ASPECTS OF THE RENORMALIZATION GROUP

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Archishman Raju

August 2018

ASPECTS OF THE RENORMALIZATION GROUP

Archishman Raju, Ph.D.

Cornell University 2018

The main part of this thesis is on the renormalization group (RG). We will explore the results of the RG in two ways.

In the first part we use information geometry, in which the local distance between models measures their distinguishability from data, to quantify the flow of information under the renormalization group. We show that information about relevant parameters is preserved, with distances along relevant directions maintained under flow. By contrast, irrelevant parameters become less distinguishable under the flow, with distances along irrelevant directions contracting according to renormalization group exponents. We develop a covariant formalism to understand the contraction of the model manifold. We then apply our tools to understand the emergence of the diffusion equation and more general statistical systems described by a free energy. Our results give an information-theoretic justification of universality in terms of the flow of the model manifold under coarse graining.

In the second part, we use dynamical systems theory to systematize the results of the RG. The results of the RG are commonly advertised as the existence of power law singularities near critical points. The classic predictions are often violated and logarithmic and exponential corrections are treated on a case-by-case basis. We use the mathematics of normal form theory to systematically group these into *universality families* of seemingly unrelated systems united by common scaling variables. We recover and explain the existing litera-

ture, predict the nonlinear generalization for universal homogeneous functions, and show that the procedure leads to a better handling of the singularity with several examples including the Random Field Ising model and the 4-d Ising model.

The RG is useful not just for systems in physics but has found application in a surprising variety of fields. In dynamical systems, it provided an nice explanation of the universality observed in the period doubling transition. We show the equivalent of so-called redundant variables in period doubling and offer a new interpretation for them. We then examine the consequences for the Ising model.

Finally, the last part of this thesis is on a very different topic. Here, we use an effective Hamiltonian to characterize particle dynamics and find escape rates in a periodically kicked Hamiltonian. We study a model of particles in storage rings that is described by a chaotic symplectic map. Ignoring the resonances, the dynamics typically has a finite region in phase space where it is stable. Inherent noise in the system leads to particle loss from this stable region. The competition of this noise with radiation damping, which increases stability, determines the escape rate. Determining this 'aperture' and finding escape rates is therefore an important physical problem. We compare the results of two different perturbation theories and a variational method to estimate this stable region. Including noise, we derive analytical estimates for the steady-state populations (and the resulting beam emittance), for the escape rate in the small damping regime, and compare them with numerical simulations.

BIOGRAPHICAL SKETCH

The author was born in Delhi, India. He spent substantial portions of his school life in Indore and Bhopal, but came back to finish school in Delhi. He finished his B.Sc. (H) in physics from St. Stephen's college at Delhi University in 2013. He then came to Cornell University to pursue his Ph.D. in physics.

After his Ph.D., he will work as a postdoctoral fellow at Rockefeller University.

This document is dedicated to my family.

ACKNOWLEDGEMENTS

There are many people who have shaped my thinking and led me to this point. The philosopher and writer Rahul Sankrityayan, when describing all those who he was indebted to, said (my translation)

"These include not only those, who have guided me or who have taught me something, but all those people, whose association was helpful as mental/intellectual support in my life journey. There are many who, without their knowledge, taught me a lot through their nature and behaviour."

This captures very well how I think of all the influences that have shaped my life up to this point. Here, I can only mention a few who have been particularly important in these past few years during my Ph.D.

First, I would like to thank my parents. In some ways everything about the person I am today, and how I think about the world can be directly traced back to them. They taught me both how to think and what to think about. I would like to thank my brother who has always been an inspiration and a teacher.

I would like to thank my advisor James Sethna. He has both presented me with fascinating research problems and allowed me the freedom to explore on my own. I have learnt a lot from his attitude towards science. He has also shown patience with those interests of mine that went beyond physics. I want to thank my collaborators on projects including Ben Machta, Sayan Choudhury, Amie Wilkinson, David Rubin and Sho Yaida. I had several useful conversations with Mark Transtrum and Chris Myers. Chris Myers and Paul McEuen both gave me very interesting questions for my A-exam which I unfortunately could not follow up on later. Thanks to Kacey Bray Acquilano for patiently answering all of the questions on the process of completing a graduate degree.

I have had many long and interesting conversations with the whole group

of people working with James Sethna. I want to thank Alex Alemi, Matt Bierbaum, Lorien Hayden, Colin Clement, Katherine Quinn, Danilo Liarte, D Zeb Rocklin, Eddie Lee and Varun Dhawan for these conversations. Many of them have been my collaborators on research projects and working with them has been a pleasure. The Boulder School for Soft Matter Physics was quite formative in my current research interests. I want to thank Hridesh Kedia and Katie Klymko who I met at the school and have since kept in touch with.

The other half of my education in the United States happened at the Church of the Advocate in North Philadelphia. I want to thank Dr. Anthony Monteiro for all that he has taught me about philosophy and history, and for introducing and explaining thinkers who now shape so much of my world view. I have looked up to Yvonne King, Yusuf Khalfani and their guidance and advice has been very important in the past few years.

I want to thank Ti-Yen Lan, Philip Burnham, Hao Shi and Andre Frankenthal for their friendship, but also for the countless lunches and coffees. I spent quite a lot of my time working with others in the graduate student union, Cornell Graduate Students United. I want to thank all those in the Rank and File Democracy Caucus but particularly Sena Aydin, Jeffrey Bergfalk, Geethika Dharmasinghe and Andi Kao for the many political discussions that we have had. The readings and events with Asian Pacific Americans for Action have taught me a lot and I want to thank all those who participated in the group, particularly Emily Dong, Eunnuri Yi and Xiao Yin Ma. I want to thank everyone in The People's School particularly Carrie Freshour, Naadhira Ali, Marc Kohlbry and Zifeng Liu. I want to thank all comrades who I have worked with in the Saturday Free School and the Desis United for Azadi in Philadelphia, particularly Meghna Chandra, Jahanzaib Choudhry, Samir Butt, Elias Gonzales, Autumn Zellers and

Brandon Do. I want to thank Sarah Gotowka, Arwa Awan, Daud El-Bakara and Douglas E Milton for many long discussions that will stay with me. Some of the above people have become close friends and made these five years very enjoyable.

Finally, I want to thank Samra Irfan for the constant support and friendship, and Purba Chatterjee for the many phone calls. Nandita Chaturvedi has been a close friend and comrade and this may not have been possible without her.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

> The material, when it is present in human society, has endless variations; the observer is himself part of the observed population, with which he interacts strongly and reciprocally. This means that the successful application of the theory needs the development of analytical power, the ability to pick out the essential factors in a given situation.
>
> D. D. Kosambi

This thesis is made out of three different, and somewhat separate projects. The first project is on interpreting the results of the renormalization group (RG) using information geometry, a field which combines statistics and geometry. Attempting to think about the information theoretic meaning of RG can perhaps clarify its meaning in contexts which are not part of physics. The second project uses normal form theory, a branch of dynamical systems theory, to classify the results of the RG into *universality families*. This leads to new results for old problems and generates a systematic machinery to do scaling collapses. As an extension of this, we examine the RG in the context of the period doubling transition. The third project looks at kicked Hamiltonian maps in the context of a particle accelerators and calculates stability boundaries and escape rates under the presence of noise.

At a first glance, these projects are quite separate [1]. Hence, I have kept them in separate sections and each section has a brief introduction of its own. However, the renormalization group is a central theme in the first two projects and is (though not in the work I report here) present even in the third one. Therefore, I want to give a general introduction to the philosophy of the renormalization group, and also locate the importance of the RG for the philosophy of science today. As a warning, Section 1.1 is only tangentially connected to the rest of the thesis but explores questions that have been a long standing interest of mine. Those who are only interested in the calculations and concrete results related to the RG presented in this thesis may prefer to skip this section, while I am sure there are some who might find it to be the only section of interest.

[1] The most superficial but somewhat amusing connection between them is the presence (in quite different ways) of Lie operators.

1.1 Philosophical considerations

There is a natural sort of universality that we are very used to in physics. At a basic level, the ability to make predictions of macroscopic phenomena without regard to microscopic details is a testament to this. The fact that we model fluids using Navier-Stokes equations which depend only on a few parameters, or the fact that we model solids using elasticity theory is well known. Indeed, statistical mechanics as a theory of the motion of gases already utilizes the ability to make predictions about *certain* phenomena when we promise to not ask certain questions. Historically, the trajectory of 19th and 20th century scientific thought was a discovery of models of physical phenomena at smaller and smaller length scales. The discovery of these models often left macroscopic models unchanged e.g. the discovery of quantum electrodynamics left the Navier Stokes equation unchanged. Nevertheless the practice of physicists at the time supported a reductionist philosophy.

In analytical western philosophy, this led to constructing the philosophy of 'physicalism' (the successor of materialism) which takes the view that everything is ultimately explained by the 'laws' of physics. More precisely, as Kim [91, 90] argues, all causal agency must rest with the 'laws' of physics. The 'laws' of physics have causal closure, any event that has a cause, has a physical cause. Positing any other cause for the event then gives a redundancy in causal explanation which must be eliminated. In this extreme notion, the laws of physics become a substitute for a Christian God that Newton had assumed with his choice of the word 'law' [2].

[2]Unfortunately, of course, this means that our discovery of the laws of physics is also causally determined by the laws of physics themselves. This not only raises basic existential

Much of the philosophy of science (eg. Popper, Kuhn) in the past two centuries has been focused on an understanding of 'fundamental physics' and often implicitly accepted the reductionist hypothesis. Attempts to deal with its logical incoherence have been made by proponents of the somewhat ill-defined term 'emergence'. Here, I propose that a serious engagement with the philosophy of statistical physics (understood broadly) or condensed matter physics may be fruitful.

'More is different' is the catch-phrase of Anderson's oft-quoted article which has become a center-piece for debates about emergence in condensed matter physics [7]. What is not quoted often though, is how the article starts

> The reductionist hypothesis may still be a topic for controversy among philosophers, but among the great majority of active scientists I think it is accepted without question. The workings of our minds and bodies, and of all the animate or inanimate matter of which we have any detailed knowledge, are assumed to be controlled by the same set of fundamental laws, which except under certain extreme conditions we feel we know pretty well.

On the other side of the debate is Weinberg, who says [153]

> In the same way, even though new concepts "emerge" when we deal with fluid or many-body systems, we understand perfectly well that

issues, it also leads to a logical paradox. If our discovery of the laws of physics is determined by those very laws, then there is a chance we will never discover all of them (since it is not in our hands). On what basis then are we supposed to trust that everything is determined by laws of physics which we may not even discover?

> hydrodynamics and thermodynamics are what they are 'because' of the principles of microscopic physics. No one thinks that the phenomena of phase transitions and chaos (to take two examples quoted by Krumhansl) could have been understood on the basis of atomic physics without creative new scientific ideas, but does anyone doubt that real materials exhibit these phenomena because of the properties of the particles of which the materials are composed?

So what is the disagreement between the two sides? Anderson goes on to say that while the reductionist hypothesis may be accepted, it does not imply a constructionist hypothesis. One cannot always take the 'laws' of physics and construct explanations for behaviour of aggregates. But, not only that, the behaviour of aggregates is in fact insensitive to these 'laws'. What does it mean to say that hydrodynamic phenomena is exhibited 'because' of the properties of the particles on which it is composed when systems with entirely different microscopic constituents exhibit the same hydrodynamic laws? As we see in the quote above, Anderson implicitly accepts this too. While constructing theories may require new creative input, there are no genuinely new causal properties at the level of hydrodynamics.

One resolution of this paradox is to argue that it doesn't really make sense to talk of causal powers of base entities when they are in their aggregate form (a free hydrogen atom is not the same as the hydrogen atom as part of a water molecule). The way the microscopic entities exert their causal powers is to interact and preserve the existence of the macroscopic causal form. The discovery of these logics of structure is an important problem in both natural and social science [94].

However, there is still a fundamental problem which was recognized long ago. The 'laws' of physics determine future states given the state at some point of time. All of the laws are either deterministic or random. This is in stark contrast to everyday experience where we contextually interact in creative ways with our environment. Here, creativity can be a very mundane sort of creativity, the kind exhibited by our language use [51, 50].

The problem of causality continues to cause debates today, and I believe part of an honest resolution of this was proposed by W.E.B Du Bois in his essay, 'Sociology Hesitant' where he said [65],

> Of the physical scientists on the one hand who say: The laws of men's deeds are physical laws, and physics studies them; of the mass of men, on the other hand, who say: Man is not wholly a creature of unchanging law, he is in some degree a free agent and so outside the realm of scientific law. Now whatever one's whims and predilections, no one can wholly ignore either of these criticisms: If this is a world of absolute unchanging physical laws, then the laws of physics and chemistry are the laws of all action of stones and stars, and Newtons and Nortons. On the other hand, for a thousand and a thousand years and today as strongly as, and even more strongly than, ever, men, after experiencing the facts of life, have almost universally assumed that in among physical forces stalk self-directing Wills, which modify, restrain, and re-direct the ordinary laws of nature. The assumption is tremendous in its import. It means, from the point of view of Science, this is a world of Chance as well as Law;

and

> Why not then flatly face the Paradox? [Why not] frankly state the Hypothesis of Law and the Assumption of Chance, and seek to determine by study and measurement the limits of each?

While the word 'Chance' could be replaced by a different one since Du Bois here is naturally not speaking of chance purely in the sense of randomness, I think that this stated assumption is a natural and necessary resolution of the paradox. The assumption of human causal agency is a natural one. A unification with microscopic laws may be possible in the future, and it may involve a complete restructuring of those laws themselves [130, 129]. However, our success in describing some microscopic theories does not in any way prove or imply that everyday observation of causal agency of humans is an illusion. The idea that microscopic theories determine the macroscopic behaviour is merely an additional metaphysical assumption which no conceivable scientific program, as of now, can confirm or deny. Hence, this metaphysical assumption should be discarded. In criticizing Al-Ghazali [5], Weinberg [154] seems to have completely misunderstood his philosophy which leaves space for such human agency.

If anything, the lesson of statistical physics is that scientific work at different scales can be done independently; unification is a possibility but not a necessity. Coordinates that describe collective behaviour have properties that the bare parts do not have, and their study might require a very different idea of science from the one that governs the bare parts themselves. This is not merely a philosophical claim but is bound to become a practical problem in the coming decades as attention shifts away from fundamental physics to biology. As the main character in the short story 'Princess Steel' says [64], "We can see the Far

Great and the Near Small but not the Great Near." From the natural sciences, physicists Nigel Goldenfeld and Leo Kadanoff say [77], "Maybe physics studies will become more like human experience". As extravagant claims about the possibilities that the availability of large amounts of data in biology has created fall short (eg. the Human genome project [101]), science of the Great Near will require a careful engagement with the philosophy of statistical physics.

There are many places in statistical physics where this becomes relevant. The first is the relationship of microscopics to thermodynamic quantities. In fact, there is no well-defined way to go from microscopic laws to statistical physics. It is easy to show, for example, that entropy remains a constant for Hamiltonian dynamics and does not increase. Jaynes proposed an interpretation of statistical mechanics under the 'subjectivist' interpretation of probability, which says that the probabilities assigned to a system quantify our ignorance about the system. Hence the observer, even in statistical physics, becomes bound up with the observed system in the sense that probabilities quantify epistemological uncertainty rather than ontological uncertainty. Attempts to construct a world which leaves no space for the human being have not been successful.

Another example is given by the ability to take continuum limits, or describe macroscopic phenomena at length scales much larger than the atomic parts that constitute it with great success in that region of validity. The diffusion equation and the Navier Stokes equations are classic examples of our ability to do this. What makes this interesting at all is that these theories describe a wide variety of microscopically different systems, the details of which are subsumed into a few parameters. The renormalization group is the most spectacular example that provides an explanation of universality near critical points. Universality at

critical points is a somewhat tricky matter because the behaviour at all scales matters at such points.

The renormalization group does not refer to a precise method or a theory. Rather, its fundamental contribution is a conceptual leap, to think about the space of *all* theories [3]. In this space, we ask what happens if we coarse grain (ignore or trace over some degrees of freedom) and rescale (scale the system back to its original size). The system is now described by a new theory with renormalized parameters. The RG is a tool which allows us to calculate how the parameters renormalize as we coarse grain. Most parameters turn out to be irrelevant and do not matter (at a first approximation) to describe the behaviour at the critical point. The RG allows us to classify these critical points into universality classes which all exhibit the same behaviour. Systems with vastly different microscopic compositions can look very similar at a phase transition eg. the liquid gas transition is in the same universality class as the Ising model of magnets.

Batterman has argued that the singularities in the RG are absolutely crucial to understanding the insensitivity of theories to microscopics [17] (see also [36, 16]). This is a strange argument, because in practice there are no physical systems which exhibit these singularities. The divergence of the correlation length in every physical system is cut off by the system size. To the contrary, we embrace the finiteness of systems in our analysis. Our goal in the first part of this thesis is modest, it is to rephrase the results of the Renormalization Group, a key tool in statistical physics, in the language of information theory. The proposed interpretation is that coarse graining preserves information about certain

[3] The word all here is not used in any precise sense, but is meant to convey the abstraction. In practice, this space is parametrized by some set of parameters.

relevant variables in a theory. Hence the variables that quantify the universal aspects of behaviour are the ones that we can still measure when we throw away information about microscopics in a physical system. Let us now give a brief introduction to the basic formalism of the RG.

1.2 Renormalization Group

The basic formalism of the RG in statistical mechanics is well known [76, 39]. We present a very short version merely to introduce some notation that we will use throughout the text. We start with a Hamiltonian $\mathcal{H}$ which is a function of some parameters θ^μ and some degrees of freedom ϕ. All physical observables can be written in terms of the partition function

$$Z = \mathrm{Tr}_\phi e^{-\beta\mathcal{H}(\phi,\theta^\mu)}, \tag{1.1}$$

where $\beta = 1/T$ is the inverse temperature [4]. ϕ consists of all the microscopic degrees of freedom, it can be a continuous or discrete function. The partition function is a function of the parameters θ^μ. The free energy is defined as

$$F = -T \log Z \tag{1.2}$$

and is similarly a function of these parameters. We now ask what happens if we partially perform the trace in the partition function. Usually, we want to trace over short distance degrees of freedom (or alternatively, high momentum ones). Though it is not guaranteed in general that this is possible, many systems can be described by a Hamiltonian of the same form. Then, we can write the partition function as a function of the new degrees of freedom $\tilde{\phi}$

$$Z = \mathrm{Tr}_{\tilde{\phi}} e^{-\beta\mathcal{H}(\tilde{\phi},\theta^{\tilde{\mu}})}. \tag{1.3}$$

[4] We are working in units where the Boltzmann constant $k_B = 1$

The system can be described with the same Hamiltonian but with *renormalized* parameters $\theta^{\tilde{\mu}}$. Notice that we are putting the $\sim$ on the index rather than on θ [5]. Calculating how the parameters are renormalized requires an actual model and is usually hard work. However, in general it is some transformation $\mathcal{R}$ so

$$\theta^{\tilde{\mu}} = \mathcal{R}^{\tilde{\mu}}(\theta^{\tilde{\nu}}). \tag{1.4}$$

The renormalized parameters $\theta^{\tilde{\mu}}$ describe the system at a longer length scale. It is typical to consider a continuous version of the transformation as a function of the log of the length scale ℓ. This gives a set of differential equations

$$\frac{d\theta^{\mu}}{d\ell} = \beta^{\mu}. \tag{1.5}$$

The right hand side of these flow equations are called beta functions. A fixed point of the transformation is given by the point where $\beta^{\mu}(\theta^{\nu}) = 0$. The RG is often discussed near the fixed point where it can be linearized. It is also typical to work in variables where this linearization is also in an eigenbasis. Then the flow of the parameters is given simply by

$$\theta^{\mu}(\ell) = \theta^{\mu}_{0}(\ell)e^{\lambda_{\mu}\ell}, \tag{1.6}$$

where λ_{μ} is the eigenvalue associated with θ^{μ}. If this eigenvalue is positive, the parameter is called relevant, if it is negative, it is called irrelevant and if it is zero the parameter is called marginal. Most physical theories are described by a few number of relevant parameters. The irrelevant parameters flow to zero at longer and longer length scales.

Physical predictions near the critical point depend only on the relevant parameters. This is the basis of *universality*, the fact that many different micro-

[5] This may be somewhat mysterious to those who have not seen it before but is standard notation in general relativity, see eg. Ref. [40]

scopic theories (which differ in their irrelevant parameters) have the same behaviour near a critical point. It is this insight which allows us to reformulate the RG as a statement about the distinguishability of parameters which we do in the next chapter.

CHAPTER 2

RENORMALIZATION GROUP AND INFORMATION GEOMETRY

> Our comforting conviction that the world makes sense rests on a secure foundation: our almost unlimited ability to ignore our ignorance.
>
> Daniel Kahneman

2.1 Introduction to information geometry

Information geometry is a branch of statistics which looks at the geometry of statistical distributions. The primary quantity in information geometry is the Fisher Information Metric which measures the distance between model distributions. The metric is defined on the *model manifold*, the set of all predictions of the model. The FIM is a Riemannian metric over parameter space. The local distance is an intrinsic measure of the distinguishability of nearby models (see [114, 138, 125] for discussion of this and related thermodynamic metrics). The FIM is defined for a parameterized statistical model $P(\{x\}|\vec{\theta})$ giving the probability P for a state $\{x\}$ and parameters θ^μ as

$$g_{\mu\nu} = -\sum_x P(x|\vec{\theta})\,\frac{\partial^2 \log(P(x|\vec{\theta}))}{\partial\theta^\mu \partial\theta^\nu}, \tag{2.1}$$

$$= -\sum_x P(x|\vec{\theta})\,\frac{\partial \log P(x|\vec{\theta})}{\partial\theta^\mu}\frac{\partial \log P(x|\vec{\theta})}{\partial\theta^\nu}, \tag{2.2}$$

where $ds^2 = g_{\mu\nu}\delta\theta^\mu\delta\theta^\nu$ defines the distinguishability of models that differ by $\delta\vec{\theta}$ from data $\{x\}$. Intuition for why this is so comes from the Cramér-Rao bound. If we have model with just one parameter θ and we have an unbiased estimator $\hat{\theta}$ based on some data, then the variance of the estimator is bounded by the inverse of the FIM.

$$\mathrm{Var}(\hat{\theta}) \geq \frac{1}{g_{\theta\theta}}. \tag{2.3}$$

The FIM thus gives an estimate of how well parameters can be estimated from the given data, making it the natural measure of distinguishability. The notion of distance comes from the Kullback-Leibler divergence which is defined as

$$D_{\mathrm{KL}}(P(x,\theta_1)||P(x,\theta_2)) = -\sum_x P(x,\theta_1)\log\frac{P(x,\theta_2)}{P(x,\theta_1)}, \tag{2.4}$$

for two different values θ_1 and θ_2 of some parameter. The K-L divergence quantifies the extra information needed to encode a distribution, if we have incorrectly

estimated the parameter to be θ_2 instead of θ_1. The KL divergence is not a metric, it is not symmetric. The FIM is the infinitesimal form of the K-L divergence and is a Riemannian metric on the model manifold. In summary, distances on the model manifold as measured by the FIM are induced by the probability distribution over the data, and give a measure of how well parameters can be estimated from the given data.

Our work in this chapter builds on previous studies showing that in models from systems biology and elsewhere this metric has a characteristic *sloppy* distribution, with eigenvalues spanning many orders of magnitude [149, 69]. Many nonlinear models have a huge number of poorly determined parameters. As an example, a typical systems biology model is a bunch of Michaelis-Mentin reactions which regulate the functions of proteins and enzymes, with rate constants that are difficult to measure. However, when these models are fit to experimental data (like the concentration of a particular protein which regulates the division of cells), predictions are not sensitive to all of these parameters. Rather, they are sensitive only to particular combinations of these parameters [79, 33]. Locally, these are given by the eigendirections of the FIM which tend to be linear combination of bare parameters. We typically don't parametrize models in a way that is most efficient. However, predictions in many nonlinear models depend only on a few parameters. This by itself is not very satisfying since any local measure of distinguishability depends on the particular parametrization chosen. In particular, we can always choose a parametrization where the metric is locally Euclidean.

However, sloppiness is a property of the model rather than of the parameterization. This is made evident by looking at the geometry of model manifolds.

Many nonlinear models typically have a hierarchy of widths, in a hyper-ribbon structure with one long direction, and other directions becoming thinner and thinner as shown in Figure 2.1. Note that the models are often finite in the space of predictions, even though the parameters are unrestricted. In fact, the boundaries of the model manifolds typically correspond to parameters going to zero, ∞ or some other extreme value.

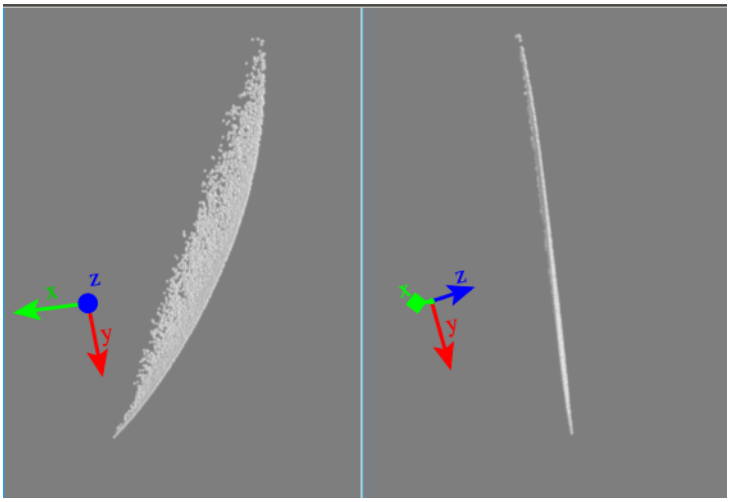

Figure 2.1: A projection of a typical model manifold in 3 dimensions. There is one very long direction, a shorter direction and a very thin direction. This hierarchy of widths continues in higher dimensions and is called a hyper-ribbon. The figure is reproduced from Ref. [149].

In physics, we are used to a very different explanation of why simple models are possible. As was covered in the introduction, the renormalization group gives a beautiful explanation of universality and effective theories. Microscopically diverse systems often yield surprisingly simple effective theories. The renormalization group (RG) describes how system parameters change as the scale of observation grows and gives a precise explanation for this emergent simplicity. Most parameter combinations are *irrelevant* and their importance decreases as the RG proceeds. Effective theories are thus determined by a small number of *relevant* parameters whose importance grows with increasing scale.

While the RG was initially used to understand systems with spatial symmetry, it has found applicability in a wide range of theories including avalanches [54], turbulence [38], differential equations [46], the central limit theorem [85, 86], period doubling [70] and quasi-periodic systems [131, 72]. Here we use information geometry [6] to reformulate the RG as a statement about how the distinguishability of microscopic parameters depends on the scale of observation. It was recently shown that renormalizable models become sloppy, but only after their data, $\{x\}$ is coarsened [68], by decimation in the Ising model and blurring in time for diffusion. While that paper was mostly numerical, they found that (1) 'relevant' directions of the FIM do not grow, but instead are almost unaffected by coarsening and (2) irrelevant directions contract at a rate given by their RG exponent.

2.2 Renormalization Group and Information Geometry [1]

In this chapter, we address this question analytically and develop a covariant formalism to describe the flow of the model manifold. Past studies have considered the connection of the RG to geometry [100, 118, 60], but not to information theory. Here we use the RG to calculate the flow of the model manifold as the observation scale changes, and connect it to the distinguishability of parameters. We show that as coarse-graining proceeds relevant directions are exactly maintained while irrelevant directions contract. Our results quantify the irreversibility of RG transformations: models which differ only in irrelevant directions rapidly become indistinguishable as the observation scale increases. Our

[1] Most of this has been published on arXiv:1710.05787 with James P. Sethna and Benjamin B. Machta

results also clarify that information about relevant directions are contained in large scale observables, since these directions do not contract as the observation scale increases. This striking feature that relevant directions are preserved is distinct from the usual way the RG is thought about, where relevant directions increase. We first describe our formalism, apply it to the diffusion equation and then to coarse graining classical statistical systems [2] described by a free energy.

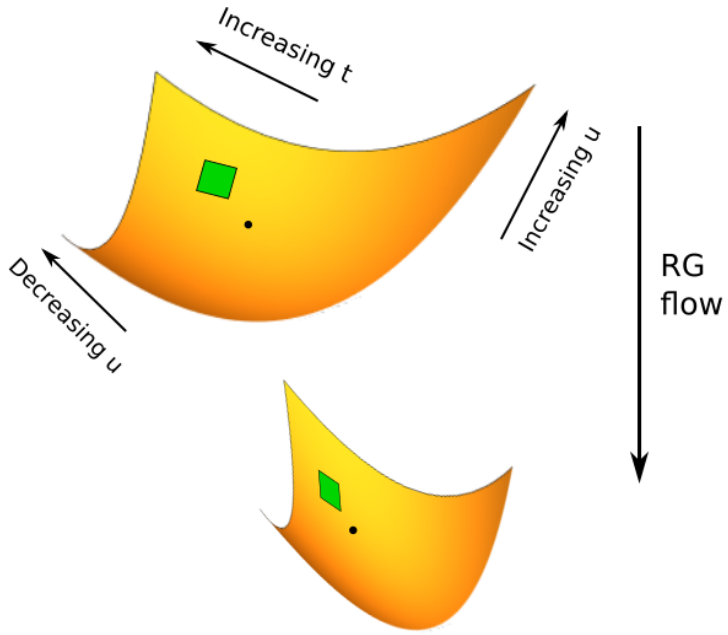

Figure 2.2: A cartoon figure of a section of the model manifold with one relevant (t) and one irrelevant (u) direction. The RG makes an initial patch (green) stay the same in the relevant direction but compress in the irrelevant directions. Information is preserved in relevant directions and lost in irrelevant directions. Finding an embedding to visualize the manifold is non-trivial [126].

To understand how coarse-graining loses information about parameters, we consider an infinitesimal RG transformation of the form $d\theta^\mu/db = \beta^\mu$ describing the flow of parameters by beta functions $\vec{\boldsymbol{\beta}} = \{\beta^\mu\}$ as the shortest length scale b is increased. Because information about microscopic degrees of freedom is

[2] We are not considering quantum systems where probabilistic interpretations are less clear [13, 23]

discarded as the flow proceeds, there is no guarantee that a forward RG flow can be uniquely reversed. However, the irreversibility of the RG is difficult to quantify through the RG equations alone. Instead, we turn to the metric tensor of eq. 2.1, a local measure of how hard it is to discriminate between models which differ by small changes of parameters. We can quantify the loss of information in parameter space as microscopic information is discarded by asking how the metric tensor changes under an infinitesimal RG transformation. Since the metric is a two-tensor and the transformation is a flow given by a vector field, the answer is given by the Lie derivative $\mathcal{L}_{\vec{\beta}}$ which defines the derivative of a covariant tensor carried along flow field $\vec{\beta}$.

