Complex Langevin Simulation*)

Keisuke OKANO,* Lothar SCHÜLKE and Bo ZHENG**)

*University of Tokuyama, Tokuyama-shi 745
FB Physik, Universität-Gesamthochschule Siegen, D-5900 Siegen, Germany

Contents

§ 1. Introduction

Extensive use of computers has been made for the numerical simulation of field theories. In Euclidean field theory, the expectation value of a certain physical quantity $\mathcal{O}(\phi)$ is given by a path-integral

$$\langle \mathcal{O}(\phi) \rangle_S \equiv \frac{\int d\phi\, \mathcal{O}(\phi) e^{-S(\phi)}}{\int d\phi e^{-S(\phi)}}, \tag{1·1}$$

where ϕ represents a field, and $S(\phi)$ is a Euclidean action of a system which is

*) Supported in part by the Deutsche Forschungsgemeinschaft, Contract No. Schu 97/5.
**) On leave of absence from Physics Department, Zhongshan University, Gaungzhou.

normally real. On the basis of the positivity of the integration measure $e^{-S(\phi)}$ standard Monte Carlo methods, Metropolis or heat bath algorithm, have been applied for the numerical simulation of the quantiy in (1·1).

It sometimes happens, however, that a Euclidean action is complex. Examples are the following important and interesting physical systems:

—Systems with *chemical potential*. It is an interesting task to investigate strongly interacting QCD systems also for non-vanishing chemical potential at finite temperature. For $SU(N)$ theories with $N>2$ the fermion determinant becomes complex and thus the effective action.

—Theories with *external charges*. The Euclidean action for a gauge theory with external static charges is complex. It is physically important to calculate the string tension and string width between a pair of static charges because they are closely connected to the confinement of quarks. The standard Monte Carlo simulation can only generate the vacuum configuration and extract the string tension from a large Wilson loop which is exponentially small. If one could directly generate configurations in the charged sector one can expect more accurate numerical results for the quantity.

—Gauge theories with *fermions*. There is no (real) c-number representation for the fermions in the standard path integral formalism. If one intends to treat the fermions directly in some special ways, rather than the fermion determinant indirectly, one meets complex actions or "negative probability".

—Effective actions including *topological terms* or *Chern-Simons theories*. Models of this kind have extensively been discussed recently. These actions cannot be made real by a Wick rotation.

—Field theories in *Minkowski space*. Although in many cases one can obtain a real Euclidean action through a Wick rotation, it is still interesting to search for a method which could solve the original problem, not the subsidiary one in Euclidean space.

In these cases the standard Monte Carlo method is not efficient. By including the imaginary part of the action into the observables, one can in principle calculate the average of a quantity $\mathcal{O}(\phi)$, i.e., with $S=S_R+iS_I$

$$\langle \mathcal{O}(\phi)\rangle_S=\frac{\langle \mathcal{O}(\phi)\exp\{-iS_I\}\rangle_{S_R}}{\langle \exp\{-iS_I\}\rangle_{S_R}}. \tag{1·2}$$

However if both the averages in the numerator and denominator are very small, the determination of $\langle \mathcal{O}(\phi)\rangle_S$ is practically not possible. Configurations following the probability distribution e^{-S_R} are *irrelevant* to the whole system. This actually happens for large volumes or for strong coupling in field theories in Euclidean space, while in Minkowski space it is even impossible since S_R is zero. It is therefore desirable to find a method by which the configurations distributed according to a "*complex probability distribution*" $\sim e^{-S}$ will directly be generated.

A possible algorithm for realizing the complex probability distribution" is the Langevin algorithm.[1] In order to calculate the quantity in (1·1), one writes a

Langevin equation

$$\frac{d\phi(t)}{dt}=-\frac{\delta S(\phi)}{\delta\phi}+\eta(t)\,, \tag{1·3}$$

where t is a fictitious time coordinate and different from the Euclidean time, and η is a Gaussian white noise. Let the solution of this equation be $\phi_\eta(t)$, then a long time average of the quantity $\mathcal{O}$ over this solution is supposed to give the above expectaion value (1·1):

$$\frac{1}{T}\int_{t_{\rm th}}^{t_{\rm th}+T} dt\,\mathcal{O}(\phi_\eta(t))\underset{T,t_{\rm th}\to\infty}{\longrightarrow}\langle\mathcal{O}(\phi)\rangle_S\,. \tag{1·4}$$

The correctness of this statement can be proved in case the action S is real (see the next section). On the other hand no general theorem has been found which states clearly what happens whenever the action S becomes complex. We can, however, solve the Langevin equation numerically independently of whether the drift force $D=-\delta S/\delta\phi$ is real or complex. It has, therefore, been proposed by Parisi[2] and Klauder[3),4)] that we may also calculate the expectation value in (1·1) for the case of a complex action S by using the Langevin equation together with the relation (1·4). It may be reasonable to expect, from the viewpoint of analyticity, that the solution still satisfies the relation (1·4) when the imaginary part of the action is not too large.

The complex Langevin simulation has been applied to a number of systems. In certain cases it turned out to work very nicely, its gain is tremendous and one can get results which can never be reached by the use of the standard Monte Carlo algorithm. In certain cases, however, it fails completely. Because of the tremendous gain obtained for certain systems many authors tried to investigate the cases of failure more carefully in order to get a clear a-priori statement what the reason for the failure is and whether it is possible to circumvent the difficulty. In § 3 we will review some typical analyses of them.

It has been known that the Langevin equation which has the asymptotic behavior (1·4) is not unique. This redundant degree of freedom in the Langevin equation is a kernel. Namely, Langevin equations with different kernels describe different relaxation dynamics, all of which, however, have the same asymptotic behavior (1·4). To adjust the kernel in order to get a smooth convergence to the equilibrium may have plenty of applications.*) Above all, investigation of the role of the kernel in the complex Langevin simulation is interesting.[9)~11)] This subject will be reviewed in §§ 2, 3 and 4.

§ 2. The fundamental framework of complex Langevin simulations

2.1. *The real Langevin equation*

In order to illustrate the fundamental idea of the Langevin simulation, we discuss a system with one degree of freedom x. The formulation, however, can straight-

*) The idea of introducing a kernel has been studied also in some other contexts in stochastic quantization, e.g., stochastic quantization of fermions,[5),6)] and attack upon the critical slowing down in numerical simulations.[7),8)]

forwardly be generalized to many degrees of freedom or to field theories. We write the Langevin equation as

$$\frac{dx(t)}{dt}=-S'(x)+\eta(t)\,, \tag{2·1}$$

here the drift force $-S'(x)$, and therefore the action S is taken to be real within this subsection 2.1. The Gaussian white noise $\eta(t)$ satisfies the following statistical property

$$\langle\eta(t)\rangle_\eta=0\,,\qquad \langle\eta(t)\eta(t')\rangle_\eta=2\delta(t-t')\,. \tag{2·2}$$

As was explained in the introduction, this Langevin equation describes a relaxation dynamics which has an asymptotic property given in (1·4). An aim of this subsection is to give an overview of the proof of this statement.

By solving the above Langevin equation (2·1) one gets a certain solution $x_\eta(t)$ depending on white noise $\eta(t)$. Solutions coming from different series of noise give a probability distribution $P(x,t)$ for x at a certain time t. An average of any quantity $\mathcal{O}(x_\eta)$ with respect to the white noise can be written as

$$\langle\mathcal{O}(x_\eta(t))\rangle_\eta=\int dxP(x;t)\mathcal{O}(x)\,. \tag{2·3}$$

In order to derive an evolution equation for the probability density $P(x;t)$ one takes the derivative with respect to the fictitious time on both sides of the equation,

$$\frac{\langle d\,\mathcal{O}(x_\eta(t))\rangle_\eta}{dt}=\int dx\,\mathcal{O}(x)\partial_t P(x;t)\,. \tag{2·4}$$

To evaluate the l.h.s. of this equation, we rewrite the Langevin equation (2·1) as

$$dx=-S'(x)dt+dW(t) \tag{2·5}$$

with

$$dW(t)=\int_t^{t+dt}d\tau\eta(\tau)\,. \tag{2·6}$$

The average of the square of dW can be reduced with the help of Eq. (2·2) to

$$\langle dW(t)^2\rangle_\eta=\int_t^{t+dt}d\tau\int_t^{t+dt}d\tau'\langle\eta(\tau)\eta(\tau')\rangle_\eta=2dt\,. \tag{2·7}$$

In order to evaluate $\langle d\,\mathcal{O}(x_\eta(t))\rangle_\eta$ up to the first order of dt one has to expand $d\,\mathcal{O}(x_\eta(t))$ up to the second order in dx, because of the property (2·7),

$$\begin{aligned}\frac{\langle d\,\mathcal{O}(x_\eta(t))\rangle_\eta}{dt}&=\langle-\mathcal{O}'(x)S'(x)+\mathcal{O}''(x)\rangle_\eta\\&=\int dx[-\mathcal{O}'(x)S'(x)+\mathcal{O}''(x)]P(x;t)\\&=\int dx\,\mathcal{O}(x)[\partial_x S'(x)+\partial_x{}^2]P(x;t)\,.\end{aligned} \tag{2·8}$$

A partial integration has been performed assuming zero boundary conditions at infinity. Together with the right-hand side of Eq. (2·4) one obtains the Fokker-Planck equation

$$\partial_t P(x;\, t) = -H_{\mathrm{FP}} P(x;\, t) \tag{2·9}$$

with the Fokker-Planck Hamiltonian

$$H_{\mathrm{FP}} = -\partial_x(\partial_x + S'(x))\,. \tag{2·10}$$

With a similarity transformation

$$\tilde{P}(x;\, t) = e^{S(x)/2} P(x;\, t) \tag{2·11}$$

the Fokker-Planck equation can be rewritten as

$$\begin{aligned} &\partial_t \tilde{P}(x;\, t) = -\tilde{H}_{\mathrm{FP}}(x)\tilde{P}(x;\, t)\,, \\ &\tilde{H}_{\mathrm{FP}}(x) = e^{S(x)/2} H_{\mathrm{FP}}(x) e^{-S(x)/2} \\ &\qquad = \left(-\partial_x + \frac{1}{2}S'(x)\right)\left(\partial_x + \frac{1}{2}S'(x)\right). \end{aligned} \tag{2·12}$$

If $S(x)$ is real, $\tilde{H}_{\mathrm{FP}}(x)$ is a self-adjoint and semi-positive definite operator, which has a complete set of eigenfunctions $\{\psi_n\}$ with non-negative eigenvalues E_n. It is easy to understand that $\psi_0 = e^{-S(x)/2}$ is the ground state with $E_0 = 0$. Writing $\tilde{P}(x, t) = e^{-\tilde{H}_{\mathrm{FP}} t}\tilde{P}(x; 0)$ and expanding $P(x, 0)$ with the above complete set, one gets

$$\tilde{P}(x;\, t) = a_0 e^{-S(x)/2} + \sum_{n=1}^{\infty} a_n \psi_n e^{-E_n t} \xrightarrow[t\to\infty]{} a_0 e^{-S(x)/2}\,. \tag{2·13}$$

Therefore $P(x;\, t)$ will have a relaxation

$$P(x;\, t) \xrightarrow[t\to\infty]{} P_{\mathrm{eq}}(x) \propto e^{-S(x)} \tag{2·14}$$

in the limit $t \to \infty$. This fact together with the ergodicity hypothesis makes the Langevin equation (2·1) a convenient tool for producing sets of solutions distributing according to $\exp\{-S\}$. It is, of course, not convenient to calculate a large set of solutions $\{x_\eta(t)\}$ for different η in order to get a reasonable average for large t. However, when t is greater than a certain time tth needed for thermalization of the system the solutions fluctuate around the thermal equilibrium and we can therefore calculate the average from one solution over a large time interval T,

$$\langle \mathcal{O}(x_\eta(t))\rangle_\eta \approx \frac{1}{T}\int_{t_{\mathrm{th}}}^{t_{\mathrm{th}}+T} d\tau\, \mathcal{O}(x_\eta(\tau)) \approx \langle \mathcal{O}(x)\rangle_S\,. \tag{2·15}$$

The last relation follows from (2·3) and (2·14).

2.2. *The complex extension*

Now, let us start to discuss the case with the complex drift force $-S'$. Due to the complexity of the drift force the solution of the Langevin equation becomes complex. Therefore, we extend $x(t)$ into the complex plane with $z(t) = x(t) + iy(t)$ and write the Langevin equation as

$$\frac{dz(t)}{dt}=-S'(z)+\eta(t)\,. \tag{2·16}$$

Note that the random force η is kept to be real.

In order to see whether the relaxation dynamics given by this Langevin equation retains the same asymptotic property as (2·15), we try to follow the same discussion as in the previous subsection. Let us introduce an effective "*complex valued probability distribution*" $P_c(x;\,t)$ by

$$\langle\mathcal{O}(z_\eta(t))\rangle_\eta=\int dx P_c(x;\,t)\mathcal{O}(x)\,. \tag{2·17}$$

The integral on the right-hand side extends over the real variable x, and $P_c(x;\,t)$ is a complex valued function. The key for the complex Langevin simulation is whether the relaxation

$$P_c(x;\,t)\xrightarrow[t\to\infty]{}P_{\rm eq}(x)\propto e^{-S(x)} \tag{2·18}$$

does really occur or not.

