Vacuum Structure and QCD Sum Rules: Introduction

M. Shifman

Theoretical Physics Institute, University of Minnesota, Minneapolis, MN 55455, USA

The story of this article is as follows. Shortly after my arrival to the United States I got a letter from Professor Hector Rubinstein who at that time was the Coordinating Editor of *Current Physics — Sources and Comments*, a serial North-Holland publication of reprint volumes on high-energy physics. Each volume was supposed to present a selection of essential source material on a given important topic compiled by an independent specialist editor, supplemented by extensive editorial commentaries and mini-reviews. Professor Rubinstein invited me to make a volume devoted to the QCD sum rules.[1] After some hesitations I agreed. The work turned out to be much more time and labor-consuming than I had expected. In several cases (Chapter 1, 8 and in some other cases) I failed to find among the published papers the one which I could consider satisfactory; then I had to write the corresponding sections myself.

Hector's idea was to put a historical introduction/foreword, which I was supposed to write, in the beginning of the book. By deadline I had only an unpolished and unfinished draft. Other urgent commitments prevented me from completing the article in time. Therefore, I had to settle for a short technical foreword which appeared in print in 1992. The original rather long Foreword has never been published. I used edited fragments of this article in various later publications. Presented below is the original draft.

⋆ ⋆ ⋆ ⋆ ⋆

The purpose of this Introduction is to give a general perspective of the development of the method which goes under an awkward name *the QCD sum rules* – one of the most productive tools for calculating hadronic parameters in Quantum Chromodynamics. The second task is to convey the historic flavor of the exciting time when the theory of hadrons was making its first steps.

It is, perhaps, worth starting from the second point. It seems fair to say that QCD was born after the talk of Gell-Mann and Fritzsch [1] (see also Ref. [2])

[1] This project was aborted in 1992. In fact, the Volume *Vacuum Structure and QCD Sum Rules* of which I was the editor, was the tenth and the last in the series.

in which the color-octet gluons were introduced. The next step is certainly the discovery of asymptotic freedom in 1973 [3]. In the first few years, roughly speaking till 1975, the theorists' attention was almost totally focussed on hard processes in which the short-distance physics plays the dominant role. This topic became hot and fashionable, piles of new papers appearing daily. The main achievement of this period is that people learned how to reliably isolate the short distance contributions governed by the small coupling constant and how to generalize electrodynamical perturbative calculations to non-Abelian theories.

A recent paper of Polyakov [4] presenting his understanding of the development of our field in the seventies is entitled "A View from the Island" which gives a good idea of our place in the scientific process in Moscow in those days. The isolation was almost complete, and we could not compete with the western theorists in most fashionable directions where the results seemed to be on the surface, for obvious reasons.

Because of the total censorship preprints and journals from the West used to come with enormous delays, and our own papers could be submitted to the Western journals only after a complicated procedure of getting clearance from half a dozen of different instances, typically a waste of a few months. Publication of preprints was an adventure due to bureaucratic limitations. For instance, one and the same paper could not be issued as a preprint and, then, in a journal, and the preprint version could not be longer than 25 typewritten pages (or 35, I do not remember exactly). So we had a lot of "fun" trying to muddle and deceive our censors by making different titles in the preprint and journal versions of one and the same work, changing the order of the authors or serializing preprint publications like a detective story in a popular newspaper (unlike the latter case, though, we had to make an impression that each successive part is not connected to the previous). Quite often these tricks created a terrible mess, to say nothing about wasted efforts. Occasionally, in the most important cases, one would risk to bypass the standard procedure by using different "illegal" channels, mostly our Western friends. By the way, any contacts with the latter were also severely constrained. There is a famous story about one of the scientific bosses in Dubna who was instructing the Soviet participants of a conference before its opening. He said: "Well..., we had to organize this conference, and even invite some foreigners. Unfortunately, to my deep regret, this time it will be impossible to completely avoid contacts..." Conferences in the West were open for a handful of specially selected, through a humiliating procedure; and even those few which took place in the Soviet Union were not always accessible. I remember, for instance, that a colleague of mine from ITEP and I were not granted permission to participate in "Neutrino –77" in Baksan. The decision to throw us out from the list of participants was made in the very last moment, and the only reason behind this decision I could think of was the fact that we both were Jewish.

