

The Mechanism of Complex Langevin Simulations

H. Gausterer¹ and Sean Lee²

Received November 9, 1992

We discuss conditions under which expectation values computed from a complex Langevin process Z will converge to integral averages over a given complex-valued weight function. The difficulties in proving a general result are pointed out. For complex-valued polynomial actions, it is shown that for a process converging to a strongly stationary process one gets the correct answer for averages of polynomials if $c_t(k) \equiv E(e^{ikZ(t)})$ satisfies certain conditions. If these conditions are not satisfied, then the stochastic process is not necessarily described by a complex Fokker-Planck equation. The result is illustrated with the exactly solvable complex frequency harmonic oscillator.

KEY WORDS: Complex Langevin methods; Fokker-Planck equation.

1. INTRODUCTION

It is well known that a straightforward application of standard simulation techniques like the Metropolis *et al.*⁽¹⁾ algorithm will fail when they are applied to problems with a complex action or Hamiltonian. This is due to the fact that there is no direct probabilistic interpretation of a distribution function of the form e^{-S} when the action S is complex valued. As an alternative to such algorithms, the complex Langevin (CL) method was first proposed by Klauder⁽²⁾ and subsequently studied by many authors. The central idea of CL is based on the fact that for a Langevin equation there is no formal restriction to a real-valued drift term. The use of a complex drift term provides CL with a genuine advantage over other methods; namely, since CL uses the entire complex action S to define a stochastic process, it can in principle converge directly to the desired distribution. This potential for circumventing the well-known "sign problems" of other

¹ Institut für Theoretische Physik, Universität Graz, A-8010 Graz, Austria.

² Department of Physics, University of Florida, Gainesville, Florida.

standard algorithms is just one reason CL continues to be a subject of great interest.

Quite independent of its utility as a numerical technique is the interesting fact that under certain conditions a system governed by a complex Hamiltonian can nevertheless be given a probabilistic interpretation. Indeed, the name complex Langevin may be slightly misleading, since the Langevin equations really describe a real diffusion process in twice as many dimensions.

Unfortunately, there is currently no complete theory of the CL method. Many conditions under which a real Langevin process can be shown to converge to a given distribution are not satisfied for a general CL process. Furthermore, from the point of view of numerical simulations, in some cases CL is known to converge to the wrong results.^(3,4) For simple actions, this truant behavior can be corrected by an appropriate choice of kernel in the Langevin equation,⁽⁵⁾ but for more general systems, in particular lattice fermion models of current interest, it is far from clear which choice of kernel is required.

The purpose of this paper is to explore in a rigorous fashion the conditions under which the CL process correctly simulates a given system with a complex Hamiltonian defined on a Euclidean space $x \in \mathbb{R}^n$. While a general theorem is still lacking, we will demonstrate a set of sufficient conditions for polynomial actions.

2. THE PROCESS

Throughout the paper we will assume that the system is described by variables $x \in \mathbb{R}^n$ with some complex action or Hamiltonian $S: \mathbb{R}^n \rightarrow \mathbb{C}$, where $\text{Re } S$ is bounded from below. For simplicity the discussion will be restricted to one degree of freedom, since the following statements allow for an immediate generalization to \mathbb{R}^n . The quantities of physical interest are of the form

$$\langle g(x) \rangle = \frac{1}{\mathcal{N}} \int_{\mathbb{R}} g(x) e^{-S(x)} dx \quad (2.1)$$

$$\mathcal{N} = \int_{\mathbb{R}} e^{-S(x)} dx \quad (2.2)$$

assuming that for the partition function \mathcal{N} we have $0 < |\mathcal{N}| < \infty$. Both $g(z)$ and $S(z)$ are assumed to be analytic in \mathbb{C} . Since $S(z)$ is analytic, this provides a local Lipschitz condition for the Langevin equation (2.8); that is, (2.8) has a unique local solution that is defined up to a random

explosion time.⁽⁶⁾ For real actions $S \in \mathbb{R}$ we define a process $\{X(\tau), \tau \geq 0\}$ by the Langevin equation