Because the FIM quantifies the observability of parameters from a fixed amount of data, we consider an RG procedure which takes place on a large but finite system, whose physical size is fixed so that the observed system length L shrinks during the RG according to $dL/db = -L$. As L is not a parameter, we must add an additional term to the usual Lie derivative defining $\bar{\mathcal{L}}_{\vec{\beta}} = \mathcal{L}_{\vec{\beta}} - L\partial_L$. This additional term can be derived as follows. The definition of Fisher Information in some finite region of length L can be written as an integral

$$g_{\mu\nu} = -\int_0^L P(x)\partial_\mu\partial_\nu P(x)dx \tag{2.5}$$

If we observe the data only at a scale $y = ax$ with $a > 1$, this gives a new sum $-\int_0^L P(y)\partial_\mu\partial_\nu P(y)dy$. Now rescaling with $x = y/a$ and letting $a = 1 + \epsilon$ gives the change in $g_{\mu\nu}$ as

$$\Delta g_{\mu\nu} = \int_L^{L(1-\epsilon)} P(x)\partial_\mu\partial_\nu P(x)dx, \tag{2.6}$$

$$= -L\partial_L g_{\mu\nu}\epsilon \tag{2.7}$$

All of these contributions give in total,

$$\bar{\mathcal{L}}_{\vec{\beta}} g_{\mu\nu} = -L\partial_L g_{\mu\nu} + \beta^\alpha \partial_\alpha g_{\mu\nu} + g_{\alpha\mu}\partial_\nu \beta^\alpha + g_{\alpha\nu}\partial_\mu \beta^\alpha. \tag{2.8}$$

This modified Lie derivative $\bar{\mathcal{L}}_{\vec{\beta}}$ gives the change in the metric as data is coarsened and it can be written as the sum of four terms. The first term derived above describes the change in the metric due to the shrinking of the effective system size. As long as the system is large enough to ignore the effects of the boundaries, the metric tensor $g_{\mu\nu} \propto L^d$, so that the first term can be written $-dg_{\mu\nu}$. The second term is the directional derivative of the metric along the $\vec{\beta}$ function, which keeps track of how the metric changes as the coordinates change. The final two terms come from the change in the parameter space distance between nearby points under flow. Together, these terms describe how the metric changes under coarse graining. Since we are working in fixed coordinates, the change in the metric reflects a change in the proper distance between nearby systems and is an invariant quantity.

To interpret this equation further, we consider the RG in its linearized form (we will discuss this assumption in detail in Chapter 3) with eigenvalues $\lambda^{(\mu)}$, so that $\beta^\mu = \lambda^{(\mu)}\theta^\mu$ and we use the notation that the summation convention is not to be used if the indices are in parentheses. With $\partial_\mu \theta^\alpha = \delta^\alpha_\mu$, this yields a simplified equation

$$\bar{\mathcal{L}}_{\vec{\beta}} g_{\mu\nu} = \lambda^{(\alpha)}\theta^\alpha \partial_\alpha g_{\mu\nu} + g_{\mu\nu}(\lambda^{(\mu)} + \lambda^{(\nu)}) - L\partial_L g_{\mu\nu}. \tag{2.9}$$

The RG is often done as a discrete operation, rather than a continuous one in which θ^μ flow to $\theta^{\tilde{\mu}}$ in one discrete iteration. In this case the new metric in the new coordinates is given by $\tilde{g}_{\tilde{\mu}\tilde{\nu}} = \langle \partial_{\tilde{\nu}} \partial_{\tilde{\mu}} \log P(\tilde{x}) \rangle$. Here and from now on we will use the $\sim$ to denote renormalized coordinates and coarse-grained data. To write the metric in bare coordinates we simply change parameters according to $\tilde{g}_{\mu\nu} = \tilde{g}_{\tilde{\mu}\tilde{\nu}} \frac{\partial \theta^{\tilde{\mu}}}{\partial \theta^\mu} \frac{\partial \theta^{\tilde{\nu}}}{\partial \theta^\nu}$. Hence, the new metric is given by $\tilde{g}_{\nu\mu} = \langle \partial_{\tilde{\nu}} \partial_{\tilde{\mu}} \log P(\tilde{x}) \rangle \frac{\partial \theta^{\tilde{\mu}}}{\partial \theta^\mu} \frac{\partial \theta^{\tilde{\nu}}}{\partial \theta^\nu}$.

We illustrate this formalism by first considering the RG flow of a generalized hopping model of diffusion. Consider a particle undergoing stochastic motion according to a kernel $K_\tau(x)$ so that

$$P(x(t+\tau)|x(t)) = K_\tau\left(x(t+\tau) - x(t)\right).$$

We assume that K_τ has a finite second moment. How well can we infer the shape of K_τ by measuring time series $\vec{x} = \{x(0), x(\tau), x(2\tau), ..., x(m\tau)\}$? We can choose to parameterize the Kernel in different ways. A convenient way to parametrize probability distributions close to a Gaussian distribution is to use Hermite polynomials. Let θ^μ parameterize an arbitrary Kernel K_τ, and H_μ be the μ^{th} Hermite polynomial. For notational convenience, we define $\bar{H}_\mu = \frac{H_\mu(y)}{\mu!(\theta^2)^{\mu/2}}$. The Kernel is then given by

$$K_\tau(x) = \frac{1}{\sqrt{2\pi(\theta^2)}} \exp\left(-y^2/2\right)\left[1 + \sum_{\mu\geq 3} \theta^\mu \bar{H}_\mu(y)\right], \tag{2.10}$$

where $y = (x - \theta^1)/(\theta^2)^{1/2}$.

The probability of a time series $\vec{x}$ is given by $P(\vec{x}) = \prod_{i=1}^{m} K_\tau\left(x(i\tau) - x((i-1)\tau)\right.$ and the FIM near $\theta = 0$ is

$$g^{(0)}_{\mu\nu} = \frac{m\delta_{\mu\nu}}{\mu!(\theta^2)^\mu}, \tag{2.11}$$

where the (0) denotes a microscopic measurement. In typical measurements, the scale of observation is much larger than the natural timescale τ. How would the slowness of a measuring apparatus influence the ability to infer K_τ? One means to address this is to consider the renormalized kernel arising from the composition of n time-steps. For example, after two steps

$$K_{2\tau}(x) = \int dx' K_\tau(x') K_\tau(x - x'). \tag{2.12}$$

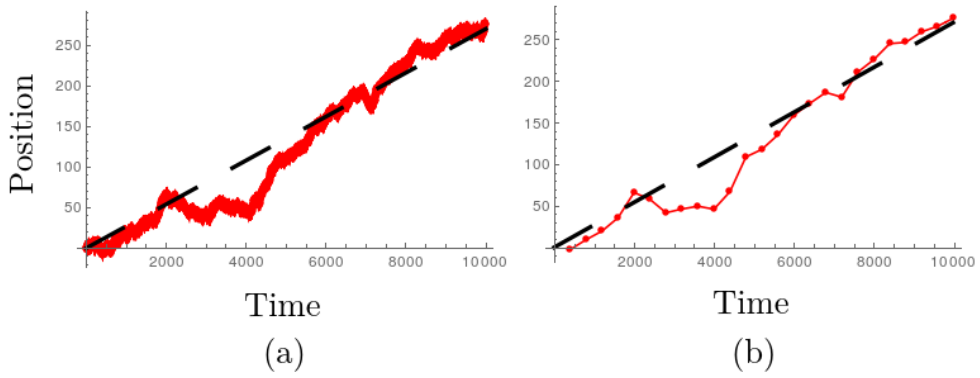

Figure 2.3: Here we show a drastic coarse graining for a random walk where most of the points in between are removed. The mean of the random walk is determined simply by the beginning and end of the walk and can be determined equally well after the coarse graining. However, less information about the fluctuations is present in the coarse grained figure (b) as opposed to the original figure (a) which contains the full microscopic data.

After n steps, the kernel can be recast into the form of Equation 2.10 provided parameters are renormalized to $\theta^{\tilde{\mu}(n)} = n^{1-\mu/2}\theta^{\mu}$ where the spatial coordinate is rescaled according to $y^{(n)} = n^{-1/2}y$. The continuous limit here gives $d\theta^{\mu}/d\log n = (1-\mu/2)\theta^{\mu}$ so that $\beta^{\mu} = (1-\mu/2)\theta^{\mu}$. The Kernel after n steps becomes

$$K_{n\tau}(x) = \frac{1}{\sqrt{2\pi(\theta^2)}} e^{-\frac{(y^{(n)})^2}{2}} \left[1 + \sum_{\mu\geq 3} n^{1-\frac{\mu}{2}}\theta^{\mu}\bar{H}_{\mu}(y^{(n)})\right]. \tag{2.13}$$

We are interested in the case where observations occur at a scale much larger than τ. In addition, for a fixed trajectory, the number of data points reduce by a factor of n, rescaling the metric in the discrete contribution corresponding to the third term in eq. 2.9. In the case of diffusion, the metric depends only on θ^2 which is itself invariant under the RG operation. Hence, the first term in eq. 2.9 is zero so that the parameter dependence of the metric doesn't contribute to its change on rescaling. This is in part because of our choice of coordinates. In different coordinates, this term could contribute to the change in the metric since the individual terms in eq. 2.9 are not covariant. However, the Lie derivative as

a whole is covariant so that the total change in the metric transforms covariantly under a change of coordinates. Finally, the contribution of the second term in eq. 2.9 is given in a straightforward way with the eigenvalues $\lambda^{(\mu)} = (1 - \mu/2)$ and the new metric near the Gaussian fixed point is given by:

$$\tilde{g}^{(n)}_{\mu\nu} = n^{1-\mu} g^{(0)}_{\mu\nu} \tag{2.14}$$

For diffusion, the mean $\theta^{(1)}$ is the only relevant variable in the sense of having RG exponent > 0, and it is exactly preserved under this coarse-graining procedure. The standard deviation is the next most distinguishable parameter, marginal in the RG sense, and under coarse-graining it becomes harder to see. All other parameters become even more indistinguishable at late times. This framework quantifies the irreversibility of the RG through the inability to distinguish irrelevant variables at long time or length scales. Figure 2.3 shows the effect of a drastic coarse graining on an instance of a random walk. The mean can be determined equally well even after most of the data is thrown away.

Our approach emphasizes that the total information depends on the total amount of data. An alternative approach, more in line with statistical physics, would be to define a Fisher information density which is independent of the system size. Relevant directions would then spread out under coarse graining with an exponent equal to the dimension of the system d, but *not* to their RG exponent. This *intensive* metric is discussed in Ref. [126]. Using the total information is not only more natural from an information-theoretic point of view but results (here and in the next section) correspond more to intuition: the information about relevant directions is preserved under coarse graining.

The diffusion equation is a simple example where everything can be analytically calculated. We now apply our formalism in more generality to statistical

systems with a free energy. Consider a Boltzmann distribution for a micro state x, $P(x) = \exp(-H(x))/Z$, with Hamiltonian of the form $H = \sum_\mu \theta^\mu \Phi_\mu(x)$, where $\Phi_\mu(x)$ are functions of the micro state x and where θ^μ are parameters. Typically, Φs will be sums or integrals over space like $\int_{\mathcal{V}} \phi^2$ in Landau theory or $\sum_i x_i x_{i+1}$ in the Ising model. Because the Hamiltonian is in an exponential family and linear in its parameters the FIM can be written as (see eq. 2.1) [68]:

$$g_{\mu\nu} = \partial_\mu \partial_\nu \log(Z) = \langle \Phi_\mu \Phi_\nu \rangle - \langle \Phi_\mu \rangle \langle \Phi_\nu \rangle \tag{2.15}$$

A renormalization group operation typically involves coarse graining by tracing over some degrees of freedom (like in our previous example for diffusion). Let $\tilde{x}^{[n]} = C^{[n]}(x)$ be the coarse-grained state observed when the bare micro state is x. This coarse-graining could involve removing high momentum states, or decimating over alternate spins (see [68] for an example). For our purposes, it is only important that the coarse graining acts as a map from bare micro-states to coarse grained ones. We can define the restricted partition function $\tilde{Z}(\tilde{x}^{[n]}) = \sum_x \delta(C^{[n]}(x), \tilde{x}^{[n]}) e^{-H(x)}$ in terms of which the probability of being in coarse grained state $\tilde{x}$ is $p(\tilde{x}) = \tilde{Z}(\tilde{x})/Z$. We choose our convention so that the renormalized Hamiltonian has the same form as the old one with renormalized parameters with any additional constant written separately

$$\tilde{H}(\theta^{\tilde{\mu}}) - G(\theta^\mu) = H(\theta^\mu). \tag{2.16}$$

We can define the Fisher metric on the bare parameter space, but for coarse-grained observables $\tilde{x}$ using eq. 2.1:

$$\tilde{g}^b_{\mu\nu} = \partial_\mu \partial_\nu \log Z - \langle \partial_\mu \partial_\nu \log \tilde{Z}^b(x) \rangle \tag{2.17}$$

To calculate derivatives of $\log \tilde{Z}^b$ it is helpful to define the expectation value of an operator $\Phi(x)$ defined at the bare micro state, conditioned on coarse-graining

to coarsened state $\tilde{x}^{[n]}$ as [68]

$$\{\Phi(x)\}_{\tilde{x}^{[n]}} = \frac{1}{\tilde{Z}(\tilde{x}^{[n]})} \sum_x \delta(C^{[n]}(x), \tilde{x}^{[n]})\Phi(x)e^{-H(x)} \tag{2.18}$$

We define correlation functions of these operators in the natural way, $\langle\{\Phi_1\}_{\tilde{x}^{[n]}}\{\Phi_2\}_{\tilde{x}^{[n]}}\rangle = \sum_{\tilde{x}^{[n]}} p(\tilde{x}^{[n]})\{\Phi_1\}_{\tilde{x}^{[n]}}\{\Phi_2\}_{\tilde{x}^{[n]}}$. In terms of this

$$\langle\partial_\mu\partial_\nu \log\tilde{Z}\rangle = \langle\left\{\Phi_\mu\Phi_\nu\right\}_{\tilde{x}^{[n]}} - \left\{\Phi_\mu\right\}_{\tilde{x}^{[n]}}\{\Phi_\nu\}_{\tilde{x}^{[n]}}\rangle \tag{2.19}$$

which is positive definite, demonstrating that coarsening reduces the FIM in all directions. The full FIM is given by:

$$\tilde{g}^n_{\mu\nu} = \langle\left\{\Phi_\mu\right\}_{\tilde{x}^{[n]}}\{\Phi_\nu\}_{\tilde{x}^{[n]}}\rangle - \langle\Phi_\mu\rangle\langle\Phi_\nu\rangle \tag{2.20}$$

To see the origin of the loss of information, we can directly calculate $\tilde{g}_{\mu\nu} = \langle-\partial_\mu\partial_\nu\tilde{H}\rangle + \partial_\mu\partial_\nu Z$. Now, from Equations 2.16 and 2.15, we get $\tilde{g}_{\mu\nu} = g_{\mu\nu} - \partial_\mu\partial_\nu G$. The change in the Fisher Information Metric is a consequence of the analytic constant that gets added during the coarse graining transformation. By exponentiating Equation 2.16 and taking a trace, we can write this as $\exp(-F(\theta^\mu)) = \exp(G(\theta^\mu))\exp(-F(\theta^{\tilde{\mu}}))$, where F is the free energy. As is common, we find it useful to divide the total free energy into a singular and an analytic piece, $F = F^{(s)} + F^{(a)}$. Using our definition of the metric in Equation 2.17, we can correspondingly write the metric $g_{\mu\nu} = g^{(a)}_{\mu\nu} + g^{(s)}_{\mu\nu}$. The singular part of the free energy is conserved along the flow. This can be written as

$$\beta^\gamma\partial_\gamma F^{(s)} = L\frac{\partial}{\partial L}F^{(s)} \tag{2.21}$$

Taking two derivatives of this equation and assuming linearity of the beta functions,

$$\bar{\mathcal{L}}_{\vec{\beta}}g^{(s)}_{\mu\nu} = 0 \tag{2.22}$$

That is, metric components of the singular part of the metric are preserved along the flow. The relevant components of the metric are dominated by the singular part of the free energy and hence information about them is preserved along the flow. This is reflected in the divergence of quantities like the specific heat and susceptibility at the critical point. The irrelevant components of the metric are dominated by the analytic part of the free energy. To see this, consider a typical scaling form of the free energy of an Ising model with one irrelevant component u, $F = L^d t^{\frac{d}{\lambda_t}} \mathcal{F}(ut^{\frac{-\lambda_u}{\lambda_t}})$. If we assume $\mathcal{F}$ is analytic, we can take two derivatives to find the component of the metric $g^{(s)}_{uu} = L^d t^{\frac{d-2\lambda_u}{\lambda_t}} \mathcal{F}''(ut^{\frac{-\lambda_u}{\lambda_t}})$. Since λ_u is negative, this expression goes to 0 at the critical point and is very small near to it. Away from the critical point, this has some finite value but is still small yielding well known corrections to scaling.

We can calculate how the analytic part of the metric changes by going back to the transformation for the free energy $F^{(a)}(\theta^{\tilde{\mu}}) = F^{(a)}(\theta^{\mu}) + G(\theta^{\mu})$. The change in the analytic part of the metric, after n coarse grainings, is given by $\Delta g^{(a)}_{\mu\nu}(\theta^{\tilde{\mu}(n)}) = -\sum_{i=1}^{n-1} \partial_\mu \partial_\nu G(b^{m\lambda_\alpha}\theta^{\alpha})$. This can be simplified to

$$\Delta g^{(a)}_{\mu\nu}(\theta^{\tilde{\alpha}(n)}) = -\sum_{m=1}^{n-1} b^{m(\lambda_\mu+\lambda_\nu)} \partial_{\tilde{\mu}} \partial_{\tilde{\nu}} G(\theta^{\tilde{\alpha}}) \tag{2.23}$$

where $\partial_{\tilde{\mu}} \partial_{\tilde{\nu}} G(\theta^{\tilde{\alpha}})$ is a positive definite quantity (as can be explicitly checked when calculation is possible, like in the 2-D Ising model on a triangular lattice [117]). Hence, the analytic part of the free energy decreases with each coarse graining. In this case, since the metric is analytic, the change in the metric because of the curvature (corresponding to the partial derivative in Equation 2.9) is less significant and the irrelevant components of the metric contract under the flow.

One can write down a partial differential equation describing the flow of the

singular part of the metric by the requirement that $\bar{\mathcal{L}} g^{(s)}_{\mu\nu} = 0$.

$$\beta^\alpha \partial_\alpha g^{(s)}_{\mu\nu} + g^{(s)}_{\alpha\mu} \partial_\nu \beta^\alpha + g^{(s)}_{\alpha\nu} \partial_\mu \beta^\alpha = d g^{(s)}_{\nu\mu} \tag{2.24}$$

This equation determines the components of the metric in the relevant direction which can be obtained by solving this equation. Such equations have recently been solved for different kinds of RG flows [127] near critical points. The change in the analytic part of the metric close to the critical point should be dominated by the change in parameter space distance in Equation 2.9 (with higher order corrections). Together, this allows us to predict the scaling of the metric for linear beta functions with correlation length ξ and coarse graining length b

$$g^{(s)}_{\mu\mu} \sim L^d \xi^{-(d-2\lambda_\mu)}, \tag{2.25}$$

$$g^{(a)}_{\mu\mu} \sim L^d b^{-(d-2\lambda_\mu)} \tag{2.26}$$

We have shown (eqs. 2.22 and 2.25) that information about the relevant components, contained in the singular part of the metric, is preserved under RG flow whereas information about the irrelevant components, contained in the analytic part of the metric, reduces (eq. 2.26). This is the information theoretic meaning of universality. The RG can be understood as a flow of the model manifold under which irrelevant components of theory becomes harder and harder to distinguish and can hence be ignored in a description of a theory at long length or time scales. While the conclusion in the case of statistical mechanics requires somewhat technical arguments, we emphasize that it was numerically foreshadowed in Ref. [68] making us more confident of the results. The essential conclusion that the physical observables (like the magnetization and correlation functions) allow you to determine the temperature of the system equally well

even after coarse graining. This is not intuitively obvious, and is our primary physical result.

We note some assumptions that have gone into the analysis. We have assumed the $\vec{\beta}$ functions that determine the RG flows are linear. In typical cases, it is possible to make analytic coordinate changes that make the $\vec{\beta}$ functions linear. The Fisher information will then no longer be given by Equation 2.15 but it should be possible to incorporate these corrections. In other cases, the $\vec{\beta}$ functions are inherently non-linear [128] which leads to logarithmic and other corrections. It should be interesting to examine the significance of these inherent nonlinearities for information theory. Universal scaling behavior typically comes with analytic and singular *corrections to scaling*. It is possible that such corrections can be organized by relevance using information theory by systematically including them in the free energy.

With the FIM as metric, the scalar curvature is small in the thermodynamic limit, scaling like ξ^d/L^d with ξ the correlation length, quantifying the intrinsic non-Gaussianity of fluctuations. By using the Fisher Information density, the divergence of the Ricci curvature can be identified with the the divergence of the correlation length as occurs near critical points [137, 138, 32].

We expect that our information theoretic understanding of coarse graining can help to identify relevant parameters of models even where explicit RG procedures are not available. In Ref [109], unimportant directions were identified as thin directions of the model manifold and were systematically removed. Here we find that relevant and irrelevant directions in parameter space behave differently under coarsening in a reparameterization invariant way, which can be seen even without an explicit parameter space flow. This suggests that relevant

directions for more general coarsening procedures can be identified with FIM eigendirections which do not contract under coarsening.

CHAPTER 3

RENORMALIZATION GROUP AND NORMAL FORM THEORY

The detached facts are not enough;
that is why we must have Science
ordered, or better, organised.

H. Poincaré

3.1 Introduction [1]

At root, ignoring challenges of implementation, the renormalization group (RG) coarse-grains and rescales the system to generate of ordinary differential equations (ODEs) for model parameters as a function of the observed log length scale ℓ. A fixed point of these flows represents a system which looks the same at different length scales; systems near criticality flow near to this fixed point. In cases where the flow can be linearized around the fixed point, the RG implies that observables near criticality are given by a power law times a universal function of an invariant combinations of variables; *e.g.* the Ising model has magnetization $m \sim t^\beta \mathcal{M}(Lt^\nu)$ for a system size L.

Surprisingly often, this scenario of universal critical exponents and scaling functions is violated; free energies and correlation lengths scale with logarithms or exponentials, and the proper form of the universal scaling functions has often remained unknown. Specifically, deviations arise in the Ising model in $d = 1$ [84], 2 [139], & 4 [99], the tricritical Ising model in $d = 3$ [152], the $d = 2$ XY model [95], the surface critical behavior of polymers [57, 67], van der Waals interactions in 3-d spherical model [56], finite size scaling of the random field Ising model (RFIM) in $d = 6$ [3], thermodynamic Casimir effects in slabs with free surfaces [59, 58], the $d = 2$, 4-state Potts model [140, 144, 24], percolation and the 6-d Potts model [146], and many other systems. Each such system has hitherto been treated as a special case.

Here we use the fact that the predictions of the RG can be written down as

[1] Some of this work has been published on the arXiv:1706.00137 with James P. Sethna, Lorien Hayden, Colin Clement, Jaron Kent-Dobias, Danilo Liarte and D. Zeb Rocklin and in forthcoming work on the 2-d Ising model with James P. Sethna and Colin Clement

a set of differential equations in the abstract space of Hamiltonians. This allows us to apply a branch of dynamical systems theory, normal form theory [116, 155] to provide a unified description applicable to all of these systems. We arrange these systems into *universality families* of theories, each defined by its normal form. Each family has *universal terms* (linear and nonlinear), whose values determine a system's universality class within the family. Finally, each family's normal form predicts the natural *invariant scaling combinations* governing universal scaling functions. Just as the topological theory of defects used homotopy theory to unify our understanding of flaws in condensed matter physics [148, 136], so we use normal form theory to systematize and simplify our understanding of universality and scaling in systems near continuous transitions.

Our machinery provides a straightforward method to determine the complete form of the critical singularity in these challenging cases. Our initial results are complex and interesting; they pose challenges which we propose to address in future work. The coordinate transformation to the normal form embodies analytic corrections to scaling, which allow us to address experimental systems as they vary farther from the critical point. Finally, bifurcation theory is designed to analyze low-dimensional dynamical systems without detailed understanding of the underlying equations; our methods should improve scaling collapses in critical phenomena like 2-d jamming [78] where there is numerical evidence for logarithms but no RG framework is available.

We give an introduction to normal form theory in Section 3.1.1. We give a survey of the previous literature on nonlinear scaling in the RG in section 3.2. We show how the application of normal form theory allows us to define universality families of fixed points in Section 3.3. We present several worked out

examples starting with the 4-d Ising model in Section 3.4 and the Random Field Ising model in Section 3.5. We then work out the application of normal form theory to the Ising model in dimensions 1, 2 and 3 in Sections 3.6– 3.8.

3.1.1 Normal Form Theory

Normal form theory [155] is a technique to reduce differential equations to their simplest possible form. This is achieved by making near-identity coordinate transformations to get rid of as many terms as possible from the equation. It was developed initially by Poincaré to integrate nonlinear systems [124, 47]. The physical behavior should be invariant under analytic changes of coordinates, and the length (or time) parameter should stay the same, which the mathematical literature addresses by perturbative polynomial changes of coordinates (attempting removal of nth order nonlinearities in the flow by using nth order or lower terms in the change of variables). To any finite order this gives an analytic change of coordinates, but it is not in general guaranteed to converge to an analytic transformation; we will thus call it a polynomial change of coordinates.

We begin by illustrating how we may often justify ignoring nonlinear terms in the RG by systematically removing them by polynomial changes of coordinates. We then notice two cases where this linearization fails – at *resonances* and at continuous bifurcations where one or more eigenvalues vanish.

Typically you have a set of differential equations of the form

$$\frac{d\theta_i}{d\ell} = f_i(\theta_j, \epsilon), \tag{3.1}$$

where ϵ is some parameter and θ_i's are called state variables. In the context of

statistical mechanics and renormalization group flows, θ_i's are typically parameters or coupling constants that enter into the free energy and ϵ is the thing that enters into the ϵ expansion (the difference in dimension from the lower or upper critical dimensions). Let us first work with the case where ϵ does not enter into the equations. The first step is to find the fixed point of the equation, use translations so the fixed point of each θ_i is at 0. The next step is to linearize about the fixed point and reduce the linear part to the simplest possible form. In general, this is the Jordan canonical form but often it is just the eigenbasis. The quantities involved are vectors and matrices so we will write the vector of variables $\boldsymbol{\theta} = \{\theta_i\}$. Then, the equation is

$$\frac{d\boldsymbol{\theta}}{d\ell} = J\boldsymbol{\theta} + \boldsymbol{f}(\boldsymbol{\theta}), \tag{3.2}$$

where $\boldsymbol{f}(\boldsymbol{\theta}) \sim O(\theta^2)$ and J is the linearized matrix of the flow. Terms of order k are defined to be made up of homogeneous polynomials of order k. So for $\boldsymbol{\theta} = (\theta_1, \theta_2, \theta_3)$, $\theta_1^2\theta_2\theta_3 \sim O(\theta^4)$. We will denote terms of order k by a lower index. So

$$\boldsymbol{f}(\boldsymbol{\theta}) = \sum_{k\geq 2} \boldsymbol{f}_k(\boldsymbol{\theta}). \tag{3.3}$$

Let the lowest non-zero term be at some order $k \geq 2$ (usually 2). Then we can write

$$\frac{d\boldsymbol{\theta}}{d\ell} = J\boldsymbol{\theta} + \boldsymbol{f}_k(\boldsymbol{\theta}) + O(\theta^{k+1}). \tag{3.4}$$

The idea is to try and remove higher order terms by making coordinate changes. To remove this we try and do a coordinate change of order k,

$$\boldsymbol{\theta} = \tilde{\boldsymbol{\theta}} + \boldsymbol{h}_k(\tilde{\boldsymbol{\theta}}). \tag{3.5}$$

This construction is similar in spirit to nonlinear scaling fields [39, 151] which try to linearize the RG flow equations with a subtle difference that we will remark on later. The higher order terms which we can remove by coordinate changes

correspond to analytic corrections to scaling. Then, to find the equations in the new variables.

$$\frac{d\boldsymbol{\theta}}{d\ell} = \frac{d\tilde{\boldsymbol{\theta}}}{d\ell} + (\mathcal{D}\boldsymbol{h}_k)\frac{d\tilde{\boldsymbol{\theta}}}{d\ell}. \tag{3.6}$$

$\mathcal{D}\boldsymbol{h}_k$ is the matrix of partial derivatives $\partial_\mu (h_k)_\nu$. Now, substituting this into the equation.

$$(1 + \mathcal{D}\boldsymbol{h}_k)\frac{d\tilde{\boldsymbol{\theta}}}{d\ell} = J(\tilde{\boldsymbol{\theta}} + \boldsymbol{h}_k(y)) + \boldsymbol{f}(\tilde{\boldsymbol{\theta}} + \boldsymbol{h}(\tilde{\boldsymbol{\theta}})) + O(\tilde{\theta}^{k+1}), \tag{3.7}$$

$$\frac{d\tilde{\boldsymbol{\theta}}}{d\ell} = (1 - \mathcal{D}\boldsymbol{h}_k)(J(\tilde{\boldsymbol{\theta}} + \boldsymbol{h}_k(\tilde{\boldsymbol{\theta}})) + \boldsymbol{f}_k(\tilde{\boldsymbol{\theta}} + \boldsymbol{h}_k(\tilde{\boldsymbol{\theta}})) + O(\tilde{\theta}^{k+1}), \tag{3.8}$$

$$= J\tilde{\boldsymbol{\theta}} - (\mathcal{D}\boldsymbol{h}_k)J\tilde{\boldsymbol{\theta}} + J\boldsymbol{h}_k + \boldsymbol{f}_k(\tilde{\boldsymbol{\theta}}) + O(\tilde{\theta}^{k+1}), \tag{3.9}$$

$$= J\tilde{\boldsymbol{\theta}} - (\mathcal{D}\boldsymbol{h}_k)J\tilde{\boldsymbol{\theta}} + (\mathcal{D}J\tilde{\boldsymbol{\theta}})\boldsymbol{h}_k + \boldsymbol{f}_k(\tilde{\boldsymbol{\theta}}) + O(\tilde{\theta}^{k+1}). \tag{3.10}$$

For the last line, notice that J is the same thing as $\mathcal{D}J\tilde{\boldsymbol{\theta}}$. Indices help here

$$\mathcal{D}J\tilde{\theta} \equiv \partial_\alpha J^\mu_\nu \theta^\nu, \tag{3.11}$$

$$= J^\mu_\nu \partial_\alpha \theta^\nu, \tag{3.12}$$

$$= J^\mu_\nu \delta^\nu_\alpha, \tag{3.13}$$

$$= J^\mu_\alpha \equiv J. \tag{3.14}$$

This is the Lie bracket (a commutator for vector fields). The final equation is

$$\frac{d\tilde{\boldsymbol{\theta}}}{d\ell} = J\tilde{\boldsymbol{\theta}} + [\boldsymbol{h}_k, J\tilde{\boldsymbol{\theta}}] + \boldsymbol{f}_k + O(\tilde{\theta}^{k+1}). \tag{3.15}$$

So, if we want to remove the term $\boldsymbol{f}_k$, we need to solve the equation $[\boldsymbol{h}_k, J\tilde{\boldsymbol{\theta}}] = -\boldsymbol{f}_k$ for $\boldsymbol{h}_k$. It's important to note that whether this equation can be solved or not depends only on the linear part of the equation. That is, within the space of transformations that we are considering, the linear part of the equation completely determines how much the equation can be simplified and how many terms can be removed. If there is a zero eigenvalue, one has to consider a broader space of transformations which we will do later.

