Topology of the $O(3)$ non-linear sigma model under the gradient flow

A thesis submitted in partial fulfillment of the requirement
for the degree of Bachelor of Science with Honors in
Physics from the College of William and Mary in Virginia,

by

Stuart Thomas

Advisor: Prof. Christopher J. Monahan

Prof. Todd Averett

Prof. Andreas Stathopoulos

Williamsburg, Virginia
May 2021

Contents

Acknowledgments

I would like to thank lots of people, and this is where I will do it...

List of Figures

List of Tables

Abstract

Abstract

Chapter 1

Introduction

Quantum field theory (QFT) is an extraordinarily successful framework which can be applied to a range of physical phenomena. Paired with the Standard Model, QFT provides the prevailing basis for all small-scale physics (that is, where general relativity does not apply) and is the fundamental tool for studying particle physics. In condensed matter physics, effective field theories model emergent effects such as phonons and quasiparticles. Compared to experiment, QFT is remarkably accurate, famously matching the experimental value for the electron g-factor to eight significant figures.[1]

However, this power comes at a cost: the study of quantum fields is rife with infinities. A naïve treatment of quantum field theory produces divergent values for physical quantities, a clearly impossible result. Since the 1950s, this issue has been resolved for a large number of models— most notably quantum electrodynamics— through perturbation theory and the so-called *renormalization group*. However, the technique fails with perturbatively nonrenormalizable theories.

One such example is the *non-linear sigma model*, a prototypical theory in both condensed matter and particle physics. In solid-state systems, this model describes Heisenberg ferromagnets and in nuclear physics, it acts as a prototype for quantum chromodynamics (QCD), exhibiting characteristic features such as a mass gap and

asymptotic freedom. [1]

In this study, we specifically consider the $O(3)$ non-linear sigma model in $1 + 1$ dimensions (one dimension of space, one dimension of time). This theory exhibits topological properties such as *instantons*, or classical field solutions at local minima of the action.

Since the non-linear sigma model cannot be renormalized perturbatively, we cannot study these topological effects with normal perturbative techniques. An alternative solution is placing the field on a discretized lattice, a technique originally used for quantum chromodynamics. In this scenario, field configurations become computationally calculable. This process introduces an nonphysical length scale a, the lattice spacing. Therefore, we expect any physical result to converge in the continuum limit, i.e. when $a \to 0$. However, this is not always the case as observables mix on the lattice, leading to divergences. As an example, states of definite angular momentum mix when discretized, a clear violation of the angular momentum commutation relations.

The gradient flow is a technique designed to remove these divergences. By dampening high-momentum fluctuations, the gradient flow reduces power-divergent mixing and make observables finite on the lattice.[2] In quantum chromodynamics, previous studies have verified the ability of the gradient flow to make observables finite[2]. Due to its success in QCD, there has been interest in using the gradient flow to finitize the topological susceptibility in the 1+1 $O(3)$ non-linear sigma model.[3].

1.1 Method Overview

To numerically study the topological qualities of the non-linear sigma model, we first implement a Markov Chain Monte Carlo simulation. We initially construct a proof-

---

[1] *citation needed*
[2] *citation needed*

of-concept Python program that models the simpler ϕ^4 model (see Sec. 2.1.2). After comparing with existing literature, we transition to a C++ simulation for efficiency, implementing the non-linear sigma model in larger lattices. Since the gradient flow has no exact solution in the non-linear sigma model, we implement a numerical solution using a fourth-order Runge-Kutta approximation. By applying the gradient flow to every configuration in the sample, we can measure its effect on the topological charge and susceptibility.

1.2 Conventions

- Throughout this paper, we use natural units, i.e. $\hbar = 1$ and $c = 1$.

- We use Einstein summation notation, an implicit sum over repeated spacetime indices. For example, if x^μ is a spacetime four-vector and x_μ is its covariant form, the term

$$x^\mu x_\mu = \sum_{\mu=0}^{4} x^\mu x_\mu$$
$$= x_0^2 - x_1^2 - x_2^2 - x_3^2.$$

Chapter 2

Theory

This thesis incorporates two main bodies of knowledge: quantum field theory and statistical simulation. Through the path integral formulation of quantum field theory, we are able to describe the physics of the former with the established mathematics of the latter.

2.1 Quantum Field Theory

In this section we outline a rough description of quantum field theory. A full introduction is beyond the scope of this paper, however we do assume knowledge of nonrelativistic quantum mechanics and classical field theory.

The fundamental hypothesis of quantum field theory (QFT) describes particles as discrete packets of energy on a quantum field. But what is a quantum field? Like in classical mechanics, a field is a function of spacetime with some mathematical object assigned to each point in space and time. In the case of the electric field, this object is a three-dimensional vector, while the electric potential is a scalar field. Classical and quantum fields have Lagrangians which define how they evolve in time. What differentiates a quantum field from a classical field is superposition: where classical fields have a definite configuration, quantum fields exist in a superposition of all possible configurations. It is possible—though nontrivial and outside the scope of this

description— to motivate the appearance of discrete particles from this superposition.

This general description allows QFT to easily incorporate special relativity. By ensuring that the Lagrangian of a theory is invariant under Lorentz transformations, we can ensure that the theory is physical.

Though most introductions to QFT use the "second quantization," we will use a alternative formulation known as the path integral formulation (for a similar introduction see Zee's textbook [4]).