$$dX(\tau) = F(X(\tau)) d\tau + dW(\tau) \quad (2.3)$$

with the drift term

$$F(x) = -\frac{1}{2} \frac{dS(x)}{dx} \quad (2.4)$$

where $W(\tau)$ is a standard Wiener process with zero mean and covariance

$$\langle W(\tau_1) W(\tau_2) \rangle = \min(\tau_1, \tau_2) \quad (2.5)$$

The probability density $f(x, \tau)$ for such a process will converge pointwise to the desired distribution function. That is,

$$\lim_{\tau \rightarrow \infty} f(x, \tau) = \hat{f}(x) \quad \text{a.e.} \quad (2.6)$$

with

$$\hat{f}(x) = \frac{1}{\mathcal{N}} e^{-S(x)} \quad (2.7)$$

As already mentioned in the introduction, for the complex case one can formally construct a Langevin equation

$$dZ(\tau) = F(Z(\tau)) d\tau + dW(\tau) \quad (2.8)$$

with the drift term

$$F(z) = -\frac{1}{2} \frac{dS(z)}{dz} \quad (2.9)$$

As above, $W(\tau)$ is a standard Wiener process. This is actually a two-dimensional process of the form

$$dX(\tau) = G(X(\tau), Y(\tau)) d\tau + dW(\tau) \quad (2.10)$$

$$dY(\tau) = H(X(\tau), Y(\tau)) d\tau \quad (2.11)$$

with $S(z) = u(x, y) + iv(x, y)$,

$$G(x, y) = -\frac{1}{2} \frac{\partial u(x, y)}{\partial x}, \quad H(x, y) = \frac{1}{2} \frac{\partial u(x, y)}{\partial y} \quad (2.12)$$

Note that the equation for dY has a zero diffusion coefficient, but is nevertheless a stochastic equation through its dependence on X .

The process $\{(X(\tau), Y(\tau)), \tau \geq 0\}$ as defined by Eq. (2.8) has a distribution density $f(x, y, \tau)$. There are now two crucial questions concerning this process. The first question is whether this so defined process converges in distribution at all to some (X, Y) ,

$$\lim_{\tau \rightarrow \infty} f(x, y, \tau) = \hat{f}(x, y), \quad \text{a.e.} \quad (2.13)$$

The second question has to do with the problem of whether $\hat{f}(x, y)$ satisfies

$$E(g(X + iY)) = \int_{\mathbb{R}^2} g(x + iy) \hat{f}(x, y) dx dy = \frac{1}{\mathcal{N}} \int_{\mathbb{R}} g(x) e^{-S(x)} dx \quad (2.14)$$

This equation contains the essence of complex Langevin. It tells us that if the process as defined above has converged in distribution, we might be able to calculate $\langle g(x) \rangle$ by an equivalent probabilistic system of twice as many dimensions.

Before we investigate Eq. (2.14) it is necessary to discuss the asymptotic behavior of $X(\tau)$, $Y(\tau)$ for $\tau \rightarrow \infty$. To do this we will examine the equivalent Fokker–Planck (FP) equation. First note that since the diffusion and drift coefficients are independent of τ , the process $(X(\tau), Y(\tau))$ is a homogeneous diffusion process. For twice continuously differentiable distribution densities with respect to x, y and once with respect to τ , there exists a FP equation.⁽⁶⁾ Then the equivalent FP equation for the above Langevin equation is

$$\frac{\partial f(x, y, \tau)}{\partial \tau} = Tf(x, y, \tau) \quad (2.15)$$

with

$$T = -G_x(x, y) - G(x, y) \frac{\partial}{\partial x} - H_y(x, y) - H(x, y) \frac{\partial}{\partial y} + \frac{1}{2} \frac{\partial^2}{\partial x^2} \quad (2.16)$$

As may be seen from the above equation, this case only requires a continuous first-order derivative with respect to y . Let us first assume that T has a unique stationary solution $\hat{f}(x, y)$,

$$T\hat{f}(x, y) = 0 \quad (2.17)$$

with

$$\hat{f}(x, y) \geq 0, \quad \text{a.e.} \quad (2.18)$$

and

$$\int_{\mathbb{R}^2} \hat{f}(x, y) dx dy = 1 \tag{2.19}$$

One can then use $\hat{f}(x, y)$ to define an invariant measure

$$\mu(\Omega) = \int_{\Omega} \hat{f}(x, y) dx dy \tag{2.20}$$

with respect to $L_{\tau} = \exp(Q\tau)$, $\tau \geq 0$, which is defined as an operator family in $L^p(\mathbb{R}^2, d\mu)$ ($p = 1, 2$). The operator Q is obtained by transforming T to $L^1(\mathbb{R}^2, d\mu)$ by