To see when the equation can be solved, we first note that the space of homogeneous polynomials is a vector space with one basis constructed in the obvious way $\theta_1^{\alpha_1}...\theta_n^{\alpha_n}$. Any term at order k can be written as a sum of such terms for which $\sum_i \alpha_i = k$. The Lie bracket can be thought of as a linear operator on this space. To find the set of possible solutions is to find the range of this linear operator. Let us take the case where the linear part is diagonalizable and so just consists of the eigenvalues λ_i. It helps to visualize the equation as a matrix

$$\begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} \begin{pmatrix} \tilde{\theta}_1 \\ \vdots \\ \tilde{\theta}_n \end{pmatrix} - \begin{pmatrix} & & \\ & \partial_i (h_k)^j & \\ & & \end{pmatrix} \begin{pmatrix} \lambda_1 \tilde{\theta}_1 \\ \vdots \\ \lambda_n \tilde{\theta}_n \end{pmatrix} = \begin{pmatrix} \lambda_1 (f_k)^1 \\ \vdots \\ \lambda_n (f_k)^n \end{pmatrix}$$

Look at any jth row on the right hand side. Let us say for simplicity that $(f_k)^j = c\tilde{\theta}_1^{\alpha_1}...\tilde{\theta}_n^{\alpha_n}$ for some set of α_i. We will have to choose an $h_k^j = \tilde{\theta}_1^{\alpha_1}...\tilde{\theta}_n^{\alpha_n}$. The first term on the left hand side will only give a contribution $\lambda_j (h_k)^j$ to the jth row. The contribution from the second term will come from the jth row of the matrix $\mathcal{D}\boldsymbol{h}_k$. This is the matrix of partial derivatives and a derivative of a polynomial just pulls down a power. The i, j element has a derivative with respect to i and then will be multiplied by $\lambda_i y_i$. So, the equation to solve is is

$$\lambda_j (h_k)^j - (\sum_i \lambda_i \alpha_i)(h_k)^j = c\theta_1^{\alpha_1}...\theta_n^{\alpha_n}. \tag{3.16}$$

Now comparing terms, we see that to solve this equation, we need to set

$$a = \frac{c}{\lambda_j - \sum_i \lambda_i \alpha_i}. \tag{3.17}$$

The usual case of power law scaling is obtained when all nonlinear terms can be removed by such coordinate transformation. The fixed point, in this case, is called hyperbolic. If we have a term for which $\lambda_j = \sum_i \lambda_i \alpha_i$, this term is called a resonance and cannot be removed from the equation for $\frac{d\theta_j}{d\ell}$. This contributes

to the singularity at the fixed point which is no longer given by power law combinations.

3.1.2 Bifurcations

Notice a special case of these equations when for some k, a particular $\lambda_k = 0$. In this case, it is possible to get an infinite number of resonances because the equation $\lambda_i = \lambda_i + \alpha_k \lambda_k$ is also true for all α_k and λ_i. This case, when one of the eigenvalues goes to 0 depending on some parameter ϵ is called a *bifurcation*. If all linear eigenvalues λ_i of the flows are distinct and non-zero, which terms can be removed using polynomial coordinate changes depends only on these λ_i. As we saw, this approach can be formulated elegantly as a linear algebra problem of the Lie bracket on the space of homogeneous polynomials. For more general cases—including bifurcations—'hypernormal form' [115, 158, 157] theory develops a systematic but somewhat more brute-force machinery to identify which terms can and cannot be removed perturbatively by polynomial changes of coordinates. Classic bifurcations include the pitchfork bifurcation, the transcritical bifurcation, the saddle node and the Hopf bifurcation.

Confusingly, bifurcation theory separately has its own 'normal form' of bifurcations. These normal forms are derived in a very different way using the implicit function theorem. The basic idea is to ask for the smallest number of terms in the equation which will preserve the qualitative behaviour of the fixed points (eg. exchange of stability of fixed points), and then map any other equation on to this simple equation using an implicit function theorem. This mapping allows for too broad a class of transformations to be useful for our purposes. An

important feature of the flows that we want to preserve is their *analyticity*, we therefore only consider polynomial changes of coordinates.

An explicit example is given by the 4-d Ising model. It is known that the magnetization $M \sim t^{1/2}(\log t)^{1/3}$ with $\log\log$ corrections. The quartic coupling u and the temperature t have flow equations which traditional bifurcation theory would simplify to

$$\frac{du}{d\ell} = -\bar{B}u^2, \tag{3.18}$$

$$\frac{dt}{d\ell} = 2t \tag{3.19}$$

Calculating the magnetization with this set of flow equations leads to the wrong power of logarithmic corrections. By allowing too broad a class of coordinate transformations, bifurcation theory hides the true singularity in the non-analytic coordinate change. We will show that normal form theory instead predicts

$$\frac{du}{d\ell} = -\bar{B}u^2 + \bar{D}u^3, \tag{3.20}$$

$$\frac{dt}{d\ell} = 2t - \bar{A}tu \tag{3.21}$$

We will present the explicit solution of this equation in Section 3.4. Here, we just note that the traditional log and $\log\log$ terms follow from the solution's asymptotic behaviour. To get these equations, we will remove higher order terms in u by using a coordinate change that is lower in order (broadening the formalism we considered in Section 3.1.1). Using lower order terms to remove higher order terms is called hypernormal form theory in the dynamical systems literature. Here, we simply use normal form theory to denote any procedure that uses only polynomial changes of coordinates to change terms in flow equations.

We will explicitly work out the case of a single variable undergoing a bifurcation for the 4d Ising model and the 2d Random Field Ising model and show how there are only a finite number of terms which cannot be changed or removed. It is worth mentioning here that there can be cases in which two variables simultaneously have 0 eigenvalues. The XY model [96] offers an example where this happens. The dATG transition in 6 dimensions has two variables that simultaneously go through a transcritical bifurcation [45, 44]. Polynomial changes of coordinates in both variables can be used here too, but because there are generically more terms at higher order than at lower order (there are many more ways to combine two variables into a sixth order polynomial than there are to combine them into a third order polynomial), we usually do not have enough freedom to remove all terms. Therefore, simultaneous bifurcations in more than one variable often have an infinite number of terms in their flow equations that cannot be removed.

3.2 Earlier work

The approach we take is inspired by Wegner's early work [151, 152], subsequent developments by Aharony and Fisher [1, 2], and by studies of Barma and Fisher on logarithmic corrections to scaling [14, 15]. Salas and Sokal's work on the 2-d Potts model [140] is similar in spirit to ours.

Wegner [151] first constructed nonlinear scaling fields which transform linearly under an arbitrary renormalization group. His construction is very similar to the coordinate changes we considered above for normal form theory. The one difference is that Wegner explicitly allows the new coordinates to depend on the

coarse graining length ℓ. We will not allow this explicit dependence on ℓ in our change of coordinates, as it doesn't seem to offer any advantage over regular normal form theory.

Eventually, the goal of using normal form theory to understand the differential equations that describe RG flow is to simplify and systematize scaling collapses. This requires a systematic way of dealing with corrections to scaling beyond the usual power laws. There are three different types of corrections to scaling that have appeared in the literature. These include logarithmic, singular and analytic corrections to scaling. Logarithms in the scaling behaviour typically occur at an upper critical dimension or in the presence of a resonance. Wegner and Riedel [152] considered the the case of a zero eigenvalues which occurs at the upper critical dimension of Ising and tri-critical Ising models. They derived the form of the scaling in terms of logarithmic corrections to scaling. However, they used perturbation theory to ignore higher order terms in the flow equations rather than only keeping those terms which cannot be removed by an application of normal form theory. Here, we will solve the full flow equations and see that the logarithmic corrections to scaling are better incorporated as part of the true singularity using normal form theory.

Analytic corrections to scaling were explored by Aharony and Fisher [2] who gave a physical interpretation of the nonlinear scaling fields (see below Equation 3.5) in terms of analytic corrections to scaling in the Ising model. Analytic corrections to scaling capture the difference between the physical variable T and H (that your thermometer or gaussmeter measures) and the symbols $\tilde{t}$ and $\tilde{h}$ in the theory of the magnet. The liquid gas transition is in the Ising universality class but a theory of the liquid gas transition has to include analytic corrections

to scaling to match with the universal predictions of the Ising model. Moreover, such corrections are also needed to explain the non-universal behaviour away from the fixed point. Analytic corrections to scaling will correspond to terms in the differential equations that can be removed by coordinate changes.

The singular corrections to scaling are also incorporated as part of the true singularity with the addition of irrelevant variables. Finally, the ability to change the renormalization scheme leads to what are called redundant variables. Here, and in the next chapter, we will argue that these variables can be seen as a gauge choice which contributes to the corrections to scaling. We will distinguish irrelevant variables based on whether they lead to gauge corrections or genuine singular corrections to scaling. The leading irrelevant variable in the 3-d Ising model for example, contributes a correction to scaling to the magnetization of the Ising model which goes as t^θ. This is a genuine singular correction to scaling and θ is a fundamentally new eigenvalue. Irrelevant variables were examined by Barma and Fisher in the 2-d Ising model. They noticed that many of the irrelevant variables in the Ising model have resonances which should lead to logarithms in the scaling behaviour. However, they were not able to observe any evidence for the presence of such logarithms in the square lattice nearest-neighbour 2-d Ising model. We will explain the lack of these logarithms using gauge corrections to scaling.

Finally Salas and Sokal, in the context of the 2-d Potts model derive the normal form of the flow equations for a transcritical bifurcation. They do not solve the full flow equations but end up approximating the solution by logarithms. In the context of QCD, Sonoda derived the solution for the flow of a coupling which undergoes a transcritical bifurcation.

Despite similar inclinations, none of these works make the complete connection to normal form theory. One advantage of our approach is precisely that it brings together this disparate literature into a unified framework. The analysis presented here is general and applicable to all kinds of situations, ranging from old problems like the nonequilibrium random field Ising model (NERFIM) [122], to newer research problems like jamming [78].

3.3 Universality Families

Traditionally, the RG has had the concept of a universality class. The universality class is essentially determined by the critical exponents which explain the scaling behaviour of a model, i.e. by linearized RG eigenvalues. Normal form theory suggests another possible classification. Each fixed point can be classified by the bifurcation or resonance that it is at. The simplest case, which is also the traditional one, is the hyperbolic universality family. In the hyperbolic case, it is possible to remove all nonlinear terms in the flow equations by changes of coordinates. Hence, the RG can be written as a linear flow to all orders in perturbation theory. Different values for the linear eigenvalues correspond to different universality classes. While traditionally, this is a statement about the linearization of the RG, here it is a statement about the only terms in the flow equations that are *universal* in the sense that they can not be removed by a coordinate change.

The need for this generalization becomes clear when we examine cases which are not traditional. In Table 3.1 we present common universality families and well-studied statistical mechanics systems governed by each. The pitchfork

Universality family	Systems	Normal form	Invariant scaling combinations
Hyper− bolic	**3-d Ising Model** (t) 3-d RFIM (w)	$dt/dl = (1/\nu)t$	Lt^{ν}
Pitch− fork	**2-d RFIM** (w) 6-d Potts model (q)	$dw/dl = w^3 + Bw^5$	$Le^{1/(2w^2)}(1/w^2+B)^{-B/2}$
Trans− critical	**4-d Ising model** (u,t) 2-d NERFIM ($-w,S$) 1-d Ising model ($-t,h$)	$du/dl = -u^2 + Du^3$ $dt/dl = 2t-Atu$	$Le^{1/u-D}(1/(Du)-1)^D = Ly^D$ $tL^2(W(yL^{1/D})/(1/(Du)-1))^{-A}$
Resonance	**2-d Ising model**	$df/dl = 2f-t^2-L^{-2}$ $dt/dl = t+AL^{-1}$	$tL+A\log L$
Higher Codimension	**2-d XY model**	$dx/dl = -y^2(1+xf(x^2))$ $dy/dl = -xy$	$y^2 - 2\int_0^x \frac{s}{1+sf(s^2)}\,ds$ $= y^2 - x^2-(2f(0)/3)x^3 + (f(0)^2/2)x^4 + O(x^5)$

Table 3.1: Normal forms and universal invariant scaling combinations for traditional and intrinsically nonlinear renormalization-group critical points. The universal scaling of most critical points are power-law combinations of the control variables, derived from the linearized normal-form equations of hyperbolic RG fixed points. Many systems have well-studied logarithmic corrections, exponentially diverging correlations, or other singularities that we attribute to intrinsic nonlinearities in the RG flow equations. In blue are new universal terms predicted by our analysis of the corresponding dynamical system normal forms, which appear not to have been hitherto discussed in the literature. In green are terms we explain which have been previously observed using other methods [151, 108, 145, 104, 121]. The normal form equations are shown for the system in bold. Other systems in the same universality family have the same equations associated with different variables (shown in parenthesis). Many of the results quoted in the table were obtained in disparate literatures (QCD, glasses, critical phenomena etc.) but are united in this common framework. Other families are possible, the flow equations for the replica symmetry breaking transition in disordered media have a simultaneous transcritical bifurcation and possibly also have a Hopf bifurcation [44]

bifurcation shows up in the 2-d Random Field Ising model; it has a cubic term in the equations for w, the ratio of the disorder to the coupling [30]. We have derived that the correct equations require an additional w^5 term [82], which was not included in previous work. The 2-d Ising model has a well known logarithmic correction to the specific heat, which Wegner associated with a t^2 resonance term in the flow equation [151]. The 1-d and 4-d Ising models have transcritical bifurcations. The 1-d Ising case is somewhat special and we will cover it later in this chapter. These cover all the important bifurcations with one variable [2].

[2]We have not studied any example of a saddle node bifurcation which would require a

When more than one variable is undergoing a bifurcation, or if more than one variable has an inherently nonlinear flow, the analysis becomes considerably more complicated. This is evidenced in the the 2-d XY model at the Kosterlitz–Thouless (KT) transition [97]. It has been shown that the simplest normal form of its flow equations (in the inverse-temperature-like variable $x \sim 1/T - 1/T_c$ and the fugacity y) has an infinite number of universal terms, which can be rearranged into an analytic function f [121] (Table 3.1). We conjecture that the very similar transition observed in randomly grown networks [37, 61] is not in the KT-universality class, but rather is in the same universality family. It is not to be expected that a percolation transition for infinite-dimensional networks should flow to the same fixed point as a 2-d magnetic system, but it is entirely plausible that they share the same normal form with a different universal function f.

Different universality classes within the same universality family, such as those of the 4-d Ising model and the 2-d NERFIM have different power laws and scaling functions. However, as shown in Table 3.1, because they both have a transcritical bifurcation the two classes have the same complicated invariant scaling combinations [3]. This hidden connection is made apparent in the shared normal form, where the quartic coupling and temperature (u, T) in the first class are associated with the (negative of) disorder strength and avalanche size $(-w, S)$ in the second [4].

Indeed, the normal form not only unites these universality classes, but al-

transition from a critical point to no critical point.

[3] A correlation length y^{-D} from Table 3.1 defined in terms of the marginal variable in both cases diverges exponentially; in terms of the temperature the correlation length is a power law

[4] The minus sign on w and t for the 1-d Ising and the NERFIM is because w and t are marginally relevant whereas u is marginally irrelevant for 4-d Ising.

lows a more precise handling of their singularity. It is usually stated that the magnetization $M \sim t^{1/2}(\log t)^{1/3}$, the specific heat $C \sim (\log t)^{1/3}$ and the susceptibility $\chi \sim (\log t)^{1/3}/t$ with $\log\log$ corrections [152]. We show in the appendix that the true singularity of the magnetization at the critical point is $M \sim t^{1/2}W(xt^{-27/25})^{1/3}$, where W is the Lambert-W function defined by $W(z)e^{W(z)} = z$, and $x[u]$ is a messy but explicit function of the irrelevant variable u. (The traditional log and log-log terms follow from the asymptotic behaviors of $W(x)$ at large and small x. The universal power 27/25 becomes manifest in the complete singularity, but is disguised into a constant factor up to leading logs.)

In the sections below, we derive in detail the scaling form for the random field Ising model (RFIM) and the 4-d Ising model. We derive the scaling forms and use to perform scaling collapses of numerical simulations. We then investigate the scaling of the Ising model in dimensions 1, 3 and 2.

3.4 4-d Ising

The study of critical points using the renormalization group was turned into a dynamical system problem by Wilson [156]. These RG calculations are done by first expressing the Ising model as a field theory with a quartic potential ϕ^4. Then by coarse-graining in momentum space and rescaling, one obtains the flow equations

$$dt/d\ell = 2t - \bar{A}tu + \bar{C}tu^2 + \bar{E}tu^3 + \bar{G}tu^4 + \bar{I}tu^5 + \bar{K}tu^6..., \tag{3.22}$$

$$du/d\ell = \epsilon u - \bar{B}u^2 + \bar{D}u^3 + \bar{F}u^4 + \bar{H}u^5 + \bar{J}u^6 + \bar{L}u^6..., \tag{3.23}$$

$$df/d\ell = (4-\epsilon)f + ..., \tag{3.24}$$

where t is the temperature, f is the free energy and u is the leading irrelevant variable (given by the quartic term in Landau theory). This is the highest order to which the flow equations are known as of now. The coefficients take the values, $\bar{A} = 1$, $\bar{B} = 3$, $\bar{C} = 5/6$, $\bar{D} = 17/3$, $\bar{E} = -7/2$, $\bar{F} \approx 32.54$, $\bar{G} \approx 19.96$, $\bar{H} \approx -271.6$, $\bar{I} \approx -150.8$, $\bar{J} \approx 2849$, $\bar{K} \approx 1355$, $\bar{L} \approx -34776$ [93, 48]. The flow equation for u in this case takes the form of a transcritical bifurcation with parameter $\epsilon = 4 - d$ tuning the exchange of stability between the Gaussian ($u = 0$) and Wilson-Fisher fixed point ($u \neq 0$).

Consider these equations for $\epsilon = 0$ which is the point at which it undergoes a transcritical bifurcation. To derive the normal form, one considers a change of variables of the form

$$t = \tilde{t} + a_1 \tilde{t}\tilde{u} + a_2 \tilde{t}\tilde{u}^2 + ..., \tag{3.25}$$

$$u = \tilde{u} + b_1 \tilde{u}^2 + b_2 \tilde{u}^3 + b_3 u^4 + ... \tag{3.26}$$

This gives the equations up to order u^4,

$$d\tilde{t}/d\ell = 2\tilde{t} - \bar{A}\tilde{t}\tilde{u} + (-\bar{A}b_1 + a_1\bar{B} + \bar{C})tu^2 + ..., \tag{3.27}$$

$$d\tilde{u}/d\ell = -\bar{B}\tilde{u}^2 + \bar{D}\tilde{u}^3 + (-b_1^2\bar{B} + b_2\bar{B} + b_1\bar{D} + \bar{E})\tilde{u}^4 + ... \tag{3.28}$$

Notice that the coefficients $\bar{A}$, $\bar{B}$ and $\bar{D}$ remain unchanged with this change of variables. The coefficients $\bar{C}$ and $\bar{E}$ are changed (though the change is independent of a_2 and b_3 because they are resonances) and in particular, can be set to 0 by an appropriate choice of coefficients. This creates a general procedure

for reducing this flow to its simplest possible form. First, all terms that are not resonances are removed in the usual way by solving Equation 3.17. Then, we perturbatively remove most of the resonances using the following procedure. First consider the u flow. Suppose the lowest order term in the flow after the u^3 term is u^n, i.e.

$$\frac{du}{d\ell} = -\bar{B}u^2 + \bar{D}u^3 + \bar{N}_n u^n + O(u^{n+1}) \tag{3.29}$$

with $n > 3$. Consider a change of variables of the form $u = \tilde{u} + b_{n-2}\tilde{u}^{n-1}$. Then

$$\begin{aligned}(1 + (n-1)b_{n-2}\tilde{u}^{n-2})\frac{d\tilde{u}}{d\ell} &= -\bar{B}(\tilde{u} + b_{n-2}\tilde{u}^{n-1})^2 + \bar{D}((\tilde{u} + b_{n-2}\tilde{u}^{n-1})^3 \\ &\quad + \bar{N}_n(\tilde{u} + b_{n-2}\tilde{u}^{n-1})^n + O(\tilde{u}^{n+1}), \qquad (3.30)\\ \frac{d\tilde{u}}{d\ell} &= \frac{-\bar{B}\tilde{u}^2 + \bar{D}\tilde{u}^3 + \bar{N}_n\tilde{u}^n - 2\bar{B}b_{n-2}\tilde{u}^n}{(1 + (n-1)b_{n-2}\tilde{u}^{n-2})} + O(\tilde{u}^{n+1}), \qquad (3.31)\\ &= -\bar{B}\tilde{u}^2 + \bar{D}\tilde{u}^3 + (\bar{N}_n - 2\bar{B}b_{n-2} + (n-1)b_{n-2}\bar{B})\tilde{u}^n + O(\tilde{u}^{n+1}). \qquad (3.32)\end{aligned}$$

Evidently, the coefficient of the $\tilde{u}^n$ term can be set to 0 with an appropriate choice

$$b_{n-2} = \frac{N_n}{\bar{B}(3-n)}. \tag{3.33}$$

So all terms of the form u^n with $n > 3$ can be removed by a change of coordinates. Incidentally, this derivation also shows why it is not possible to remove the u^3 term. Now consider the t equation with

$$\frac{dt}{d\ell} = 2t - \bar{A}t\tilde{u} + M_n t\tilde{u}^{n-1} + O(t\tilde{u}^n). \tag{3.34}$$

We consider a change of coordinates

$$t = \tilde{t} + a_{n-2}\tilde{t}\tilde{u}^{n-2}. \tag{3.35}$$

It is then straightforward to show

$$\frac{d\tilde{t}}{d\ell} = 2\tilde{t} - \bar{A}\tilde{t}\tilde{u} + (M_n + \bar{B}(n-2)a_{n-2} + a_{n-2}\bar{A})\tilde{t}\tilde{u}^{n-1} + O(t\tilde{u}^n). \tag{3.36}$$

So setting

$$a_{n-2} = -\frac{M_n}{\bar{B}(n-2) + A} \tag{3.37}$$

sets the coefficient of the tu^{n-1} term with $n > 2$ to 0.

Any term which is not of this form can be removed in the usual way by solving Equation 3.17. Hence, by considering all such polynomial change of coordinates, we can reduce this set of equations to their normal form

$$d\tilde{t}/d\ell = 2\tilde{t} - A\tilde{u}\tilde{t}, \tag{3.38}$$

$$d\tilde{u}/d\ell = -\tilde{u}^2 + D\tilde{u}^3, \tag{3.39}$$

$$d\tilde{f}/d\ell = 4\tilde{f} - \tilde{t}^2. \tag{3.40}$$

The normal form variables $\tilde{t}$, $\tilde{u}$, $\tilde{f}$ are equal to the physical variables t, u and f to linear order (up to a rescaling). Corrections to these are analytic corrections to scaling. Hence, we will henceforth simply refer to the normal form variables as t, u and v. It is important to note that we are making a particular choice for the analytic corrections to scaling by setting them equal to zero. It is possible to make a different choice for the higher order coefficients. In particular, the equation for $du/d\ell$ goes to ∞ at finite ℓ if u starts at a large enough value. Hence, it may be more useful to make a different choice for the higher order coefficients. All of these choices will agree close to the critical point but will have different behaviour away form the critical point. Later, we will consider a different choice for the higher order terms.

The 4-d Ising model has both a bifurcation and a resonance. The u^2, u^3 and *Aut* terms come from the bifurcation and cannot be removed by an analytic

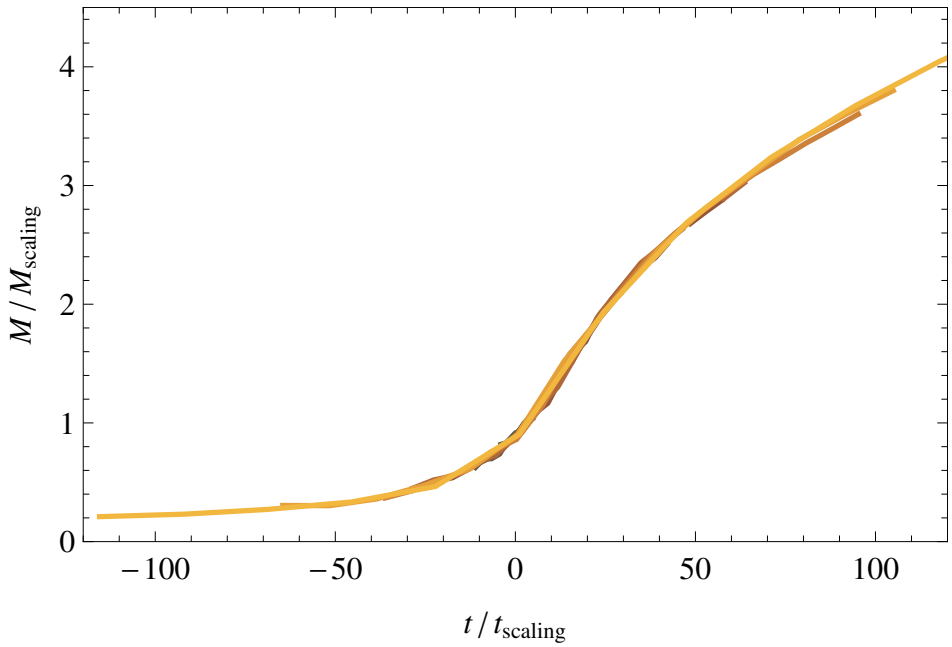

Figure 3.1: Scaling collapses for the magnetization and susceptibility using the scaling form given by the normal form equations. Simulations are done on a 4-d lattice using a Wolff algorithm for lattice sizes ranging from $L = 4$ to $L = 32$. Here $M_{\text{scaling}} = L((W(yL^{1/D}) + 1))^{1/4}$ and $t_{\text{scaling}} = L^{-2}(W(yL^{1/D})/(1/Du_0 - 1))^{1/3}((W(yL^{1/D}) + 1))^{-1/2}$. We find $u_0 = 0.4 \pm 0.1$ for the 4-d nearest-neighbor hypercubic-lattice. An estimate of the error is given by estimating u_0 with a different choice of normal form which gives $u_0 = 0.5$.

change of coordinates. The t^2 term is a consequence of an integer resonance between the temperature and free energy eigenvalue, $\lambda_t = 1/\nu = 2, \lambda_f = d = 4$. Finally, we have another degree of freedom that we have used. We can rescale u and t to set some of the nonlinear coefficients to 1. This reflects the fact that the original coefficients $\bar{A}$, $\bar{D}$ depend on an arbitrary scale of u and t that we have chosen. By choosing the scale so $\bar{B} = 1$ and the coefficient of the resonance is -1, defines $D = \bar{D}/\bar{B}^2$ and $A = \bar{A}/\bar{B}$. The resultant equations then have 2 parameters A and D which are *universal*, in a way that is similar to the eigenvalues of the RG flows 3.1.

Before examining the full solution of 3.38 – 3.40, we will first study the effect of each part of the RG flows. First, considering only the linear terms and

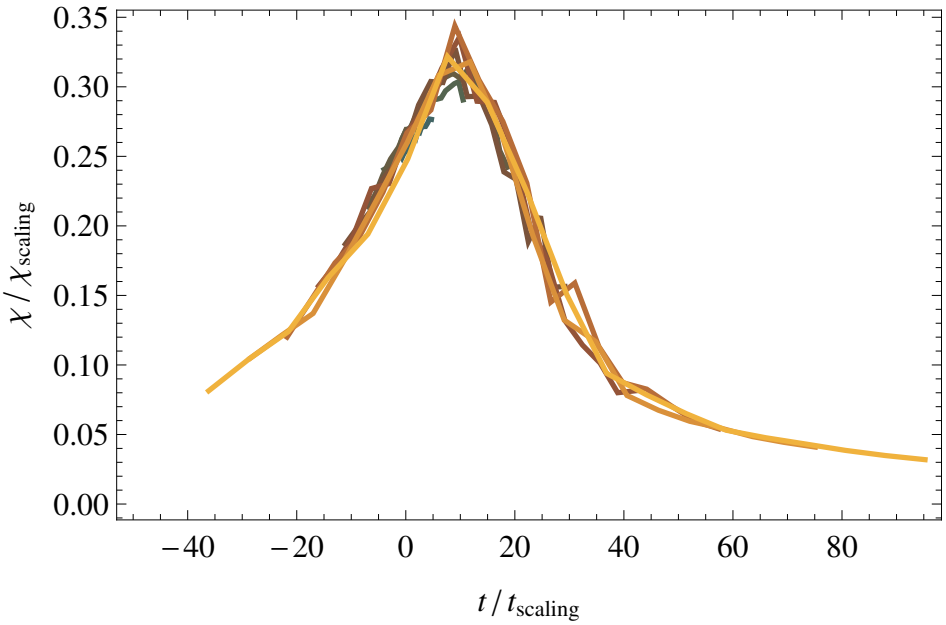

Figure 3.2: Scaling collapse for the susceptibility using the scaling form given by the normal form equations. Simulations are done on a 4-d lattice using a Wolff algorithm for lattice sizes ranging from $L = 4$ to $L = 32$. Here $\chi_{\text{scaling}} = L^2((W(yL^{1/D}) + 1))^{1/2}$ and $t_{\text{scaling}} = L^{-2}(W(yL^{1/D})/(1/Du_0 - 1))^{1/3}((W(yL^{1/D}) + 1))^{-1/2}$. We find $u_0 = 0.4$ for the 4-d nearest-neighbor hypercubic-lattice.

coarse-graining until $t(\ell^*) = 1$, the free energy is given by $f \sim t^2$. This is the mean-field result and also the traditional scaling form that RG results take in the absence of nonlinear terms in the flow equations. Second, we include the resonance between the temperature and free energy eigenvalue, which leads to an irremovable t^2 term in the flow equation for the free energy. This term cannot be removed by analytic coordinate changes, and yields the log correction to the specific heat. Third, the irrelevant variable u undergoes a transcritical bifurcation. Results in the hyper-normal form theory literature, as well as some articles in the high-energy theory literature [145, 104] recognize that the simplest form that the equation can be brought into is Equation 3.39. The solutions of 3.38 – 3.39 given in Appendix A are $u(\ell) = 1/(D(1 + W(ye^{\ell/D})))$ and $t(\ell) = t_0 e^{2\ell}(W(ye^{\ell/D})/(1/(Du_0) - 1))^{-A}$ where $y[u_0]$ is again a messy but explicit

function: $y = (1/(Du_0) - 1)\exp(1/(Du_0) - 1)$. The traditional log and log-log corrections are derived by expanding the W function for large ℓ.

