

SWANSEA UNIVERSITY

Physics Department

Numerical methods for the sign problem in Lattice Field Theory

Submitted to Swansea University in fulfilment of the requirements
for the Degree of Doctor of Philosophy



Lorenzo Bongiovanni

Swansea University 2015

Lorenzo Bongiovanni: *Numerical methods for the sign problem
in Lattice Field Theory*

Supervisors :
Gert Aarts
Biagio Lucini

Date of Submission :
25/9/2015

Abstract

The great majority of algorithms employed in the study of lattice field theory are based on Monte Carlo's importance sampling method, i.e. on probability interpretation of the Boltzmann weight. Unfortunately in many theories of interest one cannot associate a real and positive weight to every configuration, that is because their action is explicitly complex or because the weight is multiplied by some non positive term. In these cases one says that the theory on the lattice is affected by the *sign problem*. An outstanding example of sign problem preventing a quantum field theory to be studied, is QCD at finite chemical potential.

Whenever the sign problem is present, standard Monte Carlo methods are problematic to apply and, in general, new approaches are needed to explore the phase diagram of the complex theory. Here we will review three of the main candidate methods to deal with the sign problem, namely complex Langevin dynamics, Lefschetz thimbles and density of states method.

We will first study complex Langevin dynamics, combined with the gauge cooling method, on the one-dimensional Polyakov line model, and then we will apply it to pure gauge Yang-Mills theory with a topological θ term. It follows a comparison between complex Langevin dynamics and the Lefschetz thimbles method on three toy models, which are the quartic model, the U(1) one-link model with a μ dependent determinant, and the SU(2) non abelian one-link model with complex β parameter.

Lastly, we introduce the density of state method, based on the LLR algorithm, and we will employ it in the study of the relativistic Bose gas at finite chemical potential.

Contents

Introduction	6
1 Quantum field theory on the lattice	13
1.1 Path Integral approach to quantum theory	13
1.2 Monte Carlo simulations	15
1.3 An example: the Scalar Field	18
1.4 Gauge Theories on the Lattice	21
1.5 The continuum limit	23
2 Complex Langevin dynamics	25
2.1 Stochastic dynamics	25
2.2 Stochastic quantization of a field theory	27
2.3 Complex Langevin dynamics	29
2.4 Gauge Theories and Gauge Cooling	34
3 Theta term	45
3.1 Vacua of $SU(N)$ and instantons	46
3.2 Topology on the Lattice	54
3.2.1 Gradient Flow for $SU(3)$	57
3.3 Complex Langevin dynamics for $SU(3)$ gauge theory with a θ -term .	60
3.3.1 complex Langevin dynamics	62
3.3.2 Topological charge distribution	67
3.3.3 Topological charge on $SL(3, \mathbb{C})$ manifold	70
4 Lefschetz thimbles	75
4.1 Thimble equations	76
4.2 Thimbles and Langevin dynamics	78
4.3 $U(1)$ model with determinant	82
4.4 $SU(2)$ 1-link model	88
4.4.1 Angle representation	89
4.4.2 Gauge dynamics with cooling	91

5 Density of States	97
5.1 Description of the method	97
5.1.1 LLR algorithm	99
5.2 Relativistic Bose gas at finite chemical potential μ	102
5.2.1 Generalised density of states	103
5.3 The partition function from highly oscillating integral	107
5.3.1 Lefschetz thimble	109
Conclusions	112
Bibliography	117

Introduction

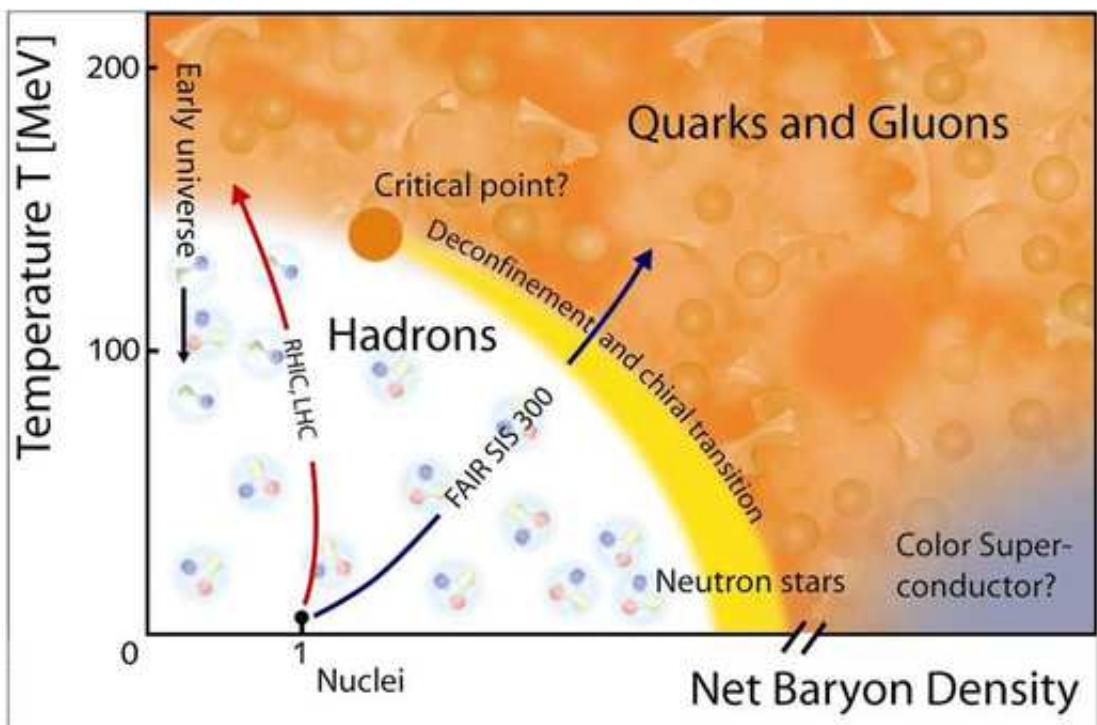


Figure 1: Qualitative representation of quantum chromodynamics (QCD) phase diagram.

Starting from the 1980s, lattice field theory has been developed and it has been proven to be a formidable tool to study quantum field theory (QFT). In general, whenever a theory manifests a non-perturbative behaviour at some energy scale, analytic quantitative solutions are very hard or impossible to obtain.

Discretization on the lattice is a well-established non-perturbative approach for Euclidian time gauge theories [1, 2, 3, 4]. In particular, much progress has been made in the study of quantum chromodynamics (QCD), i.e. the theory of quarks and gluons, which describes a big part of the high energy physics as we know it. The success of lattice field theory is due to the fact that it can be mapped into a

Statistical Mechanics ensemble and, therefore, all the techniques developed for the latter are available. The most acknowledged class of methods, successfully employed to study a vast number of models, is based on Monte Carlo's importance sampling. This is a probabilistic way of exploring the space of configurations of a system, based on the action S of each configuration. More specifically, the Boltzmann weight e^{-S} of a configuration is interpreted as the unrenormalized probability of the configuration itself and, therefore, it will contribute with this weight to the partition function. In this way one is able to build a finite representative sample of the full configuration space, and use it to compute average values of observables.

Unfortunately this way of proceeding is not applicable any more whenever the exponential of the action of the system cannot be interpreted as a probabilistic weight, that is when it is not real and positive. This phenomenon is commonly referred to as *sign problem*.

In this case, a way around is to employ methods that, even though they are still based on Monte Carlo, allow to extrapolate some information even when the action is complex. Some of these are :

- re-weighting [5], where the complex part $e^{i\phi}$ of the weight is incorporated in the observable, while the expectation value is computed only using the real part $|\rho|$ of the weight $e^{-S} = |\rho| e^{i\phi}$, namely the *phase quenched* theory

$$\langle O \rangle = \frac{\langle O e^{i\phi} \rangle_{pq}}{\langle e^{i\phi} \rangle_{pq}} \quad (1)$$

- Taylor expansion in the parameters that trigger the sign problem, so that the quantities one has to compute appear as coefficients, in a theory without sign problem.
- Analytic continuation from a part of the phase diagram where the action is real, typically where the parameters that trigger the sign problem are imaginary [6, 7, 8, 9].
- Strong coupling expansion, in which one exploits the fact that the truncated series of character expansion of the functional integral is a spin-like system with a much milder sign problem than the original one [10, 11, 12, 13].
- Mean field theory [14, 15, 16]

However, those methods are typically restricted to some areas of the theory where the parameters that trigger the sign problem are not too large. There are other approaches which aim to solve the sign problem all together. They are mostly of recent development so they sometimes lack the control which Monte Carlo methods achieved over many years. Some of those approaches are :

- complex Langevin dynamics : a stochastic quantization of the fields is adopted. The degrees of freedom are complexified and a real probability distribution for the observable is generated in the complex space (solution of the Fokker-Planck equation) [17, 18, 19].
- Lefschetz thimble : new manifolds, equivalent to the original domain of integration, are found in the complexified space, along which the imaginary part of the action is constant and, therefore, the integral is (mostly) real [20, 21, 22].
- Density of states : the density of states of a system is computed and the original path integral is reduced to a one-dimensional oscillating integral [23, 24, 25].
- Canonical approach : the canonical partition function at fixed number of particles (or in general conserved charge) is computed, and it is connected with the general partition function via the fugacity expansion [26, 27, 28].
- Reformulation of the theory in term of dual variables : in some theories it is possible an exact map into another theory with other variables where the sign problem is milder or absent[29, 30].

In this thesis we will discuss in detail some of the last group, mainly focusing on complex Langevin dynamics.

There are many theories affected by the sign problem, not only in quantum field theory but also in real time quantum mechanics and condensed matter. One of the most important and challenging is QCD at *finite density* [31, 32]. Here, in Euclidian time, the sign problem is generated by the fermionic determinant

$$Z = \int [DU] e^{-S_{YM}} \det M(\mu), \quad (2)$$

where the links U represent the gluonic degrees of freedom on the lattice, and $\det M(\mu)$ is the result of the Grassman integral over the fermionic part $\det M(\mu) = \int d\psi d\bar{\psi} \exp [\bar{\psi}(\not{D} + m + \gamma_0 \mu)\psi]$. It can be shown that

$$[\det M(\mu)]^* = \det M(-\mu^*), \quad (3)$$

which allows it to be real only if μ is zero or completely imaginary. At finite μ , then, the weight is complex and the sign problem arises together with cancellations in (2). Those are responsible, for example, for an interesting phenomenon called *Silver Blaze*, for which, at $T = 0$, all the thermodynamic observables have to be independent from the chemical potential, up to the nucleon mass, even though the determinant explicitly depends on μ . In particular, the critical baryon chemical potential is the nucleon mass minus the nuclear binding energy. The qualitative explanation for this behaviour is that, before the critical chemical potential, there is not enough energy in the system to create a nucleon at $T = 0$, at $T \neq 0$, the nucleon has still a chance to be created but it will be suppressed by a Boltzmann factor $\sim \exp(-(m_N - \mu_B)/T)$, where m_N is the nucleon mass and μ_B is the critical

baryon chemical potential. This behaviour is totally spoiled [33] if the oscillations are neglected by taking the *phase quenched* theory, i.e. substituting $\det M(\mu)$ with its absolute value $|\det M(\mu)|$ in the partition function.

In practise, the QCD phase diagram in Fig.1 can only be explored, with standard Monte Carlo methods, close to the y axis, i.e. where the ratio (μ/T) is small. In that region, one can still extrapolate information by Taylor expansion in (μ/T) , reweighting methods or by analytic continuation from imaginary μ where the theory is real [7, 34].

Recently some progress has been made in exploring the QCD phase diagram, deep into the oscillating region, thanks to *complex Langevin dynamics* [35, 36].

One of the other major open problem in quantum chromodynamics is the *strong CP* problem. The theory allows the pure gauge action to have an extra gluonic term

$$S = S_{YM} + i\theta \frac{g^2}{32\pi^2} \int dx \text{Tr} \left(G_{\mu\nu} \tilde{G}^{\mu\nu} \right), \quad (4)$$

which is proportional to the topological charge $Q_{top} = \frac{1}{32\pi^2} \int dx G_{\mu\nu}^a \tilde{G}_{\mu\nu}^a$. The fact that in nature experimental evidence constrain

$$\theta < 10^{-10}, \quad (5)$$

makes fine tuning problems arise. Some axion models for a dynamical solution to this problem have been proposed [37, 38, 39], but still a non-perturbative investigation of the theory at finite θ is required. Unfortunately, the topological θ -term is imaginary, that makes the action (4) complex. Again, the sign problem prevents standard Monte Carlo methods to explore the whole $\theta - T$ plane.

Results have been achieved by analytic continuation from imaginary θ [40, 41, 42], and recently the problem at real θ has been studied with complex Langevin dynamics [43, 44].

Many other systems affected by the sign problem have been studied, we will see some of those in the following, and some successfully solved. The main challenge, however, remains to solve the QCD related sign problem.

This thesis is divided in five chapters, and all the results shown come from my work in first person.

The first chapter reviews the discretization of a field theory on the lattice. In the second, we introduce the stochastic quantization and analyse the requirements for good control of complex Langevin dynamics. We also discuss the gauge cooling method and show results of convergence in some models.

The third chapter is about the θ -term. The first half, is a review of the instantons' theory and of the discretization of topology on the lattice. In the second part, we discuss the application of complex Langevin dynamics to this theory and show

some results. I would like to thank Dénes Sexty for providing the base code which I developed to get the results shown in this chapter, and also Ben Jäger and Felipe Attanasio for their help and discussions along the development of the same code.

In the fourth chapter, we introduce the Lefschetz thimbles method and compare it with complex Langevin dynamics in the study of some toy models.

The fifth, and last, chapter is about the density of states method. We introduce the method and discuss its application to the relativistic Bose gas at finite density. I would like to thank Roberto Pellegrini for his collaboration in developing the code we have been using to get the results shown in this chapter.

Chapter 1

Quantum field theory on the lattice

Making predictions for a non trivial Quantum Field Theory (QFT) is never an easy task. Furthermore, if it allows non perturbative interaction between the fields, analytical methods are usually not readily available. The most acknowledged way to overcome these problems is, up until now, to adopt a non-perturbative regularization of the theory on a discrete lattice of points in Euclidian space-time. Proposed by Wilson in 1974 [46], lattice QFT has become the most reliable tool for strongly interacting systems.

In this chapter we'll briefly review the main concepts behind it.

1.1 Path Integral approach to quantum theory

The path integral approach to QFT, introduced by Feynman in 1948 [45], is one of the most powerful tools when it comes to non-classical calculations. It is essential not only in order to perform most of the perturbative calculations but also to study non perturbative physics. Thanks to this formulation, in fact, a mapping of a regularized QFT into a statistical mechanic system has been made possible. More precisely, the quantum degrees of freedom on a *discretized* space-time can be identified as the ones in the canonical ensemble at temperature T and the system is statistically allowed to visit every configuration.

We will now recall the basic ideas behind the path integral and its connection with statistical mechanics when Euclidian time is introduced. Let us consider a non relativistic quantum system ((0+1) QFT) described by a Hamiltonian $\mathcal{H}(p, x)$. The matrix element relative to the evolution from the point $x_a(t_a)$ to $x_b(t_b)$ is given by

$$T_{x_a, x_b} = \langle x_a | e^{-i \frac{\mathcal{H}}{\hbar} (t_b - t_a)} | x_b \rangle ; \quad (1.1)$$

one can now insert somewhere in between t_a and t_b the operator identity $I = \int dx_c |x_c(t_c)\rangle\langle x_c(t_c)|$ that leaves, of course, unchanged the probability of transition from $|x_a\rangle$ to $|x_b\rangle$

$$T_{x_a, x_b} = \int_{x_c} dx_c \langle x_a | e^{-i\frac{\mathcal{H}}{\hbar}(t_c-t_a)} | x_c \rangle \langle x_c | e^{-i\frac{\mathcal{H}}{\hbar}(t_b-t_c)} | x_b \rangle . \quad (1.2)$$

One can think of repeating this process an infinite number of times by inserting the identity operator I at every time between t_a and t_b , integrating over all the possible values of the field $x(t)$ at that time t . Furthermore, if the Hamiltonian is quadratic in the momenta $\mathcal{H} = \frac{1}{2}p^2 + V(x)$, it is possible to carry out the Gaussian integral over the momenta p and express the eq.(1.1) in the path integral form form :

$$T_{x_a, x_b} \rightarrow C \int_{x_a}^{x_b} Dx(t) e^{\frac{i}{\hbar} \int_{t_a}^{t_b} dt \mathcal{L}} \quad (1.3)$$

where \mathcal{L} is the *Lagrangian* of the system

$$\mathcal{L}(\dot{x}, x) = \frac{1}{2} \dot{x}^2 + V(x) . \quad (1.4)$$

Physically, this procedure has the meaning of interpreting the probability of transition of a system from a state to another as *the sum over all possible intermediate paths* weighted with the oscillating phase given by their action along each one of these paths. As a result, one is not surprised to see the biggest contribution coming from the paths with stationary phase, called *semi-classical approximation*, which, for values of the parameters of the theory comparable with the ones in our everyday life, will be reduced into the well known classical equation of motion.

Let us go back now to the analogy with statistical mechanics. If we consider a canonical system described by a Hamiltonian \mathcal{H} and in thermal equilibrium with a heat bath at temperature T , we can write the partition function of the system as

$$Z(\beta) = \text{Tr}(e^{-\beta H}) = \sum_n \langle n | e^{-\beta H} | n \rangle = \sum_n e^{-\beta E_n} , \quad (1.5)$$

where $\beta = 1/k_B T$ and $|n\rangle$ are the eigenstates of the Hamiltonian with eigenvalues the values of the energy E_n . Moreover, since the trace of an operator does not depend on the basis on which it is computed, one can rewrite the trace in (1.5) choosing as the base the position $|x\rangle$

$$Z(\beta) = \text{Tr}(e^{-\beta H}) = \int_{-\infty}^{\infty} dx \langle x | e^{-\beta H} | x \rangle \quad (1.6)$$

where the similarity with (1.3) is now fairly evident. The last step we have to do in order to achieve the exact analogy of the two expressions is to rotate the eq.(1.3) into *Euclidian time (Wick rotation)*

$$\begin{cases} x_i \rightarrow x_i \\ t \rightarrow -i \tau , \end{cases} \quad (1.7)$$

to identify the inverse of the temperature β with $\frac{t_b - t_a}{\hbar}$ and to restrict ourselves only to periodic paths $x(t_a) = x(t_b) = x$. With these devices the eq.(1.3) becomes

$$Z(\beta) = \mathcal{N} \int_{x(0)}^{x(\beta\hbar)=x(0)} Dx(\tau) e^{-\int_0^{\beta\hbar} d\tau \mathcal{L}_E}. \quad (1.8)$$

The integral over the time of the euclidian lagrangian \mathcal{L}_E is referred to as the euclidian *action* of the theory

$$\mathcal{S}_E(\dot{x}, x) = \int_0^{\beta\hbar} d\tau \mathcal{L}_E(\dot{x}, x), \quad (1.9)$$

so that the most common way to express the partition function of a quantum mechanical system in euclidian time is

$$Z = \mathcal{N} \int Dx(\tau) e^{-S_E}. \quad (1.10)$$

Thanks to this formulation, it is now possible to assign to every quantum trajectory $x(\tau)$ a probability based on the action of the system along that path

$$P(x(\tau)) = \frac{e^{-S_E(x(\tau))}}{Z}. \quad (1.11)$$

When dealing with number of dimensions higher than 0, the above formulation of a QFT is often still affected by ultraviolet divergences and, therefore, it needs some regularization before physically relevant quantities can be computed. One possible regulator, largely used for non perturbative calculations, is the *lattice* discretization of the space time. This particular regularization has the property of mapping a d spatial and 1 time dimension quantum field theory in an equilibrium $(d+1)$ dimensions statistical mechanics model. Therefore, using the equivalent of (1.11), it is possible to evaluate the probability the system is likely to be in a certain configuration given the euclidian action of the configuration itself. More generally this map allows us to use all the well developed tools of statistical mechanics to study the QFT .

1.2 Monte Carlo simulations

Most of the physical relevant problems in Quantum Field theory are such that the fields in the theory are coupled to each other in a non trivial and non-perturbative

way. In a more formal way this is equivalent to saying that it is impossible to compute the path integral

$$Z = \int D\phi_i \mathcal{O}(\phi_i) e^{-S(g_i; \partial\phi_j, \phi_k)} \quad (1.12)$$

exactly or perturbatively because of the interactions in the action $S(g_i; \partial\phi_j, \phi_k)$; here we called the fields ϕ_i , their derivatives $\partial\phi_j$, and the couplings of the theory g_i ; also from now on we will always work in Euclidean space so we will omit the label E at the bottom of the operators, in this case for example $S = S_E$.

The only way to compute (1.12) is then the numerical approach. There is more than one class of numerical methods that are, in principle, able to compute the correct estimation of (1.12) but, if we assume for now that S is *real*, the most popular and well developed are for sure the Monte Carlo methods. To be precise, what a Monte Carlo algorithm is able to compute is not exactly (1.12) but actually the *average value* of the observable \mathcal{O}

$$\langle \mathcal{O} \rangle = \frac{\int D\phi_i \mathcal{O}(\phi_i) e^{-S(g_i; \partial\phi_j, \phi_k)}}{\int D\phi_i e^{-S(g_i; \partial\phi_j, \phi_k)}}. \quad (1.13)$$

The idea at the base of Monte Carlo methods is to *sample* the configuration's space of the fields ϕ_i guided by the probability $e^{-S(\phi_i)}$, in such a way that the configurations with smaller action are visited more often and vice versa. This procedure assigns a weight to the configurations, based on the frequency $f_{\{\phi_i\}}$ they have been visited, and allows one to build a good sample of the configuration space that can be used to compute any observable

$$\langle \mathcal{O} \rangle \sim \sum_{\{\phi_i\}} f_{\{\phi_i\}} \mathcal{O}(\{\phi_i\}). \quad (1.14)$$

Of course the bigger the number of configurations sampled, the more accurate is the estimation of $\langle \mathcal{O} \rangle$; more precisely, the error scales like the inverse of the square root of the number N of measurements (as long as these are independent from each other)

$$\delta \langle \mathcal{O} \rangle \sim \frac{1}{\sqrt{N}}. \quad (1.15)$$

It is crucial to observe how important is the hypothesis that the action $S(g_i; \partial\phi_j, \phi_k)$ is *real*, because if this is no longer true one cannot interpret e^{-S} as a weight to sample the configuration space and the very premises of the Monte Carlo methods fail.

Lastly, we observe that in general the fields of a theory live in a space-time continuum that, in practice, is impossible to be represented compatibly with the application of any numerical methods. The solution to this problem is to adopt a

discretized space-time commonly called *Lattice*, where each continuous direction is replaced with a discrete multiple of a unit of length a called *lattice spacing*

$$\begin{aligned} x &= n_x a_x, \quad n_x = 0, 1, \dots, N_x \\ y &= n_y a_y, \quad n_y = 0, 1, \dots, N_y \\ &\vdots \end{aligned} \tag{1.16}$$

and the boundary of this lattice are periodic in each direction (periodic boundary conditions) so that

$$\phi(n_i = N_i) = \phi(n_i = 0). \tag{1.17}$$

The original volume is now replaced by a lattice of points evenly spaced in each direction but, of course, to recover the real value of $\langle \mathcal{O} \rangle$ one has to correctly estimate the *continuum limit* which we're going to discuss later.

Discretization of the space-time has an important impact on the momentum space, in fact if the shortest wavelength can't be less than the lattice spacing $x_{min} = a$, it means the large momenta can't be bigger than $p_{max} = \pi/a$. It follows that the lattice introduces a *cut-off in the momentum space*

$$\int_{-\infty}^{\infty} \frac{dk}{2\pi} \tilde{f}(k) e^{ikx} \rightarrow \int_{-\frac{\pi}{a}}^{\frac{\pi}{a}} \frac{dk}{2\pi} \tilde{f}(k) e^{ikna} \tag{1.18}$$

restricting all the integrals on the momenta to the first Brillouin zone $[-\pi/a, \pi/a]$. Hence, all the loop integrals are finite and the functional integrals are high dimensional standard integrals. Eventually the cut-off has to be removed while approaching the continuum limit and this process determines the *renormalization group flow* of the theory of the observables. From the point of view of the statistical mechanics, the lattice structure does not play a crucial role if all the physical length scales are much larger than the the lattice spacing, which means

$$\xi \gg a \tag{1.19}$$

where ξ is the correlation length of the system. In this regime the physical quantities are not sensitive to changes in lattice spacing and therefore one can safely claim the lattice system to be in the continuum limit. Such a regime is reached in correspondence of a *second order phase transition* of the theory.

1.3 An example: the Scalar Field

In this section we're going to briefly review the case of the scalar field both because it is a quite popular and instructive example of discretization of a field theory on the Lattice and because we'll come across it again describing the *density of state* method. Also from now on we will work in the Planck units

$$\hbar = k_B = c = 1 , \quad (1.20)$$

which are the natural units of measurement to use in quantum field theory.

A possible scalar field theory, expressed in the new units of measurement, could be represented by the action

$$S = \int_0^\beta d\tau \int_V d^3x \left[\frac{1}{2} (\partial_\mu \varphi \partial^\mu \varphi + m^2 \varphi^2) + g_3 \varphi^3 + \lambda \varphi^4 + \dots \right] , \quad (1.21)$$

where φ is the scalar field and g_3, λ ,etc are the couplings. In general the action can be quite complicated, including the coupling with non perturbative global charges, however the purpose of this section is just to illustrate the process of discretization of a field theory; therefore we will limit ourself to the case of the free theory

$$S = \int_0^\beta d\tau \int_V d^3x \frac{1}{2} (\partial_\mu \varphi \partial^\mu \varphi + m^2 \varphi^2) . \quad (1.22)$$

As we mentioned in the last section the analogy with statistical mechanics requires the adoption of *periodic boundary condition* in the time direction; as long as bosonic degrees of freedom are involved, that is formally equivalent to imposing

$$\varphi(\vec{x}, \tau) = \varphi(\vec{x}, \tau + \beta) \quad (1.23)$$

on our scalar field .

The next step is to discretise the space-time volume

$$x_\mu \rightarrow a n_\mu , \quad (1.24)$$

where a has the dimensions of a length or, alternatively, of the inverse of an energy and n_μ is the integer that indicates the distance, in multiples of a , in the μ direction. Since our volume has to be finite, for practical reasons, there is going to be an N_μ such that

$$aN_\mu = L_\mu , \quad (1.25)$$

where L_μ is the edge of our volume in the μ direction, so that we have the condition

$$0 < n_\mu \leq N_\mu . \quad (1.26)$$

The most common way (but not the only one) to cope with the finite volume is to impose periodic boundary conditions also in the space directions, so that

$$\varphi(n_k + N_k) = \varphi(n_k) , \quad k = 1, 2, 3..., \text{space directions} \quad (1.27)$$

and to take in account the finite size effects by studying the scaling of the observables for increasing lattice sizes, a process known as *thermodynamic limit*.

It is good practice, then, to rescale every quantity in units of a so that everything on the lattice is adimensional except the lattice spacing. The mass m has, of course, the dimensions of an energy, while for the fields it usually depends on the space-time dimensions the theory is defined on; in our case, like most of the times, we're working with the 4 dimensions so φ has the dimensions of an energy. The lattice variable will then be defined as

$$\begin{aligned} m &\rightarrow \frac{\hat{m}}{a}, \\ \varphi(x) &\rightarrow \frac{\phi(n)}{a}, \end{aligned} \tag{1.28}$$

while the integral over the continuous 4 dimensions will be replaced by a sum over the lattice sites times the fundamental cube

$$\int d^4x \rightarrow \sum_n a^4. \tag{1.29}$$

Also, for the derivative the discretization is quite straightforward

$$\partial_\mu \phi(x) \rightarrow \frac{\phi(n + \vec{\mu}) - \phi(n)}{a}, \tag{1.30}$$

recalling the original value in the limit $a \rightarrow 0$; the notation $\varphi(n + \vec{\mu})$ means one is considering the nearest neighbour of $\varphi(n)$ in the $\vec{\mu}$ direction. It is easy to see that the way we choose to discretize the derivative is not unique and there are at least 3 equivalent ways of defining a first order discretization of the derivative

- **forward** derivative

$$\partial_\mu \phi(x) \rightarrow \frac{\phi(n + \vec{\mu}) - \phi(n)}{a} \tag{1.31}$$

- **backward** derivative

$$\partial_\mu \phi(x) \rightarrow \frac{\phi(n) - \phi(n - \vec{\mu})}{a} \tag{1.32}$$

- **symmetric** derivative

$$\partial_\mu \phi(x) \rightarrow \frac{\phi(n + \vec{\mu}) - \phi(n - \vec{\mu})}{2a}. \tag{1.33}$$

Let us note that, out of the three, only the last one maintains, on the lattice, the anti-Hermiticity proper of its continuous version; the other two transform one into the other under the Hermitian conjugate operation. For the scalar field this is however

not a problem since the square of the derivative in the action makes everything Hermitian again anyway.

The last thing to be mentioned is that also the integral measure will, of course, be affected by the discretization of space-time. The integral over all possible paths will be replaced by the product of the differential of the fields on each point of the lattice

$$\int_V [\mathcal{D}\varphi(x)] \rightarrow \int_{n \in Latt} \left(\prod_n d\phi_n \right) . \quad (1.34)$$

Following step by step all the points discussed before (1.28-1.34), we can now write down the lattice version of the free theory (1.22)

$$\begin{aligned} S_L &= \frac{1}{2} \sum_n \left[(\hat{m}^2 + 2d) \phi^2(n) - \sum_{\mu=1}^4 \phi(n) \phi(n + \vec{\mu}) \right] , \\ &= \frac{1}{2} \phi_n M_{nl} \phi_l , \end{aligned} \quad (1.35)$$

where d are the dimensions ($d = 4$ in our case) and

$$M_{nm} = (\hat{m}^2 + 2d) \delta_{nm} - \frac{1}{2} \sum_{\mu=1}^4 (\delta_{n+\vec{\mu},m} + \delta_{m,n-\vec{\mu}}) , \quad (1.36)$$

having expressed the $\sum_{\mu=1}^4 \phi(n) \phi(n + \vec{\mu})$ in (1.35) in a symmetric way. The 2-point function

$$\langle \phi_n \phi_m \rangle = M_{nm}^{-1} , \quad (1.37)$$

is easily computed using

$$\sum_l M_{nl} M_{lm}^{-1} = \delta_{nm} \quad (1.38)$$

and the Fourier transform of M_{nm}

$$\tilde{M}(k) = \int_{-\pi}^{\pi} \frac{d^4 k}{(2\pi)^4} M_{nm} e^{ik(n-m)} = \hat{m}^2 + 4 \sum_{\mu=1}^4 \sin^2\left(\frac{k_\mu}{2}\right) , \quad (1.39)$$

where, as always on the lattice, we are using the adimensional momentum $k_\mu = p_\mu a$. If we now take the Fourier transform of (1.38), using (1.39), we're able to obtain an expression for

$$M_{nm}^{-1} = \int_{-\pi}^{\pi} \frac{d^4 k}{(2\pi)^4} \frac{e^{ik(n-m)}}{\hat{m}^2 + 4 \sum_{\mu} \sin^2\left(\frac{k_\mu}{2}\right)} . \quad (1.40)$$

We're now interested in taking the limit of the lattice 2-point function M_{nm}^{-1} for $a \rightarrow 0$, holding $m, \varphi, x = na$ and $y = ma$ fixed. In general, the mass m could not be held fixed because it would take cut-off dependent contributions from the self energy and, to perform the continuum limit, one would have to take in account its renormalization. For our free theory, however, that is not the case and one can naively send $a \rightarrow 0$ in

$$M^{-1}(x - y) = \int_{-\frac{\pi}{a}}^{\frac{\pi}{a}} \frac{d^4 p}{(2\pi)^4} \frac{e^{ip(x-y)}}{m^2 + 4 \sum_\mu \frac{\sin^2(p_\mu a/2)}{a^2}}. \quad (1.41)$$

We just need to see, sending $a \rightarrow 0$, that the quantity $\frac{\sin^2(p_\mu a/2)}{a^2}$ remains finite only for $\sin^2(p_\mu a/2) \rightarrow (p_\mu a/2)^2$, and we recover for the (1.41) the well-known expression for the theory in the continuum

$$\langle \phi(x)\phi(y) \rangle = \int_{-\infty}^{\infty} \frac{d^4 p}{(2\pi)^4} \frac{e^{ip(x-y)}}{m^2 + p^2}. \quad (1.42)$$

1.4 Gauge Theories on the Lattice

In this section we are going to briefly review the regularization on the lattice of a widely studied category of physical theories, namely the *Gauge Theories*.

Let's consider a N dimensional matter field ϕ defined on the sites of our lattice, for example a set of N equivalent scalar fields of the section before, that transforms under local unitary group $\mathcal{G} \in \text{SU}(N)$ as

$$\phi(x) \rightarrow \phi'(x) = \mathcal{G}(x)\phi(x); \quad (1.43)$$

we now require its action

$$S_L = \frac{1}{2} \sum_n [(\phi(n+\mu) - \phi(n))(\phi(n+\mu) - \phi(n))^\dagger + m^2 \phi \phi^\dagger] \quad (1.44)$$

to be left invariant by such transformation. We clearly see that the part

$$S_{deriv} = \sum_{\mu=-4}^4 \phi^\dagger(n)\phi(n+\overrightarrow{\mu}), \quad (1.45)$$

coming from the derivative, cannot be invariant under the action of \mathcal{G} at least in a trivial way. The way to make up for it, in analogy with the continuum case, is to introduce a field $U_\mu(x)$, namely the *gauge* field. Its function is to *parallel transport*

the action of \mathcal{G} from $\phi(n + \vec{\mu})$ to $\phi(n)$, along the geodesic connecting the two fields. The latter, on the lattice, is the straight line connecting the site n with the one $n + \vec{\mu}$. The gauge field transforms

$$U_\mu(x) \rightarrow U'_\mu(x) = \mathcal{G}(x)U_\mu(x)\mathcal{G}^{-1}(x + \mu) , \quad (1.46)$$

so that the gauged part of action

$$S_{deriv}^g = \sum_{\mu=-4}^4 \phi^\dagger(n)U_\mu(x)\phi(n + \vec{\mu}) , \quad (1.47)$$

has now been made invariant under the action of \mathcal{G} . It is clear from the (1.46) that $U(x)$ is itself an element of $SU(N)$ and can therefore be written in the form

$$U_\mu(n) = e^{igaA_\mu(n)} \quad (1.48)$$

where g is the coupling constant of the gauge field, a is the lattice spacing and $A_\mu = A_\mu^a \lambda^a$ is an element of the Lie algebra of $SU(N)$. The generators of the algebra λ_a obey the commutation relations

$$\begin{aligned} \text{Tr}(\lambda_a \lambda_b) &= \frac{1}{2}\delta_{ab} , \\ [\lambda_a, \lambda_b] &= i \sum_{c=1}^{N^2-1} f_{abc} \lambda_c , \\ \{\lambda_a, \lambda_b\} &= \frac{1}{N}\delta_{ab} + \frac{1}{2} \sum_{c=1}^{N^2-1} d_{abc} \lambda_c , \end{aligned} \quad (1.49)$$

where f_{abc} are completely antisymmetric tensors, called the *structure constants* of the group, while d_{abc} are completely symmetric tensors.

The way we introduced the $SU(N)$ gauge fields follows what has been originally the idea from Yang and Mills, i.e. they started from a fermionic matter field to point out that the gauge field has to be provided with a dynamics and, therefore a lagrangian of its own. That means the gauge field has the right be considered by itself regardless of the interaction with any other field. Whether or not we consider the continuum theory or its discretization on the lattice, the new gauge action has, of course, to satisfy our initial requirement of being gauge invariant. The simplest gauge invariant object that can be constructed on the lattice with just gauge fields is the *plaquette*

$$\Pi_{\mu\nu} = U_\mu(n)U_\nu(n + \vec{\mu})U_\mu^\dagger(n + \vec{\nu})U_\nu^\dagger(n) , \quad (1.50)$$

which, because of (1.46), corresponds to the smallest possible closed loop. It is indeed the plaquette that is the basic element of the Wilson formulation of the lattice pure gauge action

$$S_W = \beta_G \sum_n \sum_{\mu\nu} \left(1 - \frac{1}{2N} [\text{Tr}(\Pi_{\mu\nu}) + \text{Tr}(\Pi_{\mu\nu})^\dagger] \right) , \quad (1.51)$$

where

$$\beta_G = \frac{2N}{g^2}. \quad (1.52)$$

Using the Baker-Hausdorff lemma and expanding the fields at the first order in a around the point n

$$A_\mu(n + \nu) = A_\mu(n) + a\partial_\nu A_\mu(n) + O(a^2), \quad (1.53)$$

it can be shown that the plaquette depends on the strength tensor $F_{\mu\nu}$ at the first order in a

$$\Pi_{\mu\nu} = e^{ia^2 g F_{\mu\nu} + O(a^4)}, \quad (1.54)$$

that correctly allows the Wilson action (1.51) to reduce to the well-known pure gauge action in the limit $a \rightarrow 0$

$$S_W \rightarrow \frac{1}{4} \int d^4x \text{Tr}(F_{\mu\nu}F^{\mu\nu}) + O(a^2). \quad (1.55)$$

The Wilson action (1.51) is just one of the available choices and any lattice action based on closed loops that succeeds in recovering the correct naive continuum limit (1.55) can, in principle, be used. Pure gauge actions based on more complicated closed loops are often used to achieve a weaker dependence on the lattice spacing, in order to obtain a better scaling of the observables in the continuum limit.

1.5 The continuum limit

Performing the continuum limit of a theory is a more subtle process than just consider small lattice spacings. As we mentioned before there is a very large number of lattice actions that correspond to the same continuum formulation for $a = 0$, however that is not sufficient to claim that the regularized theory processes the correct continuum limit. In particular, all the dimensional quantities which will be proportional to a non-zero power of a will go to zero or infinity.