Some words of the Soviet newspeak made me sick. One of such words was *Glavlit*. As I just said, to publish a scientific paper was much more than just typing the manuscript and mailing it to the publisher. There was a long latent period, associated with getting all sorts of clearances. First, the so-called Expert Commission (a group of authorized fellow physicists in the given institution) was supposed to study the paper and recommend its publication. According to the official rules they had to certify that no new discoveries were reported, because if they were, the Expert Commission had to recommend to classify the paper right away. Of course, people tended to stretch the official rules, otherwise not a single breakthrough paper would have ever appeared in the Soviet Union.

At the next stage the paper would go to the so called Regime Department whose task was to check that no references to classified work or undesirable persons were made, no subversive ideas put forward, and so on. With all this paperwork done, the decision to allow (disallow) publication was to be made by the Director of the Institute. This is not the end of the story, however. All materials intended for publication had to be cleared through the so-called GLAVLIT, the almighty agency whose sole obligation was to ensure total Censorship in the country. If, at the previous stages the author would have at least some minimal control over what was going on with his (her) paper, GLAVLIT was a total black box.

The process of getting all clearances could extend anywhere from weeks to many months, and the paper was officially nonexistent until the very end. The author could not even refer to it in his/her further work.

This is a long saga, and I could easily write many pages on this topic, but now it is clearly time to stop. To make the long story short I will only say that making our results known was a difficult, nervous and time-consuming part of our job. This largely predetermined the choice of topics we could work on and formed in the ITEP theory group and elsewhere a very special atmosphere, now gone forever.

ITEP of 1970's had one of the best groups in the world, an excellent collection of enthusiasts whose attitude to physics was totally "non-commercial". People were always eager to discuss with each other every interesting scientific question emerging during the seminar talks or elsewhere, and these discussions quite often would last till midnight. You could easily find experts in any conceivable field or direction who would gladly share their knowledge with you. Our common enemies and common isolation created, as a counter-reaction, very strong friendly and scientific ties, as the only way of survival, and helped develop protective values; the tacit understanding that good physics was above all was among these values.

The only drawback I can think of now, in retrospect, is the general negative attitude to field theories in the very end of 1960's and the beginning of 1970's when I just appeared in ITEP. The reason is obvious, of course: the influence of Landau and his discovery of the zero-charge property in QED [5] – the influence which was alive and very strong in the ITEP theory group in those days. The attitude to the field theory as to something absolutely not serious was so deeply rooted that the fact of the *anti*screening of the gauge coupling constant in non-Abelian theories which was reported in ITEP at least twice [6], [7] in the *sixties* has not been appreciated and recognized [8]. A restructuring in minds started only after the very same fact, the surprising fall-off of the gauge charge at large distances, became known from

Ref. [3].

So, our start was relatively slow. By 1974, however, we were fully submerged in this subject, and shortly after it became clear that the cavalry attacks do not help to solve the problem of confinement and that the wide-spread expectations of a Messiah who would come soon and teach us the mystery of the solution had to be tempered. The fact that instantons [9], a beautiful theoretical construction which was a breakthrough in the qualitative understanding of the QCD vacuum [10], turned out to be useless in the quantitative sense because of the infrared divergences was a serious blow. So we adopted a less ambitious approach (by "we" I mean Valya Zakharov, Arkady Vainshtein and myself). The idea was to start from short distances where the quark-gluon dynamics was essentially perturbative and we felt ourselves on a firm ground, and then to extrapolate to larger distances (where the hadronic states are presumably formed) including non-perturbative effects "step by step" and using some kind of an approximate procedure to extract information on hadronic properties. Of course, this idea was rather vague at first, the program as we know it now has been crystallizing gradually.