Let us use this to derive the finite size scaling form of the free energy. Early finite-size scaling work [4, 98, 111] attempted scaling collapses with logs; recent work does not attempt collapses at all [103]. Finite size scaling requires an equation for the magnetic field, h, given by $dh/d\ell = 3h$. Explicit calculations show that the coefficient of the hu term is zero; we we will explain in the next section by considering redundant variables. The free energy is then a function of three scaling variables, $u(\ell)$, $t(\ell)$ and $h(\ell)$. It is given by

$$\begin{aligned} f(t_0, u_0) = e^{-4\ell} f(t(\ell), u(\ell), h(\ell)) \\ - W(ye^{\ell/D})^{-A}\left(\frac{W(ye^{\ell/D})^{-A}}{1-A} - \frac{1}{A}\right). \end{aligned} \tag{3.41}$$

To get a finite size scaling form, we coarse-grain until $\ell = \log L$, the system size. Note that $u(L)$ cannot just be ignored, because it is a dangerous irrelevant variable. However we can account for it by taking the combination $t(L)/(u(L))^{1/2}$ and $h(L)/(u(L))^{1/4}$ as our scaling variables [28]. The scaling form of the free energy then depends on u_0 which we treat as a fit parameter in the scaling form of the susceptibility:

$$\chi = L^2\left(W(yL^{\frac{1}{D}}) + 1\right)^{\frac{1}{2}}\Phi\left(t_0 L^2\left(\frac{W(yL^{1/D})}{1/(Du_0) - 1}\right)^{-A}\right). \tag{3.42}$$

At the critical point $t = 0$, the function Φ must be analytic for finite L (since non-analyticity requires an infinite system size). $\Phi(0)$ is therefore a constant independent of L and u_0 at $t = 0$. Using this, u_0 may be estimated from χ at different values of L by fitting to its predicted dependence $\chi \propto L^2(W(y[u_0]L^{1/D}) + 1)^{1/2}$ with $y[u_0]$ defined above.

The magnetization is collapsed using the best-fit value of $u_0 = 0.4$. Figures 3.1–3.2 shows the scaling collapse of the magnetization and susceptibility. Though our collapses are not significantly better than the traditional logarithmic forms, the traditional collapses give a value $u_0 = 0.8$ which is somewhat high. The correct form of the singularity will be more apparent at larger values of u_0, because the $\log\log$ term which is the second term in the asymptotic expansion of the W function is small compared to the $\log$ except at large u_0 and small L. Changing the value of u_0 will require a model different from the nearest neighbour square lattice Ising model.

3.4.1 Gauge variables in the 4-d Ising model

So far, we have been considering the effects of changing coordinates in the control variables on the predictions of the theory. Wegner [150] considered changing coordinates in the degrees of freedom of the theory removing what he termed 'redundant' terms in the Hamiltonian. Here, and in the next chapter, we will generalize this idea and see how changing coordinates in the results of our theory simplify the RG flows. As an example, the flow equations for the Ising model are usually derived by first considering the Landau potential [39]

$$\mathcal{V} = \int_x h\phi + t\phi^2 + v\phi^3 + u\phi^4 \tag{3.43}$$

with h the field, t the distance from the critical temperature, v and u the parameters of the cubic and quartic term in the potential. We know that t is a relevant variable and u is a marginally irrelevant. It is natural that v would be a relevant variable in the RG. However, it does not enter into the flow equations or the predictions. This is because it is possible to change coordinates in the field ϕ to

remove the cubic term in the potential. To see this, consider a constant shift in the value of $\phi \to \phi - \kappa$. This gives a new potential

$$\mathcal{V} = \int_x h(\phi - \kappa) + t(\phi - \kappa)^2 + v(\phi - \kappa)^3 + u(\phi - \kappa)^4. \tag{3.44}$$

Ignoring the additive constant term this adds, we can simplify this to

$$\mathcal{V} = \int_x (h - 2t\kappa + 3v\kappa^2 - 4u\kappa^3)\phi + (t - 3v^2\kappa + 6u\kappa^2)\phi^2 + (v - 4u\kappa)\phi^3 + u\phi^4 \tag{3.45}$$

Suppose we choose κ to cancel out the cubic term so $\kappa = v/(4u)$. The new potential then is

$$\mathcal{V} = \int_x (h - tv/(4u) + v^2/(8u^2))\phi + (t - 3v^2/(8u))\phi^2 + u\phi^4. \tag{3.46}$$

So, we can set $v = 0$ if we define the new variables

$$u_h = h - \frac{tv}{u} + \frac{v^2}{8u^2}, \tag{3.47}$$

$$u_t = t - \frac{3v^2}{8u}. \tag{3.48}$$

Redundant variables are a kind of gauge choice. It is possible to chose a gauge to eliminate them. We will explore this interpretation further in the 2-d Ising model and in the next chapter. For now, we note that the choice has an interesting effect on the flow equations. In the old variables, suppose we start with the flow equations of the 4-d Ising model in variables t, u, h, v. Their normal form is given by

$$\frac{du}{d\ell} = -u^2 + Du^3, \tag{3.49}$$

$$\frac{dh}{d\ell} = 3h - Ehu + a1v^3 + a_3v^3u + a^5vt + a_6vtu, \tag{3.50}$$

$$\frac{dt}{d\ell} = 2t - Atu + a_2v^2 + a_4v^2u, \tag{3.51}$$

$$\frac{dv}{d\ell} = v - Fvu. \tag{3.52}$$

Suppose we make the change of variables given above. We can write the new set of equations in u_h and u_t. They are

$$\frac{du_t}{d\ell} = 2u_t - Auu_t + \frac{1}{8}\left((-3 - 3A + 8a_2 + 6F)v^2 + (8a_4 + 3D)uv^2\right), \tag{3.53}$$

$$\begin{aligned}\frac{du_h}{d\ell} &= 3u_h - Euu_h + \frac{1}{16u}\left(8(-1 + A + 2a_5 - E + F)u_t u_v + 8(2a_6 + D)u_t u^2 v\right) \\ &+ \frac{1}{16u}\left((1 + eA - 8a_2 + 6a_5 - E - 3F)v^3\right) \\ &+ \frac{1}{16u}\left((16a_1 - 8a_4 + 6a_6 - D)uv^3 + 16(a_3)u^2v^3\right).\end{aligned} \tag{3.54}$$

All terms in parenthesis in the above equation have factors of v next to them. However, now that we have chosen a gauge to eliminate v, the physics of the problem cannot possibly depend on it. This gives a set of 7 equations which set constraints on the coefficients of the RG equations in the original variables. We have freedom to choose 7 coefficients $a_1, a_2, a_3, a_4, a_5, a_6, F$. It turns out that the above set of linear equations do not have any solution in these 7 variables. However, they do have a solution in the variables $a_1, a_2, a_3, a_4, a_5, a_6, E$. The solution is

$$\begin{aligned} a_1 &= D/16, \\ a_2 &= 3/8(1 + A - 2F), \\ a_3 &= 0, \\ a_4 &= -3D/8, \\ a_5 &= 1/4(-1 + A - 2F), \\ a_6 &= -D/2, \\ E &= 1/2(-1 + 3A). \end{aligned}$$

In the 4-D Ising model, $A = 1/3$. This means that, if the above equations hold, $E = 0$. Hence this gauge variable explains why the equation for h does not have a resonance term Euh, explaining this result discussed above in Section 3.4.

3.4.2 Choice of normal form

Normal form theory makes a particular choice for what to do with the coefficients that can be changed by coordinate changes, it sets them equal to zero. In general, however, it is not clear that the best choice to make is to set them equal to zero. Consider the equation

$$\frac{du}{d\ell} = -u^2 + Du^3 \tag{3.55}$$

which, as we saw, has the solution $u(\ell) = 1/(D(1+W(ye^{\ell/D})))$. Here, $y = (1/(Du_0) - 1)\exp(1/(Du_0) - 1)$. Note that $u_0 > 0$ as a requirement for the stability of the free energy. If $u_0 < 1/D$, then $y > 0$, and if $u_0 <= 1/D$, $y <= 0$. Hence, the domain of attraction of the fixed point at $u_0 = 0$ has a length $1/D$. If we have a system where $u_0 > 1/D$, then this will lead to $u(\ell) \to \infty$ in a finite coarse graining length. This is reflected in the branch cut of the W function at $-1/e$. In the context of high energy physics, some have tried to find deep meaning into this pole [104].

However, for scaling purposes, we generally prefer a choice of coordinates for which there is no such unphysical behaviour. One natural choice is to instead use the equations

$$\frac{du}{d\ell} = -\frac{u^2}{1 + Du}. \tag{3.56}$$

For small u, this has the same behaviour as Equation 3.55. However, the behaviour at large u is now well behaved. The solution of this equation is

$$u(\ell) = \frac{1}{DW(e^{\ell/D}1/(Du_0)\exp(1/(Du_0)))}, \tag{3.57}$$

which is in fact somewhat simpler that the solution to Equation 3.55. Scaling collapses with this choice of normal form for the susceptibility are shown in the case of the Ising model in Figure 3.3. Better numerics are needed to tell which choice of normal form is best.

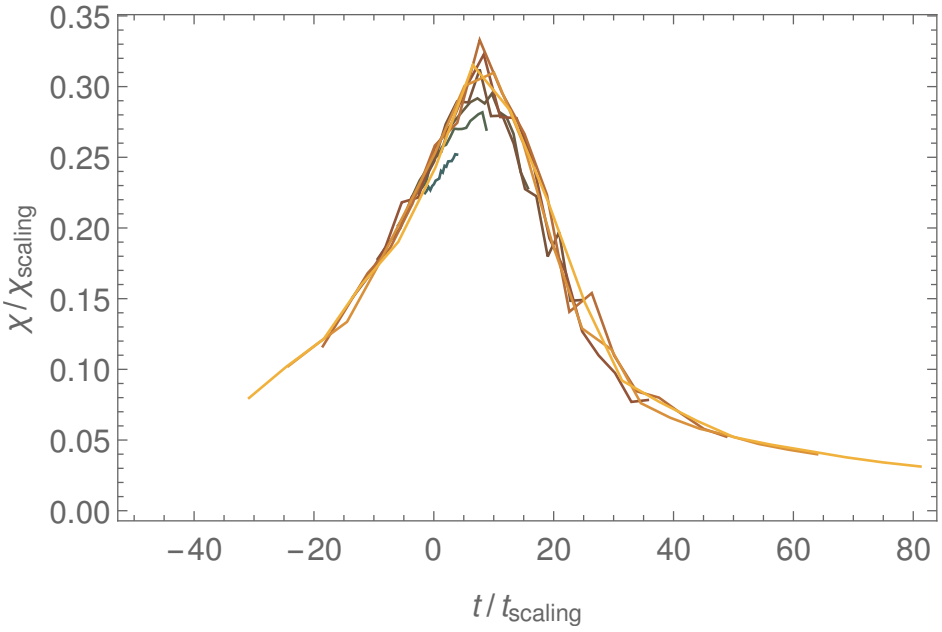

Figure 3.3: Scaling collapse for the susceptibility using the scaling form given by a different choice of normal form derived from Equation 3.56. Simulations are done on a 4-d lattice using a Wolff algorithm for lattice sizes ranging from $L = 4$ to $L = 32$. Here $\chi_{\text{scaling}} = L^2((W(yL^{1/D}))^{1/2}$ and $t_{\text{scaling}} = L^{-2}(W(yL^{1/D}))^{-1/6}/(1/Du_0))^{1/3} \exp(1/(3W(yL^{1/D})))$ with $y = 1/(Du_0)e^{1/(Du_0)}$. We find $u_0 = 0.5$ for the 4-d nearest-neighbor hypercubic-lattice using this method.

3.5 Random Field Ising model

Finding critical exponents for the random field Ising model has been a long standing challenge in physics. Some initial results used supersymmetry to prove an equivalence of the Random Field Ising model in dimensions $d+2$ with the Ising model in dimensions d [120]. It was later shown that the lower critical dimension of the Random Field Ising model is not 3 (as would be expected form such a correspondence) but rather 2 [83]. The upper critical dimension is in 6 dimensions. Here, we will look at the scaling behaviour of the Random Field Ising model at its lower critical dimensions.

Considered a spin system with a random field.

$$\mathcal{H} = -\sum_{\langle ij \rangle} J s_i s_j + \sum_i h_i s_i. \tag{3.58}$$

Here, J is the nearest neighbour coupling and h_i is a random field chosen from a Gaussian distribution with width r. A phenomenological theory for the RG was formulated by Bray and Moore [30]. It turns out to be useful to define a quantity $w = r/J$. Then, using heuristic arguments on the stability of domain walls, they derive

$$\frac{dw}{d\ell} = -\epsilon/2w + Aw^3 \tag{3.59}$$

with $\epsilon = d - 2$, and d is the dimension. Note that the flow equations have a symmetry under $w \to -w$ because the physics is invariant under $r \to -r$ about the critical point at $r = 0$. This is an example of a pitchfork bifurcation. Bray and Moore argue for this scaling form by looking at the scaling of r and J separately. The scaling of J is given by looking at the energy of a domain wall of size b^d. The energy of the domain wall is proportional to b^{d-1}. By considering the cost of roughening the domain wall because of the presence of random fields, which goes as r^2, they are able to derive the next term in the equation for J which is now

$$\frac{dJ}{d\ell} = (d-1)J + Dw^2 J + O(w^4). \tag{3.60}$$

For the random field r, the energy of a region of size b^d is proportional to $b^{d/2}$. Any corrections requires forming a domain of 'wrong spins' which, being akin to a barrier crossing problem, is exponentially suppressed. Hence the equation for r is given by

$$\frac{dr}{d\ell} = \frac{d}{2} r \tag{3.61}$$

with exponentially small corrections. These two equations together can be used to derive Equation 3.59. Bray and Moore conjecture that Equation 3.61 holds

exactly to all order in w (up to exponential corrections). However, it is possible for Equation 3.60 to have higher order terms in w and thus Equation 3.59 is only correct to order w^5. Integrating Equation 3.59, we get $\ell \sim -1/(2Aw^2) + 1/(2Aw_0^2)$. This implies that the correlation length is

$$\xi \sim e^{1/(2Aw_0^2)}. \tag{3.62}$$

For finite size systems, the system size $L \sim \exp(1/(2Aw_0^2)$. Meinke and Middleton [108] showed that their finite size data was much better fit by a function of the form $w_0^{-2y}\exp(C/w_0^2)$ where C is a constant they fit to (Δ_0 in their notation) and $y = 1.07$. We will show that this prediction is consistent with the results of normal form theory.

As we have already argued, there is no reason Equation 3.59 is true to all orders in w. Indeed the, normal form prediction for the flow equations can be derived in a straight forward way. Consider adding a term $A_n w^n$ to Equation 3.59 at $\epsilon = 0$. This is a resonance and can not be removed usually under normal form theory. Suppose we make a change of coordinates $w = \tilde{w} + a_n \tilde{w}^{n-2}$. Then, to order $O(\tilde{w}^n)$, we get

$$\frac{d\tilde{w}}{d\ell} = A\tilde{w}^3 + (3Aa_n - Aa_n(n-2) + A_n)\tilde{w}^n + O(\tilde{w}^{n+1}). \tag{3.63}$$

We can set the coefficient of $\tilde{w}^n = 0$ if we use $a_n = A_n/((n-5)A)$. This procedure fails for $n = 5$ but works for all $n > 5$. [5] Hence, the normal form of the

[5] We note that we are assuming here that the coordinate transformations respect the symmetry of the problem $w \to -w$. Otherwise, it is possible to remove the $\tilde{w}^5$ term at the cost of introducing a $\tilde{w}^4$ term.

equilibrium RFIM is given by

$$\frac{d\tilde{w}}{d\ell} = \tilde{w}^3 - D\tilde{w}^5. \tag{3.64}$$

As before, we have used the freedom to rescale w to sets the coefficient of the $\tilde{w}^3$ term to 1.

The solution of this equation gives us an expression for the correlation length

$$\xi \sim (1/w^2 - D)^{D/2} e^{1/w^2}. \tag{3.65}$$

This scaling form could explain the data in Meinke and Middleton with D as a fit parameter. Notice that for this to work, D must be positive. However, this solution has the strange property that the correlation length goes to 0 for $w^2 = 1/D$. If $w > 1/D$, $w(l)$ decreases until it reaches $1/D$. If $w < 1/D$, it increases until it reaches $1/D$. As in the 4-d Ising model, it may be more useful to consider instead the flow equation

$$\frac{d\tilde{w}}{d\ell} = \frac{\tilde{w}^3}{1 + D\tilde{w}^2}. \tag{3.66}$$

This gives the scaling form

$$\xi \sim e^{1/(2w^2)} (w^2)^{-D/2}. \tag{3.67}$$

This is exactly consistent with the scaling form Meinke and Middleton use to collapse their data. Their numerics estimate the universal value for $D \approx 2.14$ [6].

3.6 3-d Ising model

There is a sense in which the Ising model is simplest in 3 dimensions because it is part of the hyperbolic universality family. It is also the first natural application

[6] Note that they also have a fit parameter which sets the scale of the exponential. However, this parameter is not universal since it depends on the scale of w unlike D

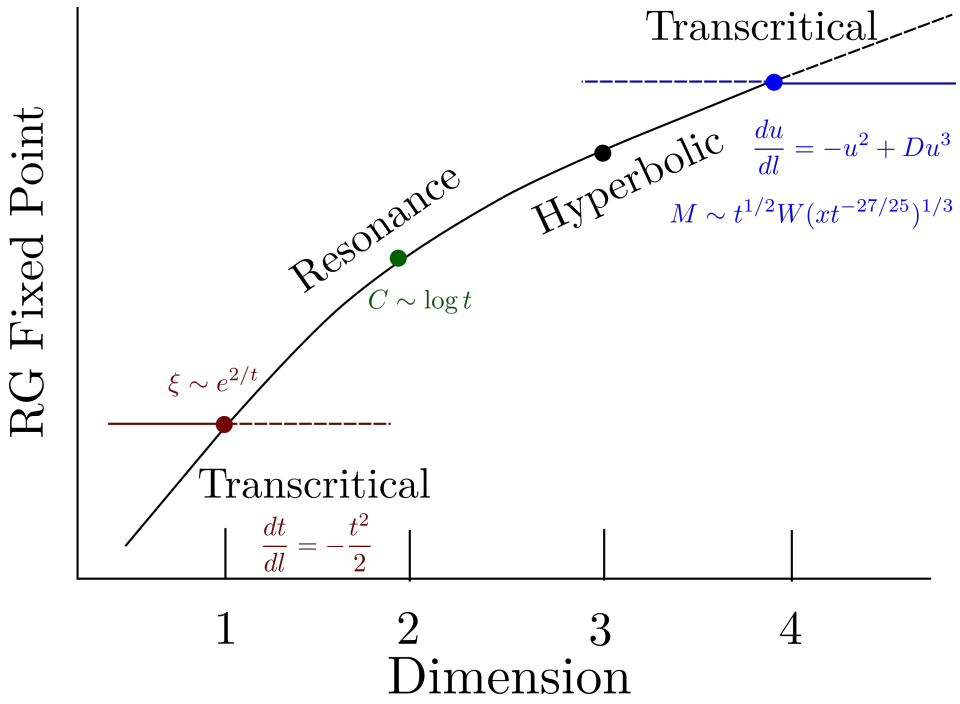

Figure 3.4: Fixed points as a function of dimension in the Ising model. There is a transcritical bifurcation in both 4 and 1 dimensions, leading to W functions and exponential correlation lengths respectively. The fixed point in 3 d is hyperbolic and the flow can be linearized. The fixed point in 2 d has a resonance which leads to a logarithmic specific heat. The challenge is to find a scaling form which interpolates between dimensions giving the correct behaviour in all of these dimensions.

of the ϵ expansion. The transcritical bifurcation at 4 dimensions leads to an exchange of stabilities of the Gaussian fixed point and the Wilson Fisher fixed point at a non-zero value of $u = u^*$. About this Wilson Fisher fixed point, the flow equations of the 3-d Ising model are in the hyperbolic universality class with linear coefficients which define the Ising universality class.

However, another approach is to consider the scaling form as a function of the dimension ϵ in a way that is well defined even at $\epsilon = 0$. Doing this, naturally requires us to keep nonlinear terms in the equation because we already know that the 4-d Ising model has nonlinear terms in its flow equations.

We want to write the flow equations about the 3-d fixed point but keep the nonlinear terms required for the scaling form to have the correct limiting behaviour in 2-d and in 4-d. We can write the normal form of the flow equations as

$$d\tilde{t}/d\ell = \lambda_t \tilde{t} - A\tilde{u}\tilde{t}, \tag{3.68}$$

$$d\tilde{u}/d\ell = \lambda_u \tilde{u} - \tilde{u}^2 + D\tilde{u}^3, \tag{3.69}$$

$$d\tilde{f}/d\ell = d\tilde{f} - \tilde{t}^2, \tag{3.70}$$

$$d\tilde{h}/d\ell = \lambda_h h. \tag{3.71}$$

We have included the nonlinear terms in u required for the correct scaling behaviour and the resonance between the temperature and the free energy. As usual, we switch notation to t, h and u with the understanding that they are different from the normal form variables by analytic corrections. Let us look at the scaling variable formed with t and u which can be obtained by solving

$$\frac{d\tilde{t}}{d\tilde{u}} = (\lambda_t \tilde{t} - A\tilde{u}\tilde{t})/(\lambda_u \tilde{u} - \tilde{u}^2 + D\tilde{u}^3). \tag{3.72}$$

The solution of this equation gives the scaling variable

$$t(u)^{-\frac{\lambda_t}{Du_1u_2}}(u-u_1)^{-\frac{(\lambda_t - Au_1)}{Du_1(u_1-u_2)}}(u-u_2)^{\frac{\lambda_u - Au_2}{D(u_2-u_1)u_2}} = \text{const} \tag{3.73}$$

where u_1 and u_2 are the two roots of the equation which to first order in λ_u are given by $u_1 = \lambda_u$ and $u_2 = 1/D - \lambda_u$. The form of the scaling variable is interesting, it is essentially given by a product of the linearized scaling variables at the three fixed points that the equation has. Taking the limit $\epsilon \to 0$, we get

$$te^{-2/u}u^{2D-A}(1-Du)^{A-2D} = \text{const} \tag{3.74}$$

which is the right scaling variable in 4-d. We have not yet been able to obtain an analytical form for the scaling variable involving t and h. This is because the equation for $u(l)$ does not seem to have a closed form solution here unlike the 4 dimensional case. Nevertheless, we are motivated by an attempt to create scaling variables which interpolate between different dimensions and have the correct scaling behaviour in many dimensions going down from 4 to 1. Once the full scaling variables are written down, a first test would be to see if these scaling variables do better of collapsing the numerical data in 3-d.

3.7 1-d Ising model

The 1-d Ising model is somewhat different because it is the lower critical dimension and does not have a phase transition. The 1-d Ising model has an exact solution which can be obtained by using transfer matrices. The partition function can be written as the trace of a transfer matrix $\mathcal{T}^N$ where N is the number of spins in the system. The matrix $\mathcal{T}_{ij} = e^{-\beta\mathcal{H}(s_i s_j)}$. Coarse graining here can be done by a well defined procedure, the coarse grained transfer matrix is defined as $\tilde{\mathcal{T}} = \mathcal{T}^b$ where b is the coarse graining length scale. Defining $\ell = \log b$ and expanding for b close to 1, we can get flow equations for the temperature T

$$\frac{dT}{d\ell} = -\frac{T^2}{2}\sinh\left(\frac{2}{T}\right)\log\left(\tanh\left(\frac{1}{T}\right)\right). \tag{3.75}$$

This is different from the flow equations we have considered so far because of the presence of non-analytic terms in the flow. The non-analytic term which multiplies the T^2 term is $= -1$ at $T = 0$. So, this equation is a transcritical bifurcation

$$\frac{dT}{d\ell} = \frac{T^2}{2} + ... \tag{3.76}$$

where the additional terms are non-analytic at $T = 0$. This can be used to derive a correlation length $\chi \sim \exp(2/T)$. To interpret the flow further, consider the change of coordinates $\kappa = \exp(-2/T)$. In these variables, the flow is

$$\frac{d\kappa}{d\ell} = (\kappa^2 - 1) \log\left(\frac{1-\kappa}{\kappa+1}\right). \tag{3.77}$$

Evidently, the flow is analytic in this variable. Solving the full flow Equation 3.75 gives $\chi \sim -1/(\log \tanh(1/T))$.

For non-zero ϵ, this argument is usually extended in what is called a Migdal-Kadanoff procedure for doing RG [89]. The flow equations are identical except for the presence of a $-\epsilon T$ term which serves as the bifurcation parameter. The $1 + \epsilon$ expansion can be summed completely because the flow equation is known to all orders. It does not do a great job with critical exponents though it gets the value of the critical temperature exactly right in 2-d (because it respects duality symmetry). Several people have tried to improve the expansion [107, 34].

The presence of non-analytic terms in the flow equations complicates the application of normal form theory. We will come back to it when discussing the Legendre transform of flow equations.

3.8 2-d Ising model

The 2-d Ising model is a particularly nice example because it has an exact solution in the absence of a magnetic field. All predictions then can be compared to the exact solution. Here, we will compare the exact solution to results from normal form theory. We will also consider how Legendre transforms affect the flow.

The only variable required to describe the 2-d Ising model in the absence of a field is the temperature t. The linear eigenvalues of the free energy and the temperature are 2 and 1 respectively. The normal form of the flow equations can be written as

$$\frac{d\tilde{f}}{d\ell} = 2\tilde{f} - \tilde{t}^2, \tag{3.78}$$

$$\frac{d\tilde{t}}{d\ell} = \tilde{t}. \tag{3.79}$$

We have used the fact that the only term which cannot be removed by traditional normal form analysis is the resonance t^2. In fact, it cannot be removed by any analytic change of variables. We have also used the freedom to rescale t to set the coefficient of the resonance equal to -1 [7]. The solution to this can be written as $\tilde{t} = \tilde{t}_0 e^{\ell}$ and the free energy

$$\tilde{f}(\tilde{t}_0, \ell) = e^{-2\ell}\tilde{f}(\tilde{t}_0 e^{\ell}) - \tilde{t}_0^2 \ell. \tag{3.80}$$

Coarse graining until $\tilde{t}(\ell) = 1$ or $l = -\log(\tilde{t}_0)$, we get

$$\tilde{f}(\tilde{t}_0) = \tilde{t}_0^2 f(1) + \tilde{t}_0^2 \log \tilde{t}_0. \tag{3.81}$$

Now, the normal form variable $\tilde{t}_0$ is some analytic function of the physical variable t_0. It is linear to first order in t_0. Hence, we can write it as $\tilde{t}_0 = t_0(1 + c(t_0))$ where c is some analytic function. Then, we can expand

$$\begin{aligned}
\tilde{f}(t_0) &= t_0^2(1 + c(t_0))^2 f(1) + t_0^2(1 + c(t_0))^2 \log t_0(1 + c(t_0)), \\
&= t_0^2(1 + c(t_0))^2 f(1) + t_0^2(1 + c(t_0))^2 \log(1 + c(t_0)) + t_0^2(1 + c(t_0))^2 \log t_0, \\
&= a(t_0) + b(t_0) \log t_0,
\end{aligned} \tag{3.82}$$

[7] The sign is set to match the exact solution of the square lattice nearest neighbour Ising model

where both $a(t_0)$ and $b(t_0)$ are some analytic functions of t_0. Meanwhile any change of coordinates which adds an analytic function of t_0 to $\tilde{f}$ can be absorbed in the definition of $a(t_0)$. Hence, we can write the final most general form of the free energy of the 2-d Ising model as $f = a(t_0) + b(t_0) \log t_0$. This is quite surprising, as discussed in forthcoming work [53]–all corrections can be absorbed by changes of variables. Indeed, the exact solution of the 2-d Ising model can be written in this form [42].

While the basic solution of the 2-d Ising model is simple, it allows us an opportunity to explore other aspects of the flow equations. We explore the effect of Legendre transforms and irrelevant variables in the following sections.