As we mentioned in section 1.2, the essential requirement for the continuum limit to exist is that every correlation length $\hat{\xi}$, defined in lattice units, has to diverge compared to the lattice spacing. Hence, the continuum field theory can only be realized at a critical point $\{g_i^*\}$ in the parameter space of the discrete theory

$$\hat{\xi}(g_i) \xrightarrow{\{g_i\} \rightarrow \{g_i^*\}} \infty. \quad (1.56)$$

For the Wilson regularization (1.51) of an $SU(N)$ pure gauge theory, the only parameter is the bare coupling g so that, in this case, the renormalization group is determined by just one equation. One can find it by requiring that any physical quantity $\mathcal{O}(a, g)$ should not depend on the regulator of the lattice spacing

$$\left[a \frac{\partial}{\partial a} - \beta(g) \frac{\partial}{\partial g} \right] \mathcal{O}(a, g) = 0 , \quad (1.57)$$

where

$$\beta(g) = -a \frac{\partial g}{\partial a} \quad (1.58)$$

is called *β -function*. If one was able to know $\beta(g)$ then it would be possible to integrate (1.58) to obtain $g(a)$ and basically know the renormalization group of the theory for every value of a . Of course $\beta(g)$ cannot be known exactly, but it can still be calculated in perturbation theory around the critical point

$$\beta(g^*) = 0 . \quad (1.59)$$

In the proximity of the fixed point we can expand the *β -function* in powers of $(g - g^*)$

$$a \frac{\partial g}{\partial a} = \beta_1(g - g^*)^3 + \beta_2(g - g^*)^5 + \dots . \quad (1.60)$$

For $SU(N)$, the first non zero term that can be compute in perturbation theory is $\beta_1 = \frac{11N}{48\pi}$. Furthermore, we know from asymptotic freedom that $g^* = 0$. We are now able to predict, at least at the first order, the dependence of the lattice spacing from the coupling

$$a = \frac{1}{\Lambda_L} e^{-\frac{1}{2\beta_0 g^2}} , \quad (1.61)$$

where Λ_L is an integration constant with the dimensions of a mass and, in general, depends on the renormalization scheme that has been adopted.

The standard way to approach the continuum limit is to consider the ratio of dimensionful quantities of the theory. Let us suppose we know a particular mass m_0 from the experiments

$$m_0 = \frac{1}{a(g)\xi(g)} , \quad (1.62)$$

where $\xi_{phy} = a(g)\xi(g)$ is the physical correlation length corresponding to the inverse of the mass m_0 ; we can use this information to set the scale of the theory on the lattice so that every other mass m_i can be computed in relation to it

$$\frac{m_i}{m_0} = \lim_{g \rightarrow g^*} \frac{\xi_0(g)}{\xi_i(g)} . \quad (1.63)$$

In practice one can consider to have reached the continuum limit when the above ratio (1.63) holds constant within the statistical error of the measurement.

Chapter 2

Complex Langevin dynamics

In this chapter we are going to review the main concepts behind complex Langevin dynamics along with its recent developments and successes. The main advantage of this method, compared to standard Monte Carlo ones, is that it does not rely on the action to assign a weight to the field configurations. This approach, as we shall see, is therefore not affected by the *sign problem* at all when it comes to numerical simulations.

2.1 Stochastic dynamics

The path integral approach described in the previous chapter, however successful, is just one way to quantize a QFT. In general, other choices are possible and, following Damgaard and Huffel [17], we are going to review one of the most robust alternatives, namely *stochastic quantization*. The idea was first introduced by Parisi and Wu [47] in 1980 and consists of considering the Euclidian QFT as an equilibrium limit of a system governed by a stochastic process. The system evolves in an additional time t_L under the effect of some drift force, determined by the system, together with random noise. When the equilibrium is reached, for $t_L \rightarrow \infty$, stochastic averages become identical to ordinary Euclidean vacuum expectation values.

The oldest and best known stochastic equation, and the one we are interested in, is the *Langevin equation* [48]

$$m \frac{d}{dt} \mathbf{v}(t) = -\alpha \mathbf{v}(t) + \boldsymbol{\eta}(t), \quad (2.1)$$

introduced in 1908 to describe the Brownian motion of a particle of mass m in a fluid with viscosity α that randomly collides with other particles of the fluid with intensity and direction $\boldsymbol{\eta}$. The latter is represented by a gaussian distributed random

noise

$$\begin{aligned}\langle \eta_i(t) \rangle &= 0, \\ \langle \eta_i(t) \eta_j(t') \rangle &= 2\lambda \delta_{ij} \delta(t - t'),\end{aligned}\tag{2.2}$$

with mean 0 and variance 2λ . This old model, however simple and classical, is worth brief analysis to gain some insight into the Langevin stochastic process. The Langevin equation (2.1) is a non-homogeneous first order linear differential equation and therefore can be analytically solved as the Green's function

$$v_i(t) = \exp\left(-\frac{\alpha}{m}t\right)v_i(0) + \frac{1}{m} \int_0^t \exp\left(-\frac{\alpha}{m}(t-\tau)\right)\eta_i(\tau)d\tau.\tag{2.3}$$

We notice that the dependence on the initial conditions $v(0)$ is lost exponentially fast with time so that we might as well assume that $v(0) = 0$ without losing any generality. Having an equation for $v(t)$ we want to calculate some physical quantity from it, for example the average kinetic energy of our Brownian particle

$$\begin{aligned}\frac{1}{2}m\langle v^2(t) \rangle &= \frac{1}{2m} \int_0^t d\tau \int_0^t d\tau' \exp\left(-\frac{\alpha}{m}(2t-\tau-\tau')\right) \langle \eta_i(\tau)\eta_j(\tau') \rangle \\ &= \frac{3\lambda}{2\alpha} \left[1 - \exp\left(-\frac{2\alpha}{m}t\right)\right].\end{aligned}\tag{2.4}$$

We note that by taking $\lambda = kT\alpha$ the correct value for the average kinetic energy $E = \frac{3}{2}kT$ is recovered for $t \rightarrow \infty$.

Since eventually we are going to be interested in numerical simulations, it is essential for us to know the behaviour of the probability distribution governing the Langevin stochastic process. Let us set $\lambda = 1$, in this case, the observables will be functions of the velocity $v(t)$

$$\begin{aligned}\langle \mathcal{O}(v(t)) \rangle_\eta &\equiv \int D\eta \exp\left(-\frac{1}{4} \int_0^t \eta^2(\tau)d\tau\right) \mathcal{O}(v(t)) \\ &= \int dv \mathcal{O}(v) P(v, t) \equiv \langle \mathcal{O}(v(t)) \rangle_P,\end{aligned}\tag{2.5}$$

where we used the fact that averages can be computed either over the noise η or over the probability distribution $P(v)$,

$$\langle f \rangle_\eta = \langle f \rangle_P.\tag{2.6}$$

Taking the time derivative of (2.5) and using the Langevin equation (2.1) (with $m = \alpha = 1$), one gets

$$\langle \frac{\partial \mathcal{O}(v)}{\partial v} \frac{dv}{dt} \rangle = \langle \frac{\partial \mathcal{O}(v)}{\partial v} (-v + \eta) \rangle = \int dv \mathcal{O}(v) \frac{\partial P(v, t)}{\partial t}.\tag{2.7}$$

Furthermore, using eq. (2.5) and integrating by parts, we can write

$$\begin{aligned} \langle \frac{\partial \mathcal{O}(v)}{\partial v} \eta \rangle &= \int D\eta \left[\frac{\partial}{\partial \eta(t)} \exp \left(-\frac{1}{4} \int_0^t \eta^2(\tau) d\tau \right) \right] \frac{\partial \mathcal{O}(v)}{\partial v} = \\ 2 \langle \frac{\partial^2 \mathcal{O}(v)}{\partial v^2} \frac{\partial v}{\partial \eta} \rangle &= \langle \frac{\partial^2 \mathcal{O}(v)}{\partial v^2} \rangle, \end{aligned} \quad (2.8)$$

where in the last step we used the expression (2.3) for $v(t)$

$$\frac{\partial v(t)}{\partial \eta(t)} = \frac{\partial}{\partial \eta(t)} \int_0^\infty \theta(t-\tau) e^{-(t-\tau)} \eta_i(\tau) d\tau = \theta(0) = \frac{1}{2} \quad (2.9)$$

adopting the middle point prescription for the Heaviside step function $\theta(t-\tau)$. At this point, we can use (2.8) and, after integrating by part, we can rewrite (2.7) as

$$\int dv \mathcal{O}(v) \left[\frac{\partial}{\partial v} \left(v + \frac{\partial}{\partial v} \right) \right] P(v, t) = \int dv \mathcal{O}(v) \frac{\partial P(v, t)}{\partial t}. \quad (2.10)$$

This gives an equation for the evolution of the probability distribution $P(v, t)$ that goes under the name of *Fokker-Planck* equation

$$\frac{\partial P(v, t)}{\partial t} = \frac{\partial}{\partial v} \left(v + \frac{\partial}{\partial v} \right) P(v, t). \quad (2.11)$$

We notice that the stationary solution $\frac{\partial}{\partial t} P = 0$ of eq.(2.11), after requiring the condition $P(v) = P(-v)$, leads to the Boltzmann distribution for the Brownian particle in equilibrium with the system

$$P^{eq} \sim \exp \left(-\frac{v^2}{2} \right). \quad (2.12)$$

The Fokker-Planck equation, as we shall see, will play a crucial role in the numerical application of the Langevin dynamics in the case of a complex field theory.

2.2 Stochastic quantization of a field theory

The idea behind stochastic quantization is to formulate the equivalent of the Langevin equation (2.1) for a field theory in such a way that the associated Fokker-Planck distribution may have the Euclidian Boltzmann distribution $\exp(-S_E)$ as the unique stationary solution.

The first step is the introduction of an additional fictitious time t_L in which the stochastic systems evolves

$$\phi(x_0, \dots, x_n) \rightarrow \phi(x_0, \dots, x_n, t_L). \quad (2.13)$$

From now on we are going to call the Langevin time t_L just t , having in mind it is different from the Euclidian time x_0 .

The second requirement is that the evolution of fields be described by the Langevin stochastic equation

$$\frac{\partial}{\partial t}\phi(x, t) = -\frac{\delta S}{\delta\phi(x, t)} + \eta(x, t), \quad (2.14)$$

where S is the Euclidian action of the field theory, which also depends on the Langevin time ,

$$S = \int dt d^nx \mathcal{L} \left(\phi(x, t), \frac{\partial}{\partial t}\phi(x, t) \right), \quad (2.15)$$

and $\eta(x, t)$ is again Gaussian white noise

$$\begin{aligned} \langle \eta(x, t) \rangle &= 0, \\ \langle \eta(x, t)\eta(x', t') \rangle &= 2 \delta^n(x - x')\delta(t - t'). \end{aligned} \quad (2.16)$$

In the same way as in the classical case, equation (2.14) is associated with a probability distribution function $P(\phi, t)$ for the fields at the Langevin time t

$$\langle \phi(x_1, t)\dots\phi(x_n, t) \rangle_\eta = \int D\phi P(\phi, t) \phi(x_1)\dots\phi(x_n), \quad (2.17)$$

which satisfies the Fokker-Planck equation, corresponding to (2.11), generalized for the field theory

$$\frac{\partial P(\phi, t)}{\partial t} = \int d^n x \frac{\delta}{\delta\phi(x, t)} \left(\frac{\delta S}{\delta\phi(x, t)} + \frac{\delta}{\delta\phi(x, t)} \right) P(\phi, t). \quad (2.18)$$

As a final remark we would like to show that the former equation (2.18) leads $P(\phi, t)$ to converge to the Euclidian Boltzmann weight of the standard path integral quantization exponentially fast with Langevin time. Let us consider, for simplicity, one degree of freedom x . The partition function of this system reads

$$Z = \int dx e^{-S(x)} \quad (2.19)$$

and the corresponding Langevin equation is

$$\frac{dx}{dt} = -\partial_x S(x) + \eta. \quad (2.20)$$

The time evolution of the associated probability distribution function $P(x)$

$$\langle \mathcal{O}(x) \rangle = \int dx \mathcal{O}(x)P(x, t) \quad (2.21)$$

is determined by the Fokker-Planck equation

$$\partial_t P(x, t) = \partial_x [\partial_x + \partial_x S(x)] P(x, t), \quad (2.22)$$

whose stationary point is easily found to be $P(x) \sim e^{-S(x)}$. Moreover, we can rewrite eq.(2.22) upon the transformation

$$P(x, t) = \psi(x, t) e^{-\frac{1}{2}S(x)}, \quad (2.23)$$

into the Schrödinger-like equation

$$\dot{\psi}(x, t) = -H_{FP} \psi(x, t) \quad (2.24)$$

where

$$H_{FP} = \left(-\partial_x + \frac{1}{2}S'(x) \right) \left(\partial_x + \frac{1}{2}S'(x) \right). \quad (2.25)$$

The operator (2.25) is self-adjoint and, if $|\lim_{x \rightarrow \infty} S'(x)| \rightarrow \infty$, the spectrum of its eigenvalues is non negative and discrete

$$H\Psi_n = E_n\Psi_n, \quad (2.26)$$

with the ground state $\Psi_0 \sim e^{-\frac{1}{2}S(x)}$ annihilating eq. (2.25). Therefore we can rewrite $\psi(x, t)$ on the base of the eigenvectors of H_{FP}

$$\psi(x, t) = c_0 e^{-\frac{1}{2}S(x)} + \sum_n c_n \Psi_n(x) e^{-E_n t} \rightarrow c_0 e^{-\frac{1}{2}S(x)} \quad (2.27)$$

and the correct distribution $P(x) \sim e^{-S(x)}$ is reached exponentially fast.

2.3 Complex Langevin dynamics

We already discussed how a complex weight prevents the application of standard Monte Carlo methods. On the other hand, stochastic processes do not rely on importance sampling, which makes them good candidates to deal with the sign problem. In this section we shall see how Langevin dynamics can be generalized to the case of complex actions $S(x)$, examining in detail the careful steps that make this method successful.

A straightforward adaptation of (2.20) for a complex action is still possible and would lead to the Langevin equation

$$\dot{x} = -\partial_x S(x) + \eta \quad (2.28)$$

and, consequently, to the FP equation

$$\partial_t \rho(x, t) = L_0^T \rho(x, t), \quad (2.29)$$

where L_0^T is the usual Fokker-Planck operator $L_0^T = \partial_x(\partial_x + \partial_x S_c(x))$ that now is complex. Eq.(2.29) is expected to have the desired complex weight

$$\rho(x) \sim e^{-S(x)} \quad (2.30)$$

as a stationary solution. However, being complex-valued, $\rho(x)$ is not suitable to be regarded as a probability distribution function (PDF) as in (2.21). Furthermore, the associated FP Hamiltonian $H_{FP}(z)$, the complex equivalent of (2.25), is not self-adjoint any more, so that a proof of exponentially fast convergence to the unique solution cannot be provided. The way to proceed then [18, 49], is to consider the real and imaginary parts of the complexified variables $z \rightarrow x+iy$ as new and independent degrees of freedom

$$\begin{aligned} \dot{x} &= K_x + \sqrt{N_R} \eta_R \\ \dot{y} &= K_y + \sqrt{N_I} \eta_I , \end{aligned} \quad (2.31)$$

with the two drifts

$$K_x = -\Re e \partial_z S(z), \quad K_y = -\Im m \partial_z S(z) . \quad (2.32)$$

The correlators between the noises η_R and η_I derive from the original prescription (2.2) on the complex noise $\eta = \eta_R + i\eta_I$ and read

$$\begin{aligned} \langle \eta_R(t)\eta_R(t') \rangle &= 2N_R\delta(t-t') \\ \langle \eta_I(t)\eta_I(t') \rangle &= 2N_I\delta(t-t') \\ \langle \eta_R(t)\eta_I(t') \rangle &= 0 \end{aligned} \quad (2.33)$$

where $N_R - N_I = 1$ and $N_I \geq 0$.

The complexification of the Fokker-Planck equation

$$\dot{P}(z, t) = \partial_z(N_R\partial_z - K_z)P(z, t) \quad (2.34)$$

can be written, for *holomorphic* observables, in terms of the two independent variables $x(t)$ and $y(t)$ in (2.31)

$$\dot{P}(x, y, t) = [\partial_x(N_R\partial_x - K_x) + \partial_y(N_I\partial_y - K_y)]P(x, y, t) \quad (2.35)$$

and has the form of a continuity equation with the probability density $P(x, y; t)$ being the charge

$$\dot{P}(x, y, t) = \partial_x J_x + \partial_y J_y , \quad (2.36)$$

and

$$J_x = (N_R\partial_x - K_x)P , \quad J_y = (N_I\partial_y - K_y)P , \quad (2.37)$$

being the currents. Eq.(2.35) generates a *real* PDF for the holomorphic observables

$$\langle \mathcal{O} \rangle_{P(t)} = \int dx dy P(x, y, t) O(x + iy) , \quad (2.38)$$

which is, in fact, the main idea of complex Langevin (CL) dynamics, i.e. to reformulate a d -dimensional complex system into a $2d$ -dimensional real one. One requirement is to consider the holomorphic continuations of the observables $\langle \mathcal{O}(x) \rangle \rightarrow \langle \mathcal{O}(z) \rangle = \langle \mathcal{O}(x + iy) \rangle$. In this sense neither the quantity $\langle \mathcal{O}(x) \rangle$ nor $\langle \mathcal{O}(y) \rangle$ have, by themselves, any meaning in the complexified space.

The reason why complex Langevin dynamics was not largely employed immediately after its introduction in the 80's is that the equation (2.35), even keeping the same form of the real case, is much harder to solve or to be proved convergent to the appropriate stationary distribution $\rho(x) \sim e^{-S(x)}$ [50, 51, 52]. On top of that, unstable solutions of (2.32) can be found in the complex plane and that was believed to inevitably spoil the dynamics when solved numerically. These two problems have been more recently addressed, allowing CL to become one of the most acknowledged methods when it comes to system affected by the sign problem.

The issue of instabilities on the lattice was the first to be successfully and consistently solved [53]. The discretized CL equations for the field ϕ are

$$\begin{aligned} \phi_x^R(n+1) &= \phi_x^R(n) + \epsilon K_x^R + \sqrt{\epsilon N_R} \eta_x^R(n) \\ \phi_x^I(n+1) &= \phi_x^I(n) + \epsilon K_x^I + \sqrt{\epsilon N_I} \eta_x^I(n) , \end{aligned} \quad (2.39)$$

where ϵ is the discrete Langevin time step and x labels the sites of the lattice. When the system is brought near an unstable trajectory $\tau(\phi^R, \phi^I)$, the drifts $K^R(\tau)$ and $K^I(\tau)$ can potentially lead the fields to infinity, in a finite Langevin time $t_L = \epsilon n$. It turns out that careful integration in the form of adaptive stepsize ϵ_n along those trajectories is enough to completely remove the problem. The idea is to keep the product $\epsilon_n \mathcal{K}$ constant, where $\mathcal{K} = f(\sqrt{K_R^2 + K_I^2})$ is a function of the drift to be chosen optimally depending on the system, in order to greatly reduce the stepsize ϵ_n along the unstable trajectories and allow the real component of the random noise η^R to kick the system away from such trajectories. For this purpose, the imaginary component of the random noise η^I is, in general, counter-productive so that is usually preferable to get rid of it. This is in perfect accord with the prescriptions (2.33) and corresponds to the choice of parameters $N_I = 0$ and $N_R = 1$.

The problem of convergence of CL is much more complicated to address. Although no definitive solution has been found yet, fundamental progress has been made to fully understand this issue. In particular *proofs of convergence* have been found to infer, from the distribution of the observables, whether the CL is expected to converge to the right result or not [54, 18, 55, 56, 57]. In the following, we are going to review the main arguments behind these criteria.

We can rewrite eq. (2.35) in a more compact way

$$\frac{\partial}{\partial t}P(x, y; t) = L^T P(x, y; t) \quad (2.40)$$

where

$$L^T = \partial_x(N_R \partial_x - K_x) + \partial_y(N_I \partial_y - K_y) . \quad (2.41)$$

It is easy to see that the adjoint of the operator L^T determines the time evolution of a function f along a solution $z(t) = x(t) + iy(t)$ of the CL equation (2.31). Let us consider the scalar product :

$$\langle P, f \rangle = \int f(x, y) P(x, y) dx dy \quad (2.42)$$

where the time evolution operator can be moved from the probability density $P(x, y; t)$ to the function f in the usual way

$$\langle L^T P, f \rangle = \langle P, Lf \rangle , \quad (2.43)$$

so that the time evolution of f is described by

$$\frac{d}{dt} \langle f(x(t), y(t)) \rangle = \langle L f(x(t), y(t)) \rangle , \quad (2.44)$$

where the brackets mean averages over the stochastic noise (2.33) and the Langevin operator L has the form

$$L = (N_R \partial_x + K_x) \partial_x + (N_I \partial_y + K_y) \partial_y . \quad (2.45)$$

Our purpose is to obtain some conditions under which the analytically continued holomorphic observables $\mathcal{O}(x + iy)$, distributed according to the real-valued PDF (2.35) $P(x, y; t)$, retain the same average values as if defined on the real manifold $\mathcal{O}(x)$, weighted with the complex function $\rho(x)$ in (2.29), i.e.

$$\begin{aligned} \langle \mathcal{O} \rangle_{P(t)} &= \frac{\int \mathcal{O}(x + iy) P(x, y; t) dx dy}{\int P(x, y; t) dx dy} , \\ \langle \mathcal{O} \rangle_{\rho(t)} &= \frac{\int \mathcal{O}(x) \rho(x; t) dx}{\int \rho(x; t) dx} . \end{aligned} \quad (2.46)$$

In other words, one would like

$$\langle \mathcal{O} \rangle_{P(t)} = \langle \mathcal{O} \rangle_{\rho(t)} \quad (2.47)$$

provided the two are identical at the start $\langle \mathcal{O} \rangle_{P(0)} = \langle \mathcal{O} \rangle_{\rho(0)}$, which is easily assured if

$$P(x, y; 0) = \rho(x; 0) \delta(y) . \quad (2.48)$$

In order to link the two expressions for the average values of \mathcal{O} in (2.47), we define for $0 \leq \tau \leq t$ the function

$$F(t, \tau) = \int P(x, y; t - \tau) \mathcal{O}(x + iy; \tau) dx dy \quad (2.49)$$

which interpolates between the two definitions

$$F(t, 0) = \langle \mathcal{O} \rangle_{P(t)}, \quad F(t, t) = \langle \mathcal{O} \rangle_{\rho(t)}. \quad (2.50)$$

While the first equality is straightforward, the second can be seen using the prescription for the initial conditions (2.48) and integrating by part

$$\begin{aligned} F(t, t) &= \int P(x, y; 0) (e^{tL} \mathcal{O})(x + iy, 0) dx dy \\ &= \int \rho(x; 0) (e^{tL_0} \mathcal{O})(x, 0) dx \\ &= \int \mathcal{O}(x, 0) (e^{tL_0^T} \rho)(x, 0) dx \\ &= \int \mathcal{O}(x, 0) \rho(x; t) = \langle \mathcal{O} \rangle_{\rho(t)}. \end{aligned} \quad (2.51)$$

We only had to assume that no boundary contribution is introduced by the integration by parts, which is a very standard requirement on $\rho(x)$ without which it cannot possibly be integrated anyway. One could now obtain the (2.47) on the condition that the interpolating function $F(t, \tau)$ is independent on τ .

$$\begin{aligned} \frac{\partial}{\partial \tau} F(t, \tau) &= - \int (L^T P(x, y; t - \tau)) \mathcal{O}(x + iy; \tau) dx dy \\ &\quad + \int P(x, y; t - \tau) (L \mathcal{O})(x + iy; \tau) dx dy \end{aligned} \quad (2.52)$$

It is evident, after integration by part, that the two pieces of the former equation are equal and opposite except for a possible *boundary contribution coming from infinity in the complex plane*.

The one above is the core argument for the *criteria of correctness*, that can be summarized like this :

the formal argument for which complex Langevin is expected to converge to the right result, might fail when the distribution of the observables

$$P(x, y; t) \mathcal{O}(x + iy) \quad (2.53)$$

does not *decay fast enough in the complex plane*, giving rise to boundary terms. On the other hand, if the distribution (2.53) is localized enough, one expects CL to converge to the right result (detailed studies of those conditions can be found here [57, 56]). In general, more care is needed when the drift $\partial_z S(z)$ has poles, since the

hypothesis of holomorphicity drops. In those cases one has to study in detail how the dynamics is influenced by the poles and whether the results are correct or not will depend on the specific case [58, 59].

Let us conclude by saying that, when the criteria of correctness are satisfied, the stationary Fokker-Plank probability $P(x, y)$ should be positive-definite, since it is the distribution built up during the Langevin process. E.g. by binning the process one sees that in a bin with size $dxdy$, $P(x, y)$ is either 0 or > 0 . This assumes the relation between the Langevin process (2.31) and the FP distribution $P(x, y)$ holds, which it should when the stochastic process converges. Whether it is normalisable, depends on how the distribution goes to zero at large x and y . In numerical simulations, we always check the behaviour of $P(x, y)$ at large x and y , since it is also necessary for the criteria of correctness, and it always seems to be fast enough. However one might still argue that long tails might not be adequately sampled. In practice, we always find that if the complex Langevin converges, the stationary Fokker-Plank probability is normalisable.

2.4 Gauge Theories and Gauge Cooling

We will now start to describe the application of Langevin dynamics to gauge theories with sign problem. The most recent developments allowed the use of CL to investigate some of the most challenging and fundamental problems in the physics of strong interactions, such as finite density heavy quark QCD [19, 49, 44, 60], full QCD [61, 62, 63, 64, 35] and QCD in the presence of a topological θ -term related to the strong CP problem [44, 43]. We will review CL dynamics in the context of gauge theories, introduce the *gauge cooling* method and analyze in some detail how this helps to control the criteria of convergence.

For nonabelian $SU(N)$ gauge theories on the lattice, the Langevin equation for the link U reads

$$U_{\mu x}(t + \epsilon) = R_{\mu x}(t)U_{\mu x}(t), \quad R_{\mu x}(t) = \exp \left[-i \sum_a \lambda_a (\epsilon D_a S[U_{\mu x}] + \sqrt{\epsilon} \eta_{a\mu x}) \right], \quad (2.54)$$

where ϵ is the discrete Langevin time step used during the simulations, λ_a are the Gell-Mann matrices ($a = 1, \dots, N - 1$) and D_a is the $SU(N)$ Lie derivative

$$D_a f(U) = \frac{\partial}{\partial \omega} f(e^{i\omega \lambda_a} U) |_{\omega=0}. \quad (2.55)$$

As usual η_a is *real* Gaussian noise satisfying the relations

$$\begin{aligned}\langle \eta_{a\mu x} \rangle &= 0, \\ \langle \eta_{a\mu x}(t) \eta_{b\mu x}(t') \rangle &= 2\delta(t - t')\delta_{ab}.\end{aligned}\tag{2.56}$$

Looking at eq.(2.54), one immediately realizes that for complex actions $S[U]$ the operator $R(t)$ takes values into the group $SL(N, \mathbb{C})$. Consequently, even if at the beginning the links U were in $SU(N)$, the whole dynamic will be naturally enlarged in the bigger $SL(N, \mathbb{C})$. The element of this group retain the property $\det U = 1$ but lose the unitarity : $U^{-1} \neq U^\dagger$. To preserve analyticity then, every observable will have to be expressed in terms of U and U^{-1} , rather than U^\dagger . For example the correct analytical continued equivalent of the scalar quantity $\text{Tr}(UU^\dagger)$, equal to N in $SU(N)$, would be $\text{Tr}(UU^{-1})$ which, in fact, is still equal to N for every matrix $U \in SL(N, \mathbb{C})$. The operator $\text{Tr}(UU^\dagger)$ itself, however, can now take any real value bigger than N , as a consequence of the non compactness of $SL(N, \mathbb{C})$. In particular this can be seen since every element $U \in SL(N, \mathbb{C})$ can be written (polar decomposition) as

$$U = PV\tag{2.57}$$

with $V \in SU(N)$ and $P = P^\dagger$ a positive definite matrix with $\det P = 1$. This property of $\text{Tr}(UU^\dagger)$ allows us to define the *Unitarity Norm* (UN) d , a quantity that measures how deep in the $SL(N, \mathbb{C})$ manifold is the configuration

$$d = \frac{1}{N} \text{Tr} (UU^\dagger - \mathbb{1}) \geq 0,\tag{2.58}$$

where evidently $d=0$ for $SU(N)$. It is possible to define the UN in many other ways that still allows to quantify the distance to real manifold $SU(N)$, but the one in eq.(2.58) is the simplest and the one we are going to use in the following.

Gauge invariance is still present in $SL(N, \mathbb{C})$

$$U_\mu(x) \rightarrow U'_\mu(x) = \Omega(x)U_\mu(x)\Omega^{-1}(x + \mu), \quad \Omega \in SL(N, \mathbb{C}),\tag{2.59}$$

however the gauge transformation $\Omega(x)$ does not leave the UN (2.58) invariant, while its analytical continuation $\text{Tr}(UU^{-1})$ is preserved. Let us note that the parameters of the gauge group are now twice as many as in the case of $SU(N)$

$$\Omega = e^{i\omega_a \lambda_a}, \quad \omega_a \in \mathbb{C}\tag{2.60}$$

and, as we are about to learn, this abundance of gauge freedom is rather harmful for CL dynamics. It has been observed for many gauge models [65, 19, 49] that the CL dynamics, if left alone, tend to explore the huge gauge freedom available, bringing the system deep in the complex manifold. That, as one might expect, eventually leads to a wide distribution that violates the criteria of correction mentioned in the previous section.

The way to deal with this problem, first introduced in [19], takes the name of *gauge cooling* (gc.), and consists in gauge transforming all links of a given configuration up to the point where the UN (2.58) is minimal. One can achieve this by choosing as a parameter of the gauge transformation the gradient of the UN itself along the gauge orbit :

$$\Omega(x) = e^{-\epsilon \alpha_{gf} f_{ax} \lambda_a}, \quad f_{ax} = 2\text{Tr} \left[\lambda_a \sum_{\mu} \left(U_{x,\mu} U_{x,\mu}^\dagger - U_{x-\hat{\mu},\mu}^\dagger U_{x-\hat{\mu},\mu} \right) \right], \quad (2.61)$$

where ϵ is of the order of magnitude of the stepsize used in the Langevin process and α_{gf} is a parameter which can still be chosen to optimize the gauge cooling. One can convince himself that the gauge cooling reduces the UN by looking at the effect an infinitesimal gauge transformation (2.61) has on (2.58). For convenience we chose the gauge cooling transformation to be active only on the even sites while being the $\mathbb{1}$ on the odd ones

$$\begin{aligned} U_{\mu}(x) &\rightarrow U'_{\mu}(x) = \Omega(x)U_{\mu}(x), \\ U_{\mu}(x-\mu) &\rightarrow U'_{\mu}(x-\mu) = U_{\mu}(x-\mu)\Omega^{-1}(x). \end{aligned} \quad (2.62)$$

With this convention the change in UN along the gauge cooling trajectory can be easily calculated

$$d' - d = -\frac{\epsilon \alpha_{gf}}{N} (f_{ax})^2 + \mathcal{O}((\epsilon \alpha_{gf})^2). \quad (2.63)$$

Depending only on $(f_{ax})^2$, the right hand side of the equation above is always negative, up to higher orders in $\epsilon \alpha_{gf}$, resulting in a monotonic decrease of the UN itself. If the original configuration is gauge-equivalent to a $SU(N)$ one, cooling will eventually transform the configuration into the unitary one. Once the unitary manifold is reached, f_{ax} vanishes, as evident from (2.61), so that gauge cooling no longer has any effect. On the other hand, if the starting configuration is not gauge-equivalent to $SU(N)$, gauge cooling will bring it as close as possible, i.e. it will minimize d (Fig.2.1).

Although it is rather difficult to show that the minimum of the unitarity norm is unique, in practice, we observed many times that if we operate a $SL(3, \mathbb{C})$ random gauge transformation on a configuration and then we apply gauge cooling, the configuration is always brought back to the same minimum. That is a strong indication that the minimum is unique. However, even if the minimum wasn't unique, that wouldn't mine the usefulness of gauge cooling, which idea is to keep the dynamic close to the real manifold.

It is also important to stress that, since the gauge cooling process is separate from the CL evolution, it is not equivalent to gauge fixing term, i.e. it is not derived from a gauge fixing of the action, and, consequently, no Fadeev-Popov determinant is needed.

Let us remark that the complex Langevin dynamics is gauge invariant, in the sense that having an infinite precision machine the results with or without gauge

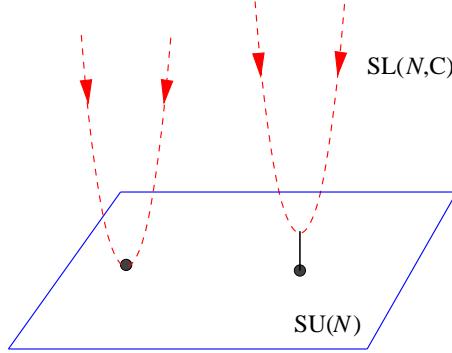


Figure 2.1: Gauge cooling in $\text{SL}(N, \mathbb{C})$ brings the link as close as possible to $\text{SU}(N)$. The orbit on the left is equivalent to a $\text{SU}(N)$ configuration, while the one on the right is not.

cooling would agree. The effect of gauge cooling is to minimizing the rounding errors by keeping the links as 'small' as possible with a gauge transformation.

One link model

Let us have a look at the gauge cooling process in a 1-link $\text{SL}(N, \mathbb{C})$ model [49], where analytical insight is possible. Since gauge cooling is completely disconnected from the CL dynamics itself, the form of the action is irrelevant to the following. After taking the continuous Langevin time limit ($\epsilon \rightarrow 0$), the eq.(2.63) becomes, to leading order in α_{gc} ,

$$\dot{d} = -\frac{\alpha_{gf}}{N} (f_a)^2 \quad (2.64)$$

with

$$f_a = 2\text{Tr} [\lambda_a (UU^\dagger - U^\dagger U)] , \quad (2.65)$$

where the μ and x indices can be dropped since there is only one link. Using the identity

$$\lambda_a^{ij} \lambda_a^{kl} = 2 \left(\delta_{il} \delta_{jk} - \frac{1}{N} \delta_{ij} \delta_{kl} \right) , \quad (2.66)$$

eq. (2.64) can be written as

$$\dot{d} = -\frac{16\alpha_{gf}}{N} \text{Tr} (UU^\dagger [U, U^\dagger]) . \quad (2.67)$$

In the case of $\text{SU}(2)$ this expression can be further simplified and it is written as

$$\dot{d} = -8\alpha_{gf} (d^2 + 2(1 - |c|^2)d + c^2 + c^{*2} - 2|c|^2) , \quad (2.68)$$

where $c = \frac{1}{2}\text{Tr } U$ and $c^* = \frac{1}{2}\text{Tr } U^\dagger$ (the case $c = 1$ refers to $U = \mathbb{1}$ which is trivial). When $c = c^*$, U is gauge-equivalent to an element of SU(2) and eq. (2.67) simplifies to

$$\dot{d} = -8\alpha_{gf} (d + 2(1 - c^2)) d. \quad (2.69)$$

This equation indeed has a *unique* fixed point at $d = 0$, which is reached exponentially fast,

$$d(t) \sim 2(1 - c^2)e^{-16\alpha_{gf}(1-c^2)t} \rightarrow 0. \quad (2.70)$$

On the other hand, if $c \neq c^*$, U cannot be gauge-equivalent to a SU(2) matrix and the stationary point is

$$d(t \rightarrow \infty) = |c|^2 - 1 + \sqrt{1 - c^2 - c^{*2} + |c|^4} > 0, \quad (2.71)$$

where the minimum distance is, as expected, larger than 0. This brief example and the analytical computations involved support what is shown in the sketch of Fig.2.1. When it comes to more complicated models, the effect of gauge cooling can only be computed numerically but the behaviour remains the same.

Adaptive gauge cooling for SU(2) Polyakov chain

Most of the cases of interest are too complicated to be studied analytically so that numerical computations are required. Here we shall discuss the case of a one-dimensional chain of N_l SU(2) gauge links (Polyakov chain), with the action

$$S = -\frac{\beta}{2}\text{Tr} (U_1 U_2 \dots U_{N_l}) , \quad \beta \in \mathbb{C} \quad (2.72)$$

and partition function

$$Z = \int \prod_{k=1}^{N_l} DU_k e^{-S[U]}. \quad (2.73)$$

One might notice that all the links except one could be transformed into the identity matrix $\mathbb{1}$ using the gauge transformation

$$U_n \rightarrow U'_n = U_n^{-1} U_n U_{n+1}. \quad (2.74)$$

The problem is, then, equivalent to the one-link model with $S = -\frac{\beta}{2}\text{Tr} (U)$, where the moments

$$\langle \text{Tr}(U^n) \rangle = \int DU e^{-S[U]} \text{Tr}(U^n) = \frac{I_n(\beta)}{I_0(\beta)} \quad (2.75)$$

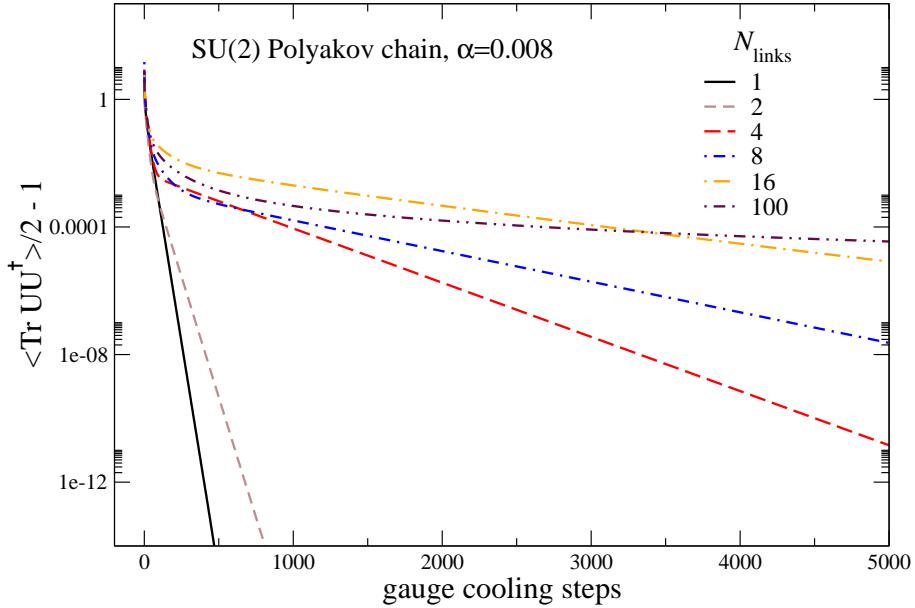


Figure 2.2: SU(2) Polyakov chain. Effect of gauge cooling on Polyakov chains with a different number of links.

are known in term of the modified Bessel functions of the first kind

$$I_n(\beta) = \frac{1}{\pi} \int_0^\pi d\theta e^{\beta \cos(\theta)} \cos(n\theta) . \quad (2.76)$$

This property is very useful in that it allows us to know the expectation value of the observable $\text{Tr}(U)$ and, at the same time, it does not lessen the interest in the original Polyakov chain model. In fact, from a numerical point of view, the full model in (2.72) still has as many degrees of freedom as the number of group parameters times the number of links, and can be used as a useful toy model to study the effect of gauge cooling in anticipation of proper four-dimensional gauge theories.

