (RS) = 0. \quad (1)$$

[Here $\mathcal{R}$ is a physical quantity with the series expansion $\mathcal{R} = a^N (1 + r_1 a + \dots)$ in the couplant $a \equiv \alpha_s / \pi$. The above symbolic equation can be translated straightforwardly into mathematical terms⁶⁾. It is implicit that the physical variables (kinematic invariants, etc.) are held at fixed values.]

The information in Eq. (1) is not already contained in the n^{th} order diagram calculations, which can at most tell us that

$$\partial \mathcal{R}^{(n)} / \partial (RS) = O(\alpha^{N+n}). \quad (2)$$

There is a huge difference between Eqs. (1) and (2). The first is a numerical statement - something equals zero -, while the second is a formal statement about asymptotic behaviour in a hypothetical limit, $a \rightarrow 0$. Equation (1) implies (2), but not vice versa; the point being that we have Eq. (2) only for a finite number of orders. Equation (1) expresses a property of the unknown, exact result, and cannot be deduced from the diagrammatic calculation of $\mathcal{R}^{(n)}$. It is therefore an extra piece of information.

An easy way to see that the PMS criterion utilizes some extra information is to consider the "QID" example of Celmaster and Sivers⁴⁾. Their little fable does illustrate several features of the RS-dependence problem - but it leaves out the most important point. It is not said that the exact results in "QID" are independent of the choice of scheme. It is not explained in what sense the different expansion parameters are formally equivalent. That is, the example, as presented, is not endowed with a renormalization group. One cannot even begin to apply the PMS criterion, because the piece of information it needs is just not there.

Do we have any other useful information besides the RG? Perhaps our accumulated experience with, and insights into, Feynman diagrams constitutes some additional "information"? ¹⁰⁾. The argument goes ^{3),4),18),19)} that momentum subtraction schemes - being defined so that, at a scale μ , the radiative corrections to some vertex are absorbed into the coupling constant - will tend to suppress the contributions of higher-order diagrams, if μ is chosen to reflect the "typical" loop momenta involved. This argument is not wrong - but it is distressingly vague. One can use this argument to understand intuitively why the PMS criterion picks out the scale $\bar{\mu}$ that it does ^{1),6)-8)}. However, I do not believe our Feynman diagram intuition is precise enough to definitely contradict, or improve upon, the minimal-sensitivity result. [Even if we could understand the leading behaviour of each diagram, there is still the mysterious phenomenon of near-cancellations between different diagrams. (See, for example, Fig. 1 of 24).] The RG is the only handle we have on such cancellations.]

The above remarks should be qualified in the special case of leading order. Since one has not then had to renormalize explicitly, the RS-dependence arises only from the scale μ/Λ , at which one chooses to evaluate the leading-order running coupling constant. The PMS criterion is then no help because the approximant is monotonic in the RS variable. One can only guess at the appropriate value of μ/Λ , and so momentum subtraction ideas are very relevant for doing good phenomenology with only leading-order calculations.

However, when higher-order calculations exist we should not be satisfied with such vagueness - otherwise our claims to predictive power become suspect. Fixing the RS by intuition alone reduces us to special pleading in each and every separate case. It may seem cynical to say so, but intuition has a habit of changing so as to give a desired result, or to accommodate new experimental data. Advocates of a mixed repertoire of scheme-fixing devices [e.g., the conclusions of 10)], are even more open to this criticism. Of course, one can argue that the danger is not serious so long as the prediction is insensitive to changes from one "reasonable" scheme to another. However, to argue thus is to concede the essential validity of the minimal sensitivity approach.

There is one case where we may have some extra information. This is the issue of π^2 terms in, say, the Drell-Yan process, whose origin is in the cancellation of infra-red divergences ²⁵⁾. Our understanding is incomplete at present, but it seems likely that these π^2 terms will exponentiate. Since this is quite independent of the RS question, it is perfectly consistent and sensible to use this "information" in conjunction with the PMS criterion. [This is discussed in Section VID of 6).] One must be clear about the motivation, however: removing

the π^2 's because one believes that they exponentiate is honourable, but doing so just to "make the next-to-leading-order coefficient look smaller" is shabby - and silly, because the size of the coefficient is an RS-dependent matter. (In some schemes, removing the π^2 terms will increase the magnitude of the coefficient !)

The question of the size of next-to-leading coefficients is to be discussed in the next section.

