On the correct convergence of complex Langevin simulations for polynomial actions

1994 J. Phys. A: Math. Gen. 27 1325

(http://iopscience.iop.org/0305-4470/27/4/025)

J. Phys. A: Math. Gen. 27 (1994) 1325–1330. Printed in the UK

On the correct convergence of complex Langevin simulations for polynomial actions†

H Gausterer
Institut für Theoretische Physik, Universität Graz, A-8010 Graz, Austria

Received 12 October 1993

Abstract. There are problems in physics and particularly in field theory which are defined by complex-valued weight functions e^{-S} where S is a polynomial action $S : \mathbb{R}^n \to \mathbb{C}$. The conditions under which a convergent complex Langevin calculation correctly simulates such integrals are discussed. All conditions on the process which are used to prove proper convergence are defined in the stationary limit.

1. Introduction

Complex Langevin (CL) methods have turned out to be quite useful in the calculation (simulation) of high dimensional integrals over complex-valued weight functions of the form e^{-S}, where S is the action or the Hamiltonian of some physical system. Since there is no formal restriction to a real-valued drift term for Langevin equations, the application of CL is convincingly simple [1]. Unfortunately one has to deal with two problems of uncertainty. The first is that it is *a priori* unknown whether the process will converge at all. The second problem is that, although the process has converged, it will not necessarily give the correct answer. This is, that long time averages of such a process do not necessarily simulate the complex valued weight function integrals. CL is known to sometimes give the wrong answer (see e.g. [2]).

Several attempts have been made to understand CL (e.g. see references [2–4]). For some simple actions the behaviour of CL can be improved by modifying the drift term with an appropriate kernel, but for general problems the choice of the kernel is not clear [5]. Recently progress has been made in the comprehension of the results which one gets from a convergent process [6, 7]. In particular the assumptions needed to guarantee correct results for convergent processes on certain compact manifolds (S_1, S_2) turn out to be surprisingly simple and easy to verify in a numerical simulation. Contrary to that, many assumptions are used to prove the behaviour of processes on $\mathbb{R}^n$ driven by polynomial actions and moreover these assumptions are rather technical [6].

For polynomial actions a lot of attention has been given to the existence of a pseudo Fokker–Planck (FP) equation which describes the dynamics of a possibly equivalent complex valued weight function [6]. In earlier investigations especially the spectrum of this operator played a major role [1]. But statements on the properties of the spectrum are not sufficient to draw conclusions on the correctness or the convergence of CL [6]. Certainly, if one can show that the pseudo FP equation exists and that the real part of the spectrum of the operator is semidefinite then CL converges but not necessarily to the desired result. Further

† Supported by Fonds zur Förderung der Wissenschaftlichen Forschung in Österreich, project P7849.

0305-4470/94/041325+06$19.50

conditions must hold (see [6]). Except for very simple cases it is hard and most unlikely to get exact information on the complete spectrum. Certainly there always exists the real FP operator for the process and the convergence of the process follows if one can prove that the operator has a unique non-negative integrable solution to the zero eigenvalue. But this is also very hard and so far there is no classification scheme for actions which have the suitable properties. So, to get information on the convergence for any problem one must check either the existence and the whole spectrum of the pseudo FP operator or the zero eigenvalue properties of the real FP operator. In practice therefore the question of convergence still remains a matter of experiment and experience.

2. Proper convergence

Let us now turn to the main purpose of the paper and examine in a rigorous fashion the conditions under which CL, if convergent, gives the right answer. To demonstrate this, several conditions at finite time have been put on the process in [6]. In this approach fewer conditions on the process are used and these conditions are put forward to $t \to \infty$. For simplicity the discussion and the formulae are restricted to the one dimensional case. All following statements allow for an immediate generalization to $\mathbb{R}^n$. It will be assumed that the system of interest is described by a complex polynomial action of degree N

$$S(x) = \sum_{n=0}^{N} a_n x^n . \tag{2.1}$$

$S : \mathbb{R} \to \mathbb{C}$ such that $e^{-S} \in \mathcal{S}(\mathbb{R})$. $\mathcal{S}(\mathbb{R})$ is the Schwartz space of C^∞ functions of rapid decrease. With $g(x)$ a polynomial of degree M it is thus guaranteed, that the quantities of physical interest