Before studying the actual dynamics of the system, we start by gauge transforming a SU(2) configuration into $\text{SL}(2, \mathbb{C})$ and applying gauge cooling to it until it is back in the original real manifold. Fig. 2.2 shows, on a logarithmic scale, the effect of gauge cooling on the distance d (2.58) increasing the number of links in the Polyakov chain. For $N_l = 1$ one finds numerical agreement with our analytical calculation of the one link model, i.e. the exponential fall off of the UN to the SU(2) fixed point $d = 0$. When $N_l > 1$ the exponential behaviour is slower and slower until it appears to become non-exponential for $N_l = 100$.

In gauge theories the number of degrees of freedom usually grows with the volume, that is why it would be preferable to find an implementation of gauge cooling that can keep the fall-off of the UN as fast as possible given the number of degrees of freedom involved. A possible solution is *adaptive gauge cooling* [49]. The idea is, given a transformation $\Omega_{gc}(x) = e^{-\epsilon \alpha_{gf} f_{ax} \lambda_a}$, to adapt the strength a_{gf} of gauge cooling depending on the proximity to the stationary point of eq.(2.63). In this

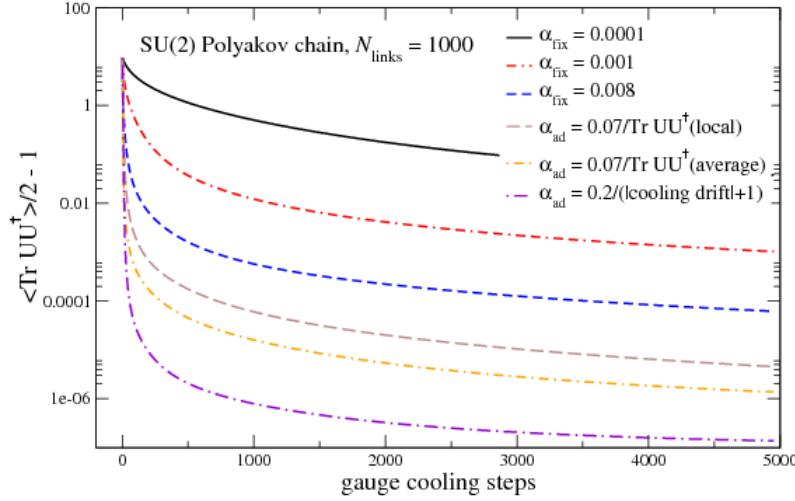


Figure 2.3: SU(2) Polyakov chain. Effect of different implementations of adaptive gauge cooling on a Polyakov chain of length $N_\ell = 1000$, plotted on a linear-log scale

way gauge cooling is expected to proceed with bigger steps when far away from the minimum of the UN, while being more sensitive when closer to it. We define

$$\alpha_{ad} = \frac{\alpha_{gf}}{D(U, U^\dagger)}, \quad (2.77)$$

where $D(U, U^\dagger)$ is a scalar function of the gauge field that can be adapted to the model under investigation. Clearly for $D(U, U^\dagger) = 1$ it is the case of fixed (non adaptive) gauge cooling.

In Fig. 2.3 and Fig. 2.4 we show the effect of different kinds of adaptive gauge

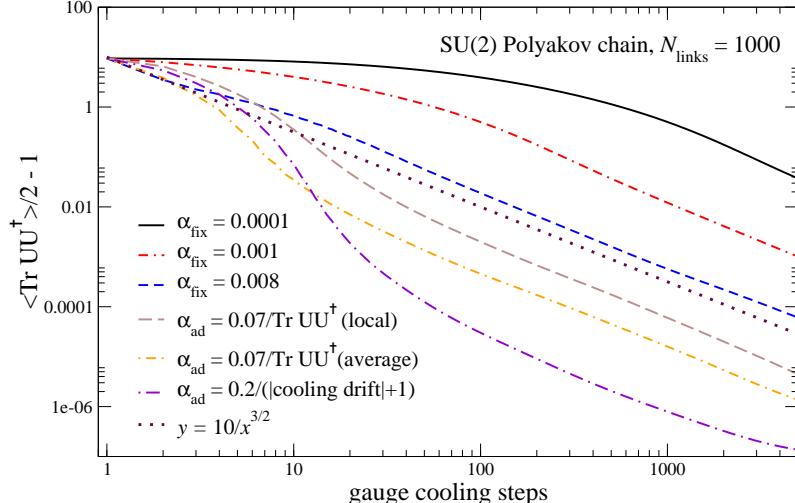


Figure 2.4: Same as FIG. 2.3 but plotted on a log-log scale .

cooling applied to the same Polyakov chain.

In particular we compare the cases

$$\begin{aligned}\alpha_{ad}(k) &= \alpha_{gf} , \\ \alpha_{ad}(k) &= \frac{\alpha_{gf}}{\frac{1}{N} \text{Tr} \left(U_k U_k^\dagger \right)} , \\ \alpha_{ad}(k) &= \frac{\alpha_{gf}}{\frac{1}{N} \langle \text{Tr} (U U^\dagger) \rangle} , \\ \alpha_{ad}(k) &= \frac{\alpha_{gf}}{\frac{1}{N_l} \sum_{a,k} |f_{ak}| + 1} .\end{aligned}\tag{2.78}$$

Near the minimum of the UN, all the above definitions of $\alpha_{ad}(k)$ tend to an effective fixed a_{gf}^{eff} ($\frac{1}{N} \text{Tr} \left(U_k U_k^\dagger \right) \rightarrow \mathbb{1}$ and $f_{ak} \rightarrow 0$) so that, close to the fixed point of the configuration, we still expect a sub-exponential decay of the kind seen in Fig. 2.2 for $N_l = 100$. On the other hand, adaptive gauge cooling will be effective at the beginning of the cooling process when α_{ad} is truly adaptive and strongly dependent on the UN. All these consideration can be verified looking at Fig. 2.3. Here we can see that, compared to the fixed gauge cooling the adaptive gauge cooling significantly helps the decrease of the UN in the first steps, while later on it stabilizes on an almost identical decay behaviour. This can be better seen looking at Fig. 2.4, where the same data are plotted on a log-log scale and a common asymptotic powerlike decay of the order $t^{\frac{3}{2}}$ (t being the number of gauge cooling steps) can be observed. In conclusion, we learned that with an adaptive implementation even just the first few cooling steps are enough to greatly reduce the distance from SU(2). In the remaining part of the chapter we are going employ this method in the actual dynamics of the system to help its convergence to the right result.

The complex Langevin dynamics simulations are carried out in such a way that between two consecutive Langevin update steps, a certain number of gauge cooling steps is applied to the links. We would like to stress again that, since the action (2.72) is complex-valued for $\beta \in \mathbb{C}$, CL dynamics is required for the full model even if the exact results are available thanks to the equivalence with the one-link model.

We will now present results for a chain with $N_l = 30$ links at $\beta = \frac{1}{2}(1 + i\sqrt{3})$. First of all, looking at Fig. 2.5, one is completely convinced of the role of gauge cooling, i.e. keeping the distance (2.58) from the configuration to the SU(2) manifold as small as possible during the whole CL evolution. Here the black line represents the UN as a function of langevin time during the CL evolution, without any gauge cooling being applied. One can then appreciate how the introduction of fixed gauge cooling (red line) has an important effect in reducing the UN of roughly two orders of magnitude. Even more, when adaptive gauge cooling is applied, the UN gets reduced by another order of magnitude. Of course the Polyakov chain, thanks to the complex action, will evolve in $SL(2, \mathbb{C})$ and will not be gauge equivalent to an SU(2) configuration, so that one expects the UN to be, in general, different than 0. Although the UN is a strong indication of the compactness of the distribution in the

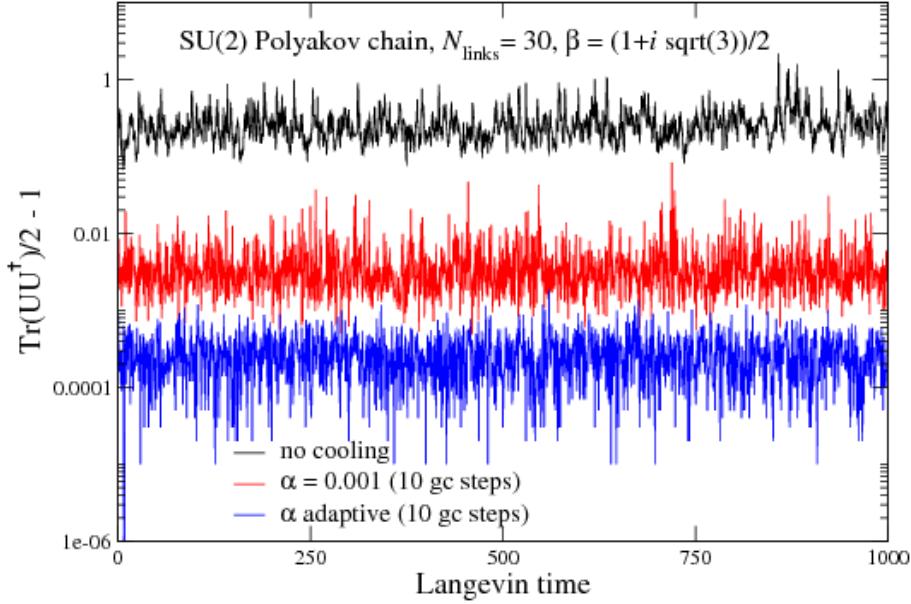


Figure 2.5: SU(2) Polyakov chain : comparison between the history of unitarity norm without gauge cooling, with fixed gauge cooling and with adaptive gauge cooling

complex $\text{SL}(2, \mathbb{C})$ manifold, the exact condition that has to be satisfied in order to satisfy the criteria of correctness, is about the distribution of the observables itself. Fig. 2.6 shows an histogram of the distribution of the action $S = \frac{\beta}{N_l} \text{Tr}(U_1 U_2 \dots U_{N_l})$ in the complex plane ($\Re(S)$ and $\Im(S)$). We can see how the effect of gauge cooling in keeping the UN as small as possible also directly affects the distribution of S . The black line (no gauge cooling applied) clearly shows that the distribution of S is quite widely spread and presents a sub-exponential decay in the complex direction. That corresponds to violation of the criteria of correctness and, as shown in Fig. 2.7 for number of gauge cooling steps = 0, to the convergence of CL to the wrong results. The application of fixed gauge cooling, 10 fixed gauge cooling steps every CL step, improves considerably the situation and makes the distribution much more localised although skirts are still present. Finally, with the adaptive choice the distribution is localised very well and drops quickly to zero outside its main support.

In Fig. 2.7 we show the value of $\langle S \rangle$ which CL dynamics converges to, as a function of the number of gauge cooling steps applied after every CL step. The black dotted lines are the correct analytical result $\frac{I_n(\beta)}{I_0(\beta)}$ for $\beta = \frac{1}{2}(1 + i\sqrt{3})$. Looking at the red dots (fixed gauge cooling), we can see how increasing the number of gauge cooling steps pushes CL dynamics towards the right result. Furthermore, using adaptive gauge cooling (blue dots), both $\Re\langle S \rangle$ and $\Im\langle S \rangle$ stabilize on the correct value after just 4 steps.

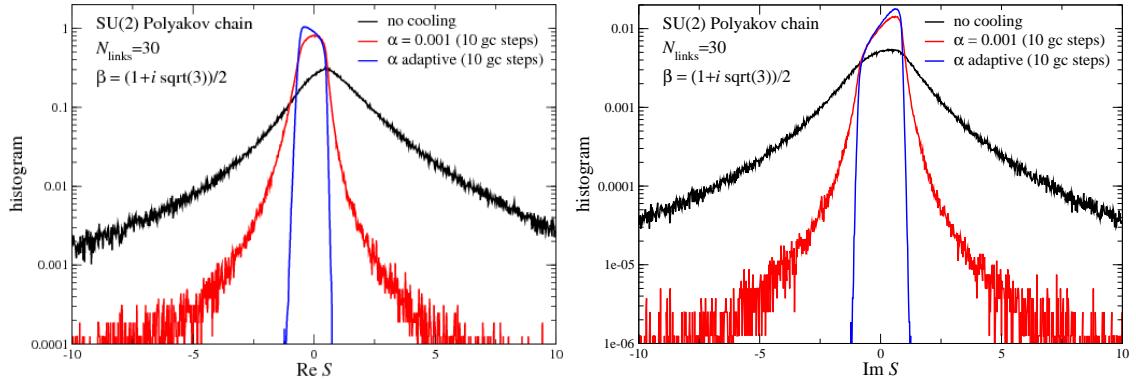


Figure 2.6: Histogram of the distribution of the real and imaginary part of $\langle S \rangle$

In this chapter, we showed how complex Langevin dynamics is a powerful method to deal with the sign problem. In particular, we discussed the existence of criteria of correctness, the meeting of which leads to the convergence of CL to the right results. Then, we introduced the concept of gauge cooling and employed it in the study of the Polyakov chain model. We explicitly showed how the effect of gauge cooling results in the satisfaction of the criteria of correctness and, consequently, leads the CL dynamics to the right results.

In the next chapter we are going to apply those concepts in the study of a more complicated and physically interesting gauge theory, i.e. Yang-Mills pure gauge with a CP violating topological θ term.

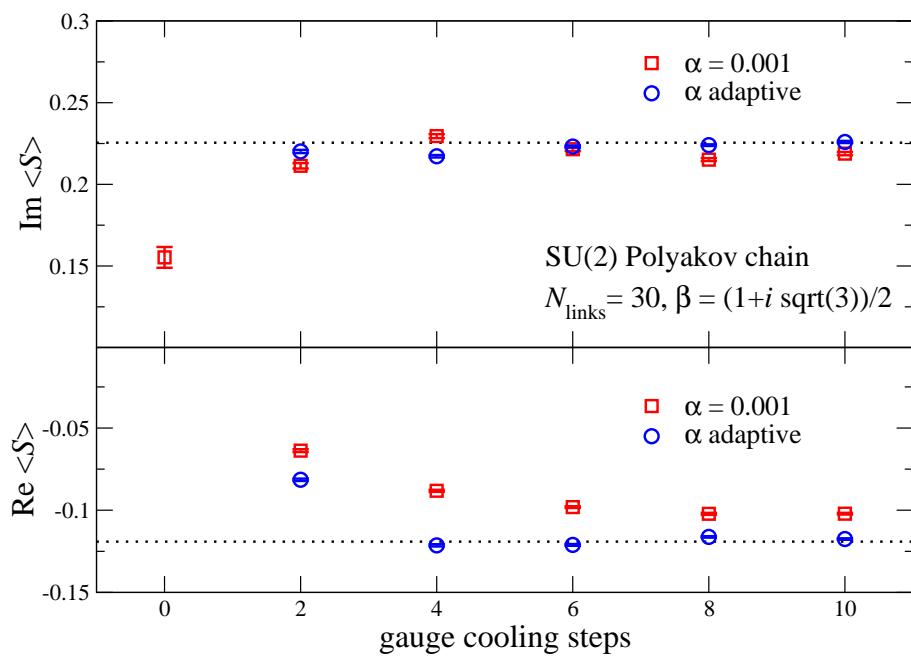


Figure 2.7: Convergence of CL to the correct result after the application of fixed and adaptive gauge cooling in the $SU(2)$ Polyakov chain

Chapter 3

Theta term

Topology in QCD started to get attention in the middle of the 70s with the famous paper by Polyakov al. [66]. In particular, they discovered classical solutions with non-trivial topology to the Yang-Mills equation, called the *instantons*. Those fields were found to represent, in Minkowsky space, tunneling events between degenerate classical vacua of the theory. Because proven to generate confinement in certain 3d models [67] instantons became immediately a very hot topic of research. However it was eventually realized that they could not be the cause in 4d Yang-Mills theories. Nevertheless, they were proven to be responsible for very important phenomena such as the anomalous breaking of the $U(1)_A$ symmetry [68] due to their interaction with the zero-modes of light quarks (index theorem). For the same reason they were found to be the cause of the large mass of the η' meson by the mechanism proposed by Witten [69] and Veneziano [70]. In the same years, the θ -vacua structure of QCD was introduced and systematically studied [71, 72, 67]. It was realized that a CP violating topological θ -term should be included in the QCD lagrangian; however, a calculation of the neutron magnetic dipole [73] constrains this coupling to being negligible ($\theta < 10^{-10}$). A definitive mechanism to explain why the θ -term has to be so small is still to be discovered and goes under the name of *strong CP problem*.

One of the most elegant solutions would be the introduction of the QCD *axion* particle. This particle would be a pseudo-Nambu-Goldstone boson which arises from the spontaneous breaking of a hidden chiral symmetry (Peccei-Quinn symmetry) of the Standard Model [37, 38]. Not only the broken symmetry would automatically suppress the CP violating part of the action, but the mass of the axion could be a candidate for the dark matter in the universe.

In this chapter we will review in some detail the instantons' picture and the structure of the QCD vacua (see [74, 75] for extensive treatment). Later on, we will discuss the study of topology on the lattice and present new results using complex Langevin dynamics.

3.1 Vacua of SU(N) and instantons

To describe the classical vacua of the pure gauge Yang Mills theory one has to find the zero energy field configurations. The gauge $SU(N)$ fields are defined in the usual way :

$$A_\mu = \sum_a A_\mu^a \frac{\lambda^a}{2}, \quad [\lambda^a, \lambda^b] = 2if^{abc}\lambda^c, \quad \text{Tr}(\lambda^a \lambda^b) = 2\delta^{ab}, \quad (3.1)$$

where $a = 1\dots N$ and λ^a are the Gell-Mann matrices. The minimum of the action

$$S = \frac{1}{4g^2} \int d^4x G_{\mu\nu}^a G_{\mu\nu}^a, \quad (3.2)$$

is given by those fields that satisfy the condition

$$G_{\mu\nu} = 0, \quad (3.3)$$

i.e. the pure gauge fields of the form

$$A_\mu = iU(x)\partial_\mu U^\dagger(x) \quad U \in SU(N). \quad (3.4)$$

A physical interpretation of their role can be given in the temporal gauge $A_0 = 0$, in which the fields of the vacuum have the time independent form

$$A_i(\mathbf{x}) = iU(\mathbf{x})\partial_i U^\dagger(\mathbf{x}), \quad (3.5)$$

where \mathbf{x} indicates only the spatial directions. In this gauge, it is possible to identify [76] that the gauge transformations that can be connected with the trivial topological sector, where $A_i(\mathbf{x}) = 0$, are only the ones that respect the condition

$$U(\mathbf{x}) \rightarrow 1 \text{ for } |\mathbf{x}| \rightarrow \infty. \quad (3.6)$$

This condition is equivalent, for what concerns $U(\mathbf{x})$, to the compactification of \mathbb{R}^3 in the S^3 hyperspherical surface and, therefore, $U(\mathbf{x})$ can be seen as an application that maps S^3 into $SU(N)$.

Those maps can be divided into disjointed *homotopy classes* and classified with an integer *winding* (or Pontryagin) number n_w

$$n_w = \frac{1}{24\pi^2} \int d^3x \epsilon^{ijk} \text{Tr} [(U^\dagger \partial_i U)(U^\dagger \partial_j U)(U^\dagger \partial_k U)], \quad (3.7)$$

which, in terms of the corresponding gauge fields, is the Chern-Simons number

$$n_{CS} = \frac{1}{16\pi^2} \int d^3x \epsilon_{ijk} \left(A_i^a \partial_j A_k^a + \frac{1}{3} f^{abc} A_i^a A_j^b A_k^c \right). \quad (3.8)$$

Let us remark that the winding number n_w (or n_{CS}) is not gauge invariant, as it is defined in the temporal gauge, and in three dimensions, only. However, the difference between two winding numbers is gauge invariant because, as we shall see, it can be related to a gauge invariant quantity, i.e. the *topological charge*.

The picture is then that Yang-Mills theories have an infinite number of distinct classical configurations with zero energy, each one can be related to an integer number n_w . Those vacua, however, are not completely isolated from one another. In fact, there are fields configurations that tunnel between two vacua with a different winding number. Those fields are classical solutions of the euclidean equations of motion. In particular, using the identity

$$S = \frac{1}{4g^2} \int d^4x G_{\mu\nu}^a G_{\mu\nu}^a = \frac{1}{4g^2} \int d^4x \left[\pm G_{\mu\nu}^a \tilde{G}_{\mu\nu}^a + \frac{1}{2} \left(G_{\mu\nu}^a \mp \tilde{G}_{\mu\nu}^a \right)^2 \right], \quad (3.9)$$

one can rewrite the action as a sum of the topological invariant $\int d^4x G_{\mu\nu}^a \tilde{G}_{\mu\nu}^a$, with a semi-positive definite part. The minimum, then, is the (anti)self-dual solution

$$G_{\mu\nu}^a = \pm \tilde{G}_{\mu\nu}^a. \quad (3.10)$$

In Euclidian space, one can prove that the (anti)self-dual solutions automatically satisfy the Yang-Mills classical equations of motion

$$D_\mu G_{\mu\nu} = 0. \quad (3.11)$$

These configurations are called *instantons* (*anti-instantons*) and, to have a finite contribution from the action, they need to be pure gauge at infinity

$$A_\mu(x) \xrightarrow{x \rightarrow \infty} iU(x)\partial_\mu U(x)^\dagger. \quad (3.12)$$

We notice that this generates again a maps from the three sphere S^3 into the gauge group $SU(N)$, which can be labelled by an integer number $Q \in \mathbb{Z}$ called *topological charge*. This can be better understood in the case of $SU(2)$ where the group itself is homeomorphic to the S^3 sphere. In this case eq.(3.12), or better its exponentialization, is an application that maps S^3 into S^3 and the corresponding homotopy group $\pi_3(S^3)$ is also isomorphic to the integers \mathbb{Z} , from which the winding number that represents how many times the first sphere wraps around the second one. For this reason two solutions with different topological charge Q , cannot be deformed one into the other with a continuous transformation without violating finiteness of the action. The same argument can be extended to every gauge group since [77] every continuous application $U : S^3 \rightarrow G$, with G Lie group, can be deformed continuously in an application $U : S^3 \rightarrow SU(2) \subset G$.

Differently from the case of the vacuum fields (3.5), this time the map arises naturally, not only in a specific gauge, which results in the topological charge Q being a gauge invariant quantity. This can also be seen from its formal definition

$$\begin{aligned} Q &= \int d^4x q(x), \\ q(x) &= \frac{1}{32\pi^2} \text{Tr} \left[G_{\mu\nu}^a \tilde{G}^{a\mu\nu} \right] \end{aligned} \quad (3.13)$$

where $\tilde{G}^{a\mu\nu}$ is the Hodge dual of $G^{a\mu\nu}$ defined as $\tilde{G}^{a\mu\nu} = \frac{1}{2}\epsilon_{\mu\nu\rho\sigma}G^{a\rho\sigma}$. One can show that the topological charge density $q(x)$ is a total derivative

$$\begin{aligned} q(x) &= \partial_\mu K^\mu , \\ K^\mu &= \frac{1}{32\pi^2}\epsilon_{\mu\alpha\beta\gamma} \left(A_\alpha^a \partial_\beta A_\gamma^a + \frac{1}{3}f^{abc}A_\alpha^a A_\beta^b A_k^\gamma \right). \end{aligned} \quad (3.14)$$

K^μ describes, at $|x| \rightarrow \infty$, the infinitesimal winding of the S^3 sphere on the $SU(N)$ gauge group, that makes the quantity

$$Q = \int_V d^4x \partial_\mu K^\mu = \int_\Sigma d\sigma_\mu K^\mu, \quad (3.15)$$

precisely the integral of the total derivative that accounts for the topological charge. Going back to the temporal gauge, we notice that the only component of K^μ that survives in (3.14) is K^0 , which is precisely the winding (or Chern-Simons) number defined before in (3.8). Then, the integral (3.15)

$$Q = \int_{-\infty}^{\infty} dt \frac{d}{dt} \int d^3x K_0 = n_{CS}(t = \infty) - n_{CS}(t = -\infty), \quad (3.16)$$

shows that field configurations with $Q \neq 0$ connect different topological vacua. From (3.9), the action of those configurations is

$$S_{inst} = \frac{8\pi^2|Q|}{g^2} \quad (3.17)$$

which generates a probability of tunneling

$$P_{tun.} \sim e^{-8\pi^2/g^2}, \quad (3.18)$$

where the coefficient of proportionality has to be evaluated using perturbation theory.

A classic example of a $SU(2)$ instanton with charge $Q = 1$ is the *BPST* (Belavin, Polyakov, Schwartz & Tyupkin) instanton [66]

$$A_\mu^a(x) = 2\eta_{a\mu\nu} \frac{(x - x_0)_\nu}{(x - x_0)^2 + \rho^2}, \quad (3.19)$$

where x_0 is the centre of the instanton, ρ is its size and $\eta_{a\mu\nu}$ is the 't Hooft symbol :

$$\eta_{a\mu\nu} = \begin{cases} \epsilon_{a\mu\nu} & \mu, \nu = 1, 2, 3, \\ \delta_{a\mu} & \nu = 4, \\ -\delta_{a\nu} & \mu = 4. \end{cases} \quad (3.20)$$

For the anti-instanton solution, with topological charge $Q = -1$, one just needs to replace $\eta_{a\mu\nu} \rightarrow \bar{\eta}_{a\mu\nu}$, which is the same tensor as (3.20) with just opposite signs in the temporal directions.

Let us remark that, despite the BPST instanton being a long range field ($A_\mu \sim 1/x$), its field strength

$$G_{\mu\nu}^a G^{a\mu\nu} = \frac{192\rho^4}{(x^2 + \rho^2)^4}, \quad (3.21)$$

is well localized ($G_{\mu\nu}^a \sim 1/x^4$) in space and time, from which the name 'instanton' derives.

Let us briefly recall the properties of the two fields :

- Fields of the vacuum :

$$A_\mu^{vacuum}(x) = iU(\vec{x})\partial_\mu U^\dagger(\vec{x}), \quad \forall x; \quad (3.22)$$

those are the solutions which annihilate the strength tensor

$$G_{\mu\nu}(A_\mu^{vacuum}) = 0, \quad (3.23)$$

and, therefore, have got a 0 action contribution.

- Instantons :

$$A_\mu^{inst}(x) \xrightarrow[x \rightarrow \infty]{} iU(\vec{x})\partial_\mu U^\dagger(\vec{x}) + O\left(\frac{1}{x}\right), \quad \lim : x \rightarrow \infty; \quad (3.24)$$

those are semi-classical, (anti)self-dual solutions of the equations of motion $G_{\mu\nu} = \pm \tilde{G}_{\mu\nu}$. They represent, in Minkowsky space, tunnelling events between two different vacua of QCD and they have a finite action contribution

$$S_{inst} = \frac{8\pi^2|Q|}{g^2}. \quad (3.25)$$

Since the infinite number of degenerate vacua are connected by some tunneling events, the correct way to express the ground state of the theory, in analogy with the Bloch's theorem, is a linear combination of all the separate topological vacua

$$|\theta\rangle = \sum_{Q=-\infty}^{\infty} e^{i\theta Q}|Q\rangle \quad (3.26)$$

where $|Q\rangle$ indicates the vacuum with topological number Q and $e^{i\theta Q}$ comes from the periodic structure of the vacua. From (3.26) one understands that, in principle, there are an infinite number of possible ground states corresponding to all the values $\theta \in [0, 2\pi)$. However, we shall see that any choice of θ isolates a sector of physical states completely disconnected from the others at different values of θ . If we consider the unitary gauge transformation R that generates a shift of one in the topological vacuum

$$R|Q\rangle = |Q+1\rangle, \quad (3.27)$$

then we realize, from (3.26), that the ground state $|\theta\rangle$ is an eigenstate of R

$$R|\theta\rangle = e^{-i\theta}|\theta\rangle . \quad (3.28)$$

Furthermore, any gauge invariant operator \mathcal{O} must commute with the gauge transformation R , i.e. $[\mathcal{O}, R] = 0$. Hence,

$$0 = \langle\theta| [\mathcal{O}, R] |\theta'\rangle = \left(e^{-i\theta'} - e^{i\theta}\right) \langle\theta|\mathcal{O}|\theta'\rangle , \quad (3.29)$$

that, when $\theta \neq \theta'$, implies

$$\langle\theta|\mathcal{O}|\theta'\rangle = 0 . \quad (3.30)$$

That means it is impossible for any physical (gauge invariant) operator to connect states with different θ . Furthermore, since $[\mathcal{H}, R] = 0$ also holds, the ground state $|\theta\rangle$ remains unchanged in time. In other words θ is a parameter of super-selection of the physical theory that, once selected, remains constant with no possibility of contact with states at different θ .

Let us note that the ground state is not CP invariant as the topological vacua $|Q\rangle$ are not ($\text{CP}(Q) = -Q$). One can explicitly see this from (3.26)

$$\text{CP}|\theta\rangle = \sum_{Q=-\infty}^{\infty} e^{-i\theta Q}|Q\rangle \neq |\theta\rangle , \quad (3.31)$$

where the strength of CP violation is characterized by the angle θ . However, experimental evidences on the neutron electric dipole moment (nEDM) limit the violation of CP in our world to a very small value. In particular, this can be translated in a constraint on the parameter θ [73]

$$\theta < 10^{-10} . \quad (3.32)$$

This unnaturally small value of θ , which is otherwise not restricted by theory, is known as the strong CP problem. One elegant solution was proposed by Peccei and Quinn [37], which makes the θ -parameter vanish dynamically. However, the mechanism also requires the appearance of a Goldstone boson, the axion, which remains to be discovered.

Axial anomaly and Atiyah-Singer index theorem

For historical reasons, we are now going to briefly mention how the non-trivial vacuum structure of QCD is essential to explain the $U_A(1)$ axial anomaly in the context of chiral symmetry breaking.

Let us start by defining an anomaly in the context of quantum field theory. One talks about anomaly when a group of transformation G leaves unchanged the action

of the theory $G(S) = S$, but not the measure of integration $G(d\mu) = d\mu'$ of the generating functional :

$$Z = \int d\mu e^{-S/\hbar} , \quad (3.33)$$

so that, overall, $G(Z) = Z' \neq Z$. As a consequence, the Noether current, classically associated to the symmetry of the action in respect of the group G , will not be conserved any more. When that happens, one says that a classical symmetry is *anomalously* broken at the quantum level. We see that in the limit $S \gg \hbar$ (classical limit) only the saddle point of the action S_{min} contributes and, therefore, the integration over the measure in the field space does not matter any more. In this limit the conservation of the Noether current is restored, as it should be.

QCD chiral symmetry is a physical example where an anomalous breaking of symmetry occurs. Let us recall the Lagrangian of N_f families of massless fermions:

$$\mathcal{L}_{massless} = \sum_{f=i}^{N_f} \bar{\psi}^f(x) i\partial \psi^f(x) = \sum_{f=i}^{N_f} \bar{q}_L^f i\partial q_L^f + \bar{q}_R^f i\partial q_R^f , \quad (3.34)$$

where we have explicitly written the left and right components of the spinor $\psi^f = \begin{pmatrix} q_R^f \\ q_L^f \end{pmatrix}$. The Lagrangian is invariant under chiral rotation, in the N_f fermions families, of both left and right components independently. The resulting symmetry is the group of transformations $U(N_f)_L \times U(N_f)_R$ called *chiral symmetry*.

In QCD all the quarks are massive so, in principle, there is no exact chiral symmetry. However, the three quarks (u, d, s) have masses much smaller than the chiral symmetry spontaneous breaking energy scale $\Lambda_\chi \sim 1\text{GeV}$. Therefore, they can be regarded as massless in first order approximation, in the sense that the spontaneous breaking of the chiral symmetry dominates over the explicit one. At first order in the quark masses, then, the chiral symmetry in QCD is

$$U(3)_L \otimes U(3)_R , \quad (3.35)$$

and can be decomposed into the irreducible representations :

$$U(3)_L \otimes U(3)_R \rightarrow SU(3)_L \otimes SU(3)_R \otimes U(1)_V \otimes U(1)_A . \quad (3.36)$$

The $U(1)_V$ symmetry corresponds to the baryon number conservation and the $U(1)_A$ is anomalously broken as we shall see in a bit. As anticipated before, however, chiral symmetry spontaneously breaks at temperatures of the order of 1GeV. In particular the group

$$SU(3)_L \otimes SU(3)_R , \quad (3.37)$$

is broken down to the vector subgroup

$$SU(3)_V , \quad (3.38)$$

which results in the classification of the hadrons into the SU(3) irreducible representations. The remaining 8 broken generators produce as many Goldstone bosons, namely the octet of pseudo-scalar mesons ($\pi^\pm, \pi^0, K^\pm, K^0, \eta$). Those mesons, despite being Goldstone modes, are massive as a result of the mass of the quarks that explicitly prevents chiral symmetry to be realized exactly.

The part which, historically, created most of the trouble was the $U(1)_A$ symmetry, i.e. the one associated with the transformation

$$q_L \rightarrow e^{i\theta\gamma_5} q_L \quad q_R \rightarrow e^{-i\theta\gamma_5} q_R . \quad (3.39)$$

That this symmetry was anomalous was first discovered by Adler [78], in the context of QED, and then extended by Jackiw and Bell for QCD [79]. The realization of such a symmetry would protect the neutral pion from the decay into two photons $\pi^0 \rightarrow \gamma\gamma$, when we know that, on the contrary, this amplitude is finite in QCD

$$\Gamma(\pi^0 \rightarrow \gamma\gamma) \approx 8.4 \text{ eV}. \quad (3.40)$$

Furthermore, it is also impossible for it to be realized in the Nambu-Goldstone way since, in this case, one expects the presence of another light Goldstone boson with a mass $m \leq \sqrt{3}m_\pi \approx 230$ MeV, which is not observed in nature. The only meson with the correct quantum numbers is the η' , however its mass is too big $m_{\eta'} = 958$ MeV. In the late 70's, Gerard 't Hooft [68, 80] showed that the axial current of singlet

$$j_\mu^5(x) = \sum_{i=1}^{N_f} \bar{\psi}_i(x) \gamma_\mu \gamma_5 \psi_i(x) \quad (3.41)$$

was not conserved at the quantum level due to the presence of field configurations with non trivial topology. In particular, the anomaly explicitly depends on the topological charge

$$\partial_\mu j_\mu^5(x) = \frac{N_f}{16\pi^2} G_{\mu\nu}^a \tilde{G}_{\mu\nu}^a = 2N_f Q(x) , \quad (3.42)$$

and cannot, therefore, be set to zero, resulting in the non conservation of the axial current.

The η' mass was also explained through the anomaly in connection with topology. More precisely the famous Witten-Veneziano formula relates it directly with the topological susceptibility in Yang-Mills theory

$$\chi^{(YM)} = \frac{f_\pi^2}{2N_f} (m_\eta^2 + m_{\eta'}^2 - 2m_K^2) . \quad (3.43)$$

We shall briefly show how exactly the instantons are related to the divergence of the axial current. Let us consider the case of massless quarks in which the chiral unbroken symmetry is exact. In this case, the fermion propagator is the inverse of the Dirac operator $S(x, y) = \langle x|(i\mathcal{D})^{-1}|y\rangle$. The Dirac operator can be expressed

in term of its eigenfunctions $i\mathcal{D}\psi_\lambda(y) = \sum_\lambda \lambda \psi_\lambda(y)$, so that the fermion propagator reads :

$$S(x, y) = \sum_\lambda \frac{\psi_\lambda(x)\psi_\lambda^\dagger(y)}{\lambda} . \quad (3.44)$$

One can then express the variation of the axial charge $Q_5 = \int$ in terms of the fermion propagator (3.44)

$$\begin{aligned} \Delta Q_5 &= Q_5(t = +\infty) - Q_5(t = -\infty) = \int d^4x \partial_\mu j_\mu^5(x) \\ &= \int d^4x N_f \partial_\mu \text{Tr}(S(x, x) \gamma_\mu \gamma_5) \\ &= N_f \int d^4x \text{Tr} \left(\sum_\lambda \frac{\psi_\lambda(x)\psi_\lambda^\dagger(x)}{\lambda} 2\lambda \gamma_5 \right) . \end{aligned} \quad (3.45)$$

Now we note that $\gamma_5 \psi_\lambda$ is also eigenfunction of the Dirac operator :

$$\mathcal{D}(\gamma_5 \psi_\lambda) = -\gamma_5 \mathcal{D}\psi_\lambda = -\lambda \gamma_5 \psi_\lambda , \quad (3.46)$$

with eigenvalue $-\lambda$. That means ψ_λ and $\gamma_5 \psi_\lambda$ are orthogonal, which makes the integral (3.45) vanish for every $\lambda \neq 0$. On the other hand, the zero modes ψ_0 give a finite contribution. Furthermore, one can choose the base in which chirality is well defined, i.e. $\gamma_5 \psi_0 = \pm \psi_0$ for right/left handed zero modes. In this way the integral (3.45) reduces to

$$\Delta Q_5 = 2N_f(n_L - n_R) , \quad (3.47)$$

where n_L and n_R are the number of positive and negative chirality zero modes of the Dirac operator. The connection with topology arose when 't Hooft realized that the Dirac operator $i(\partial_\mu + A_\mu)\gamma_\mu$ has a zero mode solution when the field A_μ is the instanton (3.19). This solution is

$$\psi_0(x) = \frac{\rho}{((x - x_0)^2 + \rho^2)^{3/2}} \phi , \quad (3.48)$$

where x_0 and ρ are, respectively, center and size of the instanton, and ϕ is a space-independent spinor. However, since $\mathcal{D}\psi_0 = 0$, it is clear that $\det([\mathcal{D}(A_\mu^I)]) = 0$ along the instanton solution A_μ^I . That means the functional integral of the vacuum

$$Z = \int DA_\mu \det(\mathcal{D}[A_\mu]) e^{S(A_\mu)} , \quad (3.49)$$

will vanish exactly for such configurations. Another way to say it is that tunnelling between vacua of the theory is suppressed in the presence of massless fermions. In

the chiral limit ($m_{ferm} = 0$), in fact, we know that the topological susceptibility has to vanish $\chi_{top}^{chiral} = 0$. However, the tunnelling amplitude is non-zero in the presence of *external quark sources*, because zero modes in the denominator of the quark propagator may cancel against zero modes in the determinant.