2.1.1   Path Integral Formulation

As stated above, we can model a quantum field theory as a superposition of all possible classical fields. Like single-particle quantum mechanics, each configuration has a probability amplitude. To measure expectation values of observables, we simply take an average over all configurations weighted by this complex amplitude. We can formalize this notion using the fundamental formula[1]

$$\langle \hat{O} \rangle = \frac{1}{Z} \int \mathcal{D}\phi \, \hat{O}[\phi] \, e^{iS[\phi]} \tag{2.1}$$

where $\langle \hat{O} \rangle$ is the expectation value of an arbitrary operator $\hat{O}$; Z is a normalization constant; S is the action functional, defined from the field's Lagrangian; and $\int \mathcal{D}\phi$ represents the eponymous path integral. Though it is possible to define this integral rigorously, for our purposes we can equate it to a sum over all possible configurations. This form is the quantum analog of the classical principle of least action and reduces to such for large values of the action. For a more pedagogical explanation, see Richard Feynman's lectures on physics.[5]

At first glance, Eq. 2.1 is remarkably similar to the statistics of the canonical ensemble. Through this similarity, we will be able to use mathematical tools from statistical mechanics to study quantum field theories. However, the factor of i in

---

[1]Since this study concerns vacua, we do not include a source term.

the exponent currently prohibits us from making this jump. To remedy this issue, we perform a "Wick rotation" which shifts spacetime into Euclidean coordinates. In normal spacetime, defined by the Minkowski metric, the Lorentz-invariant distance is given as

$$s^2 = x_0^2 - x_1^2 - x_3^2 - x_3^2 \tag{2.2}$$

where $x_0 = ct$ and $\vec{x} = (x_1, x_2, x_3)^T$. By redefining the time coordinate of a spacetime point x to be $x_4 = ix_0$, we find that the quantity

$$s_E^2 = x_1^2 + x_3^2 + x_3^2 + x_4^2, \tag{2.3}$$

is equivalently Lorentz invariant, which is representative of a four-dimensional Euclidean space. Furthermore, we find that

$$d^4 x_E = d^3 \vec{x} dx_4 \tag{2.4}$$

$$= i d^3 \vec{x} dx_0 \tag{2.5}$$

$$= i d^4 x. \tag{2.6}$$

We can use this transformation to redefine the Lagrangian $\mathcal{L}$ in Euclidean space as $\mathcal{L}_E$, replacing all x_0 with ix_0. Since a Lorentz-invariant Lagrangian must include only even powers and derivatives of x, the Euclidean Lagrangian remains real. Subsequently, we can define a Euclidean action based on the differential in Eq. 2.6:

$$S_E = \int d^4 x_E \mathcal{L}_E \tag{2.7}$$

$$= i \int d^4 x \mathcal{L}_E, \tag{2.8}$$

allowing us to redefine the path integral as

$$\langle \hat{O} \rangle = \frac{1}{Z} \int \mathcal{D}\phi \, \hat{O}(\phi) \, e^{-S_E[\phi]}. \tag{2.9}$$

By replacing the Minkowski action S with a Euclidean action S_E, we have transformed the amplitude e^{iS} to a statistical Boltzmann factor e^{-S_E}. This new form will allow us to use statistical techniques to simulate quantum fields.

2.1.2 ϕ^4 model

The simplest interacting field theory is known as the ϕ^4 model. This theory describes a spin-0 boson and consists of a real scalar field given by the action

$$S[\phi] = \int d^D x \left[\frac{1}{2} \partial^\mu \phi \partial_\mu \phi - \frac{1}{2} m_0^2 \phi^2 - \frac{\lambda}{4} \phi^4 \right]. \tag{2.10}$$

The first two terms describe a free relativistic particle of mass m_0 while the last term describes an interaction with strength λ. Note that we have generalized the action for D dimensions. Per Einstein summation notation, there is an implicit sum over spacetime dimensions $\mu \in \{0, 1, 2, 3\}$ indexing the derivative vectors[2]

$$\partial^\mu = \left(\frac{\partial}{\partial t}, \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right)$$

$$\partial_\mu = \left(\frac{\partial}{\partial t}, -\frac{\partial}{\partial x}, -\frac{\partial}{\partial y}, -\frac{\partial}{\partial z} \right).$$

Following Sec. 2.1.1, we calculate the Euclidean action in 1+1 dimensions as

$$S_E[\phi] = \int d^2 x_E \left[\frac{1}{2} \left(\partial_t \phi \right) + \frac{1}{2} \left(\partial_x \phi \right) + \frac{1}{2} m_0^2 \phi^2 + \frac{\lambda}{4} \phi^4 \right]. \tag{2.11}$$

where ∂_t and ∂_x are the two derivatives in 1+1 Euclidean spacetime.

This field features spontaneous symmetry breaking at a critical value of m_0^2. In practice this property causes the field to spontaneously align, similarly to spins aligning in a ferromagnet. The name "symmetry breaking" refers to the transformation $\phi \to -\phi$, which changes the values of observables in the aligned regime but not the disordered regime. These two phases are known as the "broken" and "symmetric" phases and their transition is well understood.

2.1.3 Non-linear sigma model

The non-linear sigma model (NLSM) is a prototypical theory for many physical phenomena, including applications in string theory. As a simple nonperturbative model,

---

[2]This canonical representation of the kinetic term $\frac{1}{2} \partial^\mu \phi \partial_\mu \phi$ is equivalent to $\frac{1}{2} \dot{\phi}^2 - \frac{1}{2} \left(\nabla \phi \right)^2$.

it provides an ideal starting point for lattice QCD studies. Specifically, the NLSM exhibits many properties shared by Yang-Mills gauge theories, such as a mass gap, asymptotic freedom and $O(2)$ renormalizability. Furthermore, it has an exact application to Heisenberg ferromagnets in condensed matter.

Unlike the phi^4 model, which consists of a real value at each point in spacetime, the $O(3)$ NLSM consists of a 3D unit vector at each point. For this reason, every transformation of the field must be "norm preserving". Per its name, the non-linear sigma model features a global symmetry under the 3D orthogonal group $O(3)$. In other words, the theory is invariant under rotating all vectors the same way. To differentiate it from the ϕ^4 model, we denote the NLSM field as $\vec{e}(x)$.

The theory is defined by the Euclidean action

$$S_E = \frac{\beta}{2} \int d^D x \partial^\mu \vec{e} \cdot \partial_\mu \vec{e} \tag{2.12}$$

subject to the constraint that $\vec{e} \cdot \vec{e} = 1$. Here, β is the inverse coupling.

2.1.4 NLSM θ term

In order to make the non-linear sigma model topologically nontrivial, we introduce a θ term into the action such that

$$S[\vec{e}] \to S[\vec{e}] - i\theta Q[\phi], \tag{2.13}$$

where Q is the topological charge (see Sec. 2.4).

