

# Chemical Potential and Sign Problem: Discussion on the Journal Club

Author:  
Lorenzo Dini

Organized by  
Prof. Gergely Endrödi

---

11/12/2020

# Contents

<b>1</b>	<b>Chemical Potential on the Lattice</b>	<b>2</b>
1.1	Introduction . . . . .	2
1.2	Linear Chemical Potential . . . . .	3
<b>2</b>	<b>Sign Problem</b>	<b>5</b>
<b>3</b>	<b>Trying to Circumvent the Sign Problem</b>	<b>6</b>
3.1	Reweighting . . . . .	7
3.2	Taylor Expansion . . . . .	9
3.3	Imaginary Chemical Potential . . . . .	10
3.4	Complex Langevin . . . . .	11
3.5	Thimbles . . . . .	14
<b>4</b>	<b>Conclusion and Summary</b>	<b>17</b>

# 1 Chemical Potential on the Lattice

## 1.1 Introduction

In this section I will summarize what Dennis introduced to us in our previous session on chemical potential on the lattice, in the last year edition of the Journal Club (see [1] for more details).

Physically, in the continuum, if we are going to consider systems with a nonzero baryon density, the usual canonical function  $Z$  that we use to describe our thermodynamical system in the vacuum is not appropriate anymore. We have to extend our function into a *gran canonical* one, that is:

$$Z(T, \mu) = \text{Tr} \left( e^{-(\hat{H} - \mu \hat{N}_q)/T} \right) \quad (1)$$

Where  $\hat{N}_q$  is the quark number operator.

If one wants to simulate systems described by this gran canonical function, then it will be necessary to introduce the chemical potential  $\mu$  on the lattice.

In principle, the chemical potential is introduced as the Lagrange multiplier associated to the conserved quantum number of the  $U(1)$  symmetry. That means that, in the continuum, I modify the action adding the relative Noether current:

$$J_0(x) \equiv \mu \bar{\psi}(x) \gamma_0 \psi(x) \quad (2)$$

And indeed, if one spatially integrates the Noether current above, he gets the Noether charge  $N_q$ , i.e. the quark number operator.

So, one could be tempted to naively introduce the chemical potential on the lattice adding a local operator to the action, similar to a mass term:

$$\hat{J}_0(n) \equiv a\mu \bar{\hat{\psi}}_n \gamma_0 \hat{\psi}_n \quad (3)$$

Where, the hatted quantities are the discretized version of the continuum observables, evaluated at site  $n$ .

I recall that, within this framework we will obtain a QCD Dirac operator (using a naive discretization of fermions) that looks like:

$$D_{nm} = \frac{1}{2} \sum_{i=1}^4 \left( \gamma_i U_i(n) \delta_{n, m-\hat{i}} - \gamma_i U_i^\dagger(n) \delta_{n, m+\hat{i}} \right) + am \delta_{nm} + a\mu \gamma_4 \delta_{nm} \quad (4)$$

But in our last discussion [1], Dennis showed us that this naive introduction leads to quadratic divergences in the energy density.

In order to avoid this, see e.g. Ref. [2] or the original paper [3], we have to modify the temporal links  $U_4(x) \rightarrow f(a\mu) U_4(x)$ . And if we take the continuum limit of the lattice version of the energy density, and we impose that is equal to the correct, continuum version, one finds that this function  $f(a\mu)$  that takes care of the chemical potential has to satisfy:

$$\begin{cases} f(0) = 1 \\ f(a\mu) = \frac{1}{f(-a\mu)} \end{cases} \quad (5)$$

And the exponential function is the easiest one that does the trick, thus  $f(a\mu) = e^{a\mu}$ .

From a more physical point of view, we see that what we really do in the continuum is to write down a theory that has gauge symmetry, via the gauge field. Then, after the gauge symmetry is correctly imposed, we perform a shift of the gauge field:

$$A_\mu \rightarrow \tilde{A}_\mu \quad | \quad \tilde{A}_0 = A_0 + i\mu. \quad (6)$$

If we do so, then in the action appears the  $U(1)$  Noether Current and we reproduce the correct grand canonical ensemble.

In the lattice, the gauge symmetry is implemented via the first-neighbor subtraction performed with the links, that play the role of the parallel transport. As we know, the links are related to the gauge fields according to  $U_\mu = e^{iaA_\mu}$ . That means that if we apply the transformation above on the temporal link, after that we defined  $U$  and  $U^\dagger$  in order to enforce the gauge symmetry, we find that the new link should be:

$$\begin{cases} \tilde{U}_4 = e^{a\mu}U_4 \\ \tilde{U}_4^\dagger = e^{-a\mu}U_4 \end{cases} \quad (7)$$

That is how the chemical potential is usually introduced in the lattice, in the exponential form. If we apply this substitution to the naive Dirac operator, we get:

$$\begin{aligned} D_{nm} = & \frac{1}{2} \sum_{i=1}^3 \left( \gamma_i U_i(n) \delta_{n,m-\hat{i}} - \gamma_i U_i^\dagger(m) \delta_{n,m+\hat{i}} \right) + am \delta_{nm} + \\ & + \frac{1}{2} \left( \gamma_4 e^{\mu a} U_4(n) \delta_{n,m-\hat{4}} - \gamma_4 e^{-\mu a} U_4^\dagger(m) \delta_{n,m+\hat{4}} \right) \end{aligned} \quad (8)$$

## 1.2 Linear Chemical Potential

The only thing that I would like to add to the previous discussion, as a possible follow-up on the introduction of the chemical potential on the lattice, is that the formulation above is not the only way in which one can introduce the chemical potential on the lattice. There are other ways, that could be more or less useful, depending on the strategy that you would use to try to circumvent the infamous sign problem (see next section).