Although the theory in the continuum is quite developed, exact calculations of many quantities (such as η' mass, tunnelling amplitudes, etc...) require a non-perturbative approach. In particular, the study of the theory dependence from the θ -term is very challenging on the lattice since it is affected by the sign problem. Already the pure gauge theory $SU(3) + \theta$ -term cannot be simulated with standard Monte Carlo methods for real θ . Our goal, in this chapter, is to study pure gauge theory at real θ .

3.2 Topology on the Lattice

The nature of topological phenomena is highly non-perturbative. Some quantities could still be computed using approximated models like the diluted instanton gas approximation or in the large N limit. For a completely non-perturbative study, however, the use of lattice field theory is required. Historically, the first main reason that brought interest in the study of topology on the lattice was the non-perturbative determination of the topological susceptibility for the Witten-Veneziano formula (3.43).

At first, one might think that the lattice should not be able to retain the topological content of the configurations. In fact, lattice regularization requires the space to be discrete where every link can be continuously deformed into the trivial one $U_\mu(x) = \mathbb{1}$. Therefore, the configuration space is simply connected; that means homotopy classes do not exist on the lattice and topology is, strictly speaking, always trivial. However, one must not forget that any operator defined on the lattice has a physical meaning only in view of its continuum limit. The topological quantities are no exception and so one just has to define a regularization of those operators such that they represent the proper physical quantities in the continuum limit, where the standard non-trivial topology is recovered. In particular, it can be shown [81, 82] that for small values of the local action density

$$s_p = \sum_{\mu=1}^3 \sum_{\nu>\mu}^4 \left(\frac{1}{6} \text{Re Tr} [\mathbb{1} - \Pi_{\mu\nu}(p)] \right) < \epsilon , \quad (\epsilon \approx 0.07) , \quad (3.50)$$

where $\Pi_{\mu\nu}$ is the 1×1 plaquette, the lattice field configurations fall again into well defined topological sectors. This regime is reached naturally when approaching the

continuum limit where the value of the bare coupling becomes small ($g \rightarrow 0$) and

$$\langle s_p \rangle = \frac{3}{8} g^2 + O(g^4) . \quad (3.51)$$

However, it is not necessary to perform the continuum limit to meet the requirement (3.50). There are, in fact, various methods which main idea consists of smoothing the field configurations towards the local minimum of the Wilson action. This procedure "cool" down the configurations to values of $S < \epsilon$, getting rid of the short range fluctuations letting the long range topological modes emerge. We are going to discuss one of these methods in more detail later in this chapter.

We will start by introducing the local operator representing the density of topological charge on the lattice. As usual, more than one operator can be built that recover the correct naive continuum limit

$$q_L(x) \rightarrow a^4 q(x) + O(a^6) , \quad (3.52)$$

and they differ for orders $O(a^6)$. One of the simpler and more common choice is the twisted double plaquette operator

$$q_L(x) = -\frac{1}{2^4 \times 32\pi^2} \sum_{\mu\nu\rho\sigma=\pm 1}^{\pm 4} \epsilon_{\mu\nu\rho\sigma} \text{Tr}[\Pi_{\mu\nu}(x)\Pi_{\rho\sigma}(x)] , \quad (3.53)$$

where, again, $\Pi_{\mu\nu}$ is the 1×1 plaquette and $\epsilon_{\mu\nu\rho\sigma}$ is the completely anti-symmetric Levi-Civita tensor extended to the negative directions with the prescription $\epsilon_{\mu\nu\rho\sigma} = -\epsilon_{(-\mu)\nu\rho\sigma}$. The presence of $\epsilon_{\mu\nu\rho\sigma}$ requires the plaquettes $\Pi_{\mu\nu}$ and $\Pi_{\rho\sigma}$ to lie on two completely orthogonal hyperplanes. To get the bare, i.e. unrenormalised, topological charge on the lattice one has, then, simply to sum over all the lattice sites

$$Q_L = \sum_x q_L(x) . \quad (3.54)$$

Taking the naive limit $a \rightarrow 0$, one can show (3.53) to classically recover the correct form of the continuum

$$q_L(x) \xrightarrow[a \rightarrow 0]{} a^4 \frac{g^2}{32\pi^2} \text{Tr} \left[G_{\mu\nu}^a \tilde{G}^{a\mu\nu} \right] + O(a^6) . \quad (3.55)$$

However, as anticipated before, the operator (3.53) needs to be correctly regulated on the lattice before it can represent the topological charge of the lattice fields. In particular a multiplicative renormalization is required

$$q_L(x) \rightarrow a^4 Z(g^2) q(x) + O(a^6) , \quad (3.56)$$

where $Z(g^2)$ is a finite function of the bare coupling g^2 , obeying $Z(g^2) \rightarrow 1$ in the limit $g \rightarrow 0$. The value of $Z(g^2)$ can be computed in perturbation theory near the continuum fixed point and, for SU(3), it is [83, 84]

$$Z(g^2) = 1 - 0.908g^2 + O(g^4) . \quad (3.57)$$

The presence of the additional renormalization (3.57) is the reason why computing the bare topological charge (3.54) will not result only in integer numbers, but any real value. The value of $Z(g^2)$ cannot be computed using perturbation theory far away from the continuum limit. In this case one has to employ some techniques that can extrapolate the topological content from a configuration in a non perturbative way. Apart from using the fermionic definition of topology, via the index theorem, there are a number of techniques (cooling, smearing, gradient flow), that successfully deal with the gluonic operator (3.53). Despite the differences between those methods, they all involve smoothing of the field in order to eliminate the short range quantum fluctuations and enhance the classical long range modes. In fact, they all result in the recovery of an almost integer value for the topological charge (3.54).

Problems with lattice renormalization of topological observables first arose when trying to compute the topological susceptibility on the lattice as a way to validate the Witten-Veneziano formula for the η' mass. The expression for the topological susceptibility on the lattice is the topological charge two point function

$$\chi_L = \sum_x \langle q_L(x)q_L(0) \rangle = \frac{\langle Q_L^2 \rangle}{V}. \quad (3.58)$$

This quantity will be affected, other than from the multiplicative renormalization, also from an additive one due to the contact terms arising when $x = 0$:

$$\chi_L = Z^2(g^2) a^4 \chi + M(g^2). \quad (3.59)$$

The additive renormalization $M(g^2)$ contains the mixing terms of χ_L with all the other operators with the same quantum numbers and dimension $\dim_O \leq \dim_{\chi_L} = 4$, i.e. the trace of the energy tensor and the identity

$$M(g^2) \sim A(g) \langle G_{\mu\nu}^a G_{\mu\nu}^a \rangle + B(g). \quad (3.60)$$

In general, those renormalizations completely dominate the value of Q_L and χ_L even in the scaling region of the continuum limit (80%-90% of the real value). That is why particular care has to be put in eliminating those factors.

From now on we are going to use the following notation for the topological charge on the lattice

$$\begin{aligned} Q &\text{ is the renormalized topological charge ,} \\ Q_L &\text{ is the bare one ,} \end{aligned} \quad (3.61)$$

where two definitions are related via (3.56)

$$Q_L = Z(g^2) Q. \quad (3.62)$$

The renormalization factor $Z(g^2)$ can be non-perturbatively estimated using one of the smoothing methods that allows to sample the the correct value of the topological charge

$$Z(g^2) = \frac{\langle QQ_L \rangle_{\theta=0}}{\langle Q^2 \rangle_{\theta=0}}, \quad (3.63)$$

where the expectation value of Q^2 has to be chosen because $\langle Q \rangle_{\theta=0} = 0$.

3.2.1 Gradient Flow for $SU(3)$

We shall now briefly describe the gradient flow method in relation to topology [85, 86]. It has been shown in [87], for $SU(3)$ gauge theory, that the smoothing process obtained with gradient flow is equivalent to the one obtained with standard cooling. However, the first one seems to be more suitable to be extended to the $SL(3, \mathbb{C})$ group in order to be employed together with complex Langevin dynamics. This is because most of the other methods (like stout smearing, action cooling) require the re-projection of some quantities onto the original unitary manifold, which is not easily extendible in the case where the original manifold is $SL(3, \mathbb{C})$. On the contrary, the gradient flow does not present this inconvenience as it does not involve any projection and, therefore, can be extended more naturally to the complexified manifold. The gradient flow equation, in the continuum, is

$$\begin{aligned}\dot{V}_\mu(x, \tau) &= -g^2 [\partial_{x,\mu} S(V(\tau))] V_\mu(x, \tau) \\ V_\mu(x, 0) &= U_\mu(x),\end{aligned}\tag{3.64}$$

where with $\dot{V}_\mu(x, \tau)$ we indicate the differentiation of the link $V_\mu(x, \tau)$ with respect to the flow time τ , $S(V(\tau))$ is the Wilson action at the time τ and $\partial_{x,\mu}$ is the Lie derivative with respect of the link $V_\mu(x, \tau)$. We should note that (3.64) is very similar to the infinitesimal classical, i.e without statistical noise, Langevin equation. They differ, though, because the coupling g cancels from the equation (3.64) in such a way that the evolution in gradient flow time is independent from the coupling. Since the variation of the field $\dot{V}_\mu(x, \tau)$ is proportional to minus the gradient of the action, it is clear that along the gradient flow, the action decreases monotonically, $\dot{S} \leq 0$, smoothing the gauge field in accord with (3.50).

The discretized version of (3.64) is :

$$\frac{V_\mu(x, \tau + \epsilon) - V_\mu(x, \tau)}{\epsilon} = - \left(\sum_a Tr [\lambda_a (\Gamma_\mu(x) - \Gamma_\mu^{-1}(x))] \lambda_a \right) V_\mu(x, \tau) \tag{3.65}$$

where $S(V(\tau))$ has been explicitly taken to be the Wilson action and

$$\Gamma_\mu(x) = \sum_{|\nu| \neq \mu, \nu = -4}^4 \Pi_{\mu\nu}(x) \tag{3.66}$$

is the sum of all the plaquettes containing the link U_μ . There are several ways to numerically integrate the equation (3.65); more refined algorithms will have better convergence in ϵ but will be computationally slower. We choose, for simplicity, to

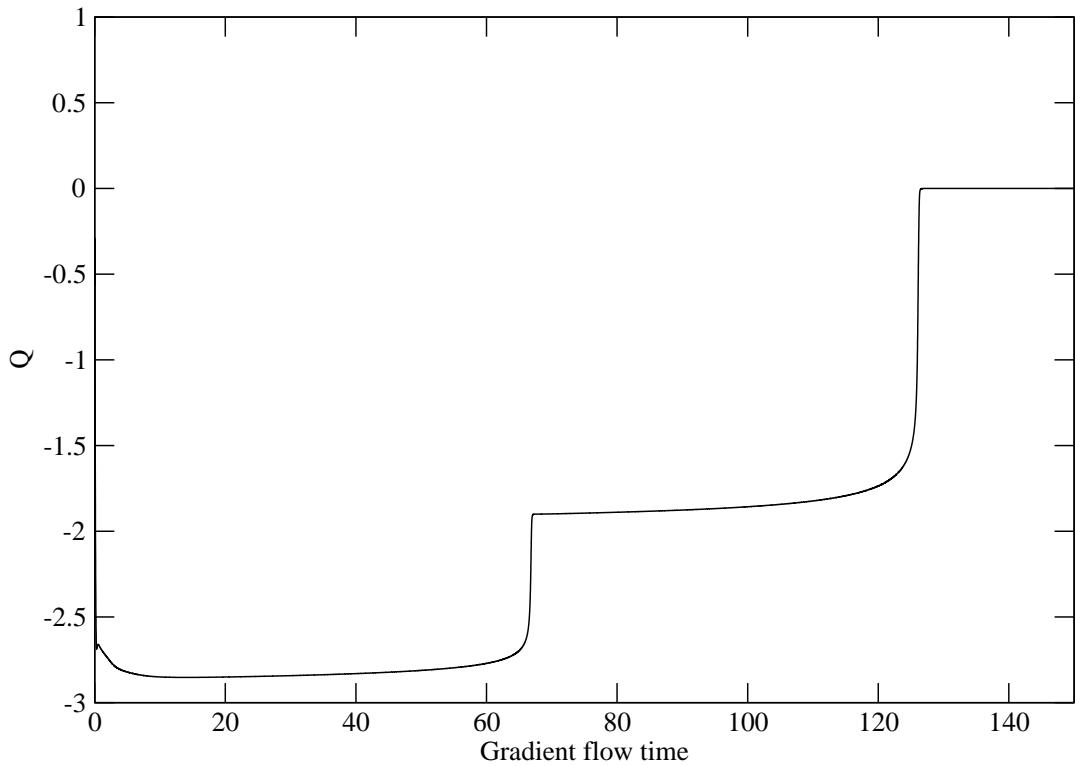


Figure 3.1: History of the topological charge of a SU(3) configuration (with topological charge $Q = -3$) along the gradient flow, for a 12^4 lattice at $\beta = 5.9$.

integrate via exponentiation, obtaining an expression analogous to Langevin without noise :

$$V_\mu(x, \tau + \epsilon) = \exp \left[-\epsilon \left(\sum_a \text{Tr} [\lambda_a (\Gamma_\mu(x) - \Gamma_\mu^{-1}(x))] \lambda_a \right) \right] V_\mu(x, \tau) \quad (3.67)$$

up to orders $O(\epsilon^2)$. In Fig. 3.1 we show the application of the gradient flow to a SU(3) configuration.

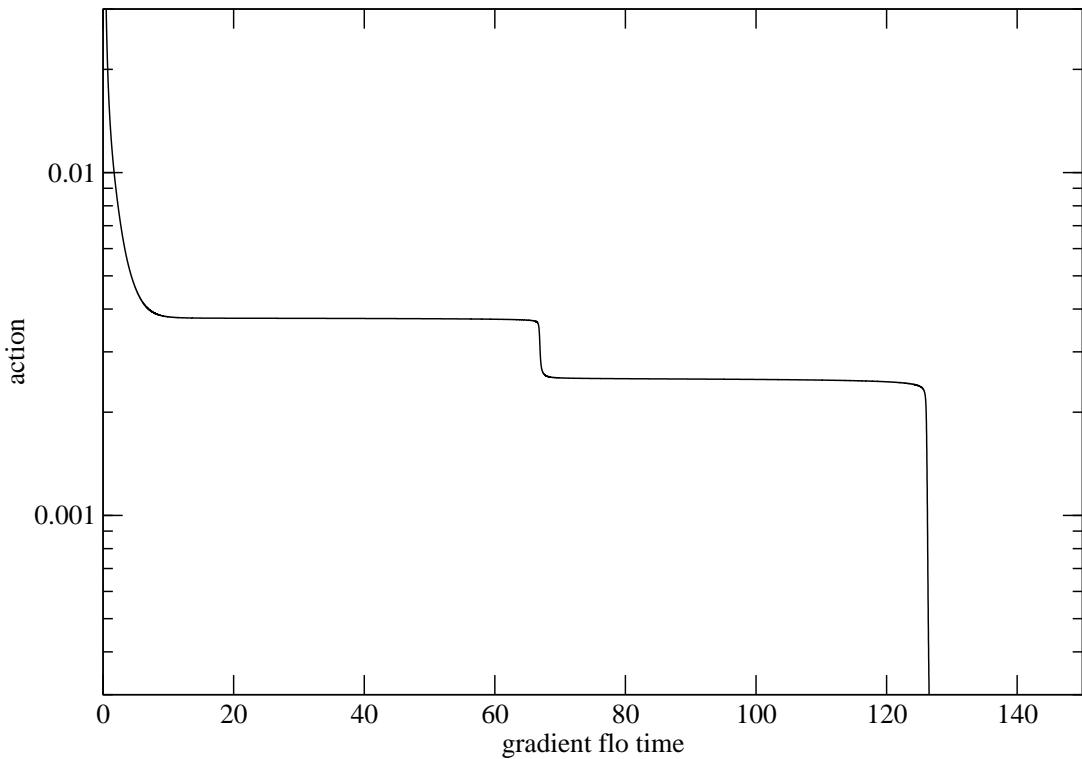


Figure 3.2: History of the action S for the same configuration of Fig. 3.1 .

As one can see, after some thermalization time, the value of the topological charge stabilizes on an (almost) integer plateau. Eventually it will fall down onto an other (almost) integer plateau and so on until the configuration reaches the real classical minimum of the Lattice Wilson action, the trivial one with all links $U_\mu(x) = \mathbb{1}$ and $Q = 0$. In this specific example we can see how the topological content of the configuration consists in 2 anti-instantons, one with $Q = -1$ and the other with $Q = -2$. Let us point out that the fact that the plateaus are not exactly integers is just a finite size effect. As soon as the instanton's size becomes much smaller than the physical volume considered, the values of the plateaus reproduce the correct integers. We can also look at what happens to the action during the gradient flow time. In Fig. 3.2 we can see how the plateaux of the action represents the metastable minima of the Wilson action. As soon as an instanton is destroyed the system tunnels back to the minimum closer to the absolute minimum of the Wilson action with $Q = 0$.

3.3 Complex Langevin dynamics for SU(3) gauge theory with a θ -term

We shall introduce now the theory we want to study. In the introduction to this chapter, we mentioned why the strong CP problem is still a hot topic to be studied and how the introduction of an axion field might be an elegant solution for it. The Peccei-Quinn argument, however, requires the θ parameter to be a dynamical variable of the theory and, because of that, one needs to be able to perform simulations at finite θ .

Other than that, the dependence of QCD on θ is of theoretical and phenomenological interest by itself. In fact, derivatives of the vacuum free energy $F(\theta)$ in respect to θ , computed at $\theta = 0$, enter various aspects of hadron phenomenology. For example, we saw that the topological susceptibility χ_{top} enters the Witten-Veneziano equation for the η' mass. The susceptibility χ_{top} can be seen both as the second moment of the topological charge distribution and as the linear response of the $\langle Q_{top} \rangle$ to the parameter θ .

In the continuum, the Lagrangian of pure gauge theory with the θ -term reads

$$\mathcal{L}_\theta = \frac{1}{4} G_{\mu\nu}^a(x) G_{\mu\nu}^a(x) - i\theta \frac{g^2}{64\pi^2} \epsilon_{\mu\nu\rho\sigma} G_{\mu\nu}^a(x) G_{\rho\sigma}^a(x), \quad (3.68)$$

where in the second part we recognise the topological charge density introduced in (3.13). Noticing that $\epsilon_{\mu\nu\rho\sigma} G_{\mu\nu}^a(x) G_{\rho\sigma}^a(x) = 4 \epsilon_{0\nu\rho\sigma} G_{0\nu}^a(x) G_{\rho\sigma}^a(x)$, it is easy to see that the Wick rotation has no effect on the topological charge. This fact is quite relevant since it results in the θ -term being imaginary also in Euclidian space

$$Z(\theta) = \int [DA] e^{-S_{YM}} e^{-i\theta Q_{top}} = \exp[-Vf(\theta)], \quad (3.69)$$

where $f(\theta) = \frac{F(\theta)}{V}$ is the free energy density. From the large- N limit, it has been conjectured [88, 89] that the free energy $F(\theta)$ is a multibranched function of the type

$$F(\theta) = N^2 \min_k H\left(\frac{\theta + 2k\pi}{N}\right). \quad (3.70)$$

This function is periodic in θ , but (if H is not constant) it cannot be smooth and, at some value of θ , there has to be a jump between two different branches. In particular only the three points $\theta = 0$ and $\theta \pm \pi$ leave (3.70) invariant under a CP transformation ($\theta \rightarrow -\theta$). Furthermore, the absolute minimum of the free energy must be at $\theta = 0$ where the integrand of the Euclidean space path integral is real and positive ($Z_{max} = e^{-Vf_{min}}$).

It would be very interesting to verify this conjecture directly. However, this is not possible because the complex integration measure (3.69) prevents, in general, the theory to be studied on the lattice using standard Monte Carlo methods. However,

some information on the free energy has been extrapolated, around $F(0)$, using imaginary θ [40, 41, 42]

$$\theta_R \equiv -i\theta_I , \quad (3.71)$$

and then analytically continue the results at real θ_R . One might expand the free energy density around the minimum at $\theta = 0$

$$\mathcal{F}(\theta) = f(\theta) - f(0) = \frac{1}{2}\chi\theta^2 s(\theta) , \quad (3.72)$$

where χ is the topological susceptibility and $s(\theta)$ is a dimensionless function of θ , which can be expanded as

$$s(\theta) = 1 + b_2\theta^2 + b_4\theta^4 + \dots . \quad (3.73)$$

As said before, the coefficients χ, b_2, b_4 , etc. are of great phenomenological interest and they are related, through derivatives of the free energy, to the connected momenta of the topological charge distribution. From (3.69), in fact, one finds :

$$\frac{d^k}{d\theta^k} f(\theta) = -(-i)^k \frac{\langle Q^k \rangle_c^\theta}{V} , \quad (3.74)$$

and, in particular, this equation can be used to relate the topological charge expectation value to the coefficients :

$$\langle Q \rangle_\theta = i V \frac{d}{d\theta} f(\theta) = i V \chi \theta (1 + 2b_2\theta^2 + 3b_4\theta^4 + \dots) . \quad (3.75)$$

It is clear, then, that the topological charge distribution is a very interesting observable to be studied at finite θ .

As we mentioned before, the definition of the topological charge on the lattice is affected by the renormalization factor $Z(\beta)$ (3.63). When simulating at finite θ , this will not only affect the measurements of topological observables, a problem solvable using cooling techniques, but also the evolution of the system itself. Since the action depends on the bare topological charge $S_\theta = i\theta Q_L$, it is necessarily affected by its renormalization. The common strategy, in the case of finite θ , is to calculate $Z(\beta)$ at $\theta = 0$ and then carry out the simulation at finite θ absorbing the renormalization into the parameter

$$S_\theta = i\theta_L Q_L = i\theta_L Z(\beta) Q = i\theta Q . \quad (3.76)$$

Here Q is the proper integer topological charge and we called θ_L the bare parameter assigned in the simulation. Eq (3.76) tells us that the actual value of the coupling with the topological charge is the renormalized

$$\theta = Z(\beta)\theta_L . \quad (3.77)$$

One needs, then, to rescale one's results in order to be dependent on the correct renormalized parameter θ .

In the following we are going to introduce the complex Langevin dynamics for pure gauge in the presence of a finite θ -term, and we are going to show our results.

3.3.1 complex Langevin dynamics

Let us discuss the set up and recall the CL equation for gauge theories

$$\begin{aligned} U_{\mu x}(t + \epsilon) &= R_{\mu x}(t)U_{\mu x}(t) , \\ R_{\mu x}(t) &= \exp \left[-i \sum_a \lambda_a (\epsilon D_a S[U_{\mu x}] + \sqrt{\epsilon} \eta_{a\mu x}) \right] , \end{aligned} \quad (3.78)$$

where both $U_{\mu x}$ and $R_{\mu x}$ belong to $\text{SL}(3, \mathbb{C})$, as it is evident from (3.78) when the action $S[U]$ is complex. Again, η_a is *real* Gaussian noise satisfying the relations

$$\begin{aligned} \langle \eta_{a\mu x} \rangle &= 0 , \\ \langle \eta_{a\mu x}(t) \eta_{b\mu x}(t') \rangle &= 2\delta(t - t')\delta_{ab} . \end{aligned} \quad (3.79)$$

The lattice action is the discretized version of (3.68), i.e. Wilson action plus the topological charge coupled with the parameter θ

$$\begin{aligned} S = \sum_x \left[\beta \sum_{\mu\nu} \left(1 - \frac{1}{2N} [\text{Tr}(\Pi_{\mu\nu}) + \text{Tr}(\Pi_{\mu\nu})^{-1}] \right) \right. \\ \left. - \frac{i\theta}{2^4 \times 32\pi^2} \sum_{\mu\nu\rho\sigma=\pm 1}^{\pm 4} \tilde{\epsilon}_{\mu\nu\rho\sigma} \text{Tr}[\Pi_{\mu\nu}(x)\Pi_{\rho\sigma}(x)] \right] , \end{aligned} \quad (3.80)$$

where with $\Pi_{\mu\nu}$ we indicate the plaquette. It is useful, at this point, to fix some notation, with which we are going to indicate the following :

$$\begin{aligned} \theta &= \text{renormalized } \theta, \text{ equivalent to the one in the continuum ,} \\ \theta_L &= \text{bare (unrenormalized) } \theta \text{ on the lattice ,} \\ \theta_I &= \text{imaginary bare } \theta , \\ \theta_R &= \text{real bare } \theta . \end{aligned} \quad (3.81)$$

We shall now briefly discuss our strategy. The results shown in this section are mainly presented in the papers [44, 43]. Before starting to simulate the complex action at real θ_R , a lot of care has to be put in the preparatory tests. As discussed in the last chapter 2, in order to control the exploration of the complexified configuration space of complex Langevin dynamics, gauge cooling is needed.

- Particularly important is the behaviour of CL at *imaginary* θ_I where the action, starting from a $\text{SU}(3)$ configuration, should remain real during all the evolution. As it is well known, also in the case of Monte Carlo methods, machine rounding errors tend to kick the dynamics off the $\text{SU}(3)$ manifold. The standard procedure that prevent this to happen is called *reunitarization* and consists in projecting back, every now and then, each link into the original $\text{SU}(3)$. Reunitarization can be applied also to Langevin dynamics, as long as the action is real, granting its convergence (as seen in Cap.2 for convergence of

real Langevin dynamics). However that would not give us any information on the stability of CL in the case when the action is complex and reunitarization is not applicable. What we do, instead, is to let free CL to explore the complex space, even at θ_I , and just use gauge cooling (see Cap.2) to control it. In this way we can study the stability of CL in reproducing the correct distribution. Furthermore, it can be observed (and it has been shown [19]) that the main concern for the correct convergence of CL, comes from the β dependence of the Wilson action rather than from the parameters that trigger the sign problem (θ in this case). Making sure CL converges correctly with the real action is, then, a strong hint for its convergence even with complex action, at least in a similar range of parameters.

- The second point we want to make sure of is for the observables to behave smoothly in the transition from θ_I (real action, $\theta^2 < 0$) to θ_R (complex action $\theta^2 > 0$). After we ascertained CL can be controlled with gauge cooling at θ_I , a smooth (without jumps) behaviour around $\theta^2 = 0$ indicates convergence also in the complex manifold, at least for small θ_R .
- Third and most important argument is the satisfaction of the criteria of correctness discussed in Cap.2. The compactness of the observables in the complex directions is always the most important property to be verified in order to ensure convergence of CL .

We shall proceed with the first point. As a proof of good convergence of CL at θ_L we choose to reproduce the topological charge distribution and confront it with the one obtained via the Hybrid Monte Carlo (HMC) method. In Fig.(3.3) we show the result of such a test, at bare $\theta_L = \pm 20i$ and $\beta = 6$ on a 12^4 lattice. Here configurations are generated using complex Langevin dynamics, including gauge cooling, to control the process. We subsequently use gradient flow to smooth them and recover their topological content. We observe the expected response as the sign of θ_L is flipped, i.e. $\langle Q \rangle_\theta$ is compatible with $-\langle Q \rangle_{-\theta}$ as predicted by (3.75). Also we checked the average value of the topological charge and we found agreement with the one computed with the HMC algorithm

$$\begin{aligned} \theta_I = 20 : \quad \langle Q \rangle_{CL} &= -5.38(\pm 0.06), \quad \langle Q \rangle_{HMC} = -5.42(\pm 0.09) , \\ \theta_I = -20 : \quad \langle Q \rangle_{CL} &= 5.46(\pm 0.06), \quad \langle Q \rangle_{HMC} = 5.48(\pm 0.09) . \end{aligned} \quad (3.82)$$

That is enough to make us confident CL converges to the right results at θ_I , at least for values of β not much smaller than 6 .

To proceed with our testing plan, we compute the plaquette at 3 values of β , for the bare parameter going from imaginary $\theta_L^2 < 0$ to real $\theta_L^2 > 0$. In Fig.3.4, we can see the smooth transition of the plaquette from imaginary θ to real ones. Good agreement with Monte Carlo results is also shown at imaginary θ_L . As expected, the plaquette does not depend much on θ , so that it keeps staying on the same value also at $\theta_L^2 > 0$. Nevertheless, this is a very useful result because it shows the good behaviour of CL when the action is complex.

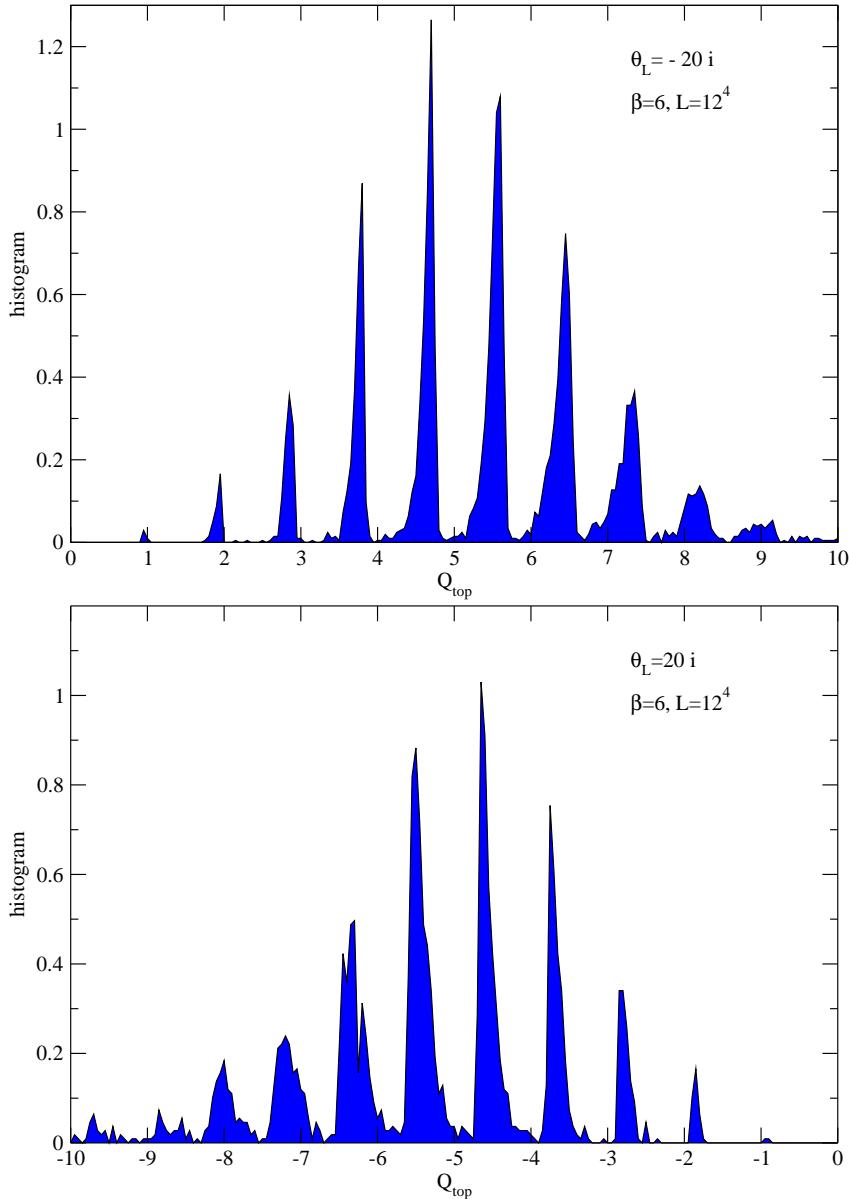


Figure 3.3: Distribution of Q_{top} at $\theta_I = -20i$ (above) and $\theta_I = 20i$ (below), on a 12^4 lattice at $\beta = 6$, obtained by generating configurations using complex Langevin dynamics with gauge cooling and subsequently applying gradient flow to recover the discrete values of the charges. The spacing between the peaks is a bit smaller than one as a finite volume effect, and this is responsible for the peaks being shifted in respect to the integers.

As mentioned before, the key point to trust the simulation at θ_R is the distribution of the observables in the complex plane. In Fig.3.5 we show the normalized histograms of the values of the real and imaginary part of the action (ReS and ImS). The plot shows, in a logarithmic scale, the decay of the observable in the complex

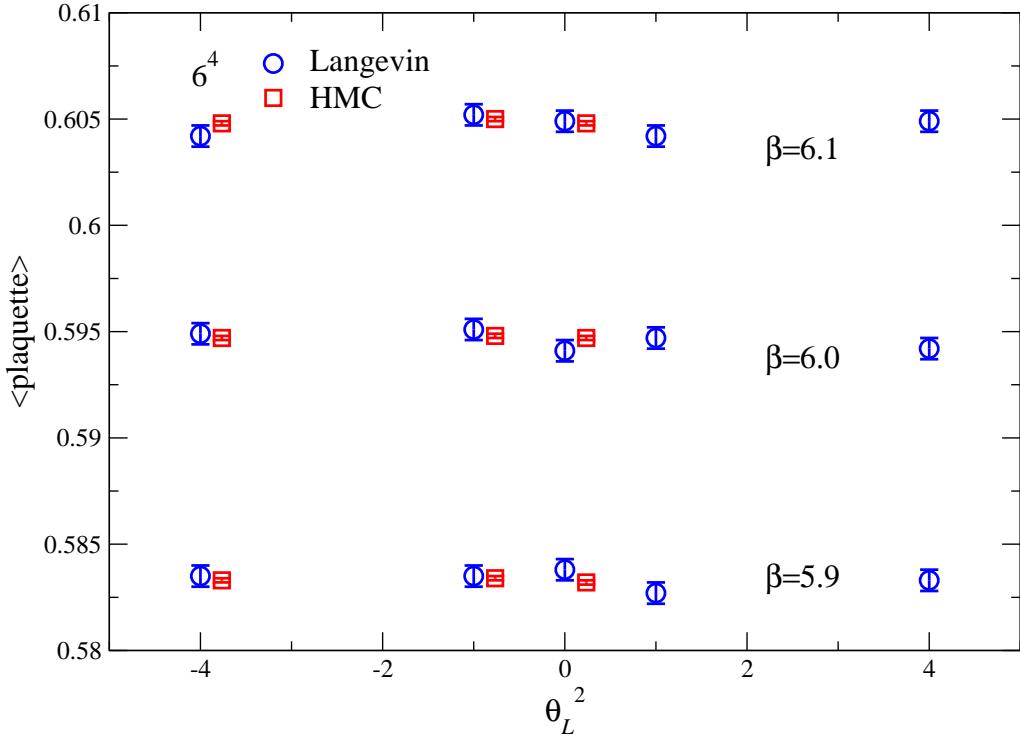


Figure 3.4: expectation values of the plaquette in $SU(3)$ Yang-Mills theory in the presence of a θ -term , for real and imaginary θ_L , using CL and HMC (imaginary θ_L only), on a 6^4 lattice at three β values. The HMC data has been shifted horizontally for clarity.

plane (often referred to as 'skirt' of the observable). The more tight the skirts are, the better CL criteria of correctness are satisfied. We can see that at larger β , gauge cooling is very effective in controlling the skirts of the distribution while it becomes more and more problematic at smaller β . Eventually the skirt will become too wide to be controlled and CL will have problems of convergence. Nevertheless in the cases presented, even for $\beta = 5.9$, CL reproduces the correct values for the action, as the results at small θ_R lie on the analytical continuation of the ones at θ_I (Fig.3.4).

Let us stress further that, at this stage, we are not yet interested in measuring physical quantities. In this sense, even if some observable is likely to be affected by finite-size effects in a 6^4 volume at $\beta = 6.1$, we are only interested in comparing CL and Monte Carlo at the same volume and β and get the same results.

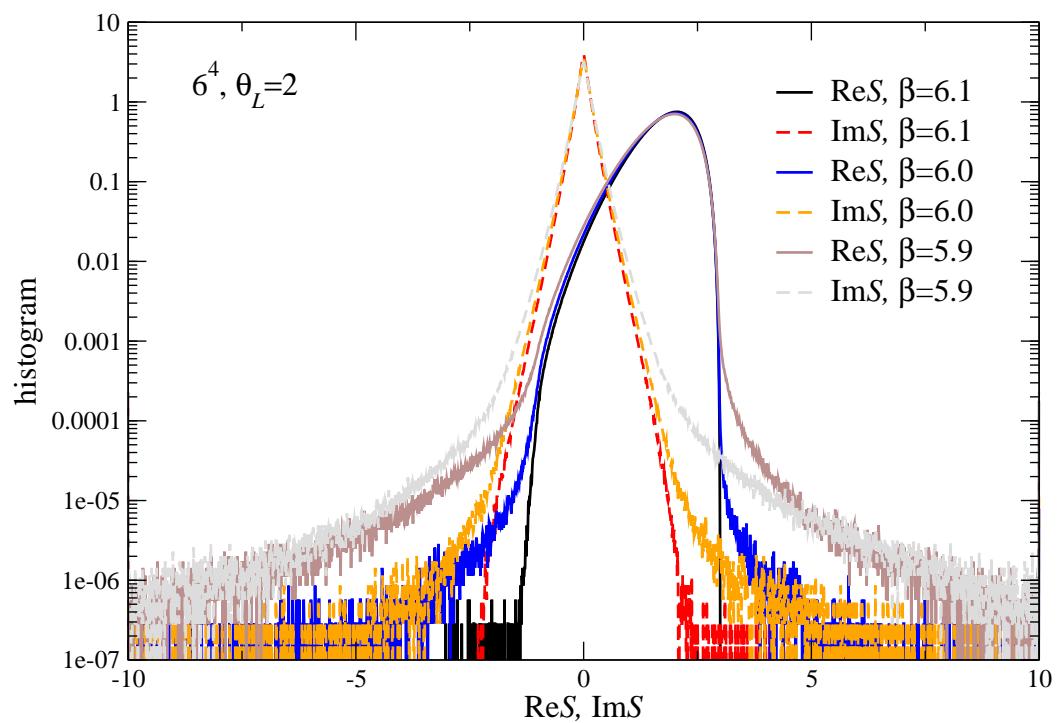


Figure 3.5: Normalized histogram of the distribution of the complex action for real $\theta_L=2$ and $\beta = 5.9, 6, 6.1$.

