On complex Langevin dynamics and zeroes of the measure I: Formal proof and simple models

Gert Aarts[*]
Department of Physics, College of Science, Swansea University, Swansea, United Kingdom
E-mail: `g.aarts@swan.ac.uk`

Erhard Seiler
Max-Planck-Institut für Physik (Werner-Heisenberg-Institut), München, Germany
E-mail: `ehs@mppmu.mpg.de`

Dénes Sexty
Department of Physics, Bergische Universität Wuppertal, Wuppertal, Germany
Inst. for Theoretical Physics, Eötvös University, Budapest, Hungary
E-mail: `sexty@uni-wuppertal.de`

Ion-Olimpiu Stamatescu
Institut für Theoretische Physik, Universität Heidelberg, Heidelberg, Germany
E-mail: `I.O.Stamatescu@thphys.uni-heidelberg.de`

In the complex Langevin approach to lattice simulations at nonzero density, zeroes of the fermion determinant lead to a meromorphic drift and hence a need to revisit the theoretical derivation. We discuss how poles in the drift affect the formal justification of the approach and then explore the various potential issues in simple models, in a manner that is applicable to heavy dense and full QCD.

---

[*]Speaker.

1. Introduction

Complex Langevin dynamics has solved the sign problem in a number of theories with a complex weight due to a nonzero chemical potential [1, 2], including heavy dense QCD [3–5], and progress for QCD with lighter quarks is underway [6–8]. However, the presence of the fermion determinant causes a theoretical problem, since the Langevin drift is then no longer holomorphic: zeroes of the determinant cause poles in the drift. In practice, this may lead to incorrect convergence, but not necessarily so [9–11]. Since the formal justification [12,13] relies (among others) on holomorphicity a reconsideration of the derivation is required. This is sketched below. The interplay between the poles and the real and positive distribution sampled during the Langevin process turns out to be essential and this is studied in a sequence of models, with the aim to extract generic lessons. This contribution is based on Ref. [14] and accompanied by Ref. [15].

2. Formal derivation revisited

We will start with revisiting the formal derivation and justification of the complex Langevin approach [12, 13] and point out where the arguments have to be amended to include meromorphic drifts. The approach hinges on the equivalence of two expectation values, one defined with respect to the original complex weight $\rho(x)$ and one with respect to the real and positive weight $P(x,y)$ on the analytically extended manifold, i.e.,

$$\langle O \rangle_{\rho(t)} = \int dx\, \rho(x,t) O(x), \qquad \langle O \rangle_{P(t)} = \int dx dy\, P(x,y;t) O(x+iy). \qquad (2.1)$$

These distributions satisfy the Fokker-Planck equations (we consider real noise only [12])

$$\dot{\rho}(x,t) = \nabla_x [\nabla_x - K(x)]\rho(x,t), \qquad \dot{P}(x,y;t) = [\nabla_x (\nabla_x - K_x) - \nabla_y K_y] P(x,y;t), \qquad (2.2)$$

with the Langevin drift terms

$$K(z) = \nabla_z \rho(z)/\rho(z), \qquad K_x = \operatorname{Re} K(z), \qquad K_y = \operatorname{Im} K(z). \qquad (2.3)$$

Success is obtained when these expectation values are equal: $\langle O \rangle_{\rho(t)} = \langle O \rangle_{P(t)}$.

The equivalence can indeed by demonstrated [12,13], provided that 1) the drift and observables are holomorphic; 2) the distribution $P(x,y)$ has fast decay at $y \to \pm\infty$. In particular, the proof requires partial integration at $|y| \to \infty$ without boundary terms.

Let us now consider the case with (at least) one zero in measure, $\rho(x = z_p) = 0$. In that case the drift $K(z) = \nabla_z \rho(z)/\rho(z)$ has pole at $z = z_p$ and is no longer holomorphic, but only meromorphic. Hence it is necessary to revisit the derivation. Note that QCD is an example that falls in this category, since after integrating the quarks, the partition function is

$$Z = \int DU \det M(U) e^{-S_{\text{YM}}}, \qquad \text{with } \det M(U) = 0 \text{ for some } U \in \text{SL}(N,\mathbb{C}). \qquad (2.4)$$

It turns out that the derivation as above goes through, provided that the region around the pole is excluded, i.e. $|z - z_p| > \varepsilon$ [14]. However, this yields the possibility of new potential boundary terms at $z \sim z_p$, besides the ones at $|y| \to \infty$. It is therefore necessary to study the behaviour of the product of the distribution and observables, $P(x,y)O(x+iy)$, around $z \sim z_p$ carefully.

