

Justification of the complex Langevin method with the gauge cooling procedure

Keitaro Nagata^{1,*}, Jun Nishimura^{1,2,*}, and Shinji Shimasaki^{1,*}

¹*KEK Theory Center, High Energy Accelerator Research Organization, 1-1 Oho, Tsukuba, Ibaraki 305-0801, Japan*

²*Graduate University for Advanced Studies (SOKENDAI), 1-1 Oho, Tsukuba, Ibaraki 305-0801, Japan*

*E-mail: knagata@post.kek.jp, jnishi@post.kek.jp, simasaki@post.kek.jp

Received September 18, 2015; Accepted October 13, 2015; Published January 1, 2016

.....
Recently, there has been remarkable progress in the complex Langevin method, which aims to solve the complex action problem by complexifying the dynamical variables in the original path integral. In particular, a new technique, called gauge cooling, has been introduced and the full QCD simulation at finite density has been made possible in the high-temperature (deconfined) phase or with heavy quarks. Here we provide an explicit justification of the complex Langevin method including the gauge cooling procedure. We first show that the gauge cooling can be formulated in the form of a modified complex Langevin equation involving a complexified gauge transformation, which is chosen appropriately as a function of the configuration before cooling. The probability distribution of the complexified dynamical variables is modified accordingly. However, this modification is shown *not* to affect the Fokker–Planck equation for the corresponding complex weight as long as observables are restricted to gauge-invariant ones. Thus we demonstrate explicitly that gauge cooling can be used as a viable technique to satisfy the convergence conditions for the complex Langevin method. We also discuss “gauge cooling” in 0D systems such as vector models or matrix models.
.....

Subject Index B01

1. Introduction

Monte Carlo calculation plays an important role in nonperturbative studies of quantum field theories. However, its usefulness becomes quite limited when the action S becomes complex, because the integrand e^{-S} in the path integral can no longer be regarded as the Boltzmann weight. This occurs in many interesting cases, such as QCD at finite density or with a theta term, gauge theories with a Chern–Simons term, chiral gauge theories, and so on. It also occurs in supersymmetric gauge theories and matrix models relevant to nonperturbative studies of superstring theory.

Amongst various approaches to this complex action problem, that based on the complex Langevin equation has recently been attracting a lot of attention. The original idea was proposed by Parisi [1] and Klauder [2] in 1983, and since then it has been applied to various systems with complex actions. A salient feature of the method is that it works beautifully in some fairly nontrivial cases, but fails completely in other cases. For a long time, theoretical understanding of this feature was missing, and that led to the gradual decline of interest in this approach. However, in 2011, one of the problems of the method in cases in which it fails was clearly identified [3,4]. The authors first derived the key relation between the complex Langevin process and the Fokker–Planck equation for the complex weight. Then it was found that the integration by parts used in the derivation may not be justified

unless the probability distribution of the complexified dynamical variables is suppressed strongly enough when they take large values.

Gauge cooling has been proposed to cure this problem in the case of gauge theories [5]. It has been applied to finite-density QCD in the heavy dense limit and shown to work in the whole parameter regime in that limit [6]. More recently, it has been applied to finite-density QCD without taking the heavy dense limit, and has been shown to work, at least in the deconfined phase [7]. This is already quite remarkable, since the cases that have been studied include a parameter region, which would be hardly accessible by other methods such as reweighting. On the other hand, it is also realized, in a solvable gauge theory with a complex coupling constant, that there exists some parameter regime in which the gauge cooling cannot completely cure the insufficient fall-off of the probability distribution [8].

In fact, there is another problem that is anticipated to occur when one applies the complex Langevin method (CLM) to QCD with light quarks at low temperature. This was realized in Ref. [9] by applying the CLM to the random matrix theory for finite-density QCD. It turned out that a naive implementation of the method fails as the quark mass is decreased (see, however, Ref. [10]). The reason for this failure was speculated to have something to do with the logarithmic singularity in the action due to the fermion determinant [9–11]. On the other hand, it was also pointed out [12] that the problem occurs due to a singular drift term, which breaks the requirement of holomorphy in the derivation of the key relation between the complex Langevin process and the Fokker–Planck equation for the complex weight [3,4]. Recently, two of the authors (J.N. and S.S.) [13] have argued that it is actually the integration by parts used in the derivation that is invalidated by the singular drift term. According to this understanding, the problem can be avoided if the probability distribution is suppressed strongly enough near the singularity. In a separate paper (K. Nagata et al., manuscript in preparation), we show that this can also be achieved by gauge cooling with appropriate choice of the quantity that should be reduced by the cooling procedure.

While intuitive arguments for justification of the gauge cooling are given in the literature (see, for instance, Sect. 5 of Ref. [14]), an explicit justification is missing. In fact, there is even some suspicion in the community that the procedure may not be fully justified. Some of the concerns that we have encountered in private communications are that: 1) gauge cooling uses a complexified gauge symmetry, which is not a symmetry of the original system; 2) the noise term in the complex Langevin equation is invariant under the original gauge transformation but not under the complexified gauge transformation; 3) the quantity that one tries to reduce by the complexified gauge transformation is not holomorphic, which may spoil the justification of the CLM. In view of this situation, here we provide an explicit justification of the CLM including the gauge cooling procedure. We first show that gauge cooling can be formulated in the form of a modified complex Langevin equation involving a complexified gauge transformation, which is chosen appropriately as a function of the configuration before cooling. The probability distribution of the complexified dynamical variables is modified accordingly. However, this modification is shown not to change the Fokker–Planck equation for the corresponding complex weight as long as the observables are restricted to gauge-invariant ones. Thus we conclude that gauge cooling can be used to realize the properties of the probability distribution that are required for its relation to the complex weight without affecting the Fokker–Planck equation.

We also discuss “gauge cooling” in 0D systems such as vector models or matrix models, which is simpler than that in lattice gauge theory. Apart from pedagogical purposes, we consider that it is useful, for instance, in studying the matrix models relevant to superstring theory [15,16].

The rest of this paper is organized as follows. In Sect. 2, we briefly review the Langevin method, starting from the well-established case of real action, and discuss the conditions for correct convergence in the case of complex action. In Sect. 3, we discuss the “gauge cooling” in 0D systems and provide its justification. In Sect. 4, we discuss the application of the CLM to lattice gauge theory. In Sect. 5, we present an explicit justification of gauge cooling in lattice gauge theory. Section 6 is devoted to a summary and discussions.

2. Brief review of the Langevin method

In this section, we briefly review the Langevin method. (For a comprehensive review on this subject, we recommend Ref. [17].) Here we consider a system of n real variables x_k ($k = 1, \dots, n$) given by the partition function

$$Z = \int dx e^{-S(x)} = \int \prod_k dx_k e^{-S(x)}, \quad (2.1)$$

where the action $S(x)$ is a function of $x = (x_1, \dots, x_n)$. We start with the well-established case of real action, which is also known as stochastic quantization. Then we discuss the case of complex action, focusing on the conditions for correct convergence.

2.1. The case of real action

When the action $S(x)$ is real, we can use the ordinary Langevin method to study this system [18]. Introducing a fictitious time t , we consider the t -evolution governed by the Langevin equation

$$\dot{x}_k^{(\eta)}(t) = -\frac{\partial S}{\partial x_k} + \eta_k(t), \quad (2.2)$$

where $\eta_k(t)$ are probabilistic variables obeying the probability distribution $e^{-\frac{1}{4} \int dt \eta_k(t)^2}$. The first and second terms on the right-hand side of the Langevin equation (2.2) are commonly called the drift term and the noise term, respectively, for historical reasons.

The probability distribution of $x^{(\eta)}(t)$ can be defined as

$$P(x, t) = \left\langle \prod_k \delta(x_k - x_k^{(\eta)}(t)) \right\rangle_\eta, \quad (2.3)$$

where the expectation value $\langle \dots \rangle_\eta$ is defined by

$$\langle \dots \rangle_\eta = \frac{\int \mathcal{D}\eta \dots e^{-\frac{1}{4} \int dt \eta_k(t)^2}}{\int \mathcal{D}\eta e^{-\frac{1}{4} \int dt \eta_k(t)^2}}. \quad (2.4)$$

Using this notation, one obtains, for instance,

$$\left\langle \eta_k(t_1) \eta_l(t_2) \right\rangle_\eta = 2\delta_{kl}\delta(t_1 - t_2). \quad (2.5)$$

One can actually show that $P(x, t)$ satisfies the Fokker–Planck (FP) equation (see Sect. 2.2 for the derivation):

$$\frac{\partial P}{\partial t} = \frac{\partial}{\partial x_k} \left(\frac{\partial S}{\partial x_k} + \frac{\partial}{\partial x_k} \right) P, \quad (2.6)$$

which has a time-independent solution

$$P_{\text{time-indep}}(x) = \frac{1}{Z} e^{-S(x)}. \quad (2.7)$$

Under quite general conditions [17], one can show that the eigenvalues of the differential operator acting on P on the right-hand side of (2.6) are strictly negative, except for the zero eigenvalue corresponding to (2.7). This implies that the probability distribution $P(x, t)$ approaches (2.7) exponentially. One can therefore obtain a vacuum expectation value (VEV) with respect to the partition function (2.1) as

$$\begin{aligned}\langle \mathcal{O}(x) \rangle &= \int dx \mathcal{O}(x) P_{\text{time-indep}}(x) \\ &= \lim_{t \rightarrow \infty} \int dx \mathcal{O}(x) P(x, t) \\ &= \lim_{t \rightarrow \infty} \left\langle \mathcal{O}\left(x^{(\eta)}(t)\right) \right\rangle_{\eta} \\ &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_{t_0}^{t_0+T} dt \mathcal{O}\left(x^{(\eta)}(t)\right).\end{aligned}\quad (2.8)$$

In the last step, the statistical average over η is replaced by the time average, assuming the ergodicity of the stochastic process, as in the usual Monte Carlo methods.

2.2. The discretized Langevin equation

When one tries to solve the Langevin equation (2.2) numerically, one has to discretize the fictitious time t and solve, for instance¹,

$$x_k^{(\eta)}(t + \epsilon) = x_k^{(\eta)}(t) + \epsilon \left(-\frac{\partial S}{\partial x_k} + \eta_k(t) \right), \quad (2.9)$$

where the probabilistic variables $\eta_k(t)$ obey the probability distribution $e^{-\frac{1}{4}\epsilon \sum_t \eta_k(t)^2}$. Let us rescale them as $\tilde{\eta}_k = \sqrt{\epsilon} \eta_k$ so that they obey the probability distribution $e^{-\frac{1}{4} \sum_t \tilde{\eta}_k(t)^2}$ and hence, in particular,

$$\left\langle \tilde{\eta}_k(t_1) \tilde{\eta}_l(t_2) \right\rangle_{\eta} = 2\delta_{kl} \delta_{t_1, t_2}. \quad (2.10)$$

With this normalization, the discretized Langevin equation (2.9) becomes

$$x_k^{(\eta)}(t + \epsilon) = x_k^{(\eta)}(t) - \epsilon \frac{\partial S}{\partial x_k} + \sqrt{\epsilon} \tilde{\eta}_k(t). \quad (2.11)$$

Below, we omit the tilde on η to simplify the notation.