3.3.2 Topological charge distribution

Having completed the preliminary tests, we shall now go back to our goal to study the topological charge dependence on θ . The behaviour we would like to reproduce is the one shown in equation (3.75). Unfortunately to do so we run into the problem of renormalizations of θ . We mentioned before, (3.77), that it would be enough just to rescale θ in the action with the renormalization factor $Z(\beta)$ computed at $\theta = 0$ (since it only depends on β). However, this procedure is not correct when both the action and the observable measured are affected by the same renormalization. In particular, if the observable is Q_L itself, this can be seen from the partition function

$$\langle Q_L \rangle = \frac{1}{Z(\theta_L)} \int [DU] Q_L e^{-S_W} e^{-i\theta_L Q_L}, \quad (3.83)$$

where the noise of Q_L measured autocorrelates with the one of Q_L in the action. In this case, one first has to measure the correct renormalized topological charge Q and, only after, he is allowed to rescale θ_L to get the correct result. However, computing the renormalized topological charge at real θ would require the implementation of some smoothing technique in $SL(3, \mathbb{C})$ which is a hard task and is still a work in progress (we will show some results in the last section). Instead, what we are going to present here is a study of the bare lattice theory.

In the bare theory, the bare topological charge is still expected to follow the behaviour of (3.75), but with different values of the coefficients χ, b_2, b_4, \dots :

$$\begin{aligned} \langle Q_L \rangle_{\theta_I} &= -V \frac{d}{d\theta_I} f(\theta_I) = -V \chi_L \theta_I (1 - 2b_2^L \theta_I^2 + 3b_4^L \theta_I^4 + \dots), \\ \langle Q_L \rangle_{\theta_R} &= i V \frac{d}{d\theta_R} f(\theta_R) = i V \chi_L \theta_R (1 + 2b_2^L \theta_R^2 + 3b_4^L \theta_R^4 + \dots). \end{aligned} \quad (3.84)$$

Here we explicitly wrote down the dependence of Q_L on imaginary and real θ_L and indicated with χ_L, b_2^L, b_4^L the bare parameters. Note that for real θ $\langle Q_L \rangle$ must be imaginary because the real part of (3.83) cancels with a CP transformation.

In Fig.3.6 we show the behaviour of the bare topological charge in the transition from θ_I to θ_R , for small values of the bare parameters θ_L . The shaky lines are the running averages of $\langle \text{Re}(Q_L) \rangle$ at θ_I , where $\langle \text{Im}(Q_L) \rangle$ is 0. We can see the averages converging to the values predicted by Hybrid Monte-Carlo. When θ_L is real, we have the steady lines that represent $\langle \text{Im}(Q_L) \rangle$, while now $\text{Re}(Q_L)$ fluctuates around 0, exactly as expected.

In general the renormalization is of the order of $Z(\beta) \simeq 0.1$ [41], so that when we set $\theta_L = 1$ we are really working at roughly $\theta \simeq 0.1$. That means in Fig.3.6 θ_L is actually quite small and we can mostly see the linear part of (3.84), proportional to the bare topological susceptibility χ_L . To be able to observe the effect of higher powers of θ_L in (3.84) one needs larger θ_L . In Fig.3.7 we show the dependence of $\langle Q_L \rangle$ on θ_I (blue line) and θ_R (red line) for bigger values of θ_L . Again, imaginary and real parts of Q_L are multiplied by a " - " sign and an i , so that they can be plotted next to each other. Already from the plot it is evident that the two curves

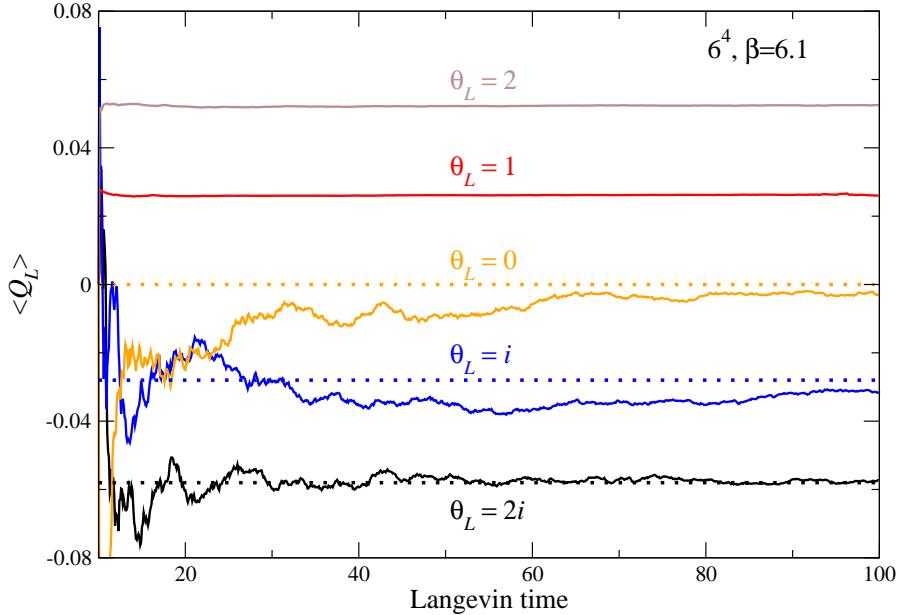


Figure 3.6: CL running averages of the bare topological charge, for $\beta = 6.1$. The imaginary part of Q_L is shown for θ_R and the real part for θ_I . The dotted lines at θ_I indicate results obtained with HMC algorithm.

have opposite curvature at high θ_L . More precisely, by fitting the data with (3.84) we obtain, for both curves

$$V\chi_L = 0.026, \quad \text{and} \quad b_2^L \simeq 10^{-5}, \quad (3.85)$$

confirming the expected common linear response and an equal and opposite curvature, in perfect accord with the opposite sign of the cubic term in (3.84).

We can now use our information about the θ dependence of $\langle Q_L \rangle$ to study the dependence of the bare topological susceptibility on β . Since an increasing value of β means lower temperature (and also smaller physical volumes), we expect the susceptibility to decrease. To see this we just have to repeat the fit in (3.85) at different values of β . In Fig.3.8 we show just the data at θ_R for fixed volume and increasing temperature (increasing β).

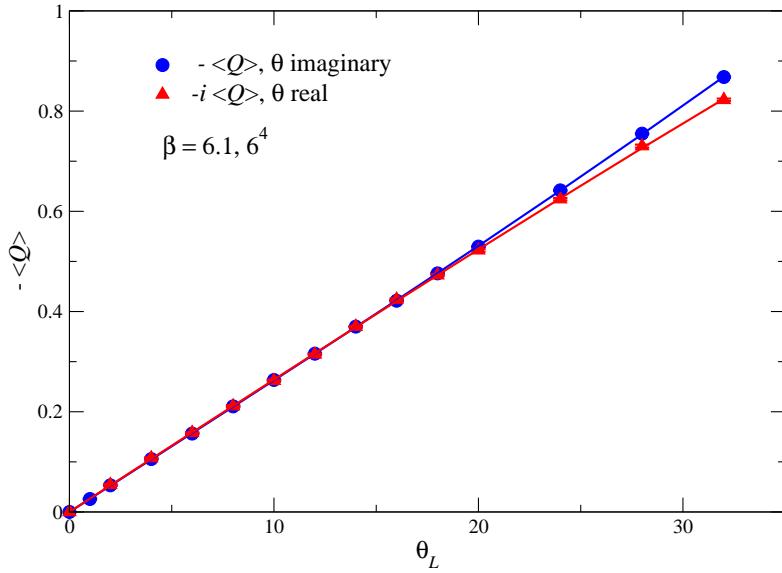


Figure 3.7: Opposite curvature in the dependence of $\langle Q_L \rangle$ on real and imaginary θ_L , at $\beta = 6.1$.

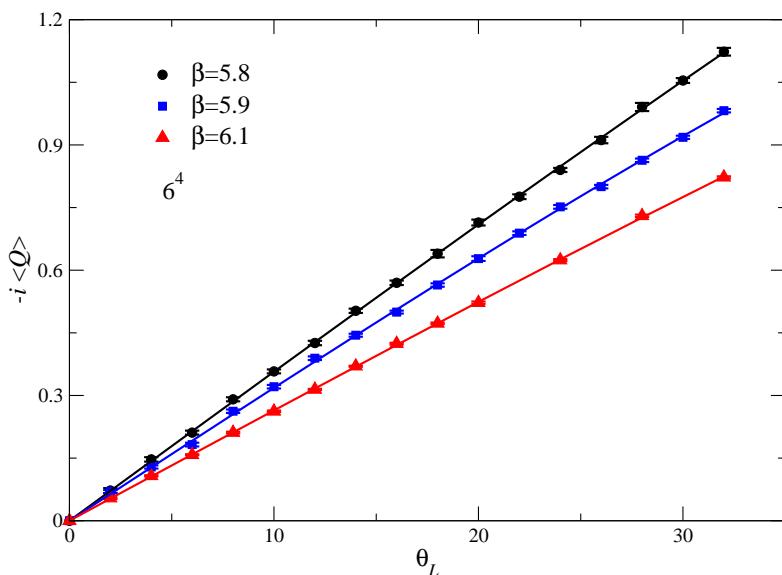


Figure 3.8: Same as Fig.3.7, at real θ_L , for three values of $\beta = 5.8, 5.9, 6.1$.

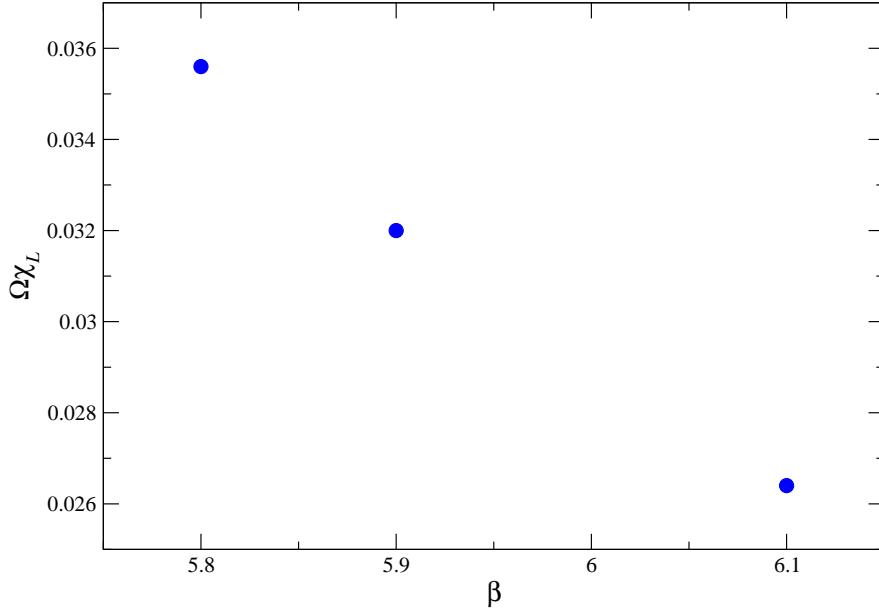


Figure 3.9: Bare topological susceptibility χ_L as a function of β .

The values for the bare topological susceptibility times the volume, i.e. the linear term of the curves in Fig.3.8, are reported in Fig.3.9. Although they are not trivially connected with the values of the physical renormalized topological susceptibility, they must have the same qualitative dependence on the temperature. In fact, we can see that χ_L decreases as β (temperature) increases, as is the case with the physical renormalized χ .

3.3.3 Topological charge on $SL(3, \mathbb{C})$ manifold

In the previous section we showed results for the Yang-Mills theory with bare topological charge. We already mentioned (Sec.(3.2.1)) the correct procedure to deal with the lattice renormalizations of topological quantities when the gauge group is $SU(3)$. However, CL at real θ works in the $SL(3, \mathbb{C})$ gauge group which greatly complicates things not only practically but also conceptually. The very definition of topology (reviewed in Sec.(3.1)) relies on the compactness of the gauge group that generates, for the pure gauge fields, a precise mapping between the spatial infinity three-sphere S_3 and the group itself. As soon as the gauge group gets enlarged

$$SU(3) \longrightarrow SL(3, C) \quad (3.86)$$

it loses its property of compactness and the Yang-Mills part of the action does not have well defined classical minima any more. Consequently, the operator

$$Q = \int d^4x F\tilde{F} \quad (3.87)$$

is not expected to assume integer values any more, nor to be invariant under small deformation of the fields .

Here we shall discuss some results on the subject that, however, are not complete yet. First of all we should keep in mind that $SU(3)$ is still sub-group of $SL(3, \mathbb{C})$, so we should expect to have some sectors where some part of the topological charge is an integer. Whether the actual topological charge is affected by this sector is another story and, in fact, we shall see that this is not the case.

It is useful, in order to understand what is going on, to bring up an example which refers to a much simpler theory. Let us consider a real scalar field theory with a topological term on a circle (S_1) :

$$0 \leq x < L , \quad 0 \leq \phi < 2\pi , \quad \phi = \phi + 2k\pi , \quad k \in \mathbb{Z} , \quad (3.88)$$

where the partition function is

$$Z = \int D\phi e^{-S+i\theta Q} , \quad (3.89)$$

and the action and the topological charge are defined as :

$$S = \frac{1}{2} \sum_x (\partial_x \phi)^2 , \quad Q = \frac{1}{2\pi} \sum_x \partial_x \phi . \quad (3.90)$$

It is easy to see that the topological charge must be an integer classifying how many times the field ϕ wraps around itself on the circle. Furthermore, again for CP reasons, the expectation value of the topological charge must be purely imaginary

$$\langle Q \rangle = \frac{1}{Z} \int D\phi e^{-S} i \sin(\theta Q) Q . \quad (3.91)$$

In order to use complex Langevin dynamics, we complexify the field

$$\phi \rightarrow \phi_R + i \phi_I , \quad (3.92)$$

where ϕ_I is unbounded, and we can write separately the real and imaginary part of the topological charge

$$Q = \sum_x \partial_x \phi_R + i \sum_x \partial_x \phi_I = Q_R + i Q_I . \quad (3.93)$$

The space still being a circle and the real part of the field ϕ_R periodic, we expect Q_R to have the exact same topological meaning of Q for the real theory. However using CL (see Cap.2), we know that the observables are distributed according to the equilibrium Fokker-Planck probability distribution function, which is *real* even when the fields are complex

$$\langle Q \rangle = \frac{1}{Z} \int D\phi P(\theta, \phi_I, \phi_R)(Q_R + i Q_I) . \quad (3.94)$$

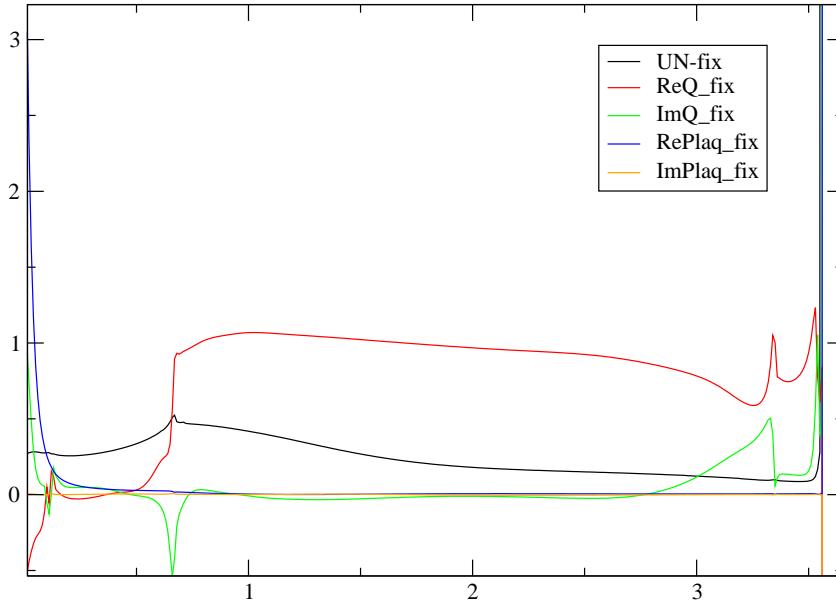


Figure 3.10: Wilson Flow on a $SL(3, \mathbb{C})$ configuration (x axis is flow time) for $\theta_L = 20$, $\beta = 6.1$, $L = 8^4$.

Since we know, from (3.91), that $\langle Q \rangle$ has to be purely imaginary for real θ , we have to conclude that $\langle Q_R \rangle = 0$. We conclude that with CL, the contribution to the average values of the topological charge at real θ comes entirely from the non-compact imaginary direction and, therefore, has no topological meaning in the enlarged manifold. Furthermore, the topological charge of the real subgroup, which still retains the usual topological properties, needs to cancel exactly for every real θ

The case of $SL(3, \mathbb{C})$ is completely analogous. Although the topological charge operator cannot be simply factorised as in (3.93), its properties must be the same. On the $SU(3)$ subgroup it retains its properties and meaning, but needs to average to 0. On the other hand, on the complex manifold, Q has no topological meaning but, provided good convergence of CL, it would average to the correct value.

We would like to observe this on a $SL(3, \mathbb{C})$ configuration. As anticipated, we are going to employ an extended version of the gradient flow discussed in Sec.((3.2.1)). That is because, out of all the smoothing methods, the gradient flow seems the one which can get more naturally extended to $SL(3, \mathbb{C})$ because it doesn't explicitly involves any projection on the real manifold. In Fig.3.10, we show the effect of the gradient flow on a thermalized configuration at real $\theta_L = 20$. In particular, it is interesting to observe the real part of the topological charge operator (red line) reaching an integer plateaux, proof that topological sectors are still present in the

$SU(3)$ subgroup. The imaginary part of Q , on the other hand, does not stabilise on any integer value and just flows to (more or less) 0.

Eventually, at flow time $t \simeq 3.6$, we can see that the gradient flow equation encountered an unbounded runaway direction, along which all the quantities diverged.

A consistent application of the gradient flow on the $SL(3, \mathbb{C})$ manifold is still under work. Its development is essential to extract the correct imaginary part of the topological charge operator from the configurations and, therefore, to obtain the correct value of the renormalized topological charge at real θ .

In this Chapter we reviewed some aspects on the instantons and the structure of the vacuum of the continuum theory. We discussed the necessity to study topology, non-perturbatively, on the lattice and the procedures required to extrapolate meaningful values of the topological observables. We mentioned why it is relevant to understand the theory at finite θ and how this theory is affected by the sign problem on the lattice. We adopted complex Langevin dynamics as a solution for this and we showed our results at real θ for the bare theory. We explained the difficulties involved in the extraction of the renormalized topological quantities in the complexified manifold in which CL dynamics takes place. Lastly, we gave an account of our progress in the direction of solving those problems.

Chapter 4

Lefschetz thimbles

In this chapter we are going to introduce another method to deal with sign problem in lattice quantum field theory, i.e. integration along the *Lefschetz thimble*. The starting idea is similar to complex Langevin dynamics. One is faced with an (highly) oscillatory integral over some real manifold, where the sign problem is severe, and tries to complexify the degrees of freedom so that in the new complex manifold the sign problem could be milder or absent. We saw in chapter 2 how complex Langevin dynamics enlarges the system into the complex manifold. There the probability distribution function, generated by the Fokker-Plank equation, is real and, provided the criteria of correctness are satisfied and the action is holomorphic, leads the observables to converge to the right expectation value.

For the Lefschetz thimble, instead, the idea is similar to the one of the saddle-point approximation. However, here the integral is not approximated by the Taylor expansion of the complex function around the dominant stationary point. Instead, the original real domain of integration is deformed into the union of new paths, of the same dimension, i.e. the Lefschetz thimbles. Each thimble is a sub-manifold of the complexified space, with the property of passing through one critical point of the action. Along them the phase of the integrand is constant, although in general it has a different value on each of them. The union of all the thimbles represents a deformation of the original real domain. If the function is holomorphic, the integral over the new paths should have the same (Cauchy theorem) results as over the real manifold, with the advantage that the phase along the thimbles is constant. The sign problem, however, has not completely disappeared. A residual sign problem is still present along each thimble, due to the curvature of the integration contour in the complexified space, i.e. the complex Jacobian coming from the parametrisation of the thimble. A second sign problem appears in the case that more than one thimble contributes. In this case one has to take in account that the phases, although constant, are different along each thimble. This is typically referred to as a global sign problem.

Witten proposed to apply this method to quantum field theory [90, 20] and, since then, much progress has been made in the direction of application to lattice field theory [91, 92, 22]. Mostly studied have been the relativistic Bose gas at finite

chemical potential [21, 93], but also real time quantum mechanics and condense matter [94, 95].

Since both complex Langevin dynamics and Lefschetz thimble work in the complexified space, it is very interesting to compare them. In the following we are going first to review the basic theory of the Lefschetz thimbles and, then, compare results for U(1) 1-link model with a determinant and SU(2) 1-link model.

4.1 Thimble equations

First we are going to introduce the basics of the complexification on Lefschetz thimbles. Let us consider n real degrees of freedom $\mathbf{x} = \{x_1, x_2, \dots, x_n\} \in \mathbb{R}^n$ that work as a domain for the complex action $S[\mathbf{x}] \in \mathbb{C}$. The partition function for this model is defined, as usual, by the path integral

$$Z = \int_{\mathbb{R}^n} d^n x e^{-S(\mathbf{x})}, \quad (4.1)$$

where $d^n x$ is the element of infinitesimal volume. As anticipated, the next step is to complexify the domain of integration, i.e. changing the variables $x_i \rightarrow x_i + iy_i$, while the action is extended to the holomorphic $S[z_i] = S[x_i + iy_i]$. Thanks to Picard-Lefschetz/Morse theory [96, 97], then, it is possible to select appropriate n dimensional sub-manifolds, i.e. the *Lefschetz thimbles*, immersed into the complex manifold, along which the integral has the same value as in the original real manifold. The new contour represent the *downward* gradient flow associated to the real part of the action

$$\frac{dz_i}{dt_i} = -\overline{\left(\frac{\partial S(\mathbf{z})}{\partial z_i} \right)} = -\frac{\partial S(\bar{\mathbf{z}})}{\partial \bar{z}_i}, \quad (4.2)$$

where t is the variable that parametrizes the thimble. Equivalently the equations for the single components are

$$\begin{aligned} \frac{dx_i}{dt_i} &= -\text{Re} \frac{\partial S(\mathbf{z})}{\partial z_i}, \\ \frac{dy_i}{dt_i} &= \text{Im} \frac{\partial S(\mathbf{z})}{\partial z_i}, \end{aligned} \quad (4.3)$$

The procedure is to start from the set Σ of critical points $\{\mathbf{z}_\sigma \mid [\partial_{z_i} S(\mathbf{z})]_{\mathbf{z}=\mathbf{z}_\sigma} = 0\}$ and construct the sub-manifold solution of (4.2) \mathcal{J}_σ that originates from those points.

Along those trajectories the imaginary part of $S[\mathbf{z}]$ is ensured to be constant and the integral with the real part is convergent :

$$\begin{aligned}\frac{d}{dt} \text{Im}S &= \frac{1}{2i} \left[\frac{\partial S(\mathbf{z})}{\partial z_i} \dot{z}_i - \frac{\partial S(\bar{\mathbf{z}})}{\partial \bar{z}_i} \dot{\bar{z}}_i \right] = 0 , \\ \frac{d}{dt} \text{Re}S &= \frac{1}{2} \left[\frac{\partial S(\mathbf{z})}{\partial z_i} \dot{z}_i + \frac{\partial S(\bar{\mathbf{z}})}{\partial \bar{z}_i} \dot{\bar{z}}_i \right] = - \left| \frac{\partial S(\mathbf{z})}{\partial z_i} \right|^2 \leq 0 .\end{aligned}\quad (4.4)$$

It is useful to look at the one-dimension case, where the original manifold \mathbb{R} gets enlarged into \mathbb{C} . Here we know (Cauchy theorem) that we can deform, with continuity, the original contour of integration as much as we want, as long as $S[z]$ is holomorphic and the new contour tends to \mathbb{R} for $x \rightarrow \pm\infty$. We have no problem, then, to accept that the integral of $e^{-S[z]}$ along any complex path, respecting those two conditions, will have the same value as on \mathbb{R} . Furthermore, we can write the thimble equations more explicitly

$$\begin{aligned}\frac{\partial}{\partial t} x &= -\frac{1}{2} \left(\frac{\partial u(x, y)}{\partial x} + \frac{\partial v(x, y)}{\partial y} \right) , \\ \frac{\partial}{\partial t} y &= -\frac{1}{2} \left(\frac{\partial u(x, y)}{\partial y} - \frac{\partial v(x, y)}{\partial x} \right) ,\end{aligned}\quad (4.5)$$

where we used the holomorphicity of $S[x + iy] = u(x, y) + iv(x, y)$, with $u(x, y)$ and $v(x, y)$ real functions, and the property of the derivatives $\frac{\partial}{\partial z} = \frac{1}{2} \left(\frac{\partial}{\partial x} - i \frac{\partial}{\partial y} \right)$. Here one can show directly that the imaginary part of the action $\text{Im}S[z(t)] = v(x(t), y(t))$ is constant. Let us take its derivative along the thimble

$$\begin{aligned}\frac{d}{dt} v(x(t), y(t)) &= \frac{\partial v}{\partial x} \dot{x} + \frac{\partial v}{\partial y} \dot{y} \\ &= -\frac{1}{2} \left(\frac{\partial v}{\partial x} \frac{\partial u}{\partial x} + \frac{\partial v}{\partial x} \frac{\partial v}{\partial y} + \frac{\partial v}{\partial y} \frac{\partial u}{\partial y} - \frac{\partial v}{\partial y} \frac{\partial v}{\partial x} \right) ,\end{aligned}\quad (4.6)$$

where we substitute the expression (4.5) for \dot{x} and \dot{y} . Using the Cauchy Riemann equations $\frac{\partial u}{\partial x} - \frac{\partial v}{\partial y} = \frac{\partial u}{\partial y} + \frac{\partial v}{\partial x} = 0$, it follows that $\dot{v}(x, y) = 0$ which means, of course, that $v(x, y)$ is constant. In 1 dimension, then, the thimbles method can be understood more intuitively using simple complex analysis. In higher dimension the argument is much more complicated but can be proven using Morse theory.

Together with the Lefschetz thimble, one can create another set of sub-manifolds passing through each critical point \mathbf{z}_σ and associated, this time, to the *upward* gradient flow. Those are path of steepest ascent \mathcal{K}_σ , generated by the equation

$$\dot{z}_i = + \overline{\partial_{z_i} S(\mathbf{z})} , \quad (4.7)$$

along which the Morse function $-\text{Re}S[\mathbf{z}]$ diverges. The unstable thimbles are essential to know which of the stable ones have to be taken into account to produce a complete equivalent the original real integration path. The prescription is

$$Z = \sum_{\sigma \in \Sigma} n_\sigma Z_\sigma \quad (4.8)$$

where n_σ is the 'intersection pairing'

$$n_\sigma = \langle \mathbb{R}^n, \mathcal{K}_\sigma \rangle . \quad (4.9)$$

of the unstable thimble, passing through the critical point \mathbf{z}_σ , with the original domain of integration. The product in (4.9) counts the intersection points (or submanifolds) of two manifolds, and the intersection number n_σ can be 0 or ± 1 , selecting which thimble contributes to the path integral and which not. Again, in the 1 dimensional case one can usually recognise the thimbles contributing just by looking at the ones that represent a continuous deformation of the original real cycle.

The integral over thimbles can be explicitly written as

$$Z_\sigma = e^{-i \text{Im} S(\mathbf{z}_\sigma)} \int_{\mathbf{z}(t) \in \mathcal{J}_\sigma} dz(t) e^{-\text{Re} S[\mathbf{z}(t)]} , \quad (4.10)$$

where the constant imaginary part $e^{i \text{Im} S[\mathbf{z}_\sigma]}$ has been factorized out of the integral and generates the global sign problem we anticipated before. The other, more subtle, residual sign problem comes from the integration measure $dz(t)$. Here, in fact, one has to take in account the Jacobian coming from integration along the complex manifold

$$\mathcal{D}z(t) = \left| \frac{\partial \mathbf{z}}{\partial t} \right| d^n t , \quad (4.11)$$

which could generate a non trivial sign problem due to the curvature of the thimble. This sign problem, however, is supposed to be milder than the one at the beginning and, therefore, can be treated without compromising the measurements [22].

All the results for Lefschetz thimbles (i.e. plots and numerical results) presented in this chapter are obtained by numerical integration with Mathematica of the Lefschetz equations (4.3).

4.2 Thimbles and Langevin dynamics

Here we show a practical example for the application of the Lefschetz thimble method to a quartic model. We are also going to compare it with the complex Langevin dynamics approach to the same model. This result appears in the papers [98].

The action we want to examine is

$$S(z) = \frac{\sigma}{2} z^2 + \frac{1}{4} z^4 + hz, \quad \sigma \in \mathbb{R}, h \in \mathbb{C}. \quad (4.12)$$

where the linear term explicitly breaks parity symmetry $z \rightarrow -z$ and generates, in this case, the sign problem. Here the original domain is \mathbb{R} and its complexification is the complex plane.

The complex Langevin dynamics have been discussed in Cap.2. Here let us just recall that the equation for z reads

$$\dot{z} = -\partial_z S(z) + \eta, \quad (4.13)$$

or explicitly

$$\begin{aligned} \dot{x} &= -\text{Re } \partial_z S(z) + \eta, \\ \dot{y} &= -\text{Im } \partial_z S(z), \end{aligned} \quad (4.14)$$

where η is real gaussian noise obeying

$$\begin{aligned} \langle \eta \rangle &= 0, \\ \langle \eta^2 \rangle &= 2. \end{aligned} \quad (4.15)$$

The Langevin equations (4.14) generate a *real* probability distribution function $P(x, y)$ over the complex plane,

$$Z = \int dx dy P(x, y), \quad (4.16)$$

which is the solution of the Fokker-Plank equation associated to (4.14).

First of all we should note how the classical (i.e. without noise η) CL equation (4.13) is the complex conjugate of the thimble equation (4.2), i.e. the imaginary part has opposite sign. That leads to some important differences regarding stability of some trajectories. In particular, thimbles that flow to infinity in the imaginary direction coincide with run-away solutions of the classical CL equations. We saw (Cap.2), however, that those kind of trajectories can always be avoided in the stochastic process (i.e. with noise η) by an appropriate choice of the CL integration stepsize.

Another obvious difference is that CL dynamics takes place in the whole complexified manifold, in this case the complex plane, doubling the degrees of freedom of the original problem. The Lefschetz thimbles, on the other hand, are a complex sub-manifold of the complexified domain, but they have the same dimension of the original domain of integration.

Those two points combined together also imply that the critical points, where $\partial_z S(\mathbf{z})|_{\mathbf{z}=\mathbf{z}_k} = 0$, have different properties in the two approaches. From the point of view of the Lefschetz thimble, a non-degenerate critical point *has* to sit on a saddle point because, from (4.4), it has to be a maximum for the stable thimble and a minimum for the unstable one.

i.e. the Hessian matrix in these points $\partial_z^2 S(z)|_{\mathbf{z}=\mathbf{z}_k^*}$ can be diagonalised and its eigenvalues are half positive and half negative. This property is essential to allow both the stable and the unstable thimble to be associated with these points.

For complex Langevin dynamics, instead, each critical point is either a relative maximum or a relative minimum. That means that the drift is either attractive or

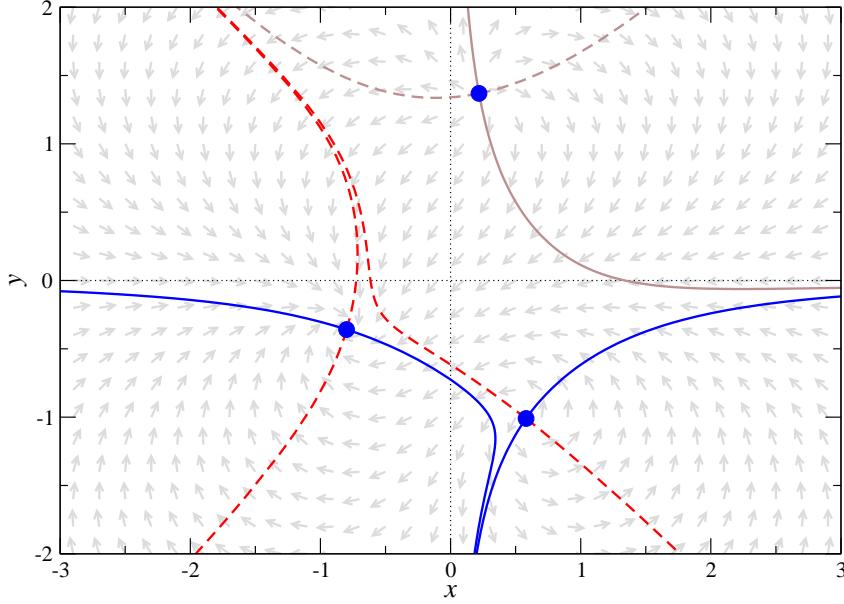


Figure 4.1: Thimbles and Langevin flow in the quartic model with $\sigma = 1$ and $h = 1 + i$: the blue circles denote the fixed points, the (normalised) arrows the classical Langevin drift, and full (dashed) lines the stable (unstable) thimbles. The two blue thimbles contribute. The third fixed point does not contribute.

repulsive. In Fig.4.1 we show both the Lefschetz thimbles (lines) and CL classical drift (grey arrows), while the blue dots are the three critical points

$$\begin{aligned} z_k &= e^{2\pi ik/3} D - e^{-2\pi ik/3} \frac{\sigma}{3D}, & k &= 0, 1, 2, \\ D &= \left(-\frac{h}{2} + \frac{h}{2} \sqrt{1 + \frac{4\sigma^3}{27h^2}} \right)^{1/3}. \end{aligned} \quad (4.17)$$

From the Lefschetz thimble point of view each of the three points have one stable and one unstable direction. For CL, on the other hand, only the first one is attractive while the other two are repulsive (as one can see in Fig.4.1).

Here we can see that the two thimble contributing, i.e. the ones whose connected unstable thimbles (red dashed lines) intersect the real axes, are clearly a deformation of the \mathbb{R} contour with the limit at $x \rightarrow \pm\infty$ fixed. That fact that the two thimbles go to infinity in the y direction should not worry since in that region of the complex plane the integrand vanishes. One can see this by noticing that for large $z = re^{i\theta}$ the dominant contribution to $\text{Re}S$ comes from the z^4 term. In particular, $\text{Re}S$ is positive in the regions $\cos(4\theta) > 0$, i.e. where $-\frac{\pi}{8} + k\frac{\pi}{2} < \theta < \frac{\pi}{8} + k\frac{\pi}{2}$, which, for $k = 3$ that is exactly the region where the two thimbles go to infinity. This is general, i.e. the stable thimbles always in the region where $\text{Re}S \rightarrow +\infty$, otherwise the integral on it would not be finite, while the unstable ones always in the regions where $\text{Re}S \rightarrow -\infty$.

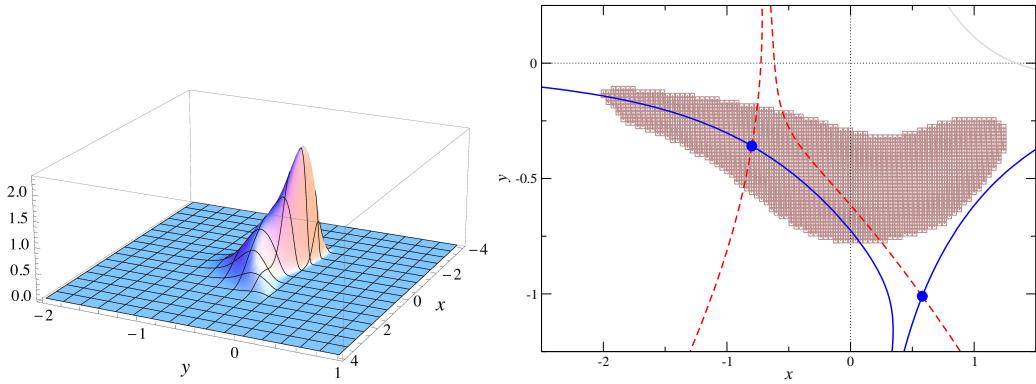


Figure 4.2: Quartic model, with $\sigma = 1$ and $h = 1 + i$. Histogram collected during a complex Langevin simulation (left) and a comparison with the thimbles (right).

For what concerns CL, this model has been studied in detail in [57] in the case of $\sigma \in \mathbb{C}$ and $h = 0$. For the purpose of this section we just have to notice that two of the three critical points are repulsive, so that we expect CL dynamics to avoid the first two and concentrate somewhere near the third one. Furthermore, as mentioned earlier, thimbles that flows to $\pm\infty$ in the imaginary direction, correspond to runaway solution of the classical CL drift due to the complex conjugate equation of motion in the two models. Those trajectories, however, are avoided when the process is made stochastic. For the same reason, when the thimbles are parallel to the real axis, their attractive flow coincide with the CL one.

We have now all the instruments to understand Fig.(4.2). It shows Lefschetz thimbles in the complex plane together with the history of CL dynamics. We can clearly see that CL tends to follow the thimbles when they flow mainly in the real direction, while it strongly departs from them when they start to pick a relevant imaginary component. We can also observe the expected cluster of CL dynamics around the only attractive point (see histogram in the left part of Fig.(4.2)) while, of course, the repulsive ones are avoided.

We have verified that integration along the two contributing thimbles yealds to the right result. In particular $Z_1 = 1.744 + i0.461$ and $Z_2 = 0.021 + i0.426$, which leads to the correct sum $Z = Z_1 + Z_2 = 1.765 + i0.887$. For a numerical comparison with complex Langevin see the table (4.1).

We can conclude this section saying that knowing the Lefschetz thimbles structure of a complexified theory, at least in simple cases, gives very precise information about the CL dynamics distribution in the complex plane. The two methods, however, are far away from being in a one to one relation especially in the presence of a repulsive fixed point.

In the following we will extend this comparison to simple gauge models.

4.3 U(1) model with determinant

It is interesting to study a simple one-link abelian model, with link $U = e^{ix}$, with a QCD-like determinant in the action. This model has been studied in the paper [99] specifically for the complex Langevin dynamics part. Here, instead, our goal is the comparison between CL dynamics and Lefschetz thimbles. The partition function of this theory is

$$Z = \int_{\mathrm{U}(1)} dU e^{-S_B} \det M = \int_{-\pi}^{\pi} \frac{dx}{2\pi} e^{\beta \cos x} [1 + \kappa \cos(x - i\mu)], \quad (4.18)$$

where β is taken real and positive like the gauge coupling and the complex weight is introduced by the determinant $[\det M(\mu)]^* = \det M(-\mu^*)$, via the chemical potential μ like in QCD. We observe, already at $\mu = 0$, that when $\kappa < 1$ the weight is real and positive, while for $\kappa > 1$ there is already a sign problem. Furthermore, after the exponentiation of $\det M$, the action we are left with is not holomorphic any more,

$$S(z) = -\beta \cos z - \ln [1 + \kappa \cos(z - i\mu)], \quad (4.19)$$

which is reflected in the drift $\partial_z S(z)$ having poles where the determinant $1 + \kappa \cos(z - i\mu)$ is zero :

$$\partial_z S(z) = \beta \sin z + \frac{\kappa \sin(z - i\mu)}{1 + \kappa \cos(z - i\mu)}. \quad (4.20)$$

Violation of holomorphicity needs to be treated very carefully both in complex Langevin dynamics and in the Lefschetz thimble method. In the first case it invalidates the formal proof of convergence and it might lead to convergence to the wrong results [100, 58]. In the thimbles case it might compromise the deformation of the original contour of integration. However, as we shall see, the thimbles usually ends on the poles so that the integral over the contour is not affected by it. Moreover, in the case of multi-branched functions, one thimble is not expected to trespass in a branch different than the one where its critical point lies, as to do so it would need to cross a singularity.