9. - APPARENT CONVERGENCE

Most discussions of the RS-dependence problem have centered on the following observation. If the next-to-leading-order correction term is as large as, or larger than, the leading term, then the result is not likely to be reliable. Therefore, it is argued, one should only trust the result in an RS in which the calculated coefficient is small. (Of course, the real issue is whether the uncalculated coefficients are small. Intuitively - but not rigorously - this is correlated with the size of the known coefficients. To avoid any confusion, I shall use the term "K-coefficients" for the known coefficients, as distinct from the unknown, higher-order coefficients.)

The argument that small K-coefficients are, in some sense, desirable, does have some validity, and it has an important rôle to play. However, the idea that one should fix the RS so as to make the K-coefficients small is not a sound basis for resolving the RS-ambiguity.

To begin with, there is no necessity to use the same RS in all cases, as explained before. Therefore, one does not have to compromise by finding a scheme in which the K-coefficients of many different processes are simultaneously small. For each physical quantity one is free to make the K-coefficients as small as one likes. This already seems suspicious. However, if one really believes that small K-coefficients are a good thing - and the smaller, the better -, then one is driven to the "fastest apparent convergence" (FAC) criterion. This says: fix the RS by requiring all the K-coefficients to vanish *).

*) This method has been proposed by Grunberg⁵⁾, and it is similar in spirit to the " Λ_N " scheme²⁾. [At least, this is how I understand Grunberg's proposal: he would not completely agree with my paraphrase here. The interested reader is therefore referred to Refs. 5), 26).] The term FAC originated independently in an early version of 6), which criticized this kind of approach. The formalism of 6) makes it easy to see how one could implement FAC at any order²⁷⁾. FAC has nothing to do with the hypothesis of a "good universal convention", contrary to remarks in 10).

I believe that the FAC criterion confuses two quite separate issues in the RS-dependence problem. The primary problem is to decide what perturbative approximations really mean if they depend on a supposedly arbitrary choice of RS. The secondary problem is to decide how to judge the reliability of the results.

The secondary problem arises even for ordinary approximations, which do not involve unphysical parameters. There is a well-established way of dealing with the problem: one looks at the rate at which the series appears to converge, as a guide to the likely error involved in omitting the higher-order terms. That is, one forms an error estimate based on the apparent convergence of the series ^{*)}.

With a non-invariant approximation, one also has to contend with the primary problem. The approximation must be given a meaning before a sensible error estimate can be made. Fixing the RS by demanding good apparent convergence is putting the cart before the horse. It artificially distorts a method of error estimation into a method for choosing the RS. Moreover, it seriously prejudices the question of how one can judge the probable error. It cannot make sense to say, "the result in this RS should be reliable, because the next-to-leading correction is small", if we have adjusted the RS specifically to achieve this!

The situation with the PMS criterion is quite different. It resolves the primary problem - specific to non-invariant approximations - by means of their distinctive feature that the exact result is known to be independent of certain variables. Non-invariant approximations are thereby reduced to the level of ordinary approximations. Moreover, this is done in a way which is wholly independent of the notion of apparent convergence: the PMS criterion does not even care that the approximation is obtained from a series. Therefore it is perfectly legitimate for one to use, subsequently, the traditional method of error estimation, based on apparent convergence. [Note that what matters is the relative size of the correction (coefficient $\times$ couplant), not just the size of the coefficient.]

It has been pointed out many times ^{7),9),27)-29)} that, in practical QCD applications, the PMS criterion often gives results rather similar to those of the FAC criterion. In fact, the PMS criterion applied to field theory corresponds,

^{*)} Note that there is no rigour here. The apparent convergence could be a false guide to what the series will actually do in higher orders. A statement such as "the error in the series $1+2x+x^2+\frac{5}{6}x^3+x^4+O(x^5)$, at $x = 0.1$, is probably around 10^{-5} ", cannot be proved, nor even properly formulated. Counter-examples are easy. Nevertheless, for centuries physicists have accepted such statements as, in some sense, true. (One should bear this in mind when considering the arguments of Section 7.)

roughly, to a particular "proportionality" between the series coefficients and the β function coefficients, as emphasized by Pennington ⁹⁾ [see also 27), 30)]. In particular, the next-to-leading order coefficient is governed by the ratio of the first two β function coefficients, and always has a magnitude around, or somewhat less than, unity. Pennington has argued that therefore the PMS criterion is a FAC-type method in disguise, and is no more justified than approaches based on enforcing good apparent convergence ⁹⁾. That is not correct, because it confuses effects with justifications. Merely because two arguments lead to similar results, one cannot conclude that they have equal validity. The arguments for the PMS criterion are quite different from those of FAC. Furthermore, there are examples (outside field theory) where the two criteria give very different results ⁶⁾. Nothing in the minimal sensitivity argument guarantees that the results will show good apparent convergence: if they do in practice, then it must be due to the properties of the approximation scheme (perturbation theory), and of the theory (QCD), in question.