The one that I would like to hint in this section is the so called linear chemical potential approach, introduced for the first time by Gavai and Sharma in Ref. [4].

The main idea is that, once we understood that the local form of the Noether current (3) is not the correct way to introduce the chemical potential on the lattice, since it does not reproduce the correct  $U(1)$  symmetry, one can fix this, without using the exponential form. Indeed, if one calculates the Noether current associated to the  $U(1)$  symmetry in the *discretized* version of the fermionic action, what we get is the Noether Current (3) in point-split form, i.e.:

$$\hat{J}_4(n) = \bar{\psi}_n U_4(n) \gamma_4 \psi_{n+\hat{4}} + \bar{\psi}_n U_4^\dagger(n-\hat{4}) \gamma_4 \psi_{n-\hat{4}} \quad (9)$$

This is the exact, conserved Noether Current associated to the  $U(1)$  symmetry on the lattice, not the one described in (3). So, if one wants to introduce the chemical potential in the exact same way that he does in the continue, once has to add this term to the discrete action, multiplied by the factor  $a\mu$ . Then, if one calculates the Dirac operator the naive fermionic discretization in this way, he obtains:

$$D_{nm}(\mu) = D_{nm}(0) + \mu a \left( \gamma_4 U_4(n) \delta_{n,m-\hat{4}} + \gamma_4 U_4^\dagger(m) \delta_{n,m+\hat{4}} \right) \quad (10)$$

And if one substitutes the Dirac operator at zero chemical potential,  $D_{nm}(0)$  he gets:

$$D_{nm} = \frac{1}{2} \sum_{i=1}^3 \left( \gamma_i U_i(n) \delta_{n,m-\hat{i}} - \gamma_i U_i^\dagger(m) \delta_{n,m+\hat{i}} \right) + am \delta_{nm} + \quad (11)$$

$$+ \frac{1}{2} \left( \gamma_4 (1 + \mu a) U_4(n) \delta_{n,m-\hat{4}} - \gamma_4 (1 - \mu a) U_4^\dagger(m) \delta_{n,m+\hat{4}} \right)$$

That is the exact same result of (8), but with  $e^{\pm\mu a} \leftrightarrow (1 \pm \mu a)$ . If we do so, since the function  $f(a\mu) = 1 \pm \mu a$  does not satisfies the relations (5), that we choose to get the correct, physical results, if we calculates the energy density on the lattice using the Dirac operator above, we found the same quadratic divergences that we discussed in last year discussion [1]. But why so? Where are these divergence in the continuum, if this time we introduced the chemical potential in a way completely equivalent to the continuum case? The answer is simple: they are there in the continuum as well, and they get eliminated during the renormalization process. In Ref. [5], indeed, the author shows that if one analytically calculates the quark number density in the continuum, he finds the exact same divergences.

So, summarizing, in the continuum we introduce the chemical potential as a coupling to the conserved Noether charge of  $U(1)$  symmetry. This term introduces some unphysical divergences: when we discretize the action on the lattice we can take advantage of the liberty in the discretization to introduce a parametrization of the chemical potential that cancels, by construction, these unphysical divergences, and this is exactly what we do with the exponential notation. In the continuum, though, we cannot modify the theory as we do in the discrete, so we will use renormalization to eliminate the unphysical divergences.

Following the linear approach, one can show that the if we introduce the chemical potential in the same way as it is done in the continuum, i.e. following Eq. (11) we will find the exact same divergences that there are in the continuum. These unphysical contributions can be calculated in the lattice, and manually subtracted from the observables. If we do so, and compare it with the case of the exponential approach, we will find compatible results: that means that, exactly as in the continuum case, there are no more divergences due to some discretization effects, since the exponential case was free of divergences by construction.

So, if one has to manually calculate the divergent contributions and subtract them to obtain the result of the exponential approach, why bother at all with this method? This will depend on the sign problem and the strategy that we will use to try to beat it. So, let us tackle this problem in the next section.

## 2 Sign Problem

In this section I will first summarize the sign problem, introduced last year in [1]. Alternatively, for more details, see e.g Ref. [2].

The easiest way to see the arise of the sign problem in the context of finite density is to use  $\gamma_5$ -hermiticity: most of the Dirac operators, indeed, enjoy the following property:

$$\gamma_5 D \gamma_5 = D^\dagger \quad (12)$$

That, in turn, implies (using Binet Theorem and  $\det(\gamma_5) = 1$ ):

$$\det(D) = \det^*(D) \quad (13)$$

Thus, the determinant that appears in the partition function that we use to simulate our system is real. That is a good news for us, since we want to interpret it as a probability distribution.

Sadly, the presence of the chemical potential term in the Dirac operator that we showed in Eq. (8) or in Eq. (11) breaks  $\gamma_5$ -hermiticity, i.e.:

$$\gamma_5 D(\mu) \gamma_5 = D^\dagger(-\mu) \quad (14)$$

Indeed, in both our prescriptions the chemical potential enters as  $\mu$  when it multiplies the link  $U_4$ , and as a  $-\mu$  in front of  $U_4^\dagger$ . This follows directly from the correct way to implement it on the lattice, having the meaning of a shift of the temporal gauge field, as we explained in Eqs. (7). Then, when I move the right side  $\gamma_5$  to the left, I get  $D^\dagger$ , except for the fact that a factor  $\mu$  is multiplying  $U_4^\dagger$  and a factor  $-\mu$  is multiplying  $U_4$ , since  $\mu$  it is not effected by  $\gamma_5$ . Hence, we obtain the equation above, from which it directly follows:

$$\det[D(\mu)] = \det^*[D(-\mu)] \quad (15)$$

That means that if  $\mu = 0$  we recover  $\gamma_5$ -hermiticity, and then a real determinant. Otherwise, for  $\mu \neq 0$  we have a complex determinant, and a complex probability distribution, that cannot be sampled anymore. This is the sign problem, in the framework of finite density.

