Theoretical understanding of the problem with a singular drift term in the complex Langevin method

Jun Nishimura$^1$, and Shinji Shimasaki$^2$

$^1$KEK Theory Center, High Energy Accelerator Research Organization, Tsukuba 305-0801, Japan
$^2$Department of Particle and Nuclear Physics, School of High Energy Accelerator Science, Graduate University for Advanced Studies (SOKENDAI), Tsukuba 305-0801, Japan

(Dated: April, 2015; preprint: KEK-TH-1816)

The complex Langevin method aims at performing path integral with a complex action numerically based on complexification of the original real dynamical variables. One of the poorly understood issues concerns occasional failure in the presence of logarithmic singularities in the action, which appear, for instance, from the fermion determinant in finite density QCD. We point out that the failure should be attributed to the breakdown of the relation between the complex weight that satisfies the Fokker-Planck equation and the probability distribution associated with the stochastic process. In fact, this problem can occur in general when the stochastic process involves a singular drift term. We show, however, in a simple example that there exists a parameter region in which the method works although the standard reweighting method is hardly applicable.

PACS numbers: 05.10.Ln, 02.70.Tt, 12.38.Gc

Introduction.— The path integral formulation plays an important role in nonperturbative studies of quantum field theories due to the possibility of using Monte Carlo methods. The basic idea is to generate field configurations with a probability $e^{-S}$ and to evaluate the path integral by taking a statistical average. It is not straightforward, however, to apply such an approach to cases with a complex action $S$ since one can no longer view $e^{-S}$ as the probability distribution. This "complex action problem" occurs, for instance, in QCD at finite density or with a theta term, Chern-Simons gauge theories and chiral gauge theories. It also appears in supersymmetric gauge theories and matrix models, which are relevant in nonperturbative studies of superstring theory.

The complex Langevin method (CLM) attempts to solve this problem by extending the idea of stochastic quantization for ordinary systems with a real action to the case with a complex action. This necessarily requires complexifying the real dynamical variables that appear in the original path integral. A stochastic process for the complexified variables is defined by the Langevin equation. We start with simple examples and discuss this issue from the viewpoint of the Fokker-Planck (FP) equation. We provide theoretical understanding of this issue from the viewpoint of the Fokker-Planck (FP) equation. We start with simple examples and discuss more general cases towards the end.

One-variable case.— Let us consider a simple example defined by the partition function

$$Z = \int dx \ w(x), \quad w(x) = (x + i\alpha)^p e^{-x^2/2},$$

(1)
where $x$ is a real variable and $\alpha$ and $p$ are real parameters. For $\alpha \neq 0$ and $p \neq 0$, the weight $w(x)$ is complex, and the idea of important sampling cannot be applied to (1) by regarding $w(x)$ as the Boltzmann weight.

Following the usual procedure in CLM, we define the drift term

$$v(x) = w(x)^{-1} \frac{\partial w(x)}{\partial x} = \frac{p}{x + i\alpha} - x ,$$

and complexify the variable as $x \mapsto z = x + iy$. The action

$$S(z) = -\log w(z) = -p \log(z + i\alpha) + z^2/2$$

after the complexification has a logarithmic singularity at $z = -i\alpha$ for $p \neq 0$, which causes the aforementioned ambiguity due to the branch cut. However, we point out that this is not an issue in general because all we need in formulating the CLM, as we show below, is the single-valuedness of the drift term $v(z)$ after the complexification and the single-valuedness of the complex weight $w(x)$ as a function of $x$. These are satisfied in the present case even for a non-integer $p$.

The complex Langevin equation corresponding to the partition function (1) can be written as

$$\frac{dz}{dt} = v(z) + \eta(t) = \frac{p}{z + i\alpha} - z + \eta(t) ,$$

where $\eta(t)$ represents a real Gaussian noise satisfying $\langle \eta(t)\eta(t') \rangle = 2\delta(t-t')$. We define the probability distribution $P(x, y; t)$ of $x(t)$ and $y(t)$ at the Langevin time $t$. Its time evolution follows the FP-like equation

$$\frac{\partial}{\partial t} P(x, y; t) = L P$$

$$= \frac{\partial}{\partial x} \left\{ -(\text{Re} v)_{z=x+iy} + \frac{\partial}{\partial x} \right\} P$$

$$+ \frac{\partial}{\partial y} \left\{ -(\text{Im} v)_{z=x+iy} P \right\} .$$

The crucial point in the CLM is that there exists a complex weight $\rho(x; t)$, which is related to the probability distribution $P(x, y; t)$ through