With this discretized version, we can derive the FP equation (2.6) in a more elementary manner than in the continuum [17]. Let us consider a test function $f(x)$ and its expectation value

$$\left\langle f\left(x^{(\eta)}(t)\right) \right\rangle_{\eta} = \int dx f(x) P(x; t) \quad (2.12)$$

¹ There are more sophisticated ways of discretization that can be used to reduce the systematic errors due to the discretization; see Ref. [19] and references therein.

at a fictitious time t . The t -evolution of this quantity is given by

$$\begin{aligned}
& \left\langle f\left(x^{(\eta)}(t+\epsilon)\right)\right\rangle_{\eta}-\left\langle f\left(x^{(\eta)}(t)\right)\right\rangle_{\eta} \\
& =\left\langle\frac{\partial f}{\partial x_k}\left(-\epsilon \frac{\partial S}{\partial x_k}\right)+\frac{1}{2} \frac{\partial^2 f}{\partial x_k \partial x_l}(\sqrt{\epsilon})^2 \eta_k(t) \eta_l(t)\right\rangle_{\eta}+O\left(\epsilon^2\right) \\
& =\epsilon \int d x\left(-\frac{\partial f}{\partial x_k} \frac{\partial S}{\partial x_k}+\frac{\partial^2 f}{\partial x_k^2}\right) P(x ; t)+O\left(\epsilon^2\right) \\
& =\epsilon \int d x f(x) \frac{\partial}{\partial x_k}\left(\frac{\partial S}{\partial x_k}+\frac{\partial}{\partial x_k}\right) P+O\left(\epsilon^2\right) .
\end{aligned} \tag{2.13}$$

Here we have used

$$\left\langle\frac{1}{2} \frac{\partial^2 f}{\partial x_k \partial x_l}(\sqrt{\epsilon})^2 \eta_k(t) \eta_l(t)\right\rangle_{\eta}=\frac{1}{2} \epsilon\left\langle\frac{\partial^2 f}{\partial x_k \partial x_l}\right\rangle_{\eta}\left\langle\eta_k(t) \eta_l(t)\right\rangle_{\eta}=\epsilon\left\langle\frac{\partial^2 f}{\partial x_k^2}\right\rangle_{\eta}, \tag{2.14}$$

which follows from the fact that the function $\frac{\partial^2 f}{\partial x_k \partial x_l}$ is evaluated at $x=x^{(\eta)}(t)$, which depends only on $\eta(0), \eta(\epsilon), \dots, \eta(t-\epsilon)$, but not on $\eta(t)$. Using (2.12), the same quantity (2.13) should be written as

$$\left\langle f\left(x^{(\eta)}(t+\epsilon)\right)\right\rangle_{\eta}-\left\langle f\left(x^{(\eta)}(t)\right)\right\rangle_{\eta}=\int d x f(x)\left(P(x ; t+\epsilon)-P(x ; t)\right) . \tag{2.15}$$

Since (2.13) and (2.15) should be equal for an arbitrary $f(x)$, one obtains

$$P(x ; t+\epsilon)-P(x ; t)=\epsilon \frac{\partial}{\partial x_k}\left(\frac{\partial S}{\partial x_k}+\frac{\partial}{\partial x_k}\right) P+O\left(\epsilon^2\right) . \tag{2.16}$$

Thus, in the $\epsilon \rightarrow 0$ limit, one obtains (2.6).

2.3. The case of complex action

Let us apply the same method to the case in which the action S is a complex-valued function of the real variables x_k ($k=1, \dots, n$). In that case, however, the first term on the right-hand side of the Langevin equation (2.2) becomes complex, which means that $x_k^{(\eta)}(t)$ becomes complex even if one starts from a real configuration $x_k^{(\eta)}(0) \in \mathbb{R}$. Let us therefore complexify the variables² as $x_k \mapsto z_k = x_k + i y_k$, and solve the complex Langevin equation

$$\dot{z}_k^{(\eta)}(t)=-\frac{\partial S}{\partial z_k}+\eta_k(t), \tag{2.17}$$

where the action S is now considered as a function of the complex variables z_k ($k=1, \dots, n$) by analytic continuation. It is important for the method that the action $S(z)$ thus obtained is a holomorphic

² In this respect, there is a closely related approach based on the so-called Lefschetz thimble [20,21], which has attracted much attention recently; see Refs. [22–26] and references therein.

function of z_k . The probabilistic variables $\eta_k(t)$ in (2.17) are, in general, complex:

$$\eta_k(t) = \eta_k^{(R)}(t) + i\eta_k^{(I)}(t), \quad (2.18)$$

and obey the probability distribution $e^{-\frac{1}{4} \int dt \left\{ \frac{1}{N_R} \eta_k^{(R)}(t)^2 + \frac{1}{N_I} \eta_k^{(I)}(t)^2 \right\}}$. The probability distribution corresponding to (2.3) is defined as

$$P(x, y; t) = \left\langle \prod_k \delta(x_k - x_k^{(\eta)}(t)) \delta(y_k - y_k^{(\eta)}(t)) \right\rangle_\eta, \quad (2.19)$$

where the expectation value $\langle \cdots \rangle_\eta$ is defined by

$$\langle \cdots \rangle_\eta = \frac{\int \mathcal{D}\eta \cdots e^{-\frac{1}{4} \int dt \left\{ \frac{1}{N_R} \eta_k^{(R)}(t)^2 + \frac{1}{N_I} \eta_k^{(I)}(t)^2 \right\}}}{\int \mathcal{D}\eta e^{-\frac{1}{4} \int dt \left\{ \frac{1}{N_R} \eta_k^{(R)}(t)^2 + \frac{1}{N_I} \eta_k^{(I)}(t)^2 \right\}}}. \quad (2.20)$$

With this notation, we have, for instance,

$$\begin{aligned} \left\langle \eta_k^{(R)}(t_1) \eta_l^{(R)}(t_2) \right\rangle_\eta &= 2N_R \delta_{kl} \delta(t_1 - t_2), \\ \left\langle \eta_k^{(I)}(t_1) \eta_l^{(I)}(t_2) \right\rangle_\eta &= 2N_I \delta_{kl} \delta(t_1 - t_2), \\ \left\langle \eta_k^{(R)}(t_1) \eta_l^{(I)}(t_2) \right\rangle_\eta &= 0. \end{aligned} \quad (2.21)$$

In what follows, we assume that

$$N_R - N_I = 1, \quad (2.22)$$

for a reason that becomes clear later. For practical purposes, one should actually use $N_R = 1, N_I = 0$, corresponding to real $\eta_k(t)$ with the distribution (2.4), to reduce the excursion in the imaginary directions [3,4], which spoils the validity of the method, as we review below.

Repeating the analysis given in Sect. 2.2, one can easily show that $P(x, y; t)$ satisfies the FP-like equation

$$\frac{\partial P}{\partial t} = \frac{\partial}{\partial x_k} \left\{ \text{Re} \left(\frac{\partial S}{\partial z_k} \right) + N_R \frac{\partial}{\partial x_k} \right\} P + \frac{\partial}{\partial y_k} \left\{ \text{Im} \left(\frac{\partial S}{\partial z_k} \right) + N_I \frac{\partial}{\partial y_k} \right\} P. \quad (2.23)$$

In fact, for observables $\mathcal{O}(x)$ that admit holomorphic extension to $\mathcal{O}(x + iy)$, one can show that, under certain conditions, there exists a complex function $\rho(x; t)$, which satisfies

$$\int dx dy \mathcal{O}(x + iy) P(x, y; t) = \int dx \mathcal{O}(x) \rho(x; t), \quad (2.24)$$

and obeys the differential equation

$$\frac{\partial \rho}{\partial t} = \frac{\partial}{\partial x_k} \left(\frac{\partial S}{\partial x_k} + \frac{\partial}{\partial x_k} \right) \rho, \quad (2.25)$$

which is formally the same as the FP equation (2.6). Although there is an important difference that the action S and ρ are now complex, the FP equation (2.25) still has a time-independent solution:

$$\rho_{\text{time-indep}}(x) = \frac{1}{Z} e^{-S(x)}. \quad (2.26)$$

The convergence to this solution in the $t \rightarrow \infty$ limit requires that all the eigenvalues of the operator acting on ρ on the right-hand side of (2.25) should have strictly negative real parts, except for the

zero eigenvalue corresponding to (2.26). While this is not guaranteed in general, unlike in the real action case, one can argue that the convergence to (2.26) should occur if the relation (2.24) holds and the solution to the FP-like equation (2.23) uniquely converges to some function. Suppose that the operator acting on ρ has an eigenvalue with a positive real part. Then the overall magnitude of ρ increases exponentially with t , and (2.24) cannot be satisfied. Also, suppose that the operator acting on ρ has an eigenvalue with a vanishing real part other than the zero eigenvalue corresponding to (2.26). Then the asymptotic behavior of ρ depends on the initial condition, and (2.24) cannot be satisfied with P having a unique asymptotic behavior. To the best of our knowledge, this argument has been given for the first time in Ref. [13] with explicit examples. Thus, provided that the relation (2.24) holds and the solution to the FP-like equation (2.23) uniquely converges to some function, we can calculate the VEV with respect to the partition function (2.1) as

$$\begin{aligned}
 \langle \mathcal{O} \rangle &= \int dx \mathcal{O}(x) \rho_{\text{time-indep}}(x) \\
 &= \lim_{t \rightarrow \infty} \int dx \mathcal{O}(x) \rho(x; t) \\
 &= \lim_{t \rightarrow \infty} \int dx dy \mathcal{O}(x + iy) P(x, y; t) \\
 &= \lim_{t \rightarrow \infty} \left\langle \mathcal{O}\left(x^{(\eta)}(t) + iy^{(\eta)}(t)\right) \right\rangle_{\eta} \\
 &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_{t_0}^{t_0+T} dt \mathcal{O}\left(x^{(\eta)}(t) + iy^{(\eta)}(t)\right). \tag{2.27}
 \end{aligned}$$

In what follows, we review the derivation³ of the key relation (2.24) given in Refs. [3,4]. At $t = 0$, we can choose

$$P(x, y; 0) = \rho(x; 0) \delta(y), \tag{2.28}$$

where $\rho(x; 0) \geq 0$ so that (2.24) holds trivially. In order to prove the relation (2.24) at arbitrary $t > 0$, we are going to show that each side of (2.24) can be rewritten as

$$\int dx dy \mathcal{O}(x + iy) P(x, y; t) = \int dx dy \mathcal{O}(x + iy; t) P(x, y; 0), \tag{2.29}$$

$$\int dx \mathcal{O}(x) \rho(x; t) = \int dx \mathcal{O}(x; t) \rho(x; 0). \tag{2.30}$$

In Eq. (2.29), we have introduced the time-dependent observables $\mathcal{O}(z; t)$ defined by solving

$$\frac{\partial}{\partial t} \mathcal{O}(z; t) = \tilde{L} \mathcal{O}(z; t), \tag{2.31}$$

$$\tilde{L} = \left(\frac{\partial}{\partial z_k} - \frac{\partial S}{\partial z_k} \right) \frac{\partial}{\partial z_k} \tag{2.32}$$

with the initial condition

$$\mathcal{O}(z; 0) = \mathcal{O}(z). \tag{2.33}$$

Let us recall that we are considering holomorphic observables $\mathcal{O}(z)$. One can actually show that the time-evolved observables $\mathcal{O}(z; t)$ remain holomorphic when $S(z)$ is a holomorphic function [3].

³ For earlier work on this issue, see Ref. [27].

The observables $\mathcal{O}(x; t)$ that appear in (2.30) are obtained by setting $y = 0$ in $\mathcal{O}(x + iy; t)$, and they satisfy the differential equation

$$\frac{\partial}{\partial t} \mathcal{O}(x; t) = L_0 \mathcal{O}(x; t), \quad (2.34)$$

$$L_0 = \left(\frac{\partial}{\partial x_k} - \frac{\partial S}{\partial x_k} \right) \frac{\partial}{\partial x_k}. \quad (2.35)$$

Since the right-hand sides of (2.29) and (2.30) are equal to each other due to (2.28), Eqs. (2.29) and (2.30) imply the desired relation (2.24).