A clear example of the problem, for a simple toy model is illustrated in Fig. (1): the probability distribution, from Gaussian becomes highly oscillatory. Every configurations in the ensemble matters, and it becomes practically impossible to use importance sampling.

Here, the only thing that I want to add to last year discussion, is to point out that the sign problem is not inherently linked to finite density, but it is a more general problem, that can arise in several fields in which we need to use Monte Carlo simulations.

For example, since last year we discussed a bit also about topology, that is a problem that arises as well in the study of the  $\theta$  term of QCD as well. If we consider the axial anomaly of  $U(1)_A$  chiral symmetry, indeed, this adds a term to the Lagrangian that, in principle, creates a sign problem. Indeed, the partition function of QCD in Euclidean space, in presence of a  $\theta$ -term is:

$$Z(\theta) = \int \mathcal{D}A \mathcal{D}\bar{\psi} \mathcal{D}\psi e^{-S_{QCD}} e^{i\theta Q} \quad (16)$$

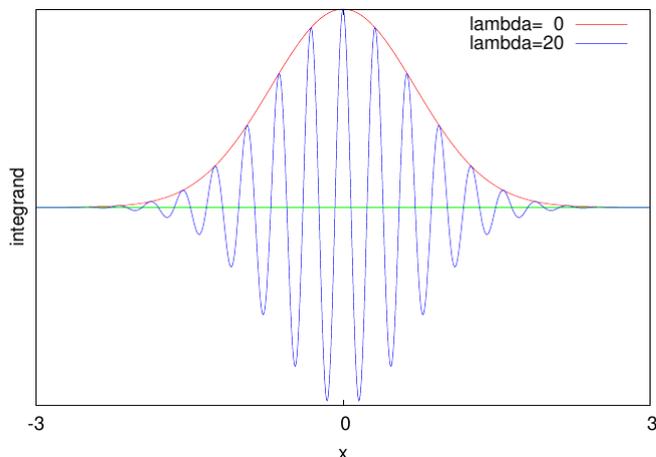


Figure 1: Consider a toy partition function of a Gaussian plus an imaginary phase:  $Z(\lambda) = \int_{-\infty}^{+\infty} dx e^{-x^2+i\lambda x}$ . In the figure is depicted the weight of the distribution with and without the factor  $\lambda$ : a graphical explanation of the sign problem

Where  $Q$  is the topological charge. Here, the presence of the  $\theta$ -term creates a complex phase term in the action, that again prevent it for an effective sampling of the full theory. To have a naive understanding on where this complex phase comes from, one can think to the equivalent system of the particle in a circle, that shares most of the topological properties of QCD, on a qualitative level: in the presence of a magnetic field piercing the circle, the system acquires a  $\theta$ -term, equivalent to the one in QCD and, in that case, the presence of the velocity in the Lagrangian gives birth to the complex phase.

Not only the sign problem appears in quantum field theory in finite density or non-zero  $\theta$ -term, but it appears also in other fields, like nuclear physics or condensed matter. That means that the sign problem is an unfortunate feature of several quantum systems and we do not expect that it could be possible to find a general solution to it. One can, still, try to search a particular solution to each case affected to this problem, and that is what I will try to outline in the following section.

### 3 Trying to Circumvent the Sign Problem

As I said before, there are some mathematical arguments that say that there is no general solution to the sign problem. So, instead of focusing on a general receipt to solve all possible sign problems, let us first try to focus on finite density QCD. In this section I will briefly go over several strategies that are used in order to try to circumvent the sign problem in this framework. As you will see, this is already quite the challenge, and this is still an open problem after decades of work on it.

In the following subsections I will go through these methods, following roughly the structure of Ref. [6].

### 3.1 Reweighting

The most logical and straightforward method to tackle the sign problem is the so called *reweighting*.

In general, in LQCD, we are interested in evaluating integrals that represents expectation values of an observable  $\mathcal{O}$ . These integrals, after integrating away the fermion fields, have the following form:

$$\langle \mathcal{O} \rangle = \frac{\int \mathcal{D}U \mathcal{O} e^{S_{YM}} \det D}{\int \mathcal{D}U e^{S_{YM}} \det D} \quad (17)$$

Where  $S_{YM}$  is the usual Yang-Mills action, and  $D$  is the fermion matrix. And, as we said in the previous section, the sign problem occurs when  $\det D$  becomes complex and then we cannot interpret  $\rho(U) \equiv e^{S_{YM}} \det D$  as a probability distribution.

Now, in principle we can choose an appropriate, well defined statistical weight  $w(U)$  instead of the oscillatory  $\rho(U)$ , and modify the previous equation in the following way:

$$\langle \mathcal{O} \rangle = \frac{\int \mathcal{D}U \mathcal{O} \frac{\rho(U)}{w(U)} w(U)}{\int \mathcal{D}U \frac{\rho(U)}{w(U)} w(U)} = \frac{\langle \mathcal{O} \frac{\rho}{w} \rangle_w}{\langle \frac{\rho}{w} \rangle_w} \quad (18)$$

The denominator in the equation above is called, then, *Average Sign*. In this way, then we can sample the link variables according to our new, real and positive weight  $w$  and insert the oscillatory phase in the observables.