2.2 Markov Chain Monte Carlo

To accomplish a statistical analysis of quantum fields, we use a Monte Carlo simulation. This method amounts to producing a large number of configurations and calculating statistics on the sample. A brute-force calculation over all possible configurations—as Eq. 2.9 suggests— is clearly infinite and computationally infeasible. However, the

8

exponential nature of the Boltzmann factor dictates that only configurations near the action minimum contribute to observable statistics. Therefore, by selecting a sample of configurations near this minimum, we are able to extract meaningful results with a finite computation.

2.2.1 The Markov Chain

The Markov chain is an effective method to identity these low-action states. Essentially, we begin with a random configuration and then make small adjustments, gradually lowering the action. Starting with a field configuration ϕ_a, we propose a new field ϕ_b. The probability of accepting this change, thereby adding the configuration to the chain, is given by the function

$$P(\phi_a \to \phi_b).$$

There are four requirements that this function must obey to produce a Boltzmann distribution of samples:

1. $P(\phi_a \to \phi_b)$ must depend only on the configurations ϕ_a and ϕ_b.

2. The probability must be properly normalized, i.e. $\sum_\phi P(\phi_a \to \phi) = 1$.

3. Every configuration must be reachable in a finite number of steps. In other words, the chain must be ergodic.

4. In order to reach equilibrium, the chain must be reversible. In other words, the probability of a $\phi_a \to \phi_b$ transition must be equal to the probability of a $\phi_b \to \phi_a$ transition. Mathematically, this condition takes the form of *detailed balance equations*:

$$P(\phi_a) \, P(\phi_a \to \phi_b) = P(\phi_b) \, P(\phi_b \to \phi_a), \tag{2.14}$$

where $P(\phi)$ is the probability of a system existing in state ϕ.

This final condition will allow us to explicitly define the transition probability using the action. From the Boltzmann distribution, we know

$$P(\phi) = \frac{1}{Z}e^{-S_E[\phi]}.$$ (2.15)

Therefore, be rearranging Eq. 2.14, we find

$$\frac{P(\phi_a \to \phi_b)}{P(\phi_b \to \phi_a)} = e^{-(S_E[\phi_b]-S_E[\phi_a])}.$$ (2.16)

This formula will provide the explicit probabilities of the Metropolis and Wolff algorithms.

2.3   Observables

To extract physics from Monte Carlo simulations, we define a set of "observables". These quantities manifest as expectation values of operators, calculated using the Euclidean path integral formula (Eq. 2.9). We can classify these observables into two categories: primary and secondary observables. Primary observables are calculated as expectation values of global operators while secondary observables are derived from these quantities.

2.3.1   Primary Observables

Each primary observable is defined on each configuration independently, meaning they do not encode ensemble statistics of the Markov Chain.[3] In the ϕ^4 model, we can develop an intuition around these quantities by visualizing the symmetric and broken phases. Fig 2.1 and Tab. 2.1 show examples of these quantities in three different configurations: one in the broken phase, one in the symmetric phase, and one at the transition.

---

[3]One slight exception is the magnetic susceptibility, which can be defined both ways.

There are two potential points of confusion here. The first lies in the definition of "broken" phase. Though the symmetric phase more closely resembles a pane of broken glass, it leaves $\phi \to -\phi$ symmetry *un*-broken, thereby giving the title "broken" to the more visually uniform configuration. An additional potential pitfall is the distinction between the *lattice average* and the *ensemble average*. The first is a mean over all lattice sites while the second is a mean over all configurations in the Markov Chain.

| (a) broken phase | (b) transition | (c) symmetric phase |
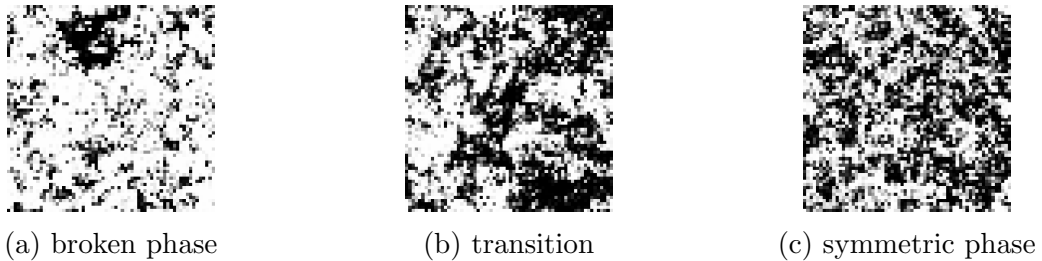
Figure 2.1: Visualization of broken phase, symmetric phase and transition. Simulation run on 64×64 lattice, plotted after 1000 sweep thermalization.

	broken	transition	symmetric
$\|\bar{\phi}\|$	0.56	0.07	0.02
χ_m	1274.05	22.37	1.98
$\S_E/L^2$	0.29	0.40	0.44

Table 2.1: Primary observables with average action per site corresponding to the particular configurations in Fig. 2.1.

Average Magnetization

The average magnetization quantifies the total alignment of the field. In the ϕ^4 theory, this value is a scalar and indicates the spontaneous symmetry breaking of the field. In the NLSM, this value is a vector and also represents a spontaneous symmetry breaking, though this quality is outside the scope of this study. In both theories, a value of zero indicates the symmetric phase while a nonzero value indicates broken symmetry. In the NLSM, a magnitude of one represents total alignment.

Mathematically, we can define this quantity in the ϕ^4 model as

$$|\bar{\phi}| \equiv \frac{1}{L^2} \sum_i^{L^2} \phi(x_i).$$

(2.17)

In the NLSM, the equivalent value is defined as

$$|\bar{\vec{e}}| \equiv \frac{1}{L^2} \sum_i^{L^2} \vec{e}(x_i).$$

(2.18)

Magnetic Susceptibility

2.3.2   Secondary Observables

Binder Cumulant

2.3.3   Bimodality

2.3.4   Jackknife Method

2.4   Topological Observables

The $O(3)$ non-linear sigma model model features topological properties originating from two properties:

1. At $x \to \infty$, the field must become uniform since the Lagrangian must vanish. This allows us to model $x \to \infty$ as a single point on the field, forming a Riemann sphere in three dimensions.