In order to show (2.29), we introduce the function

$$F(t, \tau) = \int dx dy \mathcal{O}(x + iy; \tau) P(x, y; t - \tau), \quad (2.36)$$

which interpolates each side of (2.29) with $0 \leq \tau \leq t$. Taking the derivative with respect to τ , we get

$$\frac{\partial}{\partial \tau} F(t, \tau) = \int dx dy \tilde{L} \mathcal{O}(x + iy; \tau) P(x, y; t - \tau) - \int dx dy \mathcal{O}(x + iy; \tau) L^\top P(x, y; t - \tau), \quad (2.37)$$

where $L^\top$ denotes the operator acting on P on the right-hand side of (2.23). The operator L is then defined as an operator satisfying $\langle Lf, g \rangle = \langle f, L^\top g \rangle$, where $\langle f, g \rangle \equiv \int f(x, y) g(x, y) dx dy$, assuming that f and g are functions that allow integration by parts. The explicit form of the operator L can be obtained as

$$L = \left\{ -\text{Re} \left(\frac{\partial S}{\partial z_k} \right) + N_R \frac{\partial}{\partial x_k} \right\} \frac{\partial}{\partial x_k} + \left\{ -\text{Im} \left(\frac{\partial S}{\partial z_k} \right) + N_I \frac{\partial}{\partial y_k} \right\} \frac{\partial}{\partial y_k}. \quad (2.38)$$

An important observation here is that, when L acts on a holomorphic function $f(z)$ of z_k , it can be replaced by $\tilde{L}$, since

$$\begin{aligned} Lf(z) &= \left\{ -\text{Re} \left(\frac{\partial S}{\partial z_k} \right) + N_R \frac{\partial}{\partial z_k} \right\} \frac{\partial f}{\partial z_k} + \left\{ -\text{Im} \left(\frac{\partial S}{\partial z_k} \right) + i N_I \frac{\partial}{\partial z_k} \right\} \left(i \frac{\partial f}{\partial z_k} \right) \\ &= \left\{ -\frac{\partial S}{\partial z_k} + (N_R - N_I) \frac{\partial}{\partial z_k} \right\} \frac{\partial f}{\partial z_k} \\ &= \tilde{L} f(z), \end{aligned} \quad (2.39)$$

where we have used (2.22). This implies that $\tilde{L}$ in the first term of (2.37) can be replaced by L , and hence (2.37) vanishes if one can perform integration by parts. In that case, $F(t, \tau)$ is independent of τ , and (2.29) follows.

A similar argument can be used to show (2.30). We define

$$G(t, \tau) = \int dx \mathcal{O}(x; \tau) \rho(x; t - \tau), \quad (2.40)$$

which interpolates each side of (2.30) for $0 \leq \tau \leq t$. Taking the derivative with respect to τ , we get

$$\frac{\partial}{\partial \tau} G(t, \tau) = \int dx L_0 \mathcal{O}(x; \tau) \rho(x; t - \tau) - \int dx \mathcal{O}(x; \tau) L_0^\top \rho(x; t - \tau), \quad (2.41)$$

where we have used (2.34) and (2.25). Here, the integration on the right-hand side involves the real directions x_k only, so we can perform integration by parts without any problem due to the effects of the action, which make $\rho(x; t)$ well localized. Thus (2.41) vanishes, and (2.30) follows.

On the other hand, the integration by parts that one needs to use to show that (2.37) vanishes involves the imaginary directions y_k . It can therefore be justified only if the probability distribution $P(x, y; t)$ has a sharp fall-off in the imaginary directions [3,4].

Recently, it has been pointed out that the integration by parts can also be invalidated when the drift term includes a singularity [13]. This issue is relevant, in particular, to complex action systems involving fermions, such as finite-density QCD, since the fermion determinant gives rise to a singular drift term. The CLM still works if the probability distribution $P(x, y; t)$ is suppressed strongly enough near the singularity.

3. “Gauge cooling” in 0D systems

At the end of the previous section, we discussed two possible problems, which can make the CLM give wrong results. The gauge cooling was originally proposed to cure the first problem [5] in gauge theories. In K. Nagata et al. (manuscript in preparation), we propose that it can also be applied to cure the second problem, and demonstrate that it does in the random matrix theory for finite-density QCD. In this section we consider the “gauge cooling” in 0D systems such as the random matrix theory and provide an explicit justification. Apart from pedagogical purposes, we consider that it is useful, in particular, in matrix models relevant to superstring theory [15,16]. Generalization to the lattice gauge theory is straightforward and is given in Sects. 4 and 5.

3.1. Complexified symmetry

Let us consider a system of N real variables with a symmetry under

$$x'_j = g_{jk} x_k, \quad (3.1)$$

where g is a representation matrix of a Lie group. An infinitesimal transformation is denoted as

$$\delta x_j = i \lambda_{jk} x_k. \quad (3.2)$$

Here, λ is an element of the Lie algebra, which can be expanded as

$$\lambda_{jk} = \sum_a \lambda_a (t_a)_{jk} \quad (3.3)$$

in terms of the generators t_a of the Lie group under consideration with real coefficients $\lambda_a \in \mathbb{R}$. Upon complexifying the variables $x_k \mapsto z_k = x_k + i y_k$, the symmetry of the action and the observables naturally enhances from (3.1) to

$$z'_j = g_{jk} z_k, \quad (3.4)$$

where g is an element of the Lie group that can be obtained by complexifying the original Lie group. In particular, an infinitesimal transformation of the complexified symmetry is given by

$$\delta z_j = i \lambda_{jk} z_k. \quad (3.5)$$

Here λ is an element of the Lie algebra for the complexified Lie group, which can be expanded as (3.3) but now with complex coefficients $\lambda_a \in \mathbb{C}$.

As a simple example, let us consider an $O(N)$ vector model

$$S(x) = \sigma \sum_{k=1}^N (x_k)^2 + \kappa \left\{ \sum_{k=1}^N (x_k)^2 \right\}^2 \quad (3.6)$$

with $\sigma \in \mathbb{C}$, which is invariant under (3.1) with $g \in O(N)$. An infinitesimal transformation is given by (3.2), where λ_{jk} is a purely imaginary antisymmetric $N \times N$ matrix. Upon complexification

$x_k \mapsto z_k = x_k + iy_k$, the action becomes

$$S(z) = \sigma \sum_{k=1}^N (z_k)^2 + \kappa \left\{ \sum_{k=1}^N (z_k)^2 \right\}^2, \quad (3.7)$$

which is invariant under (3.4) with $g \in \mathrm{O}(N, \mathbb{C})$; namely, with g being an $N \times N$ complex matrix satisfying $g^\top g = gg^\top = \mathbf{1}$. (The symbol $g^\top$ here represents the transpose of the matrix g .) An infinitesimal transformation is given by (3.5), where λ_{jk} is a complex antisymmetric $N \times N$ matrix.

3.2. A modified complex Langevin equation

The discretized version of the complex Langevin equation (2.17) can be written as

$$z_k^{(\eta)}(t + \epsilon) = z_k^{(\eta)}(t) - \epsilon \frac{\partial S(z)}{\partial z_k} + \sqrt{\epsilon} \eta_k(t), \quad (3.8)$$

analogously to (2.11). The probabilistic variables

$$\eta_k(t) = \eta_k^{(\mathrm{R})}(t) + i\eta_k^{(\mathrm{I})}(t) \quad (3.9)$$

obey the probability distribution $e^{-\frac{1}{4} \sum_t \left\{ \frac{1}{N_R} \eta_k^{(\mathrm{R})}(t)^2 + \frac{1}{N_I} \eta_k^{(\mathrm{I})}(t)^2 \right\}}$. The gauge cooling [5] is a procedure of making a complexified symmetry transformation (3.4) between the Langevin steps. Thus it amounts to modifying the complex Langevin equation (3.8) into

$$\tilde{z}_k^{(\eta)}(t) = g_{kl} z_l^{(\eta)}(t), \quad (3.10)$$

$$z_k^{(\eta)}(t + \epsilon) = \tilde{z}_k^{(\eta)}(t) - \epsilon \frac{\partial S(\tilde{z})}{\partial \tilde{z}_k} + \sqrt{\epsilon} \eta_k(t), \quad (3.11)$$

where g is an element of the complexified Lie group chosen appropriately as a function of the configuration before cooling. The basic idea is to determine g in such a way that the modified Langevin process (3.10), (3.11) does not suffer from the problem of the original Langevin process (3.8). Clearly, this idea will only have a chance of working if the degrees of freedom in the symmetry transformation have at least the same order of magnitude as those of the dynamical system itself. Gauge theories are one such example, but 0D models such as vector models and matrix models would be equally good, as demonstrated explicitly in the random matrix theory (K. Nagata et al., manuscript in preparation).

For instance, if the excursions in the imaginary directions are problematic in studying the model (3.6) by the CLM, one can introduce the norm⁴

$$\mathcal{N} = \sum_{k=1}^N (y_k)^2 = -\frac{1}{4} \sum_{k=1}^N (z_k - z_k^*)^2, \quad (3.12)$$

which measures the distance from the real region, and determine the transformation g in (3.10) in such a way that the norm $\mathcal{N}$ for $z^{(\eta)}(t)$ is reduced by the transformation. In K. Nagata et al. (manuscript in preparation), we propose to combine this norm with another norm to also cure the problem caused by a singular drift term. Typically, the norm that one tries to reduce is invariant under transformations in the original Lie group but not under transformations in the complexified

⁴ This is analogous to the so-called “unitarity norm” [5] proposed in the complex Langevin simulation of lattice gauge theory.

Lie group, as in the case of (3.12). The main issue that we address below is whether the modification of the Langevin process by “gauge cooling” spoils the equivalence to the path integral reviewed in Sect. 2.3.

Note that gauge cooling is a completely deterministic procedure. In particular, the transformation g in (3.10) is determined only by the configuration $z^{(\eta)}(t)$ before cooling. Therefore, for our purpose, it is convenient to regard (3.10), (3.11) as describing the t -evolution of $z_k^{(\eta)}(t)$ only⁵.

3.3. Justification for infinitesimal transformation

In this section, we discuss the justification of gauge cooling, assuming for simplicity that the asymptotic behavior of g in the $\epsilon \rightarrow 0$ limit is given by

$$g = \exp \left\{ i \epsilon \lambda(x^{(\eta)}(t), y^{(\eta)}(t)) \right\}, \quad (3.13)$$

where $\lambda(x, y)$ is an element of the Lie algebra of $O(N, \mathbb{C})$. For instance, one may use

$$\lambda_{kl}(x, y) = \alpha(x, y) \left(x_k y_l - y_k x_l \right) = \frac{1}{2} i \alpha(x, y) \left(z_k z_l^* - z_k^* z_l \right), \quad (3.14)$$

which can be obtained by calculating the gradient of the norm (3.12) with respect to the $O(N, \mathbb{C})$ transformation. The real positive function $\alpha(x, y)$ can be chosen to optimize the reduction of the norm. Note that $\lambda(x, y)$ is not a holomorphic function of z_k in general, as in (3.14).

Using (3.13) in Eqs. (3.10), (3.11) and taking the $\epsilon \rightarrow 0$ limit, we obtain the continuum complex Langevin equation for $z^{(\eta)}(t)$ as

$$\dot{z}_k^{(\eta)}(t) = -\frac{\partial S}{\partial z_k} + \eta_k(t) + i \lambda_{kl}(x^{(\eta)}(t), y^{(\eta)}(t)) z_l^{(\eta)}(t), \quad (3.15)$$

where the effect of the gauge cooling is the infinitesimal transformation represented by the last term on the right-hand side. Then we can easily find that the FP-like equation (2.23) that $P(x, y; t)$ satisfies is modified by the gauge cooling as

$$\begin{aligned} \frac{\partial P}{\partial t} = & \frac{\partial}{\partial x_k} \left\{ \operatorname{Re} \left(\frac{\partial S}{\partial z_k} - i \lambda_{kl}(x, y) z_l \right) + N_R \frac{\partial}{\partial x_k} \right\} P \\ & + \frac{\partial}{\partial y_k} \left\{ \operatorname{Im} \left(\frac{\partial S}{\partial z_k} - i \lambda_{kl}(x, y) z_l \right) + N_I \frac{\partial}{\partial y_k} \right\} P. \end{aligned} \quad (3.16)$$

This modifies the differential operator L in Eq. (2.38) into

$$\begin{aligned} L' = & \left\{ -\operatorname{Re} \left(\frac{\partial S}{\partial z_k} - i \lambda_{kl}(x, y) z_l \right) + N_R \frac{\partial}{\partial x_k} \right\} \frac{\partial}{\partial x_k} \\ & + \left\{ -\operatorname{Im} \left(\frac{\partial S}{\partial z_k} - i \lambda_{kl}(x, y) z_l \right) + N_I \frac{\partial}{\partial y_k} \right\} \frac{\partial}{\partial y_k}. \end{aligned} \quad (3.17)$$

Acting this operator L' on a holomorphic function $f(z)$, we obtain

$$\begin{aligned} L' f(z) = & \left(-\frac{\partial S}{\partial z_k} + i \lambda_{kl}(x, y) z_l + (N_R - N_I) \frac{\partial}{\partial z_k} \right) \frac{\partial}{\partial z_k} f(z) \\ = & \tilde{L} f(z) + i \lambda_{kl}(x, y) z_l \frac{\partial}{\partial z_k} f(z), \end{aligned} \quad (3.18)$$

⁵ In practice, one usually measures observables using the configuration $\tilde{z}^{(\eta)}(t)$ after cooling instead of $z^{(\eta)}(t)$. This does not cause any problem since $z^{(\eta)}(t)$ and $\tilde{z}^{(\eta)}(t)$ are related to each other by the complexified symmetry transformation (3.4), under which the observables are invariant.

where $\tilde{L}$ is defined by (2.32). The extra term compared with (2.39) represents the change of $f(z)$ under an infinitesimal $O(N, \mathbb{C})$ transformation. Note that the time-evolved observables $\mathcal{O}(z; t)$ defined by (2.31) remain invariant under the $O(N, \mathbb{C})$ transformation as long as the action $S(z)$ and the original observables $\mathcal{O}(z)$ are invariant. Therefore, the $\tilde{L}$ in the first term of (2.37) can be replaced by L' . Hence (2.37) vanishes if one can perform integration by parts for L' and the modified P . In that case, the crucial identity (2.24) holds for the modified P with the same ρ . Thus we have shown explicitly that “gauge cooling” provides the possibility of improving the property of the probability distribution $P(x, y; t)$ so that (2.24) holds, without affecting the FP equation (2.25) for $\rho(x; t)$.