Nevertheless, it is very important to test those two methods, and in general every method that aims to solve the sign problem, in the case where the action of the theory is non-holomorphic, since this is the case for full QCD. In this sense, this simple model is a very useful toy model since it can be solved analytically. In fact, the partition function and every power of the trace of the link $\langle \cos^n(x) \rangle$ are linear combination of modified Bessel functions

$$\begin{aligned} Z &= I_0(\beta) + \kappa I_1(\beta) \cosh \mu, \\ \langle \cos(x) \rangle &= \frac{1}{Z} \left[I_1(\beta) + \frac{1}{2} \kappa \cosh \mu (I_2(\beta) + I_0(\beta)) \right]. \end{aligned} \quad (4.21)$$

The goal will then be confronting our numerical results with the exact ones.

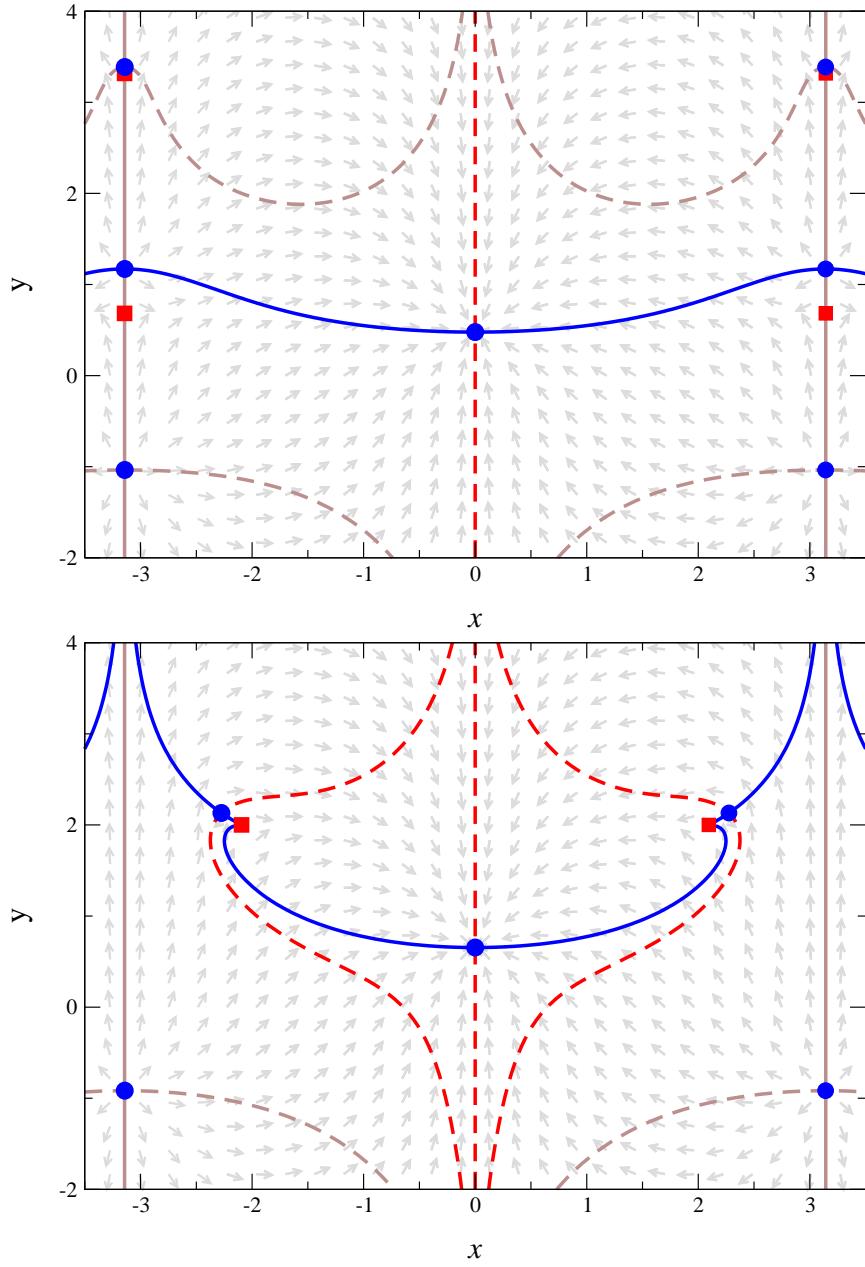


Figure 4.3: Thimbles and Langevin flow in the $U(1)$ one-link model with $\beta = 1$, $\mu = 2$ and $\kappa = 1/2$ (up) and $\kappa = 2$ (down): the blue circles indicate the fixed points, the (normalised) arrows the classical Langevin drift, and full (dashed) lines the stable (unstable) thimbles. Note that the blue lines indicate the contributing stable thimbles, with the red associated unstable ones, while the light brown lines indicate thimbles that do not contribute. The squares indicate where the flow diverges, $\text{Im}S$ jumps, and the direction of the flow along the thimble changes sign. Only the blue thimble(s) contribute.

The critical points of the action are determined by

$$\partial_z S(z) = 0 \quad (4.22)$$

and the singular points by

$$1 + \kappa \cos(z - i\mu) = 0, \quad (4.23)$$

or explicitly

$$\begin{cases} \cos(x) \cosh(y - \mu) = -1/\kappa \\ \sin(x) \sinh(y - \mu) = 0 . \end{cases} \quad (4.24)$$

Although the solutions to (4.22) are quite complicated to solve exactly, one can still see that for $\kappa \leq 1$ the stationary points have to be at $x_c = 0, = \pm\pi$, while for $\kappa > 1$ solutions at different x_c are possible. The same happens for the singular points (4.24). Here we can see explicitly that for $\kappa \leq 1$ only solutions with $x_s = \pm\pi$ and $\cosh(y_s - \mu) = 1/\kappa$ are allowed. Instead, when $\kappa > 1$, the solutions $y_s = \mu$ and $\cos(x_s) = 1/\kappa$ become acceptable as well. The situation is captured in Fig.4.3, where fixed points (singular point) are indicated with blue circles (red squares).

In this case, it is possible to gain some analytical insight about the thimbles. Again we need to separate the case of $\kappa \leq 1$ and $\kappa > 1$.

- $\kappa \leq 1$:

this is the case where critical points are only at $x_c = 0, = \pm\pi$ and poles at $x_c = \pm\pi$. We will see that we can tell which are the relevant thimbles just by looking at the action. Let us start with the critical point in $x_c = 0$. Here we found the action (4.19) along the imaginary direction

$$S(0 + iy) = -[\beta \cosh(y) + \ln(1 + \kappa \cosh(y - \mu))] \quad (4.25)$$

to be completely real and unbounded from below as $y \rightarrow \pm\infty$. This means that $\text{Im}S(0 + iy) = 0$ and the integral is divergent, i.e. the y axis is the unstable thimble. Since the y axis definitely crosses the x axis, we expect the stable thimble associated with $x_c = 0$ to be relevant for the integral.

Similar is the case of the thimble associated to $x_c = \pm\pi$. The action here reads

$$S(\pm\pi + iy) = \beta \cosh(y) - \ln(1 - k \cosh(y - \mu)). \quad (4.26)$$

Since we are in the case of $k \leq 1$, $\text{Im}S(\pm\pi + iy)$ is again constant along the y axis and its value only depends on y . When $\cosh(y - \mu) < 1/\kappa$, in fact, $\text{Im}S = 0$, otherwise $\text{Im}S = \pm\pi$. The interesting fact here is that the value of $\text{Im}S$ jumps in correspondence of the singularities at $x_s = \pm\pi$ and $\cosh(y_s - \mu) = 1/\kappa$.

About the contribution of the thimble : we can see that in the regions $x = \pm\infty, y \rightarrow \pm\infty$, the action $S \rightarrow +\infty$ so that the integral is convergent there and the thimbles are the stable ones. However, because they are exactly parallel to the y axis, they cannot be a deformation of the original real axis, i.e. they cannot contribute to the integral.

- $\kappa > 1$:

this is the case where some of the repulsive fixed points move away from $x = \pm\pi$, while also the singular drift is no longer at $x = \pm\pi$, but instead at $y = \mu$ and $x = x_s$ such that $\cos x_s = -1/\kappa$. Here the stable thimbles that flow to $y \rightarrow +\infty$ are not parallel to the y axis any more. Furthermore, they manage to connect with the stable thimble connected to $x_c = 0$ in the sense that they both end at the singularity, where the imaginary part of the action has a jump. In this way they belong to the deformation of the real axis, so that now they will need to be considered for the integral. Those thimbles, though, are too complicated to be expressed in an analytic form.

This is illustrated in Fig.4.3. In both cases the results, obtained by summing the integral over all the thimbles that contribute, agree with the exact result (4.21), provided both the residual phase factor and the global phase phase factor are correctly incorporated.

For what concerns complex Langevin dynamics the critical point in $x_c = 0$ is always attractive, while the others are repulsive. We can try to guess the CL scatter plot from the thimble structure, the same exercise we did for the quartic model.

- $\kappa \leq 1$:

close the critical point at $x_c = 0$ the thimble is mostly parallel to the x axis, so CL classical drift has the same direction of the thimble flow. Furthermore, the latter approaches the critical point from above, that means the imaginary component of CL drift, i.e. the complex conjugation of the thimble flow equations, will point towards the positive y direction. Then, we expect CL scatter plot to be clustered following the thimble around the attractive fixed point at $x_c = 0$ and, more precisely, to be mostly above it.

The other stable thimbles at $x = \pm\pi$ are parallel to the y axes, which corresponds to run-away solutions for CL. We expect then the latter to avoid those trajectories.

In Fig.4.4 (up) we see that, as expected, the CL dynamics stays over the main thimble around the fixed point at $x_c = 0$. Moreover the compactness of its distribution and the fact that it stays away from the poles (red squares) leads to the correct results (see Table (4.1)) at least for low moments $\langle \cos^n(x) \rangle$, with small n [99] .

- $\kappa > 1$:

in this case the thimble connected with $x_c = 0$ is much less flat in the y

direction. Also two of the stable thimbles, previously localised at $x = \pm\pi$, move towards the centre and join the main thimble. Altogether this necessarily leads to a much stronger classical complex Langevin drift in the y direction. The resulting scattering plot (Fig.4.4 down) is, therefore, wider spread into the complex plane. Moreover, also the poles are closer to the attractive fixed point and, this time, CL dynamics manages to reach them. In this case CL dynamics is found not to be able to reproduce the correct results (see Table (4.1)) even for the low momenta. Since the poles are immersed in the distribution the arguments in [100] and [54], relying in holomorphicity, break down.

To summarise, we find that in the presence of a determinant the flow has singular points. For Langevin dynamics, this leads to a breakdown of the formal justification and possible wrong results in practice. For the Lefschetz approach, we find that thimbles may end at singular points and the imaginary part of the action jumps by a constant. Hence, we observed for the first time in[98], if there is more than one contributing thimble, they connect either at $|z| \rightarrow \infty$ or at a singularity.

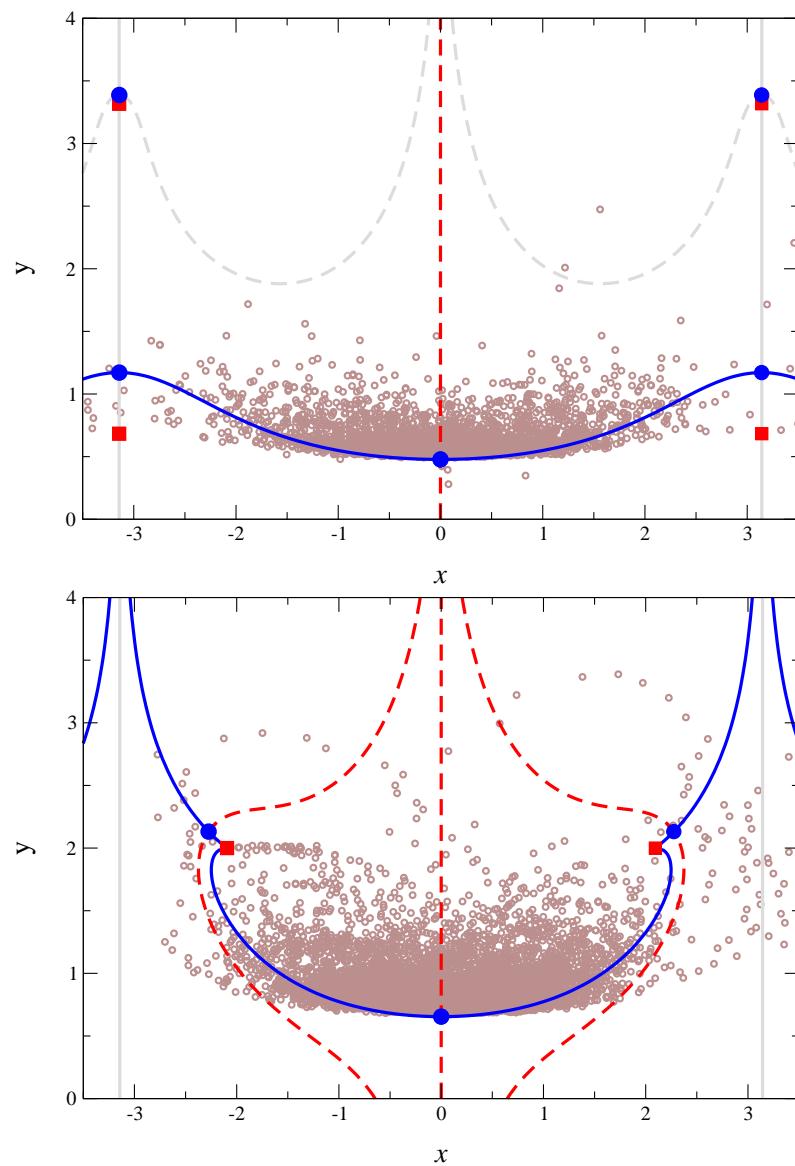


Figure 4.4: As in the previous plot, with scatter data from a complex Langevin simulation added.

4.4 SU(2) 1-link model

We now extend the analysis to the case of a SU(2) one link model,

$$Z = \int dU \exp \left[\frac{\beta}{2} \text{Tr} U \right], \quad \beta \in \mathbb{C}, \quad U \in \text{SU}(2). \quad (4.27)$$

This time the sign problem is introduced by a complex β instead of a determinant. This model is again analytically solvable due to the fact that the action only depends on the trace of the link U . It is, therefore, invariant under any change of basis of $U \rightarrow RUR^{-1}$ and this allow to diagonalize the link making its dependence from an angle ϕ explicit

$$\begin{aligned} U(\phi, \vec{n}) &= e^{i\omega_a \sigma_a} = \cos(\phi) \mathbb{1} + i \sin(\phi) n_a \sigma_a, \\ \text{Tr} U &= 2 \cos \phi, \end{aligned} \quad (4.28)$$

where σ_a are the Pauli matrices, $\phi = \sqrt{\omega_a \omega_a}$ and $n_a = \omega_a / \phi$. For convenience, we will refer to this formulation as the 'angle representation'. In this case one has also to consider the change in the Haar measure for the partition function

$$\int dU \rightarrow \int_{-\pi}^{\pi} d\phi \int \frac{d\Omega(\vec{n})}{4\pi} \sin^2(\phi), \quad (4.29)$$

where $\int \frac{d\Omega(\vec{n})}{4\pi}$ is the integration of the uniform angular measure on the unit sphere, in our case it is equal to 1 because we only have dependence from ϕ . This formulation allows the model to be solved analytically in terms of modified Bessel functions of the first kind

$$\begin{aligned} Z &= \frac{1}{2} [I_0(\beta) - I_2(\beta)], \\ \langle \text{Tr} U \rangle &= 2 \langle \cos(\phi) \rangle = \frac{1}{Z} \frac{4}{\beta} I_2(\beta). \end{aligned} \quad (4.30)$$

Diagonalizing the link U corresponds to a complete gauge fixing in the sense that the degrees of freedom of the model are reduced from three to one. A detailed study of the effect of coordinate changing on complex Langevin dynamics, including this model, has been done in [56]. There it is shown how the introduction of a Haar measure, dictated by the coordinate change, can help the convergence of the dynamics and, therefore, it is a useful tool associated with CL.

The other main way to control CL dynamics is gauge cooling (see cap.2). The latter does not require an explicit gauge fixing and it is employed when the dynamics takes place in the full gauge group.

Although, obviously, in both cases the model is the same, from the perspective of CL dynamics they are completely different. In the first case the space to be explored is the strip : $-\pi < \text{Re}\phi < \pi$, $-\infty < \text{Im}\phi < \infty$ of the complex plane ,

while in the second case it is the complex group $SL(2, \mathbb{C})$. In the following we will compare both approaches with the Lefschetz thimble method. The latter can be computed exactly in the angle representation and directly compared with CL in the same representation. Then, it can be mapped into the $\text{Tr}U$ plane to be confronted with the case in which CL dynamics takes place into the whole $SL(2, \mathbb{C})$ group.

4.4.1 Angle representation

In this approach the partition function (4.27) is written

$$Z = \int_{-\pi}^{\pi} \frac{dx}{2\pi} \sin^2 x e^{\beta \cos x} = \int_{-\pi}^{\pi} \frac{dx}{2\pi} e^{-S(x)}, \quad (4.31)$$

with

$$S(z) = -\beta \cos z - 2 \ln \sin z. \quad (4.32)$$

Similar to the determinant for the $U(1)$ model, the Haar measure introduces a logarithm in the action which makes it non-holomorphic. As usual we need to identify the critical points and the poles of the drift

$$-\partial_z S(z) = -\beta \sin z + 2 \frac{\cos z}{\sin z}. \quad (4.33)$$

Hence, the fixed points are

$$\cos z_c^\pm = -\frac{1}{\beta} \left(1 \pm \sqrt{1 + \beta^2} \right), \quad (4.34)$$

while the poles satisfy

$$\sin x \cosh y + i \sinh y \cos x = 0, \quad (4.35)$$

which are the points $x_s = 0, \pm\pi$ and $y_s = 0$.

If we integrate numerically the thimble equations we obtain the picture in Fig.4.5 (left). The plot respects the parity symmetry and, as expected, the stable thimbles contributing (solid blue lines) are the ones that form a deformation of the original contour $x \in \mathbb{R}, x \in (-\pi, \pi]$. We can also map the results into the $\text{Tr}U = 2 \cos z$ plane, see Fig.4.5 (right) to compare it with complex Langevin dynamics in the gauge formulation.

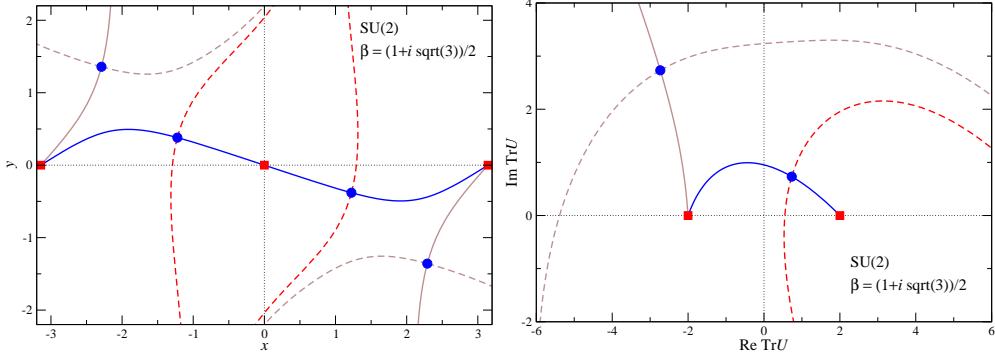


Figure 4.5: SU(2) one-link model, with $\beta = (1 + i\sqrt{3})/2$: thimbles in the xy plane (left) and the $\text{Tr}U$ plane (right). Blue circles indicate fixed points, full (dashed) lines the stable (unstable) thimbles, and red squares the singular points. The blue thimble contributes.

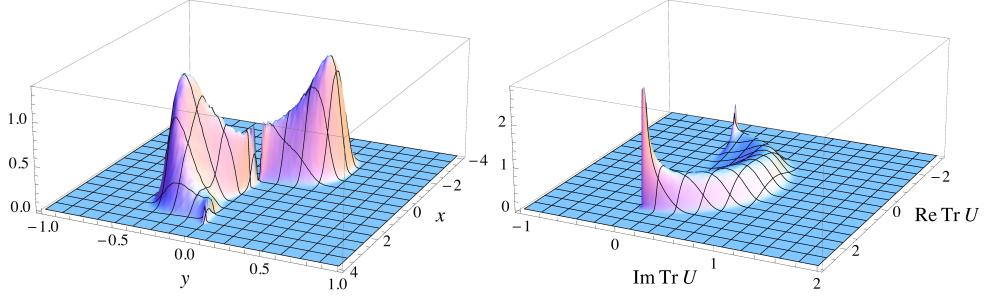


Figure 4.6: Histograms collected during a complex Langevin simulation in the gauge fixed formulation of the SU(2) one-link model, with $\beta = (1 + i\sqrt{3})/2$, in the xy plane (left) and the $\text{Tr}U$ plane (right).

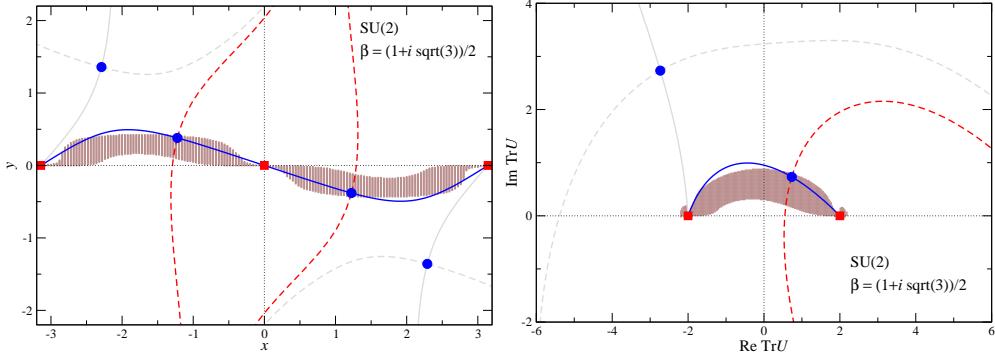


Figure 4.7: Comparison between the complex Langevin histograms and the thimbles, in the xy plane (left) and the $\text{Tr}U$ plane (right).

In Fig.4.7 we show a comparison between thimbles and the histogram of CL dynamics. We notice that CL stays very close to the thimble lines. Once more, we can appreciate how the complex conjugate drift leads CL dynamics to stay mostly below the thimble when this is concave, and mostly above when this is convex. Also they are connected with two repulsive fixed points, from the point of view of CL.

The two non-contributing thimble are ignored by CL. That is because they are far away from the two central attractive points and mostly oriented in the y direction, i.e. they coincide with run-away trajectories avoided by CL.

We can clearly see how the distribution of CL dynamics into the complex plane is compact. Furthermore we can see from Fig.4.5(left) that the poles are avoided in the angle representation, while in the matrix representation for CL (right) the action has no poles to begin with. This, as expected, is reflected in the convergence to the right results (as we can see in Table (4.1)).

4.4.2 Gauge dynamics with cooling

As anticipated, we will study the same model but from a different approach for what concerns complex Langevin dynamics. This time we will not impose any gauge fixing and let CL free to explore the $SL(2, \mathbb{C})$ space. We will use, though, gauge cooling to control the dynamics.

This approach is more expensive in terms of computing time, but has the advantage of being more easily generalised to the 4d case. Also if the gauge is not fixed, the Haar measure is not needed any more and the action is holomorphic.

Complex Langevin dynamics for gauge theories have been reviewed in Cap.2. Here we will just briefly recall the evolution equation

$$U(t + \epsilon) = R(t)U(t), \quad R = \exp [i\sigma_a (\epsilon K_a + \sqrt{\epsilon}\eta_a)], \quad (4.36)$$

where t is the (discretised) Langevin time, $K_a = -D_a S$ is the drift and σ_a are the Pauli matrices. Application of *gauge cooling* is necessary to control the width of the distribution in the complex direction. This process is orthogonal to the dynamics of the system and it can be implemented an arbitrary number of times between two consecutive CL updates. The single steps can be expressed as

$$U'(t) = \Omega_{gc}(\alpha)U(t)\Omega_{gc}^{-1}(\alpha), \quad (4.37)$$

where parameter α regulates the intensity of the cooling. The number of steps is decided in such a way that the balance is optimal between effectiveness of gauge cooling and computer time usage.

We can see in Fig.4.8 that increasing the number of cooling steps progressively improves the compactness of the distribution of $\text{Tr}U$ in the complex plane. However one can clearly see a jump at 2 gauge cooling steps. That results in $\langle \text{Tr}U \rangle_{CL}$ to converge to the right results (see Table (4.1)) for number of gauge cooling steps ≥ 2 .

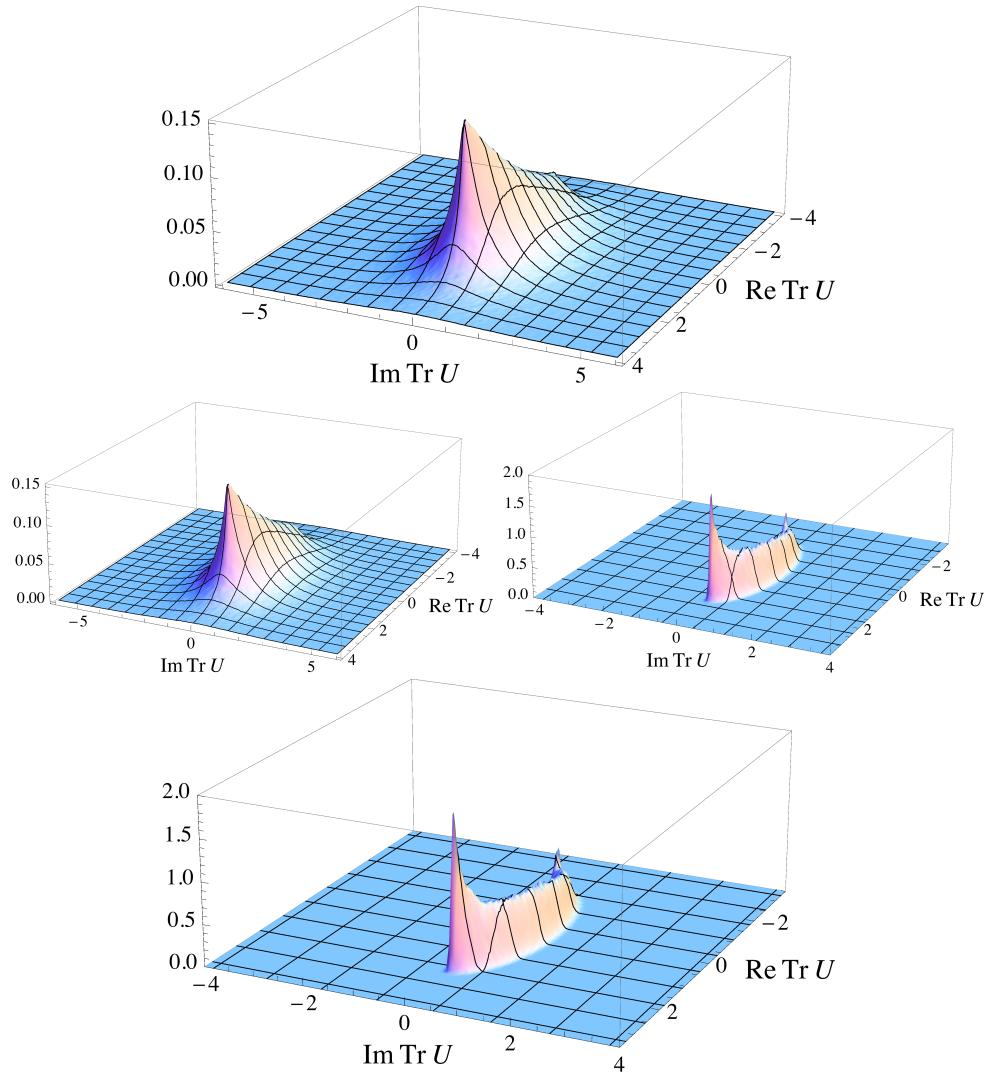


Figure 4.8: Histograms collected during a complex Langevin simulation in the matrix formulation of the SU(2) one-link model at $\beta = (1 + i\sqrt{3})/2$, with gauge cooling, using 0, 1, 2, 4 (from top to bottom) gauge cooling steps.

Lastly we are going to analyse the case when $\beta^2 = -1$ (and hence β is purely imaginary). Here the fixed point z_d , with $\cos z_d = -1/\beta = \pm i$, is degenerate in the sense that the Hessian $|\frac{\partial S}{\partial z^2}|_{z=z_d} = 0$. The choice of imaginary β is motivated by dynamics in real (Minkowskian) time, see e.g. Refs. [101, 102, 65] for complex Langevin studies. For degenerate fixed points the standard reasoning to justify the Lefschetz approach and construct the thimbles by numerical integration is not well defined. Here we give a brief analysis.

At the fixed point the action is real, with $S(z_d) = 1 - \ln 2$. We take $\beta = i$ and write

$$\frac{1}{2}\text{Tr}U = \cos z = u + iv. \quad (4.38)$$

Equating the imaginary part of the action,

$$\text{Im}S = -u - \phi, \quad \tan \phi = \frac{-2uv}{1 - u^2 + v^2}, \quad (4.39)$$

to 0, then yields the thimbles, and we find

$$v_{\pm}(u) = \frac{1}{\tan u} \left(u \pm \sqrt{u^2 - (1 - u^2) \tan^2 u} \right), \quad (4.40)$$

where $0 < u < 1$ for $v_-(u)$ and $-1 < u < 0$ for $v_+(u)$. These two branches make up the stable thimble. The unstable thimble is given by $u = 0$, for which the imaginary part of the action vanishes as well.

The thimbles are shown in Fig. 4.9 (right), using the same colour coding as above. They are a deformation of the thimbles for general complex β , shown earlier, satisfying reflection symmetry in the $\text{Re}\text{Tr}U = 0$ axis. We have verified that integrating along the thimble, with the inclusion of the residual phase, gives the correct answer. The Langevin histogram is shown in Fig. 4.9 (left). A similar histogram was obtained earlier in Ref. [65]. A comparison between the histogram and the thimbles is finally given in Fig. 4.9 (right). For this case we note that the distribution does not overlap substantially with the thimble and that again the distribution of the weight for both approaches is quite different.

In conclusion the aim of this chapter was to compare two of the most promising methods to deal with the sign problem in Lattice quantum field theory, i.e. complex Langevin dynamics and Lefschetz thimble. We examined several models that present a sign problem. Those were simple enough to be solved analytically but interesting enough to capture some important characteristics of more complicated gauge theories. We started with a quartic model to fix the ideas about Lefschetz thimbles and to set the parameters of comparison between the latter and complex Langevin

Model	Observ.	Lefsc. thimbles	complex Lang.	Exact result
quartic mod.	$\langle z \rangle$	-0.501475 - i 0.420789	-0.4993(7) - i 0.4174(2)	-0.501475 - i 0.420789
$U(1)$ ($\kappa = 0.5$)	$\langle \cos z \rangle$	0.808709	0.814(5) - i 0.002(5)	0.808709
$U(1)$ ($\kappa = 2$)	$\langle \cos z \rangle$	1.05808	1.019(7) + i 0.021(9)	1.05808
$SU(2)$ (gauge fix.)	$\langle \cos z \rangle$	0.271435 + i 0.43181	0.272(1) + i 0.4319(3)	0.271435 + i 0.43181
$SU(2)$ (gauge cool.)	$\langle \cos z \rangle$	x	0.272(2) + i 0.4315(5)	0.271435 + i 0.43181
$SU(2)$ ($\beta = i$)	$\langle \cos z \rangle$	i 0.5222229	0.001(1) + i 0.5222(4)	i 0.5222229

Table 4.1: Comparison, between CL, Lefschetz thimble and exact result, of the numerical values for the models discussed in this chapter. The 'x' symbol, in the $SU(2)$ (gauge cool.) line, means we didn't perform the Lefschetz thimble in the matrix representation, having the result already in the gauge fixed representation.

dynamics. We proceeded by examining a one-link $U(1)$ model with a determinant. Here the sign problem was introduced by chemical potential μ , in analogy with full QCD.

Lastly we looked at the case of one-link $SU(2)$ model with complex β parameter. This model could be approached in two different ways by CL. One way was to fix the gauge and study the dynamics in the lower dimensional parameter space. The other was to let CL free to explore all the complexified gauge degrees of freedom and control the dynamics with gauge cooling. Both approaches have been confronted with the thimbles method and found to converge to the right results.

About the comparison of the two methods, we found that the study of the Lefschetz thimbles is strictly connected with CL dynamics. The reason for this is that the two share the same critical points and their equation of motion are the complex conjugate of the other. Knowing the thimbles structures, therefore, allows one to have the exact information of the classical CL drift along the thimbles lines connecting the fixed points.

The fixed point in the Lefschetz thimble formulation have always an attractive direction and a repulsive one. On the other hand, for CL they can be completely attractive or completely repulsive. However, knowing the thimbles flow near the fixed points allows to establish exactly what kind of fixed point they are in the CL formulation.

We also showed, when there are poles involved, that they are the endpoint of some thimbles while, for CL dynamics, they are quite dangerous and threaten the good convergence whenever they are immersed into the dynamics.

Altogether we show that, starting from the Lefschetz thimble picture, one can make prediction on the CL dynamics with some accuracy.

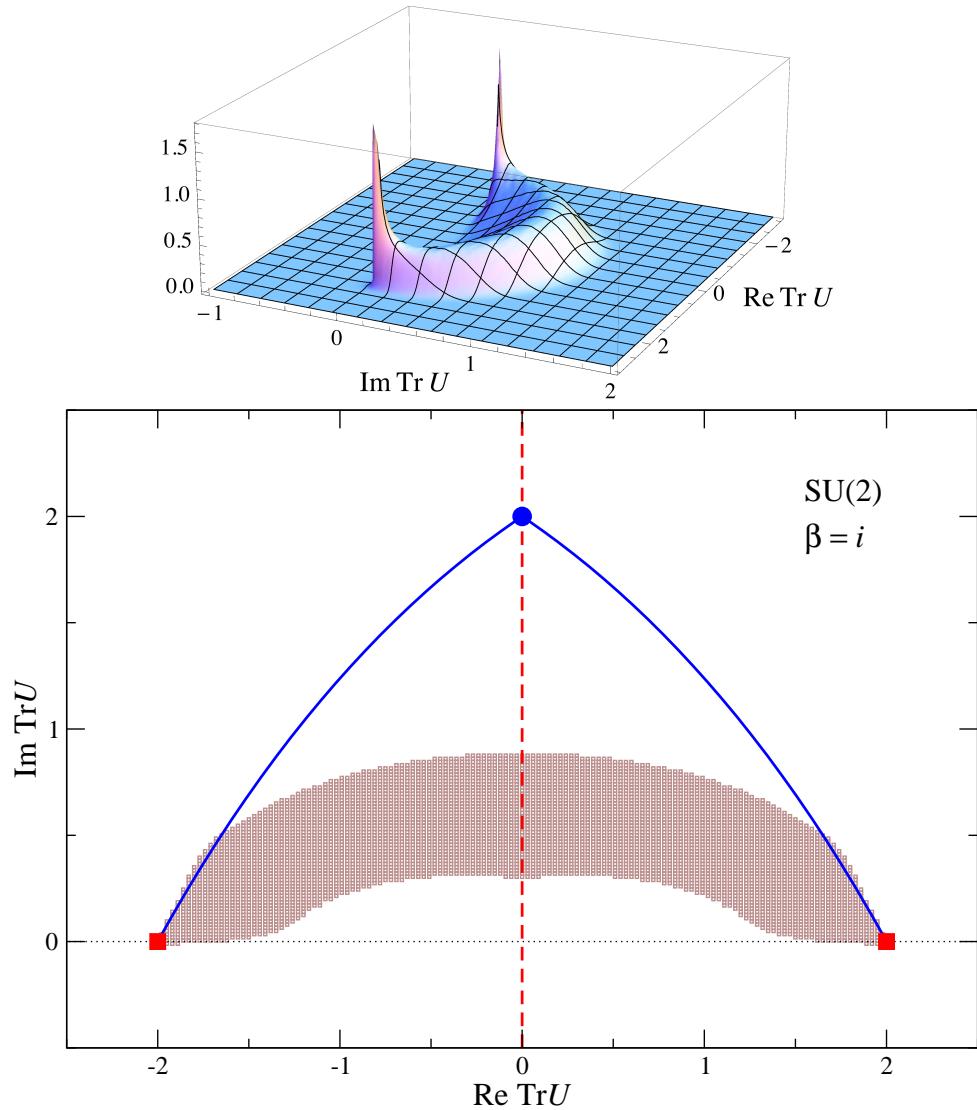


Figure 4.9: $SU(2)$ one-link model at $\beta = i$. Histogram collected during a complex Langevin simulation in the $\text{Tr } U$ plane (left) and a comparison with the thimbles associated with the degenerate fixed point at $\text{Tr } U = 2i$ (right).

Chapter 5

Density of States

In the previous chapters we have discussed complex Langevin dynamics and Lefschetz thimbles as methods to deal with the sign problem in quantum field theory. In this last chapter we are going to review another promising approach to the same problem, i.e. the *density of states* (DOS) approach. The idea is to construct the density of states of a given theory in the canonical, or grand-canonical, fashion. An efficient way has been developed in [103] for discrete systems, and then generalised to quantum field theories with continuous energy spectrum [104, 24, 25, 105, 106, 108, 109, 107]. Standard Monte-Carlo methods are highly developed tools when it comes to measure expectation values of observables on theories with a real and positive Boltzmann weight e^{-S} . However, they do not allow direct computation of some extensive quantities such as the free energy or the partition function. Furthermore, as stressed many times before, they become inefficient when the theory is affected by a sign problem.