2. The elements of the $O(3)$ non-linear sigma model are three-dimensional unit vectors, thereby existing on a three dimensional unit sphere.

With these two properties, we can view the field as a continuous mapping between two 3D spheres, denoted as S^2, and associate an integer number of wrappings to each mapping from S^2 to S^2. We can envision a tangible metaphor for this wrapping with a balloon and a baseball: by simply inserting the baseball into the balloon, we have established a mapping from every point on the balloon to every point on the baseball.
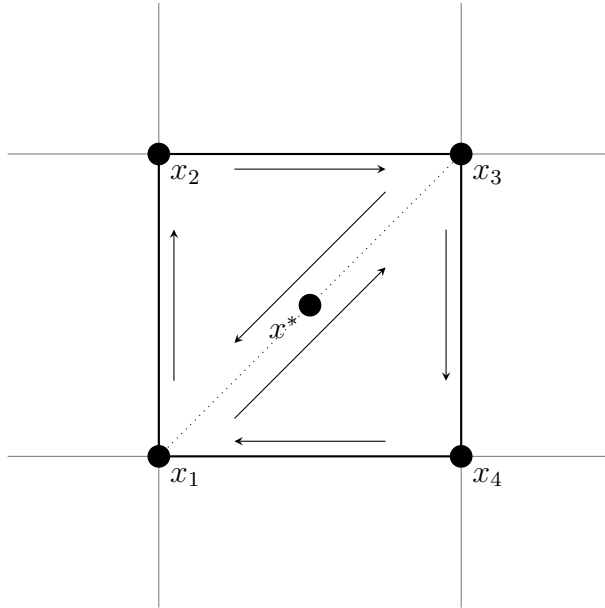
Figure 2.2: Visualization of plaquette x^*. Dotted line separated plaquette into two signed areas which are used to define the topological charge density $q(x^*)$. Arrows represent order of signed area.

We can create an equally valid map by twisting the balloon's mouth and wrapping the baseball again. In a purely mathematical world, we perform this process an infinite number of times, thereby associating every possible mapping with an integer. The group of integers is known as the *homotopy group* of the non-linear sigma model. We associate every field configuration with an element of this group, known as the *topological charge*, which we denote as Q.

In practice, this quantity is nontrivial to calculate. To begin, we define a local topological charge density q, defined for each square of adjacent lattice points. This square, known as a *plaquette*, is denoted x^*. The global charge Q is the sum of all local charges:

$$Q \equiv \sum_{x^*} q(x^*). \tag{2.19}$$

As a function of x^*, the charge density is a function of the field on the plaquette vertices, an idea visualized in Fig. 2.2. In the NLSM, the field at each of these

vertices is a point on the sphere S^2. Therefore, there is a signed area A on the sphere associated with each triplet of points (the sign changes with odd permutations of the ordering). We follow the derivation in [6] and split the plaquette into two triangles, as shown in Fig. 2.2, with the ordering determining the sign. The topological charge density is defined using this signed area as

$$q(x^*) = \frac{1}{4\pi}\left[A\Big(\vec{e}(x_1), \vec{e}(x_3), \vec{e}(x_3)\Big) + A\Big(\vec{e}(x_1), \vec{e}(x_3), \vec{e}(x_4)\Big)\right]. \qquad (2.20)$$

This sign is defined if $A \neq 0, 2\pi$, or in other words, as long as the three points on the sphere are distinct and do not form a hemisphere. In numerical calculations, these points can be ignored. Therefore, we impose that the signed area is defined on the smallest spherical triangle, or mathematically

$$-2\pi < A < 2\pi. \qquad (2.21)$$

Following [6], this yields an expression for the signed area

$$A(\vec{e}_1, \vec{e}_2, \vec{e}_1) = 2\arg\Big(1 + \vec{e}_1 \cdot \vec{e}_2 + \vec{e}_2 \cdot \vec{e}_3 + \vec{e}_3 \cdot \vec{e}_1 + i\vec{e}_1 \cdot (\vec{e}_2 \times \vec{e}_3)\Big). \qquad (2.22)$$

Under periodic boundary conditions, these triangles on the sphere necessarily wrap S^2 an integer number of times, ensuring Q is an integer.

Following this quantity, we can define a topological susceptibility χ_t

$$\chi_t \equiv \frac{1}{L^2}\Big(\langle Q^2 \rangle - \langle Q \rangle^2\Big). \qquad (2.23)$$

$$\chi_t \equiv \qquad (2.24)$$

2.5   Ultraviolet Divergences

In Section 2.1.1, we defined a fundamental equation of quantum fields using a "path integral" which encompasses an uncountably infinite configuration space. However,

we said nothing of the integral's convergence. In fact, many fundamental processes in QFT have divergent amplitudes, yielding nonsensical results. The most type of divergence stems from high-momentum states, giving them the name "ultraviolet divergences". The remedy to this catastrophe is unintuitive. Essentially, we adopt infinite values for the parameters of the Lagrangian (m_0^2 and λ in ϕ^4 theory). Since neither of these two quantities is ever measured directly, we do not have to assume that their values are finite. In practice, this technique is involved and consists of two steps: regularization and renormalization.

2.5.1 Regularization: fields on the lattice

Regularization is a process which introduces a new parameter into calculations. One example is a momentum cutoff. This technique transforms infinite momentum integrals as follows:

$$\int_0^\infty dk \to \int_0^\Lambda dk,$$

introducing Λ as a regularization parameter. This process makes results Λ-dependent, but finite. Another example is dimensional regularization, which calculates results in terms of the spacetime dimension d and analytically continues this parameter into the real numbers.

In this study, we employ lattice regularization. This process discretizes the field, modeling the field $\phi(x)$ as a lattice ϕ_i where i indexes lattice sites. The inherent parameter in this case is the lattice spacing a which measures the width of each lattice chunk.