3.4. Justification for finite transformation

In practical applications, the asymptotic behavior (3.13) of g in the $\epsilon \rightarrow 0$ limit may not be satisfied. Therefore, it is important to discuss the justification of the gauge cooling without assuming it. In this case, we cannot take the $\epsilon \rightarrow 0$ limit of the complex Langevin equation (3.10), (3.11) to arrive at the continuum version (3.15). Therefore, we have to deal with the discretized version (3.10), (3.11). While our argument becomes slightly more complicated, we can still justify the gauge cooling, as we see below.

First let us derive the discretized FP-like equation for $P(x, y; t)$ in a similar way to what we did in Sect. 2.2. Let us consider a test function $f(x, y)$ and its expectation value

$$\left\langle f\left(x^{(\eta)}(t), y^{(\eta)}(t)\right)\right\rangle_{\eta} = \int dx dy f(x, y) P(x, y; t) \quad (3.19)$$

at a fictitious time t . The t -evolution of this quantity is given by

$$\begin{aligned} & \left\langle f\left(x^{(\eta)}(t+\epsilon), y^{(\eta)}(t+\epsilon)\right)\right\rangle_{\eta} \\ &= \left\langle f\left(\operatorname{Re}\left(\tilde{z}^{(\eta)}(t)\right), \operatorname{Im}\left(\tilde{z}^{(\eta)}(t)\right)\right)\right\rangle_{\eta} \\ &+ \left\langle -\epsilon \left\{ \frac{\partial f}{\partial x_k} \operatorname{Re}\left(\frac{\partial S}{\partial z_k}\right) \right\} \right\rangle_{\tilde{z}^{(\eta)}} + \frac{1}{2} \frac{\partial^2 f}{\partial x_k \partial x_l} \Big|_{\tilde{z}^{(\eta)}} (\sqrt{\epsilon})^2 \eta_k^{(\mathrm{R})}(t) \eta_l^{(\mathrm{R})}(t) \Big\rangle_{\eta} \\ &+ \left\langle -\epsilon \left\{ \frac{\partial f}{\partial y_k} \operatorname{Im}\left(\frac{\partial S}{\partial z_k}\right) \right\} \right\rangle_{\tilde{z}^{(\eta)}} + \frac{1}{2} \frac{\partial^2 f}{\partial y_k \partial y_l} \Big|_{\tilde{z}^{(\eta)}} (\sqrt{\epsilon})^2 \eta_k^{(\mathrm{I})}(t) \eta_l^{(\mathrm{I})}(t) \Big\rangle_{\eta} + \cdots \\ &= \left\langle \left\{ :e^{\epsilon L}: f(x, y) \right\} \right\rangle_{\tilde{z}^{(\eta)}}, \end{aligned} \quad (3.20)$$

where the operator L is defined by (2.38), and the symbol $: \cdots :$ implies that the operators are ordered in such a way that derivative operators appear on the right; e.g., $:(f(x) + \partial)^2: = f(x)^2 + 2f(x)\partial + \partial^2$. We can rewrite the last expression as

$$\begin{aligned} & \left\langle \left\{ :e^{\epsilon L}: f(x, y) \right\} \right\rangle_{\tilde{z}^{(\eta)}} \\ &= \int dx dy \left\{ :e^{\epsilon L}: f(x, y) \right\} \Big|_{z=z^{(\mathrm{g})}} P(x, y; t) \\ &= \int dx dy \int d\tilde{x} d\tilde{y} \prod_k \left\{ \delta\left(\tilde{x}_k - \operatorname{Re}\left(z_k^{(\mathrm{g})}\right)\right) \delta\left(\tilde{y}_k - \operatorname{Im}\left(z_k^{(\mathrm{g})}\right)\right) \right\} \left\{ :e^{\epsilon L}: f(x, y) \right\} \Big|_{z=\tilde{z}} P(x, y; t) \end{aligned}$$

$$\begin{aligned}
&= \int dx dy \left\{ :e^{\epsilon L}: f(x, y) \right\} \tilde{P}(x, y; t) \\
&= \int dx dy f(x, y) \left(:e^{\epsilon L}: \right)^\top \tilde{P}(x, y; t),
\end{aligned} \tag{3.21}$$

where $z_k^{(g)} = g_{kl}(x, y)z_l$. In the third equality, we have relabeled $(\tilde{x}, \tilde{y})$ as (x, y) and defined a function

$$\tilde{P}(\tilde{x}, \tilde{y}; t) = \int dx dy P(x, y; t) \prod_k \left\{ \delta\left(\tilde{x}_k - \text{Re}(g_{kl}(x, y)z_l)\right) \delta\left(\tilde{y}_k - \text{Im}(g_{kl}(x, y)z_l)\right) \right\}, \tag{3.22}$$

which is nothing but the probability distribution of $\tilde{x}^{(\eta)}(t)$ and $\tilde{y}^{(\eta)}(t)$ defined similarly to (2.19). Using (3.19), the same quantity (3.21) should be written as

$$\left\langle f\left(x^{(\eta)}(t + \epsilon), y^{(\eta)}(t + \epsilon)\right) \right\rangle_\eta = \int dx dy f(x, y) P(x, y; t + \epsilon). \tag{3.23}$$

Since (3.21) and (3.23) should be equal for an arbitrary $f(x, y)$, one obtains

$$P(x, y; t + \epsilon) = \left(:e^{\epsilon L}: \right)^\top \tilde{P}(x, y; t). \tag{3.24}$$

Note that the effect of the gauge cooling comes only through $\tilde{P}(x, y; t)$ in (3.24). When we perform the gauge cooling in such a way that a norm like (3.12) is *strictly* minimized, the distribution $\tilde{P}(x, y; t)$ becomes degenerate since it is nonzero only for configurations that minimize the norm with respect to the $O(N, \mathbb{C})$ transformation. The operator on the right-hand side of (3.24) diffuses the distribution to a width of $O(\sqrt{\epsilon})$ in the degenerating direction, which is due to the noise term in the Langevin equation (3.11). In this situation, one cannot expand (3.24) with respect to ϵ and truncate the series at a finite order. Therefore we need to deal with expressions like (3.24), which make sense at finite ϵ .

Let us then define the function $F(t, \tau)$ by (2.36), where τ is now discretized similarly to t . The discretized t -evolution of the operator $\mathcal{O}(x + iy; t)$ is defined as

$$\mathcal{O}(z; t + \epsilon) = :e^{\epsilon \tilde{L}}: \mathcal{O}(z; t), \tag{3.25}$$

which reduces to the continuum version (2.31) in the $\epsilon \rightarrow 0$ limit. The initial condition is given by (2.33) as before. Using (3.24), we get

$$\begin{aligned}
F(t, \tau - \epsilon) &= \int dx dy \mathcal{O}(x + iy; \tau - \epsilon) \tilde{P}(x, y; t - \tau + \epsilon) \\
&= \int dx dy \mathcal{O}(x + iy; \tau - \epsilon) \left(:e^{\epsilon L}: \right)^\top \tilde{P}(x, y; t - \tau) \\
&= \int dx dy \left\{ :e^{\epsilon L}: \mathcal{O}(x + iy; \tau - \epsilon) \right\} \tilde{P}(x, y; t - \tau) \\
&= \int dx dy \left\{ :e^{\epsilon \tilde{L}}: \mathcal{O}(x + iy; \tau - \epsilon) \right\} \tilde{P}(x, y; t - \tau) \\
&= \int dx dy \mathcal{O}(x + iy; \tau) \tilde{P}(x, y; t - \tau) \\
&= \int dx dy \mathcal{O}\left(z^{(g)}; \tau\right) P(x, y; t - \tau) \\
&= F(t, \tau).
\end{aligned} \tag{3.26}$$

The fourth equality follows from (2.39), and, in the last equality, we have used the $O(N, \mathbb{C})$ symmetry $\mathcal{O}(z^{(g)}; \tau) = \mathcal{O}(z; \tau)$ of the observable. Therefore, $F(t, \tau)$ is constant in τ , which, in particular,

implies $F(t, 0) = F(t, t)$. Thus we have shown that (2.29) holds at finite ϵ . The rest of the arguments for the justification are the same as in Sect. 2.3.

4. Application of the CLM to lattice gauge theory

In this section, we discuss the application of the CLM to lattice gauge theory, which is defined by the partition function

$$Z = \int dU e^{-S(U)} = \int \prod_{n\mu} dU_{n\mu} e^{-S(U)}, \quad (4.1)$$

where the action S is a complex-valued function of the configuration $U = \{U_{n\mu}\}$, composed of link variables $U_{n\mu} \in \text{SU}(3)$, and the integration measure $dU_{n\mu}$ represents the Haar measure for the $\text{SU}(3)$ group. The only complication compared with the case discussed in the previous sections comes from the fact that the dynamical variables take values on a group manifold. The Langevin equation in such a case with a real action is discussed intensively in Refs. [28–32]. Using this formulation, we can easily generalize our discussions to the case of lattice gauge theory.