$$Q = \hat{f}^{-1} T \hat{f} \tag{2.21}$$

Then there is a theorem which tells us that $\{L_{\tau}, \tau \geq 0\}$ is a contraction semigroup,⁽⁷⁾ from which it follows that for any function $\phi \in L^p(\mathbb{R}^2, d\mu)$, $p = 1, 2$, one has convergence in the strong sense

$$s\text{-}\lim_{\tau \rightarrow \infty} L_{\tau} \phi = c_{\phi} \tag{2.22}$$

where c_{ϕ} is a constant. Since $\mu(\mathbb{R}^2) < \infty$, any distribution density $f(x, y, \tau)$ converges pointwise to the stationary solution

$$\lim_{\tau \rightarrow \infty} f(x, y, \tau) = \hat{f}(x, y) \text{ a.e.} \tag{2.23}$$

On the other hand, if the zero eigenvalue of T is M -fold degenerate, then there are M ergodic classes. In this case we have

$$\lim_{\tau \rightarrow \infty} f(x, y, \tau) = \sum_{i=1}^M c_i \hat{f}_i(x, y) \text{ a.e.} \tag{2.24}$$

Unfortunately, it is still an open question as to what conditions guarantee the existence of a stationary solution at all. Note that the CL process defined by the real two-dimensional equations (2.10) and (2.11) has a singular diffusion matrix. Therefore, a general statement on the existence of stationary solutions cannot be made based on the nature of the drift terms.

Further, T does not fall into the class of hypoelliptic operators of constant strength for general $S(x + iy)$, $(x, y) \in \mathbb{R}^2$. In this case a general statement on the regularity of the solutions cannot be made⁽⁸⁾; thus one cannot exclude the possibility that the solutions $T\hat{f} = 0$ exist only in the

sense of distributions (weak solutions). If this is the case, then it might be quite difficult to find (construct) the stationary density.

It is easy to see that a situation like this can occur with the very simple example $S(x) = cx^2$, $c \in \mathbb{R}^+$, which is now supposed to be solved by complex Langevin. The Langevin equation reads

$$dX(\tau) = -cX(\tau) d\tau + dW(\tau) \quad (2.25)$$

$$dY(\tau) = -cY(\tau) d\tau \quad (2.26)$$

The stationary density $\hat{f}(x, y)$ is then a weak solution and can be formally given as

$$\hat{f}(x, y) \sim e^{-cx^2} \delta(y) \quad (2.27)$$

From the above one also must conclude that in general $\hat{f}(x, y)$ cannot be assumed to be a Gibbsian density, i.e., \hat{f} cannot be assumed to be of the form $\hat{f} = \exp(\Phi)$ with Φ some real potential. This situation is acceptable, since the physics of the system is described by the complex distribution $\exp(-S)$. The CL method and thus the resulting stationary distribution merely serve as a convenient algorithm to simulate such complex densities.

3. POLYNOMIAL ACTIONS

Returning to the question of Eq. (2.14), we now demonstrate the following general result: let $S(x)$ and $g(x)$ be polynomials of degree N and M ($M \leq N-1$) and $S(x)$ such that $e^{-S} \in \mathcal{S}(\mathbb{R})$. $\mathcal{S}(\mathbb{R})$ is the Schwartz space of C^∞ functions of rapid decrease. Assume $Z(\tau)$ converges in distribution to a strongly stationary process. Then Eq. (2.14) holds if there exists a τ_0 such that for all $\tau \geq \tau_0$:

1. We have

$$|E(Z^n(\tau) e^{ikZ(\tau)})| < \infty \quad \text{for all } 0 \leq n \leq N-1, \quad k \in \mathbb{R} \quad (3.1)$$

2. The Fourier transform

$$h(x, \tau) = \frac{1}{2\pi} \int_{\mathbb{R}} c_\tau(k) e^{-ikx} dk \quad (3.2)$$

of the expectation value

$$E(e^{ikZ(\tau)}) \equiv c_\tau(k) \quad (3.3)$$

satisfies $h(x, \tau) \in C^2(\mathbb{R})$ with respect to x , and $h(x, t) \in C^1(\mathbb{R})$ with respect to τ . Furthermore, $x^{N-1}h(x, \tau) \in L^1(\mathbb{R}, dx)$.