2.5.2 Renormalization

After regularization, we redefine the Lagrangian parameters in terms of the regularization parameter, following a handful of boundary conditions. In this study, we

assert that L/ξ remains constant, where L is the side length of the system and ξ is the coherence length. In perturbation theory, this process is arduous and includes the introduction of counter-terms into the Lagrangian. In the case of the non-linear sigma model, it is impossible using counterterms[4]but can be performed numerically. In this study, we use predetermined values from [3].

At this point, we can calculate observables as functions of regularization parameters. To achieve physical values, we take the limit as the regularization parameters approach their physical values. With a momentum cutoff, we take $\Lambda \to \infty$ and with dimensional regularization we usually take $d \to 4$. With lattice regularization, we approach the continuum, taking the lattice spacing $a \to 0$.

At this point, we have surely eliminated all divergences, right? Unfortunately, this is not always the case. The topological susceptibility χ_t is one such value that diverges in the continuum limit. As we decrease the width of each lattice site, high frequency modes become more significant, leading to an ultraviolet divergence in the operator.

2.5.3 The Gradient Flow

To remove this ultraviolet divergence, we adopt a technique is "smearing", a local averaging of the field.[7] Specifically, we use a technique known as the "gradient flow" [8] which introduces a new half-dimension called "flow time", or τ.[5] The flow time parameterizes the smearing such that an evolution in flow time corresponds to suppressing ultraviolet divergences.

Specifically, the gradient flow pushes field configuration toward classical minima of the action. Additionally, renormalized correlation functions remain renormalized at nonzero flow time.[9] In 2D ϕ^4 scalar field theory, the gradient flow is defined by

---

[4]*citation needed*
[5]The term "half-dimension" indicates that $\tau > 0$.

the differential equation

$$\frac{\partial \rho(\tau, x)}{\partial \tau} = \partial^2 \rho(\tau, x) \tag{2.25}$$

where ∂^2 is the Laplacian in 4-D Euclidean spacetime and τ is the flow time. Here, ρ is the field evolved into a nonzero flow time, bounded by the condition $\rho(\tau = 0, x) = \phi(x)$. In the ϕ^4 theory, we can solve this equation exactly to find [2]

$$\rho(\tau, x) = \frac{1}{4\pi\tau} \int d^2 y\, e^{-(x-y)^2/4\tau} \phi(y). \tag{2.26}$$

This function forms a Gaussian, smoothly dampening high-momentum modes and removing ultraviolet divergences from evolved correlation functions.[10] We can visualize this by plotting the ϕ field, shown in Fig. 2.3. These plots demonstrate the reduction of high momentum modes.

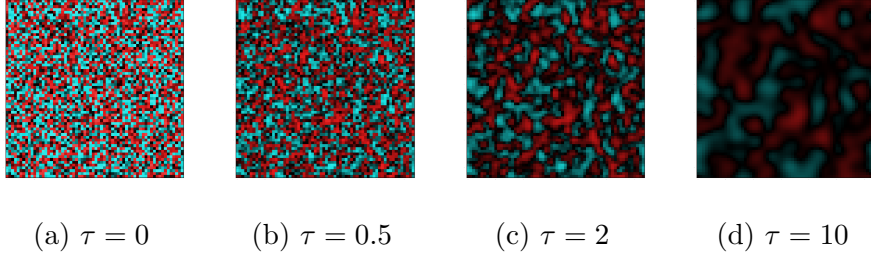(a) $\tau = 0$ (b) $\tau = 0.5$ (c) $\tau = 2$ (d) $\tau = 10$

Figure 2.3: Effect of flow time evolution on a random lattice in the symmetric phase. Red and blue indicate positive and negative values of the field in 2D spacetime.

Generally, we can choose any flow time equation that drives the field towards a classical minimum. Beyond the ϕ^4 model, we need flow equations that incorporate different types of fields. In the non-linear sigma model, it has been shown that an appropriate manifestation of the flow time can resolve divergences [10]. Therefore, it is possible to define a different flow equation for this model as well. Following [3], we can define the gradient flow in this model via the differential equation

$$\partial_\tau \vec{\rho}(\tau, x) = \left(1 - \vec{\rho}(\tau, x)\vec{\rho}(\tau, x)^T\right) \partial^2 \vec{\rho}(\tau, x), \tag{2.27}$$

17

where ∂^2 is the Laplacian operator in Euclidean space[6]. We solve this equation numerically using the boundary condition $\vec{\rho}(0, x) = \vec{e}(x)$.

---

[6]Explicitly, $\partial^2 = \frac{\partial^2}{\partial t^2} + \nabla^2$

Chapter 3

Methods

Our study of the gradient flow in the non-linear sigma model consists of a computational part and an analytical part. We begin by outlining a numerical Monte Carlo method to simulate the lattice in two and three dimensions. We verify our program with the well-studied ϕ^4 scalar field theory. We then generalize our model to a vector field to simulate the non-linear sigma model. This simulation system provides data on vacuum states with which we study the gradient flow's effect on topology.

We implement these two algorithms first in Python for the ϕ^4 model. Afterwards, we transition to C/C++ code due to the increased speed for more complicated theories.

3.1   Fields on the Lattice

To implement the lattice regularization technique, we must redefine the field action in terms of discrete positions, a process known as "discretization". We transition from x as a continuous vector in $\mathbb{R}^2$ to $x_{i,j}$ where

$$x_{i,j} = ia\hat{\mu}_t + ja\hat{\mu}_x. \tag{3.1}$$

Here, a is the lattice constant and $\hat{\mu}_0$ and $\hat{\mu}_1$ are unit vectors. This change effectively shifts the domain of the field from $\mathbb{R}^2$, which is uncountably infinite, to $\mathbb{Z}^2$,

which is countably infinite. To achieve a finite domain, we impose periodic boundary conditions, such that

$$\phi\left(x_{i+L,j}\right) = \phi\left(x_{i,j+L}\right) = \phi\left(x_{i,j}\right) \tag{3.2}$$

where L is the side length (in units of the lattice constant a) of the system. In this study we focus solely on square geometries and thus the side length L is unambiguous.