In deriving the Fokker-Planck equation, the same manipulations which have led us from Eqs. (2·3) to (2·10) can formally be applied, taking Eq. (2·17) as a definition. This leads to the same result,

$$\partial_t P_c(x;\,t)=-H_{\rm FP}P_c(x;\,t)\,,\qquad H_{\rm FP}=-\partial_x(\partial_x+S'(x)) \tag{2·19}$$

or after the similarity transformation $\tilde{P}_c(x;\,t)=e^{S(x)/2}P_c(x;\,t)$,

$$\partial_t\tilde{P}_c(x;\,t)=-\tilde{H}_{\rm FP}(x)\tilde{P}_c(x;\,t)\,,\quad \tilde{H}_{\rm FP}(x)=\Big(-\partial_x+\frac{1}{2}S'(x)\Big)\Big(\partial_x+\frac{1}{2}S'(x)\Big)\,. \tag{2·20}$$

Due to the complexity of the action S, the Fokker-Planck Hamiltonian $\tilde{H}_{\rm FP}$ is now no more self-adjoint and semi-positive definite. *A general proof* of the convergence (2·18) seems, therefore, not possible in the complex case. However, due to the analyticity arguments, if the imaginary part of the Fokker-Planck Hamiltonian is sufficiently small, it may still have a complete set of complex eigenfunctions ψ_n with positive semi-definite eigenvalues E_n. The relaxation (2·18) will be ensured in these cases and we can use the complex Langevin equation (2·16) to obtain a set of complex numbers $\{z_\eta\}$ which will be distributed according to the measure $\exp\{-S\}$. This conjecture has explicitly been checked by using a simple model with one degree of freedom in Ref. 12).

The complex valued distribution $P_c(x;\,t)$ does not have the meaning of a real probability density. We can, however, rewrite the Langevin equation (2·16) using two real variables x and y with $z=x+iy$,

$$\begin{aligned}\dot{x}&=\mathcal{R}\{-S'(z)\}+\eta\,,\\ \dot{y}&=\mathcal{I}\{-S'(z)\}\,.\end{aligned} \tag{2·21}$$

A *real* probability distribution $P(x,\,y;\,t)$ being a positive definite function of x and y and corresponding to the stochastic processes can be defined by

$$\langle \mathcal{O}(z_\eta(t))\rangle_\eta = \int dxdy P(x, y; t)\mathcal{O}(x+iy)\,. \tag{2·22}$$

Following the discussion in the previous subsection it is straightforward to derive the Fokker-Planck equation for $P(x, y; t)$

$$\partial_t P(x, y; t) = -H_{\mathrm{FP}} P(x, y; t)\,,$$
$$H_{\mathrm{FP}} = -\partial_x{}^2 - \partial_x \Re(S') - \partial_y \Im(S')\,. \tag{2·23}$$

The distribution $P(x, y; t)$ is related to the complex valued distribution $P_c(x; t)$ by

$$\int dxdy P(x, y; t)\mathcal{O}(x+iy) = \int dx P_c(x; t)\mathcal{O}(x)\,. \tag{2·24}$$

The $P(x, y; t)$, therefore, is a probabilistic distribution in the sense of (2·24). To obtain $P(x, y; t)$ by solving the Fokker-Planck equation (2·20) or even the stationary equation, $H_{\mathrm{FP}}P(x, y; t)=0$, however, is not so easy. This problem has been solved only for a simple Gaussian model with one degree of freedom.[13),14)]

Before coming to the next subsection, it is instructive to show some of mathematical formulae to relate both representations with $P(x, y; t)$ and $P_c(x; t)$.[15)] For this purpose, let us try to derive (2·19) starting from (2·23). In the derivation an identity

$$e^{-iy\partial_x}\mathcal{O}(x)e^{iy\partial_x} = \mathcal{O}(x-iy) \tag{2·25}$$

and also relations

$$e^{-iy\partial_x}\partial_x e^{iy\partial_x} = \partial_x\,,$$
$$e^{-iy\partial_x}\partial_y e^{iy\partial_x} = \partial_y + i\partial_x \tag{2·26}$$

are useful. Intergrating by part over y we find that Eq. (2·24) is fulfilled if

$$\int dy e^{-iy\partial_x} P(x, y; t) = P_c(x; t)\,. \tag{2·27}$$

Taking the derivative on both sides of Eq. (2·27) with respect to the fictitious time t, and using the Fokker-Planck equation (2·23) together with the above relations (2·25) and (2·26) one can easily arrive at Eq. (2·19).

2.3. *Including a kernel*

In previous subsections, a certain Langevin equation has been used to describe the relaxation dynamics which gives a Boltzmann-type thermal equilibrium distribution. As was pointed out in the introduction, this relaxation dynamics, however, is by no means unique. It is well known that one of these redundant degrees of freedom is described by a *kernel* $K(z)$. A Langevin equation with a kernel is written as

$$\frac{dz(t)}{dt} = -K(z)S'(z) + K'(z) + \xi(t)\,, \tag{2·28}$$

where the noise $\xi(t)$ obeys the following relations

$$\langle \xi(t)\rangle = 0\,, \qquad \langle \xi(t)\xi(t')\rangle = 2K(z)\delta(t-t')\,. \tag{2·29}$$

A particular realization of $\xi(t)$ is $\xi(t)=\sqrt{K(z)}\,\eta(t)$, where η is that given in (2·2). The kernel can either be a simple complex constant or a function of the field z, which is respectively called a constant kernel or field dependent kernel. In the case of a field dependent kernel, the noise $\xi(t)$ becomes multiplicative, for which we have to adopt a certain interpretation. In the Langevin equation (2·28), the Ito type interpretation has been taken.*)

As mentioned above, the naive expectation is that the Langevin equations (2·28) with different kernels give different types of relaxation processes, but the same equilibrium distribution. The reason for this can easily be understood from the corresponding Fokker-Planck equation

$$\partial_t P_c(x;\,t)=-H_{\rm FP}P_c(x;\,t)\,,$$
$$H_{\rm FP}=-\partial_x K(x)(\partial_x+S'(x))\,. \tag{2·30}$$

Due to the multiplicative operator (∂_x+S') in the Hamiltonian $H_{\rm FP}$ above, one can see that the Boltzmann-type distribution $P_{\rm eq}\propto e^{-S}$ is a stationary solution of the Fokker-Planck equation, i.e., $H_{\rm FP}P_{\rm eq}=0$, *independently on the choice of the kernel.* Therefore, if it is possible to adjust the kernel so that the above Fokker-Planck Hamiltonian will have a positive semi-definite eigenspectrum, we can expect the asymptotic behaviour of the solution at large fictitious time t

$$P_c(x;\,t)\xrightarrow[t\to\infty]{}P_{\rm eq}(x)\propto e^{-S(x)}\,. \tag{2·31}$$

This idea will be checked by numerical simulation in § 4.

2.4. *Associated real probability distribution*

In this subsection we discuss the real probability distribution $P(x,\,y;\,t)$ defined in (2·22) or (2·24).

Let us first consider the thermal equilibrium of Eq. (2·24)

$$\langle O(x)\rangle_S=\int dxdy\,\mathcal{O}(x+iy)e^{-S}\,. \tag{2·32}$$

It would be very interesting if one could find directly the real probability distribution $P_{\rm eq}(x,\,y)$. In such a case, based on the probability distribution $P_{\rm eq}(x,\,y)$ and using the usual Monte Carlo method one can calculate any desired average for a complex system.

A general solution for this problem is not known. The Fokker-Planck equation is, however, showing us a way to find out one of the solutions of Eq. (2·32).

The complex Langevin equation (2·28) represents two coupled Langevin equations for the real variables x and y,

$$\dot{x}=\Re\{-K(z)S'(z)+K'(z)\}+\Re\{\sqrt{K(z)}\,\eta\}\,, \tag{2·33}$$
$$\dot{y}=\Im\{-K(z)S'(z)+K'(z)\}+\Im\{\sqrt{K(z)}\,\eta\}\,. \tag{2·34}$$

*) If one wants to adopt the Stratonovich interpretation one has to modify the drift force in (2·28), see Ref. 16) for details.

In the sense of (2·22), these equations are equivalent to the Fokker-Planck equation

$$\partial_t P(x, y; t) = -H_{\mathrm{FP}} P(x, y; t) \tag{2·35}$$

with the Fokker-Planck Hamiltonian

$$\begin{aligned} H_{\mathrm{FP}} = & -\partial_x^2 (\Re\sqrt{K})^2 - \partial_y^2 (\Im\sqrt{K})^2 - 2\partial_x\partial_y \Re\sqrt{K}\Im\sqrt{K} \\ & + \partial_x \Re(-KS' + K') + \partial_y \Im(-KS' + K') . \end{aligned} \tag{2·36}$$

For a simple Gaussian model

$$S = \frac{1}{2}\sigma x^2 , \qquad \sigma = A + iB , \tag{2·37}$$

the stationary Fokker-Planck equation $H_{\mathrm{FP}}(x, y)P(x, y) = 0$ can be solved.[13),17),8)] For a constant kernel $K = 1$, the solution is given by

$$P_{\mathrm{eq}}(x, y) = \gamma \exp\left\{ -A\left[x^2 + \left(1 + 2\frac{A^2}{B^2}\right) y^2 + 2\frac{A}{B} xy \right] \right\} . \tag{2·38}$$

Here γ is a normalization constant. It can explicitly be shown that this solution satisfies also the relation (2·32).

It is much more interesting if one can directly see how the stochastic process relaxes to its equilibrium limit and how the kernel can influence the relaxation. For the above simple model, the exact solution of (2·35) with (2·36) has been given without and with a constant kernel.[17),10)] With

$$\sqrt{K} = \alpha + i\beta , \tag{2·39}$$

$$K\sigma = A + iB , \tag{2·40}$$

the solution which has an initial distribution $P_0(x, y)$ can be written as

$$P(x, y; t) = \gamma \int dx' dy' \exp\left\{ \frac{G_3 H_1^2 - 2G_2 H_1 H_2 + G_1 H_2^2}{G_2^2 - G_1 G_3} \right\} P_0(x', y') \tag{2·41}$$

with the abbreviations

$$\begin{aligned} G_1 &= u + w + e^{-2At}(-u\cos(2Bt) - v\sin(2Bt) - w) , \\ G_2 &= v + e^{-2At}(u\sin(2Bt) - v\cos(2Bt)) , \\ G_3 &= -u + w + e^{-2At}(u\cos(2Bt) + v\sin(2Bt) - w) , \\ H_1 &= x - e^{-At}(x'\cos(Bt) + y'\sin(Bt)) , \\ H_2 &= y - e^{-At}(-x'\sin(Bt) - y'\cos(Bt)) , \end{aligned} \tag{2·42}$$

where

$$u + iv = \frac{1}{\sigma} \qquad \text{and} \qquad w = \frac{|K|}{A} . \tag{2·43}$$

These formulae show that the time dependence of the distribution consists of oscillations multiplied by a factor $\exp\{-At\}$. Thus the convergence of the real probability

distribution to its equilibrium distribution is controlled by $A=\Re(K\sigma)$. If we choose $A>0$, $\exp\{-At\}$ is a damping factor and the limit $t\to\infty$ can be taken. The integrand in (2·41) becomes in particular independent of x' and y' and the integration can be performed making the distribution independent of the initial distribution $P_0(x, y)$. It is given by

$$P_{\mathrm{eq}}(x, y)=\gamma\exp\left\{\frac{1}{w^2-u^2-v^2}(-(w-u)x^2-(w+u)y^2+2vxy)\right\}. \tag{2·44}$$

Again, using Eq. (2·27) in the limit $t\to\infty$, the equilibrium form $P_c(x)=\exp\{-(\sigma/2)x^2\}$ is recovered from the above distribution.

It is interesting to specify the ordinary real probability distribution for the bottomless case with action S_1 and for the pure imaginary Minkowski case with action S_2, both with a phase rotating kernel,

$$S_1=e^{-|a|x^2/2},$$
$$S_2=e^{ibx^2/2}, \qquad K=e^{-i\theta}. \tag{2·45}$$

The distributions become

$$S_1: \quad P_{\mathrm{eq}}(x, y)=\exp\left\{|a|\cos\theta\left(\frac{x^2}{1+\cos\theta}+\frac{y^2}{1-\cos\theta}\right)\right\}, \tag{2·46}$$

$$S_2: \quad P_{\mathrm{eq}}(x, y)=\exp\left\{-\frac{b\sin\theta}{\cos^2\theta}(x^2+y^2+2xy\sin\theta)\right\}. \tag{2·47}$$

We have to choose the angle θ accordingly as to ensure that the probability density tends to zero at the boundaries. This is nothing other than the condition $A\equiv\Re(k\sigma)>0$ we met above, since $A_1=-|a|\cos\theta$ and $A_2=b\sin\theta$, respectively, for the two actions considered.

§ 3. Numerical simulation of complex Langevin systems

In this section we discuss numerical simulations of simple systems with complex actions.

As was mentioned before, for a complex action S, the Fokker-Planck Hamiltonian is no more self-adjoint and semi-positive definite. For such cases there exists no general spectral theorem.[12] Therefore we lose, in some sense, the theoretical basis by which we can a priori judge in what situation the simulation succeeds and in what situation not. Only one (but important) point which could help us may be, as discussed before, the positive semi-definiteness of the Fokker-Planck Hamiltonian if a suitable kernel is introduced.

In some situation, complex Langevin simulation turned out to fail completely, either the solution of the Langevin equation is unstable or, even worse, it converges to a completely wrong (or unphysical) one. However, whenever one meets the situation that the complex Langevin simulation gives correct results, the numerical gain obtained by this method is very big.

Because of the tremendous gains in case of success, many authors have tried to

investigate complex Langevin simulation in more detail for simple cases where the exact results are known. These investigations of simplified models are helpful in order to study the reasons for the success or failure. In this section we will discuss two typical models: The polynomial model where the action is a simple polynomial in x^2, and the $\cos\theta$-model, which is interesting in applying the method to the lattice gauge theory.