$$\int \mathcal{O}(x) \rho(x; t) dx = \int \mathcal{O}(x + iy) P(x, y; t) dx dy$$

under certain conditions $\mathbb{R}$, where $\mathcal{O}(x)$ are observables that admit holomorphic extension to $\mathcal{O}(x + iy)$. The evolution of $\rho$ follows the usual FP equation

$$\frac{\partial}{\partial t} \rho = L_0 \rho ,$$

$$L_0 = \frac{\partial}{\partial x} \left( -v(x) + \frac{\partial}{\partial x} \right) ,$$

which has a time-independent solution $\rho(x; t) \propto w(x)$ with $w(x)$ given in (1) since it is annihilated by the operator in the parenthesis in (9). Thus the necessary and sufficient condition for being able to calculate the expectation value of $\mathcal{O}$ with respect to (1) by the CLM is:

i) The relation (7) between $\rho$ and $P$ holds.

ii) The solution $\rho(x; t)$ of the FP equation (8) asymptotes to $w(x)$ as $t \to \infty$ up to some constant factor.

As possible observables in the present example (1), we consider $\mathcal{O} = x^k$, where $k$ is a positive integer. Assuming the ergodicity of the stochastic process, the right-hand side of (7) at $t = \infty$ can be evaluated by taking the time-average of $z(t)^k$, where $z(t)$ is obtained by solving (4). We find numerically that this method gives correct results only for sufficiently large $|\alpha|$ for each $p$. In what follows, we clarify the reason why it fails at small $|\alpha|$.

* Spectrum of $(-L_0)$. — First we have investigated numerically the eigenvalue spectrum of the “FP Hamiltonian” $(-L_0)$ defined by (9) assuming that the complex weight $\rho(x)$ falls off rapidly as $|x| \to \infty$.

As is clear from what we wrote above, we have an eigenfunction $\rho(x) = w(x)$ with zero eigenvalue for arbitrary $p$ and $\alpha$. When $p$ is a positive odd integer and $\alpha = 0$, we have another zero mode $\rho(x) = |x|^p e^{-x^2/2}$ for $p > 1$, negative eigenvalues appear in the small $|\alpha|$ region. (Note that, when $\alpha = 0$, we have an eigenfunction $\rho(x) = x e^{-x^2/2}$, which corresponds to the smallest eigenvalue $\lambda = -(p-1)$ for any $p$.) Thus we find that the desired solution $\rho(x; t) \propto w(x)$ is obtained in the long-time limit of the FP equation (8) at arbitrary $\alpha$ for $p < 1$ and at sufficiently large $|\alpha|$ for $p > 1$.

In the parameter region where we have negative modes, the complex weight diverges as $\rho(x; t) \propto e^{i\lambda_{\min}|x|} \rho_{\min}(x)$, where $\lambda_{\min}$ is the smallest eigenvalue and $\rho_{\min}(x)$ is the corresponding eigenfunction. Clearly this behavior is incompatible with the relation (7) considering that $P(x, y; t) \geq 0$ and $\int dx dy P(x, y; t) = 1$. This implies that the relation (7) between $\rho$ and $P$ must be broken at least in this region.

* The relation between $\rho$ and $P$. — Let us then consider what can go wrong with (7). In deriving (7), one uses

$$\int \mathcal{O}(x + iy) P(x, y; t) dx dy = \int \mathcal{O}(x + iy; t) P(x, y; 0) dx dy ,$$

where $\mathcal{O}(x + iy; t)$ is defined by solving

$$\frac{d}{dt} \mathcal{O}(x + iy; t) = L^\top \mathcal{O}(x + iy; t)$$

with the initial condition $\mathcal{O}(x + iy; 0) = \mathcal{O}(x + iy)$. The symbol $L^\top$ represents an operator satisfying $\langle L^\top f, g \rangle = \langle f, Lg \rangle$, where $\langle f, g \rangle = \int f(x, y)g(x, y) dx dy$, assuming that $f$ and $g$ are regular functions with sufficiently fast fall-off as $|x|, |y| \to \infty$.