3.8.1 Flow equations for results

The results of the RG, like the magnetization and the susceptibility, are given by derivatives of the free energy. One can derive the equations governing the flow of the results. This is important because eventually we want to be able to treat results and control variables on a more equal footing. Consider the usual parametrization θ^μ and an equation for the free energy of the form

$$\frac{df}{d\ell} = df + \Pi \tag{3.83}$$

where d is the dimension of space and Π is a possible inhomogeneous term in the free energy which depends only on the parameters θ^μ. The beta functions are defined as $\frac{d\theta^\mu}{d\ell} = \beta^\mu$. Since the free energy is a function of all these parameters, we can rewrite this equation as

$$\beta^\gamma \partial_\gamma f = df + \Pi. \tag{3.84}$$

Taking one derivative on one side gives

$$\partial_\alpha(\beta^\gamma\partial_\gamma f) = d\partial_\alpha f + \partial_\alpha\Pi, \tag{3.85}$$

$$\partial_\alpha\beta^\gamma\partial_\gamma f + \beta^\gamma\partial_\alpha\partial_\gamma f = d\partial_\alpha f + \partial_\alpha\Pi. \tag{3.86}$$

We will use the notation $\partial_\alpha f = M_\alpha$. Then, after commuting partial derivatives in the second term on the left hand side, we can write this equation as

$$\frac{dM_\alpha}{d\ell} = d(M_\alpha - \partial_\alpha\beta^\gamma M_\gamma) + \partial_\alpha\Pi. \tag{3.87}$$

The equations for the first derivative are closed and don't require information about the free energy. For the second derivative, we use the notation, $\partial_\mu\partial_\nu f = \chi_{\mu\nu}$. Then taking one more derivative as before gives

$$\partial_\mu\partial_\nu\beta^\gamma + \partial_\nu\beta^\gamma\partial_\mu\partial_\gamma f + \partial_\mu\beta^\gamma\partial_\nu\partial_\gamma f + \beta^\gamma\partial_\mu\partial_\nu\partial_\gamma f = d\partial_\mu\partial_\nu f + \partial_\mu\partial_\nu f. \tag{3.88}$$

Rearranging and commuting partial derivatives gives

$$\frac{d\chi_{\mu\nu}}{d\ell} = d\chi_{\mu\nu} - (\partial_\nu\beta^\gamma\chi_{\mu\gamma} + \partial_\mu\beta^\gamma\chi_{\nu\gamma}) - \partial_\mu\partial_\nu\beta^\gamma M_\gamma + \partial_\mu\partial_\nu\Pi. \tag{3.89}$$

Evidently the equations for the second derivatives depend on the equations for the first derivative if the beta functions are nonlinear. If the beta functions are linear and of the form $\beta^\mu = \lambda^{(\mu)}\theta^\mu$ (brackets mean no summation convention), then both equations reduce to a much simpler and recognizable form (apart from the inhomogeneous term)

$$\frac{dM_\alpha}{d\ell} = (d - \lambda^{(\alpha)})M_\alpha + \partial_\alpha\Pi, \tag{3.90}$$

$$\frac{d\chi_{\mu\nu}}{d\ell} = (d - \lambda^{(\mu)} - \lambda^{(\nu)})\chi_{\mu\nu} + \partial_\mu\partial_\nu\Pi. \tag{3.91}$$

We can use this to derive the equation for the specific heat C in two dimension which in this notation is the equation for χ_{tt} with $\Pi = -t^2$. It then follows,

$$\frac{dC}{d\ell} = -2 \tag{3.92}$$

The inverse specific heat is

$$\frac{dC^{-1}}{d\ell} = 2C^{-2} \tag{3.93}$$

and has a transcritical bifurcation in two dimensions. This raises a question, is it legitimate to talk about a bifurcation in two dimensions for the Ising model if it happens in the space of results rather than the space of control variables? Intriguingly, though perhaps unrelated, a bifurcation has been observed in 2 dimensions using methods of conformal bootstrap [53]. A natural framework which interchanges between results and control parameters is given by Legendre transforms.

3.8.2 Legendre Transforms

The variables f and t are a particular choice of thermodynamic coordinates. Thermodynamics has a natural invariance under Legendre Transforms. Here, we consider the effect of Legendre transforms on the RG flow of the 2-d Ising model. We consider the Legendre transform

$$S = f - t\partial_t f, \tag{3.94}$$

$$E = \partial_t f. \tag{3.95}$$

Note that the free energy is usually defined as $-T \log Z$ and we have been exploring the dependence on T since we are usually interested in the behaviour close to the critical point. We will continue to ignore this dependence when taking derivatives with respect to t. It is not much more difficult to keep the dependence explicitly, it changes none of the essential conclusions of this section. We want to derive the flow equations for $dE/d\ell$ and $dS/d\ell$. It is straightforward

to derive

$$\frac{dS}{d\ell} = 2S - t^2, \tag{3.96}$$

$$\frac{dE}{d\ell} = E - 2t. \tag{3.97}$$

Naturally, we want the flow equations to only depend on E and hence substitute the function $t(E)$. Te free energy f is a non-analytic function of t (since it has a term proportional to $t^2 \log t$). Hence, E defined as one derivative of the free energy is also a non-analytic function of t (with a term proportional to $t \log t$). Inverting this function gives another non analytic function $t(E)$. Hence, the flow equations for S and E are non-analytic in E.

This has been explicitly verified in our group for the case of the 2-d Ising model [53]. In fact, the problem runs even deeper than this. Even for a simple hyperbolic fixed point, adding analytic corrections (any nonlinear terms in the flow equations) leads to a non-analyticity in the equation for the Legendre transformed coordinates.

We are forced to conjecture that t (and h etc.) are uniquely specified as the correct control variables for RG. It is possible that it is more natural to consider removing degrees of freedom in the canonical ensemble (t and f), then in a microcanonical one (E and S) [8].

[8] In fact, there is an interesting connection here with information geometry. It is much more natural to talk about the Fisher Information metric in the canonical ensemble. To be able to talk about the uncertainty in a thermodynamic quantity, one has to be able to exchange that quantity with the environment. Otherwise, calculating the Fisher Information Metric can give ill-defined answers. This is further motivation to consider that a particular thermodynamic ensemble may be more suitable for some purposes.

3.8.3 Irrelevant variables

Irrelevant variables are the source of singular corrections to scaling. Despite the exact solution available for the 2-d Ising model, its scaling form has still not been completely resolved. The scaling form which has been conjectured before [2, 42] naturally follows from normal form theory. An interesting unresolved issue is the presence of higher powers of logarithms in the susceptibility which are not found in the free energy [119, 43]. This is usually attributed to the presence of irrelevant variables. Here we show than the irrelevant variables which are derived from conformal field theory [42] would in fact lead to higher powers of logarithms in the free energy which are not observed. Hence, they cannot explain the higher powers of logarithms in the susceptibility. It is possible that there are other irrelevant variables in the 2-d square lattice nearest neighbour Ising model with a field which are not predicted by conformal field theory but can capture the higher powers of logarithms in the susceptibility, as they turn on with a field.

Below, we keep our analysis somewhat general, but the direct application is to the 2-d Ising model. We want to examine the scaling form in the presence of irrelevant variables in the case where it is possible to have a resonance between a relevant and irrelevant variable. Let us say we have two relevant variables t and h with positive eigenvalues λ_t, λ_h and one irrelevant variables u with a negative eigenvalue λ_u. We assume one of the relevant variables is related to the irrelevant one by the equation

$$p\lambda_t + q\lambda_u = 0, \tag{3.98}$$

for some positive integers p and q which are the smallest integers for which the equation is satisfied. We are trying to find the hyper-normal form of the equa-

tions. First, assume that we have removed all other terms by regular normal-form theory. There are still an infinite number of resonances in this theory, all of which satisfy the equation

$$\lambda_x + n(p\lambda_t + q\lambda_u) = \lambda_x \tag{3.99}$$

where x can be any variables and n can be any positive integer. To see which one of the terms can be removed by hyper-normal form theory, consider defining the variable $w = t^p u^q$. This gives an efficient way to write the resonances. So, we would expect the equations to be of the form

$$\frac{dt}{d\ell} = \lambda_t t + a_1 tw + a_2 tw^2 + a_3 tw^3 + ..., \tag{3.100}$$

$$\frac{du}{d\ell} = \lambda_u u + b_1 uw + b_2 uw^2 + ..., \tag{3.101}$$

$$\frac{dh}{d\ell} = \lambda_h h + c_1 hw + c_2 hw^2 + ..., \tag{3.102}$$

As we show in Appendix C, the normal form of this equation can be written as

$$\frac{dt}{d\ell} = \lambda_t t + a_1 tw + a_2 tw^2, \tag{3.103}$$

$$\frac{du}{d\ell} = \lambda_u u + b_1 uw, \tag{3.104}$$

$$\frac{dh}{d\ell} = \lambda_h h + c_1 hw. \tag{3.105}$$

There is some freedom in choosing the normal form. In particular, we can choose to keep the w^2 term in the u equation and remove it from the t equation.

Let us go back to the original flow equations and consider the effects of an irrelevant variable u to the scaling form. The most significant irrelevant variable in the 2-d Ising model has eigenvalue -2. So we want to solve the set of equations

$$\frac{df}{d\ell} = 2f - t^2, \tag{3.106}$$

$$\frac{dt}{d\ell} = t, \tag{3.107}$$

$$\frac{du}{d\ell} = -2u. \tag{3.108}$$

Before we move on to the full normal form, let us first calculate what order the corrections appear in with these equations. We can solve for the free energy

$$f(t_0, u_0, \ell) = e^{-2l} f(t_0 e^{\ell}, u_0 e^{-2\ell}) - t_0^2 \ell. \tag{3.109}$$

Coarse graining until $t(l) = 1$ or $l = -\log t_0$, we get

$$f(t_0, u_0) = t_0^2 \mathcal{F}(u_0 t_0^2) + t_0^2 \log t_0. \tag{3.110}$$

The function $\mathcal{F}$ is analytic and so has an expansion $\mathcal{F}(x) = a_0 + a_1 x + ...$ So

$$f(t_0, u_0) = t_0^2 (a_0 + a_1 t_0^2 u_0 + ...) + t_0^2 \log t_0. \tag{3.111}$$

That is, the irrelevant variable should add a correction of order t_0^4 and leave the logarithm unchanged. So, the irrelevant variable should add what looks like an analytic correction at order t_0^4. It is to be noticed that this means it is impossible to distinguish between what is technically a singular correction coming from an irrelevant variable and an analytic correction coming from a definition of t purely from the form of the free energy.

For now, the combination $u_0 t_0^2$ has eigenvalue 0 and should add resonances to the flow equations. Remember that the best way to handle the resonances is to define the variable $g = ut^2$ and find its flow equations. The normal form as shown above will be

$$\frac{df}{d\ell} = 2f - t^2, \tag{3.112}$$

$$\frac{dt}{d\ell} = t + agt, \tag{3.113}$$

$$\frac{du}{d\ell} = -2u + bgu + dg^2 u. \tag{3.114}$$

There is some freedom in choosing the normal form. In particular, we can remove the g^2 term in the u equations at the cost of adding a g^2 term in the t

equations. However, it makes sense to keep the t equation as simple as possible. The equations for t and u give an equation for g

$$\frac{dg}{d\ell} = Bg^2 + Dg^3, \tag{3.115}$$

where $B = 2a + b$ and $D = d$. It is easier to work with the variable g instead of working with u. Since we plan to coarse grain until $t = 1$, it is also useful that $g(t = 1) = u(t = 1)$ and so the two can be used interchangeably after coarse graining. It is possible to solve these full set of equations. However, the solution is cumbersome and hard to interpret. Let us first work with a much easier example with $a = d = 0$. This still has the main features of the corrections the equations generate but is much easier to interpret. We will then solve the full set of equations. So, first

$$\frac{df}{d\ell} = 2f - t^2, \tag{3.116}$$

$$\frac{dt}{d\ell} = t, \tag{3.117}$$

$$\frac{dg}{d\ell} = Bg^2. \tag{3.118}$$

Then the solution for $t(\ell) = t_0 e^{\ell}$ and $g(\ell) = g_0/(1 + Blg_0)$. So, coarse graining until $t(\ell) = 1$, we get $g(t = 1) = g_0/(1 - Bg_0 \log t_0)$. Finally the free energy

$$f_0(t_0, u_0) = t_0^2 \mathcal{F}(g_0/(1 - Bg_0 \log t_0)) + t_0^2 \log t_0. \tag{3.119}$$

Expanding the scaling function $\mathcal{F}(x) = a_0 + a_1 x + ...$ gives

$$f_0 = t_0^2(a_0 + a_1 g_0 + a_1 B g_0^2 \log t_0 + a_2 g_0^2 + O(g_0^3)) + t_0^2 \log t_0, \tag{3.120}$$

or in the original variables

$$f_0 = a_0 t_0^2 + a_1 t_0^4 u_0 + t_0^6 u_0(a_1 \log t_0 + a_2) + t_0^2 \log t_0 + O(t_0^8). \tag{3.121}$$

So a resonance should add a $\log t_0$ to the 6th order term. However, this is indistinguishable from an analytic correction at 6th order to t_0. So the first sign

of a resonance which is genuinely incompatible with the 'true' from of the free energy is of the form $t_0^8(\log t_0)^2$. The 2-d Ising model has no such term in the free energy. It presents an interesting challenge to see how an irrelevant variable would lead to a log in the susceptibility but not in the free energy or magnetization. The basic problem with finding a resonance with an irrelevant variable which could lead to such behaviour can be seen in the solution of $g(l)$ above. One could conjecture than an irrelevant variable exists which has a resonance with h instead of t. The amplitude of the log would then involve h_0 and go to 0 when h_0 goes to 0. Any irrelevant variable which has a resonance with t would lead to powers of log in the free energy. Neither of these two cases are consistent with the observed behaviour of the 2-d Ising model. Explaining the behaviour of the susceptibility with a scaling analysis is thus, so far, an unresolved problem.

We can show that the behaviour of the full normal form is similar though the algebra is considerably more complicated. We start with the equations

$$\frac{dt}{d\ell} = t + agt, \tag{3.122}$$

$$\frac{dg}{d\ell} = Bg^2 + Dg^3. \tag{3.123}$$

The nice thing is these are just the equations of a transcritical bifurcation. In fact, these equations are identical to those of the 4-d Ising model. The only differences in the calculations here are that the irrelevant variable is not dangerous, we will not rescale $B = 1$ since we want to check the solution for $B = 0$ and finally we are coarse graining until $t = 1$ rather than until $l = L$. First we solve g in terms of l. It helps to define an auxiliary variable $s = B/(Dg) + 1$. Then

$$g(\ell) = -\frac{B}{D(W(-s_0 e^{-s_0} e^{\ell B^2/D}) + 1)}. \tag{3.124}$$

By solving the equations for dg/dt, we can get

$$g(t=1) = \frac{B}{D(\tilde{a}W(1/(\tilde{a})s_0 e^{s_0/\tilde{a}} t_0^{B^2/(aB-D)}) - 1)}, \tag{3.125}$$

where $\tilde{a} = (aB/D - 1)$. A useful identity is that $W(xe^x y) = x + \log y$ for large x. Also note that g_0 small corresponds to s_0 large. Hence, for small g_0, expanding this solution gives

$$g(t=1) = \frac{g_0}{1 - B\log t_0 g_0}. \tag{3.126}$$

So we see an expansion will generate similar corrections to the ones in the toy example. However, the full solution will be of the form

$$f_0 = t_0^2(W(1/(\tilde{a})s_0 e^{s_0/\tilde{a}} t_0^{B^2/(aB-D)})/s_0)^{-a/B}\mathcal{F}(g(t=1)) - t_0^2 \int (W(-s_0 e^{-s_0} e^{\ell B^2/D})/s_0)^a dl. \tag{3.127}$$

A detailed expansion shows that the first sign of the resonance will now be $t_0^4 \log t_0$, and $t_0^6(\log t_0)^2$ with $\log\log t_0$ and other corrections. The 2-d Ising model has many irrelevant variables with integer eigenvalues. However, it is known that the exact solution does not have any powers of logs or any W functions. Either coincidentally all of the resonances have zero coefficients or there is a deeper underlying reason why the resonances don't show up in the scaling form of the free energy. This leads us to conjecture that the irrelevant variables predicted by conformal field theory are actually gauge variables which can be removed by coordinate changes. As we show in the next chapter, irrelevant variables are distinguished into those that contribute genuine singular corrections to scaling with new power law exponents and those that contribute no new power laws but can simply be written as linear combinations of already known exponents. Similarly, here the powers of log would be a genuine correction to scaling and the fact that they do not appear in the exact solution seems to be evidence that the irrelevant variables are in fact gauge variables. Further

evidence that this is given by the fact that the irrelevant variables predicted by conformal field theory in the 2-d Ising model are in fact 'descendant operators' given by acting derivatives on 'primary operators' (relevant variables).

CHAPTER 4

GAUGE VARIABLES AND PERIOD DOUBLING

> It is high time to abandon the mainstream and take to the turbulent waters of truly dynamic analysis.
>
> Joan Robinson

4.1 Introduction

There are several different ways to coarse grain a physical system. Coarse graining in momentum space leaves you at a fixed point which is different from a fixed point that coarse grains in real space. Any anisotropy due to the lattice won't vanish in the real space renormalization group (RG) whereas it will be washed out if the coarse graining is done in momentum space. Presumably, the momentum space and real space RG lead to different fixed points. So far, we have been considering corrections to scaling which leave the fixed point unchanged. An interesting question in the renormalization group (RG) is if the fixed point itself can be moved, and whether it is uniquely defined. If this is possible, then one can ask how to characterize the space of all RG fixed points. As a corollary, one can ask if any critical point can be made into an RG fixed point.

Here, we use the example of the period doubling transition to try and give definitive answers to these questions. As we will see, answering this question leads us naturally to a complete charecterization of the *corrections to scaling*. It is well known that the RG has corrections to scaling beyond the universal power law behaviour observed at critical points. However, actual calculations are usually difficult in the case of statistical systems like the Ising model. The period doubling transition offers a nice example which is simple enough that we can derive the complete form of the corrections to scaling but also sufficiently complicated to give non-trivial insights into those corrections. Our work here is part of a much larger effort to systematize corrections to scaling in critical phenomena by using normal form theory, which was covered in the last chapter. While we will focus on the period doubling transition to chaos, we will learn lessons

with applications to critical phenomena in general.

The renormalization group has been useful in describing systems outside of physics. The period doubling transition is a famous example of its application to dynamical systems theory. The form of the RG was first worked out by Feigenbaum [70] who showed how the behaviour of a a class of iterated maps had universal characteristics. Since then, this kind of analysis has been extended and applied to other maps [22, 131]. The archetypal example is that of the logistic map defined by

$$g(x) = 4\mu x(1 - x) \tag{4.1}$$

This map has one fixed point at 0 and another at x^* for small values of μ. As the value of the parameter μ is increased, a stable 2-cycle appears. As it is increased further, a stable 4-cycle appears and so on. The bifurcation diagram for the period doubling transition is showing in the Figure 4.1. There is a sequence of μ_n which give give a stable 2^n cycle. Eventually, this sequence converges to a point u_∞ where you get a transition to chaotic motion. The sequence of μ_n converge geometrically with $\Delta\mu^n \propto \delta^{-n}$, with δ being a universal constant. It is easier to calculate the scaling form of the superstable points which occur roughly midway between two bifurcation points. It is also possible to consider the scaling of the 'results' variable x, the distance between the 2^{n-1} cycle and the line $x = 0.5$ has a leading order behaviour $\Delta x_n \propto 1/\alpha^n$ (see Figure 4.1). The interesting thing about the bifurcation diagram is its universality and self-similarity. Any other one-humped map with a quadratic maximum shows the same sequence of bifurcations. It is conventional to translate the map so that the maximum is at the origin rather than at $x = 1/2$. For all of the RG calculations that follow, we will assume this has been done. The renormalization group description of this period doubling transition is as follows.

Consider the coarse graining operation $g(g(x))$ followed by a rescaling, so $\mathcal{T}[g(x)] = \alpha g(g(x/\alpha))$. This transformation (with α being a universal value) gives a function which looks very similar to the original $g(x)$. The fixed point of such a transformation occurs when

$$\mathcal{T}[g^*(x)] = g^*(x). \tag{4.2}$$

The fixed point is given by the whole function $g^*(x)$. The operator $\mathcal{T}$ can be linearized close to the fixed point. Its largest eigenvalue δ explains the leading scaling behaviour $\Delta\mu_n \propto 1/\delta^n$. Here, we will consider corrections to this result with an aim to answer some of the questions raised in the first paragraph. We will then move on to describing the complete form of the corrections to scaling and draw some conclusions for statistical mechanics.

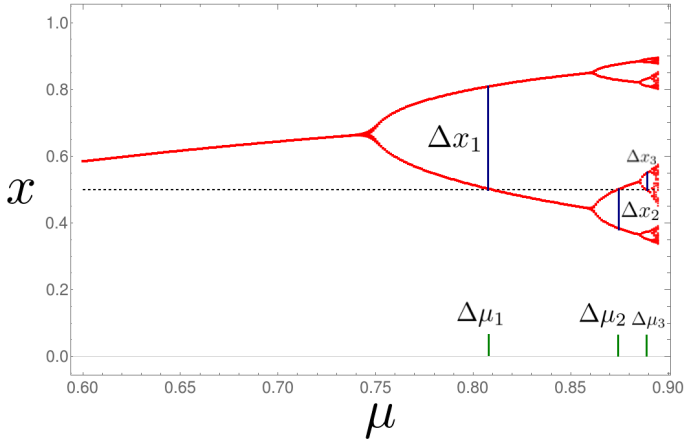

Figure 4.1: A figure of the bifurcation diagram of the period doubling transition. This is a plot of the fixed points and n cycles at different values of x as a function of the parameter μ of the logistic map given in the text. The intersection of the bifurcation diagram with the dashed line at $x = 0.5$ gives the super-stable orbits. There are two predictions in the RG, the spacing between subsequent superstable points given by $\Delta\mu_n$ and the vertical distance between the 2^{n-1} cycle and the line $x = 0.5$ for which we derive the scaling form. The vertical width of the lines in blue give the Δx_n and the horizontal coordinate of the green markers gives the values of $\Delta\mu_n$ (see main text for definitions).

4.2 Gauge corrections in period doubling

When considering analytic corrections to scaling in the RG, it is usual to consider making an analytic change in the parameter μ. Since μ is just a parameter of the specific function that we are considering, redefining it will not change the fixed point. Rather than redefining the horizontal coordinate in Figure 4.1, it is natural to also consider re-definitions of the vertical coordinate x. This is somewhat different because it is a redefinition of the degrees of freedom rather than parameters. We choose to call this a *gauge correction to scaling* as it involves a change in the way we measure (gauge) distance x.

Let the change of coordinate be $y = \phi(x)$. This induces a map $\tilde{g}(y) = \phi(g(\phi^{-1}(y)))$. Naturally, this leads to a new fixed point function $\tilde{g}^*(y) = \phi(g(\phi^{-1}(y)))$. This answers one of the questions we started with. It is possible to move the fixed point in the renormalization group. Correspondingly, it is possible to define a new renormalization group $\tilde{\mathcal{T}}$ whose fixed point is given by this new fixed point function. This is how changes of coordinates lead to new fixed points of the renormalization group. This situation has a parallel in statistical mechanics. Early results by Swendsen considered changing the form of the RG to change any critical point to a fixed point [147]. However, Fisher and Randeria [73] argued that the fixed point was uniquely determined, up to so called redundant variables, as the point that has no singular corrections to scaling. The singular corrections to scaling came from irrelevant variables with universal critical exponents. Given that we can choose ϕ arbitrarily, what is the space of equivalent fixed points that the RG for period doubling can take? We can answer this question by acting the old renormalization group on the new fixed point (see Ref. [71]), generating the full space of 'redundant' (gauge) vari-

ables.

$$\mathcal{T}[\tilde{g}^*](x) = \alpha(\phi(g^*(g^*(\phi^{-1}(\alpha^{-1}x))))) \tag{4.3}$$

To make progress, we consider an infinitesimal near identity change of coordinates of the form $\phi(x) = x+\epsilon\psi(x)$. The inverse transformation $\phi^{-1}(x) = x-\epsilon\psi(x)$. At the fixed point $\alpha(g^*(g^*(x/\alpha))) = g^*(x)$. Taking a derivative of this equation gives $g^{*\prime}(g^*(x/\alpha)))g^{*\prime}(x/\alpha) = g^{*\prime}(x)$. We can expand to linear order in ϵ

$$\begin{aligned}
\mathcal{T}[\tilde{g}^*](x) &= \alpha(\phi(g^*(g^*(\phi^{-1}(x/\alpha))))), && (4.4)\\
&= \alpha(\phi(g^*(g^*(x/\alpha - \epsilon\psi(x/\alpha))))), && (4.5)\\
&= \alpha((g^*(g^*(x/\alpha - \epsilon\psi(x/\alpha))))) + \epsilon\psi(g^*(x)/\alpha), && (4.6)\\
&= \alpha((g^*(g^*(x/\alpha - \epsilon\psi(x/\alpha)))) + \epsilon\alpha\psi(g^*(x)/\alpha)), && (4.7)\\
&= \alpha(g^*(g^*(x/\alpha)) - g^{*\prime}(x)\epsilon(\psi(x/\alpha)) + \epsilon\alpha(\psi(g^*(x)/\alpha))), && (4.8)\\
&= g^*(x) + \alpha\epsilon\left(\psi(g^*(x)/\alpha) - (\psi(x/\alpha))\, g^{*\prime}(x)\right). && (4.9)
\end{aligned}$$

Let $\psi(x)$ have a Taylor series $\psi(x) = \sum_p \psi_p x^p$. We then get

$$\mathcal{T}[\tilde{g}^*](x) - g^*(x) = \sum_p \alpha^{1-p}(g^*(x)^p - g^{*\prime}(x)x^p). \tag{4.10}$$

Hence the space of such equivalent fixed points have eigenvalues α^{1-p} with eigenfunctions given by $g^*(x)^p - g^{*\prime}(x)x^p$. The odd eigenvalues are given numerically by $\{1.00, 0.16, 0.0255, ...\}$ [1]. It is instructive to compare this with all of the

[1]The period doubling literature often restricts itself to an even subspace of functions which does not see half of the eigenvalues. This is why we are reporting only half of the eigenvalues here. In principle it is possible to include all of the eigenvalues and eigenfunctions. Feigenbaum had conjectured early on that all of the eigenvalues and eigenfunctions in period doubling could

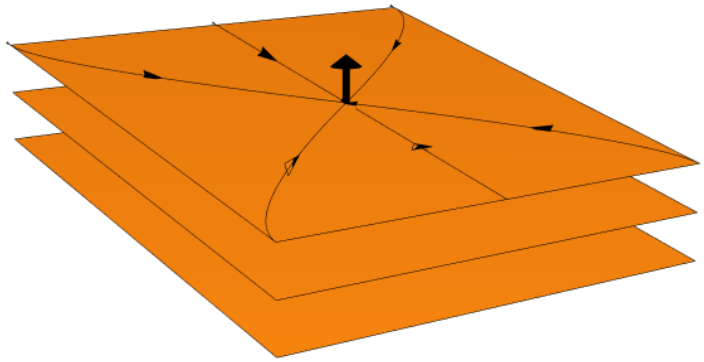

Figure 4.2: One gauge variable parametrizes a line of fixed points all of which can be reached by a change of coordinates. Each such fixed point has an associated critical manifold. Here, we are showing a series of critical manifolds each with their own fixed point and RG trajectories (trajectories are only shown on the top manifold here).

known eigenvalues and eigenfunctions near the fixed point which can be obtained using a systematic perturbation theory. The eigenvalues near the fixed point are given by $\{4.67, \underline{1.00}, \underline{0.160}, -0.124, -0.0573, \underline{0.0255}, ...\}$ [52]. Hence, some of the eigenvalues of the period doubling RG are simply by powers of α ('gauge' eigenvalues). Others are fundamentally new numbers which cannot be written down in terms of already known eigenvalues (singular irrelevant eigenvalues). The only relevant eigenvalues is the first one given by $\delta \approx 4.67$.

The equivalent of the coordinate changes we consider here in statistical mechanics were named redundant variables and explored in detail by Wegner [150]. They arise when considering a change of coordinates in the degree of freedom. So, for example, the cubic term in the Hamiltonian of the Ising model does not contribute to the scaling behaviour because it can be removed by a change of coordinates [39]. The statistical mechanics literature therefore usually ignores the effect of these variables on the scaling behaviour. Our analysis

be reached by changes of coordinates [71] (equivalent to Swendsen's assumptions); the failure of his conjecture is Randeria and Fisher's counter argument.

of period doubling suggests an alternative approach. Redundant variables are (usually) irrelevant variables which contribute to the corrections to scaling of the results of the RG. However, they do not lead to any new eigenvalues. Their 'gauge' eigenvalues can be written in terms of some linear combination (or in this discrete case, by some product) of already known eigenvalues. Other irrelevant variables that have fundamentally new eigenvalues contribute genuine *singular* corrections to scaling. Hence, some of the variables in the renormalization group are like a gauge choice. Having fixed a gauge, it contributes to the observed behaviour. Thus we discriminate between genuine *singular* corrections to scaling and *gauge* corrections to scaling, both of which come from irrelevant variables. As we saw in the previous chapter, this is important to understand the scaling behaviour of the 2-d Ising model. In fact the cubic term in the Hamiltonian of the Ising model is usually quoted to have an eigenvalues $d - \lambda_h$ where λ_h is the eigenvalue of the linear term, consistent with our classification.

In Wegners's work, he showed that the eigenvalues of redundant variables can be changed by choosing an appropriate RG transformation. In period doubling, doing this requires choosing the form of the rescaling to depend on the function being rescaled. This is possible to do in principle but seems to be an unnecessary complication in practice. We thus show the parallel with his work in Appendix E.

We now move on to deriving the full corrections to scaling for the RG of period doubling.

4.3 Corrections to scaling in period doubling

Corrections to scaling in period doubling have been considered before [106, 31, 55, 133]. While an ad-hoc form of the corrections was presented in Ref. [106], the corrections to scaling coming from the irrelevant eigenvalues within the linear RG was derived in Refs. [31, 133]. Here, we derive a more complete form of the corrections to scaling to compare how analytic, singular and gauge corrections appear in the physical predictions. The period doubling fixed point is hyperbolic and as we saw in the previous chapter, we can generally choose coordinates so that the RG is linear if we are at a hyperbolic fixed point [2]. Along with this, we set the gauge corrections to zero by choosing coordinates appropriately. We thus start on the submanifold with no gauge corrections to scaling and with *normal form* coordinates which linearize the RG flow.

We denote the linearization of $\mathcal{T}$ by $\mathcal{T}_L$. Let the eigenvalues and eigenfunctions of $\mathcal{T}_L$ be given by λ_p and Φ_p. The critical point is at the value of $\mu = \mu_\infty$. Let $\Delta\mu = \mu - \mu_\infty$. In our coordinates,

$$\mathcal{T}[g_\mu - g^*](x) = \Delta\mu \sum_p a_p \lambda_p \Phi_p(x) \tag{4.11}$$

Now, let us act with the operator n times, so

$$\mathcal{T}^n[g_\mu - g_{\mu\infty}](x) = \Delta\mu \sum_p a_p \lambda_p^n \Phi_p(x) \tag{4.12}$$

If g has a 2^n cycle, with $\mu = \mu_n$, then the application of $\mathcal{T}^n$ has a defined value at $x = 0$, so

$$\sum_p \Delta\mu_n a_p \lambda_p^n \Phi_p(0) = c, \tag{4.13}$$

where c is a constant. We note in passing that any difference between the fixed point function g^* and $g_{\mu\infty}$ which would lead to a constant term that we are not

[2] This is different from linearizing the RG at the fixed point.

considering here can be absorbed in the constant c. We redefine constants $\tilde{a}_p$ to absorb $\Phi_p(0)$ and c. This gives

$$\sum_p (\tilde{a}_p \Delta\mu_n)\lambda_p^n = 1. \tag{4.14}$$

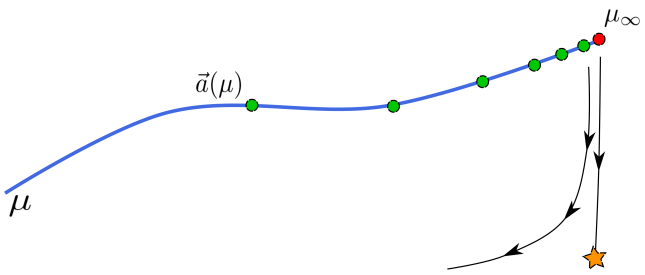

Figure 4.3: The blue line shows the function parameterized by μ with a fixed point at μ_∞ shown as a red circle. The n cycles are shown as green circles on the line. The vector $\vec{a}(\mu)$ gives the amplitude of the irrelevant eigenfunctions as a function of μ. The orange star is the fixed point in the space with redundant variables set to 0. The RG flows are shown in black.