3.1. *The polynomial model*

This model has been discussed first by Klauder and Petersen.[12] The action is given by

$$S=\frac{1}{2}\sigma x^2+\frac{1}{4}\lambda x^4 \tag{3·1}$$

with σ and λ in general complex. Langevin equation for this system is given by

$$\dot{z}=-\sigma z-\lambda z^3+\eta\,. \tag{3·2}$$

From the general discussion in the previous section we may expect that a Langevin simulation should lead to the correct result if the real part of λ is positive and the imaginary parts of σ and λ are not too big. It is interesting to study the eigenvalues of the associated Fokker-Planck Hamiltonian (see Eq. (2·12)),

$$\tilde{H}_{\mathrm{FP}}=\frac{1}{4}(-4\partial_x{}^2-2\sigma-6\lambda x^2+\sigma^2x^2+2\sigma\lambda x^4+\lambda^2x^6)\,, \tag{3·3}$$

and in particular their movement with increasing imaginary part of the parameter σ or λ. In Ref. 12) the eigenvalues have been calculated numerically for three different cases for $\lambda=2$ and $\sigma=1+is$, $\sigma=is$, and $\sigma=-1+is$, with changing s from 0 to 40. The results are plotted in Fig. 1.

The trajectories show that in all three cases the real parts of the eigenvalues decrease with increasing s. However, in case of (a) ($\Re\sigma=1>0$) they remain positive while in case of (b) ($\Re\sigma=0$) they vanish for large s. In case of (c) where $\Im\sigma=-1$ is negative, the real parts of the eigenvalues vanish for some intermediate value of s and become negative for larger values of this parameter. For $\Re\sigma>0$, where the Fokker-Planck Hamiltonian exhibits positive semi-definite eigenvalues, the simulated results always agree with the exact ones within the error. We expect, therefore, that the Langevin simulation will fail for $\Re\sigma<0$ and sufficiently large $\Im\sigma$. It turned out that this is actually the case.[12),10] The situation is illustrated in Table I, where numerical results for $\langle z^2\rangle_\eta$ calculated by the use of the Langevin equation have been given for $\sigma=-1+is$ for $s=0, 1, \cdots$ and $\lambda=2$ fixed.

In calculating $\langle z^2\rangle$ in Table I, total fictitious time $T=100$ has been used for the average in each run, and 64 runs with different initial configurations have been taken for estimating the statistical errors.

One observes that the agreement of the simulated results with those from the direct integration of

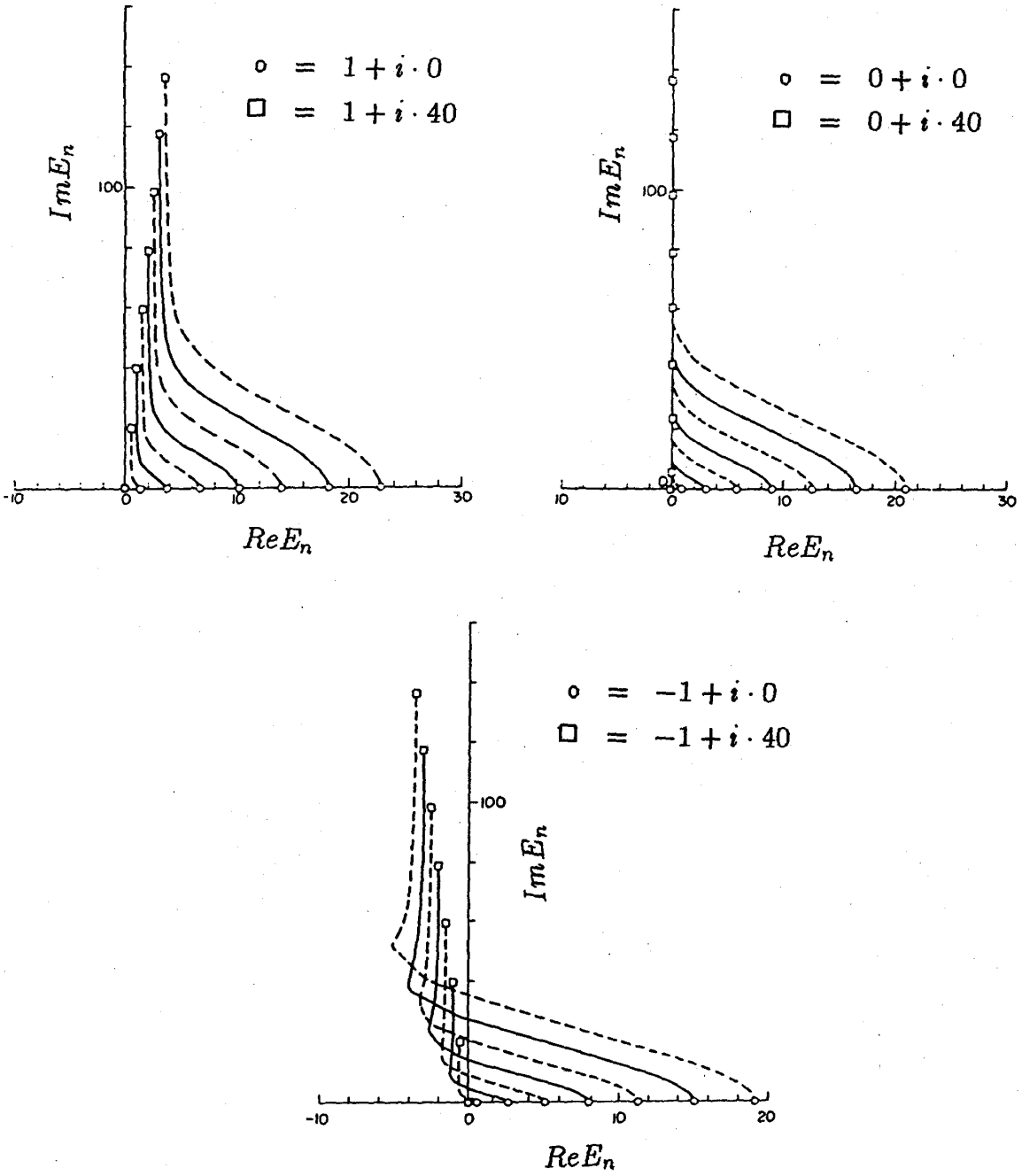

Fig. 1. Numerical results of eigenvalues E_n; $H_{FP}\psi_n=-E_n\psi_n$ for H_{FP} in (3·3) (Klauder and Petersen Ref. 12)).

Table I. Results of the Langevin simulation compared to the exact one in the case $\mathrm{Re}(\sigma)<0$.

σ	$\langle z^2\rangle$	$\langle z^2\rangle_{\mathrm{exact}}$
-1	$(0.627+0.000i)\pm(0.051+0.000i)$	$(0.645+0.000i)$
$-1+i$	$(0.620-0.197i)\pm(0.047+0.028i)$	$(0.605-0.198i)$
$-1+2i$	$(0.574-0.408i)\pm(0.048+0.065i)$	$(0.485-0.365i)$
$-1+4i$	$(0.522-0.973i)\pm(0.066+0.119i)$	$(0.107-0.441i)$
$-1+40i$	$(0.0096-0.315i)\pm(0.0062+0.076i)$	$(-0.0006-0.025i)$

$$\langle x^2\rangle=\frac{\int dx x^2 e^{-S(x)}}{\int dx e^{-S(x)}} \tag{3·4}$$

becomes worse if s increases.

An origin of the failure of the complex Langevin simulation observed in the above example has also been investigated from a somewhat different viewpoint: the *runaway solutions*. In case of a system with a real action, the configuration updated by using the Langevin equation is mainly confined to the neighbourhood of the minimum of the action. This is not always the case for complex actions. One sometimes encounters *runaway solutions,* i.e., in updating configurations z, using a fixed time step, the configuration gets more and more pushed towards large values of $|z|$ such that finally one meets overflow in the computer.

In order to understand more clearly the runaway solutions the Langevin equation (3·2) without noise term, has been investigated[13]

$$\dot{z}=-\sigma z-\lambda z^3 . \tag{3·5}$$

This can exactly by solved and the solution is

$$z(t)=\frac{z_0}{\sqrt{e^{2\sigma t}+\frac{\lambda}{\sigma}z_0^2(e^{2\sigma t}-1)}}, \qquad t\geq 0 \tag{3·6}$$

with $z_0=z(t=0)$. The asymptotic behaviour of this solution is quite different depending on whether $\Re\sigma>0$ or $\Re\sigma<0$.

For $\Re\sigma>0$ the solution tends to zero for $t\to\infty$, i.e., $z=0$ is a point of attraction. On the other hand, for $\Re\sigma<0$ the solution accumulates to "Higgs vacua", $\pm\sqrt{-\sigma/\lambda}$.

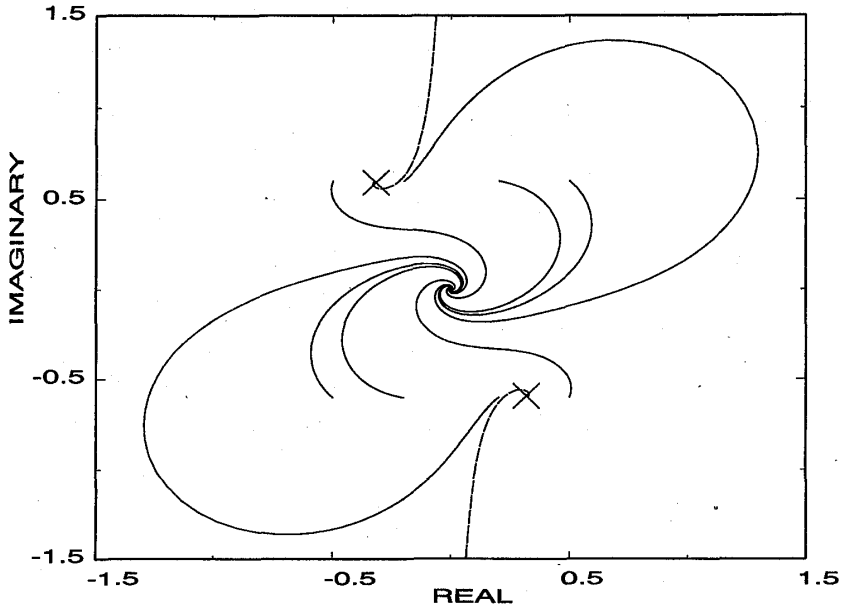

Fig. 2. Flow pattern according to Eq. (3·6) with $\sigma=0.5+0.75i$ and $\lambda=2$. The singular lines $z_0(\tau)$ starting at the values $\pm\sqrt{-\sigma/\lambda}$ have been plotted by broken lines.

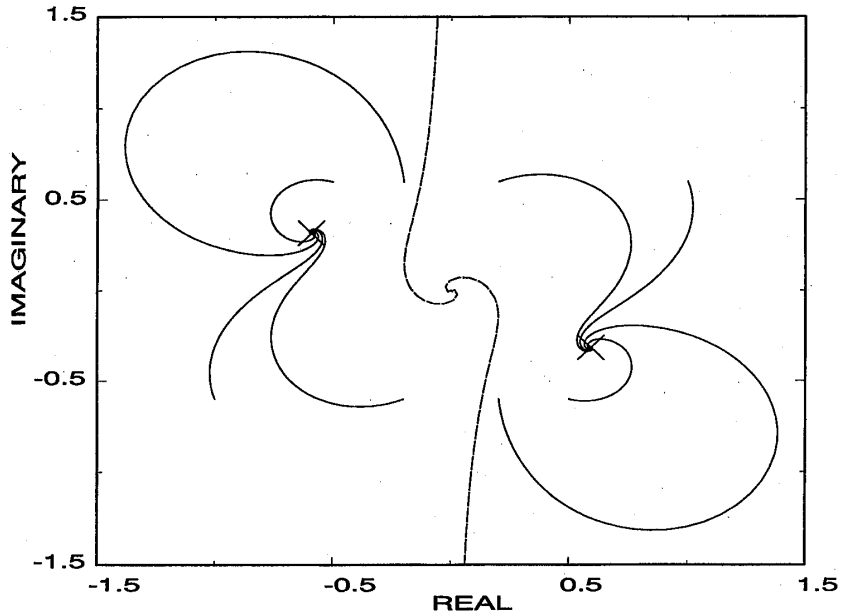

Fig. 3. Flow pattern according to Eq. (3·6) with $\sigma=-0.5+0.75i$ and $\lambda=2$. The singular lines $z_0(\tau)$ starting both at $z=0$ have been plotted by broken lines.

See Figs. 2 and 3.[13)]

In each case, however, there exists an exception if z_0 is chosen such that the denominator in Eq. (3·6) would vanish. This is when one starts from the initial points parametrized by a parameter τ as

$$z_0(\tau)=\pm\sqrt{\frac{-\sigma/\lambda}{1-\exp\{-2\sigma\tau\}}}. \tag{3·7}$$

The solution (3·6) starting from the above initial point is

$$z(t)=\pm\sqrt{\frac{-\sigma/\lambda}{1-\exp\{2\sigma(t-\tau)\}}}. \tag{3·8}$$

This shows that if one solves Eq. (3·5) starting from an initial point (3·7) which lies on the dotted lines in Figs. 2 and 3, the solution drifts along the line and diverges at time $t=\tau$. (Note that those two path's which start in the vicinity of the dotted lines in both figures are extending over very large values of z, until they finally reach their points of convergence.) In the case of the Langevin equation, i.e., with an additional random force η on the r.h.s. of (3·5), this runaway solution is, in principle, not a problem because the random force works to kick the configuration out of the lines. In the practical sense, however, a serious problem exists: When the solution tends to blow up, the drift term in the Langevin equation becomes very big. One needs, in such a case, a smaller time step, otherwise one is no more solving the differential equation. As the result a relatively big amount of computer time will be required to estimate the effect of those solutions which do a long trip in the region with very big imaginary part of z.

To throw away these runs is a simple way to circumvent the problem.[12)] This method has actually been adopted to get the results given in Table I. A justification of this method may be found in the following observation. As was discussed above,

when the configuration travels around the region which is far from the origin, the drift force becomes big. Therefore, the absolute time during which the configuration stays at large z should be small. Because of this reason the contribution from each blow-up solution to the average is expected to be small. Following this, whenever one encounters the solution in which $|z|$ becomes bigger than a certain bound, one discards this run and restarts from other initial point. If one solves the Langevin equation in this way and does not meet so many runaways then the simulation may give correct results. On the other hand, in case one meets a lot of runaway solutions the simulation will fail,[12),13)] see also Ref. 19).