If the "optimized" coefficients had turned out to be large, I should have worried, not about the soundness of the PMS criterion, but about the health of QCD perturbation theory. As it is, it seems that the great scare over large coefficients, e.g., 31), was not warranted ¹⁾. The "intrinsic" size of next-to-leading coefficients is, in fact, of order one. The historical problem was that the RS's used (i.e., the combinations of convention and scale choices), corresponded to unrealistically small effective coupling constants for the processes considered. The theory retaliated by giving a large next-to-leading order coefficient. Thus, the initial over-optimism about the size of the effective coupling constant was the direct cause of the later over-pessimism about the size of the coefficients. [This general conclusion seems to be widely accepted ^{1), 4), 23)}.]

I consider the appearance of small K -coefficients in the "optimization" method to be good news for QCD perturbation theory, - and it is certainly no grounds for disparaging the PMS criterion. Equally, however, it is important to understand that small K -coefficients are not claimed as a "success" of the "optimization" method. If this is a success, then the credit belongs to QCD perturbation theory: if properly interpreted, and within its known limitations, it shows every sign of being a well-behaved, usable approximation scheme.

10. - ADMISSIBLE EVIDENCE

As mentioned above, small K -coefficients are not evidence of the success of any method for resolving the RS ambiguity. The so-called "successes" of momentum subtraction¹⁰⁾ are all of this inadmissible kind. The only real measure of success is whether one's RS-fixing method gives a reliable and accurate estimate of the exact result. This, of course, cannot be demonstrated in QCD - because the exact results are unknown. (We do not even have any third-order results for physical quantities, which might indicate which interpretation of second order was best.) Consequently, the only way to produce admissible evidence is to study analogous examples, in which the exact results are known. The PMS criterion has been tested in a number of examples, drawn from pure mathematics, quantum mechanics, and field theory^{6),32),*)}, and it has an impressive record. Of course, one can debate the significance of such evidence, but the point is that it is the only kind of evidence to be had.

Is it fair to use examples far removed from field theory? Since "the salient features of prescription dependence have nothing to do with the complexities of QCD"⁴⁾, I believe it is perfectly legitimate to do so. The minimal sensitivity argument is very abstract, and is claimed to be valid for any kind of non-invariant approximation. If the argument were basically unsound, then one would expect it to fail badly in the first few cases. In fact, the PMS criterion works very well in the cases studied.

Moreover, a general pattern emerges, which reinforces the idea that a "non-invariant approximation" is a useful, abstract concept. Typically, one finds the lowest-order approximation (and all its derivatives) to be monotonic in the unphysical variable, and therefore irredeemably ambiguous. In next order one finds a single maximum or a single minimum^{**)}. In the order beyond one finds either a maximum and a minimum close together, or a point of inflexion. Qualitatively, the sequence of shapes can be described as "linear", "quadratic", "cubic", etc.

*) Although it is not remarked upon in that paper, Fig. 11 in Appendix A of 4) is another illustration of the correlation between accuracy and insensitivity to unphysical variables.

**) Since the PMS criterion suggests that the exact result should be close to the stationary point, it is no surprise to find that it often lies just above the maximum, or below the minimum. To say that the PMS criterion works, in the examples studied, only because such "Rayleigh-Ritz"-type inequalities are lurking nearby¹⁰⁾, is to miss the point: the PMS criterion predicts that such inequalities are common - and that, even when not exact, they still hold in an approximate sense.

In field theory the situation is a little more complicated, since each new order introduces an extra unphysical variable. However, one can see a "linear", "quadratic",... pattern for each unphysical variable. (In a way, this multitude of variables is an advantage, since there is then only one stationary point, and hence no residual ambiguity in the PMS result.) These points are discussed in more detail in 6).

