METHODOLOGICAL NOTES

Stochastic trajectory generation by computer

É. V. Shuryak

Institute of Nuclear Physics, Siberian Branch of the USSR Academy of Sciences, Novosibirsk

Usp. Fiz. Nauk **143**, 309–317 (June 1984)

In recent years the Feynman formulation of quantum theory with the use of functional integrals became the basis of numerical methods, mainly in the theory of gauge fields. This approach is discussed in detail in the present study using examples of quantum mechanical problems. In the case of several degrees of freedom D this approach is substantially simpler and more powerful than the traditional approaches based on the Schrödinger equation: an average computer can handle problems with $D \sim 10$. Possible applications are briefly discussed.

CONTENTS

1. Introduction 448
2. Quantum mechanics and statistical physics 449
3. Statistical systems on a computer 450
4. Examples 450
5. Conclusion 452
References 453

1. INTRODUCTION

The traditional formulation of quantum mechanics, to which specialists in atomic, molecular, nuclear physics, and related areas have become accustomed, is based on the concept of a wave function, satisfying the Schrödinger equation. However, the application of this approach in the case of several degrees of freedom, where separation of variables is not possible, encounters many difficulties. Therefore in practice it is necessary to resort to various simplifying approximations, which are difficult to justify or verify.

The use of a Schrödinger wave function in quantum field theory, where the number of degrees of freedom is infinite, is obviously impractical, so from the very beginning one is forced to search for other methods.

An alternative formulation of quantum theory, using functional integrals, was suggested by Feynman.[1] With its help he first derived the celebrated rules of diagram techniques, which found wide application far outside the scope of quantum electrodynamics. The use of this approach in quantum field theory is a matter of routine at the present time.

Unfortunately, only functional integrals of Gaussian type can be calculated analytically, which strongly restricts the usefulness of the method.

In recent years a qualitative trend has occurred in this area—integrals over quantum fields have been investigated numerically by computer. The basic direction of developing this approach has become the study of contemporary strong interaction theory, quantum chromodynamics, particularly the calculation of the pion and proton masses, etc., from first principles. Though a fair amount of work has been done, the computational power and the skill of the theoreticians presently available for a direct solution of this problem are still inadequate.

The purpose of the present note is not a review of these studies, but an attempt to draw attention of the large community of theoreticians to the methodological aspect of the problem, as there is no doubt in my mind that the power and simplicity of this approach guarantee a large number of applications in many areas of quantum physics. Naturally, to achieve this purpose it is advisable to discuss the basics of the method, using very simple examples.

Concluding the introduction, we would like to note the psychological and organizational trends toward active application of numerical methods in elementary particle theory. It can be concluded with confidence that similar developments will also take place in other areas.

Today it is not necessary to convince anyone of the advantages of using computers; almost every physicist uses them on a daily basis. However, in most cases these numerical calculations play an auxiliary role (for example, numerical estimation of integrals, etc.), and no specific knowledge is required for them (for example, a simple Monte-Carlo method is used, based on random "tossing" of the argument).

However, to solve quite complicated problems deeper knowledge is required of *computational algorithms*, which may qualitatively change the whole approach to the problem. A good example is the subject of the present study. Mathematically the subject is estimation of integrals (more accurately, modeling of the integrand functions) with a very large number of variables (10^2–10^6), which is quite impractical when using ordinary methods of integration. An adequate algorithm was suggested 30 years ago,[3] as it turned out by physicists. Nevertheless, its wide use in the context mentioned is only starting.

Theoretical physicists did not confine themselves to becoming proficient in the use of appropriate algorithms. They took an active part in the *process of developing these algorithms*, in the same way that theoretical physics of the last century was instrumental in developing methods of solution of differential equations of "mathematical physics." Certain organizational approaches were developed, such as hiving

 0038-5670/84/060448-06$01.80

off a separate section of the journal "Nuclear Physics" under the name "Field theory and statistical systems" almost completely devoted to numerical methods.

2. QUANTUM MECHANICS AND STATISTICAL PHYSICS

Theoretical physicists naturally solve very different problems under the title of the present section. In the former case one calculates the probability amplitude of a certain dynamic process evolving in time. In the latter case—a standard problem is the calculation of the probability of realization of some configuration of a system in a steady (equilibrium) state. There exists, however, some connection between these problems, although quite a formal one, but also quite fruitful, on which we plan to dwell.