Let us make the following remark on the time evolution of holomorphic observables [14],

$$\dot{O}(z;t) = \tilde{L}O(z;t), \qquad \tilde{L} = [\nabla_z + K(z)]\nabla_z, \qquad (2.5)$$

with the solution

$$O(z;t) = e^{\tilde{L}t}O(z;0) = \sum_k \frac{t^k}{k!}\tilde{L}^k O(z;0). \qquad (2.6)$$

Since $\tilde{L}$ has a pole at $z = z_p$, one may expect $O(z;t)$ to have an essential singularity at $z = z_p$. However, this potential disaster is counteracted by the vanishing of $P(x,y)$ as $z \to z_p$, as well as by the nontrivial angular dependence of $P(x,y)$ around $z = z_p$ (see below), which soften the singularity.
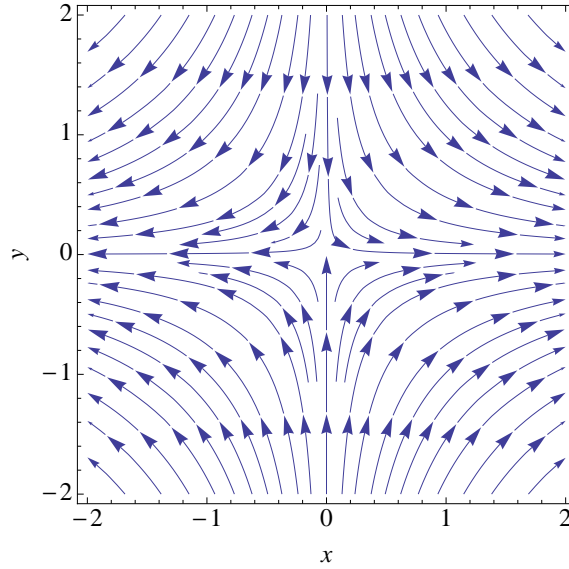
3. Poles and the distribution

The flow pattern around a pole has a generic structure. Consider a zero at $z = z_p$ of order n_p,

$$\rho(x) = (x - z_p)^{n_p} e^{-S(x)}. \qquad (3.1)$$

The drift is then given by

$$K(z) = \frac{\rho'(z)}{\rho(z)} = \frac{n_p}{z - z_p} - S'(z). \qquad (3.2)$$

In Fig. 1, we show the corresponding classical flow pattern, for $z_p = 0$ and $S(z) = 0$. The attractive and repulsive directions are generic and lead to particular angular dependence. We note that due to this, multiple circlings of the pole are not expected.

Figure 1: Generic flow pattern around a pole at $z = z_p = 0$, with attractive and repulsive directions.

Given the formal justification, it is necessary to better understand the behaviour of the distribution (and observables) around the pole. Logically, there are three possibilities: 1) the pole is outside the distribution; 2) the pole is on the edge of the distribution; 3) the pole is inside the distribution. We will now encounter these cases in various models.

4. Pole and distribution in simple model

We start with a simple but often studied example, with the distribution

$$\rho(x) = (x - z_p)^{n_p} e^{-\beta x^2}, \qquad\qquad \beta \in \mathbb{R}, \qquad z_p = x_p + i y_p \in \mathbb{C}. \qquad (4.1)$$

Following the analysis of Ref. [16], it is easy to arrive at some essential and rigorous properties of distribution $P(x,y)$. For real β, the distribution is nonzero in a horizontal strip only. Hence the decay at $|y| \to \infty$ poses no problem and this possibility of breakdown is avoided. Depending on the parameters (β, y_p, n_p), the pole is either located exactly on the edge of the strip or outside the strip. This is illustrated in Fig. 2.
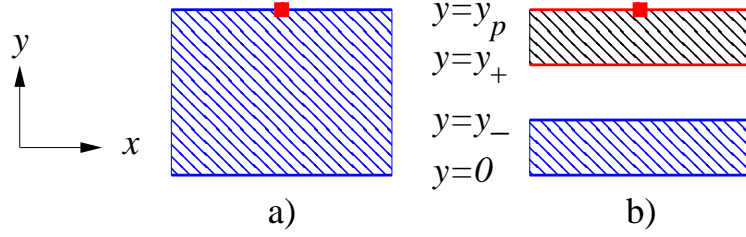
Figure 2: Strips in the xy plane where $P(x,y) \neq 0$. The pole is indicated with the red square. The red striped region in b) is transient only.