In all the examples it is essential to allow the unphysical variables to take on different values in different circumstances. For instance, in the anharmonic oscillator case one could not fit more than one or two energy levels with a fixed value of Z [see Eqs. 2.13-2.15 of 6)]. This point can be made very strongly: all previous experience of non-invariant approximations in theoretical physics suggests that the worst possible mistake is to assume that the unphysical variables should have fixed, universal values 33),6),32). That is why I am so sceptical of fixed coupling constants, or universal conventions.

The FAC criterion does not do too badly in the examples, mainly because it does allow the unphysical variables to change with the circumstances. In many cases it gives results similar to the PMS results. However, there are several cases where FAC fails rather badly. There is nothing particularly artificial about these examples: the approximation schemes involved are quite healthy, and the PMS results are accurate and well-convergent. The generic reason for these failures seems to be the phenomenon illustrated in the Figure. Quite often - for reasons associated with the limitations of the approximation method - the curves in adjacent orders tend to be parallel as they curve away from the flat region. Then, although the two approximations are close where they are flat, they intersect only at a very different height. The apparent convergence is good in the flat region, but it is fastest in a quite different place. Empirically 6),32), it seems that when PMS and FAC differ significantly in this way, it is FAC which fails. This phenomenon seems not to occur in field theory, where PMS and FAC generally give similar results: nevertheless, its existence is evidence that the motivation behind FAC is less sound.

In judging the reliability of the approximate result, it is not really relevant how well PMS and FAC agree. [I say this as a caveat to 9), 27).] In cases resembling the Figure one would grossly overestimate the error: in other examples one would be misled into underestimation. A better motivated, and empirically more reliable, method is to estimate the error from the apparent convergence rate of the PMS results.

11. - TRUNCATION OF THE β FUNCTION

In a recent paper ⁸⁾, Politzer has discussed the application of the "optimization" method to factorization and mass scheme dependence. He also raises an objection to the method, which I now want to answer. The complaint is that in second order, for example, only two terms of the β function

$$\mu \frac{\partial a}{\partial \mu} \equiv \beta(a) = -b a^2 (1 + c_1 a + c_2 a^2 + c_3 a^3 + \dots) \quad (3)$$

are kept, while the higher-order coefficients $c_2, c_3, \dots$ are arbitrarily set to zero. This criticism has been repeated, in strong terms, by Celmaster ¹⁰⁾. However, that clearly invites the retort: what does one do with $c_2, c_3, \dots$ at second order in momentum subtraction? A glance at Refs. 3), 4) shows that they are, of course, set to zero.

This is not merely a debating point, but is the crux of the matter. For an approximation to be useful, it obviously must be constructed out of known quantities. Thus, any conceivable method can only utilize a finite number of the coefficients of $\mathcal{R}$ (the physical quantity in question), and a finite number of β function coefficients - since this is all we ever know. Therefore, we traditionally define our second-order approximation by truncating the series

$$\mathcal{R} = a^N (1 + r_1 a + \dots) \quad (4)$$

setting $r_2 = r_3 = \dots = 0$, and by replacing the couplant a ($\equiv \alpha_s/\pi$) with its second-order approximation $a^{(2)}$. The approximate couplant $a^{(2)}$ satisfies a truncated form of the β function equation:

$$\mu \frac{\partial a^{(2)}}{\partial \mu} = -b (a^{(2)})^2 (1 + c_1 a^{(2)}). \quad (5)$$

The setting to zero of $c_2, c_3, \dots$ is on exactly the same footing as the zeroizing of $r_2, r_3, \dots$: it is a matter of defining - within a general, fixed RS - what the perturbative approximant is. Only after the approximant has been defined can one ask about its sensitivity to changes in RS ^{*)}. Thus the problem of monotonicity in

*) When investigating the behaviour of perturbative approximants as a function of RS, one must remain consistent with one's own definitions. The perturbative approximant of each scheme is defined by truncating $\mathcal{R}$ and β in that scheme. This implies that the approximant in one scheme when re-expanded in powers of the coupling constant of another scheme is a known, well-defined, infinite series. To casually re-truncate this series at will is to render the whole exercise rather pointless, since one is not then comparing the results of different schemes at all. This criticism applies, for example, to the analysis leading to Fig. 1 of 9). (The second stationary point of curve II, for instance, is entirely spurious.) To be fair, the version of 9) to be published does make clear the unsatisfactory and approximate nature of the analysis: also 9) is only interested in comparing schemes which are relatively close to the "base" scheme.

$c_2, c_3, \dots$, encountered in 8), does not arise: the approximant is constructed to be independent of $c_2, c_3, \dots$ - and it has to be, because these coefficients are unknown quantities.