Even if in principle this approach solves the sign problem, most of the time it is not possible to apply it successfully. To understand why, let us first choose a value of the weight. What usually people do, since the goal is to separate the oscillatory phase of the determinant, is to consider  $\det D \equiv |\det D| e^{i\theta}$ , and choose  $w = |\det D|$ . If one does that, then the theory described by this probability distribution is often called *Phase Quenched*. The average sign defined in equation above, then, with this choice, becomes (see also Ref. [7] for more details):

$$\left\langle \frac{\rho}{w} \right\rangle_w \rightarrow \left\langle \frac{|\det D| e^{i\theta}}{|\det D|} \right\rangle_{pq} = \langle e^{i\theta} \rangle_{pq} = \frac{\int \mathcal{D}U |\det D| e^{i\theta}}{\int \mathcal{D}U |\det D|} = \frac{Z}{Z_{pq}} \quad (19)$$

Where the subscript  $pq$  means evaluated in the phase quenched theory. We recall that the partition function can be written as a function of the free energy on the lattice,  $f$ , and since  $f = -(T/V) \log Z$ , then, substituting  $Z$  in equation above I get:

$$\left\langle \frac{\rho}{w} \right\rangle_{pq} = e^{-V \Delta f / T} \quad (20)$$

Where  $\Delta f = f - f_{pq}$ , that is the difference between the two free energies, the one relative to our physical system and the other relative to the phase quenched one.

From here we can instantly see a two-folded problem. If  $\Delta f \neq 0$ , then the average sign at the denominator in Eq. (18) goes to zero exponentially fast, when the volume increases. That means that the signal to noise ratio drops and to achieve a good statistical precision the number of configurations has to

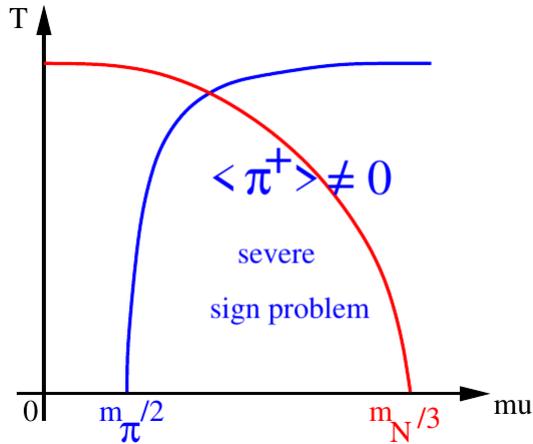


Figure 2: Sketch of the phase diagrams of QCD with two flavors (red) and phase quenched QCD (blue), from Ref. [6]. The region  $m_\pi/2 < \mu < m_N/3$  (where  $m_N$  is the nucleon mass) is the so-called Silver Blaze region, where there is an overlap problem and the sign problem is severe.

grow exponentially with the volume as well. Since we are often interested in the extrapolation to the thermodynamic limit, this becomes unfeasible quite soon. This is also sometimes referred as *overlap problem*, since it occurs when the two theories (the original and the phase quenched one) differs in a certain region.

The reweighting is a strategy so common and natural, that sometimes even the definition of the sign problem is slightly changed accordingly to this strategy: that is why sometimes people says that a system suffers from a sign problem when the numerical effort grows exponentially with the volume. If  $\Delta f = 0$ , indeed, this is not the case and the approach above actually solves the sign problem. Sadly, as you probably already guessed, this is not the case for most of the interesting systems, including QCD at finite density.

Clearly, then, the sign problem depends on the magnitude of  $\Delta f$ . If  $\Delta f$  is not zero, but still extremely small, then the system suffers from a sign problem, but it may be possible to simulate it anyway, using the reweighting technique, even up to considerably large volume. In this case we speak about a *mild* sign problem. If  $\Delta f$  is large, instead, then the sign problem becomes *severe*.

In this last case, the problem occurs even in smaller volume. Consider a case in which, for example, the phase quenched system goes through a phase transition at a certain  $\mu_c^{(pq)} < \mu_c$ , where  $\mu_c$  is the critical chemical potential of the original theory. Around  $\mu_c^{(pq)}$  the free energy has a gap, and the large value of  $\Delta f$  kills the signal-to-noise ratio even at fixed volume. Physically speaking, indeed, the phase quenched system depends critically on the chemical potential, while the original theory is mostly independent on the chemical potential before the phase transition. Looking at Eq. (18) we see that the critical dependence of the phase quenched system has to be exactly canceled by the oscillatory phase in the observable, since the original ensemble does not depend on  $\mu$ . Then the observable oscillate drastically in that region and you need an enormous amount of statistic to sample it correctly. In addition, the error is hard to estimate: because of this mechanism there is a huge systematic error in this

region of the parameters. This is the so called *Silver Blaze region* and people often speak about this mechanism as the Silver Blaze problem (see e.g. Ref. [7]). An example of this exact problem is showed in Fig. (2), for 2 flavors QCD.

### 3.2 Taylor Expansion

If the sign problem is severe, then we cannot use the reweighting to circumvent it. Then, there are a pair of standard techniques that indeed are based on a similar mechanism.

The first one is the Taylor Expansion method. If one is interested to analyze the thermodynamic properties of QCD, indeed, one can focus on the pressure, that is defined as:

$$P = \frac{T}{V} \log Z(\mu/T) \quad (21)$$

Since most of the other thermodynamical observables can be obtained starting from the pressure (like, for example, the energy and entropy density). Now, since we know that the chemical potential is introduced as an external temporal component of the gauge field, and since the action is invariant under time reversal, then:

$$Z(\mu/T) = Z(-\mu/T) \quad (22)$$

That means that the partition function is an even function of the chemical potential. For small  $\mu/T$  I can, then, Taylor expand the pressure in the following way:

$$P = T^4 \sum_{n=0}^{\infty} c_{2n}(T) \left(\frac{\mu}{T}\right)^{2n} \quad (23)$$

Where, just the even power of  $\mu/T$  appears. The coefficients  $c_{2n}$  are the generalized quark number susceptibilities, that are indeed proportional to derivatives of the free energy  $f = (T/V)\log Z$ :

$$c_{2n}(T) = \left. \frac{\partial^{2n} \log Z}{\partial \mu^{2n}} \right|_{\mu=0} \quad (24)$$

