

A Wonderful Piece of Scientific Research

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 | iii |
| List of Figures | iv |
| List of Tables | v |
| Abstract | v |
| 1 Introduction | 1 |
| 1.1 Conventions | 2 |
| 2 Theory | 3 |
| 2.1 Quantum Field Theory | 3 |
| 2.1.1 Path Integral Formulation | 4 |
| 2.1.2 Ultraviolet Divergences | 6 |
| 2.2 Markov Chain Monte Carlo | 6 |
| 2.2.1 The Markov Chain | 6 |
| 2.3 Gradient Flow | 7 |
| 2.4 Non-linear sigma model | 9 |
| 2.5 Topological Charge | 10 |
| 3 Methods | 11 |

| | | |
|----------|--|-----------|
| 3.1 | Monte Carlo Simulations | 11 |
| 3.1.1 | Metropolis Algorithm | 12 |
| 3.1.2 | Wolff Cluster Algorithm | 12 |
| 3.1.3 | Checkerboard algorithm | 13 |
| 3.1.4 | Autocorrelation times and thermalization | 13 |
| 3.2 | Gradient Flow | 14 |
| 3.2.1 | Runge-Kutta Algorithm | 14 |
| 4 | Results | 16 |
| 5 | Statistical Analysis | 17 |
| 6 | Conclusion & Outlook | 18 |
| A | My Great Computer Program | 19 |
| A.1 | Code sample | 19 |

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 model 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 framework of particle physics. In condensed matter physics, effective field theories can 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.[?]

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* which provides a prototypical theory in both condensed matter and particle physics. In solid-state systems, this model describes Heisenberg ferromagnets. In nuclear physics, the non-linear sigma model acts as a prototype for quantum chromodynamics, exhibiting characteristic

features such as a mass gap and asymptotic freedom.

In this study, we specifically consider the $O(3)$ nonlinear 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 a nonphysical length scale a , the lattice spacing. Therefore, we expect any physical result to converge in the continuum limit, i.e. when $a \rightarrow 0$. 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 to remove these divergences. By dampening high-momentum fluctuations, the gradient flow reduces power-divergent mixing and makes observables finite on the lattice.[1]

1.1 Conventions

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

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 an alternative formulation known as the path integral formulation (for a similar introduction see Zee’s textbook [2]).

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 formula¹

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

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 generalizes the classical principle of least action, a relationship that Richard Feynman pedagogically explains in his lectures on physics [3].

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 the exponent currently prohibits us from making this jump. To remedy this issue,

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

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

$$s = x_0^2 - x_1^2 - x_2^2 - x_3^2 \quad (2.2)$$

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

$$s_E = x_1^2 + x_2^2 + x_3^2 + x_4^2, \quad (2.3)$$

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

$$d^4x_E = d^3\vec{x}dx_4 \quad (2.4)$$

$$= id^3\vec{x}dx_0 \quad (2.5)$$

$$= id^4x. \quad (2.6)$$

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

$$S_E = \int d^4x_E \mathcal{L}_E \quad (2.7)$$

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

allowing us to redefine the path integral as

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

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} .

2.1.2 Ultraviolet Divergences

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 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 identify 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

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

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

1. $P(\phi_a \rightarrow \phi_b)$ must depend only on the configurations ϕ_a and ϕ_b .
2. The probability must be properly normalized, i.e. $\sum_{\phi} P(\phi_a \rightarrow \phi) = 1$.
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 $\phi_a \rightarrow \phi_b$ transition must be equal to the probability of a $\phi_b \rightarrow \phi_a$ transition. Mathematically, this condition takes the form of *detailed balance equations*:

$$P(\phi_a) P(\phi_a \rightarrow \phi_b) = P(\phi_b) P(\phi_b \rightarrow \phi_a), \quad (2.10)$$

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

$$P(\phi) = \frac{1}{Z} e^{-S_E[\phi]}. \quad (2.11)$$

Therefore, by rearranging Eq. 2.10, we find

$$\frac{P(\phi_a \rightarrow \phi_b)}{P(\phi_b \rightarrow \phi_a)} = e^{-(S_E[\phi_b] - S_E[\phi_a])}. \quad (2.12)$$

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

2.3 Gradient Flow

To remove ultraviolet fluctuations, we introduce a momentum cutoff, effectively reducing the predictive power of our theory by ignoring very high momentum states. In a gauge theory, such as QCD, a hard cutoff violates gauge invariance. Therefore, renormalization techniques must be more nuanced in the reduction of high-frequency divergences.

One such technique is “smearing,” a local averaging of a gauge field in order to remove diverging fluctuations [4]. In this study, we use a technique known as the “gradient flow” [5] which introduces a new half-dimension called “flow time”.

The flow time parameterizes the smearing; an evolution in flow time corresponds to suppressing ultraviolet divergences. In a 2D ϕ^4 scalar field theory, given by the action

$$S_\phi[\phi] = \frac{1}{2} \int d^2x \left[(\partial\phi)^2 + m^2\phi^2 + \frac{\lambda}{2} \right], \quad (2.13)$$

we can define the evolution in flow time as

$$\frac{\partial \rho(\tau, x)}{\partial \tau} = \partial^2 \rho(\tau, x) \quad (2.14)$$

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 $\rho(\tau = 0, x) = \phi(x)$. When solved exactly, this function forms a Gaussian which forces operators to converge. Therefore, in all orders of perturbation theory, any correlation function in the gauge field that has been evolved in flow time becomes immune to these divergences[6].