It is rather difficult to identify the work which, for us, was the first crucial step. With some reservations I might say that the first step has been done in Ref. [11]. Perhaps now the sum rule for the charmed particle photoproduction obtained in [11] does not seem very impressive but this analysis carried important elements which where later laid in the foundation of the sum rule method. A spectacular success came after we united our efforts with V. Novikov, L. Okun and M. Voloshin. It turned out that a whole variety of the charmonium parameters are predictable essentially from pure duality, and for about a year we played the game of constructing the charmonium widths and mass ratios from simple numbers. In 1977 a review report [12] was submitted. At about that time it became clear that the progress was limited; the method presented in [12] could not be reliably generalized to such typical representatives of the hadronic family as, say, ρ mesons or nucleons. And the desire to get access to these hadrons was strong.

The remainder of this story, including its culmination – introduction of the gluon condensate [13] in fall 1977 – is described elsewhere [14]. It was a hot summer of 1977, just before the vacations. Our big six-head strong collaboration ceased to exist. We — Valya, Arkady and myself — were leisurely discussing something when the first hints appeared. The conjecture was that the vacuum was actually something like a gluon "medium," and the basic particle properties are due to the quark interaction with this medium which can be conveniently parametrized by certain quark and gluon condensates. We worked out the first implications of the gluon condensate in fall 1977. At first, we were discouraged by a wrong sign of the gluon condensate term in our "show-case" sum rule, the one for the ρ meson. Then we suddenly understood that this sign could be compensated by the four-quark condensate — a real breakthrough. The accuracy of our results turned out to be much higher that anyone could expect *a priori*.

The basic elements of the approach were already visible in the first JETP Letters publication [13], with new important findings (e.g. borelization) coming later, one by one. We worked at a feverish pace for the whole academic year, accumulating a large number of results for the hadronic parameters.

I can't help mentioning an eposide that occurred in the spring of 1978 when we were done with this work. The episode could have been funny were it not so nerve-wrecking for me. When we decided that this stage of the work was over, I collected all my drafts (hundreds sheets of paper with calculations and formulas), organized them in a proper order, selected all formulas and expressions we might need for the paper and the future work, carefully rewrote them in a thick notebook (remember, we had no access to photocopying machines!), destroyed the drafts, put the the notebook in my briefcase and went home. It was about midnight, and I was so exhausted that I fall asleep while on the metro train. A loud voice announcing my stop awoke me, and I jumped out of the train, leaving the briefcase were it was, on the seat. By the time I realized what happened the train was gone, and gone with it forever my calculations ... I have never recovered my briefcase with the precious notebook... After a few agonizing days it became clear that the necessary formulas and expressions had to be recovered anew. Fortunately, Arkady had kept lots of his own calculations. He never throws out anything. Therefore, the problem with his drafts was to dig out "informative" sheets of paper from the "noise". (This was hindered by the fact that he was in Novosibirsk while we were in Moscow.) Some of my drafts survived in the drawers of a huge desk which I had inherited from Sudakov. Moreover, many crucial calculations were discussed so many times by us, over and over again, that I remembered them by heart. Nevertheless, I think it took a couple of uneasy weeks to recover the contents of the lost notebook.

In late spring 1978 the question of how all this wealth could be published became of prime concern to us. As usual, we had a lot of funny adventures in preparing the manuscript, typing it and issuing preprints. Needless to say, it was a serial publication, see above. As for the journal article, Nuclear Physics was a natural candidate, but previously we had bad luck with this journal: our paper on penguins [15] was buried in the editorial office for more than two years. We were too tired, however, to invent anything new and decided to try our luck again. The report occupying the whole issue of Nuclear Physics [16] appeared a year later.

Beginning in 1980, the approach we suggested was tested, with a remarkable success, in analyzing practically every static property of all established low-lying hadronic states. "Classic" mesons were supplemented by baryons [17]. The approach was then developed in various directions — three-point functions, inclusion of external electromagnetic and other auxiliary fields, light-cone modifications and so on. This allowed one to expand the range of applicability, shifting the emphasis from the calculation of masses and coupling constants of "classic" resonances to such problems as magnetic moments [18], form factors [19] at intermediate momentum transfers, weak decays [20], structure functions [21] of deep inelastic scattering at intermediate x, and many others. A wealth of low-energy parameters – dozens, if not hundreds – were obtained with a typical theoretical uncertainty of 10 to 30% essentially from the gluon and quark condensates and a couple of other vacuum mean values introduced later. At the initial stage the method was essentially unchallenged since the lattice calculations were lagging far behind.