The setting to zero of $c_2, c_3, \dots$ and $r_2, r_3, \dots$ is a matter of the definition of perturbation theory. It is hard to understand why this definition should be controversial, since it has been universally adopted in the past ^{2)-4),*)}. [The analogous procedure at leading order, of keeping only the leading term of $\beta(a)$, is also standard practice; e.g., 34), 12).] It is ironic that the "optimization" method should be criticized on the point where it is most conventional.

Of course, since truncating $\mathcal{R}$ and β is a matter of definition, it can be said to be arbitrary. However, if one regards this as a problem, then it is a problem common to all RS-fixing methods. It would have arisen even if our approximants had not been dependent on RS, and so is quite separate from the RS-dependence problem.

The only remaining point to discuss is why one should truncate the β function at the same order as $\mathcal{R}$. One might consider approximants $\mathcal{R}^{(i,j)}$; keeping i terms of $\mathcal{R}$, and j terms of β ³⁵⁾. For example, at present we actually know three terms of β in the MS scheme ³⁶⁾, while we have only two terms of any physical quantity. We could therefore construct an $\mathcal{R}^{(2,3)}$ approximation. Would we expect greater accuracy thereby? Would we be using more information?

The answer to both these questions is: no. Firstly, a little analysis shows that the error in an $\mathcal{R}^{(i,j)}$ approximation is $O(a^{N+k})$, where $k = \min(i,j)$; i.e., the expected error is governed by whichever truncation is the most drastic. It is natural, therefore, to concentrate on the "diagonal" ($i = j$) approximants, as giving the greatest accuracy for the least work. Secondly, the quantity of physical information we have is determined by the number of RS-invariants $\rho_1, \rho_2, \dots$ ^{17),6)} that are known. This is also governed by $\min(i,j)$. A knowledge of c_n without knowledge of r_n , or vice versa, means that we have only some scheme-dependent information - that is, information about the scheme, not about the physics. [Thus, the calculation in 36), while useless on its own, is a vital pre-requisite for any third-order prediction.]

*) In fact, earlier authors often made further approximations by expressing the solution to (5) as a truncated series in $1/\ln(\mu/\Lambda)$. This is unnecessary, and causes further difficulties, and has been criticized elsewhere ^{4),6),7)}.

I conclude that there is nothing to be gained by considering "non-diagonal" ($i \neq j$) approximations: indeed, it may be inconsistent to do so. In the minimal sensitivity philosophy, such approximants have no numerical meaning. An $i \neq j$ approximant is monotonic in at least one of the unphysical variables, so there is no flat region where it can be trusted ^{8),35)}.

12. - PROBLEMS OF ORDINARY APPROXIMATIONS

The arguments of the last section, in defence of truncating the β function, can be summarized as follows: (i) something of the kind is necessary in order to define a usable approximation, (ii) truncation is the most natural, "minimal" definition, and has always been used previously, (iii) the criticism applies equally to any RS-fixing method, and is not specific to "optimization".

This is not entirely satisfactory, of course, and it would be good to have a more positive answer. The point is, though, that this kind of problem arises even with ordinary approximations, which are free from the presence of unphysical variables. One is dealing with philosophical difficulties that are literally centuries old.

Consider the simplest possible example: the power series expansion of a function

$$f(x) = A + Bx + Cx^2 + Dx^3 + \dots \quad (6)$$

Imagine that we have calculated A, B, C , while $D, E, \dots$ remain unknown to us. Traditionally, one would form "the second-order approximation"

$$f^{(2)}(x) \equiv A + Bx + Cx^2. \quad (7)$$

That is, one sets $D, E, \dots$ to zero. This is just a definition. It can be criticized on the grounds of arbitrariness, since there are many other approximants one could form from the same information: for instance,

$$[AB + (B^2 - AC)x] / [B - Cx], \quad (8)$$

or

$$(1 + 2x)^{1/2} [A + (B - A)x + (C - B + \frac{3}{2}A)x^2], \quad (9)$$

to take just two random examples from an infinite set. All these approximants are consistent with the known terms of $f(x)$, but they represent different guesses as to the values of the unknown coefficients $D, E, \dots$

If we do not have the slightest idea of what $D, E, \dots$ will be, then - appealing to the principle of equal a priori probabilities - they are as likely to be positive as negative. In some sense, then, their "expectation values" are zero. This is about as far as one can go in justifying the traditional choice (7). [Of course, occasionally we do have other scraps of information. For example, if we know that $f(x)$ tends to a constant as $x \rightarrow \infty$, then it may be more sensible to use the Padé approximant (8), which shares this property. Alternatively, if we know that $f(x)$ had a factor $\sqrt{1+2x}$, then the form (9) might be a good choice.]