As a simple quantum-mechanical process we choose particle propagation from point x_i to point x_f during a time t_0. The probability amplitude is written as follows:

$$G(x_i, x_f, t_0) = \left\langle x_f \mid \exp\left(-\frac{i}{\hbar}\hat{H}t_0\right) \mid x_i \right\rangle = \int Dx(t) \exp\left\{\frac{i}{\hbar} S[x(t)]\right\}. \quad (1)$$

The first statement uses the traditional state vectors and the Hamiltonian operator $\hat{H}$, and the second—the Feynman path integral and the action $S[x]$.

If the potential $V(x)$ is time-independent, a transition is possible to imaginary (or, as is also called, Euclidean) time $\tau_0 = -it_0$, so that oscillating functions in (1) are replaced by exponentially decaying ones. Here an analogy occurs with the Boltzmann factor $\exp(-\hat{H}/T)$, where the role of temperature is displayed by $\hbar/\tau_0$.

The development of this analogy is aided by the construction of a statistical analog system, for which the Euclidean time τ_0 is conveniently discretized. We divide its range of variation $(0, \tau_0)$ into $N+1$ segments of length $a = \tau_0/(H+1)$, and we characterize the trajectory by the set of N points[1]) $x_k = x(\tau = ak)$, $k = 1, \ldots N$. By the definition of the path integral, it equals the limit as $N \to \infty$, $a \to 0$ of the N-fold integral over x_k. However, we do not yet let N tend to infinity, and attempt to use the language of statistical mechanics for finite N.

In our example the analog system is the one-dimensional lattice with step a, at whose sites are given "spins" x_k, $k = 0, 1, \ldots, N+1$, taking on arbitrary real values. Replacing time derivatives by finite differences, we obtain the Hamiltonian ("Euclidean action") in the form

$$\overline{H} = \sum_{k=0}^{N} \left[\frac{m}{2a}(x_{k+1} - x_k)^2 + aV(x_k)\right]. \quad (2)$$

Naturally, the derivative can also be approximated more accurately, in which case a different system is obtained, in which not only neighboring "spins" interact with each other.

In several cases, for example, in gauge field theory, the construction of an analog system[4] is not a simple matter, since it is desirable to retain the exact symmetry of the problem even at finite N. The choice of a "lattice action", similar to (2), is also a quite important and highly nontrivial problem.

In order to become somewhat more accustomed to the statistical language, and to understand how to use properly the propagator (1) at imaginary times, we discuss several very simple problems.

For a "spin" lattice the existence of a finite correlation length τ_{cor} is natural, as well as the tendency at $\tau_0 \gg \tau_{cor}$ of the distribution over x (in particular, the mean energy density per unit length) to some equilibrium value. What corresponds to these quantities in a quantum problem?

We recall in this case the standard expansion of the propagator (1) in terms of steady-state eigenfunctions

$$G(x_f, x_i, \tau_0) = \sum_n \psi_n^*(x_i)\psi_n(x_f) \exp\left(-\frac{E_n\tau_0}{\hbar}\right). \quad (3)$$

As $\tau_0 \to \infty$ the ground state, obviously, "survives," so that its characteristics are described by the long "lattice." In particular, the energy density is nothing other than E_0, the ground state energy. The correlation length characterizes the distance up to excited levels (the so-called spectral "gap"). In particular, if for some operator $A(x)$ (with $\langle 0|A|0\rangle = 0$) one measures the correlation function, then in virtue of

$$\langle A(\tau) A(0)\rangle \xrightarrow[\tau\to\infty]{} |\langle 1 \mid A \mid 0\rangle|^2 \exp\left[(E_0 - E_1)\frac{\tau}{\hbar}\right] \quad (4)$$

its decay with τ makes it possible to infer both the value of the "gap" and the matrix elements of the given operator. It is in just this manner that masses and other characteristics of hadrons are calculated in the method of sum rules[5] and lattice calculations in QCD.

Finally, we apply periodic boundary conditions $x_i = x_f = x$, and integrate (3) over x:

$$\int dx\, G(x, x, \tau_0) = \sum_n \exp\left(-\frac{E_n\tau_0}{\hbar}\right). \quad (5)$$

We have literally obtained the partition function at a temperature $T = \hbar/\tau_0$. Thus, a lattice of finite lengths provides information on the behavior of a quantum system at a finite temperature! It is in just this manner that the behavior of matter at high temperatures $T = 200$ MeV is investigated, at which a transition occurs to the quark-gluon plasma state, and hadrons "melt."