3.2. *The cosθ model*

In the numerical simulation of lattice gauge theory, it sometimes happens that one has to calculate a correlation function between two Wilson loops which are separated from each other by a large distance. In such a case, usual Monte Carlo technique suffers from a large statistical error because such a quantity is exponentially small. One idea to escape from the problem is to include one of the Wilson loop factors $\mathrm{tr}(W)$ into the effective action, $\mathrm{tr}(W)e^{-S}\equiv e^{-S_{\mathrm{eff}}}$. By the use of this S_{eff}, one can generate a set of configurations whose elements follow the effective distribution $e^{-S_{\mathrm{eff}}}=\mathrm{tr}(W)e^{-S}$, by which one may simulate the correlation function more efficiently.

In order to investigate whether the idea really works well or not, a simple toy model which can, however, be a good representative of lattice gauge theory has been invented. The model has been investigated in detail by many authors.[13),20)~24)] Let us consider the following average of $\cos\theta$ whose result is known in an analytical form:

$$\langle\cos\theta\rangle=\frac{\int_0^{2\pi}d\theta\cos^2\theta\exp\{\beta\cos\theta\}}{\int_0^{2\pi}d\theta\cos\theta\exp\{\beta\cos\theta\}}=\frac{I_0(\beta)}{I_1(\beta)}-\frac{1}{\beta}. \tag{3·9}$$

One can rewrite this as

$$\langle\cos\theta\rangle=\frac{\int_0^{2\pi}d\theta\cos\theta\exp\{-S_1\}}{\int_0^{2\pi}d\theta\exp\{-S_1\}}, \tag{3·10}$$

where the effecive action S_1 has become complex,

$$S_1=-\beta\cos\theta-i\theta. \tag{3·11}$$

Following the general discussion in the previous sections the Langevin equation by which we can simulate (3·9) is given by

$$\dot{\theta}=-\beta\sin\theta+i+\eta. \tag{3·12}$$

A result of a simulation of $\langle\cos\theta\rangle_\eta$ for $\beta=0.1,\cdots,5$ is shown in Fig. 4 (open squares) together with the curve which represents the exact result (3·9). All the simulated values lie exactly upon the curve, which is showing a success of the complex Langevin simulation in this case.

On the other hand, the above expectation value (3·9) can also be rewritten as

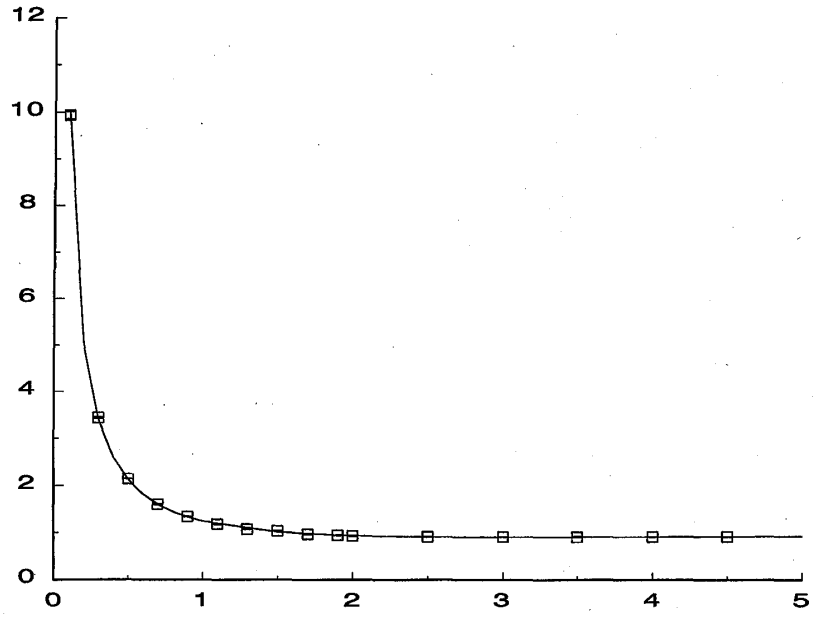

Fig. 4. The real part of $\langle\cos\theta\rangle_\eta$ obtained from the Langevin equation (3·12) for various values of β (open squares). The curve gives the exact values calculated from (3·9).

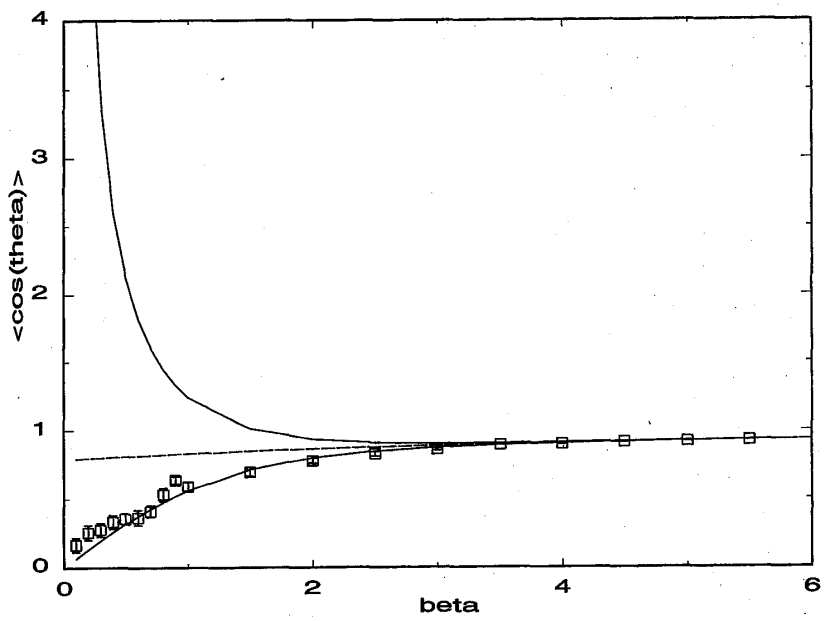

Fig. 5. The real part of $\langle\cos\theta\rangle_\eta$ obtained from the Langevin equation (3·15) for various values of β (open squares). The upper curve is calculated from Eq. (3·9), the lower curve is the result of Eq. (3·21), while the intermediate curve is calculated from Eq. (3·22)) (segregation).

$$\langle\cos\theta\rangle=\frac{\int_0^{2\pi}d\theta\cos\theta\exp\{-S_2\}}{\int_0^{2\pi}d\theta\exp\{-S_2\}} \tag{3·13}$$

with an effective action S_2 given by

$$S_2=-\beta\cos\theta-\ln(\cos\theta)\,. \tag{3·14}$$

The expectation value should, therefore, also be obtained from a solution of the

corresponding Langevin equation

$$\dot{\theta}=-\beta\sin\theta-\tan\theta+\eta\,. \tag{3·15}$$

A curious but very interesting point found is that, in the strong coupling region, i.e., for small β ($\beta\lesssim 2.5$), the Langevin equation (3·15) fails completely to give the correct result, as can be seen from Fig. 5. The exact analytical result (3·9) indicates that the expectation value $\langle\cos\theta\rangle$ should rise proportional to $1/\beta$ for $\beta\sim 0$. The simulated result, however, fails completely to give this $1/\beta$ singularity. In order to obtain such a behavior the quantity θ must exhibit an appreciable imaginary part during the updation procedure. If the configuration generated by the Langevin equation always stays near the real axis we can never expect this kind of singularity.

For the case of S_1, the big imaginary part of the solution can be expected from the following simple argument. The stationary points of the drift force of the corresponding Langevin equation (3·12) are given by

$$\theta=\theta_n\equiv 2n\pi+i\xi\,,\qquad \sinh\xi=\frac{1}{\beta}\,. \tag{3·16}$$

It is possible to show that these stationary points are attractive. For this, one calculates the drift for small deviation $\delta\theta=\theta-\theta_n$ around the point $\theta=\theta_n$. The result is

$$-\beta\cosh\theta_n\delta\theta=-\sqrt{1+\beta^2}\,\delta\theta\,, \tag{3·17}$$

which proves the above statement.

For the case of S_2, however, the situation becomes different. This can be seen by repeating the same analysis for the Langevin equation (3·15). The drift term has many stationary points but mostly on the real axis. Only in the case $\beta<1$, one finds stationary points whose imaginary part is non-trivial,

$$\theta_n=(2n+1)\pm i\xi\,,\qquad \cosh\xi=\frac{1}{\beta}\,. \tag{3·18}$$

The drift around these points, however, becomes

$$-\{\beta(\cos\theta_n+1)+\tan^2\theta_n\}d\theta=-(\beta+2)(\beta-1)\delta\theta \tag{3·19}$$

showing us that the stationary points (3·18) are not stable because $-(\beta+2)(\beta-1)>0$ for $\beta<1$. The configuration cannot, therefore, stay in the imaginary region for a long period. Therefore, the simulaion fails.

The difference of the drift pattern for the two cases discussed above can clearly be seen from Figs. 6 and 7.

This is, however, only an empirical reason why the simulation succeeds in one case and fails in the other case. The more fundamental reason for the failure may be found in the non-analyticity of the action S_2, Eq. (3·14). When θ varies, $\cos\theta$ changes its sign at $\theta=\pi/2$ and therefore the imaginary part of the action abruptly jumps from 0 to $i\pi$. It seems to be difficult to pass this behaviour to the Langevin equation, which turned out to be the principal reason of the failure.

In order to check this point, let us consider somewhat different model which does

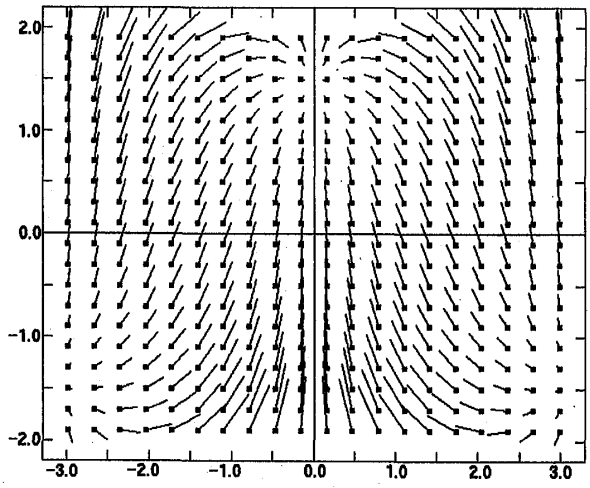

Fig. 6. Drift pattern for the drift of the Langevin equation (3·12) which corresponds to the action (3·11), $S_1=-\beta\cos\theta-i\theta$, for $\beta=0.5$. The point $\theta=1.32i$ is a point of attraction.

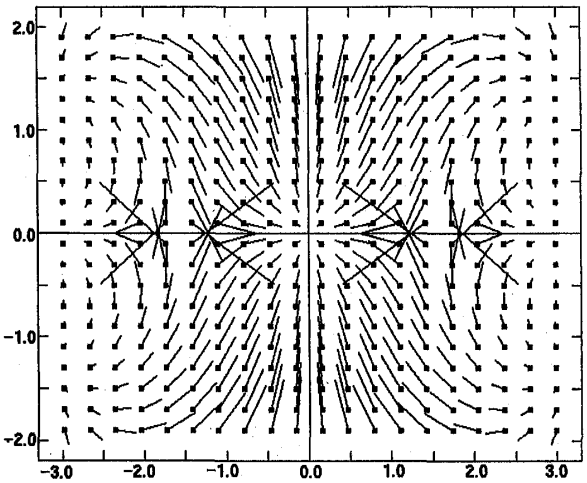

Fig. 7. Drift pattern for the drift of the Langevin equation (3·15) which corresponds to the action (3·14), $S_2=-\beta\cos\theta-\ln(\cos\theta)$, for $\beta=0.5$. The point $\theta=0$ is a point of attraction.

not have this discontinuity,[23]

$$S_a=-\beta\cos\theta-\ln(|\cos\theta|)\,. \tag{3·20}$$

For real θ, this action differs from S_2 in the region of $\cos\theta<0$. The difference, however, is only a constant $i\pi$, which will not give any additional contribution to the drift force $\delta S/\delta\theta$ in the Langevin equation. We claim that the Langevin equation (3·15) simulates a system with the (real) action S_a in (3·20).

In fact,[23),24)] the result of $\langle\cos\theta\rangle_\eta$ simulated by the Langevin equation (3·15) turns out to be equal to an average over the Boltzmann distribution given by the action S_a but not by S_2, which is

$$\langle\cos\theta\rangle_{S_a}=\frac{\int_0^\pi d\theta\cos\theta|\cos\theta|\exp\{\beta\cos\theta\}}{\int_0^\pi d\theta|\cos\theta|\exp\{\beta\cos\theta\}}\,. \tag{3·21}$$

The lower curve in Fig. 5 is showing the values predicted from this formula.

In the literature,[21] an alternative discussion can be found about the reason why the Langevin equation (3·15) fails to give the desired results (3·9). There, a violation of the ergodicity due to the so-called segregation theorem is blamed for that. The theorem states that in the case of a real stochastic process, if there exists a zero of the probability distribution, $P(\theta=\theta_0)=0$, the configuration cannot pass through the point $\theta=\theta_0$ within a finite time interval. Actually, in the above example, by the zeros of S_2 the whole interval $[0, 2\pi]$ is divided into two regions $[\pi/2, 3\pi/2]$ and $[0, \pi/2)\oplus(3\pi/2, 2\pi]$. (Note that $\theta=0$ is equal to $\theta=2\pi$ because of the periodicity.) If one performs the simulation on the real axis, therefore, the updated configuration will stick to one of the above two regions and give the average, e.g.,

$$\langle\cos\theta\rangle_{\text{seg}}=\frac{\int_0^{\pi/2}d\theta\cos^2\theta\, e^{\beta\cos\theta}}{\int_0^{\pi/2}d\theta\cos\theta\, e^{\beta\cos\theta}}. \tag{3·22}$$

The dotted line in Fig. 5 is calculated from this formula.