Any kind of approximation necessarily involves guesswork about higher orders. Forming approximants by truncation, as in (7), is indeed an arbitrary definition. However, in the total absence of other information, it seems the most natural - certainly the most conventional - thing to do. In perturbation theory, two such truncations (in $\mathcal{R}$ and β) are involved, as has long been recognized.

There is a similar point to be made about the "non-linearity" of the "optimization" method. The "optimized" value of a sum or ratio is not the same as the sum or ratio of the terms optimized separately ^{6),8)}. Thus, it is significant what one chooses to calculate. Again, this is a difficulty present even with ordinary approximations. If f, g, h are three functions satisfying $f = g/h$, and $f^{(n)}, g^{(n)}, h^{(n)}$ are their n^{th} order truncated power series, then $f^{(n)} \neq g^{(n)}/h^{(n)}$, in general. [True, one does have $f^{(n)} = g^{(n)} + h^{(n)}$ when $f = g + h$, but even this is a special property of truncated power series, not shared by, say, Padé approximants.] Thus, it is always significant what one chooses to calculate. This problem is closely related to the one discussed above. For example, if one decides to calculate f^2 , instead of f , as a power series, then this is equivalent to using a second-order approximation to f of the form

$$[A^2 + 2ABx + (2AC + B^2)x^2]^{1/2}. \quad (10)$$

In conclusion, there are indeed residual ambiguities in "optimized perturbation theory". However, these ambiguities have nothing to do with scheme dependence, and are inherent in any approximation method. Such difficulties are centuries old, and physicists have learned to live with them. I think it is very

important for these problems to be given serious consideration, but it is hard to give much general guidance beyond saying that one should calculate what seems most natural, and is most directly measured in experiments ^{*)}.

In what sense, then, is the RS-dependence problem solved by the PMS criterion? The answer is that the ambiguity due to the unphysical variables has been dealt with: the original non-invariant approximation has been reduced to the level of an ordinary approximation. Any problems that remain are, at least, ordinary ones.

13. - CONCLUSIONS

It cannot be overstressed that the best RS can be different under different circumstances. This fact is already embodied in the familiar notion of a "running" or "effective" coupling constant. The full renormalization group urges us to go forward with this idea. There is no universal best scheme or convention - not even in QED. All searches for one are doomed.

The essential feature of the RS-dependence problem is that we know that the exact result is independent of RS. Only the minimal sensitivity approach makes any use of this key fact, and attempts to reconcile it with the presence of RS-dependence in the finite-order results.

*) Actually, a virtue of the PMS method is that it does put constraints on the choice of what to optimize. Since one is relying on the RS invariance of the full perturbation series, one can only work with quantities which are (i) physical quantities, and (ii) "purely" perturbative. Thus, in deep-inelastic scattering one must deal with the Q^2 dependence of the moments, $F_n \equiv -(Q/M_n)(dM_n/dQ)$, rather than with the moments themselves 6)-8), 28). Similarly, in quarkonium decays, one cannot directly "optimize" quantities, such as $\Gamma(T \rightarrow \text{hadrons})$, having the form $|\psi(0)|^2 S(\alpha_s)$. Since $|\psi(0)|^2$ is not a physical quantity, the series $S(\alpha_s)$ is not RS invariant by itself. In fact, the calculation of $\Gamma(T \rightarrow \text{hadrons})$ involves re-absorbing the "instantaneous Coulomb exchange" part of the radiative corrections into $|\psi(0)|^2$ 24), 29), 37). This introduces a kind of factorization scheme ambiguity, which is eliminated by considering the ratio $\Gamma(T \rightarrow \text{hadrons})/\Gamma(T \rightarrow e^+e^-)$ 24), 29). The procedure used in 26), whereby FAC is applied to $\Gamma(T \rightarrow \text{hadrons})$ and $\Gamma(T \rightarrow e^+e^-)$ separately, would not be valid for the PMS method. This is another caveat to the view that FAC and PMS are almost equivalent.