3. We have

$$\lim_{\tau \rightarrow \infty} c_\tau(k) \in \mathcal{S}(\mathbb{R}) \tag{3.4}$$

Before proving this result, we note that condition 3 is also a necessary condition for Eq. (2.14), since, by assumption, $e^{-S} \in \mathcal{S}$. Also, requiring $c_\tau(k) \in \mathcal{S}(\mathbb{R})$ would be a simplification of condition 2. However, this requirement for all $c_\tau(k)$ might be too restrictive.

Under the above assumptions we have

$$\langle e^{ikx(\tau)} \rangle = E(e^{ikZ(\tau)}) \tag{3.5}$$

where $\langle g(x(\tau)) \rangle$ is given by

$$\langle g(x(\tau)) \rangle = \int_{\mathbb{R}} g(x) h(x, \tau) dx \tag{3.6}$$

From assumption 2 it follows that $c_\tau(k) \in C^{N-1}(\mathbb{R})$, and thus we can conclude that for $f(x) = e^{ikx}$

$$E\left(\frac{df(Z(\tau))}{dZ(\tau)} \frac{dS(Z(\tau))}{dZ(\tau)}\right) = \left\langle \frac{df(x(\tau))}{dx(\tau)} \frac{dS(x(\tau))}{dx(\tau)} \right\rangle \tag{3.7}$$

$$E\left(\frac{d^2f(Z(\tau))}{dZ^2(\tau)}\right) = \left\langle \frac{d^2f(x(\tau))}{dx^2(\tau)} \right\rangle \tag{3.8}$$

and that the surface terms in the above integral expressions vanish. Thus $h(x, \tau)$ obeys the pseudo-FP equation with a complex drift term

$$\frac{\partial h(x, \tau)}{\partial \tau} = \frac{1}{2} \frac{\partial}{\partial x} \left[\frac{\partial S(x)}{\partial x} + \frac{\partial}{\partial x} \right] h(x, \tau) = \tilde{T}h(x, \tau) \tag{3.9}$$

The operator \tilde{Q} , which is \tilde{T} transformed to $L^1(\mathbb{R}, d\tilde{\mu})$ by

$$\tilde{Q} = e^S \tilde{T} e^{-S} \tag{3.10}$$

where $d\tilde{\mu}$ is given by

$$d\tilde{\mu}(x) = \left| \frac{1}{\mathcal{N}} e^{-S(x)} \right| dx \tag{3.11}$$

has a zero eigenvalue with eigenfunction

$$\hat{h}(x) = 1 \tag{3.12}$$

However, we also note that \tilde{Q} has a second solution of zero eigenvalue in $L^1(\mathbb{R}, d\tilde{\mu})$ given by

$$\hat{h}_s(x) = \int_{x_0}^x e^{S(y)} dy \tag{3.13}$$

Although $e^{-S(x)}\hat{h}_s(x)$ is in general not positive definite, and hence automatically excluded when S is real, it is not *a priori* clear that $h(x, \tau) \rightarrow e^{-S(x)}\hat{h}_s(x)$ a.e. is excluded for complex S .

But note that

$$\hat{h}_s(x) = \mathcal{O}\left(\frac{e^{S(x)}}{\partial_x S(x)}\right) \quad \text{for } |x| \rightarrow \infty \tag{3.14}$$

Therefore one has

$$e^{-S(x)}\hat{h}_s(x) = \mathcal{O}\left(\frac{1}{x^{N-1}}\right) \quad \text{for } |x| \rightarrow \infty \tag{3.15}$$

where N is the degree of the polynomial $S(x)$. Thus $e^{-S(x)}\hat{h}_s(x) \notin \mathcal{S}(\mathbb{R})$, which contradicts (3.4). So $e^{-S(x)}\hat{h}_s(x)$ cannot be the limit of $h(x, \tau)$. As was true with the real- S case, the finite measure $\tilde{\mu}$ implies that all solutions of (3.9) satisfying (3.2) converge pointwise to the desired result.