The results of a simulation follow closely the curve drown from Eq. (3·21) rather than that from Eq. (3·22). Details about this point will be reported elsewhere.[24]

In order to confirm the fact that the failure of the complex Langevin equation observed for the action S_2 is really due to the sudden appearance of the imaginary part, let us consider another example with a somewhat different action,

$$S_3=-\beta\cos\theta-\ln(\cos^2\theta). \tag{3·23}$$

The structure of this action is quite similar to that of S_2, except for that $\cos^2\theta$ does never become negative and therefore no imaginary part is arising from $\ln(\cos^2\theta)$. It can easily be shown that

$$\langle\cos\theta\rangle_{S_2}=1/\left\langle\frac{1}{\cos\theta}\right\rangle_{S_3}. \tag{3·24}$$

The Langevin equation with the action S_3, i.e.,

$$\dot{\theta}=-\beta\sin\theta-2\tan\theta+\eta \tag{3·25}$$

can be used to simulate the average $\langle\cos\theta\rangle$ for the system given by an action S_2. As the discontinuity in the imaginary part of the action existing in S_2 is not present in S_3, one can expect to get the correct result for the average.[23] Actually this turns out to be true, which can be seen from Fig. 8.

In this secion, we saw the typical success and failure of the complex Langevin simulation by the use of simple toy models with one degree of freedom. In the

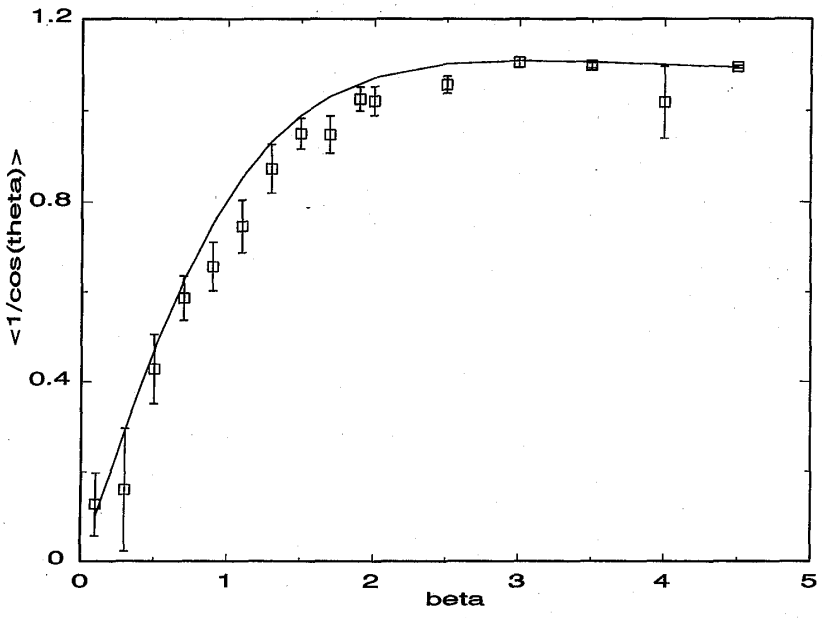

Fig. 8. The average $\langle 1/\cos\theta\rangle$ calculated from the Langevin equation (3·25) (squares). The curve represents the exact results.

discussion of more practial problems in lattice gauge theory (see § 5 for details), however, one encounters quite similar situations as those in this section. Roughly speaking, the complex action in $U(1)$ gauge theory has a quite similar structure to S_1. A beautiful success and a big numerical gain have been reported in this case for space time dimension $d=2$. On the other hand, the complex action in non-Abelian gauge theory has a similar structure as S_2, and it has been claimed that the complex Langevin simulation fails completely for $SU(2)$ and $SU(3)$ models.[25),20)]

§ 4. Kernel control

As we have discussed in § 2, the Langevin equation admits the use of a kernel without changing the thermal equilibrium of the corresponding Fokker-Planck distribution. In fact, according to Eq. (2·30) one can see that a distribution proportional to $\exp\{-S\}$ satisfies the stationality condition, $H_{\text{FP}}\exp\{-S\}=0$, independently of what kind of function is assigned for the kernel. One can, therefore, make use of the freedom to choose the kernel in order to accelerate the convergence or to improve convergence in the cases where the naive Langevin equation fails. Two types of kernels, constant kernel and field dependent kernel, will be discussed in the following two subsections separately.

4.1. *Constant kernel and solutions in different Riemann sheets*

In order to see explicitly the role of the kernel, we go back to the discussion of the polynomial model. As we saw in § 3.1, the Langevin simulation fails in situations where the eigenvalues of the associated Fokker-Planck equation exhibit negative real parts. The use of a kernel can change this situation: If we choose a kernel

$$K=e^{-i\phi}, \tag{4·1}$$

ϕ being a constant phase, it rotates the eigenvalues of the Fokker-Planck Hamiltonian by the angle $-\phi$ in the complex plane. An inspection of Figs. 1(b) and (c) shows that by a suitably chosen angle it can be achieved that the real parts of all the eigenvalues become positive. The use of a constant kernel of this type in the polynomial model has been investigated in detail in Ref. 10).

4.1.1. A complex Gaussian model

For simplicity let us start to discuss a Gaussian model with the action,

$$S_2=\frac{1}{2}\sigma x^2, \tag{4·2}$$

where σ may have any value in the complex plane. The Minkowski action (σ purely imaginary) and the bottomless action ($\Re\sigma<0$) are included as special cases. We consider the average of x^2, which can exactly be calcuated in the case $\Re\sigma>0$,

$$\langle x^2\rangle=\frac{\int dx x^2\exp\{-\sigma x^2/2\}}{\int dx\exp\{-\sigma x^2/2\}}=\frac{1}{\sigma}. \tag{4·3}$$

The above result can analytically be continued to the whole complex plane ($\sigma \neq 0$), even though the integrals in the numerator and in the denominator do not converge separately for $\Re\sigma < 0$. With the constant kernel K the corresponding Langevin equation reads (see Eq. (2·28))

$$\dot{z} = -K\sigma z + \sqrt{K}\,\eta\,. \tag{4·4}$$

The Fokker-Planck Hamiltonian for this system is given by

$$H = KH_0\,, \qquad H_0 = -\partial_x(\partial_x + S')\,. \tag{4·5}$$

After a similarity transformation in (2·12), it is nothing other than that of the harmonic oscillator. One can, therefore, exactly solve its eigenvalue problem ($Hf_n = E_n f_n$) and the result for E_n is

$$E_n = nK\sigma\,. \tag{4·6}$$

From this form, it is clear that one can easily make the real part of the eigenvalues positive by a suitable choice of the kernel.

Although Eq. (4·4) is exactly solvable, there remains some interest in doing numerical simulations to check whether the above idea is really working or not. Let us consider, therefore, the system with action (4·2) for different values σ chosen to lie on the unit circle. By a naive Langevin equation ($K \equiv 1$) the simulation is only possible for those systems where σ lies in the right half of the complex plane ($\Re\sigma > 0$). For $\Re\sigma \leq 0$, the numerical solution of the Langevin equation always blows up and one cannot get any stable results.

In contrast to that, it has been checked that the Langevin equation with a kernel

$$K = \sigma^* \tag{4·7}$$

allows to perform a numerical simulation for *all* values of σ. It is interesting to remark that the numerical results of $\langle z^2 \rangle_\eta$ obtained turned out to coincide with those values predicted by the analytical formula $1/\sigma$ even in the region $\Re\sigma < 0$.

It should be noted that the result $1/\sigma$ for $\Re\sigma \leq 0$ is defined only in the sense of analytic continuation and its physical meaning is not clear. It is very interesting that in spite of this fact the Langevin equation allows us to get a sample $\{z_1, z_2, \cdots\}$ which has an average $\langle z^2 \rangle_\eta$ coinciding with that value.

4.1.2. Unphysical solution and λx^4-model

Before starting the discussion of the general polynomial model, we give here a notion of an *unphysical solution*. A Langevin simulation with a complex kernel turned out to pick up unphysical solutions, by which one sometimes arrives at a result far from the desired one. In order to explain the unphysical solution we discuss here the following power model,

$$S_4 = \frac{1}{4}\lambda x^4\,. \tag{4·8}$$

For this action the average $\langle x^2 \rangle$ can also be calculated exactly for $\Re\lambda > 0$,*)

*) The analysis can be performed analogously for models with actions with a general power x^n, where the average $\langle x^2 \rangle$ is multi-valued, $\langle x^2 \rangle \propto (1/\lambda)^{2/n}$, see Refs. 10) and 23).

$$\langle x^2\rangle=\gamma\sqrt{\frac{1}{\lambda}}, \qquad \gamma=2\frac{\Gamma\left(\frac{3}{4}\right)}{\Gamma\left(\frac{1}{4}\right)}. \tag{4·9}$$

The result can be continued analytically into the first Riemann sheet of λ, where the branch cut is taken along the negative imaginary axis, i.e.

$$-\pi<\arg(\lambda)\equiv\phi_0<\pi\,. \tag{4·10}$$

This restriction on the argument of λ makes the r.h.s. of Eq. (4·9) unique. This is required from the fact that $\langle x^2\rangle$ on the l.h.s. of Eq. (4·9) is a unique quantity which should be positive when λ becomes real. If this restriction is not given, (4·9) gives a two-fold solution corresponding to λ in the first Riemann sheet (the physical solution) and second Riemann sheet (unphysical solution).

An interesting point to note is that *the Langevin equation with a constant kernel picks up both of the solutions*. In order to show this, let us use the notation

$$\lambda_k=|\lambda|e^{i(\phi_0+2k\pi)}, \qquad k=0,1 \tag{4·11}$$

to specify λ in the first ($k=0$) and second ($k=1$) Riemann sheet. The Langevin equation for this model is

$$\dot{z}=-K\lambda z^3+\sqrt{K}\,\eta\,. \tag{4·12}$$

Dividing both sides of Eq. (4·12) by $\sqrt{K}$ and rewriting it by introducing $r\equiv K^{-1/2}z$, we get

$$\dot{r}=-K^2\lambda r^3+\eta\,. \tag{4·13}$$

Since η is taken to be real, a stable solution is guaranteed if the coefficient of the drift

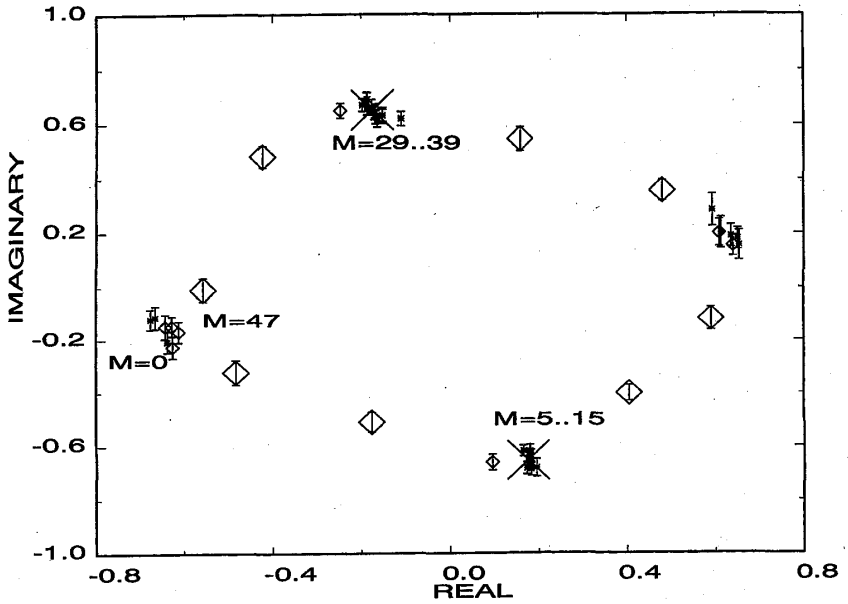

Fig. 9. Average $\langle z^2\rangle_\eta$ obtained from the Langevin equation (4·12) with fixed $\lambda=e^{(5/6)\pi i}$ and $K=e^{-(M/24)\pi i}$, $M=0,\cdots,47$.

term is adjusted to be real and positive. We have two solutions for the kernel K which satisfy this condition, that is

$$K \equiv K_k = e^{-i(\phi_0/2+k\pi)}, \qquad k=0, 1. \tag{4·14}$$

By this choice of the kernel we obtain real and stable solutions for r. Using the relation between r and z, we conclude that the Langevin equation has two solutions $z=z_k$,

$$z_k^2 = r^2 K_k = r^2 e^{-i(\phi_0/2+k\pi)} = z_0^2 e^{ik\pi}, \qquad k=0, 1. \tag{4·15}$$

Here, z_0 corresponds to the physical solution and z_1 to the unphysical one.

From the above discussion it is already clear why the Langevin equation with a complex kernel picks up not only the physical solution but also the unphysical one. This may, however, more vividly be seen from the result of numerical simulation. In Fig. 9 results for $\langle z^2 \rangle_\eta$ for the model (4·8) with $\lambda = e^{(5/6)\pi i}$ have been given.