On the other hand, the density of state method is naturally suitable for calculating extensive quantities. Moreover, as we shall see in the following, it automatically reduces the sign problem of the full quantum field theory to one in a one-dimensional oscillating integral.

5.1 Description of the method

Let us consider the partition function

$$Z(\beta) = \int \mathcal{D}\phi \exp(-\beta S[\phi]), \quad (5.1)$$

where ϕ is a generic field. As always, we can identify the Euclidean action $S[\phi]$ with the 'energy' of the field configuration. In analogy with the canonical ensemble in

Statistical Mechanics one can define the number of states having energy E

$$\rho(E) = \int [\mathcal{D}\phi] \delta(S[\phi] - E), \quad (5.2)$$

where $\delta()$ is the Dirac delta function. The partition function can then be expressed as the integral of the density over the energy, weighted with the Boltzmann factor :

$$Z(\beta) = \int dE \rho(E) e^{-\beta E}. \quad (5.3)$$

Any expectation value that can be expressed as an explicit function of the energy, can also be computed

$$\langle O(E) \rangle = \frac{1}{Z} \int dE O(E) \rho(E) e^{-\beta E}. \quad (5.4)$$

In practice it is not possible to know $\rho(E)$ for every value of E in a continuum system. Therefore, in the simulations, we will divide the energy spectrum in small intervals $[E_0 - \Delta, E_0 + \Delta]$ of width 2Δ , with E_0 the central value of. Moreover, we would like to have an analytic expression for $\rho(E)$ inside each interval so that we can employ a piecewise definition of the density of states

$$\hat{\rho}(E) = \sum_i \rho_i(E), \quad (5.5)$$

where ρ_i is defined only inside the interval $[E_i - \Delta, E_i + \Delta]$.

In principle, if the intervals are small enough, it is perfectly acceptable to replace a regular function with its linear approximation, up to order Δ^2 . However, we know that $\rho(E)$ can vary over many orders of magnitude depending on the range of energies considered. It is more convenient, then, to linearly approximate the natural logarithm of the DOS, in each interval, and then exponentiate the piecewise function to construct $\rho(E)$.

The piecewise defined logarithm has the form

$$\ln \rho(E) = C(E_i) + a(E_i)(E - E_i) + \mathcal{O}(\Delta^2), \quad E \in [E_i - \Delta, E_i + \Delta] \quad (5.6)$$

so that the DOS can be written as

$$\rho(E) = C_{exp}(E_i) e^{a(E_i)(E - E_i)} + \mathcal{O}(\Delta^2), \quad E \in [E_i - \Delta, E_i + \Delta]. \quad (5.7)$$

The parameter $a(E_i)$, the slope of the linear approximation of $\ln \rho(E)$, is the core focus of this approach and is the only quantity that has to be numerically computed. We shall discuss it in more detail below. The factor $C(E_i)$ is a constant in each interval, whose function is to set $\ln \rho(E)$ at the right value in each interval.

5.1.1 LLR algorithm

We shall now describe the LLR (Langfeld-Lucini-Rago or linear local relaxation) algorithm. It has specifically been created in order to compute the density of states of systems with a continuous energy spectrum.

As mention before, one is interested in computing the quantity $a(E_i)$ in the interval $[E_i - \Delta, E_i + \Delta]$. The way to achieve this is to insert the inverse of (5.7) into the partition function, and tune the value of $a(E_i)$ to $a^*(E_i)$ in such a way that the density of states is annihilated in the interval, i.e.

$$\rho_{exact}(E)e^{-a^*(E_i)E} = 1, \quad (5.8)$$

up to order Δ^2 . This automatically gives an estimate of the number of states itself

$$e^{a^*(E_i)E} = \rho_{est}(E), \quad E \in [E_i - \Delta, E_i + \Delta]. \quad (5.9)$$

To carry out this process, we need to introduce a probe function $f(E)$, of which the expectation value can be measured in the interval

$$F_i(a) = \langle\langle f(E) \rangle\rangle_{i,a} = \frac{1}{Z_a(E_i)} \int_{E_i-\Delta}^{E_i+\Delta} dE f(E) \rho(E) e^{-aE}, \quad (5.10)$$

where $\langle\langle \dots \rangle\rangle_{i,a}$ means the average value in the interval E_i at fixed parameter a , and $Z_a(E_i)$ is the partition function restricted to the same interval which also depends on the parameter a

$$Z_a(E_i) = \int_{E_i-\Delta}^{E_i+\Delta} dE \rho(E) e^{-aE}. \quad (5.11)$$

Let us note that $Z_a(E_i)$ is just the normalization in the Monte Carlo restricted to the intervals, and has nothing to do with the partition function of the system which is defined in (5.3).

At this point, the idea is to choose $f(E)$ in such a way that its expectation value is a good indicator to establish when $\rho(E)e^{-a^*(E_i)E} = 1$. The simplest and more efficient choice is the energy itself or, more precisely, $f(E) = E - E_i$

$$\delta E(a(E_i)) = \langle\langle E - E_i \rangle\rangle_a = \frac{1}{Z_a(E_i)} \int_{E_i-\Delta}^{E_i+\Delta} dE (E - E_i) \rho(E) e^{-aE}. \quad (5.12)$$

From the point of view of the algorithm, we compute $\delta E(a(E_i))$ with a standard Monte Carlo in each interval $[E_i - \Delta, E_i + \Delta]$. Every time, before proposing a new configuration to the Monte Carlo, we check if the energy E_{new} of the configuration is inside the interval $[E_i - \Delta, E_i + \Delta]$ and if it is not, we reject the configuration. A slightly better implementation, that is not manifestly non ergodic, is to consider 'gaussian tails' attached to the extremities of each interval. In this way a configuration, which energy is out of the given interval, can still be proposed to the Monte

Carlo with a probability of $\exp[E_{new} - (E_i \pm \Delta)]^2$, where the \pm depends whether E_{new} is bigger than $E_i + \Delta$ or lower than $E_i - \Delta$.

Going back to the description of the LLR method, one can see that, when (5.8) is satisfied, equation (5.10) implies that

$$\delta E(a^*(E_i)) = 0. \quad (5.13)$$

We are, then, looking for an iterative process that changes $a(E_i)$ towards the zero of $\delta E(a)$. This is a classic one-dimensional root finding problem, with the complication that the value of our function, the energy, is known as a result of a Monte-Carlo simulation, i.e. it cannot be computed directly but only estimated via noisy observation.

Stochastic approximation methods are algorithms developed precisely for this purpose. The Robbins-Monro algorithm [110] is the first of these class, and it is a simple adaptation of the famous Newton-Raphson root finding method for deterministic functions. According to the Robbins-Monro method, it is possible to achieve convergence in the root finding of a noisy function.

The $n + 1$ iteration of Robbins-Monro can be written as

$$a_{n+1} = a_n - d_n \delta E(a_n), \quad (5.14)$$

where the coefficient d_n must obey the constraints

$$\begin{aligned} \sum_{n=0}^{\infty} d_n &= \infty, \\ \sum_{n=0}^{\infty} d_n^2 &< \infty. \end{aligned} \quad (5.15)$$

It is easy to see that any series

$$\frac{d}{\sqrt{n}} < d_n \leq \frac{d}{n}, \quad (5.16)$$

satisfy the requirements (5.15), if d is a finite constant. The faster is the decay of d_n the faster the convergence [111]

$$\lim_{n \rightarrow \infty} a_n = a^*. \quad (5.17)$$

For this reason we might as well choose the fastest allowed

$$d_n = \frac{d}{n}. \quad (5.18)$$

Qualitatively speaking, the coefficient d_n has the function to kill the statistic fluctuations introduced by the noise. Furthermore, it is possible to prove that $\sqrt{n}(a_n - a^*)$ is asymptotically normal with variance

$$\sigma_a^2 = \frac{d^2 \sigma_x^2}{2dF'(a^*) - 1}, \quad (5.19)$$

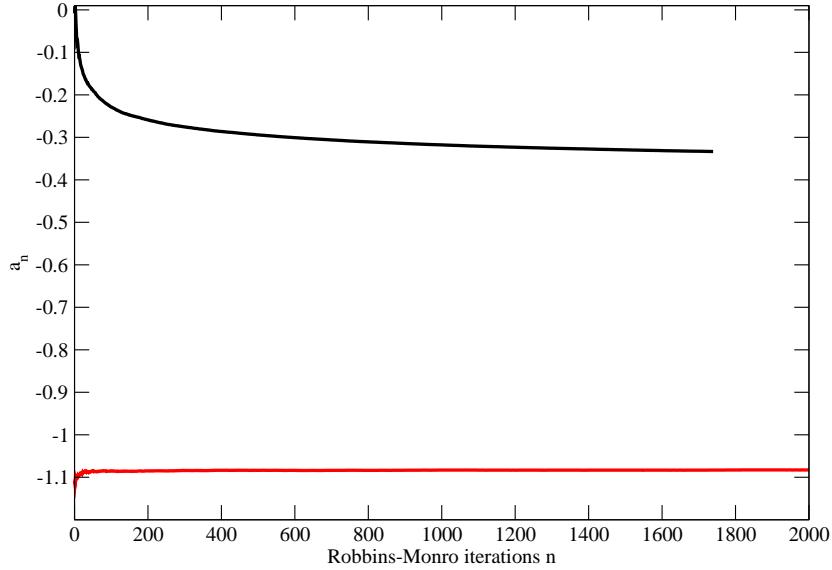


Figure 5.1: An example demonstrating the effect of the thermalization in the Robbins-Monro root finding method. In both cases the initial value is $a_0 = 0$. The black line represents the application of Robbins-Monro from the beginning, while the red line has been thermalized with 30 steps of Newton-Raphson.

where σ_χ^2 is the variance of the noise. The optimal value for d can be found simply by minimizing the variance $\frac{\partial^2 \sigma_a^2}{\partial d} = 0$, that leads to

$$d = \frac{1}{F'(a^*)}. \quad (5.20)$$

In our case, from (5.12), we found

$$F'(a^*) = \frac{\partial}{\partial a} \delta E(a)|_{a=a^*} = \frac{\Delta^2}{3}, \quad (5.21)$$

so that the explicit iteration of the Robbins-Monro reads

$$a_{n+1} = a_n - \frac{3}{(n+1)\Delta^2} \delta E(a_n). \quad (5.22)$$

Let us conclude this part with an important remark. Even though the proof of convergence of the Robbins-Monro method is valid for any initial value a_0 , in practice starting from a random a_0 would introduce an offset impossible to eliminate due to the fact that at large n the contribution of d_n is suppressed. The solution of this problem is introducing some steps of *thermalization* for the a_n . For example, a simple way we adopted is to start with the Newton-Raphson algorithm, i.e. $d_n = d$. In this case the a_n are rapidly brought around the value a^* since there is no suppression of the coefficient d_n . Once the a_n fluctuate around a fixed value then one can switch

to the Robbins-Monro method. An example of what we discussed can be seen in Fig.5.1. Clearly starting from a value a_0 far away from a^* prevents the Robbins-Monro to reach the right result in a sensible time (black line), even though the correct asymptotic behaviour should still be reached. Let us note that the number of iterations in the Fig.5.1 is already two thousands. On the other hand, if we start with just 30 steps of Newton-Raphson as thermalization, the Robbins-Monro stabilizes very fast on the asymptotic value of a^* (red line), which yields for the correct results.

Once the $a^*(E_i)$ are known for a range of energies large enough, the density of states can be build and the other observables, depending on E , can be easily computed.

Clearly, this all procedures requires $\rho(E)$ to refer to a real weight e^{-S} for (5.12) to be able to be computed with a Monte Carlo. As we shall see in the following section, the way to extend the method to the complex case is to separate the imaginary part of the action from the real one, and compute the DOS relative only to the imaginary part.

5.2 Relativistic Bose gas at finite chemical potential μ

In this section we are going to discuss the application of the LLR method to the case of a four dimensional complex scalar field theory at finite density, i.e. the Bose gas. A finite real chemical potential introduces a sign problem in this theory. This model has been extensively studied in literature with different approaches like complex Langevin dynamics [112, 15], Lefschetz thimble [21], dual formulation on flux tubes [114, 115, 116] and extended mean field theory [113].

Because it is one of the simplest quantum field theories and because of the great amount of results present in the literature, the relativistic Bose gas at finite density is one of the best models to test new algorithms that deal with sign problem.

We will start by recalling the Euclidian action in the continuum

$$S = \int d^4x \left[(\partial_4 - \mu)\phi^\dagger(\partial_4 + \mu)\phi + \partial_i\phi^\dagger\partial^i\phi + m^2|\phi|^2 + \lambda|\phi|^4 \right], \quad (5.23)$$

from which it is clear how the chemical potential enters as an imaginary temporal component of a gauge field A_0 . This fact is important when it comes to the discretization on the lattice.

The (5.23) can be written as

$$S = \int d^4x \left[\partial_\nu\phi^\dagger\partial^\nu\phi + (m^2 - \mu^2)|\phi|^2 + \lambda|\phi|^4 + \mu(\phi^\dagger(\partial_4\phi) - (\partial_4\phi^\dagger)\phi) \right], \quad (5.24)$$

where the chemical potential explicitly couples to the conserved charge $j_4 = \phi^\dagger \partial_4 \phi - \partial_4 \phi^\dagger \phi$ associated to the global U(1) symmetry. Furthermore, it determines a shift in the mass parameter which, for finite λ and μ large enough, is responsible for the symmetry broken phase (Mexican hat potential).

However, the naive discretization of (5.24) would not work because it loses this characteristic of the chemical potential. Instead, the correct way to discretize the action is to write :

$$S = \sum \left[(2d + m^2) \phi_x^* \phi + \lambda (\phi_x^* \phi)^2 - \sum_{\nu=1}^4 (\phi_x^* e^{-\mu \delta_{\nu,4}} \phi_{x+\hat{\nu}} + \phi_{x+\hat{\nu}}^* e^{\mu \delta_{\nu,4}} \phi_x) \right]. \quad (5.25)$$

The sum is over the sites of an $N_s^3 N_t$ lattice, with periodic boundary conditions in all the directions, $d = 4$ is the number of dimensions, and m is the bare mass. Let us note how the piece $\phi_x^* e^{-\mu} \phi_{x+\hat{4}} + \phi_{x+\hat{4}}^* e^{\mu} \phi_x$ is complex. We can rewrite (5.25) in terms of the real and imaginary part of the field $\phi = \frac{1}{\sqrt{2}}(\phi_1 + i\phi_2)$

$$S = \sum_x \left[\frac{1}{2} (2d + m^2) \phi_{a,x}^2 + \frac{\lambda}{4} \phi_{a,x}^4 - \sum_{i=1}^3 \phi_{a,x} \phi_{a,x+\hat{i}} - \cosh(\mu) \phi_{a,x} \phi_{a,x+\hat{4}} + i \sinh(\mu) \epsilon_{ab} \phi_{a,x} \phi_{b,x+\hat{4}} \right], \quad (5.26)$$

where a summation over repeated indices is intended. The completely antisymmetric tensor ϵ_{ab} follows the prescriptions $\epsilon_{12} = -\epsilon_{21} = 1$ and $\epsilon_{11} = \epsilon_{22} = 0$.

The operator number of particles on the lattice is

$$n_x = \frac{d}{d\mu} \ln Z|_{\mu=0} = (\delta_{ab} \sinh \mu - i \epsilon_{ab} \cosh \mu) \phi_{a,x} \phi_{b,x+\hat{4}}, \quad (5.27)$$

which corresponds to the Noether current at $\mu = 0$.

Since here we are mostly interested in the sign problem, we will fix, once and for all, the parameters $m = \lambda = 1$, and every result we are going to show will follow the same prescription.

5.2.1 Generalised density of states

Our goal is to build the density of states for this theory. First of all we will consider separately real and imaginary part of the action

$$S[\phi] = S_R[\phi, \mu] + i\kappa(\mu) S_I[\phi], \quad (5.28)$$

where

$$S_I[\phi] = \sum_x [\phi_{1,x} \phi_{2,x+\hat{4}} - \phi_{2,x} \phi_{1,x+\hat{4}}]. \quad (5.29)$$

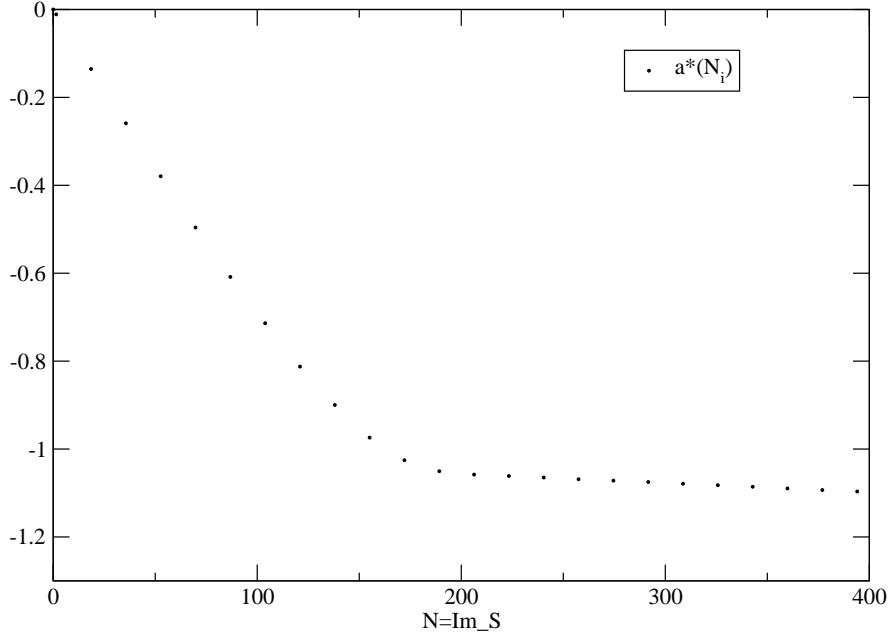


Figure 5.2: Distribution of the $a^*(N)$ for $L = 8^4$ and $\mu = 0.9$.

Then we will generalise the procedures reviewed in Sec.5.1 in analogy with the Grand Canonical ensemble. Let us note that in (5.28) $\kappa(\mu) = \sinh(\mu)$ and $S_R[\phi, \mu]$ has a built-in dependence on $\cosh(\mu)$.

The idea is to elect $S_I[\phi] = N_I$ as 'imaginary' number of particle so that we can write the partition function

$$Z = \sum_{N_I} \int d\phi \delta(S_I[\phi] - N_I) e^{-S_R[\phi, \mu]} e^{-i\kappa(\mu)S_I[\phi]}. \quad (5.30)$$

It is important to note that, in this formulation, what we called N_I is not the actual number of particles, but just the value of the imaginary part of the action. For this reason it will not take integer values and, in fact, can be any number $N_I \in \mathbb{R}$. Furthermore, N_I is an extensive quantity and therefore can be written as $N_I = n_I \Omega$, where Ω is the 4 dimensional volume.

We are, then, able to numerically compute the DOS at fixed N_I

$$\rho(N_I; \mu) = \int d\phi \delta(S_I[\phi] - N_I) e^{-S_R[\phi, \mu]}. \quad (5.31)$$

Since the imaginary number of particle is a continuum, we can adopt for the DOS, the same definition as for the Canonical case (5.7) with the only difference that now the variable is N_I . To ease the notation, from now on we are going to rename $N \equiv N_I$. In complete analogy with the energy case, the logarithm of the DOS is a piecewise function defined in the intervals $N \in [N_i - \Delta, N_i + \Delta]$, where N_i is the centre of the i -th interval.

The partition function is then the Fourier transform of the DOS

$$Z = \int_{-\infty}^{\infty} dN \rho(N; \mu) e^{-i\kappa(\mu)N}. \quad (5.32)$$

The LLR algorithm is also fundamentally the same as the canonical case. As before, we want to numerically compute the $a^*(N_i)$ in (5.9), and the observable we use in the process is the imaginary part of the action $N = S_I[\phi]$

$$\Delta N(a) = \langle \langle N - N_i \rangle \rangle_a = \frac{1}{Z_i} \int_{N_i - \Delta}^{N_i + \Delta} dN (N - N_i) \rho(N; \mu) e^{-aN}. \quad (5.33)$$

As explained before, the root finding process for $a^*(N_i)$ is implemented separately in each interval $[N_i - \Delta, N_i + \Delta]$.

Let us remark that under charge conjugation the imaginary part of the action changes sign $S_I[\phi] \rightarrow -S_I[\phi^\dagger]$ while the real part remains the same, for this reason the density of states has to be *even* in N

$$\rho(N; \mu) = \rho(-N; \mu). \quad (5.34)$$

For the same argument $a(N)$ has to be *odd*

$$a(N) = \frac{d}{dN} \ln \rho(N; \mu) = -a(-N), \quad (5.35)$$

which means that we really need only informations from $N \in [0, \infty]$, and the partition function can be written as

$$Z = 2 \int_0^{\infty} dN \rho(N; \mu) \cos(\kappa(\mu)N). \quad (5.36)$$

In Fig.5.2 we show an example of the typical distribution of

$$a^*(N_i) = \frac{d}{dN} \ln \rho(N; \mu)|_{N=N_i}, \quad \mu = 0.9, \quad (5.37)$$

for the Bose gas. We observe an initial approximately linear slope, which means the dominant contribution to $\rho(N; \mu)$ is gaussian, as we can see integrating (5.37). Such gaussian behaviour for a generalised DOS has been already observed for the \mathbb{Z}_3 model with complex action in [106, 107]. Here, though, we can see the $a^*(N_i)$ suddenly change trend at a larger value of N , and stabilize to a less steep slope.

This behaviour of $a^*(N)$ is quite general, and appears the same at different values of μ and volumes. In fact, it is useful to look, in Fig.5.3, at the dependence of $a^*(N)$ from the chemical potential to gain some insight on the severity of the sign problem. We can see how the initial linear behaviour is common at every chemical potential, but the higher is the value of μ the earlier they will depart from it to settle on the other slope. The latter has, again, the same inclination at every μ and it just differs only for the shift.

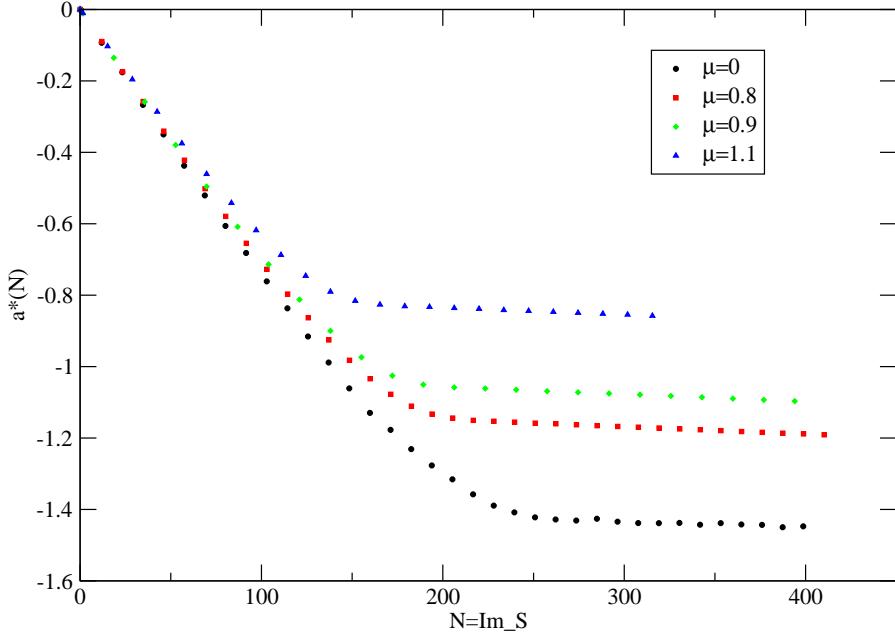


Figure 5.3: Dependence of $a^*(N)$ on the chemical potential μ for $L = 8^4$.

The behaviour in Fig.5.3, together with the obvious increase in frequency of the $\cos(\kappa(\mu)N)$ in (5.36), contribute to the severity of the sign problem. In fact, let us suppose for a moment that the oscillation was constant at every μ

$$Z = \int dN \rho(N; \mu) \cos(\omega N), \quad (5.38)$$

with

$$\frac{d}{d\mu} \omega = 0. \quad (5.39)$$

The DOS at zero chemical potential $\rho(N; \mu = 0)$, generated by the black dots in Fig.5.3, would suppress the oscillation $\cos(\omega N)$ more than $\rho(N; \mu = 1.1)$ (generated by the blue triangles), because it keeps decaying with the fastest slope even after the one at $\mu = 1.1$ is settled to the slower slope. Consequently $Z(\mu = 0)$ would be bigger than $Z(\mu = 1.1)$ even if the oscillation was not increasing with the chemical potential, which means the dependence of the DOS from the chemical potential alone favours the sign problem.

In our case, furthermore, the $\cos(\kappa(\mu)N)$ oscillates faster and faster increasing μ , which obviously results in even more severe cancellations.

5.3 The partition function from highly oscillating integral

In this last section we are going to discuss how extract the value of the partition function knowing $\rho(N; \mu)$. At first sight, it would seem that (5.36) can be easily computed just by using some algorithm that operates a Fourier transform (FT). However this is, unfortunately, not the case because of the statistical noise imprinted in the DOS. The argument for this can be formulated as follows.

We already know that our DOS is a fast decaying function, which is both in $L^1(\mathbb{R})$ and $L^2(\mathbb{R})$, since its main contribution is gaussian. Let us assume that we can consider $\rho(N; \mu)$ to be the sum of the 'true' distribution plus noise. We know (Plancherel theorem) that the Fourier transform \widehat{f} , of an $L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ function f , is again $L^2(\mathbb{R})$, i.e. \widehat{f} is a fast decaying function in our frequency $\kappa = \sinh(\mu)$. On the other hand, the FT of the noise is again noise in κ , which will soon be dominant over the fast decaying signal, and will completely spoil the computation of the partition function. Unfortunately, this happens at relatively small values of μ so that prevents the observation of interesting phenomena like the Silver Blaze.

A way out seems to be fitting the $a^*(N)$. The only requirement is that the fitting function has to be C^∞ . In fact, using a function which presents an arbitrary discontinuity in the derivative would spoil the dependence of the partition function Z from the chemical potential μ . More precisely, a discontinuity in the m -th derivative of a function $f(x)$, dominates the FT: $f(x) = \widehat{f}(k) \sim k^{-m}$ with a polynomial decay.

A good choice for the fitting functions are the polynomials :

$$a^*(N) = \frac{d}{dN} \ln \rho(N; \mu) \simeq \sum_{i=0}^m c_{2i+1} N^{2i+1}, \quad (5.40)$$

where only odd powers are considered since we know that $a^*(N)$ is odd in N . We can see in Fig.5.4 that a polynomial of order seven $p(x) = ax + bx^3 + cx^5 + dx^7$, is already good enough to fit the data very well. In practise, we use the Bayesian evidence for model comparison to determine the best polynomial fit that avoids over-fitting [117, 118, 119].

Proceeding like this, we are able to shift the noise in the coefficients c_{2i+1} in such a way that it does not affect the Fourier transform of

$$\rho(N; \mu) = \exp \left(\sum_{i=0}^m \frac{c_{2i+1}(\mu)}{2i+2} N^{2i+2} \right), \quad (5.41)$$

which is, in fact, an $L^2(\mathbb{R})$ function. The noise appears in the result using some error propagating method (like Bootstrap or Jackknife).

So, after the fit, we are left with the partition function

$$Z = 2 \int_0^\infty dN \exp \left(\sum_{i=0}^m \frac{c_{2i+1}(\mu)}{2i+2} N^{2i+2} \right) \cos(\kappa(\mu)N), \quad (5.42)$$

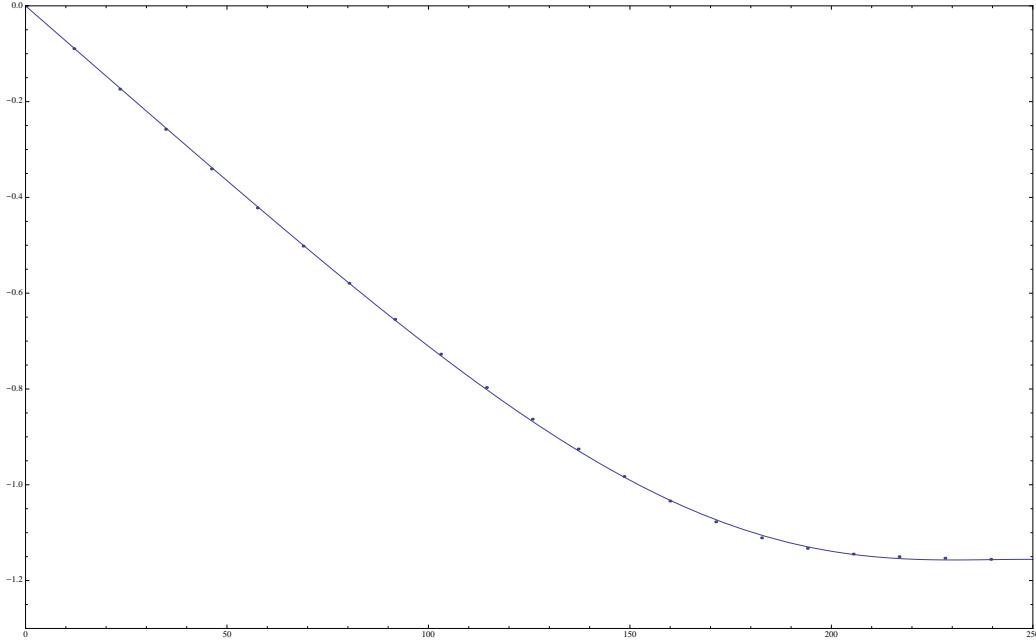


Figure 5.4: Fit of $a^*(N)$ with an odd polynomial of order 7 $a(N) = c_1N + c_3N^3 + c_5N^5 + c_7N^7$, for $\mu = 0.8$ and $L = 8^4$.

which is an highly oscillating integral.

Problems of this kind are treated in mathematical literature (see [120, 121]), especially when the integral is restricted in an interval. In our specific case, we would like to adopt and confront two different methods :

- multi-precision numerical integration
- Lefschetz thimbles method.

The first method is direct and consists in computing the FT in (5.42) numerically. However, due to the cancellation in the integral, one has to be able to know the value of the function at any given point with a precision at least of the same order of magnitude of Z . The problem here is that not only the cancellations increase with μ , but they become also exponentially more severe with the volume. The classic argument for this is to look at the average phase factor

$$\langle e^{i\phi} \rangle = \frac{Z(\mu)}{Z_{pq}} \sim \exp \left(-V \frac{f}{T} \right), \quad (5.43)$$

where f is the free energy density and Z_{pq} is the phase quenched partition function

$$\begin{aligned} weight &= |weight| e^{i\phi} \\ Z_{pq} &= \int dN |weight|. \end{aligned} \quad (5.44)$$

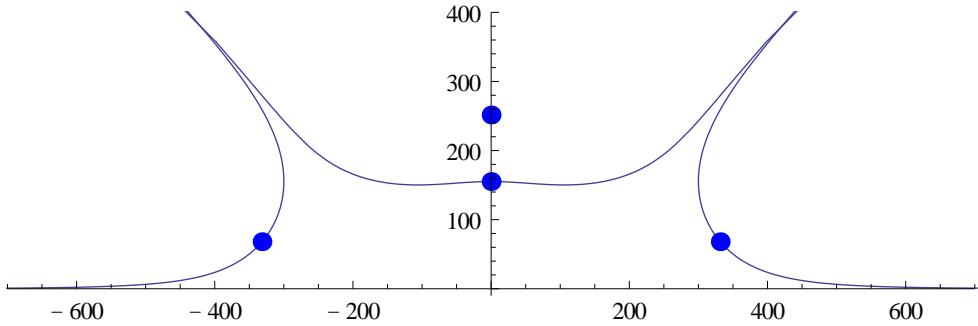


Figure 5.5: Contributing Lefschetz thimbles for order seven polynomial fit of the $a^*(N)$.

If the precision we need to computing the integral (5.42) numerically grows exponentially with the volume of our lattice, it means that multi-precision will eventually not be enough. To give an idea, already at volume 8^4 and $\mu = 1$ the precision required is more or less 50 digits.

This method, then, do not represent an efficient solution to our problem per se, but it will be very useful to be compared with the other two at small volumes.

5.3.1 Lefschetz thimble

Another very successful way to deal with oscillating integral is the Lefschetz thimbles method. We already described it in detail in Cap.4, and here we will recall just the very main steps.

After having performed the polynomial fit and, therefore, having an explicit form for (5.41), we can define the action-like function

$$S[N] = - \sum_{i=0}^m \frac{c_{2i+1}(\mu)}{2i+2} N^{2i+2} + i\kappa(\mu)N, \quad (5.45)$$

including in it the oscillating term. We proceed by complexify the variable $N \rightarrow z = x + iy$, and we look for the critical points of $S[z]$, i.e. $\{z_\sigma \in \mathbb{C} : \partial_z S[z]|_{z=z_\sigma} = 0\}$. Starting from each one of those critical points z_σ we can build a Lefschetz thimble J_σ with equations

$$\begin{aligned} \dot{x} &= -\text{Re}(\partial_z S[z]), \\ \dot{y} &= +\text{Im}(\partial_z S[z]), \end{aligned} \quad (5.46)$$

where t is the parametrization of the thimbles. In this case, they are one-dimensional trajectories in the complex plane, along which the imaginary part of the action $\text{Im}S[z]$ is constant. In the end, we are able to write the partition function as the sum of integrals over the thimbles

$$Z = \sum_{\sigma} m_{\sigma} e^{-i\text{Im}S[z_\sigma]} \int_{J_\sigma} \frac{\partial z}{\partial t} dt e^{-\text{Re}S[z(t)]}, \quad (5.47)$$

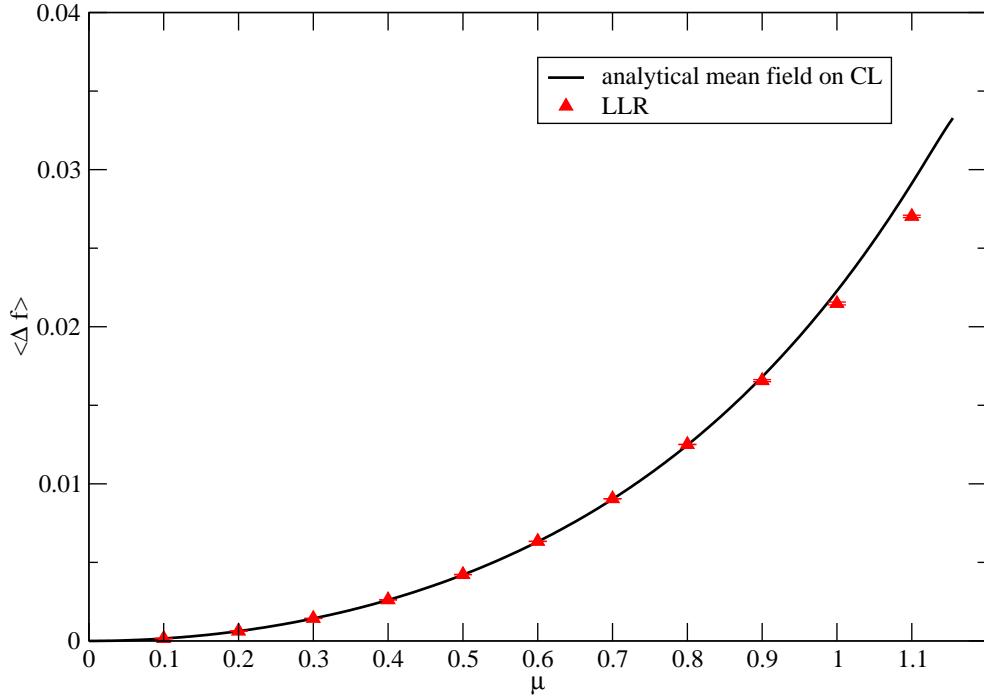


Figure 5.6: Plots of the average free energy difference $\langle \Delta f \rangle = -\frac{T}{V} \log \frac{Z(\mu)}{Z_{pq}}$. Comparison between LLR method for density of states (red dots) and analytical mean field theory result applied to complex Langevin dynamics (black line). Simulation on a 8^4 lattice with the choice of parameters : $m = \lambda = 1$.

where m_σ the intersection number that tell us if the thimble contributes or not. The union of all contributing thimbles forms a deformation of the original real axis. In Fig.5.5 we show a typical example of new domain of integration composed by the contributing Lefschetz thimbles in the complex plane.

Along this new contour we are left with a global phase $e^{-i\text{Im}S[z_\sigma]}$, which is usually harmless even though it is different on each thimble and, therefore, has to be carefully taken into account. Together with the global phase, there is a residual sign problem coming from the Jacobian $\frac{\partial z}{\partial t}$. The latter, in general, is much milder than the original sign problem and is seldom a problem and also in our case it does not create any problem to the numerical integration along the thimbles.