These two cases are distinguished by [14]

a) $y_p^2 < 2n_p/\beta$: pole on edge $\Rightarrow P(x,y) \neq 0$ when $0 < y < y_p$;

b) $y_p^2 > 2n_p/\beta$: pole outside strip $\Rightarrow P(x,y) \neq 0$ when $0 < y < y_- < y_p$.

For case b), with the pole outside, the standard justification still holds, since the pole is avoided as the upper strip is a transient. Indeed, complex Langevin dynamics reproduces the exact results in this case. Case a), with the pole on the edge, is more interesting. Here the results depend on the properties of the distribution, determined by the parameter values.

We demonstrate this with an example, using $z_p = i, n_p = 2, \beta = 1.6, 3.2, 4.8$, and show a comparison of the exact results and the results obtained with CL, for the observable $\langle z^n \rangle$, with $n = 1, 2, 3, 4$, in Fig. 3. It can be seen that there is no agreement for $\beta = 1.6$, but good agreement for $\beta = 3.2$ and 4.8. Recall that for all three parameter values, $P(x,y)$ is expected to be nonzero for $0 < y < y_p = 1$, i.e. all the way up to the pole. Given the formal justification, this different behaviour should be visible in properties of the distribution $P(x,y)$.

This is demonstrated in Fig. 4, where the partially integrated distribution $P_y(y) = \int dx P(x,y)$ is shown on a linear (left) and logarithmic (right) scale, for $\beta = 1.6$ (CL incorrect) and 3.2 (CL consistent). We observe that for $\beta = 1.6$, the distribution is nonzero right up to the pole and seems to go to zero linearly. In such a case, boundary terms at $z = z_p$ due to partial integration will contribute and complex Langevin dynamics is not valid, as discussed in Sec. 2. On the other hand, for $\beta = 3.2$, the decay is much faster, possibly exponentially, and hence partial integration poses no problem for observables z^n. Consistent with the formal justification, complex Langevin dynamics then reproduces the correct results.

We conclude that it is possibly to reconcile the properties of the distribution, the formal justification and the success/failure of the complex Langevin process, in the presence of a pole.
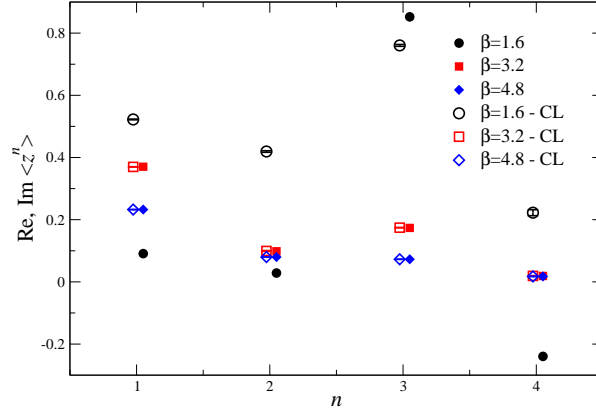
Figure 3: Observables $\langle z^n \rangle$ versus n, obtained with complex Langevin (CL, open symbols) and exact results (smaller filled symbols), for three β values.
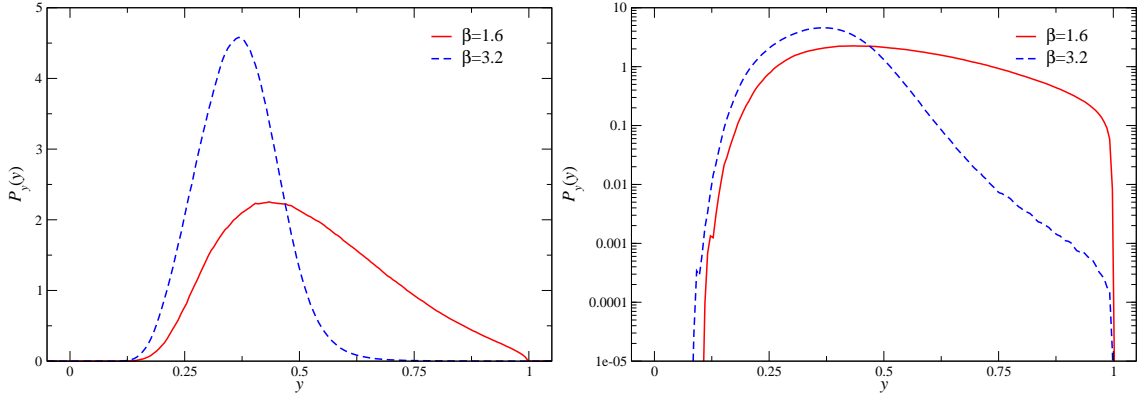
Figure 4: Partially integrated distributions $P_y(y)$ on a linear (left) and logarithmic (right) scale.