Each point in this figure corresponds to the result for $\langle z^2 \rangle_\eta$ simulated by the use of the Langevin equation (4·12) with a certain kernel K. There, kernels which are parametrized as

$$K = e^{-(M/24)\pi i}, \qquad M=0, \cdots, 47, \tag{4·16}$$

have been used. For this λ, and from the general discussion above (Eq. (4·14)), one can predict which kernel leads to the physical or unphysical stable result. It is, for the physical one, $K=K_0 \equiv \exp\{-5\pi i/12\}$, i.e. $M=10$, while for the unphysical one $K=K_1 \equiv \exp\{-17\pi i/12\}$, i.e. $M=34$. This is indeed the case as can be read off from Fig. 9. When one changes the kernel from $M=0$ to $M=4$, the corresponding result moves in the complex $\langle z^2 \rangle_\eta$-plane counter-clockwise to approach the exact result whose point is marked by a cross in the figure. For a broad range of kernels, however, around $K \sim K_0$, i.e., for $M=5, \cdots, 15$, the simulated results accumulate around the exact

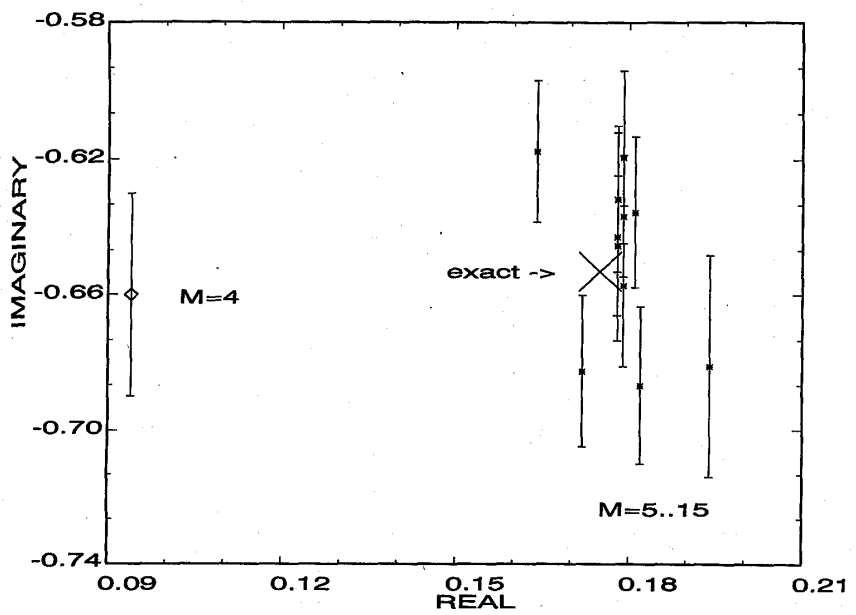

Fig. 10. Enlarged view on the cluster of results in Fig. 9 for $M=4, \cdots, 15$.

physical value. Corresponding results for $M=16,\cdots,28$ once more scatter around the complex plane, while the results cluster again also for a broad range of kernels around $M=29,\cdots,39$. This is exactly the region of $K\approx K_1$, and the result obtained is the unphysical one. In order to show the clustering of values around the physical result, a plot in an enlarged scale around that region is given in Fig. 10.

For each of the result in Fig. 9, 2000 iterations of the Langevin equation with a fixed time step $dt=0.01$ have been performed. An average and the statistical error have been estimated by repeating 20 times the same kind of runs with different initial points. When one meets blowups the run has been discarded and replaced with a new run. For the case where one meets only a small number of blowups (<10), the result has been drawn by a small diamond. For the case where one meets a larger number of blowups (up to 44), the result has been drawn by a big diamond.

4.1.3. Polynomial model

Let us now apply the idea of using a kernel to the polynomial model (3·1). The fundamental idea has already been explained in the first part of § 4.1: by the constant kernel (4·1) the spectrum of the Fokker-Planck Hamiltonian is just rotated in a complex plane by an angle $-\phi$. If it is possible to choose the angle ϕ such that all of $\Re(K\lambda_n)$ become positive (see Fig. 1), we might expect to get convergence and a correct result. It is interesting to check this simple idea. In Fig. 11 (for $\lambda=2$ and $\sigma=4i$) and in Fig. 12 (for $\lambda=2$ and $\sigma=-1+4i$), results of the average $\langle z^2\rangle_\eta$ obtained using the kernel

$$K=e^{-i(M\pi/24)}, \qquad M=0,\cdots,24 \tag{4·17}$$

are shown.

Both figures show typical examples in which the naive Langevin simulation, which corresponds to the case $M=0$, fails. The open diamonds are the real part of

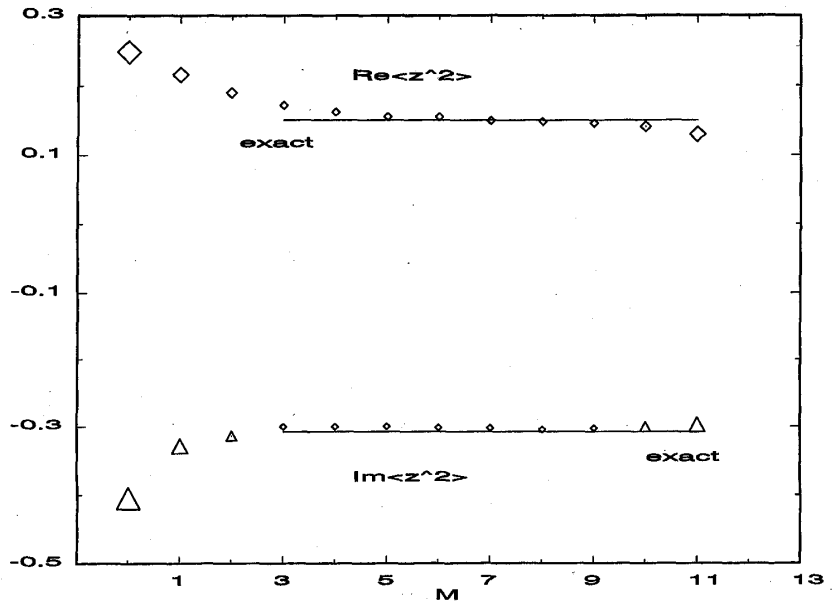

Fig. 11. Results of a simulation of $\langle z^2\rangle_\eta$ versus M for the polynomial model with $\lambda=2$ and $\sigma=4i$. The open diamonds and triangles are the real and imaginary part of a simulation with constant kernel.

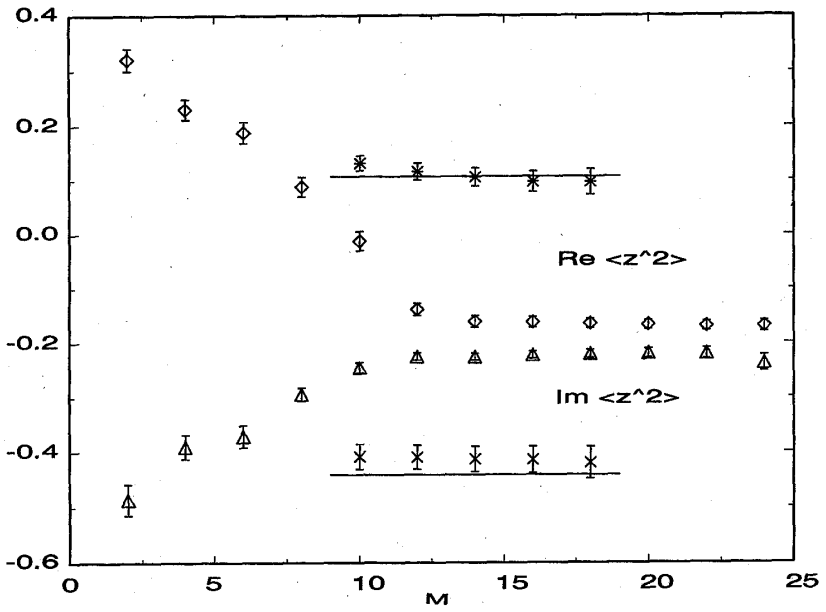

Fig. 12. Results of a simulation of $\langle z^2\rangle_\eta$ versus M for the polynomial model with $\lambda=2$ and $\sigma=-1+4i$. The open diamonds and triangles are the real and imaginary part of a simulation with constant kernel, crosses that for a simulation with the field dependent kernel (4·20). The lines show the exact results.

$\langle z^2\rangle_\eta$, the triangles the imaginary part. The solid lines represent the exact results from direct integrations.

For the case $\lambda=2$ and $\sigma=4i$ (see Fig. 11) one can see that there exists a broad range of M for which the simulated results are stable. Within this range the result coincides with the exact one. This is a typical success of the idea.

We sometimes encounter, however, a mysterious situation where the simulation fails completely. A typical example of the failure can be seen in Fig. 12, where $\lambda=2$ and $\sigma=-1+4i$. In this figure, the simulated results also converge very nicely for $M=12, \cdots, 24$. The results, however, are different from the exact one.

Here we again encounter the *unphysical solution in the second Riemann sheet.* In the case of the polynomial model, the average $\langle x^2\rangle$ cannot, in general, be calculated analytically. It is, however, possible to perform a perturbative calculation of the integrals by expanding $e^{-(\sigma/2)x^2}$.[10]

$$\langle x^2\rangle=\frac{\int dx x^2 e^{-(\sigma/2)x^2-(\lambda/4)x^4}}{\int dx e^{-(\sigma/2)x^2-(\lambda/4)x^4}}=\frac{4\sqrt{\lambda}N_1\Gamma\left(\frac{3}{4}\right)-\sigma N_2\Gamma\left(\frac{1}{4}\right)}{-2\sqrt{\lambda}\,\sigma D_1\Gamma\left(\frac{3}{4}\right)+2\lambda D_1\Gamma\left(\frac{1}{4}\right)}, \tag{4·18}$$

where the abbreviations N_1, N_2, D_1, D_2 are, with $\kappa=\sigma^2/(4\lambda)$, defined by

$$N_1=\sum_{n=1}^{\infty}\frac{\kappa^n}{(2n)!}\prod_{i=0}^{n}(4i-1), \qquad N_2=\sum_{n=0}^{\infty}\frac{\kappa^n}{(2n+1)!}\prod_{i=0}^{n}(4i+1),$$

$$D_1=\sum_{n=1}^{\infty}\frac{\kappa^n}{(2n+1)!}\prod_{i=1}^{n}(4i-1), \qquad D_2=\sum_{n=1}^{\infty}\frac{\kappa^n}{(2n+1)!}\prod_{i=1}^{n}(4i-3). \tag{4·19}$$

In this formula, the physical result is obtained by taking λ within the first Riemann

sheet, while λ from the second sheet gives the unphysical one. Taking, for example, the value $\sigma=-1+4i$ and $\lambda=2e^{i0}$, we obtain $\langle x^2\rangle=0.107-0.441i$ which is equal to the exact result predicted by a direct integration. If we take, on the other hand, $\lambda=2e^{2\pi i}$, we arrive at the different result $\langle x^2\rangle=-0.160-0.223i$. This is nothing other than the value of convergence observed in Fig. 12. We can, therefore, conclude that the stable solution obtained by the use of the constant kernel in the range of $M=12, \cdots, 24$ is nothing other than the unphysical solution. It is worthwhile to mention that even if we use a kernel with bigger rotation angles, $M=25, \cdots, 48$, no other stable solution has been obtained in this case.

One may ask why, in the above example, only the unphysical solution has been picked up. In order to get some insight for the answer to this question, let us recall the discussion of the simple power models S_2 and S_4 given in the previous two subsections. For each model, we can establish the best choice of the kernel leading to maximum convergence of the Langevin simulation. Let us simply apply these results separately to the quadratic part $(\sigma/2)x^2$ and the quartic part $(\lambda/4)x^4$ of the polynomial model. Let us investigate whether the range of the kernel which is adjusted for the quadratic part of the model overlaps with the (physical or unphysical) range of the kernel adjusted for the quartic one or not. It is also natural to wonder whether this kind of argument does have some sense or not. This is because the Langevin equation, in general, is a nonlinear equation and the effect of the kernel to its solution is expected to be not simply linear. It turned out, a posteriori, however, that the simple argument is helpful at least for the polynomial model.

We are now interested in the polynomial model with $\sigma=-1+4i$ and $\lambda=2$. Following the discussions in the previous subsections, for the $(\sigma/2)x^2$ part with $\sigma=-1+4i$, the kernel around $K\approx K_\sigma\equiv\exp\{-i\phi_\sigma\}$ with $\phi_\sigma=\pi/2+\tan^{-1}(1/4)\approx 104^\circ$ (stable zone for σ) leads to the best convergence. On the other hand, for the $(\lambda/4)x^4$ part with $\lambda=2$, $K\approx K_{\lambda,\mathrm{phys}}\equiv\exp\{-i\phi_{\lambda,\mathrm{phys}}\}=1$, or $\phi\approx\phi_{\lambda,\mathrm{phys}}=0$ (physically stable zone for λ) gives the best convergence for the physical solution, while $K\approx K_{\lambda,\mathrm{unphys}}\equiv\exp\{-i\phi_{\lambda,\mathrm{unphys}}\}=-1$, or $\phi_{\lambda,\mathrm{unphys}}=180^\circ$ (unphysically stable zone for λ) for the unphysical solution. An important point to be mentioned is that the stable zone for σ rather overlaps with the unphysical solution. It is also very probable that we cannot get rid of this situation when we stick to the constant kernel (phase rotation with a constant angle which is independent on the field z).

4.2. *Field dependent kernel*

At the end of the last subsection we have seen how the (constant) kernel works effectively for the full polynomial model, if the parameters σ and λ are such that the stable zone for σ overlaps with the physically stable zone for λ. This is not always the case. As a nice example we also have discussed the polynomial model with a particular choice of the parameters σ and λ for which the simulation leads to only a stable but unphysical result (Fig. 12).