Indeed, the first coefficient is proportional to the actual quark number susceptibility:

$$c_2 \propto \frac{T}{V} \frac{\partial^2 \log Z}{\partial \mu^2} = \frac{\partial n_q}{\partial \mu} = \chi_q \quad (25)$$

The important result here is in Eq. (24): since the coefficients of the Taylor series are evaluated at chemical potential equal to zero, in principle one can explore the phase diagram of QCD at small value of  $\mu/T$ , successfully circumventing the sign problem, at least for small values of the chemical potential. Sadly, this is not the only disadvantage in this technique. To hint to these technical difficulties, let us first look to the general form of these coefficients. Since the only dependence on the chemical potential is in the fermionic matrix,  $\partial \log Z / \partial \mu = \partial \log(\det D) / \partial \mu$ . Now, writing  $\det(D) = e^{\text{Tr}(\log D)}$ , then the derivative above becomes:

$$\frac{\partial \det D}{\partial \mu} = \text{Tr} \left[ D^{-1} \frac{\partial D}{\partial \mu} \right] \quad (26)$$

In principle this trace can be calculated numerically with the method of the noisy estimator. Still, if one increase the order of the expansion, i.e. increasing the order of the derivative, it gets, for example:

$$\frac{\partial^2 \det D}{\partial \mu^2} = \text{Tr} \left[ D^{-1} \frac{\partial^2 D}{\partial \mu^2} \right] - \text{Tr} \left[ D^{-1} \frac{\partial D}{\partial \mu} D^{-1} \frac{\partial D}{\partial \mu} \right] \quad (27)$$

$$\begin{aligned} \frac{\partial^3 \det D}{\partial \mu^3} = & \text{Tr} \left[ D^{-1} \frac{\partial^3 D}{\partial \mu^3} \right] - 3 \text{Tr} \left[ D^{-1} \frac{\partial D}{\partial \mu} D^{-1} \frac{\partial^2 D}{\partial \mu^2} \right] \\ & + 2 \text{Tr} \left[ D^{-1} \frac{\partial D}{\partial \mu} D^{-1} \frac{\partial D}{\partial \mu} D^{-1} \frac{\partial D}{\partial \mu} \right] \end{aligned} \quad (28)$$

And so on. It is already clear that the number of terms grows extremely fast and then the numerical effort grows equally faster.

The number of terms is not the only technical difficulty: the more we increase the order, the more important finite-size effects become. In addition, again most of this trace depend on the volume: at large volume, then one needs more and more random sources to evaluate the traces.

Some of these problems can, at least in part, be solved using clever tricks. I will not dive into a full description of these tricks, but I will just hint at some of them.

Concerning the growing of coefficients, for example, one can use the prescription on how we introduce the chemical potential on the lattice as an advantage: indeed, if one use the standard exponential notation, the form of the coefficients is the one of Eqs. (27-28). If we use the linear prescription, though, described in the first section, then all the terms with a second order derivative (and all the higher orders) will become zero. Then, just the last terms of Eqs. (27-28) survive.

Finally, to deal with the increasing of random vectors to evaluate the traces, one can use some technical tools like smart algorithms for preconditioning and deflation, to try to speed up the evaluation of the traces.

### 3.3 Imaginary Chemical Potential

Another method, that shares most of the properties and difficulties of the Taylor Expansion above is the imaginary chemical potential approach.

It is based on the fact that, looking at Eq. (14), if we assume that the chemical potential is a complex, non physical, parameter, then  $\gamma_5 D(\mu) \gamma_5$  is equal to  $D^\dagger$  except for the sign of the chemical potential *and* a complex conjugate factor on the chemical potential, that is still not effected on the  $\gamma_5$ . Then, Eq. (14) becomes:

$$\gamma_5 D(\mu) \gamma_5 = D^\dagger(-\mu^*) \quad (29)$$

That means that if the chemical potential is purely imaginary, i.e.  $\mu = i\eta$ , then  $\gamma_5$ -hermiticity is recovered, and the determinant is once again a real, positive number. Then we can use our standard Monte Carlo technique again.

Still, to recover the physical results one has to analytical continue the imaginary  $\mu$  result to the real value. To do so, once again note that observables like pressure and so on depend on the complex variable:

$$z \equiv \left(\frac{\mu}{T}\right)^2 = -\left(\frac{\eta}{T}\right)^2 \quad (30)$$

Then one can simulate the theory with these negative values of  $z$  and then expand the observables in a power series in  $z$  and then extrapolate to positive, physical  $z$ .

Then the approach becomes similar to the Taylor Expansion method discussed before. Again, one finds himself to deal with the same advantages and disadvantages explained in the previous section.

Both of these methods are confined in a regime of small  $\mu$  and they seem to work properly and agree to each other, if compared. Still, they are both numerically demanding, and fail (at least for now and probably in the near future) to explore regimes of large chemical potential, that can be interesting to study extreme and exotic systems like neutron stars and superconductive color fluids. Then, one can try to search a radically different way to solve the sign problem completely, abandoning the expansion methods and the Monte Carlo technique, and searching for a quite different approach. This is what it is done in the methods that simply accept that the action is complex now, and try to use this property as a feature instead of an obstacle, diving in the complex plane. In the last two subsections I will try to hint to two quite popular ways to solve the sign problem using the properties of the complex plane.

### 3.4 Complex Langevin

A completely different approach, that does not require any sampling according to a probability distribution, is the so called *Complex Langevin* evolution (for a recent review see e.g. Ref. [8]).