4.1. Description of the method

When the action S is complex, the drift term in the Langevin equation makes the link variables evolve into $\text{SL}(3, \mathbb{C})$ matrices (i.e., 3×3 general complex matrices with the determinant one) even if one starts from a configuration of $\text{SU}(3)$ matrices. Let us therefore complexify the link variables as $U_{n\mu} \in \text{SL}(3, \mathbb{C})$, and solve the complex Langevin equation

$$\dot{U}_{n\mu}^{(\eta)}(t) = i \sum_a \left(-\mathcal{D}_{an\mu} S(\mathcal{U}) + \eta_{an\mu}(t) \right) t_a U_{n\mu}^{(\eta)}(t), \quad (4.2)$$

where the action $S(\mathcal{U})$ is now considered as a holomorphic function of the complexified configuration $\mathcal{U}_{n\mu}$, and t_a are the generators of the $\text{SU}(3)$ group normalized by $\text{tr}(t_a t_b) = \delta_{ab}$. The probabilistic variables $\eta_{an\mu}(t)$ are defined similarly to (3.9). The derivative operator $\mathcal{D}_{an\mu}$ is defined as⁶

$$\mathcal{D}_{an\mu} = \frac{1}{2} \left(\mathcal{D}_{an\mu}^{(\text{R})} - i \mathcal{D}_{an\mu}^{(\text{I})} \right), \quad (4.3)$$

$$\mathcal{D}_{an\mu}^{(\text{R})} f(\mathcal{U}) = \left. \frac{\partial}{\partial x} f(e^{ix t_a} U_{n\mu}) \right|_{x=0}, \quad (4.4)$$

$$\mathcal{D}_{an\mu}^{(\text{I})} f(\mathcal{U}) = \left. \frac{\partial}{\partial y} f(e^{-y t_a} U_{n\mu}) \right|_{y=0}. \quad (4.5)$$

Here $f(\mathcal{U})$ are functions on the complexified group manifold, which are not necessarily holomorphic, and x and y in Eqs. (4.4) and (4.5) are real parameters. Note that, for a holomorphic function $f(\mathcal{U})$, we have $\bar{\mathcal{D}}_{an\mu} f(\mathcal{U}) = 0$, where

$$\bar{\mathcal{D}}_{an\mu} = \frac{1}{2} \left(\mathcal{D}_{an\mu}^{(\text{R})} + i \mathcal{D}_{an\mu}^{(\text{I})} \right), \quad (4.6)$$

and hence

$$\mathcal{D}_{an\mu}^{(\text{R})} f(\mathcal{U}) = \mathcal{D}_{an\mu} f(\mathcal{U}), \quad \mathcal{D}_{an\mu}^{(\text{I})} f(\mathcal{U}) = i \mathcal{D}_{an\mu} f(\mathcal{U}). \quad (4.7)$$

⁶ The derivative operators defined in Eqs. (4.3) and (4.6) may be regarded as analogues of $\frac{\partial}{\partial z} = \frac{1}{2} \left(\frac{\partial}{\partial x} - i \frac{\partial}{\partial y} \right)$ and $\frac{\partial}{\partial \bar{z}} = \frac{1}{2} \left(\frac{\partial}{\partial x} + i \frac{\partial}{\partial y} \right)$, respectively.

Then we define the probability distribution

$$P(\mathcal{U}; t) = \left\langle \prod_{n\mu} \delta(\mathcal{U}_{n\mu}, \mathcal{U}_{n\mu}^{(\eta)}(t)) \right\rangle_{\eta}, \quad (4.8)$$

where the delta function is defined by

$$\int d\mathcal{U} f(\mathcal{U}) \delta(\mathcal{U}_{n\mu}, \tilde{\mathcal{U}}_{n\mu}) = f(\tilde{\mathcal{U}}) \quad (4.9)$$

for any function $f(\mathcal{U})$. The integration measure that appears on the left-hand side represents the Haar measure for the $\text{SL}(3, \mathbb{C})$ group normalized appropriately. One can show that the probability distribution $P(\mathcal{U}; t)$ obeys the FP-like equation (see Appendix B for the derivation):

$$\frac{\partial P}{\partial t} = \mathcal{D}_{an\mu}^{(R)} \left\{ \text{Re}(\mathcal{D}_{an\mu} S(\mathcal{U})) + N_R \mathcal{D}_{an\mu}^{(R)} \right\} P + \mathcal{D}_{an\mu}^{(I)} \left\{ \text{Im}(\mathcal{D}_{an\mu} S(\mathcal{U})) + N_I \mathcal{D}_{an\mu}^{(I)} \right\} P. \quad (4.10)$$

In fact, for observables $\mathcal{O}(U)$ that admit holomorphic extension to $\mathcal{O}(\mathcal{U})$, one can show under certain conditions that there exists a complex function $\rho(U; t)$, which satisfies

$$\int d\mathcal{U} \mathcal{O}(\mathcal{U}) P(\mathcal{U}; t) = \int dU \mathcal{O}(U) \rho(U; t), \quad (4.11)$$

and obeys the FP equation

$$\frac{\partial}{\partial t} \rho(U; t) = D_{an\mu} (D_{an\mu} S(U) + D_{an\mu}) \rho(U; t). \quad (4.12)$$

Here we have defined the derivative operator $D_{an\mu}$, which acts on a function $f(U)$ of the unitary gauge configuration as

$$D_{an\mu} f(U) = \left. \frac{\partial}{\partial x} f(e^{ix t_a} U_{n\mu}) \right|_{x=0}. \quad (4.13)$$

Note that the FP equation (4.12) has a time-independent solution

$$\rho_{\text{time-indep}}(U) = \frac{1}{Z} \exp(-S(U)). \quad (4.14)$$

As we argued in Sect. 2.3, the convergence to (4.14) should occur if the relation (4.11) holds and the FP-like equation (4.10) uniquely converges to some function. In that case, we can calculate the VEV with respect to the partition function (4.1) as

$$\begin{aligned} \langle \mathcal{O} \rangle &= \int dU \mathcal{O}(U) \rho_{\text{time-indep}}(U) \\ &= \lim_{t \rightarrow \infty} \int dU \mathcal{O}(U) \rho(U; t) \\ &= \lim_{t \rightarrow \infty} \int d\mathcal{U} \mathcal{O}(\mathcal{U}) P(\mathcal{U}; t) \\ &= \lim_{t \rightarrow \infty} \left\langle \mathcal{O}(\mathcal{U}^{(\eta)}(t)) \right\rangle_{\eta} \\ &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_{t_0}^{t_0+T} dt \mathcal{O}(\mathcal{U}^{(\eta)}(t)). \end{aligned} \quad (4.15)$$

4.2. Proof of the key relation

Let us briefly discuss how one can derive the relation (4.11). At $t = 0$, we choose

$$P(\mathcal{U}, ; 0) = \int dU \rho(U; 0) \prod_{n\mu} \delta(\mathcal{U}_{n\mu}, U_{n\mu}) \quad (4.16)$$

with $\rho(U; 0) \geq 0$ so that (4.11) holds trivially. In order to prove the relation (4.11) at arbitrary $t > 0$, we are going to show that each side of (4.11) can be rewritten as

$$\int d\mathcal{U} \mathcal{O}(\mathcal{U}) P(\mathcal{U}; t) = \int d\mathcal{U} \mathcal{O}(\mathcal{U}; t) P(\mathcal{U}; 0), \quad (4.17)$$

$$\int dU \mathcal{O}(U) \rho(U; t) = \int dU \mathcal{O}(U; t) \rho(U; 0). \quad (4.18)$$

In Eq. (4.17), we have introduced the time-dependent observables $\mathcal{O}(\mathcal{U}; t)$ defined by solving

$$\frac{\partial}{\partial t} \mathcal{O}(\mathcal{U}; t) = \tilde{L} \mathcal{O}(\mathcal{U}; t), \quad (4.19)$$

$$\tilde{L} = \left(\mathcal{D}_{an\mu} - \mathcal{D}_{an\mu} S(\mathcal{U}) \right) \mathcal{D}_{an\mu} \quad (4.20)$$

with the initial condition

$$\mathcal{O}(\mathcal{U}; 0) = \mathcal{O}(\mathcal{U}). \quad (4.21)$$

Let us recall that we are considering holomorphic observables $\mathcal{O}(\mathcal{U})$. One can actually show that the time-evolved observables $\mathcal{O}(\mathcal{U}; t)$ remain holomorphic when $S(\mathcal{U})$ is a holomorphic function [3]. The observables $\mathcal{O}(U; t)$ that appear in (4.18) are obtained by setting $\mathcal{U} = U$ in $\mathcal{O}(\mathcal{U}; t)$, and they satisfy the differential equation

$$\frac{\partial}{\partial t} \mathcal{O}(U; t) = L_0 \mathcal{O}(U; t), \quad (4.22)$$

$$L_0 = \left(D_{an\mu} - D_{an\mu} S(U) \right) D_{an\mu}. \quad (4.23)$$

Since the right-hand sides of (4.17) and (4.18) are equal to each other due to (4.16), Eqs. (4.17) and (4.18) imply the desired relation (4.11).

In order to show (4.17), we introduce the function

$$F(t, \tau) = \int d\mathcal{U} \mathcal{O}(\mathcal{U}; \tau) P(\mathcal{U}; t - \tau), \quad (4.24)$$

which interpolates each side of (4.17) with $0 \leq \tau \leq t$. Taking the derivative with respect to τ , we get

$$\frac{\partial}{\partial \tau} F(t, \tau) = \int d\mathcal{U} \tilde{L} \mathcal{O}(\mathcal{U}; \tau) P(\mathcal{U}; t - \tau) - \int d\mathcal{U} \mathcal{O}(\mathcal{U}; \tau) L^\top P(\mathcal{U}; t - \tau), \quad (4.25)$$

where $L^\top$ denotes the operator acting on P on the right-hand side of (4.10). The operator L is then defined as an operator satisfying $\langle Lf, g \rangle = \langle f, L^\top g \rangle$, where $\langle f, g \rangle \equiv \int f(\mathcal{U}) g(\mathcal{U}) d\mathcal{U}$, assuming that f and g are functions that allow integration by parts. The explicit form of the operator L can be

obtained as

$$L = \left\{ -\text{Re}(\mathcal{D}_{an\mu} S(\mathcal{U})) + N_R \mathcal{D}_{an\mu}^{(R)} \right\} \mathcal{D}_{an\mu}^{(R)} + \left\{ -\text{Im}(\mathcal{D}_{an\mu} S(\mathcal{U})) + N_I \mathcal{D}_{an\mu}^{(I)} \right\} \mathcal{D}_{an\mu}^{(I)}. \quad (4.26)$$

An important observation here is that, when L acts on a holomorphic function $f(\mathcal{U})$, it can be replaced by $\tilde{L}$, since

$$\begin{aligned} Lf(\mathcal{U}) &= \left\{ -\text{Re}(\mathcal{D}_{an\mu} S(\mathcal{U})) + N_R \mathcal{D}_{an\mu} \right\} \mathcal{D}_{an\mu} f(\mathcal{U}) \\ &\quad + \left\{ -\text{Im}(\mathcal{D}_{an\mu} S(\mathcal{U})) + i N_I \mathcal{D}_{an\mu} \right\} i \mathcal{D}_{an\mu} f(\mathcal{U}) \\ &= \left\{ -\mathcal{D}_{an\mu} S(\mathcal{U}) + (N_R - N_I) \mathcal{D}_{an\mu} \right\} \mathcal{D}_{an\mu} f(\mathcal{U}) \\ &= \tilde{L} f(\mathcal{U}), \end{aligned} \quad (4.27)$$

where we have used (4.7) and (2.22). This implies that $\tilde{L}$ in the first term of (4.25) can be replaced by L , and hence (4.25) vanishes if one can perform integration by parts. In that case, $F(t, \tau)$ is independent of τ , and (4.17) follows.

A similar argument can be used to show (4.18). We define

$$G(t, \tau) = \int dU \mathcal{O}(U; \tau) \rho(U; t - \tau), \quad (4.28)$$

which interpolates each side of (4.18) for $0 \leq \tau \leq t$. Taking the derivative with respect to τ , we get

$$\frac{\partial}{\partial \tau} G(t, \tau) = \int dU L_0 \mathcal{O}(U; \tau) \rho(U; t - \tau) - \int dU \mathcal{O}(U; \tau) L_0^\top \rho(U; t - \tau), \quad (4.29)$$

where we have used (4.22) and (4.12). Here the integration on the right-hand side involves the real directions only, which are compact in the present case, so we can perform integration by parts without any problem to show that (4.29) vanishes. Thus $G(t, \tau)$ is independent of τ , and (4.18) follows.

On the other hand, the integration by parts in (4.25) is justified only if the probability distribution $P(\mathcal{U}; t)$ has a sharp fall-off in the noncompact imaginary directions. The gauge cooling [5] was originally proposed to solve this problem. As we mentioned in Sect. 2.3, the integration by parts can also be invalidated when the drift term includes a singularity [13]. This problem is anticipated to occur when one applies the CLM to finite-density QCD at low temperature with light quarks. The CLM still works if the probability distribution $P(\mathcal{U}; t)$ is suppressed strongly enough near the singularity. We consider that the gauge cooling is also useful in solving this problem as in the case of random matrix theory (K. Nagata et al., manuscript in preparation).

5. Gauge cooling in lattice gauge theory

In this section, we discuss gauge cooling in lattice gauge theory and provide the justification of the CLM including gauge cooling. The argument is a straightforward generalization of that given in Sect. 3 for the 0D model.