We include a few words concerning the limit $N \to \infty$, $a \to \infty$, corresponding to the transition from a lattice to the continuous case. Naturally, the correlation length τ_{cor}, having a literal physical meaning, must tend to some limiting value.

In terms of the lattice model this behavior implies that τ_{cor} becomes arbitrarily large in comparison with the lattice step a. The appearance of such long-wave fluctuations in statistical systems is possible only near phase transition points, recalling the classical effect of critical opalescence. Moreover, it is precisely at phase transition points that *dynamic universality* occurs, independence of the system behavior from the specific structure of matter on the atomic scale.

In the context being discussed of quantum path integrals universality has a very simple meaning—the answer

[1]) We note that $x_0 = x_i$, $x_{N+1} = x_f$ are fixed.

must not depnd on the specific shape of the trajectory approximation and on the action over a small size step in time a.

3. STATISTICAL SYSTEMS ON A COMPUTER

Thus, having formulated the lattice analog of the system, we turn now to its numerical analysis. An obvious problem is generated at the first step: what order of N (number of "spin" variables) would be realistically needed?

We assume that we are interested in the ground state. In this case it is required that

$$\tau_0 \gg \tau_{\text{cor}}. \tag{6}$$

On the other hand, it is obvious that only for a sufficiently small step

$$\tau_{\text{cor}} \gg a \tag{7}$$

can one expect a realistic description of the system. Assuming that each inequality holds within an order of magnitude, we obtain a minimum admissible $N \sim 10^2$. It is easily verified that the usual Monte-Carlo method, random tossing, is practically useless in computing integrals of this multiplicity. The point selection must not be random!

Proceeding to the description of more effective statistical algorithms, we must start from the fact that instead of the calculation itself of the integral

$$\int W(x)\,dx \tag{8}$$

we generate an ensemble of points $\{x\}$ with weight $W(x)$. With its help it is not difficult to calculate mean values of the form

$$\langle f\rangle \equiv \frac{\int f(x)\,W(x)\,dx}{\int W(x)\,dx}, \tag{9}$$

as simple arithmetical means over the ensemble. Usually $W(x)$ is much more complicated than $f(x)$, which is what justifies this procedure.

In the case of a single variable x one usually uses the Neuman method. On part of the x, y plane (corresponding to the region of variation of x and $W(x)$) we choose a point with uniform weight. If $y < W(x)$ (i.e., the point is located under the curve), we choose it in the ensemble $\{x\}$, while in the opposite case—we reject it. Repeating this procedure many times over, we obtain a set of points with the given weight.

In the multidimensional case the effectiveness of this method is low, as the parameter region consists of a negligible fraction of the whole volume. The decisive modification of the method is the replacement of random selection of the points x by a random shift in its preceding value. In particular, the Monte-Carlo method[3] consists of successive application of the procedure described above to each coordinate x_k in turn. The magnitude of the shift is determined from the condition that the change in $W(x)$ is of order unity.

The small-step variant has been much discussed lately; in it the process of random motion of points (in some "machine" time) is described by the Langevin equation:

$$\dot{x}_k = -\frac{\delta \overline{H}}{\delta x_k} + \eta_k(t), \quad \langle \eta_k(t)\,\eta_{k'}(t')\rangle = 2\delta_{kk'}\delta(t-t'), \tag{10}$$

where $\eta_k(t)$ is a Gaussian random force with a delta-function correlator. Although a small step leads to an increase in calculation time, better correlation functions are obtained; consult Ref. 6 on this point.

Finally, sometimes one uses the "thermal bath" method, in which the x_k value is sought for fixed remaining variables x_i, $i \neq k$, directly "inverting" the weight $W(x)$.

We add that successive use in turn of all variables is called a system iteration (or "scrambling"), while its successive multiple repetition leads to relaxation of the initial configuration to some equilibrium ensemble, signaled by the approach of mean values to their limiting values.

Errors may arise due to insufficient relaxation, insufficient statistics of configurations chosen for the ensemble, and, finally, specific effects of system discretization. Clearly, an optimally organized calculation holds all these errors to a single, preassigned level.

As seen from the discussion above, these methods are simple in essence. Their realization in the form of a numerical program is accessible to all. The fact that they started to be used in quantum mechanics only now is a clear consequence of inertia in thinking and lack of training.