This method is based on the *stochastic quantization* of a system, instead of the functional quantization that we usually use in quantum field theory. I will try to highlight the main idea of this approach. Let us suppose that we have a partition function of a generic field theory:

$$Z = \int \mathcal{D}U e^{-S[\phi]} \quad (31)$$

Now, we can introduce an additional dimension, a kind of fictitious stochastic time, and evolve the field  $\phi$  through this time, following the so-called *Langevin Equation*:

$$\frac{d\phi}{dt} = -\frac{\delta S[\phi]}{\delta \phi} + \eta \quad (32)$$

Where  $\eta$  is a random noise term, often normally distributed around zero. Qualitatively, the evolved field can be interpreted as a quantized one: indeed, the meaning of equation above is that the field is drifted to the minimum, classical value of the action. In addition, we explicitly introduced some quantum fluctuation around the minimum of the action. If the noise is properly tuned one can reproduce the correct, physical, quantum physics.

More quantitatively, one can show that, after some equilibration time, the distribution of the  $d + 1$  field  $\phi(t)$  is the same that it would have been described by the probability distribution of  $d$  dimensional field  $\phi$  expressed in Eq. (31), where the extra dimension is just the fictitious stochastic time.

Practically speaking, this works in the same exact way of a Markov Chain in Monte Carlo methods: we start from a random configuration, a first ensemble of fields  $\phi$ . Then we let the system evolves following the Langevin equation and we let it equilibrate towards the equilibrium distribution (that we know to be the same described by our original theory). Then we compute the expectation value of the observable  $\mathcal{O}[\phi(t)]$  on a set of configurations  $\{\phi(t)\}$ , each of them produced at a fixed stochastic time  $t$ . That's it, or at least, that is the main idea.

Now, what happens if the action that drives our Langevin Equation is complex? This time, we do not need to face a conceptual problem. Indeed, the Langevin equation does not require the definition of any probability: you do not need to perform any Metropolis test, you just need to evolve the fields according a complex action. Practically speaking, without diving into the details, what people usually do is to complexify the degrees of freedom, such as the fields  $\phi \rightarrow \phi_R + i\phi_I$ , and then let evolve the real and imaginary part according to the real and imaginary part of the action. Then we use two sets of Langevin equations, one for each components of the new, complex field:

$$\frac{d\phi_R}{dt} = -\frac{\delta\Re(S[\phi])}{\delta\phi} + \eta_R \quad (33)$$

$$\frac{d\phi_I}{dt} = -\frac{\delta\Im(S[\phi])}{\delta\phi} + \eta_I \quad (34)$$

Each of the equations has its own drift term (the derivative of the real or imaginary part of the action) and its own random noise. Usually the real, physical one is taken normally distributed around zero, while the imaginary one is set to zero. In this way, the imaginary part of the fields gets driven directly to the fix point of the action, while the quantum fluctuations are inserted just in the physical, real part.

Of course, this is not the end of the story. Since now the manifold in which the action lives is a complex one, there is not a mathematical proof that the CL equation would actually converge to the right, physical fixed point, or actually even converge at all. Indeed, it is possible that the action has some complex, unphysical minima, and that, depending on the starting conditions, the fields are driven towards the wrong one. It can also happen that the action has some poles in the complex plane, that could lead to breakdown of the evolution algorithm. There are even some more subtle, mathematical problems, related to the boundary conditions: in real space the action, as we all know, goes to zero at infinity to describe a probability distribution. In principle, that is not necessary true in a complex manifold. This can create some unphysical runaway trajectories that drifts towards infinity, far away from the physical, true fixed point.

Still, this is a very active topic and a lot of research groups are currently fighting again all these problems: people tried and are still trying a lot of different strategies to avoid these difficulties, from adaptive step size of the numerical

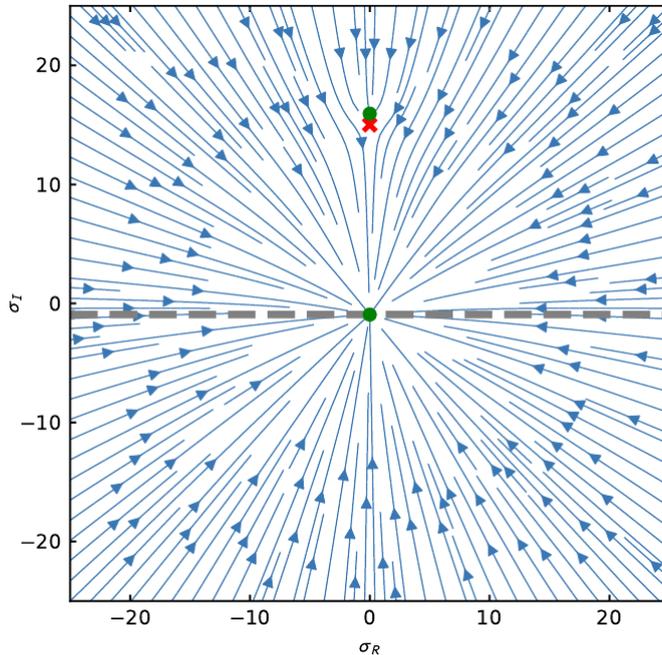


Figure 3: Flow diagram of the  $\phi^4$  theory. Here the action is written as a function of some auxiliary variables  $\sigma$ , obtained by a shift of the field variable  $\phi$ . The action then becomes  $S(\sigma) = \frac{\lambda}{24}\sigma^2 - \frac{1}{2}\log \frac{\lambda}{12\mu + 2i\lambda\sigma}$ . In this form it is easy to find the minima (fixed points, in green) and the pole at  $\sigma_p \equiv i6\frac{\mu}{\lambda}$ . For more technical details see Ref. [8], where this figure comes from.

integration of the Langevin Equation, to the so-called gauge cooling technique, that modifies the complexified links to keep them near the real solutions and away from these excursion trajectories.