5.1. Complexified gauge symmetry

The lattice gauge theory is invariant under the SU(3) gauge transformation. For instance, the plaquette action

$$S_{\text{plaquette}}(U) = -\beta \sum_n \sum_{\mu \neq \nu} \text{tr} \left(U_{n\mu} U_{n+\hat{\mu}, \nu} U_{n+\hat{\nu}, \mu}^{-1} U_{n\nu}^{-1} \right) \quad (5.1)$$

is invariant under

$$U'_{n\mu} = g_n U_{n\mu} g_{n+\hat{\mu}}^{-1}, \quad (5.2)$$

where $g_n \in \text{SU}(3)$. An infinitesimal transformation is denoted as

$$\delta U_{n\mu} = i \left(\lambda_n U_{n\mu} - U_{n\mu} \lambda_{n+\hat{\mu}} \right). \quad (5.3)$$

Here λ_n is an element of the Lie algebra, which can be expanded as

$$(\lambda_n)_{jk} = \sum_a \lambda_{na} (t_a)_{jk} \quad (5.4)$$

in terms of the generators t_a of $\text{SU}(3)$ with real coefficients $\lambda_{na} \in \mathbb{R}$.

When one complexifies the variables $U_{n\mu} \mapsto \mathcal{U}_{n\mu} \in \text{SL}(3, \mathbb{C})$, the symmetry of the action and the observables is naturally enhanced to the $\text{SL}(3, \mathbb{C})$ gauge symmetry that can be obtained by complexifying the original Lie group. For instance, the plaquette action (5.1) becomes

$$S_{\text{plaquette}}(\mathcal{U}) = -\beta \sum_n \sum_{\mu \neq \nu} \text{tr} \left(\mathcal{U}_{n\mu} \mathcal{U}_{n+\hat{\mu},\nu} \mathcal{U}_{n+\hat{\nu},\mu}^{-1} \mathcal{U}_{n\nu}^{-1} \right), \quad (5.5)$$

which is invariant under

$$\mathcal{U}'_{n\mu} = g_n \mathcal{U}_{n\mu} g_{n+\hat{\mu}}^{-1} \quad (5.6)$$

with $g_n \in \text{SL}(3, \mathbb{C})$. An infinitesimal transformation is given by

$$\delta \mathcal{U}_{n\mu} = i \left(\lambda_n \mathcal{U}_{n\mu} - \mathcal{U}_{n\mu} \lambda_{n+\hat{\mu}} \right). \quad (5.7)$$

Here λ is an element of the Lie algebra for the complexified Lie group, which can be expanded as (5.4) but now with complex coefficients $\lambda_{na} \in \mathbb{C}$.

5.2. A modified complex Langevin equation

The discretized version of the complex Langevin equation (4.2) can be written as

$$\mathcal{U}_{n\mu}^{(\eta)}(t + \epsilon) = \exp \left\{ i \sum_a \left(-\epsilon \mathcal{D}_{an\mu} S(\mathcal{U}) + \sqrt{\epsilon} \eta_{an\mu}(t) \right) t_a \right\} \mathcal{U}_{n\mu}^{(\eta)}(t). \quad (5.8)$$

The gauge cooling [5] modifies the complex Langevin equation (5.8) into

$$\tilde{\mathcal{U}}_{n\mu}^{(\eta)}(t) = g_n \mathcal{U}_{n\mu}^{(\eta)}(t) g_{n+\hat{\mu}}^{-1}, \quad (5.9)$$

$$\mathcal{U}_{n\mu}^{(\eta)}(t + \epsilon) = \exp \left\{ i \sum_a \left(-\epsilon \mathcal{D}_{an\mu} S(\tilde{\mathcal{U}}) + \sqrt{\epsilon} \eta_{an\mu}(t) \right) t_a \right\} \tilde{\mathcal{U}}_{n\mu}^{(\eta)}(t), \quad (5.10)$$

where g_n is an element of the complexified Lie group. The basic idea is to determine g_n in such a way that the modified Langevin process (5.10) does not suffer from the problem of the original Langevin process (5.8).

For instance, if the excursions in the imaginary directions are problematic, one can introduce a positive semidefinite quantity [33] (we call it the “norm” in this paper):

$$\mathcal{N} = \sum_{n\mu} \text{tr} \left(\mathcal{U}_{n\mu}^\dagger \mathcal{U}_{n\mu} - \mathbf{1} \right), \quad (5.11)$$

which measures the distance from the unitary region, and determine the transformation g_n in (5.10) in such a way that the norm $\mathcal{N}$ for $\mathcal{U}^{(\eta)}(t)$ is reduced by the transformation. Typically, the norm

that one tries to reduce is invariant under transformations in the original Lie group but not under transformations in the complexified Lie group, as in the case of (5.11). Below, we demonstrate that the modification of the Langevin process by gauge cooling does not spoil the equivalence to the path integral reviewed in the previous section.

Note that gauge cooling is a completely deterministic procedure. In particular, the transformation g_n in (5.10) is determined only by the configuration $\mathcal{U}^{(\eta)}(t)$ before cooling. Therefore, for our purpose, it is more convenient to regard (5.10) as describing the t -evolution of $\mathcal{U}^{(\eta)}(t)$ only⁷.

5.3. Justification for infinitesimal transformation

In what follows, we assume for simplicity that the asymptotic behavior of g_n in the $\epsilon \rightarrow 0$ limit is given by

$$g_n = \exp \left\{ i \epsilon \lambda_n \left(\mathcal{U}^{(\eta)}(t) \right) \right\}, \quad (5.12)$$

where $\lambda_n(\mathcal{U})$ is an element of the Lie algebra of $\text{SL}(N, \mathbb{C})$. For instance, one may use

$$\lambda_n(\mathcal{U}) = i \alpha(\mathcal{U}) \sum_{\mu} \left\{ \left(\mathcal{U}_{n\mu} \mathcal{U}_{n\mu}^\dagger - \mathcal{U}_{n-\hat{\mu},\mu}^\dagger \mathcal{U}_{n-\hat{\mu},\mu} \right) - (\text{trace part}) \right\}, \quad (5.13)$$

which can be obtained by calculating the gradient of the norm (5.11) with respect to the $\text{SL}(3, \mathbb{C})$ gauge transformation of the configuration $\mathcal{U}$. The real positive function $\alpha(\mathcal{U})$, which is not necessarily holomorphic, can be chosen to optimize the reduction of the norm. Note that $\lambda_n(\mathcal{U})$ is not a holomorphic function of $\mathcal{U}_{n\mu}$ in general, as in (5.13).

Using (5.12) in Eqs. (5.9), (5.10) and taking the $\epsilon \rightarrow 0$ limit, we obtain the continuum complex Langevin equation for $\mathcal{U}^{(\eta)}(t)$ as

$$\begin{aligned} \dot{\mathcal{U}}_{n\mu}^{(\eta)}(t) = & i \sum_a \left(-\mathcal{D}_{an\mu} S(\mathcal{U}) + \eta_{an\mu}(t) \right) t_a \mathcal{U}_{n\mu}^{(\eta)}(t) \\ & + i \left\{ \lambda_n \left(\mathcal{U}^{(\eta)}(t) \right) \mathcal{U}_{n\mu}^{(\eta)}(t) - \mathcal{U}_{n\mu}^{(\eta)}(t) \lambda_{n+\hat{\mu}} \left(\mathcal{U}^{(\eta)}(t) \right) \right\}, \end{aligned} \quad (5.14)$$

where the effect of the gauge cooling is represented by the last term on the right-hand side. Then we can easily find that the FP-like equation (4.10) that $P(\mathcal{U}; t)$ satisfies is modified by the gauge cooling as

$$\begin{aligned} \frac{\partial P}{\partial t} = & \mathcal{D}_{an\mu}^{(\text{R})} \left\{ \text{Re} \left(\mathcal{D}_{an\mu} S(\mathcal{U}) - \mathcal{C}_{an\mu} \right) + N_{\text{R}} \mathcal{D}_{an\mu}^{(\text{R})} \right\} P \\ & + \mathcal{D}_{an\mu}^{(\text{I})} \left\{ \text{Im} \left(\mathcal{D}_{an\mu} S(\mathcal{U}) - \mathcal{C}_{an\mu} \right) + N_{\text{I}} \mathcal{D}_{an\mu}^{(\text{I})} \right\} P, \end{aligned} \quad (5.15)$$

$$\mathcal{C}_{an\mu} = \text{tr} \left\{ t_a \left(\lambda_n(\mathcal{U}) - \mathcal{U}_{n\mu} \lambda_{n+\hat{\mu}}(\mathcal{U}) \mathcal{U}_{n\mu}^{-1} \right) \right\}. \quad (5.16)$$

This modifies the differential operator L in Eq. (4.26) into

$$\begin{aligned} L' = & \left\{ -\text{Re} \left(\mathcal{D}_{an\mu} S(\mathcal{U}) - \mathcal{C}_{an\mu} \right) + N_{\text{R}} \mathcal{D}_{an\mu}^{(\text{R})} \right\} \mathcal{D}_{an\mu}^{(\text{R})} \\ & + \left\{ -\text{Im} \left(\mathcal{D}_{an\mu} S(\mathcal{U}) - \mathcal{C}_{an\mu} \right) + N_{\text{I}} \mathcal{D}_{an\mu}^{(\text{I})} \right\} \mathcal{D}_{an\mu}^{(\text{I})}. \end{aligned} \quad (5.17)$$

⁷ In practice, one usually measures observables using the configuration $\tilde{\mathcal{U}}^{(\eta)}(t)$ after cooling instead of $\mathcal{U}^{(\eta)}(t)$. This does not cause any problem, since $\mathcal{U}^{(\eta)}(t)$ and $\tilde{\mathcal{U}}^{(\eta)}(t)$ are related to each other by the complexified symmetry transformation (5.6), under which the observables are invariant.

Acting this operator L' on a holomorphic function $f(\mathcal{U})$, we obtain

$$\begin{aligned} L' f(\mathcal{U}) &= \left\{ -\mathcal{D}_{an\mu} S(\mathcal{U}) + \mathcal{C}_{an\mu} + (N_R - N_I) \mathcal{D}_{an\mu} \right\} \mathcal{D}_{an\mu} f(\mathcal{U}) \\ &= \tilde{L} f(\mathcal{U}) + \mathcal{C}_{an\mu} \mathcal{D}_{an\mu} f(\mathcal{U}), \end{aligned} \quad (5.18)$$

where $\tilde{L}$ is defined by (4.20). The extra term compared with (4.27) represents the change of $f(\mathcal{U})$ under an infinitesimal $\text{SL}(3, \mathbb{C})$ gauge transformation. Note that the time-evolved observables $\mathcal{O}(\mathcal{U}; t)$ defined by (4.19) remain invariant under the $\text{SL}(3, \mathbb{C})$ gauge transformation as long as the action $S(\mathcal{U})$ and the original observables $\mathcal{O}(\mathcal{U})$ are invariant. Therefore, the $\tilde{L}$ in the first term of (4.25) can be replaced by L' . Hence (4.25) vanishes if one can perform integration by parts for L' and the modified P . In that case, the crucial identity (4.11) holds for the modified P with the same ρ . Thus we have shown explicitly that gauge cooling provides the possibility of improving the property of the probability distribution $P(\mathcal{U}; t)$ so that (4.11) holds, without affecting the FP equation (4.12) for $\rho(\mathcal{U}; t)$.

5.4. Justification for finite transformation

In practical applications, the asymptotic behavior (5.12) of g in the $\epsilon \rightarrow 0$ limit may not be satisfied. Therefore, it is important to discuss the justification of the gauge cooling without assuming it. In this case, we cannot take the $\epsilon \rightarrow 0$ limit of the complex Langevin equation (5.9), (5.10) to arrive at the continuum version (5.14), and therefore we have to deal with the discretized version (5.9), (5.10). While our argument becomes slightly more complicated, we can still justify the gauge cooling, as we see below.