Now, to include any corrections from the nonlinear regime of the RG, we assume that the coefficients $\tilde{a}_p$ are analytic functions of $\Delta\mu$ with a series expansion $\tilde{a}_p = \tilde{a}_p^{(0)} + \tilde{a}_p^{(1)}\Delta\mu_n + ...$ This is pictorially represented in Figure 4.3. Along the line parametrized by μ the function $g_\mu(x)$ has different amplitudes of the coefficients $a_p(\mu)$. So, the final expression now implicitly gives the scaling behaviour of $\Delta\mu_n$

$$\Delta\mu_n = \frac{1}{(\sum_p(\tilde{a}_p^{(0)} + \tilde{a}_p^{(1)}\Delta\mu_n + ...)\lambda_p^n)}. \tag{4.15}$$

The above equation can be solved order by order in $\Delta\mu$. The most useful result is directed at an experimentalist, what are the terms that give the corrections to scaling and how many independent coefficients are there to fit to the results? The expression in Equation 4.15 is best studied perturbatively. The lowest order

expression is

$$\Delta\mu_n = \frac{1}{(\tilde{a}_1^{(0)}\delta^n)}. \tag{4.16}$$

We use the freedom to rescale $\Delta\mu$ to set $\tilde{a}_1^{(0)} = 1$. We will organize terms by powers of δ^{-n}. Then the next order expression is

$$\Delta\mu_n = \frac{1}{\delta^n}\left(1 - \frac{\tilde{a}_1^{(1)}}{\delta^n} + \sum_{p\neq 1} \tilde{a}_p^{(0)}\lambda_p^n/\delta^n\right). \tag{4.17}$$

So far, the number of independent coefficients equal the number of new terms. At the next order, we get a much more complicated expression. The expression for $\Delta\mu_n$ is

$$\begin{aligned}\Delta\mu_n =& \frac{1}{\delta^n}\left(1 - \frac{1}{\delta^n}\left(\sum_{p\neq 1}\lambda_p^n a_p^{(0)} + a_1^{(1)}\right)\right) \\ &+ \frac{1}{\delta^{3n}}\left(\sum_{p,q\neq 1}\lambda_p^n\lambda_q^n a_q^{(0)}a_p^{(0)} + 3\sum_{p\neq 1}\lambda_p^n(a_p^{(0)}a_1^{(1)} - a_p^{(1)}) + (2a_1^{(1)^2} - a_0^{(2)})\right)\end{aligned} \tag{4.18}$$

As can be seen, corrections to this order add more terms than coefficients (e.g. if we kept only two of the irrelevant eigenvalues, it would lead to 6 new terms but only three new coefficients). The coefficients of the various terms have a somewhat complicated relationship between them. We expect this to be true at higher orders as well though we have not yet found a general expression for the corrections to scaling at arbitrary order. At each order, the corrections to scaling with the relationship between the various terms can be derived perturbatively. There are several things to notice in this expression for $\Delta\mu_n$. One, there are certain terms that go as $\frac{1}{\delta^{an}}$ for integer a. These are all analytic corrections to the relevant variable μ. Second, the correction to scaling does not involve the result in any way and so it should be independent of the gauge corrections to scaling. Changing coordinates in x should not affect the expression for $\Delta\mu_n$. Hence, powers of α should not be observed in the scaling. The corections due to the eigenvalue α^0 coincides with analytic corrections to scaling to μ and those

analytic corrections can still appear. This is exactly what was observed to lowest order in Ref. [31].

We can similarly derive a form for the corrections to the scaling of $g^{2^{n-1}}$ which we call Δx_n following Ref. [106] (see Figure 4.1). Asymptotically, these are just given by α^{-n}. To derive the corrections, we notice that $\alpha^n g^{2^{n-1}}(\alpha^{1-n}x)$ is the same as acting the operator $\mathcal{T}$ and so has a similar expansion

$$\alpha^n g^{2^{n-1}}(\alpha^{1-n}x) = g^*(x) + \Delta\mu_n \sum_p a_p \lambda_p^{n-1} \Phi_p(x). \tag{4.19}$$

Evaluating this at $x = 0$ gives

$$\Delta x_n = \alpha^{-n}(1 + \Delta\mu_n \sum_p a_p \lambda_p^{n-1} \Phi_p(0)). \tag{4.20}$$

Substituting the form of $\Delta\mu_n$ and using the series of expansion of a_p gives the full form of the corrections to scaling of Δx_n. In this case, the gauge corrections to scaling *should* affect the value of Δx_n and contribute to the observed behaviour as is indeed seen to lowest order in Ref. [31]. In Ref. [55], a change in scaling behaviour of Δx_n was seen under a change of coordinates though a RG explanation was not given. The interpretation becomes clear here, a change of coordinates will affect the gauge corrections to scaling of Δx_n and hence change the scaling behaviour.

Since the leading scaling behaviour of Δx_n is α^{-n} and all of the gauge eigenvalues are α^{1-p} for integer p, the gauge corrections can simply be generated as analytic corrections in Δx_n. This has a parallel in the 2-d Ising model where the corrections to scaling coming from irrelevant variables predicted by conformal field theory cannot be distinguished from analytic corrections. All of the conformal field theory predictions are for 'descendant operators' which are obtained by taking derivatives of primary (relevant) operators. These operators

have integer eigenvalues. Normal form theory suggests that they generically should lead to logarithmic powers which are not observed in the square lattice 2-d Ising model. Meanwhile, the leading genuine singular correction to scaling coming from an operator with eigenvalues $-4/3$ seems to have zero amplitude in the 2-d square lattice Ising model. Thus, Barma and Fisher [14, 15] had to use a double-Gaussian model to find evidence for a genuine singular correction. Our analysis here would suggest that the variables predicted by conformal field theory lead are contributing gauge corrections to scaling whereas the irrelevant variable with eigenvalue $-4/3$ that Barma and Fisher observe is a genuine singular correction to scaling.

4.4 Conclusion

We have examined some deep questions about the renormalization group in the context of period doubling. We showed that there is some freedom to move the fixed point of the RG associated with gauge transformations in the coordinates of the map. We have also derived the full form of the corrections to scaling of the period doubling transition. In doing so, we propose a strategy for simplifying the corrections to scaling in the RG. One first goes to the sub-manifold with no gauge correction to scaling and to normal form coordinates. Predictions of the RG which do not involve the traditional 'results' variables are unaffected by the gauge corrections. For the results however, the gauge coorections do contribute in a manner similar to but yet distinct from other universal singular corrections to scaling. The analysis of the RG in period doubling gives a new way to distinguish between genuine singular corrections to scaling and gauge corrections which can be removed by coordinate changes. Rather than changing

coordinates to get rid of the gauge corrections, they can be kept in the analysis and are distinguished by the fact that they lead to no new universal eigenvalues but still contribute to the corrections to scaling in the results.

We conjecture that the difference between such gauge eigenvalues, and the universal eigenvalues associated with the RG lies in the fact that these gauge eigenvalues are some combination of already known values of the RG. The 2-d Ising model is an interesting example where this conjecture can be tested. In period doubling, the degree of freedom x and the parameter μ are both one dimensional and corrections to scaling coming from changing variables in either of them are easily derived.

In statistical mechanics, the degree of freedom and parameters are in different spaces and difficult to compare. However, one usually only observes the derivatives of the free energy which are average values of the degrees of freedom and are the predictions or results of the RG. Any corrections to scaling are thought to be captured simply from analytic corrections made on the intensive variables like temperature and magnetic field. A more complete theory of the corrections to scaling in statistical mechanics would include analytic corrections to the extensive variables like free energy and its derivatives as well.

CHAPTER 5

FINDING STABILITY BOUNDARIES IN KICKED MAPS

Examples teach no less than rules.
And errors more than correct but
abstruse proofs.

V. I. Arnold

5.1 Introduction [1]

In this chapter, we move away from the study of the RG to the study of nonlinear maps. The study of the physics of nearly integrable systems has a rich and fascinating history. It has applications in fields varying from planetary science to accelerator physics. In both accelerators and planetary motion, the survival of particles under billions of revolutions under a nonlinear Hamiltonian is subtle; indeed, Hamiltonian chaos was first discovered [124] in the context of the three-body problem in planetary systems. The solution of the two-body problem is well known and a part of textbooks. Adding a third body makes an exact analytical solution impossible, one must resort to perturbation theory. The perturbation theory turns out to have chaotic resonances, when the frequency of the orbits are rationally related. These chaotic resonances have been thoroughly studied [8, 25, 132], and cause 'small denominator' problems [11] that prevent otherwise useful perturbative calculational techniques from converging. The motion of two bodies around the Sun, for example, lives on a torus with the two frequencies being around the two radii of the torus. The Kolmogorov-Arnold-Moser theorem (KAM theorem) proved that most torii are stable when the interaction between the two bodies is turned on. The word 'most' here is quantified by how irrational the ratio of the two frequencies of the bodies are [2]. Even though the KAM theorem was originally formulated in the context of planetary motion, it turned out to find application in the design of accelerators. An accelerating particle in a storage ring has magnets which lead

[1] Most of this chapter has been published on arXiv:1707.09336 with James P. Sethna, Sayan Choudhury, Amie Wilkinson and David Rubin

[2] Any irrational number has a continued fraction expansion which also quantifies how irrational it is

to its oscillatory motion in the plane perpendicular to the direction along the ring. There are three frequencies in the problem, the frequency of motion along the ring, and the frequency of motion in the horizontal and vertical direction on the perpendicular plane. If these frequencies are rationally related, it can lead to resonances and chaos which is bad for accelerator design. The motion around the ring in accelerators is time periodic and can be written as a map which gives the integrated dynamics of one time period. Figure 5.1 shows these resonances in a map (which we will define and use later in this chapter).

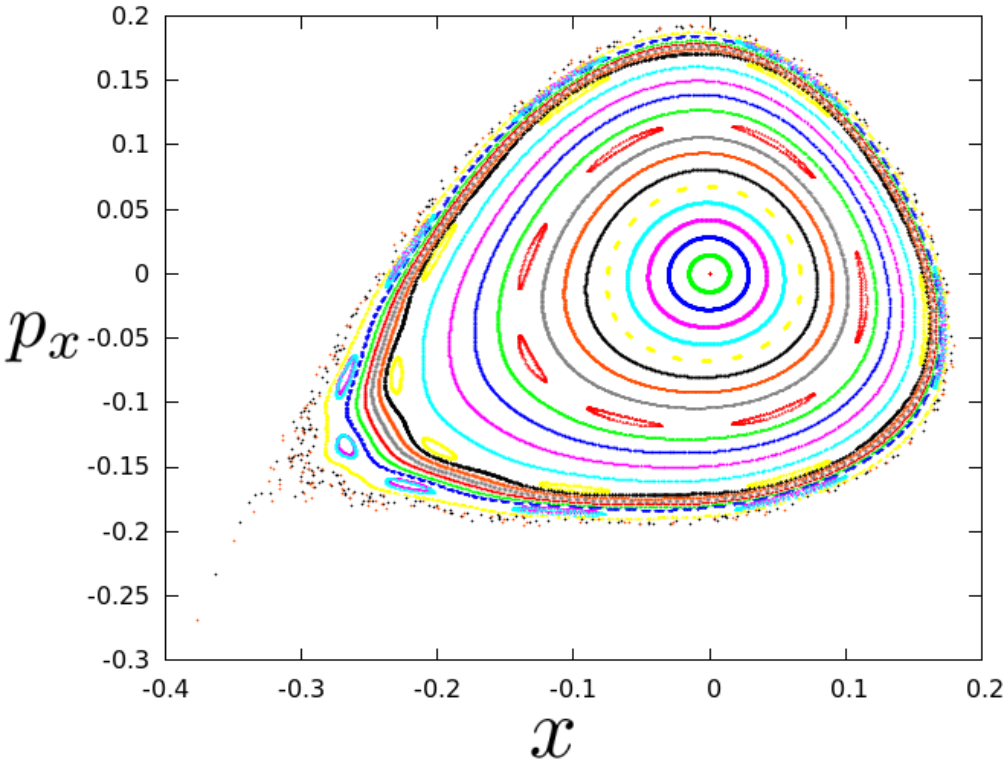

Figure 5.1: A Poincaré section of a map that we will use in this chapter. The red islands located in between smooth contours are resonances. These resonant regions are islands of stability with vivid structure (not visible here). A great amount of effort has been spent on studying and understanding these resonances. In this chapter, we will study how ignoring these resonances allows the construction of *effective integrals of motion* which can be used to predict practically stable regions.

Here, however, we shall investigate how ignoring the chaos – developing effective integrals of the motion – can be used to capture the behavior important to the design and optimization of particle accelerators, and more generally for time-periodic Hamiltonian systems with islands of long-term stability in phase space.

The resonances show up as vivid islands of stability in phase space, with a self-similar structure [3]. A time dependent Hamiltonian with one position and one momentum is technically a 3/2 degree of freedom (DoF) system, with time counting as a 1/2 degree of freedom. Hence the map in Figure 5.1 has 3/2 DoF and the motion is bounded by torii that are stable. A particle which starts anywhere in the stable region of phase space seen in the figure stays there if its dynamics is governed by the map. In 5/2 and 7/2 DoF systems however, particles can diffuse throughout the resonant chaotic regions on the energy surface in a process called Arnol'd diffusion. Hence the stable regions that we will calculate in this work are not currently mathematically well defined objects. Nevertheless, they are physically very useful not just for calculating regions that are practically stable (because particles take a long time to escape from them), but also, as we will show, to calculate other physical quantities like the escape rate, emittance etc. This raises an interesting mathematical question, why do the estimates we make of the stable region work so well, and how do we classify regions which are so nearly stable? We are inspired by the accuracy of estimates like the Stirling's formula which has a series with zero radius of convergence but is still immensely useful.

Fig. 5.2 provides an illustration of our results. In section 5.2 we introduce

[3]Interestingly, and as a tenuous connection to the previous parts of the thesis, it is possible to apply a renormalization group analysis to analyze these resonances

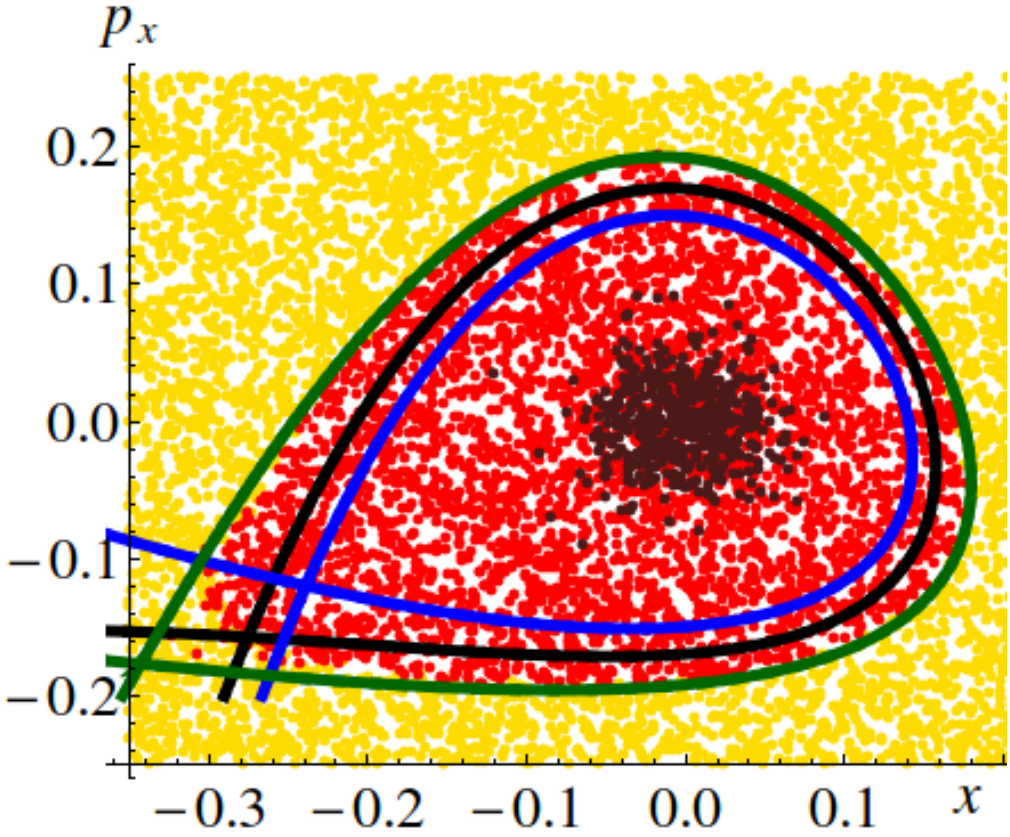

Figure 5.2: **Phase space regions** for a 1d map (one position, one momentum) describing orbits passing through a cross section in an accelerator. Yellow points escape to infinity (the walls of the chamber); red points are stable for infinite time. We use three methods, the normal form method (NF, blue), the Baker-Campbell-Hausdorff expansion (BCH, black), and a variational method (VAR, green) to estimate the *aperture* of stable orbits. The black points represent the *bunch* in phase space formed at long times when particles are subject to noise and damping. The phase-space extent of this bunch is the *emittance* which characterizes the brightness of the beam. Our methods can also estimate the rate of escape of bunch particles from the aperture (not shown).

traditional toy models for the dynamics of accelerators, and three strategies for calculating approximate invariants: the normal form method (NF), the Baker Campbell Hausdorff (BCH) expansion, and a variational method. In section 5.3 we use the three invariants (curves in Fig. 5.2) to approximate the 'aperture' – the region in phase space where lifetimes are effectively infinite. In section 5.4

we use our invariants to describe the equilibrium distribution of particles in a noisy environment, allowing the characterization of emittance (the phase-space volume of the bunch). In section 5.5 we generalize theories of chemical reaction rates to estimate the particle loss rate in the accelerator, and also provide a controlled, corrected form for the equilibrium distribution. These calculations, while useful in this context, are also generally applicable for kicked noisy maps. Our analysis in sections 5.4 and 5.5 is confined to a 1d map, but higher dimensional systems are discussed both to motivate our approximations and to outline how our methods could be generalized.

Our methods will bypass the chaotic resonances that have fascinated mathematicians and dynamical systems theorists in the last century. The aperture of stable orbits for maps of more than one dimension is mathematically a strange set, presumably with an open dense set of holes corresponding to chaotic resonances connected by Arnol'd diffusion [10]. The fact that accelerator designers characterize their apertures as simple sets motivates our use of invariants. Our approximate invariants ignore these holes, except insofar as resonances determine the outer boundary of stability. Designers avoid strong resonances, facilitating the use of our methods. Our focus, therefore, is on accurate estimates of the qualitative stability boundary, and on calculating the resulting emittance and escape rates in the presence of noise.

5.2 Map and Invariants

Particle orbits in accelerators are often well represented by a Poincaré recurrence maps. These maps usually describe nearly harmonic systems with nonlineari-

ties coming from sextupole and other higher order magnets. Most trajectories near the central 'reference orbit' are stable for infinite time (*i.e.*, live on KAM tori [92, 112, 9]); at farther distances where the nonlinearities are large orbits escape to hit the chamber walls. In accelerators it is found that the region of practically stable orbits is well described as a simple region called the 'dynamic aperture' (sometimes surrounded by a cycle of islands with similar properties). We shall refer to the stable region in phase-space as the 'aperture' in this paper.

In practice, this aperture is often found numerically by running the map for different initial conditions and seeing which particles escape. Here, we use two different kinds of perturbation theory, the normal form method (NF) and the Baker Campbell Hausdorff (BCH) expansion to try and estimate the aperture. We also use a variational method that improves on both of these methods. Our general strategy is to find one or more approximate invariants of the map and find its saddle points. The contour at one of the saddle points gives our approximation to the aperture.

As our toy example in 1-d, we will study a harmonic Hamiltonian with a kick,

$$H = \frac{p^2}{2m} + \frac{m\omega^2 x^2}{2} + \frac{Kx^3}{6} \sum_n \delta(t - n\tau). \tag{5.1}$$

Here, ω is the frequency of the particle (as it wiggles perpendicular to the direction of motion) and τ is the period between kicks. Here, the form of the kick models the action of sextupole magnets in particle accelerators [19]. Henceforth, we will set $m = 1$. The dynamics then corresponds to the classic Hénon map

$$x_{n+1} = x_n \cos\omega\tau + \frac{p_n}{\omega} \sin\omega\tau, \tag{5.2}$$

$$p_{n+1} = p_n \cos\omega\tau - \omega x_n \sin\omega\tau - \frac{K}{2} x_{n+1}^2. \tag{5.3}$$

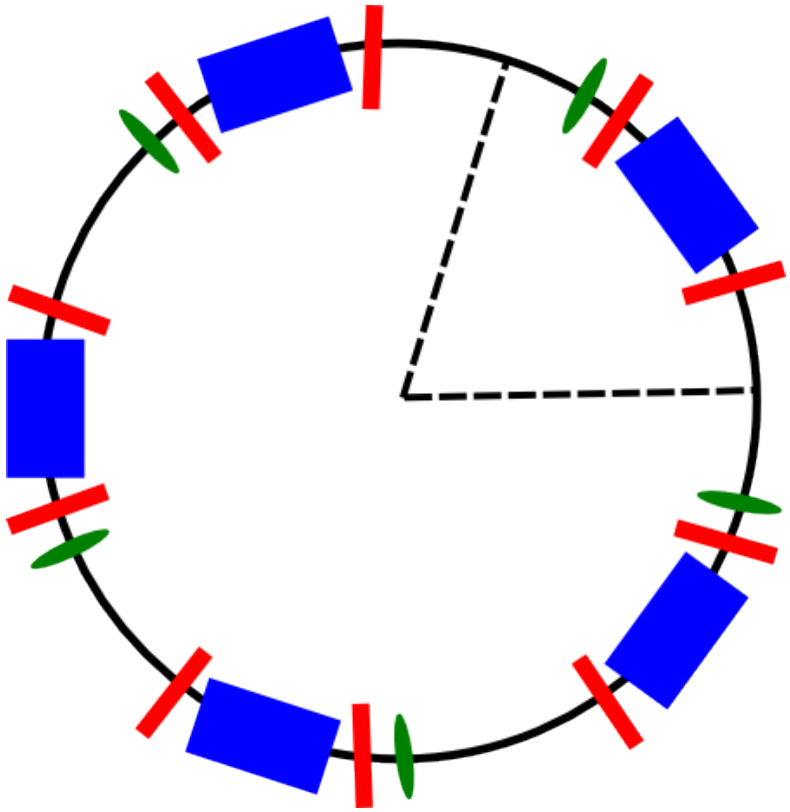

Figure 5.3: The figure above shows the toy accelerator ring that we model in this paper. The ring has *linear elements* which are the dipole (in blue) and quadrupole (in red). These can be represented as a harmonic oscillator. The nonlinear sextupole (in green) provides a kick at periodic intervals. Our calculations here work for a periodic array, and our map models the section between the dashed lines for a particle moving counter-clockwise. In any real accelerator, the sextupoles and other elements would have different strengths along the ring. Our methods would still work but the actual calculations would be messier.

Such kicked systems have become a paradigmatic example of chaos in both classical and quantum systems. In accelerators, one often only has access to the map and not to the original time-dependent Hamiltonian. We will formally denote the linear part of map without the kick by $\mathcal{M}$. We will denote the action of the kick, the non-linear part by $\mathcal{K}$. So, $\mathcal{M}(x_n) = x_{n+1}, \mathcal{K}(\mathcal{M}(p_n)) = p_{n+1}$. When acting on some function f of x_n and p_n, we have $\mathcal{M}(f(x_n, p_n)) = f(\mathcal{M}(x_n), \mathcal{M}(p_n))$.

We note that the aperture is not a simple region in any dimension greater

than one. Yet, we are inspired to use perturbation theories which give approximate invariants to capture simple regions which remain stable in practice. There is a large literature on constructing invariants for nonlinear systems. For symplectic maps, the NF [29, 113] is the most commonly used method. The NF gives approximate invariants up to a certain order in K but fails to capture the resonances. Resonant NF theory [18] can be used to include the effect of the resonances. Lie-algebraic techniques, which include the BCH expansion, can also be used to calculate invariants [63, 62, 41, 102]. Finally, people have tried numerical non-perturbative variational methods to capture the aperture by fitting polynomials or Fourier coefficients of generating functions to trajectories [27, 20, 88]. Many of these methods usually concentrate on getting the detailed structure of the aperture, including the islands, often at the cost of getting the boundary correctly [143].

Since accelerators are designed to avoid these large resonances, our focus is on getting accurate estimates of the stability boundary. As we will see later, this is particularly crucial to calculate the escape rate. It is known that both the NF and BCH lead to asymptotic series [21, 142]. Traditionally, the way to make sense of higher order terms in asymptotic series is by resumming them [81]. It is conceivable that detailed studies of a single resonance and the interaction between multiple resonances [49] could be used to create a resummation method which both captures the effect of the resonances and gives an accurate stability boundary.

While there are an infinite number of functions which can serve as approximate invariants (because any function of an invariant is invariant), one natural approach is to construct an effective Hamiltonian which also captures the dy-

namics of the system. Periodically driven systems are a class of time-dependent Hamiltonian systems for which an effective time-independent Hamiltonian (and consequently the aperture) can be obtained by an exact analytical formalism known as the Floquet formalism. The Floquet formalism is well known in classical accelerator physics [66] but the concept of a Floquet Hamiltonian is best understood using quantum mechanics. Thus in the spirit of Ref [35], we first obtain the effective Hamiltonian for the quantum version of the Hamiltonian given by Eq.(5.1) and then take the classical limit. The Floquet formalism involves calculating the evolution operator after n periods $U(nt)$ which is given by :

$$U(nT) = \mathcal{T} \exp\left(-i \int_0^{nT} H(t)/\hbar\right) = U(T)^n. \tag{5.4}$$

Thus, the evolution operator for 1 period is defined by :

$$U(T) = \mathcal{T} \exp\left(-i \int_0^{T} H(t)\right) = \exp(-iH_{\text{eff}}T/\hbar), \tag{5.5}$$

where H_{eff} is the effective Hamiltonian and $\mathcal{T}$ is a time-ordering operator. For the Hamiltonian in Eq.(5.1), this equation is particularly simple and we obtain :

$$U(T) = \exp\left(-i\frac{Kx^3}{6\hbar}\right)\exp\left(-i(\frac{p^2}{2} + \frac{\omega^2 x^2}{2})\tau\right). \tag{5.6}$$

The effective Hamiltonian we get using Floquet theory, we will call H_{BCH}. It is given by :

$$H_{\text{BCH}} = i\hbar \log\left(\exp\left(-i\frac{Kx^3}{6\hbar}\right)\exp\left(-i(\frac{p^2}{2} + \frac{\omega^2 x^2}{2})\tau\right)\right). \tag{5.7}$$

Now employing the Baker-Campbell-Hausdorff expansion and going to the classical problem in the usual way [142], we obtain the effective Hamiltonian (up to second order):

$$H_{\text{BCH}} = \frac{1}{2}\left(x^2\omega^2 + p^2\right) + \frac{K^2\tau x^4 - 2Kx\left(x^2\left(\tau^2\omega^2 - 4\right) - 6p\tau x - 2p^2\tau^2\right)}{48\tau} + O(\tau^3). \tag{5.8}$$

One major difference between the classical and quantum problems is that an effective Hamiltonian always exists for the quantum case, but the effective Hamiltonian description breaks down near resonances for the classical case. As has been argued in [142], for the quantum problem, the Baker-Campbell-Hausdorff expansion also breaks down near resonances, even though an effective Hamiltonian exists.

The more commonly used perturbation theory is the normal form method. In the context of canonical systems, this is called the Birkhoff Normal Form. Here, we do not construct canonical transformations which take the Hamiltonian to a normal form but instead directly construct (multiple) invariants of the map. The essence of the normal form method is to convert the problem of finding an invariant to a linear algebra problem on the space of homogeneous polynomials. This can be done by noticing that the action of $\mathcal{K}$ on any polynomial is to increase its order by 1. Let us start with the invariant of the linear part of the map $\mathcal{M}$ which is just the second-order polynomial $I_2 = \omega^2 x^2 + p^2$. Now, if we choose a third order polynomial I_3 so that the action of $\mathcal{M}$ on I_3 exactly cancels the action of $\mathcal{K}$ on I_2, we get an approximate invariant up to third order. We can continue this process order-by-order to get higher and higher order approximate invariants. The NF Hamiltonian to 3rd order is given by

$$H_{\mathrm{NF}} = \frac{1}{2}\left(p^2 + x^2\omega^2 + \frac{Kp^2 x\sin(\omega\tau)}{2\omega + 4\omega\cos(\omega\tau)} + \frac{1}{2}Kpx^2\right) \tag{5.9}$$

$$+ \frac{1}{2}\left(\frac{Kx^3\omega\tau\left(\left(\cos^3(\omega\tau)+1\right)\cot(\omega\tau) + \sin(\omega\tau)\cos^2(\omega\tau)\right)}{4\cos(\omega\tau)+2}\right) + \mathcal{O}(K^2). \tag{5.10}$$

The fourth order NF Hamiltonian loses the saddle point. All expansions have been truncated at the order which best describes the aperture (see sup-

plementary material for details). The NF and BCH Hamiltonians are both time-independent Hamiltonians trying to capture the one-period dynamics of the map. One method perturbs in the nonlinearity and the other perturbs in the inverse-frequency of the kick. If the series generated by perturbation theory were to converge, both would converge to the same effective Hamiltonian. However, because the series are asymptotic, they are most useful when truncated to a low order. The effectiveness of such low order truncations will depend on the particular problem. Finally, we can improve on the estimates of both of these perturbative methods numerically. One way to do this is to start from a point that lies on the aperture predicted by perturbation theory, and generate a trajectory using the map. Inspired by the form of the Hamiltonians obtained using perturbation theory, one can then simply fit a fourth order polynomial whose quadratic terms are constrained to be $p^2/2 + x^2\omega^2/2$ to the trajectory. The fit is generated by minimizing the variation of this polynomial over 10000 iterates of the map. The variational Hamiltonian we obtain for parameter values $\tau = 1$, $K = 6$, $\omega = 0.96$ is given by

$$H_{\mathrm{VAR}} = \frac{p^2}{2} + \frac{1}{2}x^2\omega^2 + a_1x^3 + a_3px^2 + a_4p^2x + b_1x^4 + b_3px^3 + b_5p^2x^2, \tag{5.11}$$

with best fit parameters $a_1 = 0.73, a_3 = 1.47, a_4 = 0.56, b_1 = 0.34, b_3 = 1.15, b_5 = 0.45$.