The problem can be circumvented by the use of a field dependent kernel.[11] The idea is to choose a kernel which becomes approximately equal to K_σ in those regions where the term $(\sigma/2)x^2$ dominates in the action, but becomes equal to $K_{\lambda,\mathrm{phys}}$ in regions where the term $(\lambda/4)x^4$ dominates in the action. For real x, the σ-term dominates in

the region of small x^2, while the λ-term dominates in the region of large x^2. Therefore, a reasonable choice of the field dependent kernel which satisfies the above condition should have the form

$$K(x)=C_\sigma f(x^2)e^{-i\phi_\sigma}+C_\lambda[1-f(x^2)]e^{-i\phi_\lambda}, \tag{4·20}$$

where C_σ and C_λ are constants, and $f(x^2)$ has the property

$$f(x^2)=\begin{cases}1 & \text{for small } x^2,\\ 0 & \text{for large } x^2.\end{cases} \tag{4·21}$$

The practical form of $f(x^2)$ should, in principle, be irrelevant for the results in the thermal equilibrium limit. For the sake of efficiency of the simulation, however, we found that the following form is preferable,

$$f(x^2)=e^{-(x^2/a)}, \tag{4·22}$$

where a is the position of the "Higgs vacuum" given by $a=-\sigma/\lambda$, and with $C_\sigma=1/|\sigma|$, $C_\lambda=1/|\lambda|$.

Let us apply this field dependent kernel to solve the problem of the unphysical solution found in the polynomial model with $\sigma=-1+4i$, and $\lambda=2$. In this case, ϕ_σ and $\phi_{\lambda,\mathrm{phys}}$ are, respectively, 104° and 0°. In order to show that the idea really works, the simulation of $\langle z^2\rangle$ has been done with a fixed $\phi_\lambda=0$ ($\phi_{\lambda,\mathrm{phys}}$) but varying ϕ_σ; i.e.,

$$\phi_\sigma=\frac{\pi}{24}M, \qquad M=1,\cdots,24; \qquad \phi_\lambda=0. \tag{4·23}$$

The results obtained are drawn in Fig. 12, which looks completely different from those results obtained by using the constant kernel. The simulated real and imaginary parts of $\langle z^2\rangle$ for $M=10,\cdots,16$ are lying exactly on the solid lines which represent the exact physical values.

Let us show another interesting results for this model. In the above simulation, we can fix $\phi_\sigma=104°$, but change ϕ_λ in a broad range which covers not only $\phi_{\lambda,\mathrm{phys}}=0$ but also $\phi_{\lambda,\mathrm{unphys}}=180°$ i.e.,

$$\phi_\sigma=14\frac{\pi}{24}, \qquad \phi_\lambda=\frac{\pi}{24}N, \qquad N=0,\cdots,47. \tag{4·24}$$

The results of this simulation are shown in Fig. 13.

It is interesting to observe that now the simulation reproduces the correct physical results for values of N near the value expected for λ in the first Riemann sheet, namely $\phi_\lambda\approx\phi_{\lambda,\mathrm{phys}}=0$ ($0\leq N\leq 8$ and $40\leq N\leq 47$), but the unphysical result for $\phi_\lambda\approx\phi_{\lambda,\mathrm{unphys}}=\pi$ ($16\leq N\leq 32$). In between these regions there are some small bands of instability where solutions do not exist or are unreliable. For this figure each point has been obtained from 64 runs and each run consisting of 4000 updations with a fictitious time step of $dt=0.005$.

A different choice for the function $f(x^2)$, e.g.,

$$f(x^2)=\frac{a^2}{x^2+a^2} \tag{4·25}$$

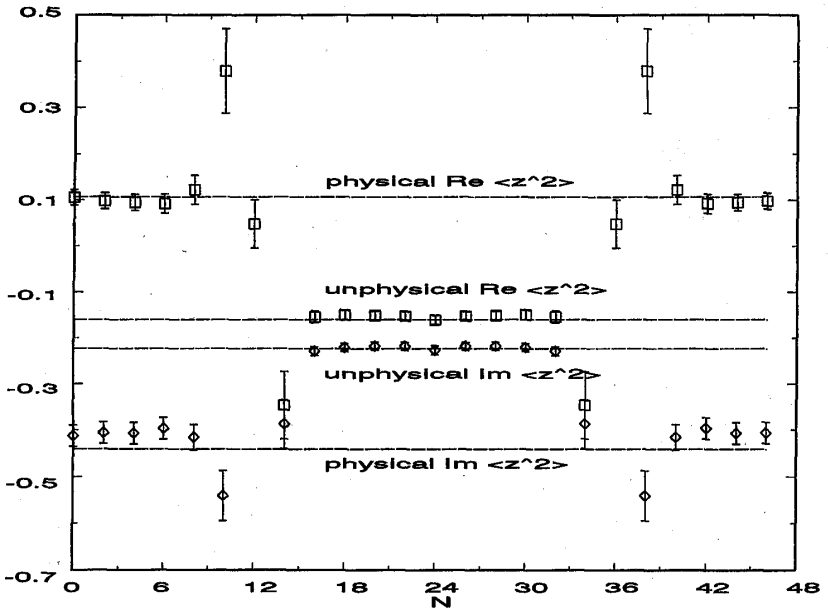

Fig. 13. Results of a simulation of $\langle z^2 \rangle$ versus N for the polynomial model with $\lambda=2$ and $\sigma=-1+4i$. The open squares and diamonds are the real and imaginary part of a simulation with the field dependent kernel (4·20), (4·24). The dotted lines show the exact results.

has also been tried. The efficiency seems not so good as that from the choice (4·22); the stable regions are smaller and the sudden jump from the physical to the unphysical result seen in Fig. 13 changed to a more smooth transition. The results obtained, however, are again showing the success of the idea of using the field dependent kernel.

The field dependent kernel even works in more pathological examples, where the action is linear or cubic in the variable inspired by topological field theories or Chern-Simons theories. The aim is to calculate the average

$$\langle \mathcal{O}(x) \rangle = \frac{\int dx\, \mathcal{O}(x) \exp\{-S(x)\}}{\int dx \exp\{-S(x)\}} \tag{4·26}$$

by Langevin simulation, with $S(x)$ linear or cubic in x.

In case

$$S_1(x) = \sigma x \tag{4·27}$$

one gets from Eq. (4·26)

$$\langle x^n \rangle = \frac{1}{Z(\sigma)} (-\partial_\sigma)^n Z(\sigma) = \frac{n!}{\sigma^n} \tag{4·28}$$

where we have defined by $Z(\sigma)$ the denominator in (4·26). The integral defining Z does not exist a priori, but it can suitably be regularized in case of pure imaginary $\sigma = i\sigma_I$. Since the integration extends from minus infinity to plus infinity, we write formally

$$Z(i\sigma_I) = \int dx e^{-i\sigma_I x} = \frac{1}{|\sigma_I|} Z_0, \qquad Z_0 = \int dx e^{-ix}, \tag{4·29}$$

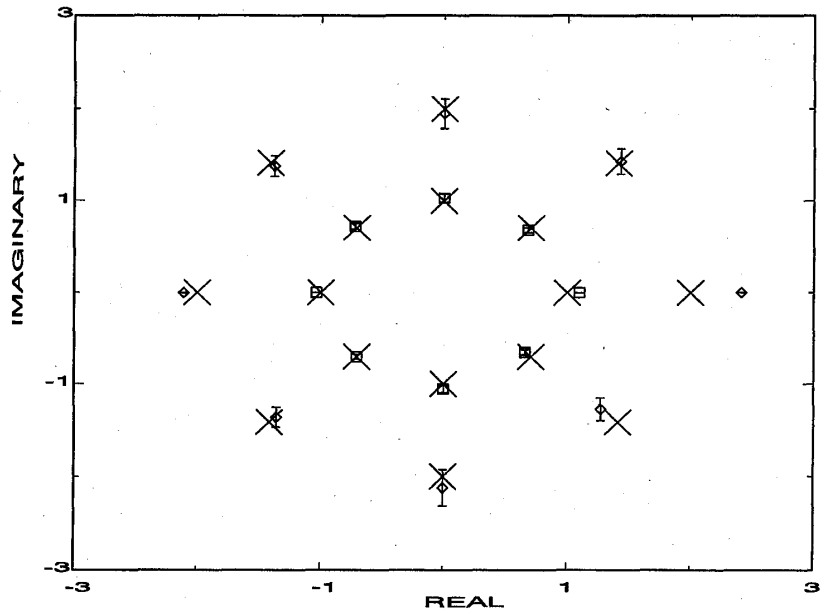

Fig. 14. Results for $\langle z\rangle$ of a Langevin simulation of (4·32) with complex σ varied in eight steps on the unit circle (squares), and for $\langle z^2\rangle$ where σ is varied on the upper half circle (diamonds). The exact results from Eq. (4·28) are marked by $\times$.

where we conclude that Z_0 is independent of σ. This result allows for the calculation of the right-hand side of (4·28) if we extend it to the whole complex plane analytically.

The Langevin equation associated to the action (4·27) reads

$$\dot{z}=-\sigma+\eta\,. \tag{4·30}$$

This equation does not converge. However, it can be made convergent with the help of the field dependent kernel

$$K(z)=\sigma^* z\,, \tag{4·31}$$

where the star denotes the complex conjugate. One obtain the Langevin equation (see (2·28))

$$\dot{z}=-|\sigma|^2 z+\sigma^*+\sqrt{\sigma^* z}\,\eta\,. \tag{4·32}$$

The result of a simulation $\langle z\rangle$ in this system has been plotted, for eight points $\sigma=N\pi/4$, $N=0,\cdots,7$, in Fig. 14 (squares). Furthermore the average $\langle z^2\rangle$ was taken (diamonds) for $\sigma=N\pi/8$. The crosses are the exact values from Eq. (4·28). For these points 50 runs were performed and each run with 10000 updations and with constant time step of size $dt=0.002$. The figure shows that a system with the linear action (4·27) can successfully be simulated.

Similar results one gets for the cubic action

$$S_3(x)=\frac{1}{3}\lambda x^3\,. \tag{4·33}$$

In this case, for pure imaginary $\lambda=i\lambda_I$ the partition function $Z(i\lambda_I)$ can be calculated exactly:

Table II. Results of a simulation of $\langle z^3\rangle$ with the cubic action (4·33) using the kernel (4·37).

exact	simulation
1.000	$1.0610-0.0002i$
	$\pm 0.0396\pm 0.0002i$
$0.707-0.707i$	$0.6800-0.6847i$
	$\pm 0.0237\pm 0.0210i$
$-i$	$0.0010-0.9615i$
	$\pm 0.0054\pm 0.0348i$
$-0.707-0.707i$	$-0.7180-0.7311i$
	$\pm 0.0203\pm 0.0295i$
-1	$-0.9730+0.0159i$
	$\pm 0.0324\pm 0.0093i$
$-0.707+0.707i$	$-0.7100+0.7205i$
	$\pm 0.0205\pm 0.0296i$
i	$0.0090+1.0753i$
	$\pm 0.0053\pm 0.0366i$
$0.707+0.707i$	$0.6830+0.6881i$
	$\pm 0.0208\pm 0.0181i$

$$Z(i\lambda_I)=\int_{-\infty}^{\infty}dx e^{-i(\lambda_I/3)x^3}=\left(\frac{3}{|\lambda_I|}\right)^{1/3}Z_0\,,$$

$$Z_0=\int dx e^{-ix^3}=2\Gamma\left(\frac{4}{3}\right)\cos\left(\frac{\pi}{6}\right). \quad (4\cdot 34)$$

The result for $Z(i\lambda_I)$ may be extended into the entire complex plane by analytic continuation. Therefore we obtain for the average

$$\langle x^3\rangle=\frac{1}{Z(\lambda)}(-3\partial_\lambda)Z(\lambda)=\frac{1}{\lambda}. \quad (4\cdot 35)$$

The corresponding Langevin equation

$$\dot{z}=-\lambda z^2+\eta \quad (4\cdot 36)$$

does not converge either. For instance with $\lambda>0$, for negative x the drift term will continuously push the configuration to even smaller values. However, if we introduce the kernel

$$K(z)=\frac{z}{\sqrt{\lambda}}, \quad (4\cdot 37)$$

the modified Langevin equation

$$\dot{z}=-\sqrt{\lambda}\,z^3+\frac{1}{\sqrt{\lambda}}+\sqrt{\frac{z}{\sqrt{\lambda}}}\eta \quad (4\cdot 38)$$

can be successfully be used to simulate the systems. Results have been given, for λ varying on a unit circle, in Table II.

Here we like to mention that the idea which leads to the introduction of the particular form of the kernel in (4·37) traces back to the discussion of the action (4·8) and the kernel (4·15) in § 4.1. If the value of λ is taken in the second Riemann sheet in the kernel in (4·37), one will also arrive at the unphysical solution, which gives the same value for $\langle x^3\rangle$ but different results for some other quantities, e.g., $\langle x^2\rangle$.

§ 5. Practical applications of complex Langevin simulation

All of the examples we have discussed in the previous sections are simple toy models. With help of them we have had some experience in what situation the complex Langevin simulation succeeds, or fails. In this subsection we briefly review the application to some realistic problems in lattice field theory.

5.1. *Spin-type effective model of the gauge theory with fermions*

As was written in the introduction, one can find many systems of physical interest whose effective actions are complex. One interesting example is the gauge theory

with fermions at finite temperature and baryon density, for which the complex Langevin simulation has been applied.[26)~29)] For the free fermions a naive extension of the continuum field theory gives rise to the following action on the lattice

$$S=a^3\sum_x\Big[ma\bar{\psi}_x\psi_x+\mu a\bar{\psi}_x\gamma_4\psi_x+\frac{1}{2}\sum_{\nu=1}^{4}(\bar{\psi}_x\gamma_\nu\psi_{x+\hat{\nu}}-\bar{\psi}_{x+\hat{\nu}}\gamma_\nu\psi_x)\Big]. \tag{5·1}$$

Here m and μ, respectively, represent the fermion mass and the chemical potential. If one postulates that the fermion field ψ follows an antiperiodic boundary condition along the imaginary time direction $\psi_{(x,x_4)}=-\psi_{(x,x_4+N_T)}$, the system (5·1) is supposed to be in the temperature T given by

$$T=(aN_T)^{-1}. \tag{5·2}$$

In Ref. 30), however, it has been shown that for $T\to 0$ this action leads to a divergent energy density $\epsilon\to(\mu/a)^2$ in the continuum limit for massless fermions, which is different from the expected μ^4. The energy density behaves correctly at $T\to 0$ if one modifies the action to

$$S=a^3\sum_x\Big[ma\bar{\psi}_x\psi_x+\frac{1}{2}\sum_{j=1}^{3}(\bar{\psi}_x\gamma_j\psi_{x+\hat{j}}-\bar{\psi}_{x+\hat{j}}\gamma_j\psi_x)$$

$$+\frac{1}{2}(e^{\mu a}\bar{\psi}_x\gamma_4\psi_{x+\hat{4}}-e^{-\mu a}\bar{\psi}_{x+\hat{4}}\gamma_4\psi_x)\Big]. \tag{5·3}$$

In the classical continuum limit, $e^{\pm\mu a}\approx 1\pm\mu a$, the naive version (5·1) is recovered from this expression.