The PMS criterion is based on commonsense, not on a mathematical theorem: but rigorous justification of any resolution of the RS ambiguity is out of the question. Lack of rigour is therefore no basis for dismissing a proposal as unsound. In fact, there is considerable evidence from soluble examples that testifies to the soundness of the minimal sensitivity argument.

There are two separate issues in the RS-dependence problem. One is to decide what RS-dependent perturbation theory means, and the other is to ascertain how well it seems to converge in practice. The idea of fixing the RS by requiring the next-to-leading coefficients to be small, or zero, is pre-judging the second question in order to answer the first. It fixes the RS only by assuming that perturbation theory is always intrinsically well-convergent. With the PMS criterion, by contrast, one fixes the RS by independent means. It is then legitimate to judge the reliability of the approximation in the traditional way - by examining the apparent convergence of the series.

I have answered some objections to the "optimization" procedure raised by Politzer⁸⁾; the point being, not that these criticisms are invalid, but that they are not specific to the "optimization" method. In particular, the truncation of the β function is a matter both of definition, and of necessity. It is also standard practice.

The discussion here has concentrated entirely on the theoretical issues. (A good discussion of the practical aspects, and applications, of "optimization" is provided by Duke and Kimel²⁸⁾.) Phenomenologically, there are significant differences between the "optimized" result and the result in a randomly chosen RS. However, one can see that intuitive ideas about which schemes are "reasonable" have usually led in the right direction. In effect several of these vague, intuitive ideas are subsumed by "optimization". In no way is the method belittled by the fact that other methods may often arrive at somewhat similar results, in practice. It is the whole point of minimal sensitivity that there are many nearby schemes which give almost the same result.

The RS-dependence problem is primarily a question of theoretical principle and integrity, and it reveals a basic incompleteness in the conventional formulation of perturbation theory. The idea of "optimization" is to use the (full) renormalization group - in the simplest, and most obvious way -, thereby giving a precise meaning to the concept of an "effective coupling constant". That is, although seemingly radical, "optimization" is really a development of two very old, and very fruitful, ideas.

ACKNOWLEDGEMENTS

In refining the arguments used here, the questions and comments of numerous physicists have been invaluable. I am very grateful to my colleagues at Madison for their moral support. I also wish to thank W. Celmaster, G. Grunberg, M.R. Pennington, and D. Sivers for correspondence, and A.C. Davis, J. Ellis, H.F. Jones, R. Petronzio, and G.M. Shore for comments and advice.

REFERENCES

- 1) C.H. Llewellyn Smith - in "High Energy Physics-1980", Proceedings of the XXth International Conference, Madison, WI; edited by L. Durand and L.G. Pondrom (AIP, New York, 1981), p. 1346.
- 2) A.J. Buras - Revs.Modern Phys. 52 (1980) 199;
A.J. Buras, E.G. Floratos, D.A. Ross and C.T. Sachrajda - Nuclear Phys. B131 (1977) 308;
W.A. Bardeen, A.J. Buras, D.W. Duke and T. Muta - Phys.Rev. D18 (1978) 3998.
- 3) W. Celmaster and R.J. Gonsalves - Phys.Rev. D20 (1979) 1420; Phys.Rev. D21 (1980) 3112.
- 4) W. Celmaster and D. Sivers - Phys.Rev. D23 (1981) 227.
- 5) G. Grunberg - Phys.Letters 95B (1980) 70.
- 6) P.M. Stevenson - Phys.Rev. D23 (1981) 2916.
- 7) P.M. Stevenson - Phys.Letters 100B (1981) 61; in "Perturbative Quantum Chromodynamics, Tallahassee, 1981", edited by D.W. Duke and J.F. Owens (AIP, New York, 1981).
- 8) H.D. Politzer - Nuclear Phys. B194 (1982) 493.
- 9) M.R. Pennington - Durham Preprint (May 1981), to be published in Phys.Rev.D.
- 10) W. Celmaster - Northeastern University Preprint NUB 2531 (Talk given at the 1981 Symposium on Particle Physics in Visegrad, Hungary, and at the Particle Theory Workshop, DESY).
- 11) M. Gell-Mann and F. Low - Phys.Rev. 95 (1954) 1300;
N.N. Bogoliubov and D.V. Shirkov - "Introduction to the Theory of Quantized Fields" (Wiley-Interscience, New York, 1959).
- 12) H.D. Politzer - Physics Reports 14C (1974) 129.
- 13) A. Buras - in "High Energy Physics-1980", op.cit., p. 789.
- 14) E.C.G. Stueckelberg and A. Peterman - Helv.Phys.Acta 26 (1953) 449.
- 15) A. Peterman - Physics Report 53C (1979) 157;
A. Peterman - in "Field Theory, Quantization and Statistical Physics", edited by E. Tirapegui (D. Reidel Publishing Company, Dordrecht, 1981).
- 16) M.R. Pennington and G.G. Ross - Phys.Letters 86B (1979) 372;
H.D. Politzer - Nuclear Phys. B146 (1978) 283.
- 17) A. Thellung - Helv.Phys.Acta 25 (1952) 307.
- 18) E. Braaten and J.P. Leveille - Phys.Rev. D24 (1981) 1369.
- 19) P. Pascual and R. Tarrach - Nuclear Phys. B174 (1980) 123 [E B181 (1981) 546];
R. Coquereaux - Phys.Rev. D23 (1981) 1365.