5.3 Aperture

In order to obtain the aperture from the Hamiltonian (or any invariant) we obtain its saddle point. This is given by simultaneously solving the equation $\partial H_{\text{eff}}/\partial x = 0$ and $\partial H_{\text{eff}}/\partial p = 0$. The energy contour corresponding to the saddle point gives the aperture. Examples of the aperture that we obtain for different parameters are shown in Figs 5.4–5.5. We show a comparison between the results of the BCH expansion, the NF and our numerical fit.

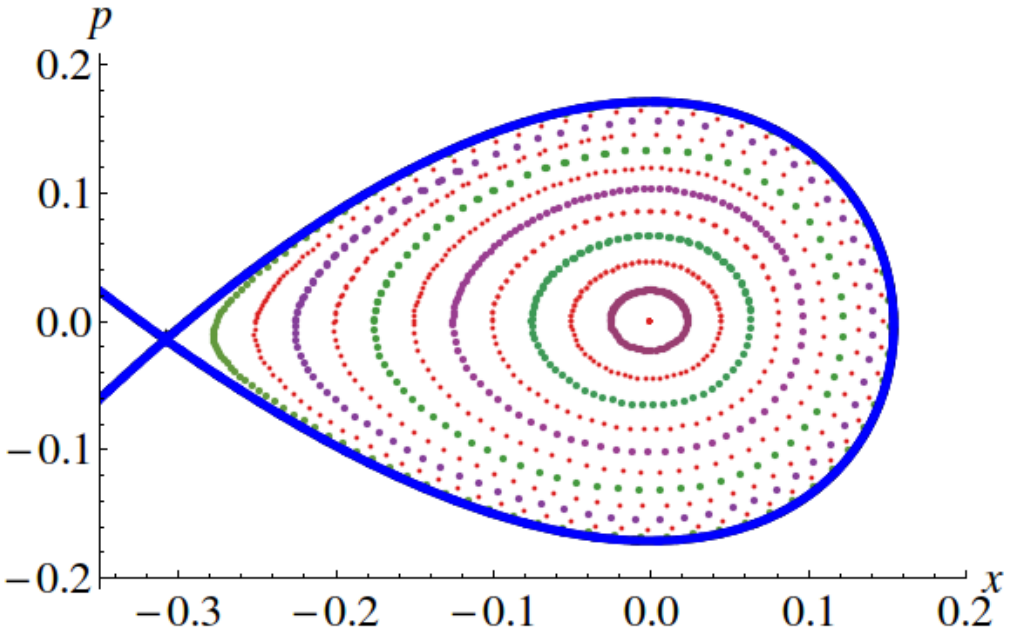

Figure 5.4: A plot of the aperture obtained using the NF(blue line), BCH (black line) and numerically (green line) on the Poincaré cross-section of the dynamics generated by the map. Parameters used here are $\tau = 0.1$, $K = 0.6$, $\omega = 0.96$. The results of the three are practically indistinguishable.

We generalize the map to two dimensions by adding a harmonic oscillator in the y variable and including a kick of the form $Kxy^2/2$ in the y momentum [75]. The form of the kick again is a sextupole, and is chosen to satisfy Laplace's equations. We can use the BCH expansion in exactly the same way to construct an effective Hamiltonian. The NF, on the other hand, gives multiple invariants

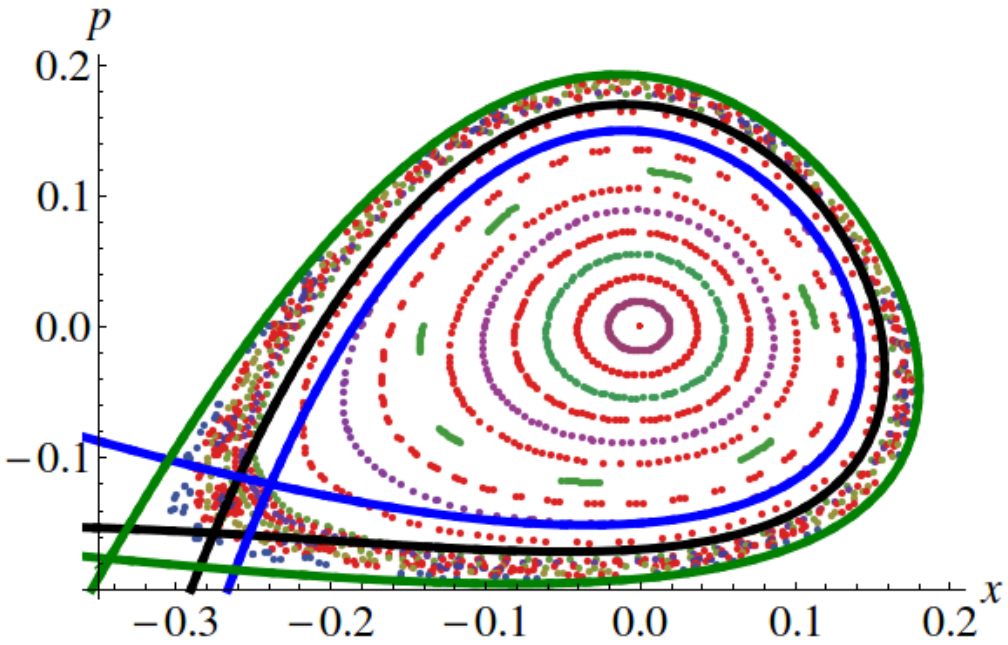

Figure 5.5: A plot of the aperture obtained using the NF(blue line), BCH (black line) and numerically (green line) on the Poincaré cross-section of the dynamics generated by the map. Parameters used here are $\tau = 1$, $K = 6$, $\omega = 0.96$

in higher dimensions. The aperture we obtain is shown in Figures 5.6– 5.7. In two dimensions, the NF gives two invariants for this map. The aperture then becomes a curve in the space of the two invariants. In Figure 5.8, we show a plot of the initial conditions that escape to infinity (in yellow) and those which stay bounded (in red) in the space of two approximate NF invariants. The curve which sets the boundary can be found by fixing a value of one of the invariants, and finding the constrained saddle point of the other. This problem can be solved using a Lagrange multiplier (see supplementary information) and the solution is shown as a dashed blue line in Figure 5.8. It is clear that this curve is not a very good approximation of the boundary.

As comparison, we also show the boundary that we get by simply adding the two invariants to get an approximate 'energy' shown using the solid blue line. Remarkably, the approximate 'energy' given by this linear combination of

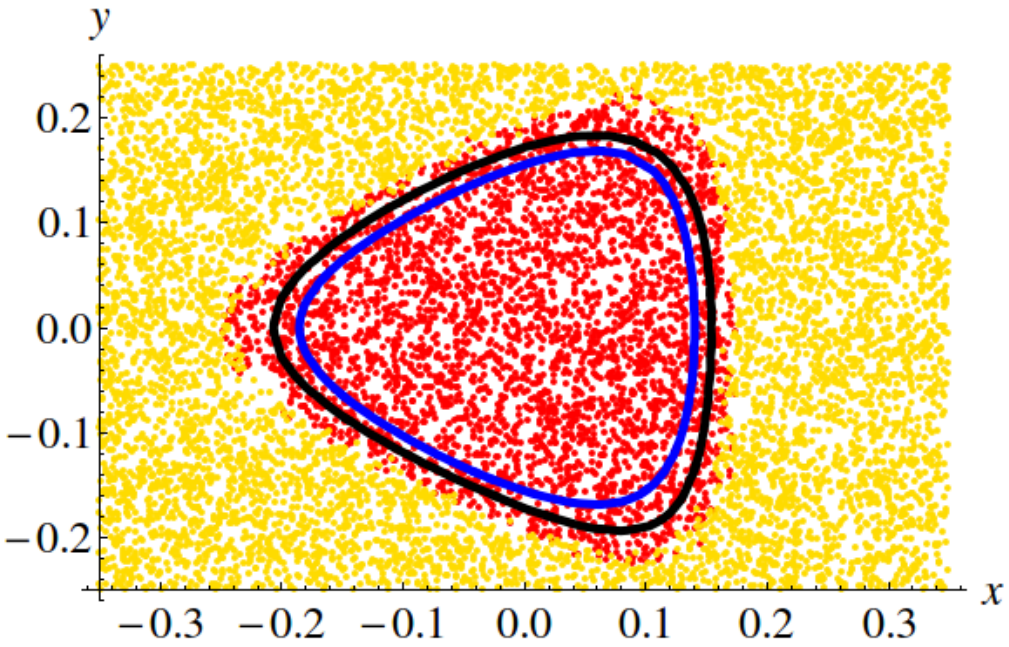

Figure 5.6: The 2-d map used is a generalization of the 1-d map and is given in the appendix. The yellow points are initial conditions which escape after a fixed number of turns, while the red points remain bounded. We use two perturbation theories, NF (in blue) and BCH (in black) to estimate the boundary between the two (see text). The figure shows two cross-sections of the map in $p_x - y$ plane.

the two invariants seems to represent the geometry of the problem better than the blue curve. It is possible that the aperture in higher dimensions is controlled only by the effective Hamiltonian obtained by simply adding the two invariants. Indeed, this is the combination we use to find the NF aperture in Figures 5.6– 5.7. This NF effective Hamiltonian is the analogue of our BCH effective Hamiltonian in two dimensions.

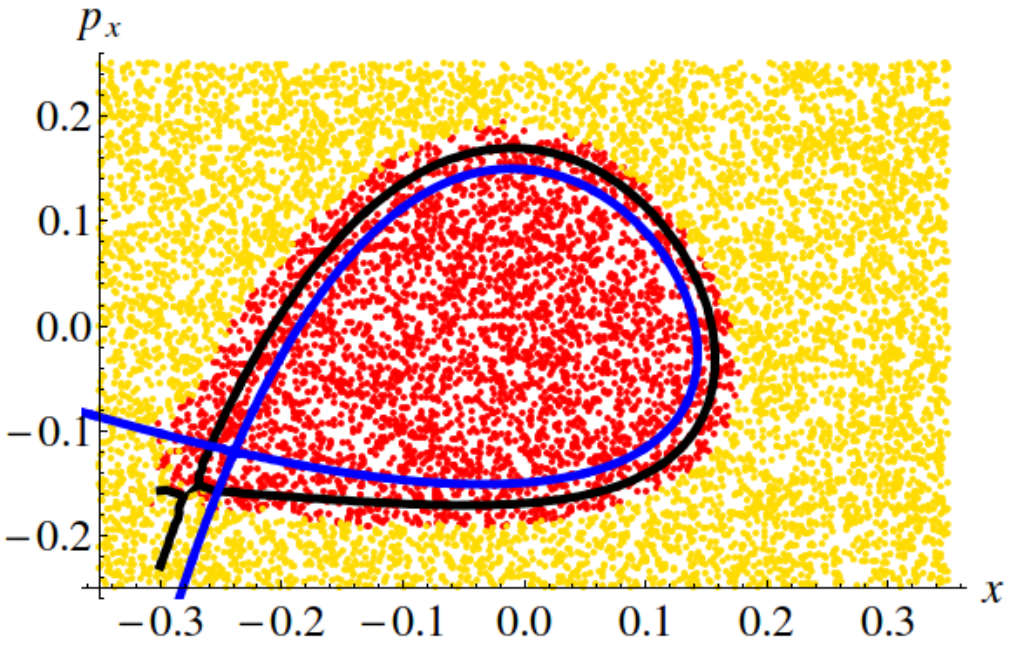

Figure 5.7: The 2-d map used is a generalization of the 1-d map and is given in the appendix. The yellow points are initial conditions which escape after a fixed number of turns, while the red points remain bounded. We use two perturbation theories, NF (in blue) and BCH (in black) to estimate the boundary between the two (see text). The figure shows two cross-sections of the map in $p_x - X$ plane.

5.4 Noise and Emittance

The effective Hamiltonian can be used to calculate the aperture but it is also useful to study the effect of noise on the dynamics. We are inspired to extend calculations done in the context of chemical reactions to describe particles escaping stability boundaries. There are many sources of noise in accelerators. These include residual gas scattering [110], photon shot noise [87] and intra-beam scattering [123]. Each of these have different forms but they all have the effect of changing the phase-space coordinates of the particle. We will only model the particle loss that occurs as a result of the particles crossing the barrier set by the dynamic aperture (other sources of particle loss exist in real accelerators). We

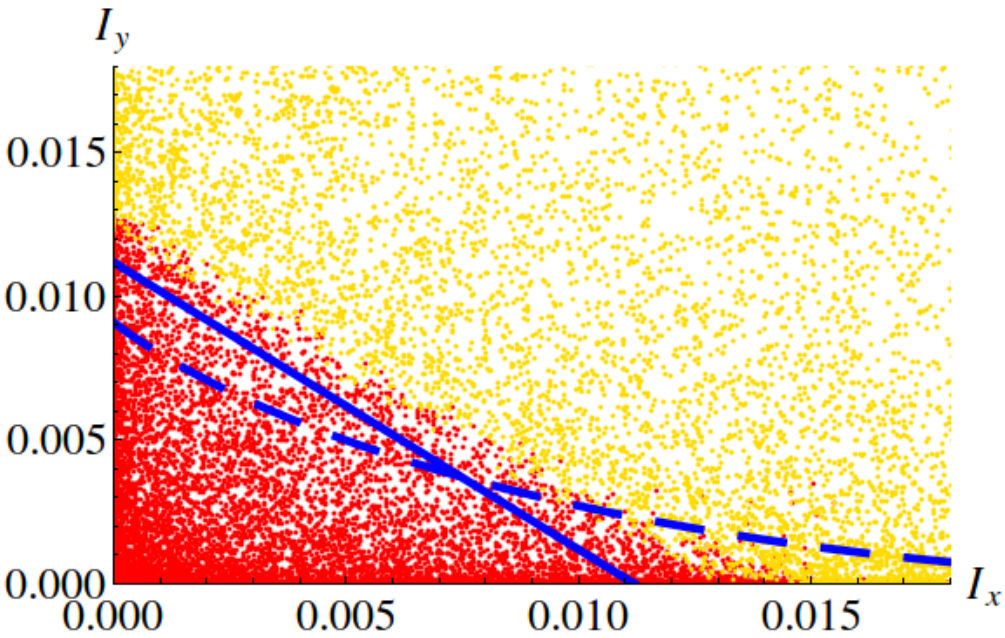

Figure 5.8: A plot of the two approximate invariants I_x and I_y obtained using the NF. Red points are initial conditions which stay bounded and yellow points are those which escape. The solid blue curve is the energy contour given by $I_x + I_y = c$ where c is the saddle point energy of $I_x + I_y$. The dashed blue curve is plotted by holding one invariant constant and finding the saddle of the other.

will model the noise phenomenologically with uncorrelated Gaussian noise and linear damping. This assumption has been used earlier to model noise in accelerators [159, 12]. A more realistic treatment of the noise would be multiplicative and could be added in principle though some parts of the calculation will then have to be done numerically. For time-independent Hamiltonian dynamics with a barrier, the effect of noise is well known at least since Kramer, who used the flux-over-population method to calculate the escape rate of particles both in the strongly damped and weakly damped case. There has been some work on the escape rate for maps in the strong damping regime [134]. The noise, whatever its form, cannot be directly added to the effective Hamiltonian that we calculate in the previous section and must instead be added to the original dynamics. We will do this here only for one dimension (one position and one momentum in

the map). For every time period, before the kick, the equations of motion then are

$$\dot{x} = p, \tag{5.12}$$

$$\dot{p} = -\omega^2 x - \gamma p + \xi(t). \tag{5.13}$$

We take the noise $\xi(t)$ to be delta-correlated Gaussian noise specified by its two-point function $\xi(t)\xi(t') = 2\gamma T\delta(t - t')$. Integrating the equations of motion over one time period and adding the kick gives the noisy map

$$x_{n+1} = e^{-\frac{\gamma\tau}{2}}\left(\frac{(2p_n + \gamma x_n)\sin(\tau\omega_r)}{2\omega_r} + x_n\cos(\tau\omega_r)\right) + \xi_{Xn}, \tag{5.14}$$

$$p_{n+1} = e^{-\frac{\gamma\tau}{2}}\left(p_n\left(\cos(\tau\omega_r) - \frac{\gamma\sin(\tau\omega_r)}{2\omega_r}\right) - \frac{x_n\left(\gamma^2 + 4\omega_r^2\right)\sin(\tau\omega_r)}{4\omega_r}\right) - K\frac{x_{n+1}^2}{2} + \xi_{Pn}, \tag{5.15}$$

where the integrated noise terms have zero mean and correlation functions

$$\langle\xi_{Pn}\xi_{Pm}\rangle = \frac{e^{-\gamma\tau}\left(\gamma^2\cos(2\tau\omega_r) - \gamma^2 + 2\gamma\omega_r\sin(2\tau\omega_r) - 4\omega_r^2\right) + 4\omega_r^2}{4\omega_r^2}T\delta_{nm}, \tag{5.16}$$

$$\langle\xi_{Xn}\xi_{Xm}\rangle = \tag{5.17}$$

$$\frac{4\omega_r^2 - e^{-\gamma\tau}\left(\gamma^2(-\cos(2\tau\omega_r)) + \gamma^2 + 2\gamma\omega_r\sin(2\tau\omega_r) + 4\omega_r^2\right)}{\omega_r^2\left(\gamma^3 + 4\gamma\omega_r^2\right)}\gamma T\delta_{nm},$$

$$\langle\xi_{Xn}\xi_{Pn}\rangle = \frac{e^{-\gamma\tau}\sin^2(\tau\omega_r)}{\omega_r^2}\gamma T\delta_{nm} \tag{5.18}$$

and $\omega_r^2 = \omega^2 - \frac{\gamma^2}{4}$.

We can use our effective Hamiltonian to calculate the spread of the particle bunch in phase space when noise is added to the map. As we show in Figure 5.9, a Boltzmann distribution with a temperature in the effective distribution does a good job of describing the equilibrium distribution of the particles near the center. This is only an approximation to the true equilibrium distribution which

we will calculate in the next section. The fact that accelerator designers characterize their bunches with effective temperatures for vertical and horizontal directions [141] is one motivation for the development of invariants that act as vertical and horizontal Hamiltonians, weakly coupled by noise in 2-d.

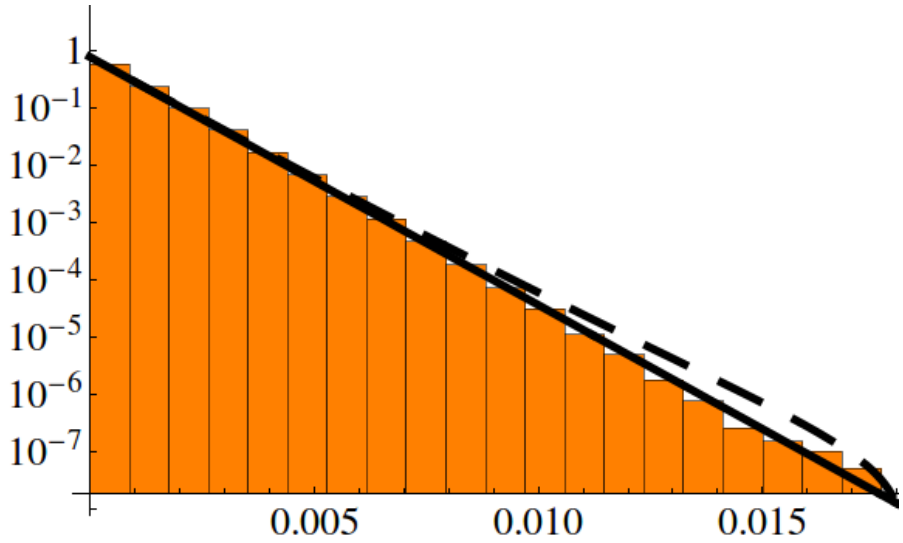

Figure 5.9: A histogram of the effective energy on a logarithmic scale of the particles which do not escape shows that a Boltzmann distribution in effective energy (solid line) given by our variational Hamiltonian serves as a good approximation to the equilibrium distribution. It is interesting to note that an improved estimate of the equilibrium distribution (dashed line) using Equation 5.31 actually does a worse job of capturing the numerical statistics. This might be because our variational Hamiltonian does not capture the dynamics (and resonances) accurately. Our escape rate calculations do not actually utilize the full form of the distribution because of the approximations we make. Parameter values used here in the simulation are $\tau = 1$, $K = 6$, $\omega = 0.96$, $\gamma = 0.005$, $T = 0.001$. We show the comparison to a typical harmonic approximation of the Hamiltonian in the supplementary material.

5.5 Steady State and Escape Rate

Equations (5.14-5.15) are completely general. To make progress, we make the assumption that we are in the weak damping regime (γ small or $1/\gamma$ large compared to all other time-scales in the system). This is usually a realistic assumption for storage rings [141]. Henceforth, we will work only to linear order in γ. We can then calculate the slow diffusion of the effective Hamiltonian under the noisy dynamics. This diffusion takes place on the Poincaré section. Hence, we've converted a non-equilibrium problem to an equilibrium problem on the Poincaré section. (The procedure fails in the resonant regions near where the frequencies are rationally related; transport across these resonances is dominated by chaos and escape rates from islands. Our numerics here are partially testing whether ignoring these resonances is valid.) Calculating the escape rate and the steady state probability distribution requires us to first know the drift and diffusion coefficient of the effective Hamiltonian. To find this, we change variables in the usual way [135]

$$D_E(x,p) = \sum_{\alpha=x,p} \frac{\partial H}{\partial \alpha} D_\alpha + \sum_{\alpha,\beta=x,p} \frac{\partial^2 H}{\partial \alpha \partial \beta} D_{\alpha\beta}, \tag{5.19}$$

$$D_{EE}(x,p) = \sum_{\alpha,\beta=x,p} \frac{\partial H}{\partial \alpha} \frac{\partial H}{\partial \beta} D_{\alpha\beta}. \tag{5.20}$$

The diffusion coefficients in x, p are defined using the correlation coefficients we calculated earlier. So, for example, $D_{xx} = \frac{\langle \xi_{X_n} \xi_{X_n} \rangle}{2\tau}$. There is a slight subtlety in making these change of variables because of the fact that our slow variable is γ. Using the notation developed earlier, we note that the difference in the

Hamiltonian evaluated after one period is given by

$$\begin{aligned}\Delta H =& H(\mathcal{M}(x_n) + \xi_{Xn}, \mathcal{K}(\mathcal{M}(p_n) + \xi_{P_n})) \\ & - H(x_n, p_n) \\ =& H(\mathcal{M}(x_n) + \xi_{Xn}, \\ & \mathcal{K}(\mathcal{M}(p_n)) + \xi_{P_n} - K\mathcal{M}(x_n)\xi_{Xn}) \\ & - H(x_n, p_n) \\ =& H(\mathcal{M}(x_n), \mathcal{K}(\mathcal{M}(p_n))) - H(x_n, p_n) \\ & + \frac{\partial H}{\partial x}\xi_{X_n} + \frac{\partial H}{\partial p}(\xi_{Pn} - K\xi_{Xn}\mathcal{M}(x_n)) \\ =& \frac{\partial H}{\partial x}\xi_{Xn} + \frac{\partial H}{\partial p}(\xi_{Pn} - K\xi_{Xn}\mathcal{M}(x_n)),\end{aligned} \tag{5.21}$$

where we have kept only linear terms in γ (and not written the second order contribution to the drift). We have also ignored the higher order terms of the perturbation theory in the effective Hamiltonian and assumed it to be a faithful characterization of the dynamics of the map. Note that the partial derivatives must be evaluated at the new points of the unkicked map and the fact that the kick takes place after the evolution requires us to be more careful about the noise in the p direction.

These drift and diffusion coefficients have to be averaged over the other canonical variable which acts as a time coordinate for the effective Hamiltonian. Calling this variable s, we then see that.

$$D_E(E) = \frac{1}{S}\oint D_E(x, p)ds, \tag{5.22}$$

$$= \frac{1}{S}\oint \frac{D_E(x, p)}{\frac{\partial H}{\partial p}}dx, \tag{5.23}$$

$$D_{EE}(E) = \frac{1}{S}\oint \frac{D_{EE}(x, p)}{\frac{\partial H}{\partial p}}dx, \tag{5.24}$$

where $S = \oint ds$. Even though an exact analytical expression for $D_{EE}(x, p)$ is easy

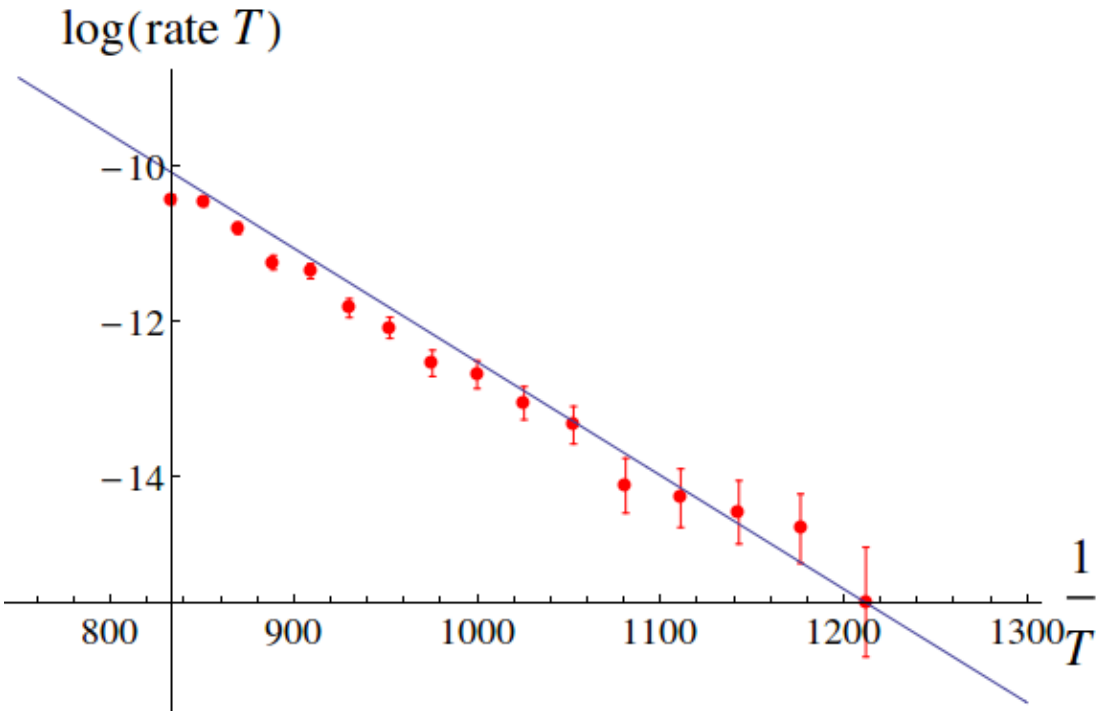

Figure 5.10: We compare our analytical predictions for the escape rates with numerical results. Because of the form of the rate given in Equation 5.33, we can plot $\log(kT)$ vs $1/T$ to get a straight line. Error bars are drawn from Poisson statistics. In case where perturbation theory works, with parameter values $\tau = 0.1, K = 0.6, \gamma = 0.005$, the effective Hamiltonian obtained from either perturbation theory does a good job of capturing the aperture

to calculate using standard computer-algebra software, these integrals have to be typically evaluated numerically. Having averaged over the fast variable, we can now make a stochastic differential equation using the prescription

$$\frac{dE}{dt} = D_E(E) + \xi_E(t), \tag{5.25}$$

where $\langle \xi_E(t)\xi_E(t')\rangle = 2D_{EE}(E)\delta(t-t')$. We now go from a Langevin equation to a Fokker-Planck equation in the usual way.

$$\frac{\partial \rho}{\partial t} = -\frac{\partial}{\partial E}(\rho D_E) + \frac{\partial^2}{\partial E^2}\rho D_{EE}. \tag{5.26}$$

The solution to this equation with a steady state current, a source at $E = 0$ and a sink at the barrier energy gives the approximate equilibrium probability

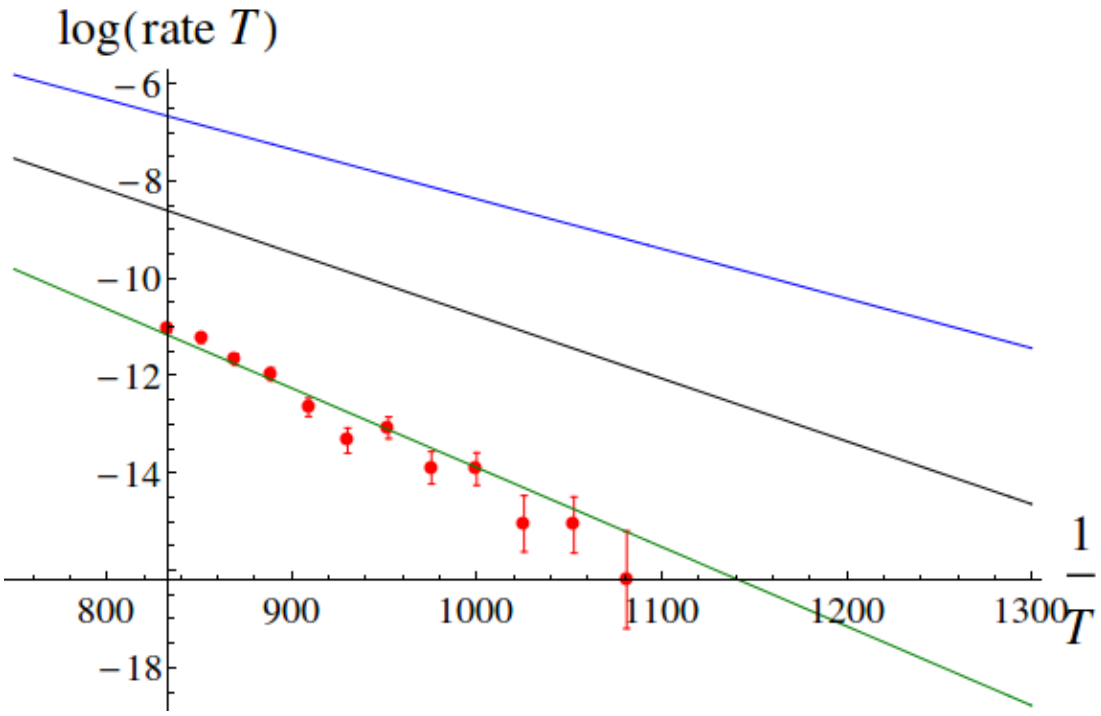

Figure 5.11: Here, parameter values are $\tau = 1, K = 6, \gamma = 0.005$. Whereas both the effective Hamiltonian generated from BCH (black line) and from NF (blue line) fail to capture the escape rate (mostly because they get the wrong E_b), we show that one can use a variational method to improve the estimate of the barrier energy and get a good estimate of the escape rate (green line)

distribution. This distribution is a Boltzmann distribution to linear order in the energy with higher order corrections in E.