First let us derive the discretized FP-like equation for $P(\mathcal{U}; t)$ in a similar way to what we did in Sect. 2.2. Let us consider a test function $f(\mathcal{U})$ and its expectation value

$$\left\langle f(\mathcal{U}^{(\eta)}(t)) \right\rangle_\eta = \int d\mathcal{U} f(\mathcal{U}) P(\mathcal{U}; t) \quad (5.19)$$

at a fictitious time t . The t -evolution of this quantity is given by

$$\begin{aligned} \left\langle f(\mathcal{U}^{(\eta)}(t + \epsilon)) \right\rangle_\eta &= \left\langle f(\tilde{\mathcal{U}}^{(\eta)}(t)) \right\rangle_\eta \\ &\quad + \left\langle -\epsilon \left\{ \mathcal{D}_{an\mu}^{(R)} f \operatorname{Re}(\mathcal{D}_{an\mu} S) \right\} \right\rangle_{\tilde{\mathcal{U}}^{(\eta)}} \\ &\quad + \frac{1}{2} \mathcal{D}_{an\mu}^{(R)} \mathcal{D}_{an\mu}^{(R)} f \Big|_{\tilde{\mathcal{U}}^{(\eta)}} (\sqrt{\epsilon})^2 \eta_{an\mu}^{(R)}(t) \eta_{an\mu}^{(R)}(t) \Big\rangle_\eta \\ &\quad + \left\langle -\epsilon \left\{ \mathcal{D}_{an\mu}^{(I)} f \operatorname{Im}(\mathcal{D}_{an\mu} S) \right\} \right\rangle_{\tilde{\mathcal{U}}^{(\eta)}} \\ &\quad + \frac{1}{2} \mathcal{D}_{an\mu}^{(I)} \mathcal{D}_{an\mu}^{(I)} f \Big|_{\tilde{\mathcal{U}}^{(\eta)}} (\sqrt{\epsilon})^2 \eta_{an\mu}^{(I)}(t) \eta_{an\mu}^{(I)}(t) \Big\rangle_\eta + \cdots \\ &= \left\langle \left\{ :e^{\epsilon L}: f(\mathcal{U}) \right\} \right\rangle_{\tilde{\mathcal{U}}^{(\eta)}}, \end{aligned} \quad (5.20)$$

where the operator L is defined by (4.26). We can rewrite the last expression as

$$\begin{aligned}
 \left\langle \left\{ :e^{\epsilon L}: f(\mathcal{U}) \right\} \right\rangle_{\tilde{\mathcal{U}}^{(\eta)}} &= \int d\mathcal{U} \left\{ :e^{\epsilon L}: f(\mathcal{U}) \right\} \Big|_{\mathcal{U}=\mathcal{U}^{(g)}} P(\mathcal{U}; t) \\
 &= \int d\mathcal{U} \int d\tilde{\mathcal{U}} \prod_{n\mu} \delta(\tilde{\mathcal{U}}_{n\mu}, \mathcal{U}_{n\mu}^{(g)}) \left\{ :e^{\epsilon L}: f(\mathcal{U}) \right\} \Big|_{\tilde{\mathcal{U}}} P(\mathcal{U}; t) \\
 &= \int d\mathcal{U} \left\{ :e^{\epsilon L}: f(\mathcal{U}) \right\} \tilde{P}(\mathcal{U}; t) \\
 &= \int d\mathcal{U} f(\mathcal{U}) (:e^{\epsilon L}:)^\top \tilde{P}(\mathcal{U}; t),
 \end{aligned} \tag{5.21}$$

where $\mathcal{U}_{n\mu}^{(g)} = g_n(\mathcal{U}) \mathcal{U}_{n\mu} g_{n+\hat{\mu}}^{-1}(\mathcal{U})$. In the third equality, we have relabeled $\tilde{\mathcal{U}}$ as $\mathcal{U}$ and defined a function

$$\tilde{P}(\tilde{\mathcal{U}}; t) = \int d\mathcal{U} P(\mathcal{U}; t) \prod_{n\mu} \delta(\tilde{\mathcal{U}}_{n\mu}, g_n(\mathcal{U}) \mathcal{U}_{n\mu} g_{n+\hat{\mu}}^{-1}(\mathcal{U})), \tag{5.22}$$

which is nothing but the probability distribution of $\tilde{\mathcal{U}}^{(\eta)}(t)$ defined similarly to (4.8). Using (5.19), the same quantity (5.21) should be written as

$$\left\langle f(\mathcal{U}^{(\eta)}(t + \epsilon)) \right\rangle_\eta = \int d\mathcal{U} f(\mathcal{U}) P(\mathcal{U}; t + \epsilon). \tag{5.23}$$

Since (5.21) and (5.23) should be equal for an arbitrary $f(\mathcal{U})$, one obtains

$$P(\mathcal{U}; t + \epsilon) = (:e^{\epsilon L}:)^\top \tilde{P}(\mathcal{U}; t). \tag{5.24}$$

Let us then define the function $F(t, \tau)$ by (4.24), where τ is now discretized similarly to t . The discretized t -evolution of the operator $\mathcal{O}(x + iy; t)$ is defined as

$$\mathcal{O}(\mathcal{U}; t + \epsilon) = :e^{\epsilon \tilde{L}}: \mathcal{O}(\mathcal{U}; t), \tag{5.25}$$

which reduces to the continuum version (4.19) in the $\epsilon \rightarrow 0$ limit. The initial condition is given by (2.33) as before. Using (5.24), we get

$$\begin{aligned}
 F(t, \tau - \epsilon) &= \int d\mathcal{U} \mathcal{O}(\mathcal{U}; \tau - \epsilon) \tilde{P}(\mathcal{U}; t - \tau + \epsilon) \\
 &= \int d\mathcal{U} \mathcal{O}(\mathcal{U}; \tau - \epsilon) (:e^{\epsilon L}:)^\top \tilde{P}(\mathcal{U}; t - \tau) \\
 &= \int d\mathcal{U} \left\{ :e^{\epsilon L}: \mathcal{O}(\mathcal{U}; \tau - \epsilon) \right\} \tilde{P}(\mathcal{U}; t - \tau) \\
 &= \int d\mathcal{U} \left\{ :e^{\epsilon \tilde{L}}: \mathcal{O}(\mathcal{U}; \tau - \epsilon) \right\} \tilde{P}(\mathcal{U}; t - \tau) \\
 &= \int d\mathcal{U} \mathcal{O}(\mathcal{U}; \tau) \tilde{P}(\mathcal{U}; t - \tau) \\
 &= \int d\mathcal{U} \mathcal{O}(\mathcal{U}^{(g)}; \tau) P(\mathcal{U}; t - \tau) \\
 &= F(t, \tau).
 \end{aligned} \tag{5.26}$$

The fourth equality follows from (4.27), and, in the last equality, we have used $\mathcal{O}(\mathcal{U}^{(g)}; \tau) = \mathcal{O}(\mathcal{U}; \tau)$ due to the $\text{SL}(N, \mathbb{C})$ symmetry. Therefore, $F(t, \tau)$ is constant in τ , which, in particular, implies $F(t, 0) = F(t, t)$. Thus we have shown that (4.17) holds at finite ϵ . The rest of the arguments for the justification are the same as in Sect. 4.2.

6. Summary and discussions

In this paper, we have provided an explicit justification of the CLM with the gauge cooling procedure. As we have reviewed in detail, the CLM relies crucially on the relation between the probability distribution P associated with the complex Langevin process and the complex weight ρ associated with the original path integral problem. This relation holds if and only if the probability distribution P satisfies the following two properties. One is that it is strongly suppressed for complexified configurations that have large imaginary parts [3,4]. The other is that the distribution is strongly suppressed for complexified configurations that make the drift term large [13]. Since the gauge cooling modifies the probability distribution P , one may hope to make it satisfy the above two properties by appropriately choosing the complexified symmetry transformation to be used in the cooling procedure. What we have shown in this paper is that the modification of the probability distribution P due to the gauge cooling does not alter the FP equation that the complex weight ρ obeys if the relation between P and ρ holds at all.

For a long time, it has been thought that the convergence of the FP equation for the complex weight ρ is not guaranteed, unlike in the real action case. However, once the relation between the probability distribution P and the complex weight ρ is established, one may argue [13] that the unique convergence of the probability distribution P already implies the unique convergence of ρ to the desired complex weight e^{-S} . Therefore, if one can satisfy the above two properties of the probability distribution P by using the gauge cooling appropriately, the CLM is guaranteed to give the correct results.

While the gauge cooling certainly enlarges the range of applicability of the CLM, it remains to be seen how powerful it is in studying various interesting systems with complex actions. In this regard, our results for the random matrix theory using gauge cooling with a new type of norm (K. Nagata et al., manuscript in preparation) look very promising.

Acknowledgements

The authors would like to thank J. Bloch, K. Fukushima, and D. Sexty for valuable discussions. We are also grateful to E. Seiler for correspondence on the first version of this paper. K.N. was supported by JSPS Grants-in-Aid for Scientific Research (Kakenhi) Grants No. 00586901, MEXT SPIRE, and JICFuS. The work of J.N. was supported in part by a Grant-in-Aid for Scientific Research (No. 23244057) from the Japan Society for the Promotion of Science.

Funding

Open Access funding: SCOAP³.

Appendix A. Derivation of Eq. (3.20)

In this appendix, we derive Eq. (3.20) by performing the integration over the Gaussian noise explicitly.

Let us first rewrite (3.20) as

$$\begin{aligned} & \left\langle f\left(x^{(\eta)}(t+\epsilon), y^{(\eta)}(t+\epsilon)\right) \right\rangle_{\eta} \\ &= \sum_{pqrs} \frac{\epsilon^{p+q+r+s}}{p!q!(2r)!(2s)!} \left\langle \left(u_k \frac{\partial}{\partial x_k}\right)^p \left(v_k \frac{\partial}{\partial y_k}\right)^q \left(\eta_k^{(R)}(t) \frac{\partial}{\partial x_k}\right)^{2r} \left(\eta_k^{(I)}(t) \frac{\partial}{\partial y_k}\right)^{2s} f(x, y) \right\rangle_{\tilde{z}^{(\eta)}_{\eta}} \end{aligned} \quad (\text{A1})$$

where we have defined

$$u_k = -\text{Re} \left(\frac{\partial S}{\partial z_k} \right) \Big|_{\tilde{z}^{(\eta)}}, \quad v_k = -\text{Im} \left(\frac{\partial S}{\partial z_k} \right) \Big|_{\tilde{z}^{(\eta)}}. \quad (\text{A2})$$

As in Sect. 2.2, $\tilde{z}^{(\eta)}$ depends only on $\eta(0), \eta(\epsilon), \dots, \eta(t - \epsilon)$, but not on $\eta(t)$. Therefore, the integration over $\eta(t)$ can be performed separately using the following formula:

$$\begin{aligned} \left\langle \left(a_k \eta_k^{(\text{R})}(t) \right)^{2r} \right\rangle_\eta &= \left\{ 4N_{\text{R}}(a_k)^2 \right\}^r \frac{\Gamma(r + \frac{1}{2})}{\Gamma(\frac{1}{2})}, \\ \left\langle \left(b_k \eta_k^{(\text{I})}(t) \right)^{2s} \right\rangle_\eta &= \left\{ 4N_{\text{I}}(b_k)^2 \right\}^s \frac{\Gamma(s + \frac{1}{2})}{\Gamma(\frac{1}{2})}. \end{aligned} \quad (\text{A3})$$

Thus we arrive at

$$\begin{aligned} &\left\langle f \left(x^{(\eta)}(t + \epsilon), y^{(\eta)}(t + \epsilon) \right) \right\rangle_\eta \\ &= \sum_{pqrs} \frac{\epsilon^{p+q+r+s}}{p!q!r!s!} \left\langle \left(u_k \frac{\partial}{\partial x_k} \right)^p \left(v_k \frac{\partial}{\partial y_k} \right)^q \left(N_{\text{R}} \left(\frac{\partial}{\partial x_k} \right)^2 \right)^r \left(N_{\text{I}} \left(\frac{\partial}{\partial y_k} \right)^2 \right)^s f(x, y) \Big|_{\tilde{z}^{(\eta)}} \right\rangle_\eta \\ &= \left\langle \exp \left\{ \epsilon \left(u_k \frac{\partial}{\partial x_k} + v_k \frac{\partial}{\partial y_k} + N_{\text{R}} \left(\frac{\partial}{\partial x_k} \right)^2 + N_{\text{I}} \left(\frac{\partial}{\partial y_k} \right)^2 \right) \right\} f(x, y) \Big|_{\tilde{z}^{(\eta)}} \right\rangle_\eta \\ &= \left\langle \left\{ :e^{\epsilon L}: f(x, y) \right\} \Big|_{\tilde{z}^{(\eta)}} \right\rangle_\eta. \end{aligned} \quad (\text{A4})$$

Appendix B. Derivation of Eq. (4.10)

In this appendix, we derive the FP-like equation (4.10) in the case of lattice gauge theory to make this paper self-contained. Here we deal with continuous t for simplicity, but one can make a similar analysis with discretized t , as we have done in Sect. 2.2.