If one is interested in visualizing mathematically the effects of the drift equation, and the potential causes of problems, he can have a look to Fig. (3). There, it is visually depicted the flow of the field variables in a  $\phi^4$  toy model, described by the action  $S(\phi) = \frac{\mu}{2}\phi^2 + \frac{\lambda}{4!}\phi^4$ . We can see that in the complex plane, that simple action has two fix points (the green circles): one physical, in the real axis, and another, purely imaginary one. In addition, there is a pole (red cross) just below the imaginary critical point. Now, this already shows a lot of potential problems: the fields could be locally driven to the unphysical fix point, or even crash against the pole, initially attracted by the local minimum. In this case, the problems solve by themselves, since even if a field is dragged toward the imaginary fixed point, the real fluctuations will move it back on tracks, towards the physical fixed point, avoiding the pole. But, in more complicated systems, this can be an issue, as well as some excursion trajectories towards infinity due to non zero boundary conditions.

In conclusion, a full description of these problems and their solutions is not the goal of this discussion: my goal is just to give the idea of the main concept behind the CL evolution, the most dangerous problems and the hints of the solutions. For more details see e.g. Ref. [8] and references therein.

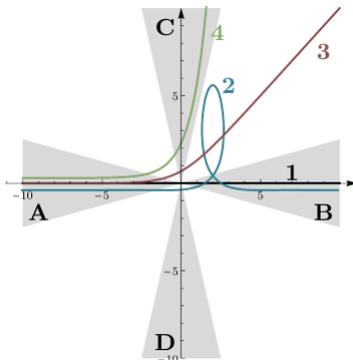


Figure 4: "Good regions" for the integral in Eq. (38), from Ref. [9]

### 3.5 Thimbles

Another strategy that involves the advantages of the complex plane is the one based on the so called *Lefschetz Thimbles*.

If the complex Langevin dynamics used the stochastic quantization as a strategy for avoid the sampling of the probability, the thimbles approach uses the Cauchy Theorem and the deformation of the contour of the integration to try so solve the sign problem. For a recent review see e.g. Ref. [9].

The main idea of the contour deformation is that, in the complex plain, one can smoothly deform the path of integration, without changing the value of the integral. That is due to the Cauchy Theorem that states, I briefly recall it, that the integral of an analytic function of a complex variable  $f(z)$ , over a close curve  $C$  vanishes, i.e.:

$$\oint_C f(z) = 0 \quad (35)$$

Then, indeed, if we have an integral over a contour  $\Gamma_1$  from  $A$  to  $B$ , if I take another contour  $\Gamma_2$  that starts and end from and to the same points, then:

$$\int_{\Gamma_1} f(z) = \int_{\Gamma_2} f(z) \quad (36)$$

Since:

$$\int_{\Gamma_1} f(z) - \int_{\Gamma_2} f(z) = \oint_C f(z) = 0 \quad (37)$$

Where  $C = \Gamma_1 - \Gamma_2$  is the close curve that goes from  $A$  to  $B$  along  $\Gamma_1$  and from  $B$  to  $A$  through  $\Gamma_2$ . Then, the last equality above is zero for the Cauchy Theorem.

Of course, this holds just in a domain where  $f(z)$  is analytic, and then the Cauchy Theorem it is applicable. Then, we can deform the contour of the integration until we find a divergence.

A clear example of this is the easy integral:

$$\int_{-\infty}^{+\infty} d\phi e^{-\phi^4} \quad (38)$$

Here, if we complexify the variable  $\phi$ , we find four "good regions" depicted in Fig. (4) where the integral above converges. The real integral in Eq. (38) is the path 1 in the figure. For the reasoning above, it can be smoothly deformed into the path 2, without changing the result. Along path 3, instead, the integral diverges. Then, if one deforms path 1, up to path 4 (that starts in  $A$  and ends in  $C$ ) the integral converges, but you cannot apply the Cauchy theorem anymore, and indeed the two integrals are different.

In this case, then, you can divide all the possible paths in four so called *homology classes*, where all the integrals coincide. It is possible to show that, in general, the number of possible classes is equal to the number of saddle points of the function.

In this case, the difference between path 1 and 4 it is easy to visualize simply because they do not end in the same region: still, one can generalize the Cauchy Theorem in more than two dimension, using the fact that the theorem holds until  $f(z)$  is analytic in the region in which you deform the contour. Then, in a multi-dimensional manifolds, you can smoothly deform it until you reach a region where there is a pole of the function: in other word, you can deform the manifold in any other that belongs on the same homology class.

In our finite density case, the manifold where the functional integral is defined is  $\mathbb{R}^N$ : we can then, in principle, try to find another manifold equivalent to  $\mathbb{R}^N$  in which the sign problem is not a problem anymore. In order to find this, I again introduce a fictitious flow time, similar to the one of the CL case. Then I flow each point  $\phi \in \mathbb{R}^N$ , according to the so called *holomorphic flow* ( $\mathcal{F}$ ) equation:

$$\frac{d\phi}{dt} = \frac{\overline{\partial S}}{\partial \phi} \quad (39)$$

Now, one can show that the action along this holomorphic flow has two important properties: first of all, the real part of  $S$  increase monotonically. That assures that the integral is well defined, and it converges even faster on this flowed manifold than in  $\mathbb{R}^N$ . Secondly, and this is even more important for our purposes, that the imaginary part of the action remains constant along the flow. That means, that the imaginary phase of the integrand that appears in finite density stays constant. Hence, in principle, the phase is not highly oscillatory as it is in  $\mathbb{R}^N$ , and then the flowed manifold is not affected by the sign problem, even if it is equivalent to the real multidimensional axis.