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 divergencies [6]. Therefore, it is possible to define a different flow equation for this model as well. Our goal is to understand this renormalizability in the non-linear sigma model using the gradient flow. Specifically, we seek to define vacuum condensates, or nonperturbative vacuum expectation values.

To explore this technique, we first plan to produce numerical calculations using a non-linear sigma model. Using a Monte Carlo method in 1+1 dimensions, we can calculate the correlation functions of quantum field theories. With more computing power, expanding this simulation to 2+1 dimensions would provide an analysis more relevant to quantum chromodynamics. Following this numerical calculation, we seek

to find and understand an implementation of the gradient flow in this model using analytic methods. This understanding can then be verified by the numerical solution.

2.4 Non-linear sigma model

The non-linear sigma model is a prototypical theory for many physical phenomena, including applications in string theory. As a simple nonperturbative model, it provides an ideal starting point for lattice QCD studies. Specifically, the non-linear sigma model 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. The theory is defined by the Euclidean action

$$S_E = \frac{1}{2g^2} \int d^D x \partial^\mu \vec{\phi} \cdot \partial_\mu \vec{\phi} \quad (2.15)$$

subject to the constraint that $\vec{\phi} \cdot \vec{\phi} = 1$. In this report, we primarily consider $O(3)$, where $\vec{\phi}$ is a 3 dimensional vector.

Following [7], we can define the gradient flow in this model via the differential equation

$$\partial_\tau \vec{\rho}(\tau, x) = (1 - \vec{\rho}(\tau, x) \vec{\rho}(\tau, x)^T) \nabla^2 \vec{\rho}(\tau, x), \quad (2.16)$$

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

In order to use this theory on the lattice, we discretize the action. We replace the derivatives with a simple difference and impose periodic boundary conditions on the system. For a more in-depth discussion, see Section ??.

2.5 Topological Charge

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

1. At $x \rightarrow \infty$, the field must become uniform since the Lagrangian must vanish. This allows us to model $x \rightarrow \infty$ 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. 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 .

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 2 and 3 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 Monte Carlo Simulations

We implement a Markov Chain Monte Carlo method following Schaich's thesis [8]. 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 $P(\mu \rightarrow \nu)$ 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.

3.1.1 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

$$P(\mu \rightarrow \nu) = \begin{cases} e^{-(S_\nu - S_\mu)} & S_\nu > S_\mu \\ 1 & \text{otherwise} \end{cases} \quad (3.1)$$

where S_ν and S_μ are the lattice actions of configuration before and after the change. 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.

3.1.2 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. In the non-linear sigma model, this value is

¹citation needed

derived as

$$P_{add} = 1 - e^{-\beta \vec{e}_a \vec{e}_b}. \quad (3.2)$$

3.1.3 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).

3.1.4 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 [9] 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 Gradient Flow

On the lattice, the definition of the gradient flow (Eq. 2.16) becomes

$$\partial_\tau \vec{e}(\tau)_x = (1 - \vec{e}(\tau)_x \vec{e}(\tau)_x^T) \nabla^2 \vec{e}(\tau)_x, \quad (3.3)$$

where the Laplace operator ∇^2 is defined as

$$\nabla^2 \vec{e}(t)_x = \sum_{\mu=1}^d (\vec{e}(t)_{x+\hat{\mu}} - \vec{e}(t)_{x-\hat{\mu}}) - 2d\vec{e}(t)_x.$$

3.2.1 Runge-Kutta Algorithm

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

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

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

$$f(\vec{e}) = \partial_\tau \vec{e}(\tau)_x = (1 - \vec{e}(\tau)_x \vec{e}(\tau)_x^T) \nabla^2 \vec{e}(\tau)_x, \quad (3.4)$$

following from Eq. 3.3. To the fourth order, this approximation becomes

$$k_1 = hf(\vec{e}(\tau)_x) \quad (3.5)$$

$$k_2 = hf(\vec{e}(\tau)_x + \frac{k_1}{2}) \quad (3.6)$$

$$k_3 = hf(\vec{e}(\tau)_x + \frac{k_2}{2}) \quad (3.7)$$

$$k_4 = hf(\vec{e}(\tau)_x + k_3) \quad (3.8)$$

$$\vec{e}(\tau + h)_x = \vec{e}(\tau)_x + \frac{k_1}{6} + \frac{k_2}{3} + \frac{k_3}{3} + \frac{k_4}{6} + O(h^5). \quad (3.9)$$

This method is usually superior to Euler's method and the midpoint method [10].

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

Chapter 5

Statistical Analysis

Chapter 6

Conclusion & Outlook

Appendix A

My Great Computer Program

A.1 Code sample

Bibliography

- [1] 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.
- [2] Anthony Zee. *Quantum Field Theory in a Nutshell*, volume 7. Princeton university press, 2010.
- [3] 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.
- [4] 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.
- [5] Christopher Monahan and Kostas Orginos. Locally smeared operator product expansions in scalar field theory. *Phys. Rev. D*, 91(7):074513, April 2015.
- [6] 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.

- [7] 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.
- [8] David A Schaich and William Loinaz. *Lattice Simulations of Nonperturbative Quantum Field Theories*. PhD thesis, Amherst College, May 2006.
- [9] Ulli Wolff. Monte Carlo errors with less errors. *Computer Physics Communications*, 176(5):383, March 2007.
- [10] 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.