In the ϕ^4 model, we can specify a discrete action using the Euclidean action from Eq. 2.11. We begin by redefining the derivative operator as a difference:

$$\partial_\mu \phi = \frac{\phi\left(x + a\hat{\mu}\right) - \phi(x)}{a} \tag{3.3}$$

We can then define the kinetic term

$$\frac{1}{2}\left(\partial_t \phi\right)^2 + \frac{1}{2}\left(\partial_x \phi\right)^2 \to \frac{1}{2a^2}\left[\left(\phi(x + a\hat{t}) - \phi(x)\right)^2 + \left(\phi(x + a\hat{x}) - \phi(x)\right)^2\right] \tag{3.4}$$

$$\to \frac{1}{2a^2}\left[\phi^2(x + a\hat{t}) + \phi^2(x + a\hat{x}) + 2\phi^2(x) \right. \\ \left. - 2\phi(x + a\hat{t})\phi(x) - 2\phi(x + a\hat{x})\phi(x)\right] \tag{3.5}$$

Since we will eventually sum over all sites x, the periodic boundary conditions imply that an overall shift in x does not effect the final action. Therefore, we can combine the first two terms with the third term to produce

$$\frac{1}{2}\left(\partial_t \phi\right)^2 + \frac{1}{2}\left(\partial_x \phi\right)^2 \to \frac{1}{a^2}\left[2\phi^2(x) - \phi(x + a\hat{t})\phi(x) - \phi(x + a\hat{x})\phi(x)\right] \tag{3.6}$$

Unlike the kinetic term, the mass and interaction terms remain unchanged under the discretization procedure. The only remaining change is a shift from an integral to a sum. This takes the form

$$\int dt dx \to a^2 \sum_i \tag{3.7}$$

such that the final discretized action becomes

$$S_{\text{lat}}[\phi] = \sum_i \left[-\phi(x_i + a\hat{t}) - \phi(x_i + a\hat{x}) + \left(2 + \frac{1}{2}m_0^2\right)\phi^2(x_i) + \frac{1}{4}\lambda\phi^4(x_i)\right] \tag{3.8}$$

Likewise, we can discretize the non-linear sigma model. In this case, the derivative term becomes

$$\frac{1}{2}\left(\partial_t \vec{e}\right)^2 + \frac{1}{2}\left(\partial_x \vec{e}\right)^2 \to \frac{1}{a^2}\left[2 - \vec{e}(x + a\hat{t})\cdot\vec{e}(x) - \vec{e}(x + a\hat{x})\cdot\vec{e}(x)\right]. \qquad (3.9)$$

Note that we have used the identity $\vec{e}\cdot\vec{e} = 1$. Inserting this into equation yields the discretized action

$$S_{\text{lat}}[\vec{e}] = \sum_i \left[2 - \vec{e}(x + a\hat{t})\cdot\vec{e}(x) - \vec{e}(x + a\hat{x})\cdot\vec{e}(x)\right]. \qquad (3.10)$$

Finally, we redefine the gradient flow in on the lattice. Since the gradient flow is solved exactly in the ϕ^4 model, we rely on a Fast Fourier Transform. *TODO: explicit expression for this* In the NLSM, the definition of the gradient flow (Eq. 2.27) becomes

$$\partial_\tau \vec{e}(\tau, x) = \left(1 - \vec{e}(\tau, x)\vec{e}(\tau, x)^T\right)\partial^2 \vec{e}(\tau, x), \qquad (3.11)$$

where the Laplacian operator ∂^2 is defined as

$$\partial^2 \vec{e}(\tau, x) = \vec{e}(\tau, x + a\hat{t}) + \vec{e}(\tau, x - a\hat{t}) + \vec{e}(\tau, x + a\hat{x}) + \vec{e}(\tau, x - a\hat{x}) - 2d\vec{e}(t)_x.$$

3.2   Monte Carlo Simulations

We implement a Markov Chain Monte Carlo method following Schaich's thesis [11]. This implementation utilizes a "random walk," i.e. a set of random steps through phase space, to determine statistical values such as correlation functions across the lattice. By the definition of the Markov chain, the probability of adoption of each state, and therefore its inclusion in the Monte Carlo calculation, depends only on the current state and the proposed state. This probability is denoted as $P(\mu \to \nu)$ where μ and ν are the existing and proposed lattice configurations respectively. We use a combination of the Metropolis and Wolff algorithms to determine this value.

3.2.1 Metropolis Algorithm

We primarily use the Metropolis algorithm for the calculation of new Markov chain configurations. We begin with a so-called "hot start," where each field value at each lattice site is randomly selected. Then we propose a new value for a single lattice point, which is accepted with a probability

$$P(\phi_a \rightarrow \phi_b) = \begin{cases} e^{-(S[\phi_b]-S_[\phi_a])} & S[\phi_b] < S_[\phi_a] \\ 1 & \text{otherwise} \end{cases} \tag{3.12}$$

where ϕ_a is the initial configuration and ϕ_b is the proposed configuration. This process is performed for each point on the lattice, making up a "sweep." Repeating this sweep many times pushes the lattice toward the action minimum.

3.2.2 Wolff Cluster Algorithm

Though the Metropolis algorithm will slowly find the absolute minimum of the theory, the presence of local minima can greatly prolong the convergence. Both the ϕ^4 model and the non-linear sigma model feature "kinetic" terms with gradients of ϕ. Therefore, the presence of large similarly-valued regions in the lattice can lead to a local minimum. The Wolff algorithm helps reduce the presence of these clusters through two steps: identifying a cluster and flipping it along some arbitrary vector. In the case of ϕ^4 theory, this flipping takes the form of a simple sign change. In the non-linear sigma model we choose a random unit vector $\vec{r}$ and consider the projection of the field on this vector. When the cluster flips, each site is flipped along this direction.[1] To identify the cluster, the algorithm uses a recursive algorithm defined by the probability of adding a new site.