This is the main idea of the thimbles: to clarify the shape of the flowed manifold, then, I now define the *thimble* around a critical, fix point of the action  $\phi_c$  as a small region  $\mathcal{B}$  around  $\phi_c$  asymptotically flowed according to the downward flow  $\tilde{\mathcal{F}}$  (along the steepest descent):

$$\frac{d\phi}{dt} = -\frac{\overline{\partial S}}{\partial \phi} \quad (40)$$

Then, the thimble is  $\mathcal{T} = \tilde{\mathcal{F}}_{t \rightarrow \infty}(\mathcal{B})$ . In the same way one define the *dual thimble* attached to  $\phi_c$ , as the collection of all points that, once they are asymptotically flowed according the upward flow  $\mathcal{F}$  in Eq. (39), reach the critical point  $\phi_c$ .

It is possible to see a visual representation of the thimbles and their effect related to the flowed manifold  $\mathcal{M} = \mathcal{F}(\mathbb{R}^N)$  in Fig. (5). This is a simple case where we have a one dimensional action that describes a toy model of a fermionic

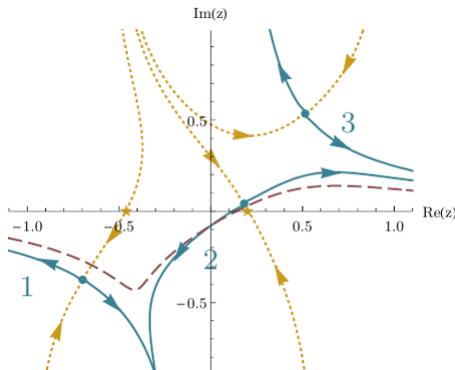


Figure 5: Thimbles, Dual thimbles and critical points of the action  $S = \phi/G - \log[(p^2 + i\mu)^2 + (\phi + m)^2]$ , for  $G = 1.1e^{i0.05}$ ,  $p = 1$ ,  $\mu = 0.3$  and  $m = i0.1$ . For more details see Ref. [9], from where this picture is taken.

action once the fermions are integrated away. Still, one can see that the action in the complex plane has three critical points (the blue dots in the figure). Then, one can construct three thimbles, one for each fix point (the blue line): they start from the critical points and go through the steepest descent of the action. Finally one can define a dual thimbles for each critical points (yellow line): they are indeed all the points that, following the upward flow, they reach the critical points.

If one then take the real axis, i.e. the physical domain of our partition function, and it starts to flow it according to  $\mathcal{F}$ , it follows the direction of the dual thimbles, approaching the thimbles (the dashed blue curve). We stress that, in this case, the flowed manifold  $\mathcal{M}$  approaches the union of the two thimbles 1 and 2. This is indeed a general statement: it is possible to show that only the thimbles associated to the dual thimbles that cross the physical manifold (marked with a yellow star in Fig. (5)) contribute to the shape of the flowed manifold. Qualitatively, this is clear, since the real axis is flowed along the direction defined by the dual thimbles. If one dual thimble does not cross the original manifold, then its effect does not occur, and the manifold will never approach the related thimble.

In conclusion, thanks to the thimbles, it is possible to define a manifold where the imaginary phase of the integrand is constant and then without the sign problem, or with a very mild one.

Of course, this is again not the end of the story: in more complicated systems it is hard to understand how many thimbles contribute to create the flowed manifold, and finding these thimbles becomes more and more complicated. In addition, the change of variables during the flow could create a residual phase due to the Jacobian of the transformation, and this is a model-dependent problem. Nevertheless, this is a very active field and people are struggling against the problems of this approach. For a review of the recent progresses, see again Ref. [9].

## 4 Conclusion and Summary

To conclude this discussion, let me summarize the ideas that I tried to introduce to all of you.

First of all, I reminded how to insert the chemical potential on the lattice, hinting about the linear approach as an alternative method to the exponential, traditional one.

Secondly, I quickly discussed why the chemical potential creates the infamous sign problem once we included it as a parameter on the lattice.

Finally, in the last and largest part of the talk, we went together through the most famous and important methods to partially circumvent the sign problems: we started with the more traditional methods to explore the small chemical potential regions, like the Taylor Expansion and imaginary chemical potential method. In the last part, instead, I hinted to some more ambitious methods to solve the sign problem entirely, using different strategies, involving the properties of the complex plane. For each of these different strategies I tried to highlight the main ideas, the major advantages and the open issues.

I hope that this talk will be the starting point of a nice discussion about methods to attack and maybe, hopefully one day overcome, the sign problem.

## References

- [1] [https://www2.physik.uni-bielefeld.de/fileadmin/user\\_upload/endrodi/template/JournalClub/Bollweg\\_chemical\\_potential\\_notes.pdf](https://www2.physik.uni-bielefeld.de/fileadmin/user_upload/endrodi/template/JournalClub/Bollweg_chemical_potential_notes.pdf).
- [2] C. Gattringer and C. B. Lang, *Quantum chromodynamics on the lattice*, vol. 788. Berlin: Springer, 2010.
- [3] P. Hasenfratz and F. Karsch, “Chemical potential on the lattice,” *Physics Letters B*, vol. 125, no. 4, pp. 308 – 310, 1983.
- [4] R. V. Gavai and S. Sharma, “Anomalies at finite density and chiral fermions,” *Phys. Rev. D*, vol. 81, p. 034501, Feb 2010.
- [5] R. V. Gavai and S. Sharma, “Divergences in the quark number susceptibility: The origin and a cure,” *Physics Letters B*, vol. 749, p. 8–13, Oct 2015.
- [6] P. de Forcrand, “Simulating qcd at finite density,” 2010.
- [7] G. Aarts, “Complex langevin dynamics and other approaches at finite chemical potential,” 2013.
- [8] C. E. Berger, L. Rammelmüller, A. C. Loheac, F. Ehmman, J. Braun, and J. E. Drut, “Complex langevin and other approaches to the sign problem in quantum many-body physics,” 2019.
- [9] A. Alexandru, G. Basar, P. F. Bedaque, and N. C. Warrington, “Complex paths around the sign problem,” 2020.