5. Towards more realistic models

The next step is to carry over the essence of this analysis to more realistic models, and devise diagnostics which are also applicable in QCD. To do this, we first consider the U(1) one-link model with the complex distribution [3]

$$\rho(x) = [1 + \kappa \cos(x - i\mu)]^{n_p} \exp(\beta \cos x). \tag{5.1}$$

The conclusions can be summarised (loosely speaking) by stating that CL works when $\kappa < 1$ and fails when $\kappa > 1$ [3, 9]. Hence we take here $\kappa = 2 > 1$ here. We also fix $\beta = 0.3, \mu = 1$, and vary the order of the zero, $n_p = 1, 2, 4$.

The distribution has zeroes at $z_p = \pm x_p + i\mu$ and is again nonzero in a strip only, $y_- < y < y_+$, so that the behaviour at $|y| \to \infty$ is under control. We study the observables $\langle e^{ikz} \rangle$ with $k = \pm 1, \pm 2$. The applicability of CL is found to depend strongly on n_p, and we find incorrect results for $n_p = 1, 2$ but correct results for $n_p = 4$ [14]. In this case, the poles lie within the strip. However, the poles pinch the distribution, i.e. approximately disconnected regions appear and the poles act as a bottleneck. In order to present this in a way that is easily extendable to more complicated theories, where the complexified configuration space is not accessible, we show the determinant
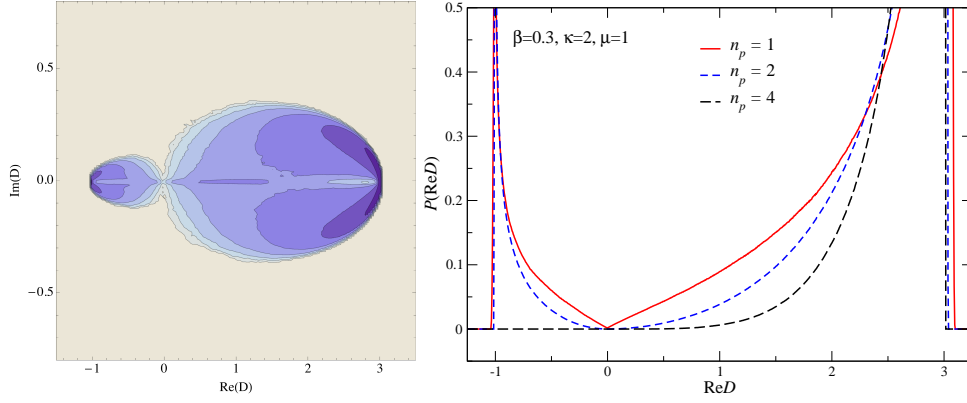
Figure 5: Left: histogram of the determinant factor D in the U(1) model for $n_p = 2$ (on a logarithmic scale). Right: partially integrated distribution for Re D, for three values of n_p.

factor $D = 1 + \kappa \cos(x - i\mu)$ in the complex plane instead. Here D is defined such that the 'full determinant' appears as D^{n_p} in Eq. (5.1).

The result for $n_p = 2$ is shown in Fig. 5 (left). We observe that the pole pinches the distribution and, even though the dynamics takes place in the complex plane, there are no multiple circlings around the pole at the origin. The distribution is zero when $D = 0$. The way this occurs is shown in Fig. 5 (right) for Re D, i.e. we have integrated over Im D. As in the model discussed above, we note that the manner in which the distribution goes to zero is essential: for $n_p = 1, 2$, it is too slow for partial integration to work without boundary terms, while for $n_p = 4$, the distribution is in fact zero when Re $D < 0$, partial integration can be applied, and the formal justification is valid.

We note that it is now easy to divide the configuration space into two disconnected regions, with Re $D \lessgtr 0$. These regions can be treated separately with constrained partition functions $Z_{\pm}$ and relative weights $w_{\pm} = Z_{\pm}/(Z_+ + Z_-)$. We find that Z_- typically contributes incorrect results. However, we also find that $w_- \ll w_+$, which is beneficial for the CL approach.
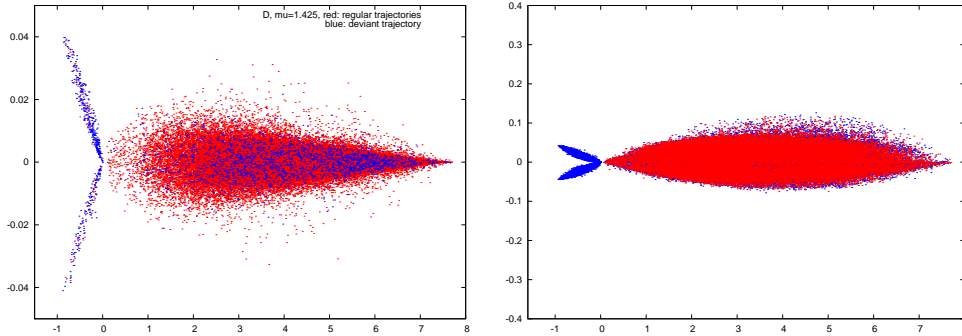
Figure 6: Scatter plot of complex determinant in the SU(3) one-link model, for two choices of parameters representing different aspects of HDQCD.