We define a cluster C as a continuous set of aligned sites which is generated probabilistically. In ϕ^4 theory, we define "aligned" as having the same sign. In the

---

[1] *citation needed*

22

non-linear sigma model, we choose a random unit vector $\vec{r}$ and use it to project the field vector. The sign of this quantity defines "aligned" in the $O(3)$ NLSM. In order to find the probability p_i of adding a site x_i to the cluster C, we use the detailed balance equation (Eq. 2.14) which relates the Markov chain probability $P(\phi_a \to \phi_b)$ to the actions of each configuration. In the context of the Wolff cluster algorithm, we will notate this probability as $P(\phi \to f_C(\phi))$ and express it as

$$P(\phi \to f_C(\phi)) = \left(\prod_{i \in C} p_i\right)\left(\prod_{j \in N} (1 - p_j)\right) \tag{3.13}$$

where N is the set of aligned neighboring sites. Since $p_i = 0$ for sites that are not aligned with the cluster, they are not included in this expression. Now if we consider the probability of flipping the cluster back, i.e. $P(f_C(\phi) \to \phi)$, we achieve a similar expression:

$$P(f_C(\phi) \to \phi) = \left(\prod_{i \in C} p_i\right)\left(\prod_{j \in M} (1 - p_j)\right) \tag{3.14}$$

where M is the set of initially unaligned sites (they become aligned after the first flip).

Using the detailed balance equations, we can now relate these quantities with the actions $S[\phi]$ and $S[f_C(\phi)]$:

$$\frac{\prod_{j \in N} (1 - p_j)}{\prod_{j \in M} (1 - p_j)} = e^{-(S[f_c(\phi)] - S[\phi])}. \tag{3.15}$$

TODO: add end of derivation

$$p_i = 1 - e^{-\beta \vec{e}_a \cdot \vec{e}_b}. \tag{3.16}$$

Fig. 3.1 shows a real demonstration of this process. We can see a large cluster of negative field values becoming positive. Note that periodic boundary conditions apply, so the small island of black at the bottom is actually part of the larger cluster.

Furthermore, this visualization demonstrates the probabilistic nature of the Wolff algorithm. Since states are added probabilistically, there are some small holes in the cluster. These will be removed by following Metropolis sweeps.
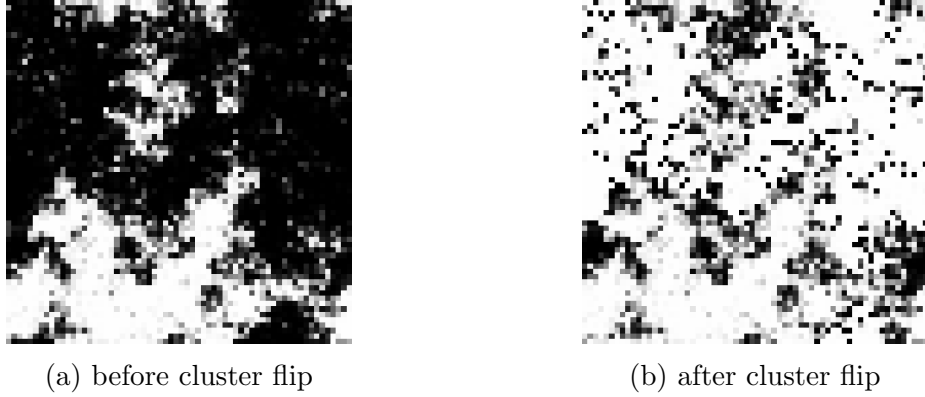

(a) before cluster flip   (b) after cluster flip

Figure 3.1: An example of the Wolff cluster algorithm in the ϕ^4 model. White represents negative values of ϕ while black represents positive. $\lambda = 0.5$, $m_0^2 = -0.9$

3.2.3 Checkerboard algorithm

In order to parallelize the Metropolis algorithm, we use a checkerboard algorithm. Since the Lagrangian density at each site does not depend on any sites diagonal to it, the lattice can be split into "white" sites and "black" sites, like the tiles on a checkerboard. Each white site is independent of every other white site and likewise with black sites. Therefore, we can split the sites of each color into separate parallel processing nodes and independently run the Metropolis algorithm, ensuring that no site affects the Lagrangian density at any other site. We use this method to parallelize the code through the Message Passing Interface (MPI).

3.2.4 Autocorrelation times and thermalization

One important aspect is the correlation between different states of the Markov chain. Ideally, each sample of the lattice would be completely independent, but this is not

the case. Therefore, we must sweep the lattice sufficiently to produce an effective Monte Carlo result. We use Wolff's automatic windowing procedure [12] and the magnetic susceptibility χ_m to estimate the autocorrelation and determine an appropriate number of sweep between measurements for each theory. Furthermore, the initial values of our Markov chain will be highly correlated with the initial random start. Therefore, we also wait a set number of sweeps before taking any measurements so that the lattice can approach a minimum of the action.

3.2.5 Runge-Kutta Algorithm

In order to calculate the gradient flow, we numerically solve the differential equation using a fourth-order Runga-Kutta approximation. This algorithm refines the Euler method

$$\vec{e}(\tau + h)_x \approx \vec{e}(\tau)_x + h f(\vec{e}(\tau)_x).$$

where $f(\vec{e})$ is defined for convenience as

$$f(\vec{e}) = \partial_\tau \vec{e}(\tau)_x = \left(1 - \vec{e}(\tau)_x \vec{e}(\tau)_x^T\right) \partial^2 \vec{e}(\tau)_x, \tag{3.17}$$

following from Eq. 3.11. To the fourth order, this approximation becomes

$$k_1 = h f\left(\vec{e}\left(\tau\right)_x\right) \tag{3.18}$$

$$k_2 = h f\left(\vec{e}\left(\tau\right)_x + \frac{k_1}{2}\right) \tag{3.19}$$

$$k_3 = h f\left(\vec{e}\left(\tau\right)_x + \frac{k_2}{2}\right) \tag{3.20}$$

$$k_4 = h f\left(\vec{e}\left(\tau\right)_x + k_3\right) \tag{3.21}$$

$$\vec{e}(\tau + h)_x = \vec{e}(\tau)_x + \frac{k_1}{6} + \frac{k_2}{3} + \frac{k_3}{3} + \frac{k_4}{6} + O(h^5). \tag{3.22}$$

This method is usually superior to Euler's method and the midpoint method [13].