Let us consider a test function $f(\mathcal{U})$ and its expectation value:

$$\left\langle f \left(\mathcal{U}^{(\eta)}(t) \right) \right\rangle_\eta = \int d\mathcal{U} f(\mathcal{U}) P(\mathcal{U}; t). \quad (\text{B1})$$

Taking the derivative with respect to the fictitious time t , we get

$$\frac{d}{dt} \left\langle f \left(\mathcal{U}^{(\eta)}(t) \right) \right\rangle_\eta = \int d\mathcal{U} f(\mathcal{U}) \frac{d}{dt} P(\mathcal{U}; t). \quad (\text{B2})$$

The left-hand side can be evaluated as follows:

$$\begin{aligned} \frac{d}{dt} \left\langle f \left(\mathcal{U}^{(\eta)}(t) \right) \right\rangle_\eta &= \left\langle \left\{ -\text{Re} \left(\mathcal{D}_{an\mu} S(\mathcal{U}^{(\eta)}(t)) \right) + \eta_{an\mu}^{(\text{R})}(t) \right\} \mathcal{D}_{an\mu}^{(\text{R})} f \left(\mathcal{U}^{(\eta)}(t) \right) \right\rangle_\eta \\ &\quad + \left\langle \left\{ -\text{Im} \left(\mathcal{D}_{an\mu} S(\mathcal{U}^{(\eta)}(t)) \right) + \eta_{an\mu}^{(\text{I})}(t) \right\} \mathcal{D}_{an\mu}^{(\text{I})} f \left(\mathcal{U}^{(\eta)}(t) \right) \right\rangle_\eta. \end{aligned} \quad (\text{B3})$$

Here we use the following formula (see, e.g., Ref. [17]):

$$\left\langle g \left(\mathcal{U}^{(\eta)}(t) \right) \eta_{an\mu}^{(\text{R})}(t) \right\rangle_\eta = \left\langle 2N_{\text{R}} \frac{\delta}{\delta \eta_{an\mu}^{(\text{R})}(t)} g \left(\mathcal{U}^{(\eta)}(t) \right) \right\rangle_\eta = \left\langle N_{\text{R}} \mathcal{D}_{an\mu}^{(\text{R})} g \left(\mathcal{U}^{(\eta)}(t) \right) \right\rangle_\eta, \quad (\text{B4})$$

$$\left\langle g \left(\mathcal{U}^{(\eta)}(t) \right) \eta_{an\mu}^{(\text{I})}(t) \right\rangle_\eta = \left\langle 2N_{\text{I}} \frac{\delta}{\delta \eta_{an\mu}^{(\text{I})}(t)} g \left(\mathcal{U}^{(\eta)}(t) \right) \right\rangle_\eta = \left\langle N_{\text{I}} \mathcal{D}_{an\mu}^{(\text{I})} g \left(\mathcal{U}^{(\eta)}(t) \right) \right\rangle_\eta. \quad (\text{B5})$$

Using (B4) and (B5) in (B3), we get

$$\begin{aligned}
 \frac{d}{dt} \left\langle f \left(\mathcal{U}^{(\eta)}(t) \right) \right\rangle_{\eta} &= - \left\langle \text{Re} \left(\mathcal{D}_{an\mu} S(\mathcal{U}^{(\eta)}(t)) \right) \mathcal{D}_{an\mu}^{(R)} f \left(\mathcal{U}^{(\eta)}(t) \right) \right\rangle_{\eta} + \left\langle N_R \mathcal{D}_{an\mu}^{(R)} \mathcal{D}_{an\mu}^{(R)} f \left(\mathcal{U}^{(\eta)}(t) \right) \right\rangle_{\eta} \\
 &\quad - \left\langle \text{Im} \left(\mathcal{D}_{an\mu} S(\mathcal{U}^{(\eta)}(t)) \right) \mathcal{D}_{an\mu}^{(I)} f \left(\mathcal{U}^{(\eta)}(t) \right) \right\rangle_{\eta} + \left\langle N_I \mathcal{D}_{an\mu}^{(I)} \mathcal{D}_{an\mu}^{(I)} f \left(\mathcal{U}^{(\eta)}(t) \right) \right\rangle_{\eta} \\
 &= \int d\mathcal{U} P(\mathcal{U}; t) \left[- \text{Re} \left(\mathcal{D}_{an\mu} S(\mathcal{U}) \right) \mathcal{D}_{an\mu}^{(R)} f(\mathcal{U}) + N_R \mathcal{D}_{an\mu}^{(R)} \mathcal{D}_{an\mu}^{(R)} f(\mathcal{U}) \right. \\
 &\quad \left. - \text{Im} \left(\mathcal{D}_{an\mu} S(\mathcal{U}) \right) \mathcal{D}_{an\mu}^{(I)} f(\mathcal{U}) + N_I \mathcal{D}_{an\mu}^{(I)} \mathcal{D}_{an\mu}^{(I)} f(\mathcal{U}) \right] \\
 &= \int d\mathcal{U} f(\mathcal{U}) \left[\mathcal{D}_{an\mu}^{(R)} \left\{ \text{Re} \left(\mathcal{D}_{an\mu} S(\mathcal{U}) \right) + N_R \mathcal{D}_{an\mu}^{(R)} \right\} P \right. \\
 &\quad \left. + \mathcal{D}_{an\mu}^{(I)} \left\{ \text{Im} \left(\mathcal{D}_{an\mu} S(\mathcal{U}) \right) + N_I \mathcal{D}_{an\mu}^{(I)} \right\} P \right]. \tag{B6}
 \end{aligned}$$

Plugging this expression in (B2), and using the fact that (B2) should hold for an arbitrary $f(\mathcal{U})$, we get Eq. (4.10).

References

- [1] G. Parisi, Phys. Lett. B **131**, 393 (1983).
- [2] J. R. Klauder, Phys. Rev. A **29**, 2036 (1984).
- [3] G. Aarts, E. Seiler, and I. O. Stamatescu, Phys. Rev. D **81**, 054508 (2010) [[arXiv:0912.3360](#) [hep-lat]] [[Search INSPIRE](#)].
- [4] G. Aarts, F. A. James, E. Seiler, and I. O. Stamatescu, Eur. Phys. J. C **71**, 1756 (2011) [[arXiv:1101.3270](#) [hep-lat]] [[Search INSPIRE](#)].
- [5] E. Seiler, D. Sexty, and I. O. Stamatescu, Phys. Lett. B **723**, 213 (2013) [[arXiv:1211.3709](#) [hep-lat]] [[Search INSPIRE](#)].
- [6] G. Aarts, E. Seiler, D. Sexty, and I. O. Stamatescu, Phys. Rev. D **90**, 114505 (2014) [[arXiv:1408.3770](#) [hep-lat]] [[Search INSPIRE](#)].
- [7] D. Sexty, Phys. Lett. B **729**, 108 (2014) [[arXiv:1307.7748](#) [hep-lat]] [[Search INSPIRE](#)].
- [8] H. Makino, H. Suzuki, and D. Takeda, [[arXiv:1503.00417](#) [hep-lat]] [[Search INSPIRE](#)].
- [9] A. Mollgaard and K. Splittorff, Phys. Rev. D **88**, 116007 (2013) [[arXiv:1309.4335](#) [hep-lat]] [[Search INSPIRE](#)].
- [10] A. Mollgaard and K. Splittorff, Phys. Rev. D **91**, 036007 (2015) [[arXiv:1412.2729](#) [hep-lat]] [[Search INSPIRE](#)].
- [11] J. Greensite, Phys. Rev. D **90**, 114507 (2014) [[arXiv:1406.4558](#) [hep-lat]] [[Search INSPIRE](#)].
- [12] E. Seiler, “Langevin with meromorphic drift: problems and partial solutions”, lecture at EMMI Workshop: SIGN 2014, 18–21 February 2014.
- [13] J. Nishimura and S. Shimasaki, Phys. Rev. D **92**, 011501 (2015) [[arXiv:1504.08359](#) [hep-lat]] [[Search INSPIRE](#)].
- [14] G. Aarts, L. Bongiovanni, E. Seiler, D. Sexty, and I. O. Stamatescu, Eur. Phys. J. A **49**, 89 (2013) [[arXiv:1303.6425](#) [hep-lat]] [[Search INSPIRE](#)].
- [15] S. W. Kim, J. Nishimura, and A. Tsuchiya, Phys. Rev. Lett. **108**, 011601 (2012) [[arXiv:1108.1540](#) [hep-th]] [[Search INSPIRE](#)].
- [16] K. N. Anagnostopoulos, T. Azuma, and J. Nishimura, J. High Energy Phys. **1311**, 009 (2013) [[arXiv:1306.6135](#) [hep-th]] [[Search INSPIRE](#)].
- [17] P. H. Damgaard and H. Huffel, Phys. Rept. **152**, 227 (1987).
- [18] G. Parisi and Y. S. Wu, Sci. Sin. **24**, 483 (1981).
- [19] M. Fukugita, Y. Oyanagi, and A. Ukawa, Phys. Rev. D **36**, 824 (1987).
- [20] E. Witten, [[arXiv:1001.2933](#) [hep-th]] [[Search INSPIRE](#)].
- [21] M. Cristoforetti, F. Di Renzo, and L. Scorzato, Phys. Rev. D **86**, 074506 (2012) [[arXiv:1205.3996](#) [hep-lat]] [[Search INSPIRE](#)].
- [22] M. Cristoforetti, F. Di Renzo, A. Mukherjee, and L. Scorzato, Phys. Rev. D **88**, 051501 (2013) [[arXiv:1303.7204](#) [hep-lat]] [[Search INSPIRE](#)].

- [23] H. Fujii, D. Honda, M. Kato, Y. Kikukawa, S. Komatsu, and T. Sano, J. High Energy Phys. **1310**, 147 (2013) [[arXiv:1309.4371](#) [hep-lat]] [[Search INSPIRE](#)].
- [24] A. Mukherjee and M. Cristoforetti, Phys. Rev. B **90**, 035134 (2014) [[arXiv:1403.5680](#) [cond-mat.str-el]] [[Search INSPIRE](#)].
- [25] F. Di Renzo and G. Eruzzi, [[arXiv:1507.03858](#) [hep-lat]] [[Search INSPIRE](#)].
- [26] K. Fukushima and Y. Tanizaki, [[arXiv:1507.07351](#) [hep-th]] [[Search INSPIRE](#)].
- [27] J. Ambjorn and S. K. Yang, Phys. Lett. B **165**, 140 (1985).
- [28] J. Alfaro and B. Sakita, CCNY-HEP-82-16, C82-09-09.1.
- [29] I. T. Drummond, S. Duane, and R. R. Horgan, Nucl. Phys. B **220**, 119 (1983).
- [30] A. Guha and S. C. Lee, Phys. Rev. D **27**, 2412 (1983).
- [31] M. B. Halpern, Nucl. Phys. B **228**, 173 (1983).
- [32] G. G. Batrouni, H. Kawai, and P. Rossi, J. Math. Phys. **27**, 1646 (1986).
- [33] G. Aarts and I. O. Stamatescu, J. High Energy Phys. **0809**, 018 (2008) [[arXiv:0807.1597](#) [hep-lat]] [[Search INSPIRE](#)].