$$\langle g(x) \rangle \equiv \frac{1}{\mathcal{N}} \int_{\mathbb{R}} g(x)\, e^{-S(x)}\, dx \tag{2.2}$$

$$\mathcal{N} = \int_{\mathbb{R}} e^{-S(x)}\, dx \tag{2.3}$$

do exist, provided $0 < |\mathcal{N}|$. If S would be real valued everything would now be straightforward ergodic theory and the long-time averages computed with the Langevin equation would reproduce the ensemble average of the system.

In the complex case analytic continuation leads to the following stochastic differential equation:

$$dZ(t) = F(Z(t))\, dt + dW(t) \tag{2.4}$$

with the drift term

$$F(z) = -\frac{1}{2}\frac{dS(z)}{dz} . \tag{2.5}$$

$W(t)$ is a standard Wiener process with zero mean and covariance

$$E\,(W(t_1)W(t_2)) = \min(t_1, t_2) . \tag{2.6}$$

Equation (2.4) is the so called CL equation. This equation has a locally unique solution which is defined up to a random explosion time [9]. In particular (2.4) describes a two-dimensional diffusion process:

$$\mathrm{d}X(\tau) = G(X(\tau), Y(\tau))\,\mathrm{d}\tau + \mathrm{d}W(\tau) \tag{2.7}$$

$$\mathrm{d}Y(\tau) = H(X(\tau), Y(\tau))\,\mathrm{d}\tau\,. \tag{2.8}$$

With $S(z) = u(x, y) + \mathrm{i}v(x, y)$, we have

$$G(x, y) = -\frac{1}{2}\frac{\partial u(x, y)}{\partial x} \qquad H(x, y) = \frac{1}{2}\frac{\partial u(x, y)}{\partial y}\,. \tag{2.9}$$

Special to this process is that (2.8) looks like a deterministic equation due to the zero diffusion coefficient. Nevertheless this is a stochastic equation through the dependence on $X(t)$. The singular diffusion matrix causes a lot of problems. So, contrary to the real action case it is in general not possible to determine from the drift and diffusion terms whether there exists a unique stationary distribution density for this process [8]. As already mentioned in the introduction there is no general proof on the existence of a stationary distribution density. For the moment let us assume that for the process $X(t), Y(t)$ there exists a unique stationary distribution density $\hat{f}(x, y)$. The idea behind CL then is that

$$\lim_{t\to\infty} E\,(g(X(t) + \mathrm{i}Y(t))) = \int_{\mathbb{R}^2} g(x + iy)\hat{f}(x, y)\,\mathrm{d}x\,\mathrm{d}y = \frac{1}{N}\int_{\mathbb{R}} g(x)\mathrm{e}^{-S(x)}\,\mathrm{d}x \tag{2.10}$$

might hold.

Assume:

(1) S is a complex valued polynomial action of degree N such that

$$\mathrm{e}^{-S} \in \mathcal{S}(\mathbb{R}) \tag{2.11}$$

and

$$\left|\int_{\mathbb{R}} \mathrm{e}^{-S(x)}\,\mathrm{d}x\right| > 0\,. \tag{2.12}$$

(2) For

$$c(k, t) \equiv E(\mathrm{e}^{\mathrm{i}kZ(t)}) = \int_{\mathbb{R}^2} \mathrm{e}^{\mathrm{i}k(x+\mathrm{i}y)} f(x, y, t)\,\mathrm{d}x\,\mathrm{d}y \tag{2.13}$$

the limit $t \to \infty$ exists pointwise and

$$\lim_{\tau\to\infty} c_\tau(k) \equiv c_\infty(k) \in \mathcal{S}(\mathbb{R})\,. \tag{2.14}$$

(3) Further

$$\lim_{t\to\infty} \left|E(Z^n(t)\,\mathrm{e}^{\mathrm{i}kZ(t)})\right| < \infty \qquad \text{for all} \quad 0 \leqslant n \leqslant N-1 \quad k \in \mathbb{R}\,. \tag{2.15}$$

Equation (2.10) then holds at least for $g(z)$ a polynomial of degree $M \leqslant N - 1$. Moreover (2.10) holds for any higher moment $E(Z^n(t))$, $n \geqslant N$ which exist for $t \to \infty$.