- 20) D. Espriu and R. Tarrach - Phys.Letters 102B (1981) 163;
D. Espriu, P. Pascual and R. Tarrach - J.Phys. G7 (1981) 1591.
- 21) M.R. Pennington and G.G. Ross - Nuclear Phys. B179 (1981) 324;
L.F. Abbott and R.M. Barnett - Ann.Phys.(N.Y.) 125 (1980) 276.
- 22) M. Haruyama, A. Kanazawa and O. Abe - Progr.Th.Phys. 65 (1981) 2067;
Preprint EPHOU-MCA 81/5 (1981);
M. Haruyama - Preprint MCA 80/0011 (November 1981).
- 23) R. Barbieri, L. Caneschi, G. Curci and E. d'Emilio - Phys.Letters 81B (1979) 207.
- 24) P.B. Mackenzie and G.P. Lepage - Phys.Rev.Letters 47 (1981) 1244.
- 25) G. Parisi - Phys.Letters 90B (1980) 295;
G. Curci and M. Greco - Phys.Letters 92B (1980) 175.
- 26) G. Grunberg - Ecole Polytechnique Preprint A474.1281 (December 1981).
- 27) J. Kubo and S. Sakakibara - Dortmund Preprint DO-TH 81/07 (July 1981).
- 28) D.W. Duke and J.D. Kimel - Florida State Preprint FSU-HEP-811125 (December 1981).
- 29) R. Gupta; and P.B. Mackenzie and G.P. Lepage - in "Perturbative Quantum Chromodynamics, Tallahassee (1981)", edited by D.W. Duke and J.F. Owens (AIP, New York, 1981).
- 30) J.C. Wrigley - Cavendish Laboratory Preprint HEP 81/10 (September 1981).
- 31) R. Barbieri, G. Curci, E. d'Emilio and E. Remiddi - Nuclear Phys. B154 (1979) 535.
- 32) P.M. Stevenson - Phys.Rev. D24 (1981) 1622.
- 33) W.A. Sanders and J.O. Hirschfelder - J.Chem.Phys. 42 (1965) 2904;
A. Dalgarno and A.L. Stewart - Proc.Roy.Soc.London A257 (1960) 534;
W. Caswell - Ann.Phys.(N.Y.) 123 (1979) 153;
I.G. Halliday and P. Suranyi - Phys.Letters 85B (1979) 421; Phys.Rev. D21 (1980) 1529;
J.W. Dash and S.J. Harrington - Phys.Letters 57B (1975) 78;
L.P. Kadanoff and A. Houghton - Phys.Rev. B11 (1975) 377;
S.M. Wandzura - J.Opt.Soc.Am. 71 (1981) 321.
- 34) See, for instance:
H.D. Politzer - Phys.Rev. D9 (1974) 2174;
G. Sterman and S. Weinberg - Phys.Rev.Letters 39 (1977) 1436.
- 35) A. Peterman - CERN Preprint TH. 3166 (September 1981).
- 36) O.V. Tarasov, A.A. Vladimirov and A.Yu. Zharkov - Phys.Letters 93B (1980) 429.
- 37) W. Celmaster - Phys.Rev. D19 (1979) 1517.

FIGURE CAPTION

An illustration of a failing of the FAC criterion. The approximants in two adjacent orders are shown as functions of the unphysical variable u . The convergence between $(n-1)^{\text{th}}$ and n^{th} orders is good for the PMS results, but becomes fastest only at a very different value, because of a tendency for the curves to be parallel [cf. Figs. 2, 4 of 6)].