4. EXAMPLES

As already mentioned, the basic applications of the method are related to four-dimensional lattice models of gauge theories, first developed by M. Creutz.[7] At the present time the "attack of QCD" with the use of the most powerful computers in the world had led to studies of statistical systems with a number of degrees of freedom reaching 10^5–10^6 (i.e., of quantum systems with $\sim 10^4$ degrees of freedom!). Of course, this in itself is a big success for the theory. However, after extracting the fourth-order root we obtain lattices of 0(10) steps along each axis, which, as has already been mentioned, is insufficient. For further progress in this problem we need new machines or new ideas.

Fortunately, physics includes a large number of much simpler problems, anticipating a solution. To these belong, for example, problems of quantum mechanics with several degrees of freedom, which we now propose to discuss. As we see, they require neither supercomputers, nor complicated programs.

Calculations of this kind in quantum mechanical problems are only getting started. The first was the study of M. Creutz and B. Freedman[8] of Brookhaven (USA), devoted to linear and nonlinear one-dimensional oscillators. A number of examples of increasing complexity is discussed in a study of O. V. Zhirov and the present author,[9] to which I shall primarily refer in the following.

One must also mention the first studies, no longer of merely an illustrative, but also of an applied nature. B. Bunk and U. Wolffe[11] investigated a one-dimensional rotor in an external (gravitational) field, with a detailed study of tunnel "rotation." In our own study[10] we calculated the Green's functions of a nonrelativistic quark and antiquark, propagating in complex vacuum fields. For obvious reasons we do not dwell on these studies.

We start with one of the simplest and extremely instructive problems of quantum mechanics, a nonlinear oscillator with two wells. After transforming to imaginary time the action is

$$S=\int d\tau\left[\frac{\dot{x}^2}{4}+(x^2-f^2)^2\right]. \tag{11}$$

It was mentioned in Section 3 that the ground state energy of the system can be calculated by using a sufficiently long "spin crystal." However, in attempting direct averaging of the Hamiltonian we encounter a divergence as $a\to 0$ of the mean kinetic energy, estimated by means of

$$\langle\dot{x}^2\rangle=\frac{\langle(x_{k+1}-x_k)^2\rangle}{a^2}\sim\frac{\hbar}{am}. \tag{12}$$

A different (finite) equation was suggested in Ref. 8, but it is much simpler to use the virial theorem

$$E_0=\lim_{\tau_0\to\infty}\left\langle\sum_{\mu=1}^{D}\frac{1}{2}x_\mu\frac{\partial V}{\partial x_\mu}+V(x)\right\rangle. \tag{13}$$

Figures 1 and 2 show the ground state energy and the probability of finding the particle at point x, calculated over the trajectory ensemble $\{x_k\}$; the latter is compared with $|\psi_0(x)|^2$. To obtain these results we needed $N\sim 100$ and several hundred iterations, which took about a minute on the quite average ES-1040 computer.

One of the most important methodological problems is the systematic error related to the introduction of a finite step in time a and the specific nature of the lattice action. As noted in Ref. 9, for sufficiently large a and "naive" action this error is basically due to the specific "jump" across the potential barrier in one step. Replacing in the action $\Sigma_k V(x_k)$ by $\int V d\tau$ along a broken path, we eliminate this effect, and sharply enhance the accuracy of the calculation. Similar "lattice instantons" are also well-known in more complicated problems, in which the choice of optimal action is of primary importance.

Naturally, these calculations are important mostly in the methodological sense: in the one-dimensional case it is simpler to solve the Schrödingr equation directly. However, in the presence of D variables the difference methods require a grid with k^D nodes, where $K\sim 10$–100. Clearly, the possibilities of this approach are severely restricted.

The statistical methods of path integration in the case of D degrees of freedom simply increase the number of "randomly sampled" variables, so that the bulk of calculations increases with D only linearly (not taking into account, it must be said, the increased complexity in the action).

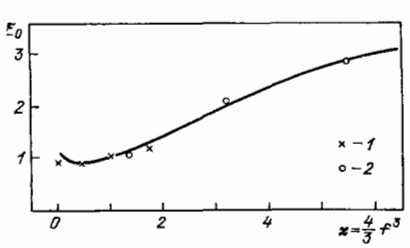

FIG. 1. The ground state energy of a two-well oscillator as a function of the parameter f, Eq. (11): 1) from Ref. 8, 2) from Ref. 9.