From assumption 2 we know that there is a $t_0 < \infty$ such that $c(k,t)$ exists and from assumption 3 that there is a $t_1 < \infty$ such that $E(Z^n(t)\,\mathrm{e}^{\mathrm{i}kZ(t)})$ exists. Applying the Itô rule one gets with $F(z)$ as defined in (2.5)

$$\frac{\partial E(\mathrm{e}^{\mathrm{i}kZ(t)})}{\partial t} = \mathrm{i}kE\left(\mathrm{e}^{\mathrm{i}kZ(t)}F(Z(t))\right) - \frac{k^2}{2}E\left(\mathrm{e}^{\mathrm{i}kZ(t)}\right). \tag{2.16}$$

Due to assumptions 2 and 3, equation (2.16) exists for $t' = \max(t_0, t_1)$. As a side result we get that, if $c(k,t) \in C^{N-1}(\mathbb{R})$ with respect to k, (2.16) can be understood as the dynamical equation for $c(k,t)$:

$$\frac{\partial c(k,t)}{\partial t} = -\frac{\mathrm{i}k}{2}\sum_{n=1}^{N} n a_n \left(-\mathrm{i}\frac{\partial}{\partial k}\right)^{n-1} c(k,t) - \frac{k^2}{2}c(k,t). \tag{2.17}$$

Note that if assumption 3 does not hold, (2.16) can also not be defined in the sense of distributions. This is because we are not simply dealing with Fourier transforms but with their possibly not existing analytic continuations.

Let us now define $\hat{h}(x)$ as

$$\hat{h}(x) = \frac{1}{2\pi}\int_{\mathbb{R}} c_\infty(k)\,\mathrm{e}^{-\mathrm{i}kx}\,\mathrm{d}k. \tag{2.18}$$

From assumption 2 follows that $\hat{h}(x) \in \mathcal{S}(\mathbb{R})$. Using (2.18) and assumption 3

$$\lim_{t\to\infty} E(Z^n(t)\,\mathrm{e}^{\mathrm{i}kZ(t)}) = \int_{\mathbb{R}} x^n\,\mathrm{e}^{\mathrm{i}kx}\hat{h}(x)\,\mathrm{d}x \tag{2.19}$$

for $0 \leqslant n \leqslant N-1$ and $k \in \mathbb{R}$. Applying the above result to (2.16) one obtains in the limit $t \to \infty$

$$0 = \mathrm{i}k\int_{\mathbb{R}} \mathrm{e}^{\mathrm{i}kx}F(x)\hat{h}(x)\,\mathrm{d}x - \frac{k^2}{2}\int_{\mathbb{R}} \mathrm{e}^{\mathrm{i}kx}\hat{h}(x)\,\mathrm{d}x. \tag{2.20}$$

Integrating the right-hand side of (2.20) by parts gives that $\hat{h}(x)$ is a $L^1(\mathbb{R}, \mathrm{d}x)$ zero eigenvalue solution of a FP type differential operator with a complex drift term (pseudo FP operator).

$$\frac{1}{2}\frac{\partial}{\partial x}\left[\frac{\partial S(x)}{\partial x} + \frac{\partial}{\partial x}\right]\hat{h}(x) \equiv \mathcal{T}\hat{h}(x) = 0. \tag{2.21}$$

$\mathcal{T}$ has two zero eigenvalue solutions. One is

$$\hat{h}_1(x) \sim \mathrm{e}^{-S(x)} \in \mathcal{S}(\mathbb{R}) \tag{2.22}$$

which fits to assumption 2, since as the Fourier transform of $c_\infty(k)$ it must be a Schwartz function. For the second solution

$$\hat{h}_2(x) \sim \mathrm{e}^{-S(x)}\int_{x_0}^{x} e^{S(y)}\,\mathrm{d}y \tag{2.23}$$

one can show that

$$\hat{h}_2(x) = \mathcal{O}\left(\frac{1}{x^{N-1}}\right) \text{ for } |x| \to \infty . \tag{2.24}$$