The results obtained with multi-precision and with Lefschetz thimble agree with each other. In Fig.5.6 we show the average free energy density difference between full and phase quenched theory $\langle \Delta f \rangle = -\frac{T}{V} \log \frac{Z(\mu)}{Z_{pq}}$, for the 8^4 lattice. We confront our results (red triangles) with the analytical computation (black line) of mean field complex Langevin dynamics [112], which the author shared with us. We notice a small discrepancy starting from $\mu = 1$. An explanation for that could be a systematic error introduced by the fit in our method or an inaccuracy of the mean field theory at high μ or both together. Along this line of thoughts, an interesting observation

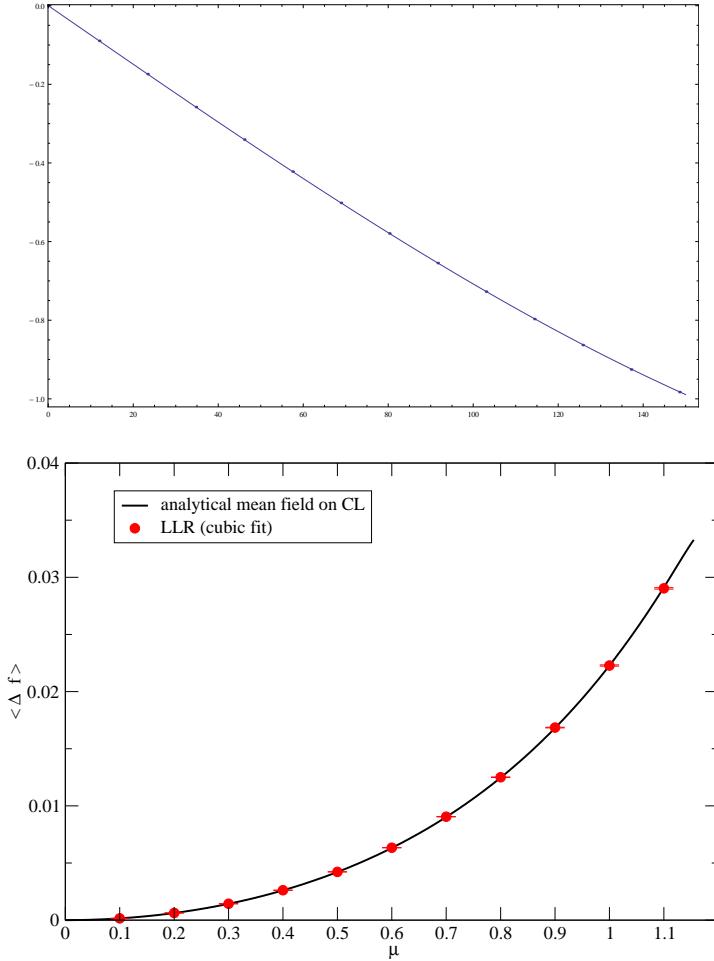


Figure 5.7: *up* : equivalent of Fig.5.4 but with a cubic fit of the data, $a(N) = c_1 N + c_3 N^3$, up until the change in slope, which in this case ($\mu = 0.8$) happens around $N \simeq 150$. *down* : same as Fig.5.6 where, this time, the DOS has been generated by the cubic fit of the $a(N)$ shown above.

is that if we fit the data only up until the change in slope (as in Fig.5.7 up) with a cubic polynomial $a(N) = c_1 N + c_3 N^3$, then we get values of $\langle \Delta F \rangle$ compatible with the mean field expected value (Fig.5.7 down).

We would, also, like to stress that the typical value of the partition function for the last values of μ is of the order of the $10^{-30} - 10^{-40}$. That means the LLR method for computing the DOS is able to achieve a remarkable precision.

In conclusion, in this Chapter we described the density of states method, based on the LLR algorithm, as a candidate for the study of theories affected by the sign problem. We showed how this allows to reduce a complicated four-dimensional

theory in a one dimensional integral. The oscillating integral, though, is still very problematic even in one dimension, and forces us to implement a fit of the $a^*(N)$ to get rid of the noise. This inevitably introduces a systematic error which we try to control with, for example, Bayesian methods.

Lastly we showed some results for the average of the logarithm of the phase. We found good agreement with a mean field calculation up to $\mu = 1$ where the two methods give slightly different results.

We also found that the DOS method is capable of extremely precise direct measurements as, for example, partition functions of the order of 10^{-40} or smaller.

Conclusions

The topic of this thesis is centred on the sign problem in lattice field theory. In particular, our efforts have been concentrated on three different approaches, namely complex Langevin dynamics, Lefschetz thimble and density of states, and their application to various theories and toy models affected by the sign problem.

Particular focus has been put on complex Langevin dynamics. We dedicate Cap.2 to review stochastic quantization as an alternative to the usual path integral formulation of the QFT, and to discuss the formal proofs of correctness for complex Langevin dynamics. We also introduced the concept of gauge cooling as a method to control CL dynamics into the complexified space. We employed gauge cooling in the one dimensional Polyakov chain model, where the sign problem is triggered by a complex β in the action. We showed that, indeed, gauge cooling is able to control the dynamics in the complex directions leading CL to converge to the right results even where, without gauge cooling, it would have failed.

After that, we moved on to applying complex Langevin dynamics to the Yang-Mills pure gauge theory with a topological θ -term. This probably represents the most challenging chapter (Cap.3) of the thesis. We showed that the use of CL dynamics, helped by the gauge cooling technique, allows us to obtain results at real θ , at least for the bare (unrenormalised) theory. We put particular care in testing the criteria of correctness to ensure the results at real θ are reliable. We produced, at real θ , the behaviour of the bare topological charge $Q_{bare}(\theta)$, expected by analytical continuation from imaginary θ where the action is real and Monte Carlo methods can be employed. Lastly, we used this information to study the dependence of the bare topological susceptibility χ_{top} from θ .

Another very interesting method developed in order to deal with the sign problem is the one based on the Lefschetz thimbles. The starting point of this method, as well as complex Langevin dynamics, is the complexification of the fields in the theory. Moreover, the equations of the thimbles are the complex conjugate of the classical, i.e. without noise, CL equations. A comparison which enhances the similarity and the differences between those two approaches is, therefore, very interesting. In Cap.4, we compare CL dynamics and Lefschetz thimble in the study of a quartic model, U(1) one-link model with a complex determinant, and SU(2) non abelian one-link model with an imaginary β parameter in the action. Here we showed the quartic model mostly as an introduction on the Lefschetz thimbles method. The second model is interesting because the determinant can be included

as a non-holomorphic term in the action, in analogy with what happens in QCD at finite density. The third model presents similarities with gauge theories in real time dynamics, i.e. imaginary β . From the CL point of view, it can be tested both with complete gauge fixing and gauge cooling. Furthermore, the gauge fixed representation produces the Haar measure, which also can be included as a non holomorphic term in the action. This study led us to observe directly the connection between CL and Lefschetz thimbles and, furthermore, we could establish the role of the singularities in the thimbles approach, i.e. they are always the end point of a thimble.

The last approach to the sign problem we have dealt with is the density of states. We reviewed how, in this method, one numerically computes the canonical density of states relative to the action, simplifying the quantum field theory to a one dimensional problem. We showed that, when the action is complex, one can regard the imaginary part of the action $\text{Im}(S) = N$ as the grand-canonical variable. It is possible, then, to build a positive generalized density of states $\rho(N)$ by using only the real part of the action $\text{Re}(S)$. In the end, we are left with a one dimensional oscillating integral in N . We applied the generalized DOS method to the relativistic Bose gas theory at finite density. We considered two different ansatz for the fit of the DOS and, in both cases, we showed our computation of free energy density as a function of the chemical potential μ .

In this thesis we described in detail three approaches to the sign problem. Of those three, Lefschetz thimbles and density of states method are somehow more recent and still in a testing phase. Which means promising results have been achieved in several toy models and simple field theories, like the Bose gas at finite density, but the step to full gauge theories and QCD has yet to be done. Complex Langevin dynamics, on the other hand, has already completed the testing part, developing formal criteria of correctness which guides the trustworthy of its results. Furthermore, its application to gauge theories with a sign problem has already started in the last couple of years, obtaining promising results even for full QCD. The most recent efforts are mainly aimed to deal with non holomorphic drifts and to achieve full control of CL dynamics in the whole space of the parameter, using gauge cooling or some alternative method. The successful completion of both this challenges would elect CL dynamics as the first consistent method to quantitatively study the phase diagram of any gauge theory affected by the sign problem, in particular QCD at finite density.

List of Figures

1	Qualitative representation of quantum chromodynamics (QCD) phase diagram.	7
2.1	Gauge cooling in $SL(N, \mathbb{C})$ brings the link as close as possible to $SU(N)$. The orbit on the left is equivalent to a $SU(N)$ configuration, while the one on the right is not.	37
2.2	$SU(2)$ Polyakov chain. Effect of gauge cooling on Polyakov chains with a different number of links.	39
2.3	$SU(2)$ Polyakov chain. Effect of different implementations of adaptive gauge cooling on a Polyakov chain of length $N_\ell = 1000$, plotted on a linear-log scale	40
2.4	Same as FIG. 2.3 but plotted on a log-log scale	40
2.5	$SU(2)$ Polyakov chain : comparison between the history of unitarity norm without gauge cooling, with fixed gauge cooling and with adaptive gauge cooling	42
2.6	Histogram of the distribution of the real and imaginary part of $\langle S \rangle$. .	43
2.7	Convergence of CL to the correct result after the application of fixed and adaptive gauge cooling in the $SU(2)$ Polyakov chain	44
3.1	History of the topological charge of a $SU(3)$ configuration (with topological charge $Q = -3$) along the gradient flow, for a 12^4 lattice at $\beta = 5.9$	58
3.2	History of the action S for the same configuration of Fig. 3.1	59
3.3	Distribution of Q_{top} at $\theta_I = -20i$ (above) and $\theta_I = 20i$ (below), on a 12^4 lattice at $\beta = 6$, obtained by generating configurations using complex Langevin dynamics with gauge cooling and subsequently applying gradient flow to recover the discrete values of the charges. The spacing between the peaks is a bit smaller than one as a finite volume effect, and this is responsible for the peaks being shifted in respect to the integers.	64
3.4	expectation values of the plaquette in $SU(3)$ Yang-Mills theory in the presence of a θ -term , for real and imaginary θ_L , using CL and HMC (imaginary θ_L only), on a 6^4 lattice at three β values. The HMC data has been shifted horizontally for clarity.	65

3.5	Normalized histogram of the distribution of the complex action for real $\theta_L=2$ and $\beta = 5.9, 6, 6.1$	66
3.6	CL running averages of the bare topological charge, for $\beta = 6.1$. The imaginary part of Q_L is shown for θ_R and the real part for θ_I . The dotted lines at θ_I indicate results obtained with HMC algorithm.	68
3.7	Opposite curvature in the dependence of $\langle Q_L \rangle$ on real and imaginary θ_L , at $\beta = 6.1$	69
3.8	Same as Fig.3.7, at real θ_L , for three values of $\beta = 5.8, 5.9, 6.1$	69
3.9	Bare topological susceptibility χ_L as a function of β	70
3.10	Wilson Flow on a $SL(3, C)$ configuration (x axis is flow time) for $\theta_L = 20$, $\beta = 6.1$, $L = 8^4$	72
4.1	Thimbles and Langevin flow in the quartic model with $\sigma = 1$ and $h = 1 + i$: the blue circles denote the fixed points, the (normalised) arrows the classical Langevin drift, and full (dashed) lines the stable (unstable) thimbles. The two blue thimbles contribute. The third fixed point does not contribute.	80
4.2	Quartic model, with $\sigma = 1$ and $h = 1 + i$. Histogram collected during a complex Langevin simulation (left) and a comparison with the thimbles (right).	81
4.3	Thimbles and Langevin flow in the $U(1)$ one-link model with $\beta = 1, \mu = 2$ and $\kappa = 1/2$ (up) and $\kappa = 2$ (down): the blue circles indicate the fixed points, the (normalised) arrows the classical Langevin drift, and full (dashed) lines the stable (unstable) thimbles. Note that the blue lines indicate the contributing stable thimbles, with the red associated unstable ones, while the light brown lines indicate thimbles that do not contribute. The squares indicate where the flow diverges, $\text{Im}S$ jumps, and the direction of the flow along the thimble changes sign. Only the blue thimble(s) contribute.	83
4.4	As in the previous plot, with scatter data from a complex Langevin simulation added.	87
4.5	$SU(2)$ one-link model, with $\beta = (1+i\sqrt{3})/2$: thimbles in the xy plane (left) and the $\text{Tr}U$ plane (right). Blue circles indicate fixed points, full (dashed) lines the stable (unstable) thimbles, and red squares the singular points. The blue thimble contributes.	90
4.6	Histograms collected during a complex Langevin simulation in the gauge fixed formulation of the $SU(2)$ one-link model, with $\beta = (1+i\sqrt{3})/2$, in the xy plane (left) and the $\text{Tr}U$ plane (right).	90
4.7	Comparison between the complex Langevin histograms and the thimbles, in the xy plane (left) and the $\text{Tr}U$ plane (right).	90
4.8	Histograms collected during a complex Langevin simulation in the matrix formulation of the $SU(2)$ one-link model at $\beta = (1+i\sqrt{3})/2$, with gauge cooling, using 0, 1, 2, 4 (from top to bottom) gauge cooling steps.	92

4.9	SU(2) one-link model at $\beta = i$. Histogram collected during a complex Langevin simulation in the $\text{Tr}U$ plane (left) and a comparison with the thimbles associated with the degenerate fixed point at $\text{Tr}U = 2i$ (right).	95
5.1	An example demonstrating the effect of the thermalization in the Robbins-Monro root finding method. In both cases the initial value is $a_0 = 0$. The black line represents the application of Robbins-Monro from the beginning, while the red line has been thermalized with 30 steps of Newton-Raphson.	101
5.2	Distribution of the $a^*(N)$ for $L = 8^4$ and $\mu = 0.9$	104
5.3	Dependence of $a^*(N)$ on the chemical potential μ for $L = 8^4$	106
5.4	Fit of $a^*(N)$ with an odd polynomial of order 7 $a(N) = c_1N + c_3N^3 + c_5N^5 + c_7N^7$, for $\mu = 0.8$ and $L = 8^4$	108
5.5	Contributing Lefschetz thimbles for order seven polynomial fit of the $a^*(N)$	109
5.6	Plots of the average free energy difference $\langle \Delta f \rangle = -\frac{T}{V} \log \frac{Z(\mu)}{Z_{pq}}$. Comparison between LLR method for density of states (red dots) and analytical mean field theory result applied to complex Langevin dynamics (black line). Simulation on a 8^4 lattice with the choice of parameters : $m = \lambda = 1$	110
5.7	<i>up</i> : equivalent of Fig.5.4 but with a cubic fit of the data, $a(N) = c_1N + c_3N^3$, up until the change in slope, which in this case ($\mu = 0.8$) happens around $N \simeq 150$. <i>down</i> : same as Fig.5.6 where, this time, the DOS has been generated by the cubic fit of the $a(N)$ shown above.111	

Bibliography

- [1] M. Creutz, L. Jacobs and C. Rebbi, “*Monte Carlo Study of Abelian Lattice Gauge Theories*,” Phys. Rev. D **20** (1979) 1915.
- [2] M. Creutz, “*Monte Carlo Study of Quantized SU(2) Gauge Theory*,” Phys. Rev. D **21** (1980) 2308.
- [3] D. Weingarten, “*Monte Carlo Evaluation of Hadron Masses in Lattice Gauge Theories with Fermions*,” Phys. Lett. B **109** (1982) 57.
- [4] G. Marchesini, B. R. Webber, G. Abbiendi, I. G. Knowles, M. H. Seymour and L. Stanco, “*HERWIG: A Monte Carlo event generator for simulating hadron emission reactions with interfering gluons. Version 5.1 - April 1991*,” Comput. Phys. Commun. **67** (1992) 465.
- [5] E. P. Munger and M. A. Novotny, “*Reweighting In Monte Carlo And Monte Carlo Renormalization Group Studies*,”
- [6] R. T. Scalettar, E. Y. Loh, J. E. Gubernatis, A. Moreo, S. R. White, D. J. Scalapino, R. L. Sugar and E. Dagotto, “*Phase Diagram Of The Two-dimensional Negative U Hubbard Model*,” NSF-ITP-89-06.
- [7] M. P. Lombardo, “*Finite density (might well be easier) at finite temperature*,” Nucl. Phys. Proc. Suppl. **83** (2000) 375 [hep-lat/9908006].
- [8] O. Philipsen, “*QCD thermodynamics at finite density*,” hep-ph/0110051.
- [9] P. de Forcrand and O. Philipsen, “*QCD phase diagram for small densities from simulations at imaginary mu*,” Nucl. Phys. Proc. Suppl. **119** (2003) 535 [hep-lat/0209084].
- [10] G. Munster, “*Strong Coupling Expansions for the Mass Gap in Lattice Gauge Theories*,” Nucl. Phys. B **190** (1981) 439 [Nucl. Phys. B **200** (1982) 536] [Nucl. Phys. B **205** (1982) 648].
- [11] J. Langelage, G. Munster and O. Philipsen, “*Strong coupling expansion for Yang-Mills theory at finite temperature*,” PoS LAT **2007** (2007) 201 [arXiv:0710.0512 [hep-lat]].

- [12] J. Langelage, S. Lottini and O. Philipsen, “*Centre symmetric 3d effective actions for thermal $SU(N)$ Yang-Mills from strong coupling series*,” JHEP **1102** (2011) 057 [JHEP **1107** (2011) 014] [arXiv:1010.0951 [hep-lat]].
- [13] M. Fromm, J. Langelage, S. Lottini and O. Philipsen, “*The QCD deconfinement transition for heavy quarks and all baryon chemical potentials*,” JHEP **1201** (2012) 042 [arXiv:1111.4953 [hep-lat]].
- [14] F. Green and F. Karsch, “*Mean Field Analysis of $SU(N)$ Deconfining Transitions in the Presence of Dynamical Quarks*,” Nucl. Phys. B **238** (1984) 297.
- [15] G. Aarts, “*Complex Langevin dynamics at finite chemical potential: Mean field analysis in the relativistic Bose gas*,” JHEP **0905** (2009) 052 [arXiv:0902.4686 [hep-lat]].
- [16] O. Akerlund, P. de Forcrand, A. Georges and P. Werner, “*Dynamical Mean Field Approximation Applied to Quantum Field Theory*,” Phys. Rev. D **88** (2013) 125006 [arXiv:1305.7136 [hep-lat]].
- [17] P. H. Damgaard and H. Huffel, “*Stochastic Quantization*,” Phys. Rept. **152** (1987) 227.
- [18] G. Aarts, F. A. James, E. Seiler and I. O. Stamatescu, “*Complex Langevin: Etiology and Diagnostics of its Main Problem*,” Eur. Phys. J. C **71** (2011) 1756 [arXiv:1101.3270 [hep-lat]].
- [19] E. Seiler, D. Sexty and I. O. Stamatescu, “*Gauge cooling in complex Langevin for QCD with heavy quarks*,” Phys. Lett. B **723** (2013) 213 [arXiv:1211.3709 [hep-lat]].
- [20] E. Witten, “*A New Look At The Path Integral Of Quantum Mechanics*,” arXiv:1009.6032 [hep-th].
- [21] M. Cristoforetti *et al.* [AuroraScience Collaboration], “*New approach to the sign problem in quantum field theories: High density QCD on a Lefschetz thimble*,” Phys. Rev. D **86** (2012) 074506 [arXiv:1205.3996 [hep-lat]].
- [22] H. Fujii, D. Honda, M. Kato, Y. Kikukawa, S. Komatsu and T. Sano, “*Hybrid Monte Carlo on Lefschetz thimbles - A study of the residual sign problem*,” JHEP **1310** (2013) 147 [arXiv:1309.4371 [hep-lat]].
- [23] K. N. Anagnostopoulos and J. Nishimura, “*New approach to the complex action problem and its application to a nonperturbative study of superstring theory*,” Phys. Rev. D **66** (2002) 106008 [hep-th/0108041].
- [24] Z. Fodor, S. D. Katz and C. Schmidt, “*The Density of states method at non-zero chemical potential*,” JHEP **0703** (2007) 121 [hep-lat/0701022].

- [25] K. Langfeld, B. Lucini and A. Rago, “*The density of states in gauge theories*,” Phys. Rev. Lett. **109** (2012) 111601 [arXiv:1204.3243 [hep-lat]].
- [26] S. Kratochvila and P. de Forcrand, “*QCD at small baryon number*,” Nucl. Phys. Proc. Suppl. **140** (2005) 514 [hep-lat/0409072].
- [27] A. Alexandru, M. Faber, I. Horvath and K. F. Liu, “*Lattice QCD at finite density via a new canonical approach*,” Phys. Rev. D **72** (2005) 114513 [hep-lat/0507020].
- [28] S. Kratochvila and P. de Forcrand, “*The Canonical approach to finite density QCD*,” PoS LAT **2005** (2006) 167 [hep-lat/0509143].
- [29] S. Chandrasekharan, “*A New computational approach to lattice quantum field theories*,” PoS LATTICE **2008** (2008) 003 [arXiv:0810.2419 [hep-lat]].
- [30] Y. D. Mercado, C. Gattringer and A. Schmidt, “*Surface worm algorithm for abelian Gauge-Higgs systems on the lattice*,” Comput. Phys. Commun. **184** (2013) 1535 [arXiv:1211.3436 [hep-lat]].
- [31] G. Aarts, “*Developments in lattice quantum chromodynamics for matter at high temperature and density*,” Pramana **84** (2015) 5, 787 [arXiv:1312.0968 [hep-lat]].
- [32] P. de Forcrand, “*Simulating QCD at finite density*,” PoS LAT **2009** (2009) 010 [arXiv:1005.0539 [hep-lat]].
- [33] M. A. Stephanov, “*Random matrix model of QCD at finite density and the nature of the quenched limit*,” Phys. Rev. Lett. **76** (1996) 4472 [hep-lat/9604003].
- [34] M. D’Elia and M. P. Lombardo, “*Finite density QCD via imaginary chemical potential*,” Phys. Rev. D **67** (2003) 014505 [hep-lat/0209146].
- [35] D. Sexty, “*Progress in complex Langevin simulations of full QCD at non-zero density*,” Nucl. Phys. A **931** (2014) 856 [arXiv:1408.6767 [hep-lat]].
- [36] Z. Fodor, S. D. Katz, D. Sexty and C. Trk, “*Complex Langevin dynamics for dynamical QCD at nonzero chemical potential: a comparison with multi-parameter reweighting*,” arXiv:1508.05260 [hep-lat].
- [37] R. D. Peccei and H. R. Quinn, “*CP Conservation in the Presence of Instantons*,” Phys. Rev. Lett. **38** (1977) 1440.
- [38] R. D. Peccei and H. R. Quinn, “*Constraints Imposed by CP Conservation in the Presence of Instantons*,” Phys. Rev. D **16** (1977) 1791.
- [39] E. Berkowitz, M. I. Buchoff and E. Rinaldi, “*Lattice QCD input for axion cosmology*,” Phys. Rev. D **92** (2015) 3, 034507 [arXiv:1505.07455 [hep-ph]].

- [40] H. Panagopoulos and E. Vicari, “*The 4D $SU(3)$ gauge theory with an imaginary θ term,*” JHEP **1111** (2011) 119 [arXiv:1109.6815 [hep-lat]].
- [41] M. D’Elia and F. Negro, “ *θ dependence of the deconfinement temperature in Yang-Mills theories,*” Phys. Rev. Lett. **109** (2012) 072001 [arXiv:1205.0538 [hep-lat]].
- [42] M. D’Elia and F. Negro, “*Phase diagram of Yang-Mills theories in the presence of a θ term,*” Phys. Rev. D **88** (2013) 3, 034503 [arXiv:1306.2919 [hep-lat]].
- [43] L. Bongiovanni, G. Aarts, E. Seiler and D. Sexty, “*Complex Langevin dynamics for $SU(3)$ gauge theory in the presence of a theta term,*” PoS LATTICE **2014** (2014) 199 [arXiv:1411.0949 [hep-lat]].
- [44] L. Bongiovanni, G. Aarts, E. Seiler, D. Sexty and I. O. Stamatescu, “*Adaptive gauge cooling for complex Langevin dynamics,*” PoS LATTICE **2013** (2014) 449 [arXiv:1311.1056 [hep-lat]].
- [45] R. P. Feynman, “*Space-time approach to nonrelativistic quantum mechanics,*” Rev. Mod. Phys. **20** (1948) 367.
- [46] K. G. Wilson, “*Confinement of Quarks,*” Phys. Rev. D **10** (1974) 2445.
- [47] G. Parisi and Y. s. Wu, “*Perturbation Theory Without Gauge Fixing,*” Sci. Sin. **24** (1981) 483.
- [48] P. Langevin, CR. Acad. Sci. (Paris) **146** (1908) 530
- [49] G. Aarts, L. Bongiovanni, E. Seiler, D. Sexty and I. O. Stamatescu, “*Controlling complex Langevin dynamics at finite density,*” Eur. Phys. J. A **49** (2013) 89 [arXiv:1303.6425 [hep-lat]].
- [50] J. R. Klauder and W. P. Petersen, “*Spectrum Of Certain Nonselfadjoint Operators And Solutions Of Langevin Equations With Complex Drift,*”
- [51] J. Ambjorn and S. K. Yang, “*Numerical Problems in Applying the Langevin Equation to Complex Effective Actions,*” Phys. Lett. B **165** (1985) 140.
- [52] J. Ambjorn, M. Flensburg and C. Peterson, “*The Complex Langevin Equation and Monte Carlo Simulations of Actions With Static Charges,*” Nucl. Phys. B **275** (1986) 375.
- [53] G. Aarts, F. A. James, E. Seiler and I. O. Stamatescu, “*Adaptive stepsize and instabilities in complex Langevin dynamics,*” Phys. Lett. B **687** (2010) 154 [arXiv:0912.0617 [hep-lat]].
- [54] G. Aarts, E. Seiler and I. O. Stamatescu, “*The Complex Langevin method: When can it be trusted?,*” Phys. Rev. D **81** (2010) 054508 [arXiv:0912.3360 [hep-lat]].

- [55] G. Aarts, F. A. James, E. Seiler and I. O. Stamatescu, “*Complex Langevin dynamics: criteria for correctness*,” PoS LATTICE **2011** (2011) 197 [arXiv:1110.5749 [hep-lat]].
- [56] G. Aarts, F. A. James, J. M. Pawłowski, E. Seiler, D. Sexty and I. O. Stamatescu, “*Stability of complex Langevin dynamics in effective models*,” JHEP **1303** (2013) 073 [arXiv:1212.5231 [hep-lat]].
- [57] G. Aarts, P. Giudice and E. Seiler, “*Localised distributions and criteria for correctness in complex Langevin dynamics*,” Annals Phys. **337** (2013) 238 [arXiv:1306.3075 [hep-lat]].
- [58] A. Mollgaard and K. Splittorff, “*Complex Langevin Dynamics for chiral Random Matrix Theory*,” Phys. Rev. D **88** (2013) 11, 116007 [arXiv:1309.4335 [hep-lat]].
- [59] J. Greensite, “*Comparison of complex Langevin and mean field methods applied to effective Polyakov line models*,” Phys. Rev. D **90** (2014) 11, 114507 [arXiv:1406.4558 [hep-lat]].
- [60] G. Aarts, L. Bongiovanni, E. Seiler, D. Sexty and I. O. Stamatescu, “*Complex Langevin simulation for QCD-like models*,” PoS LATTICE **2013** (2014) 451 [arXiv:1310.7412 [hep-lat]].
- [61] D. Sexty, “*Simulating full QCD at nonzero density using the complex Langevin equation*,” Phys. Lett. B **729** (2014) 108 [arXiv:1307.7748 [hep-lat]].
- [62] G. Aarts, E. Seiler, D. Sexty and I. O. Stamatescu, “*Simulating QCD at nonzero baryon density to all orders in the hopping parameter expansion*,” Phys. Rev. D **90** (2014) 11, 114505 [arXiv:1408.3770 [hep-lat]].
- [63] G. Aarts, F. Attanasio, B. Jger, E. Seiler, D. Sexty and I. O. Stamatescu, “*Exploring the phase diagram of QCD with complex Langevin simulations*,” PoS LATTICE **2014** (2014) 200 [arXiv:1411.2632 [hep-lat]].
- [64] G. Aarts, B. Jger, E. Seiler, D. Sexty and I. O. Stamatescu, “*Systematic approximation for QCD at non-zero density*,” PoS LATTICE **2014** (2014) 207 [arXiv:1412.5775 [hep-lat]].
- [65] J. Berges and D. Sexty, “*Real-time gauge theory simulations from stochastic quantization with optimized updating*,” Nucl. Phys. B **799** (2008) 306 [arXiv:0708.0779 [hep-lat]].
- [66] A. A. Belavin, A. M. Polyakov, A. S. Schwartz and Y. S. Tyupkin, “*Pseudoparticle Solutions of the Yang-Mills Equations*,” Phys. Lett. B **59** (1975) 85.
- [67] A. M. Polyakov, “*Quark Confinement and Topology of Gauge Groups*,” Nucl. Phys. B **120** (1977) 429.

- [68] G. 't Hooft, "Symmetry Breaking Through Bell-Jackiw Anomalies," Phys. Rev. Lett. **37** (1976) 8.
- [69] E. Witten, "Current Algebra Theorems for the $U(1)$ Goldstone Boson," Nucl. Phys. B **156** (1979) 269.
- [70] G. Veneziano, "U(1) Without Instantons," Nucl. Phys. B **159** (1979) 213.
- [71] R. Jackiw and C. Rebbi, "Vacuum Periodicity in a Yang-Mills Quantum Theory," Phys. Rev. Lett. **37** (1976) 172.
- [72] C. G. Callan, Jr., R. F. Dashen and D. J. Gross, "The Structure of the Gauge Theory Vacuum," Phys. Lett. B **63** (1976) 334.
- [73] R. J. Crewther, P. Di Vecchia, G. Veneziano and E. Witten, "Chiral Estimate of the Electric Dipole Moment of the Neutron in Quantum Chromodynamics," Phys. Lett. B **88** (1979) 123 [Phys. Lett. B **91** (1980) 487].
- [74] T. Schfer and E. V. Shuryak, "Instantons in QCD," Rev. Mod. Phys. **70** (1998) 323 [hep-ph/9610451].
- [75] R. Rajaraman, "Solitons And Instantons. An Introduction To Solitons And Instantons In Quantum Field Theory," Amsterdam, Netherlands: North-holland (1982) 409p
- [76] C. G. Callan, Jr., R. F. Dashen and D. J. Gross, "Toward a Theory of the Strong Interactions," Phys. Rev. D **17** (1978) 2717.
- [77] R. Bott, "An application of the Morse theory to the topology of Lie-groups," Bull. Soc. Math. **84** (1956) 251.
- [78] S. L. Adler, "Axial vector vertex in spinor electrodynamics," Phys. Rev. **177** (1969) 2426-2438.
- [79] J. S. Bell, R. Jackiw, "A PCAC puzzle: $\pi 0 \rightarrow \gamma \gamma$ gamma in the sigma model," Nuovo Cim. **A60** (1969) 47-61.
- [80] G. 't Hooft, "Computation of the Quantum Effects Due to a Four-Dimensional Pseudoparticle," Phys. Rev. D **14** (1976) 3432 [Phys. Rev. D **18** (1978) 2199].
- [81] M. Luscher, "Topology of Lattice Gauge Fields," Commun. Math. Phys. **85** (1982) 39.
- [82] A. Phillips and D. Stone, "Lattice Gauge Fields, Principal Bundles and the Calculation of Topological Charge," Commun. Math. Phys. **103** (1986) 599.
- [83] M. Campostrini, A. Di Giacomo and H. Panagopoulos, "The Topological Susceptibility on the Lattice," Phys. Lett. B **212** (1988) 206.

- [84] M. Campostrini, A. Di Giacomo, H. Panagopoulos and E. Vicari, “*Topological Charge, Renormalization and Cooling on the Lattice*,” Nucl. Phys. B **329** (1990) 683.
- [85] M. Luscher, “*Properties and uses of the Wilson flow in lattice QCD*,” JHEP **1008** (2010) 071 [arXiv:1006.4518 [hep-lat]].
- [86] M. Luscher, “*Topology, the Wilson flow and the HMC algorithm*,” PoS LATICE **2010** (2010) 015 [arXiv:1009.5877 [hep-lat]].
- [87] C. Bonati and M. D’Elia, “*Comparison of the gradient flow with cooling in $SU(3)$ pure gauge theory*,” Phys. Rev. D **89** (2014) 10, 105005 [arXiv:1401.2441 [hep-lat]].
- [88] E. Witten, “*Large N Chiral Dynamics*,” Annals Phys. **128** (1980) 363.
- [89] E. Witten, “*Theta dependence in the large N limit of four-dimensional gauge theories*,” Phys. Rev. Lett. **81** (1998) 2862 [hep-th/9807109].
- [90] E. Witten, “*Analytic Continuation Of Chern-Simons Theory*,” arXiv:1001.2933 [hep-th];
- [91] M. Cristoforetti, F. Di Renzo, A. Mukherjee and L. Scorzato, “*Monte Carlo simulations on the Lefschetz thimble: Taming the sign problem*,” Phys. Rev. D **88** (2013) 5, 051501 [arXiv:1303.7204 [hep-lat]].
- [92] A. Mukherjee, M. Cristoforetti and L. Scorzato, “*Metropolis Monte Carlo integration on the Lefschetz thimble: Application to a one-plaquette model*,” Phys. Rev. D **88** (2013) 5, 051502 [arXiv:1308.0233 [physics.comp-ph]].
- [93] M. Cristoforetti, F. Di Renzo, G. Eruzzi, A. Mukherjee, C. Schmidt, L. Scorzato and C. Torrero, “*An efficient method to compute the residual phase on a Lefschetz thimble*,” arXiv:1403.5637 [hep-lat].
- [94] Y. Tanizaki and T. Koike, “*Real-time Feynman path integral with Picard-Lefschetz theory and its applications to quantum tunneling*,” arXiv:1406.2386 [math-ph].
- [95] A. Mukherjee and M. Cristoforetti, “*Lefschetz thimble Monte Carlo for many body theories: application to the repulsive Hubbard model away from half filling*,” arXiv:1403.5680 [cond-mat.str-el].
- [96] R. Bott, “*Lectures on Morse theory, old and new*,” Bull. Amer. Math. Soc. (N.S.) **7** (1982) 331 [<http://projecteuclid.org/euclid.bams/1183549637>]
- [97] L. Nicolaescu, “*An Invitation to Morse Theory*,” 2nd ed. (Springer-Verlag, 2011).

- [98] G. Aarts, L. Bongiovanni, E. Seiler and D. Sexty, “*Some remarks on Lefschetz thimbles and complex Langevin dynamics*,” JHEP **1410** (2014) 159 [arXiv:1407.2090 [hep-lat]].
- [99] G. Aarts and I.-O. Stamatescu, “*Stochastic quantization at finite chemical potential*,” JHEP **0809** (2008) 018 [arXiv:0807.1597 [hep-lat]].
- [100] G. Aarts and K. Splittorff, “*Degenerate distributions in complex Langevin dynamics: one-dimensional QCD at finite chemical potential*,” JHEP **1008** (2010) 017 [arXiv:1006.0332 [hep-lat]].
- [101] J. Berges and I. -O. Stamatescu, “*Simulating nonequilibrium quantum fields with stochastic quantization techniques*,” Phys. Rev. Lett. **95** (2005) 202003 [hep-lat/0508030].
- [102] J. Berges, S. Borsanyi, D. Sexty and I. -O. Stamatescu, “*Lattice simulations of real-time quantum fields*,” Phys. Rev. D **75** (2007) 045007 [hep-lat/0609058].
- [103] F. Wang, D. P. Landau, “*Efficient, Multiple-Range Random Walk Algorithm to Calculate the Density of States*,” Phys.Rev.Lett **86** (2001) 2050 [<http://link.aps.org/doi/10.1103/PhysRevLett.86.2050>].
- [104] V. Azcoiti, G. Di Carlo, A. Galante and V. Laliena, “*New proposal for numerical simulations of theta vacuum - like systems*,” Phys. Rev. Lett. **89** (2002) 141601 [hep-lat/0203017].
- [105] K. Langfeld, J. Pawłowski, B. Lucini, A. Rago and R. Pellegrini, “*Towards a density of states approach for dense matter systems*,” PoS LATTICE **2013** (2014) 198 [arXiv:1310.8231 [hep-lat]].
- [106] K. Langfeld and B. Lucini, “*Density of states approach to dense quantum systems*,” Phys. Rev. D **90** (2014) 9, 094502 [arXiv:1404.7187 [hep-lat]].
- [107] B. Lucini and K. Langfeld, “*A novel density of state method for complex action systems*,” PoS LATTICE **2014** (2014) 230 [arXiv:1411.0174 [hep-lat]].
- [108] R. Pellegrini, K. Langfeld, B. Lucini and A. Rago, “*The density of states from first principles*,” PoS LATTICE **2014** (2015) 229 [arXiv:1411.0655 [hep-lat]].
- [109] K. Langfeld, B. Lucini, A. Rago, R. Pellegrini and L. Bongiovanni, “*The density of states approach for the simulation of finite density quantum field theories*,” J. Phys. Conf. Ser. **631** (2015) 1, 012063 [arXiv:1503.00450 [hep-lat]].
- [110] H. Robbins, S. Monro, “*A Stochastic Approximation Method*,” The Annals of Mathematical Statistics **22** (3): 400

- [111] J. Sacks, “*Asymptotic Distribution of Stochastic Approximation Procedures*,” Ann. Math. Statist. Volume 29, Number 2 (1958), 373-405.
- [112] G. Aarts, “*Can stochastic quantization evade the sign problem? The relativistic Bose gas at finite chemical potential*,” Phys. Rev. Lett. **102** (2009) 131601 [arXiv:0810.2089 [hep-lat]].
- [113] O. Akerlund, P. de Forcrand, A. Georges and P. Werner, “*Extended Mean Field study of complex ϕ^4 -theory at finite density and temperature*,” Phys. Rev. D **90** (2014) 6, 065008 [arXiv:1405.6613 [hep-lat]].
- [114] M. G. Endres, “*Avoiding the boson sign problem at finite chemical potential*,” PoS LAT **2006** (2006) 133 [hep-lat/0609037].
- [115] M. G. Endres, “*Method for simulating O(N) lattice models at finite density*,” Phys. Rev. D **75** (2007) 065012 doi:10.1103/PhysRevD.75.065012 [hep-lat/0610029].
- [116] C. Gattringer and T. Kloiber, “*Spectroscopy in finite density lattice field theory: An exploratory study in the relativistic Bose gas*,” Phys. Lett. B **720** (2013) 210 [arXiv:1212.3770 [hep-lat]].
- [117] D. Spiegelhalter,N. G. Best,B. P. Carlin, A. van der Linde “*Bayesian measures of model complexity and fit*,” J. R. Statist. Soc. B (2002) **64** , Part 4, pp. 583
- [118] V.E. Johnson, “*A Bayesian χ^2 test for goodness-of-fit*,” Mathematics - Statistics, 62C10 (Primary) 62E20. (Secondary) [arXiv:math/0508593]
- [119] C. Bailer-Jones, “*Bayesian parameter estimation and model comparison*,” Astrostats 2013 Lecture 1, 2013
- [120] A. Iserles,S.P. Nrsett, S. and Olver, “*Highly Oscillatory Quadrature: The Story so far (Numerical Mathematics and Advanced Applications)*,” Springer Berlin Heidelberg
- [121] A. Iserles,S.P. Nrsett, “*Efficient quadrature of highly oscillatory integrals using derivatives*,” The Royal Society 1364-5021 <http://dx.doi.org/10.1098/rspa.2004.1401>

This figure "SwanseaLogo.jpeg" is available in "jpeg" format from:

<http://arxiv.org/ps/1603.06458v2>

This figure "SwanseaUniLogo.png" is available in "png" format from:

<http://arxiv.org/ps/1603.06458v2>