To increase the efficiency of this algorithm, we implement the step-doubling algorithm to adaptively adjust h. If the error of a Runge-Kutta step is greater than the

tolerance, the same step is repeated with half the step size. Alternatively, if the error is less than half of the tolerance, the step size is doubled for the next calculation. Finally, if the step size is greater than the distance to the next measurement, that distance is used as the step size, using the normal value afterwards. Otherwise, the algorithm proceeds with the consistent step size.

Chapter 4

Results

4.1 ϕ^4 model

We initially implemented the ϕ^4 model using the Monte Carlo method described in Sec. 3 to verify the results of our system. According to previous studies ([2], [11]), the ϕ^4 model exhibits a symmetric and broken phase depending on its parameters m_0^2 and λ, specifically occurring at $m_0^2 = -0.72$ for $\lambda = 0.5$. We verify this result by plotting four observables: the lattice average $|\langle\bar{\phi}\rangle|$, the lattice variance χ, the Binder cumulant U and the bimodality B in Fig. 4.1.
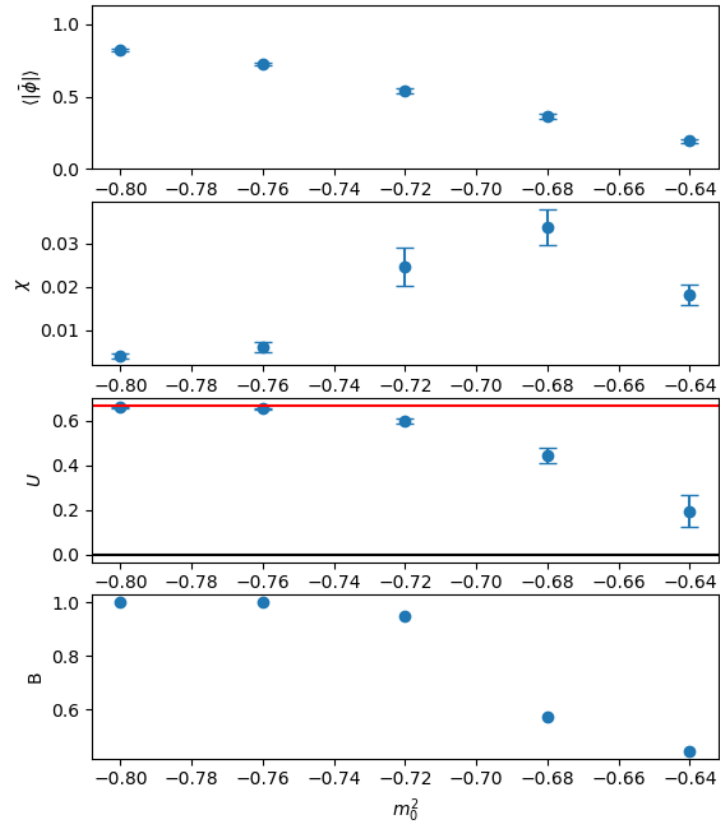
Figure 4.1: Simulation of the phase transition in the ϕ^4 model: Magnetization $\langle|\bar{\phi}|\rangle$, magnetic susceptibility χ_m, Binder cumulant U and bimodality B for various values of the mass squared m_0^2.

Chapter 5

Statistical Analysis

Chapter 6

Conclusion & Outlook

Appendix A

My Great Computer Program

A.1   Code sample

Bibliography

[1] Victor F. Weisskopf. Development of field theory in the last 50 years. *Physics today*, 34(11):69, 1981.

[2] Christopher Monahan. The gradient flow in simple field theories. In *Proceedings of The 33rd International Symposium on Lattice Field Theory — PoS(LATTICE 2015)*, page 052, Kobe International Conference Center, Kobe, Japan, July 2016. Sissa Medialab.

[3] Wolfgang Bietenholz, Philippe de Forcrand, Urs Gerber, Héctor Mejía-Díaz, and Ilya O. Sandoval. Topological Susceptibility of the 2d O(3) Model under Gradient Flow. *Phys. Rev. D*, 98(11):114501, December 2018.

[4] Anthony Zee. *Quantum Field Theory in a Nutshell*, volume 7. Princeton university press, 2010.

[5] Richard P. Feynman, Robert B. Leighton, and Matthew L. Sands. The Principle of Least Action. In *The Feynman Lectures on Physics*, volume II of *World Student Series*. Addison-Wesley Pub. Co., Reading, Mass., 1963.

[6] B Berg and Martin Lüscher. Definition and statistical distributions of a topological number in the lattice $O(3)$ σ-model. *Nuclear Physics B*, 190(2):412–424, 1981.

[7] Stefan Solbrig, Falk Bruckmann, Christof Gattringer, Ernst-Michael Ilgenfritz, Michael Müller-Preussker, and Andreas Schäfer. Smearing and filtering methods in lattice QCD - a quantitative comparison. *arXiv:0710.0480 [hep-lat]*, October 2007.

[8] Christopher Monahan and Kostas Orginos. Locally smeared operator product expansions in scalar field theory. *Phys. Rev. D*, 91(7):074513, April 2015.

[9] Martin Lüscher. Chiral symmetry and the Yang–Mills gradient flow. *J. High Energ. Phys.*, 2013(4):123, April 2013.

[10] Hiroki Makino and Hiroshi Suzuki. Renormalizability of the gradient flow in the 2D $O(N)$ non-linear sigma model. *Progress of Theoretical and Experimental Physics*, 2015(3):33B08–0, March 2015.

[11] David A Schaich and William Loinaz. *Lattice Simulations of Nonperturbative Quantum Field Theories*. PhD thesis, Amherst College, May 2006.

[12] Ulli Wolff. Monte Carlo errors with less errors. *Computer Physics Communications*, 176(5):383, March 2007.

[13] William T Vetterling, William T Vetterling, William H Press, William H Press, Saul A Teukolsky, Brian P Flannery, and Brian P Flannery. *Numerical Recipes: Example Book c*. Cambridge University Press, 1992.