At present, after the spectacular advances of lattice QCD, the numerical aspect of the sum rule approach is somewhat overshadowed, in many instances, although even now the QCD sum rules are quite competitive in problems with more complicated kinematics. At the same time, theoretical insights and tools developed in connection with this method did not loose its significance, rather they gained more in the recent years. First and foremost, one should name the operator product expansion (OPE).

The idea of factorization of short and large distances, the central point of OPE, dates back to classical Wilson's work where it was put forward in connection with theories of strong interaction with conformal invariance at short distances. Shortly after Wilson formulated a very general procedure of the renormalization-group flow (see e.g. [22]) which became known as Wilsonean renormalization group. Wilson's formulation makes no reference to perturbation theory, it applies both to strongly and weakly coupled theories. The focus of Wilson's work was on statistical physics, where the program is also known as the block-spin approach. Starting from the microscopic degrees of freedom at the shortest distances a, one "roughens" them, step by step, by constructing a sequence of effective (composite) degrees of freedom at distances $2a$, $4a$, $8a$, and so on. At each given step i one constructs an effective Hamiltonian, which fully accounts for dynamics at distances shorter than a_i in the coefficient functions.

Surprisingly, in high-energy physics of 1970's the framework of OPE was narrowed down to a very limited setting. On the theoretical side, it was discussed almost exclusively in perturbation theory, as is seen, for instance, from Refs. [23]. On the practical side, in applications, people did not go beyond the leading power effects. The OPE-based analysis of Christ *et al.* [24] of deep inelastic scattering (DIS) was implemented in QCD at the level of leading twists. Asymptotic freedom implies that the power asymptotic behavior is determined by canonic dimensions of the operators involved. The asymptotic behavior is modified by perturbative logarithms. This explains both the Bjorken scaling in DIS as well as logarithmic deviations from it. Similar logarithmic corrections were found in weak decays (penguins!). In essence, that was all one could find on the theoretical market of the day.

We were the first to adapt [26] the general Wilson construction to QCD, extending OPE to systematically include power suppressed effects. A consistent Wilsonean approach requires introduction of a normalization point μ which plays the role of a running parameter separating "hard" contributions included in the coefficient functions and "soft" contributions residing in local operators occurring in the expansion. The degree of locality is regulated by μ itself.

Prevalent at those days was a misconception that the OPE coefficients are determined exclusively by perturbation theory while the matrix elements of the operators involved are purely nonperurbative. Attempts to separate perturbation theory from "purely nonperturbative" condensates gave rise to multiple paradoxes and inconsistencies (see e.g. [25]) which questioned the very possibility of using the OPE-based methods in QCD. We had to spend much effort untangling these spurious paradoxes each time the claim of the failure of OPE has been made [26]. Surprisingly, such claims are being made till present. This became a perennial problem reminding of *perpetuum mobile* seekers.

Certainly, Wilsonean OPE, being just a book-keeping procedure, cannot fail. The question of its practical implementation is a different story. The coefficient functions summarize the hard contributions, which, to a good approximation, are given by perturbation theory. However, strictly speaking, perturbation theory does not exhaust QCD dynamics even at short distances. If there are small-size nonperturbative fluctuations, the coefficient functions acquire nonperturbative parts. An instructive example is provided by instantons. We found [26] that in most cases their contribution is suppressed and can be neglected. That is why our sum rules were so successful in the family of the "classic" mesons. But not in all cases! We managed to identify some exceptional channels (which will *not* be discussed in this book). It is clear that if instantons can produce nonperturbative parts of the coefficient functions, this can be possibly done by other small-size field fluctuations. The work on identifying such novel field fluctuations is on today's agenda.