The above three conditions appear to be a necessary set which must be satisfied in order to connect the Langevin process Z to the complex Fokker–Planck equation (3.9). These conditions also guarantee the correct convergence of the first $N - 1$ moments. More generally, however, any higher moment $E(Z^n(\tau))$ which does exist will converge correctly if $h(x, \tau)$ is such that $|\langle x^n(\tau) \rangle| < \infty$.

Two conclusions from this analysis are especially worth noting. First of all, we have seen that the complex Fokker–Planck equation has been derived under certain assumptions about the Langevin process. If these criteria are not met, in particular if the function $c_r(k)$ is ill behaved, then there is no longer necessarily a relation between the CL equation and the complex Fokker–Planck equation. This is true even if the spectrum of the complex FP equation is such that all solutions converge to the desired stationary solution. Second, our result implies that the success or failure of the CL method only depends on the properties of the two eigenfunctions of the zero eigenvalue of \tilde{Q} and not on the requirement that the real part of the spectrum of \tilde{Q} , $\text{Re } \sigma(\tilde{Q}) \leq 0$. Thus, we see that, although it is certainly more convenient if $\text{Re } \sigma(\tilde{Q}) \leq 0$, an analysis of the spectrum is neither necessary nor sufficient for studying the behavior of the CL process.

4. AN EXACTLY SOLVABLE EXAMPLE

For the Gaussian model $S(x) = \omega w^2$, $\omega = a + ib \in \mathbb{C}^+$, the two-dimensional real Fokker–Planck equation has the correct unique stationary density to which all initial solutions converge.⁽⁹⁾ The transition density is given by

$$f(x, y, \tau | x_0, y_0, 0) \sim \exp[-r^T(\tau) \Sigma^{-1}(\tau) r(\tau)] \tag{4.1}$$

with

$$r_1(\tau) = x - m_1(x_0, y_0, \tau), \quad r_2(\tau) = y - m_2(x_0, y_0, \tau) \tag{4.2}$$

For the detailed form of the matrix $\Sigma(\tau)$ and the vector $m(x_0, y_0, \tau)$, the reader is referred to ref. 9. As mentioned above, the transition density converges to the stationary density

$$\lim_{\tau \rightarrow \infty} f(x, y, \tau | x_0, y_0, 0) \sim \exp \left\{ -2a \left[x^2 + 2 \frac{a}{b} xy + \left(1 + \frac{2a^2}{b^2} \right) y^2 \right] \right\} \tag{4.3}$$

With the transition density we have

$$c_\tau(k) = E(e^{ikZ(\tau)}) = \int_{\mathbb{R}^2} e^{ikz} dx dy \int_{\mathbb{R}^2} f(x, y, \tau | x_0, y_0, 0) f(x_0, y_0, 0) dx_0 dy_0 \tag{4.4}$$

from which it follows that for $f(x_0, y_0, 0) = \delta(x_0 - x') \delta(y_0 - y')$, $c_\tau(k)$ has the form

$$c_\tau(k) \sim e^{-k^2 \alpha(a, b, \tau)} \tag{4.5}$$

By examining the form of the matrix $\Sigma(\tau)$, one can see that there always exists a τ_0 such that for all $\tau \geq \tau_0$, the real-valued function $\alpha(a, b, \tau) > 0$. Clearly, $c_\tau(k) \in \mathcal{S}(\mathbb{R})$, and conditions 1–3 hold.

Since in this case the ground state of T is unique, by the preceding analysis it follows that all initial states converge to the proper probability density.

It is interesting to note that for quadratic polynomials it can be demonstrated that all solutions to the complex Fokker–Planck equation which are in $L^2(\mathbb{R}, d\tilde{\mu})$ converge to the ground state of \tilde{Q} or to zero. Note that all square-integrable solutions are also absolutely integrable (L^1) in this case.

In this example the complex Fokker–Planck equation can be solved exactly. For the above range of the complex coupling ω , the spectrum of \tilde{Q} satisfies $\text{Re}[\sigma(\tilde{Q})] \leq 0$, and the eigenfunctions of \tilde{Q} defined on $L^2(\mathbb{R}, d\tilde{\mu})$ form a basis.