Starting from the modified version of fermions in (5·3), one can easily include the Wilson term to overcome the fermion doubling, and introduce gauge fields.[30)] After integrating out the fermion degrees of freedom, and taking the leading term for the fermion determinant in strong coupling expansion and heavy quark limit for $N_T<4$,[31),32)] one gets a partition function

$$Z=\int dW_x e^{-S_{\rm eff}} \tag{5·4}$$

with the following effective action

$$S_{\rm eff}=-\beta\sum_{(x,l)}(\mathrm{tr}\,W_x\mathrm{tr}\,W_{x+l}^{\dagger}+\mathrm{c.c.})-h_1\sum_x\mathrm{tr}\,W_x-h_2\sum_x\mathrm{tr}\,W_x^{\dagger}. \tag{5·5}$$

Here, the Polyakov loop W_x takes values in the $SU(N)$ group, and

$$h_1=he^{\mu N_T},\qquad h_2=he^{-\mu N_T},\qquad h=2(2K)^{N_T}. \tag{5·6}$$

K is the hopping parameter ($K\to 0$ when $m\to\infty$), β is an effective coupling related to the original coupling g^2 and T, and we have set the lattice spacing $a=1$. The effective system is a spin-type model with an action $S_{\rm eff}$ which is complex in case $\mu\neq 0$. As a first step to understand the state of strongly interacting matter, it is of considerable interest to analyze the phase structure of this effective model of the gauge theory with fermions in the strong coupling and heavy quark limits.

For a $U(1)$ theory, the order parameters $\langle W_x\rangle$ and $\langle W_x^{-1}\rangle$ are measured in

complex Langevin simulation, and the results turned out to approach the mean-field ones.[28]

For a $SU(3)$ theory, one can check that for $\beta=0$ (in this case the partition function factorizes and only one-link integrations remain), the Langevin simulation produces the same values for $\langle \mathrm{tr}\, W_x\rangle$ and $\langle \mathrm{tr}\, W_x^{-1}\rangle$ as the exact integration even for a quite big μ which makes the imaginary part of the action comparable with the real part.[26] It is reported that by calculating the nearest-neighboring correlation $\langle \mathrm{tr}\, W_x \mathrm{tr}\, W_{x+\hat{\imath}}^{-1}\rangle$[26] or searching for metastability signals,[29] the phase diagram in the (β, h, μ) space established from Langevin simulation is, at least qualitatively, consistent with the mean-field prediction.[33] In all these simulations, including $U(1)$ theory, the complex Langevin equation converges remarkably well without observing any blow-up solutions.

The spin-type model in (5·5) is also modified for studying the chiral symmetry restoration at finite baryon density.[27]

5.2. *Gauge theories with static external charges*

In Refs. 34) and 20) Ambjørn et al. report about the simulation of gauge theories with static external charges. The charges are placed on Wilson lines L_1 and L_2, respectively. In the Abelian case the partition function is

$$Z=\int \prod_{n,\nu} d\theta_{n,\nu} \exp(\beta \sum_{\square} \cos\theta_{\square})\, W(L_1)\, W^{-1}(L_2) \tag{5·7}$$

with

$$\theta_{\square}=\theta_{n,\mu}+\theta_{n+\hat{\mu},\nu}-\theta_{n+\hat{\nu},\mu}-\theta_{n,\nu} \qquad \text{and} \qquad W(L)=\prod_{l\in L} e^{i\theta_l}. \tag{5·8}$$

$\hat{\mu}$ and $\hat{\nu}$ indicate the space-time directions. The product in $W(L)$ extends over all links on the Wilson line L in the time direction.

The corresponding formula for the non-Abelian case with group $SU(N)$ is given by

$$Z=\int \prod_{n,\nu} dU_{n,\nu} \exp\left\{\frac{\beta}{2N}\mathrm{tr}(U_{\square}+U_{\square}^{-1})\right\} \frac{1}{N}\mathrm{tr}\, W(L_1) \frac{1}{N}\mathrm{tr}\, W(L_2)^{-1}, \tag{5·9}$$

where

$$U_{\square}=U_{n,\mu}U_{n+\hat{\mu},\nu}U_{n+\hat{\nu},\mu}^{-1}U_{n,\nu}^{-1}, \tag{5·10}$$

and $W(L)$ is the product of all the link variables which belong to the line L. Care has to be taken to use U^{-1} instead of $U^{\dagger}$ in the last formulas, since in complex Langevin simulation U will no longer take its values in $SU(N)$ but rather in its complex extension $SL(N, C)$. Only if U^{-1} is used as a variable, the action is an analytic function of $U\in SL(N, C)$.

With growing lattice size the probability for getting runaway solutions increases. This causes a big problem concerning the computer time because when one encounters a runaway solution in one of the links, all of the link variables updated within the same run have to be thrown away. In order to escape from this problem, a procedure for adjusting the fictitious time step has been invented: in case one meets a too big $d\theta$

at a particular lattice link, the time step at that link is divided into smaller ones such that the $d\theta$ in each updation does not exceed a certain value. Updation of this link is repeated with fixed surrounding link variables as far as to reach the predefined time step dt.

In case of $U(1)$ lattice gauge theory in two dimensions the simulation is a great success. The expectation value $\langle\cos\theta_{\square}\rangle$ corresponding to the energy density can be calculated exactly,

$$\langle\cos\theta_{\square}\rangle_{\text{outside charge}}=I_1(\beta)/I_0(\beta)=\langle\cos\theta_{\square}\rangle_{\text{vacuum}},$$

$$\langle\cos\theta_{\square}\rangle_{\text{inside charge}}=I_0(\beta)/I_1(\beta)-1/\beta \qquad (5\cdot11)$$

$\langle\cos\theta_{\square}\rangle_{\text{inside charge}}$ grows linearly as $1/\beta$ when β tends to zero. For a lattice of size 20×21, the two charges being separated by a distance $L_2-L_1=10$, the Langevin equation easily generates relevant configurations for the complex system, and nicely produces the linear rising potential between the two charges. On the other hand, by conventional methods one would have to measure a Wilson loop with expectation value $[I_1(\beta)/I_0(\beta)]^{10\times20}=O(10^{-111})$ for $\beta=0.5$![20] The success of complex Langevin simulation is tremendous here! Within reasonable computer time enough configurations distributed according to the density e^{-S} are simulated to allow for a correct calculation of the average.

In case of $U(1)$ lattice gauge theory in more than two dimensions the result is correct only in the weak coupling region ($\beta\gtrsim2.2$). In the strong coupling region, however, the results are wrong: The energy density between two charges is not a constant, and also does not behave as $1/\beta$ as expected. The reason might be the lack of a fix point for $\theta_{\square}$ with non-zero imaginary part in the complex plane like those in the $\cos\theta$ model in § 3.2 and the $U(1)$ theory in two dimensions. The situation is even worse for the non-Abelian $SU(N)$ where no sign for an energy density is seen between two charges.

This failure has led to the more detailed investigation of the quantum mechanical example discussed in § 3.2.

5.3. *Other topics*

There exist other applications of the complex Langevin equations, e.g., to fermionic systems, field theories in Minkowski space, Chern-Simons or topological field theories.

In order to treat fermion fields and condensed matter problems such as electrons or in statistical mechanics or spins, Klauder has (see e.g., Ref. 3)) introduced the coherent state representation of electrons and other spin systems. After eliminating these spin degrees of freedom one encounters a complex effective action which can be simulated using the complex Langevin equation.[35),36)] The result seems to be promising and an extension to lattices of higher dimension has been reported quite recently.[37)]

The idea to use the complex Langevin equation for an investigation of field theories in Minkowski space is challenging but still not yet established. Many analytical approaches can be found.[38)~40),15),41),10)] The numerical application has been

done only for the harmonic and unharmonic oscillators in quantum mechanics (i.e., for one dimensional field theory).[42]

The complex Langevin equation with a kernel (see § 4) has been used recently in investigation of Chern-Simons theories. These topics will be covered by other contributions in this volume.

Acknowledgements

One of the authors (L. Sch.) gratefully acknowledges financial support from the Japan Society for the Promotion of Science (JSPS) and the German Academic Exchange service (DAAD) for a stay in Japan, during which this paper has been completed.

References

1) G. Parisi and Y. S. Wu, Sci. Sin. **24** (1981), 483.
2) G. Parisi, Phys. Lett. **B131** (1983), 393.
3) J. R. Klauder, *Recent Developments in High Energy Physics*, ed. H. Mitter and C. B. Lang (Springer, Wien, New York, 1983), p. 251.
4) J. R. Klauder, Phys. Rev. **A29** (1984), 2036.
5) J. D. Breit, S. Gupta and A. Zaks, Nucl. Phys. **B233** (1984), 61.
6) P. H. Damgaard and K. Tsokos, Nucl. Phys. **B235** (1984), 75.
7) G. Parisi, *Progress in Gauge Field Theory*, ed. G. 't Hooft, J. Jaffe, H. Lehmann, P. K. Mitter, I. N. Singer and R. Stora (Plenum, New York, 1984), p. 531.
8) G. G. Batrouni, G. R. Katz, A. S. Kronfeld, G. P. Lepage, B. Svetitsky and K. G. Wilson, Phys. Rev. **D32** (1985), 2736.
9) K. Okano and L. Schülke, Talk presented by K. Okano at the *International Conference on Field Theory and General Relativity* at Logan, Utah, June 1988.
10) H. Okamoto, K. Okano, L. Schülke and S. Tanaka, Nucl. Phys. **B324** (1989), 684.
11) K. Okano, L. Schülke and B. Zheng, Phys. Lett. **B258** (1991), 421.
12) J. R. Klauder and W. P. Petersen, J. Stat. Phys. **39** (1985), 53.
13) J. Ambjørn and S. -K. Yang, Phys. Lett. **B165** (1985), 140.
14) H. Nakazato and Y. Yamanaka, Phys. Rev. **D34** (1986), 492.
15) H. Nakazato, Prog. Theor. Phys. **77** (1987), 20.
16) F. Langouche, D. Roekaerts and E. Tirapegui, *Functional Integration and Semiclassical Expansions* (D. Reidel Publishing Company, 1982).
17) M. Namiki, *Delta Function and Differential Equations* (Iwanami Publisher, Tokyo, 1982) (in Japanese).
18) B. Söderberg, Nucl. Phys. **B295** [FS21] (1988), 396.
19) T. Matsui and A. Nakamura, Phys. Lett. **B194** (1987), 262.
20) J. Ambjørn, M. Flensburg and C. Peterson, Nucl. Phys. **B275** [FS17] (1986), 375.
21) J. Flower, S. W. Otto and S. Callahan, Phys. Rev. **D34** (1986), 598.
22) R. W. Haymaker and J. Wosiek, Phys. Rev. **D34** (1989), 969.
23) L. Schülke and B. Zheng, Int. J. Mod. Phys. **C3** (1992), 195; Proceedings of the workshop on *FERMION ALGORITHMS*, HLRZ Jülich, April 10-12, 1991.
24) K. Fujimura, K. Okano, L. Schülke, K. Yamagishi and B. Zheng, Univ. of Siegen and Tokuyama Univ. Preprint 1992.
25) J. Ambjørn and S. -K. Yang, Nucl. Phys. **B275** (1986), 18.
26) F. Karsch and H. W. Wyld, Phys. Rev. Lett. **55** (1985), 2242.
27) E. M. Ilgenfritz, Phys. Lett. **B181** (1986), 327.
28) N. Bilić, H. Gausterer and S. Sanielevici, Phys. Lett. **B198** (1987), 235.
29) N. Bilić, H. Gausterer and S. Sanielevici, Phys. Rev. **D37** (1988), 3684.

30) P. Hasenfratz and F. Karsch, Phys. Lett. **B125** (1983), 308.
31) R. V. Gavai, Phys. Rev. **D32** (1985), 519.
32) J. Engels and H. Satz, Phys. Lett. **B159** (1985), 151.
33) N. Bilić, Phys. Lett. **B184** (1987), 89.
34) J. Ambjørn, M. Flensburg and C. Peterson, Phys. Lett. **B159** (1985), 335.
35) H. Gausterer and J. R. Klauder, Phys. Lett. **B164** (1985), 127.
36) H. Gausterer and J. R. Klauder, Phys. Rev. Lett. **56** (1986), 306.
37) J. R. Klauder and S. Lee, Talk presented by S. Lee at the *International Winter School* at Schladming, Austria, 1992.
38) E. Gozzi, Phys. Lett. **150B** (1985), 119.
39) H. Hüffel and H. Rumpf, Phys. Lett. **B148** (1984), 104.
40) H. Hüffel and P. V. Landshoff, Nucl. Phys. **B260** (1985), 545.
41) J. Sakamoto, Prog. Theor. Phys. **80** (1988) 190.
42) D. J. E. Callaway, F. Cooper, J. R. Klauder and H. A. Rose, Nucl. Phys. **B262** (1985), 19.