So much was said about the construction of consistent OPE in QCD because, after this was done in connection with the QCD sum rules, it gained a life of its own! The very same OPE constitutes the basis of the heavy quark mass expansions (Voloshin + me, Yadernaya Fizika). The heavy quark theory accounts for most significant developments in QCD in the recent years. This is a branch of QCD where there still exists a direct live feedback from experiment, which gives a special weight to any advancement in theoretical understanding. One of the most elegant results established [27] in the heavy quark physics, on the basis of OPE, is the absence of the $1/m_Q$ correction to the total inclusive decay widths of the heavy-flavor hadrons. This theorem will undoubtedly make its way into text books, let alone its practical importance for the precision determination of V_{cb} from data.

On the technical side, the progress of the OPE-based methods we are witnessing now would hardly be possible without the computational technique based on the background field method. Back in 1978, when we entered this field, people used a very awkward approach, a direct evaluation of the Feynman graphs. It worked okay as far as the leading power terms in OPE were concerned. We needed to go beyond the leading terms. It became clear almost immediately, that one cannot go further without perfecting the computational technique. The background field formalism was known for a long time in QED, were it had been developed mainly by J. Schwinger. It is extremely efficient and ideally suited for OPE-based calculations. Somehow the QCD practitioners paid no attention to it. So, we adapted it to make it usable in the QCD environment. In fact, two versions were engineered: one is based on the Fock-Schwinger gauge for the background field, another is more general. They nicely supplement each other in various problems. The review paper [28] (see

Chapter 3) summarizes our work on the background field formalism; exploiting it became a routine practice from then on.

References

[1] H. Fritzsch and M. Gell-Mann, *in* Proc. XVI Int. Conf. on High Energy Physics, Batavia 1972 (J.D. Jackson and A. Roberts, eds.), Vol. 2, p. 135.

[2] H. Fritzsch, M. Gell-Mann and H. Leutwyler, *Phys. Lett.* **B47** (1973) 365.

[3] D. Gross and F. Wilczek, *Phys. Rev. Lett.* **30** (1973) 1343;
H. D. Politzer, *Phys. Rev. Lett.* **30** (1973) 1346.

[4] A. Polyakov, *Preprint* PUTP-1353, 1992 [to be published in Proc. Conf. *The Rise of the Standard Model: Particle physics in the 1960's and 1970's.* Ed. L. Hoddeson, L.Brown, M.l Riordan, and M.Dresden (Cambridge Univ. Press)].

[5] L.D. Landau, in *Niels Bohr and the Development of Physics*, ed. W. Pauli (Mc-Graw Hill, New York, 1955).

[6] V.S. Vanyashin and M.V. Terentyev, *ZhETF* **48** (1965) 565 [*Sov. Phys. - JETP* **21** (1965) 375].

[7] I.B. Khriplovich, *Yad. Fiz.* **10** (1969) 409 [*Sov. J. Nucl. Phys.* **10** (1970) 235].

[8] The paper of Vanyashin and TerentyevVanyashin and Terentyev is entitled *"The vacuum polarization of a charged vector field"*. Let me quote a passage from the concluding part of the paper: "Unlike the ordinary electrodynamics where the renormalized charge is smaller than the bare one our case gives $e^2 > e_0^2$... As is well-known, in electrodynamics the restriction $0 \leq Z \leq 1$ on the Z factor (charge renormalization) follows from the most general principles... In our case it is impossible, generally speaking, to get the restriction $0 \leq Z \leq 1$ starting from these principles... Nevertheless, the inequality $Z > 1$ seems extremely undesirable." The first coefficient of the Gell-Mann-Low function obtained in [6] (for $SU(2)$) is 20/3, instead of 22/3 which we would get now in the massless Yang-Mills. The difference is due to the contribution of the longitudinal polarizations, which in the modern language would correspond to the Higgs fields. Vanyashin and Terentyev worked before the Higgs mechanism; they introduced the W-boson mass by hand. This is okay at one loop.