This contradicts assumption 2. So, the only possible solution is the one proportional to e^{-S} and

$$\lim_{t\to\infty} E(Z^n(t)\, e^{ikZ(t)}) = \frac{1}{\mathcal{N}} \int_{\mathbb{R}} x^n\, e^{ikx}\, e^{-S(x)}\, dx \tag{2.25}$$

for $0 \leqslant n \leqslant N-1$ and $k \in \mathbb{R}$. If further $E(Z^n(t))$, $n \geqslant N$ for $t \to \infty$ exist then

$$\lim_{t\to\infty} E(Z^n(t)) = \left.\frac{d^n c_\infty(k)}{dk^n}\right|_{k=0} = \frac{1}{\mathcal{N}} \int_{\mathbb{R}} x^n\, e^{-S(x)}\, dx . \tag{2.26}$$

Let us now briefly discuss the assumptions. Polynomial actions are very natural since most physical systems defined on $\mathbb{R}^n$ have polynomial actions. Since these actions must be bounded from below it follows that $e^{-S} \in \mathcal{S}$. Condition 2 must be there otherwise the solution $\hat{h}_2(x)$ cannot be excluded. With the correctness requirement on CL that

$$\lim_{t\to\infty} E(e^{ikZ(t)}) = \frac{1}{\mathcal{N}} \int_{\mathbb{R}} e^{ikx}\, e^{-S(x)}\, dx \tag{2.27}$$

this condition is also a necessary condition. Assumption 3 looks technical, but is so far required to relate $\hat{h}_i(x)$, the Fourier transform of $c_\infty(k)$, to the Fokker–Planck type operator $\mathcal{T}$. This condition finally allows us to show the correctness of CL. It would be nice to eliminate assumption 3 by showing that it follows from assumption 2. Unfortunately the integral transform defined by (2.13) is not an injective mapping. To the authors knowledge the nature of this integral transform has not been analysed in the literature. At present, without more detailed information on the probability density (in general not available), it is perhaps impossible to draw a conclusion on the properties of the function from the properties of its image. In a numerical simulation certainly such mathematical criteria is hard to verify exactly. Nevertheless experience tells us that when plotting such expectation values ($c_\infty(k)$, $E(Z^n\, e^{ikZ})$) one gets a very clear sign of the quality of the result [10].

3. Conclusions

The criteria under which a convergent CL simulation leads to correct results have been significantly simplified. The assumptions used in the present proof are much closer to a numerical verification than the one used in reference [6]. Unfortunately a complete theory of CL is still lacking. However, the situation that it was generally neither *a priori* nor *a posteriori* possible to prove convergence to the desired result has been ameliorated in as far as a simple *a posteriori* proof is now possible.

References

[1] Klauder J R 1983 *Recent Developments in High Energy Physics* ed H Mitter and C B Lang (New York: Springer)
[2] Klauder J R and Petersen W P 1985 *J. Stat. Phys.* **39** 53
[3] Ambjorn J and Yang S K 1985 *Phys. Lett.* **165B** 140
[4] Haymaker R W and Peng Y 1990 *Phys. Rev.* D **41** 1269
[5] Okamoto H, Okano K, Schülke L and Tanaka S 1989 *Nucl. Phys.* B **324** 684
Okano K, Schülke L and Zheng B 1991 *Phys. Lett.* **258B** 421
[6] Gausterer H and Lee S 1993 *J. Stat. Phys.* **73** *UNIGRAZ-UTP 29-09-92* to be published
[7] Lee S *Nucl. Phys.* B to be published
[8] Risken H 1984 *The Fokker–Planck Equation* (Berlin: Springer)
Gardiner C W 1983 *Handbook of Stochastic Methods* (Berlin: Springer)
[9] Arnold L 1974 *Stochastic Differential Equations* (New York: Wiley)
[10] Lee S 1993 *Thesis* University of Florida Gainesville