In order to prove (10), one considers

$$F(\tau) = \int \mathcal{O}(x + iy; \tau) P(x, y; t - \tau) dx dy$$

The relation between $\rho$ and $P$ holds.
for $0 \leq \tau \leq t$, which interpolates both sides of Eq. (10). Taking the derivative with respect to $\tau$, one gets

$$\frac{d}{d\tau} F(\tau) = \int \left\{ L^\dagger \mathcal{O}(x + iy; \tau) \right\} P(x, y; t - \tau) dx dy$$

$$- \int \mathcal{O}(x + iy; \tau) LP(x, y; t - \tau) dx dy .$$

(13)

Naively, the two terms cancel through integrating by parts, which implies that $F(\tau)$ is independent of $\tau$ and hence (10). In order to justify the partial integration, however, one should be able to neglect the boundary terms. This requires that $P(x, y; t)$ decreases sufficiently fast as $|x|, |y| \to \infty$. In the present example, this condition is satisfied, but we should be careful of the singularity at $(x, y) = (0, -\alpha)$. In order for the boundary terms to be neglected, it is required that the limits

$$\lim_{x \to 0} \left[ x \int \frac{\mathcal{O}(z; \tau)}{|z + i\alpha|^2} P(x, y; t - \tau) dy \right] ,$$

(14)

$$\lim_{y \to -\alpha} \left[ (y + \alpha) \int \frac{\mathcal{O}(z; \tau)}{|z + i\alpha|^2} P(x, y; t - \tau) dx \right]$$

(15)

should exist for arbitrary $t$ and $\tau$. Note, in particular, that $\mathcal{O}(z; \tau)$ obtained by solving (11) is highly singular at $z = -i\alpha$. For instance, let us take the $n$-th derivative of the above expressions with respect to $\tau$ at $\tau = 0$. Using (11), we obtain terms such as

$$\lim_{x \to 0} \left[ x \int \frac{(L^\dagger)^n \mathcal{O}(z)}{|z + i\alpha|^2} P(x, y; t) dy \right] ,$$

(16)

which diverges as $\sim \frac{1}{|x|(|\alpha| - 1)}$ for $n \geq 1$ if $P(x, y; t)$ is nonzero at $(x, y) = (0, -\alpha)$.

In order to describe the actual situation, let us define the radial distribution

$$\varphi(r) = \frac{1}{2\pi r} \int P(x, y, \infty) \delta(\sqrt{x^2 + (y + \alpha)^2} - r) dx dy .$$

(17)

around the singular point $(x, y) = (0, -\alpha)$. For small $|\alpha|$, we observe that $\varphi(r) \sim r$ at small $r$ as far as $p$ is not very large. In this case, (16) still diverges for sufficiently large $n$, and the relation (17) between $\rho$ and $P$ can be violated. Indeed we find that the CLM yields wrong results in such a parameter region.

**Results for large $p$.**— In the partition function (1), it is the prefactor $(x + i\alpha)^p$ that causes the complex action problem. In view of this, one might think that the CLM fails when the phase of $(z(t) + i\alpha)^p$ rotates frequently during the time evolution by (14). We find that this is not necessarily the case.

In order to demonstrate this point, we present our results for large $p$. Fig. 1 shows that the CLM reproduces the exact results for $|\alpha| \gtrsim 14$ at $p = 50$. From Fig. 2 we find for $\alpha = 14$ that $\varphi(r) = 0$ at $r \lesssim 6$ although we observe that the phase of $(z(t) + i\alpha)^{50}$ rotates frequently during the stochastic process.

We would also like to mention that the complex action problem is extremely severe at $\alpha = 14$ and $p = 50$. As a standard quantity which measures the severeness of the complex action problem, let us consider

$$R = \left\langle \frac{w(x)}{|w(x)|} \right\rangle_0 = \left\langle \frac{(x + i\alpha)^p}{|x + i\alpha|^p} \right\rangle_0 = \frac{Z}{Z_0} ,$$

(18)

where $Z_0$ is the partition function of the phase-quenched model $Z_0 = \int dx |w(x)|$ and the expectation value $\left\langle \cdot \right\rangle_0$ is taken with respect to it. In the present case, both $Z$ and $Z_0$ can be calculated analytically by performing the Gaussian integration. We find that $R \sim -7.4 \times 10^{-5}$ at $\alpha = 14$ and $p = 50$. One can imagine how hard it is to obtain correct results if one performs Monte Carlo simulation of the phase quenched model and ap-
plies the standard reweighting formula to obtain the expectation values with respect to the original partition function [11]. Thus the advantage of the CLM over the reweighting method can be appreciated even in this simple one-variable case.