In Khriplovich's work a non-covariant (Coulomb) gauge was used which is ghost-free. In this gauge the charge renormalization is given by that of the D_{00} component of the vector Green function. The logarithmically divergent part of D_{00} is presented in eq. (61) of ref. [7]. After taking into account an unconventional normalization of the coupling constant there we read off the famous 22/3! Moreover, you additionally learn from this expression that 22/3 is actually $8-\frac{2}{3}$. The logarithms are of two different types. One is non-dispersive, corresponds to the instantaneous Coulomb interaction and is responsible for antiscreening (*eight*). The second logarithm is normal dispersive; in full accordance with the general principles it gives the negative contribution (*minus two thirds*) characteristic of screening.

[9] A. Belavin, A. Polyakov, A. Schwarz and Yu. Typkin, *Phys. Lett.* **B59** (1975) 85.

[10] V. Gribov, 1975, unpublished;
C. Callan, R. Dashen and D. Gross, *Phys. Lett.* **B63** (1976) 334.

[11] M. Shifman, A. Vainshtein and V. Zakharov, *Phys. Lett.* **B65** (1976) 255.

[12] V. Novikov, L. Okun, M. Shifman, A. Vainshtein, M. Voloshin and V. Zakharov, *Phys. Rep.* **41** (1978) 1.

[13] A. Vainshtein, V. Zakharov and M. Shifman *Pisma v ZhETF* **27** (1978) 60 [*JETP Lett.* **27** (1978) 55].

[14] M. Shifman, *Current Contents (Eng. Techn. Appl. Sci.)* **23** (1992) No 34, p. 9.

[15] M. Shifman, A. Vainshtein and V. Zakharov, *Nucl. Phys.* **B120** (1977) 316.

[16] M. Shifman, A. Vainshtein and V. Zakharov, *Nucl. Phys.* **B147** (1979) 385; 448; 519.

[17] B.L Ioffe, *Nucl. Phys.* **B188** (1981) 317; (E) **B191** (1981) 591;
Y. Chung, H.G. Dosch, M. Kremer and D. Schall, *Phys. Lett.* **B102** (1981) 175; *Nucl. Phys.* **B197** (1982) 55.

[18] B.L Ioffe and A.V. Smilga, *JETP Lett.* **37** (1983) 298; *Nucl. Phys.* **B232** (1984) 109;
I. Balitsky and A. Yung, *Phys. Lett.* **B129** (1983) 328.

[19] V. Nesterenko and A. Radyushkin, *Phys. Lett.* **B115** (1982) 410;
B.L Ioffe and A.V. Smilga, *Nucl. Phys.* **B216** (1983) 373.

[20] B. Blok and M. Shifman, *Sov. J. Nucl. Phys.* **45** (1987) 135; 301; 522; *Sov. J. Nucl. Phys.* **46** (1987) 1310;
for a review and a representative list of references see A. Khodjamirian and R. Rückl, hep-ph/9801443.

[21] V. Belyaev and B.L. Ioffe, *Int. J. Mod. Phys.* **A6** (1991) 1533.

[22] K. Wilson and J. Kogut, *Phys. Reports* **12**, 75 (1974).

[23] K. Symanzik, *Commun. Math. Phys.* **23**, 49 (1971);
C. Callan, *Phys. Rev.* **D5**, 3302;
W. Zimmermann, in *Lectures on Elementary Particles and Quantum Field Theory*, Eds. S. Deser, M. Grisaru, and H. Pendelton (MIT Press, Cambridge, MA., 1971), Vol.1.

[24] N. Christ, B. Hasslacher, and A.H. Mueller, *Phys. Rev.* **D6** 3543 (1972).

[25] F. David, *Nucl. Phys.* **B209**, 433 (1982); *Nucl. Phys.* **B234**, 237 (1984).

[26] V.A. Novikov, M.A. Shifman, A.I. Vainshtein, and V.I. Zakharov, *Nucl. Phys.* **B249** 445 (1985).

[27] I.I. Bigi, N.G. Uraltsev, and A.I. Vainshtein, *Phys. Lett.* **B293** 430 (1992).

[28] V.A. Novikov, M.A. Shifman, A.I. Vainshtein, and V.I. Zakharov, *Fortsch. Phys.* **32**, 585 (1985).