Finally we discuss the SU(3) effective one-link model, designed to understand heavy dense QCD [14]. Here it is again straightforward to analyse the determinant and we found the same structure as above, as demonstrated in Fig. 6. We find that the zero pinches the distribution, which results in two disjoint areas. It is hence possible to analyse each region separately. When increasing

5

the order of the zero (increasing n_p), we find that there is a stronger drift towards and then away from pole, hence stronger pinching, and typically better agreement with expected results. The analysis of determinant is easily extended to heavy dense QCD, while in full QCD it is numerically more intensive. This is further discussed in Refs. [14, 15].

6. Summary

When using complex Langevin dynamics to solve QCD at nonzero chemical potential, the Langevin drift has poles where the fermion determinant vanishes and is no longer holomorphic. It is then necessary to revisit the formal justification of the approach. We found that the usual derivation still holds, but correctness of the results depends crucially on the behaviour of the distribution around the pole. Similar arguments are also provided in Ref. [17]. Subsequently we analysed a number of models and found common features in all of these, namely that the poles will pinch the distribution and result in disjoint regions, which can be analysed separately. When the zero is of order n_p, e.g. when the determinant can be written as $[\det D]^{n_p}$, we found that larger n_p typically yields better results. This conclusion seems not specific to simple models, but also correct in e.g. heavy dense QCD [14, 15].

References

[1] G. Aarts, PoS LATTICE **2012** (2012) 017 [arXiv:1302.3028 [hep-lat]].

[2] D. Sexty, PoS LATTICE **2014** (2014) 016 [arXiv:1410.8813 [hep-lat]].

[3] G. Aarts and I. O. Stamatescu, JHEP **0809** (2008) 018 [arXiv:0807.1597 [hep-lat]].

[4] E. Seiler, D. Sexty and I. O. Stamatescu, Phys. Lett. B **723** (2013) 213 [arXiv:1211.3709 [hep-lat]].

[5] G. Aarts, F. Attanasio, B. Jäger and D. Sexty, JHEP **1609** (2016) 087 [arXiv:1606.05561 [hep-lat]].

[6] D. Sexty, Phys. Lett. B **729** (2014) 108 [arXiv:1307.7748 [hep-lat]].

[7] G. Aarts, E. Seiler, D. Sexty and I. O. Stamatescu, Phys. Rev. D **90** (2014) no.11, 114505 [arXiv:1408.3770 [hep-lat]].

[8] D. K. Sinclair and J. B. Kogut, PoS LATTICE **2015** (2016) 153 [arXiv:1510.06367 [hep-lat]].

[9] A. Mollgaard and K. Splittorff, Phys. Rev. D **88** (2013) no.11, 116007 [arXiv:1309.4335 [hep-lat]].

[10] K. Splittorff, Phys. Rev. D **91** (2015) no.3, 034507 [arXiv:1412.0502 [hep-lat]].

[11] K. Nagata, J. Nishimura and S. Shimasaki, JHEP **1607** (2016) 073 [arXiv:1604.07717 [hep-lat]].

[12] G. Aarts, E. Seiler and I. O. Stamatescu, Phys. Rev. D **81** (2010) 054508 [arXiv:0912.3360 [hep-lat]].

[13] G. Aarts, F. A. James, E. Seiler and I. O. Stamatescu, Eur. Phys. J. C **71** (2011) 1756 [arXiv:1101.3270 [hep-lat]].

[14] G. Aarts, E. Seiler, D. Sexty and I. O. Stamatescu, in preparation.

[15] G. Aarts, E. Seiler, D. Sexty and I. O. Stamatescu, PoS(LATTICE2016) 092.

[16] G. Aarts, P. Giudice and E. Seiler, Annals Phys. **337** (2013) 238 [arXiv:1306.3075 [hep-lat]].

[17] K. Nagata et al, PoS(LATTICE2016) 067.

PoS(LATTICE2016)036