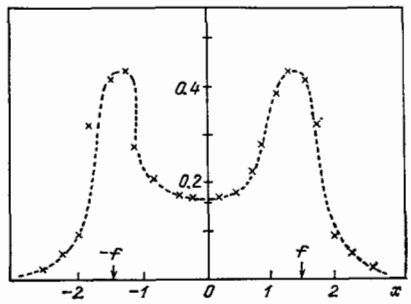

FIG. 2. The distribution with respect to x for the oscillator (11) with $f=1.4$.[8] The curve shows the exact value of $|\psi_0(x)|^2$.

Figure 3 shows estimates of the ground state energy of simple one- and two-electron atoms, obtained in Ref. 9. Simplification of the problem, i.e., separation of angular variables, was not carried out, so that trajectories were calculated, respectively, in three- and six-dimensional spaces. Naturally, as in all numerical methods, it is desirable to "smooth out" the singularity of the Coulomb potential at small distances. Another problem arising in this case is "ionization" of the atom for a sufficiently long statistical set. The point is that, as noted in Section 3, a finite τ_0 implies a nonvanishing temperature $T=\hbar/\tau_0$. This feature of the Coulomb potential, level crowding toward zero, facilitates ionization. The simplest way of preventing this is to forbid the electrons to get out to distances larger than some R, i.e., explicitly to cut off the "tails" of the wave functions.

An attempt was also undertaken in Ref. 9 to make the first calculations of properties of light nuclei, up to He^4 inclusively. As these calculations (up to 12 degrees of freedom) showed, an important feature of thse nuclear systems is their "friability," the binding energy per particle and the separa-

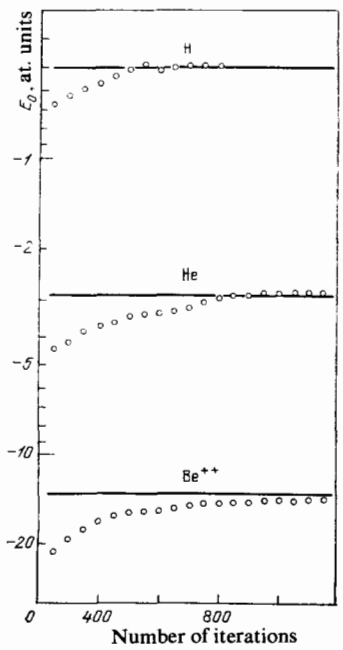

FIG. 3. The ground state energy of one- and two-electron atoms as a function of the number of iterations.

tions between levels are of the order of several MeV, which is much less than the characteristic parameters of the potential. This implies that the requirements on $T = \hbar/\tau_0$ are particularly strong, otherwise "evaporation" of nucleons in the course of the computation process is observed. The quantity N must be of the order of several hundred, and the total number of "spin" variables per lattice reaches 10^4 (which is by only one to two orders of magnitude less than in QCD computations!), and requires computer time of the order of hours on ES-1040. However, a definite ensemble relaxation is observed for very different potentials, with a rigid "core" or without it.

Finally, in several cases one needs the propagation amplitude in absolute normalization. The simplest method is to average a factor related to the potential, normalizing by the well-known free particle propagator

$$G(x_t, x_1, \tau_0) = G_0(x_t, x_1, \tau_0)\left\langle \exp\left[-\int_0^{\tau_0} V\, d\tau\right]\right\rangle. \qquad (14)$$

This method can also be applied to cases in which the particle has internal degrees of freedom, spin or color, so that the potential V is a matrix.[2)] Recently progress was achieved[10] in the quite difficult problem of sum rules for heavy quarkonia, in which a $q\bar{q}$-pair propagates in complicated color vacuum fields, with a Coulomb interaction between the members of the pair.

It is clear, however, that averaging the factor, strongly differing from unity, is a difficult matter, requiring more statistics for suppression of statistical noise. This difficulty can be avoided by adiabatic inclusion of the interaction, as suggested in Ref. 9. We introduce a parameter ξ, varing from 0 to 1:

$$S_\xi = \int d\tau\left[\frac{m\dot{x}^2}{2} + \xi V(x)\right]. \qquad (15)$$

The mean value of the integrated potential is the derivative of the partition function with respect to ξ; integrating this relation, we have

$$G = G_0 \exp\left[-\int_0^1 d\xi \left\langle \int d\tau\, V(x)\right\rangle_\xi\right], \qquad (16)$$

where the subscript ξ at the brackets denotes averaging with the action (15). Replacing the averaging of the exponent by the averaging of its index makes it possible to use this method even when the difference between G and G_0 is substantial; see, for example, Fig. 4.