To see this, note that

$$\tilde{Q} = \frac{1}{2}e^{S/2}(-H)e^{-S/2} \tag{4.6}$$

where H is now the ordinary Schrödinger operator for the complex frequency extension of the harmonic oscillator with a zero-energy ground state. With the help of the dilatation operator, which is given by⁽¹⁰⁾

$$U_\theta\psi(x) = e^{\theta/2}\psi(e^\theta x) \tag{4.7}$$

one can now show the completeness of the eigenfunctions of H . First note that the domain $D(U_\theta)$ is dense in $L^2(\mathbb{R}, dx)$. Second, the eigenfunctions of the real harmonic oscillator are in $D(U_\theta)$ for $\text{Re}[\exp(2\theta)] > 0$. The dilatation operator maps the eigenfunctions of the real harmonic oscillator to the eigenfunctions of the complex harmonic oscillator. Now for all functions $\phi(x) \in D(U_\theta)$ and $\psi_n(x)$ the n th eigenfunction of the real harmonic oscillator, we have

$$(\phi, U_\theta\psi_n) = (U_{\bar{\theta}}^{-1}\phi, \psi_n) = 0, \quad \text{for all } n \tag{4.8}$$

if and only if $U_{\bar{\theta}}^{-1}\phi = 0$, which implies $\phi = 0$. But since $U_{\bar{\theta}}^{-1}\phi$ for $\phi \in D(U_\theta)$ is dense, the eigenfunctions of the complex harmonic oscillator form a complete basis. Thus

$$\text{s-lim}_{\tau \rightarrow \infty} h(x, \tau) = \hat{h}(x) \text{ or } 0 \tag{4.9}$$

for all solutions $h(x, \tau) \in L^2(\mathbb{R}, d\tilde{\mu})$ of

$$\frac{\partial h(x, \tau)}{\partial \tau} = \tilde{Q}h(x, \tau) \tag{4.10}$$

Again, from $\tilde{\mu}(\mathbb{R}) < \infty$ one can conclude that $h(x, \tau)$ converges pointwise.

5. CONCLUSIONS

To date there is no comprehensive theory of the complex Langevin method. General results are difficult to prove because many theorems on differential operators and diffusion processes do not apply to T and the Langevin equations. However, for the case of polynomial actions we have demonstrated a set of sufficient conditions to guarantee convergence of the CL method to the correct results.

An extension of these results from polynomial actions to the more general case of an analytic S requires a more detailed understanding of the solutions $f(x, y, \tau)$ of T . Also of interest would be similar results for compact manifolds. Work is currently progressing in these directions.

ACKNOWLEDGMENTS

S.L. would like to thank Wes Petersen of the IPS-ETH Zurich for illuminating discussions. Both authors are grateful to John Klauder of the University of Florida for many helpful suggestions and reading the manuscript. S.L. also wishes to thank the University of Graz and IPS-ETH Zurich for their kind hospitality during the time in which this work was completed. H.G. thanks Bernd Thaller for many useful discussions and his help on the complex frequency harmonic oscillator. This work was supported by Fonds zur Förderung der Wissenschaftlichen Forschung in Österreich project P7849.

REFERENCES

1. N. Metropolis *et al.*, *J. Chem. Phys.* **21**:1087 (1953).
2. J. R. Klauder, in *Recent Developments in High Energy Physics*, H. Mitter and C. B. Lang, eds. (Springer, Vienna, 1983).
3. J. R. Klauder and W. P. Petersen, *J. Stat. Phys.* **39**:53 (1985).
4. H. Q. Lin and J. E. Hirsch, *Phys. Rev. B* **34**:1964 (1986).
5. H. Okamoto, K. Okano, L. Schülke, and S. Tanaka, *Nucl. Phys. B* **324**:684 (1989); K. Okano, L. Schülke, and B. Zheng, *Phys. Lett. B* **258**:421 (1991); K. Okano, L. Schülke, and B. Zheng, Siegen Preprint Si-91-8.
6. L. Arnold, *Stochastic Differential Equations* (Wiley, New York, 1974).
7. K. Yosida, *Functional Analysis* (Springer, Berlin, 1968).
8. L. Hörmander, *The Analysis of Linear Partial Differential Operators I, II* (Springer, Berlin, 1983).
9. R. W. Haymaker and Y. Peng, *Phys. Rev. D* **41**:1269 (1990).
10. H. L. Cycon, R. G. Froese, W. Kirsch, and B. Simon, *Schrödinger Operators* (Springer, Berlin, 1987).