Non-logarithmic singularities.— Another interesting implication of our argument is that the possible failure of the CLM is not specific to logarithmic singularities in the action. Indeed we have found that the CLM can fail for the weight \( w(x) = e^{-S(x)} \) with the action

\[
S(x) = \beta(x + ia)^{-2} + x^2/2. \tag{19}
\]

Note that the action \( S(z) \) after complexification does not involve a logarithmic singularity, which means, in particular, that there is no issue of ambiguity associated with the branch cut. Nevertheless we find that the CLM fails at \( |\alpha| \lesssim 1.2 \) for \( \beta = 1 \) and at \( |\alpha| \lesssim 1.7 \) for \( \beta = -1 \). On the other hand, from the studies of the eigenvalue spectrum of \((-L_0)\), we find that \( \rho(x; t) \propto e^{-S(x)} \) is obtained in the long-time limit of the FP equation (8) for arbitrary \( \alpha \) with \( \beta = \pm 1 \). Therefore, the failure of the CLM at small \( |\alpha| \) should be attributed to the violation of (7) due to the singularity in the drift term \( v(z) = 2\beta(z + ia)^{-3} - z \). This is confirmed also from the behavior of the radial distribution [12].

Two-variable case.— Our argument applies also to the case with multiple variables. To make this clear, let us consider a case with two variables given by

\[
Z = \int dx_1 dx_2 \, w(x_1, x_2),
\]

\[
w(x_1, x_2) = (x_1 + ix_2)^p e^{-|x_1|^2/2 - |x_2|^2/2},
\]

where \( x_1 \) and \( x_2 \) are real variables. The parameter \( \alpha \) is real, while \( p \) is a positive integer.

We have studied numerically the eigenvalue spectrum of the operator \((-L_0)\) assuming that the complex weight \( \rho(x_1, x_2) \) falls off rapidly as \( (x_1)^2 + (x_2)^2 \to \infty \). First we obtain the desired zero mode \( \rho(x_1, x_2) = w(x_1, x_2) \) for arbitrary \( p \) and \( \alpha \). When \( \alpha = 0 \), we have another zero mode \( \rho = |x_1 + ix_2|^p e^{-|x_1|^2/2 - |x_2|^2/2} \) for any \( p \). For \( p > 1 \), negative modes appear at small \( |\alpha| \). (Note that, when \( \alpha = 0 \), we have an eigenfunction \( \rho = (x_1 + ix_2)^p e^{-|x_1|^2/2 - |x_2|^2/2} \), which corresponds to the smallest eigenvalue \( \lambda = -(p - 1) \) for any \( p \).) Thus we can make an argument analogous to the one-variable case [11]. Indeed we find for \( p = 1, 2, 3 \) that the CLM with complexified variables \( z_1 \) and \( z_2 \) gives wrong results at small \( |\alpha| \). This can be understood from the behavior of the radial distribution for \( r = |z_1 + iz_2| \).

Implications to finite density QCD.— Let us discuss the implication of our argument to finite density QCD, which involves the complex fermion determinant \( \det(D + m) \) in the partition function, where \( D \) represents the Dirac operator and \( m \) is the quark mass. The determinant can be written as the product of the eigenvalues \( \lambda_k \) of \( (D + m) \). The drift term of the complex Langevin equation involves \( \sum_k (\lambda_k)^{-1} \partial \lambda_k \), where \( \partial \) represents the derivative with respect to the complexified gauge field.

According to our argument, the problem we discussed does not appear as far as the distribution of \( \lambda_k \) is practically zero at the origin, even if the phase of fermion determinant rotates frequently during the stochastic process. (See our results for large \( p \) in the one-variable case.) This is consistent with the results of recent QCD simulations at finite density, where the distribution of \( \lambda_k \) has the desired property due either to large quark mass [11] or to high temperature [12]. On the other hand, the eigenvalues of \( D \) obtained in the CLM are expected to accumulate at the origin in the chiral limit (corresponding to the \( m \to 0 \) limit) when the chiral symmetry is spontaneously broken [13]. Therefore, the CLM may have problems in that parameter regime unless some new idea is invoked.

Concluding remarks.— We consider that this work provides the missing piece for the establishment of the CLM. The new insights gained here are expected to be useful, in particular, in developing the method further in cases with singularities in the drift term.

Acknowledgments.— The authors would like to thank D. Sexty for valuable discussion. The work of J. N. was supported in part by Grant-in-Aid for Scientific Research (No. 23244057) from Japan Society for the Promotion of Science.