5. CONCLUSION

The examples provided show convincingly the big potential and universality of the method for solving multidimensional quantum problems, particularly ones not allowing direct solution of the Schrödinger equation. Talking about its potential applications is quite difficult, but several general considerations can be brought forward.

An obvious application area is physics of systems of several particles, few-nucleon systems of nuclear physics,

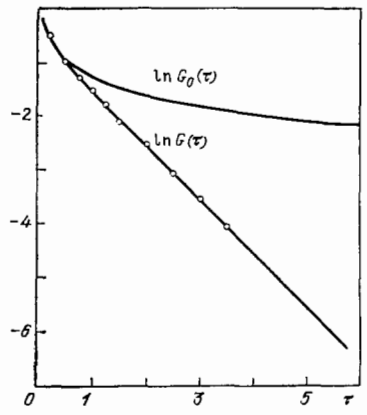

FIG. 4. The Green's function for a linear oscillator, obtained by Eq. (16) (marks), compared with the exact value (lower curve). For comparison we show the behavior of G_0 (upper curve).[9]

molecules, etc. In this connection it is particularly important to stress the possibility of studying molecules at a finite temperature—traditionally this requires unjustifiably long calculations of energy spectra and explicit evaluations of the partition function. The thermal "ionization" of atoms observed in the calculations described above, as well as nuclear disintegration, are, obviously, interesting applications.

Obviously, one can talk not only about such "monomolecular" reactions, but also about the probability of binary collisions of molecules with definite chemical reactions; one must start with some atomic trajectory leading from some given initial state to a given final state, and the computer chooses the optimal trajectory. In other words we are referring to the penetration probability through some complicated multidimensional barrier (in the coordinate space of all atoms).

We now make some remarks on what the given method cannot do. A fundamental feature of the statistical methods of ensemble generation is the probabilistic interpretation of weight functions, which cannot undergo a change of sign. Precisely therefore they cannot be used in ordinary (non-Euclidean) time, and nonstationary problems with explicit time dependence cannot be studied either. For the same reason we obtain information mostly about the ground state and the low-lying excited states, but not on all energy levels. Further, the presence of identical particles requires symmetrization or antisymmetrization methods. The latter (fermion) case gives nonpositive weights, so that working with them is also impossible. It is precisely the necessity of including the Pauli principle that restricted our calculations of atoms and nuclei.

Finally, the last comment, which is practically quite important. In the present study we have been dealing not with some exotic computational method, requiring a supercomputer and demonstrating the achievements of contemporary technology. As stressed earlier in the text, the computational tools used in the cases surveyed are practically accessible to most physicists. Neither are special programming systems required—what was discussed above is entirely sufficient for the writing by an individual of a workable

[2)] We recall that the statistical methods of ensemble generation cannot be applied to nonpostive weight functions.

model program, and the knowledge of detail comes with some experience. In brief, the topic that has been discussed is one of the computational methods required in daily work of a theoretical physicist.

The author is grateful to O. V. Zhirov for collaborating in the calculations discussed above.

[1]R. P. Feynman and A. R. Hibbs, Quantum Mechanics and Path Integrals, McGraw-Hill, 1965 (Russ. Transl., Mir, M. 1968).
[2]A. A. Belavin, A. M. Polyakov, A. S. Schwartz, and Yu. S. Tyupkin, Phys. Let. **B59**, 85 (1975).
[3]N. Metropolis, A. Rosenbluth, M. Rosenbluth, A. Teller, and E. Teller, J. Chem. Phys. **21**, 1087 (1953).
[4]K. Wilson, Phys. Rev. **D10**, 1974 (1974).
[5]M. A. Shifman, A. I. Vainshtein, and V. I. Zakharov, Nucl. Phys. **147**, 385 (1979).
[6]G. Parisi, Nucl. Phys. **B180** (FS2), 337 (1981).
[7]M. Creutz, Phys. Rev. **D21**, 2308 (1980).
[8]M. Creutz and B. Freedman, Ann. Phys. (N. Y.) **132**, 427 (1981).
[9]E. V. Shuryak and O. V. Zhirov, Preprint, Nuclear Physics Institute (INP) 83-47, Novosibirsk, 1983.
[10]E. V. Shuryak, Preprint INP83-113, Novosibirsk, 1983.
[11]B. Bunk and U. Wolffe, Nucl. Phys. **B215** (FS7), 496 (1983).

Translated by Nathan Jacobi