We now use the flux-over population method to solve for the escape rate. The flux-over population method involves solving the above equations with a constant steady state current and dividing by the density to find the escape rate [80]. That is, we want to solve the differential equation

$$-D_E(E)\rho + \frac{\partial}{\partial E}(D_{EE}\rho) = J, \tag{5.27}$$

$$\frac{-D_E + D'_{EE}}{D_{EE}}\rho + \frac{\partial \rho}{\partial E} = \frac{J}{D_{EE}}, \tag{5.28}$$

Define

$$v(E) = \int \frac{-D_E + D'_{EE}}{D_{EE}} dE, \tag{5.29}$$

$$= \int -\frac{D_E}{D_{EE}} dE + \log D_{EE}. \tag{5.30}$$

Then it can be checked that the solution is

$$\rho(E) = e^{-v(E)} \int \frac{e^{v(E)}}{D_{EE}} J dE. \tag{5.31}$$

This equation gives the full form of the equilibrium probability distribution. The inverse of the escape rate is given by $k^{-1} = \frac{\int_0^{E_b} \rho(E)}{J}$. This simplifies to

$$k^{-1} = \int_0^{E_b} \frac{e^{\int \frac{D_E}{D_{EE}}}}{D_{EE}} dE \int_E^{E_b} e^{\int -\frac{D_E}{D_{EE}}} dE'. \tag{5.32}$$

The first integral (over E') is dominated by its value at E_b. The second integral (over E) is dominated by its value at 0. Estimating the integral by its value at these boundaries gives us a first approximation to k (which is valid for $E_b \gg T$). Doing this requires us to evaluate the integral $\int D_E/D_{EE} dE$. Going back to Equation 5.19, we see that the drift coefficient has two terms. The first term is independent of T while the second term is linear in T. Hence, the above integral has a part which depends on T and contributes to the exponent. This is well behaved everywhere. The other part contributes to the pre-factor and has a singularity at $E = 0$ because the diffusion coefficient vanishes there. Hence estimating the escape rate requires one to find the finite limit $e^{-\int D_E/D_{EE}}/D_{EE}$ converges to at 0. This was done numerically by evaluating the integral for finite ϵ and then taking ϵ very small. In general, this means the escape rate has the form

$$k = \frac{a_0 \gamma}{T} e^{-f(E_b)/T}, \tag{5.33}$$

where a_0 is a pre-factor and $f(E_b)$ is some function of the barrier. Both of these are independent of the damping and the temperature.

We note that the escape rate calculation is independent of the perturbation theory used to generate the Hamiltonian. Since it depends exponentially on the energy barrier, it is very sensitive to the position of the barrier. We show a comparison of the prediction of the escape rate with simulations in Figure 5.11.

Calculating the escape rate in higher dimensions is more complicated. In 2-d, there are two slow variables given by the invariant in the vertical and horizontal direction. The noise in different directions is typically very different [141]. There are three possible approaches which we will explore in future work. One is to derive and solve the full diffusion equation in 2-d. Secondly, one can derive the decay rate in the limit where the coupling between the two directions is small. Finally, we can solve the full non-equilibrium problem using methods like transition path theory [105] which were developed to deal with inherently non-equilibrium systems.

5.6 Conclusion

We have here compared two different perturbation theories and numerically improved them to calculate the aperture of a harmonic map with a nonlinear kick. We have also calculated the emittance and escape rate in 1-d. Our method is expected to work for systems without dangerous resonances and weak damping. A lot of effort has gone into understanding the resonances which inevitably prohibit the presence of a simple aperture. However, the existence of relatively stable 'islands' of KAM tori in phase space, embedded in a sea of unstable, chaotic trajectories, is a commonly observed phenomenon. Here our aperture is such an island, and our exploration of techniques for calculating it is a special case of

a general mathematical challenge.

The perturbation theories we have been using are asymptotic series and do badly after a certain order. One way to incorporate higher order terms is by resumming the series these expansions generate. It will be useful and interesting to explore methods to resum the series that the BCH and NF methods generate in a way which is able to capture both the presence of resonances and also the presence of the aperture.

We aim to extend our escape rate calculations to higher dimensions. Arnol'd diffusion is another aspect of high-dimensional chaotic motion which we have not addressed here. Several recent analytical methods exist to try and estimate the time scale of Arnol'd diffusion which utilize the multiple invariants mentioned previously [74]. Alternatively, this time scale can also be estimated by more direct methods [26]. It would be interesting to examine the competition between the two time scales of ordinary and Arnol'd diffusion in different parts of phase space giving a much more comprehensive picture of the escape process.

APPENDIX A

DERIVATION OF MAGNETIZATION OF THE 4-D ISING MODEL

To get the magnetization of 4-d Ising model in the infinite system, we have to solve the following two equations

$$\frac{dt}{d\ell} = 2t - Atu, \tag{A.1}$$

$$\frac{du}{d\ell} = -u^2 + Du^3. \tag{A.2}$$

Dividing the two equations by each other gives

$$\frac{dt}{du} = \frac{2t - Atu}{-u^2 + Du^3}, \tag{A.3}$$

which has the solution

$$\log \frac{t}{t_0} = 2\left(\frac{1}{u} - \frac{1}{u_0}\right) + (2D - A)\log\left(\frac{1/(Du) - 1}{1/(Du_0) - 1)}\right). \tag{A.4}$$

We want to coarse grain till $t(\ell) = 1$ or equivalently $t(u) = 1$. It also helps to define $s = 1/(Du) - 1$. This is just a convenient variable for calculations. Then

$$-\log t_0 = 2D(s - s_0) + (2D - A)\log s/s_0, \tag{A.5}$$

where $s_0 = 1/(Du_0) - 1$. This is almost the standard form for the equation of a Lambert-W function defined by $W(z)e^{W(z)} = z$ or equivalently, $\log W(z) + W(z) = \log z$. The solution is

$$s = (2D - A)/(2D)W(xt_0^{1/(A-2D)}), \tag{A.6}$$

where $x = (2D)/(2D - A)s_0 e^{2D/(2D-A)s_0}$. We also have

$$\frac{du}{-u^2 + Du^3} = dl, \tag{A.7}$$

which gives

$$\ell = \frac{1}{u} - \frac{1}{u_0} + D\log\left((1/(Du) - 1)/(1/(Du_0) - 1)\right), \tag{A.8}$$

or

$$\ell = D(s - s_0 + \log s/s_0). \tag{A.9}$$

We are only interested in the dependence of the magnetization on t_0 because u_0 is an irrelevant variable, and so we can ignore the dependence on it. However, we have to be careful because u is a dangerous irrelevant variable and contributes to the leading singularity of the magnetization. One quick way to get the magnetization is to use the result from mean field theory

$$M \sim \frac{e^{-\ell}}{\sqrt{u}}, \tag{A.10}$$

$$\sim \frac{e^{-D(s-s_0)}}{D(s+1)}\left(\frac{s}{s_0}\right)^{-D}, \tag{A.11}$$

$$\sim e^{-Ds}s^{-D}\sqrt{s+1}, \tag{A.12}$$

$$\sim \exp((A-2D)/2W(xt_0^{1/(A-2D)}))\times$$
$$\sqrt{1+W(xt_0^{1/(A-2D)})}W(xt_0^{1/(A-2D)})^{-D}, \tag{A.13}$$

$$\sim t_0^{1/2}W(xt_0^{1/(A-2D)})^{-A/2}(1+W(Yt_0^{1/(A-2D)}))^{1/2}, \tag{A.14}$$

where we have used the identity $e^{aW(x)} = x^a/W(x)^a$ which follows from the definition of the W function. Finally, near the critical point as $t_0 \to 0$, the W function goes to infinity. So, ignoring the 1, we get

$$M \sim t_0^{1/2}W(xt_0^{1/(A-2D)})^{1/2-A/2}. \tag{A.15}$$

For the 4-d Ising model, $A = 1/3$, $D = 17/27$, giving

$$M \sim t_0^{1/2}W(Yt_0^{-27/25})^{1/3}, \tag{A.16}$$

which is the result quoted in the main text.

APPENDIX B

CHANGING THE LENGTH PARAMETER ℓ IN NORMAL FORM THEORY

This form of the equation does motivate another change of variables

$$(1 + Du)\frac{du}{d\ell} = -u^2, \tag{B.1}$$

$$\frac{du}{d\tilde{\ell}} = -u^2 \tag{B.2}$$

where $\tilde{\ell}$ is defined so $\frac{d\tilde{\ell}}{d\ell} = 1/(1 + Du)$. So far, we have been considering changes of variables in our coordinates but have left the flow parameter ℓ unchanged. This parameter usually corresponds to a physical length or momentum scale. However, there is nothing in principle which stops us from allowing ℓ to depend on the coordinates. This would be somewhat strange from a physical point of view but is not disallowed. The t equation is changed to

$$\frac{dt}{d\tilde{\ell}} = 2t - (A - 2D)tu - 2ADtu^2. \tag{B.3}$$

However, we can now make another change of variables in t which removes the $2ADtu^2$ term (since that will leave the flow equations for u unchanged) and so the new normal form (after renaming $\tilde{t}$ to t is)

$$\frac{dt}{d\tilde{\ell}} = 2t - (A - 2D)tu. \tag{B.4}$$

This is a simpler set of equations. Currently, we have not been able to distinguish between these two possible normal forms. This is because for the 4-D Ising model, our scaling form that we derived in the previous section implicitly rely on the equation for L

$$\frac{dL}{d\ell} = -L. \tag{B.5}$$

We have treated L as a special variable and not included it in our normal form calculations. However, if we include changes in the coarse graining length ℓ, it would be naturally modified to

$$\frac{dL}{d\tilde{\ell}} = -L(1 + Du). \tag{B.6}$$

Since the extra term is a resonance, it cannot be removed by a change of variables. In effect, this leaves du/dL and dt/dL unchanged. Hence, this analysis does not seem to matter for the finite scaling analysis of the 4-d Ising model. However, it is quite possible that such an analysis could be carried out in other cases where it does fundamentally change the form of the scaling and then it would be interesting to test it.

APPENDIX C

NORMAL FORM FOR IRRELEVANT VARIABLES WITH RESONANCES

Here, we derive the normal form for the set of equations

$$\frac{dt}{d\ell} = \lambda_t t + a_1 tw + a_2 tw^2 + a_3 tw^3 + ..., \tag{C.1}$$

$$\frac{du}{d\ell} = \lambda_u u + b_1 uw + b_2 uw^2 + ..., \tag{C.2}$$

$$\frac{dh}{d\ell} = \lambda_h h + c_1 hw + c_2 hw^2 + ..., \tag{C.3}$$

The equation for dw/dl follows

$$\frac{dw}{d\ell} = pt^{p-1}u^q \frac{dt}{d\ell} + qt^p u^{q-1} \frac{du}{d\ell}, \tag{C.4}$$

$$= pt^{p-1}u^q(\lambda_t t + a_1 tw^2 + a2tw^2 + ...) + qt^p u^{q-1}(\lambda_u u + b_1 uw + b_2 uw^2 + ...), \tag{C.5}$$

$$= (pa_1 + qb_1)w^2 + (pa_2 + qb_2)w^3 + (pa_3 + qb_3)w^4 + ... \tag{C.6}$$

The linear term cancels as a result of Equation 3.98. For all variables, except for t and u, this also gives an efficient way to remove the resonances because the problem maps on to that of a transcritical bifurcation. In the case of a transcritical bifurcation, we already know the normal form of the h equation should be

$$\frac{dh}{d\ell} = \lambda_h h + c_1 hw. \tag{C.7}$$

The change of variables to achieve this form is the same as that used in the transcritical bifurcation, $h = u_h + C_1 u_h w + ...$ t and u are a little more complicated because changing t or u also changes w. Moreover, note in the equation for $\frac{dw}{d\ell}$, it is terms of the form $a_2 tw^2$ which give the third order term in the equation for the transcritical bifurcation. We know already that this third order term cannot be removed from the hyper-normal form of the transcritical bifurcation.

In practice, this means we have to change u and t simultaneously to try and remove terms. The analysis is general but it helps to consider the simple case

where $\lambda_t = 1$, $\lambda_u = -1$ so $p = 1, q = 1$. Let's keep terms up to cubic order in w. So we start with the equations

$$\frac{dt}{d\ell} = t + a_1 tw + a_2 tw^2 + a_3 tw^3, \tag{C.8}$$

$$\frac{du}{d\ell} = -u + b_1 uw + b_2 uw^2 + b_3 uw^3, \tag{C.9}$$

$$\tag{C.10}$$

where $w = tu$ and try the change of variables

$$t(\tau, \upsilon) = \tau + A_1 \tau\omega + A_2 \tau\omega^2, \tag{C.11}$$

$$u(\tau, \upsilon) = \upsilon + B_1 \upsilon\omega + B_2 \upsilon\omega^2, \tag{C.12}$$

$$\omega = \tau\upsilon \tag{C.13}$$

Making this change of variables gives the equations

$$\frac{d\tau}{d\ell} = \tau + a_1 \tau\omega + (a_2 - A_1 b_1 + a_1 B_1)\tau\omega^2 + (a_3 - f)\tau\omega^2, \tag{C.14}$$

$$\frac{d\upsilon}{d\ell} = -\upsilon + b_1 \upsilon\omega + (b_2 + A_1 b_1 - a_1 B_1)\upsilon\omega^2 + (b_3 - g)\tau\omega^2, \tag{C.15}$$

$$\tag{C.16}$$

f and g are some polynomial expressions which depend on all parameters. The first thing to notice is that terms like $\tau\omega$ cannot be removed. For the second order (in ω) terms, we can remove *only one* of either $\tau\omega^2$ or $\upsilon\omega^2$. This is because choosing $A_1 = \frac{a_1 B_1 + a_2}{b_1}$ cancels B_1 in the term $\upsilon\omega^2$. We no longer have the freedom to choose B_1 to cancel the second order term in the υ equation.

This problem does not extend to the third order terms because they now depend on a free parameter B_1, along with A_2 and B_2 which we can choose to cancel both $\tau\omega^2$ and $\upsilon\omega^2$ terms. This argument extends to higher order terms and to arbitrary eigenvalues. Hence the normal form of the equations is

$$\frac{dt}{d\ell} = \lambda_t t + a_1 tw + a_2 tw^2, \tag{C.17}$$

$$\frac{du}{d\ell} = \lambda_u u + b_1 uw, \tag{C.18}$$

$$\frac{dh}{d\ell} = \lambda_h h + c_1 hw. \tag{C.19}$$

APPENDIX D

WEGNER'S RESULTS ON REDUNDANT VARIABLES

Here we give a brief derivation of Wegner's analysis of redundant operators. Wegner [150] considers a Hamiltonian $H(s_q)$ as a function of the Fourier transform of a soft-spin variable s_q.

1. *Coarse Graining*

   Wegner's insight was to notice that coarse graining can be represented by a change of variables $s_q \rightarrow s_q + \delta\psi_q(s)$. He considered an infinitesimal change of variables, and so the Hamiltonian changes as

   $$H(s_q) \rightarrow H(s_q + \delta\psi_q(s)), \tag{D.1}$$

   $$= H(s_q) + \delta\partial_s H\psi_q \tag{D.2}$$

   The partition function is given by

   $$Z = \int \prod_q dS_q e^{-H(s_q)} \tag{D.3}$$

   Under this change of variables, the measure of the integral also changes as $dS_q \rightarrow dS_q(1 + \partial_s\psi\delta)$. So, the total change in the partition function

   $$Z = \int \prod_q dS - qe^{-H(s_q)-\delta(\partial_s H\psi-\partial_s\psi)} \tag{D.4}$$

   This can be represented as an operator $\mathcal{G}$ which depends on ψ as

   $$\mathcal{G}(\psi)H = \prod_q \partial_s H\psi_q - \partial_s\psi_q \tag{D.5}$$

   Note that ψ itself can depend on H.

2. *Rescaling* The spin operator is rescaled to account for the changed length (momentum) scale as

   $$s_q \rightarrow (1+\delta)^{d/2} s_{q+\delta q}, \tag{D.6}$$

   $$= s_q + \delta(d/2 s_q + q\nabla s_q) \tag{D.7}$$

where the ∇ denotes a derivative with respect to q. And so the Hamiltonian changes as

$$H(s_q) \to H(s_q) + \sum_q \delta \partial_s H(d/2 s_q + q \nabla s_q) \tag{D.8}$$

Because the term in paranthesis only gives a constant when differentiated with respect to s_q (which we ignore), we can represent the dilitation as $\mathcal{G}(D)$ where $D = d/2 s_q + q \nabla s_q$. This dilitation is different from a change of variables because of the presenece of ∇s_q which changes the volume of the system.

The combination of these two operators gives the RG trajectory

$$\frac{dH}{dl} = (\mathcal{G}(\psi) + \mathcal{G}(D))H = \mathcal{G}(\psi + D)H \tag{D.9}$$

where we use linearity of the operator $\mathcal{G}$. Near the fixed point, the flow equations can be linearized

$$\frac{d\Delta H}{dl} = \mathcal{L}\delta H \tag{D.10}$$

where $\mathcal{L}$ is the linearization of $\mathcal{G}$ around the fixed point $\mathcal{G}(\psi^* + D)H^* = 0$. Now, Wegner considers a perturbation $H^* + \mu\mathcal{G}(\phi)H^*$ with some arbitrary ϕ. The operator $\mu\mathcal{G}(\phi)H^*$ is just a change of variables and is hence redundant. Then it can be shown that

$$\mathcal{G}(\psi^* + D)(H^* + \mu\mathcal{G}(\phi)H^*) = \mu\mathcal{G}(\{\psi^* + D, \phi\}) \tag{D.11}$$

That is, the action of the RG operator on a redundant perturbation near the fixed point leads to another redundant operator. Therefore, the space of redundant operators is a subspace of the set of total operators. Acting on a redundant operator with an RG operation leads to another redundant operator. However, because the RG depends on H, a perturbation $\mathcal{G}(\phi)$ also perturbs the form of the

RG. So

$$\mathcal{G}(\psi^* + \mu\Psi + D)(H^* + \mu\mathcal{G}(\phi)H^*) = \mu\mathcal{G}(\{\psi^* + D, \phi\} + \Psi) \tag{D.12}$$

Wegner's argument is that we can choose $\mathcal{G}(\phi)$ to have any eigenvalue by chosing Ψ appropriately. Hence, the eigenvalues of redundant operators are not fixed and can be changed by changing the RG.

APPENDIX E

CHANGING GAUGE EIGENVALUES IN PERIOD DOUBLING

Here, we show how the treatment of redundant variables in Wegner [150] can be translated into period doubling. Wegner took a general Hamiltonian $H(s_q)$ as a function of the Fourier transform of a soft-spin variable s_q and then considered the effect of changing the definition of s_q on the RG operator $\mathcal{G}$. A detailed version of his argument is in Appendix D. Here, we just note that there is a direct map between period doubling and statistical mechanics

$$x \to s_q,$$

$$g \to H,$$

$$\mathcal{T} \to G.$$

In period doubling, we denote the change of coordinate by $y = \phi(x)$. This induces a map $\tilde{g}(y) = \phi(g(\phi^{-1}(y)))$. We follow the treatment in the mean text, acting the old RG on the new fixed point.

$$\mathcal{T}[\tilde{g}^*](x) = \tilde{\alpha}(\phi(g^*(g^*(\phi^{-1}(\tilde{\alpha}^{-1}(x))))))) \tag{E.1}$$

As before the infinitesimal version of the transformation $\phi(x) = x + \epsilon\psi(x)$. The inverse transformation $\phi^{-1}(x) = x - \epsilon\psi(x)$. The difference here is that we allow the rescaling to depend on x through $\tilde{\alpha}(x) = \alpha(x + \epsilon\Lambda(x))$ and $\tilde{\alpha}^{-1}(x) = x/\alpha - \epsilon\Lambda(x/\alpha)$. At the fixed point $\alpha(g^*(g^*(x/\alpha)) = g^*(x)$. Taking a derivative of this equation

gives $g^{*\prime}(g^*(x/\alpha)))g^{*\prime}(x/\alpha) = g*'(x)$. We can expand to linear order in ϵ

$$\mathcal{T}[\tilde{g}^*](x) = \tilde{\alpha}(\phi(g^*(g^*(\phi^{-1}(x/\alpha - \epsilon\Lambda(x/\alpha)))))), \tag{E.2}$$

$$= \tilde{\alpha}(\phi(g^*(g^*(x/\alpha - \epsilon\Lambda(x/\alpha) - \epsilon\psi(x/\alpha))))), \tag{E.3}$$

$$= \tilde{\alpha}((g^*(g^*(x/\alpha - \epsilon\Lambda(x/\alpha) - \epsilon\psi(x/\alpha))))) + \epsilon\psi(g^*(x)/\alpha), \tag{E.4}$$

$$= \alpha((g^*(g^*(x/\alpha - \epsilon\Lambda(x/\alpha) - \epsilon\psi(x/\alpha)))) + \epsilon\alpha\psi(g^*(x)/\alpha)) + \alpha\epsilon\Lambda(g^*(x)/\alpha), \tag{E.5}$$

$$= \alpha(g^*(g^*(x/\alpha)) - g^{*\prime}(x)\epsilon(\Lambda(x/\alpha) + \psi(x/\alpha)) + \epsilon\alpha(\psi(g^*(x)/\alpha) + \Lambda(g^*(x)/\alpha)), \tag{E.6}$$

$$= g^*(x) + \alpha\epsilon\left(\Lambda(g^*(x)/\alpha) + \psi(g^*(x)/\alpha) - (\Lambda(x/\alpha) + \psi(x/\alpha))\, g^{*\prime}(x)\right). \tag{E.7}$$

Setting $\Lambda(x) = 0$ gives back the result in the main text. If $\psi(x)$ has a Taylor series $\psi(x) = \sum_p \psi_p x^p$, we get

$$\mathcal{T}[\tilde{g}^*](x) - g^*(x) = \sum_p \alpha^{1-p}(g^*(x)^p - g^{*\prime}(x)x^p). \tag{E.8}$$

However, we are free to choose $\Lambda(x)$ to be whatever we want. In particular, if it has a Taylor series $\Lambda(x) = \sum_p \Lambda_p x^p$, we can set $\Lambda_p = -\beta_p + \alpha^{p-1}$, that sets all the eigenvalues to 1 with the same eigenfunction. Allowing a rescaling that depends on the deformation allows us to set some of the eigenvalues to whatever we want. This is different from universal RG eigenvalues which are thought to be independent of the renormalization scheme. In the main text, we don't consider this broad class of RGs and confine ourselves to the case where $\Lambda(x) = 0$.

APPENDIX F

DETAILS OF FIGURES FOR STABILITY BOUNDARIES IN KICKED MAPS

Here we give a few more details on some of the figures in the paper. The saddle points for the 1-d map given in the main text are calculated by truncating the perturbation theory at a certain order. We have truncated the BCH Hamiltonian to second order. The next order correction does a worse job of capturing the dynamic aperture as can be seen from Figure F.2. We have truncated the NF expansion at 3rd order. If we keep the next order term, the Hamiltonian given below no longer has a saddle point in the region where we expect the boundary of the aperture to be as is evident from Figure F.2. The next order Hamiltonians in the two cases are given by:

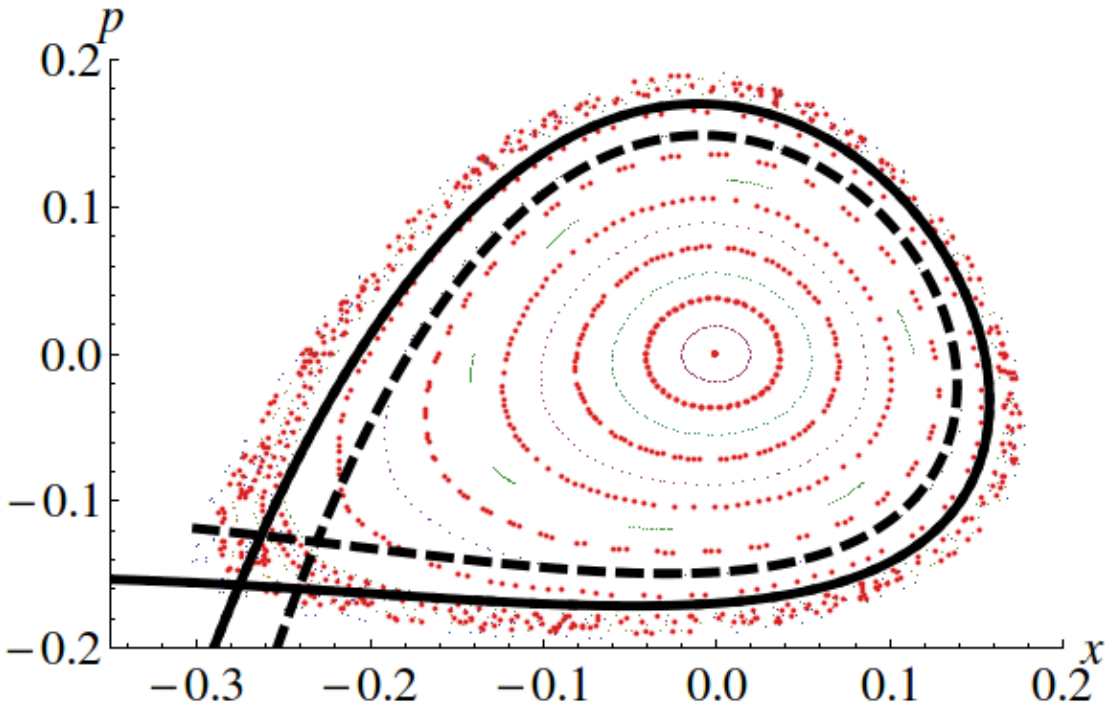

Figure F.1: Here we show the effect of including the next order term in the BCH expansion. A comparison of the 2nd (black line) and 3rd order (black dashed line) Hamiltonians obtained using BCH. The aperture obtained using the 3rd order Hamiltonian is smaller in size.

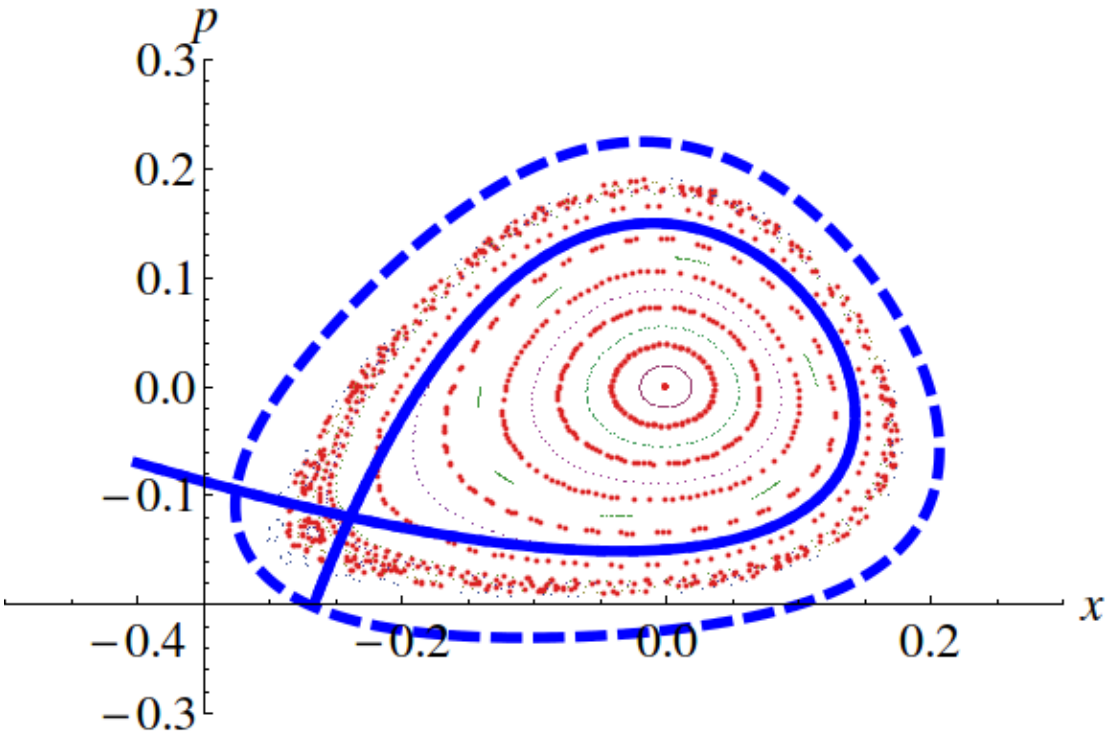

Figure F.2: Here we show the effect of including the next order term in the NF expansion. The 3rd order (blue line) NF Hamiltonian has an aperture but the 4th order (blue dashed line) Hamiltonian no longer has a saddle point close to the boundary of the actual aperture as the given contour shows.

$$H_{\mathrm{NF}} = \frac{1}{2}(p^2 + x^2\omega^2) + \frac{1}{2}\left(-\frac{K^2 p^4 \cos(\omega\tau)(\omega + \tan(\omega\tau))^2}{8\left(\omega^3 + \omega\right)^2 (2\cos(\omega\tau) + 1)}\right) \tag{F.1}$$

$$- \left(\frac{K^2 p^2 x^2 \sec(\omega\tau)\left(\omega^4 + \left(3\omega^4 + 1\right)\cos(2\omega\tau) + 4\omega^3 \sin(2\omega\tau) - 1\right)}{16\left(\omega^3 + \omega\right)^2 (2\cos(\omega\tau) + 1)}\right)$$

$$- \frac{1}{2}\left(\frac{K^2 x^4 \sec(\omega\tau)\left(2\omega^3 \sin(2\omega\tau) - \omega^2 - \left(3\omega^2 + 1\right)\cos(2\omega\tau) - 1\right)}{16\left(\omega^2 + 1\right)^2 (2\cos(\omega\tau) + 1)}\right)$$

$$+ \frac{1}{2}\left(\frac{K p^2 x \sin(\omega\tau)}{2\omega + 4\omega\cos(\omega\tau)} + \frac{1}{2}K p x^2\right)$$

$$+ \frac{1}{2}\left(\frac{K x^3 \omega\left(\left(\cos^3(\omega\tau) + 1\right)\cot(\omega\tau) + \sin(\omega\tau)\cos^2(\omega\tau)\right)}{4\cos(\omega\tau) + 2} + \frac{K^2 p x^3 \sin(\omega\tau)}{4\omega + 8\omega\cos(\omega\tau)}\right)$$

$$+ O(K^3) \tag{F.2}$$

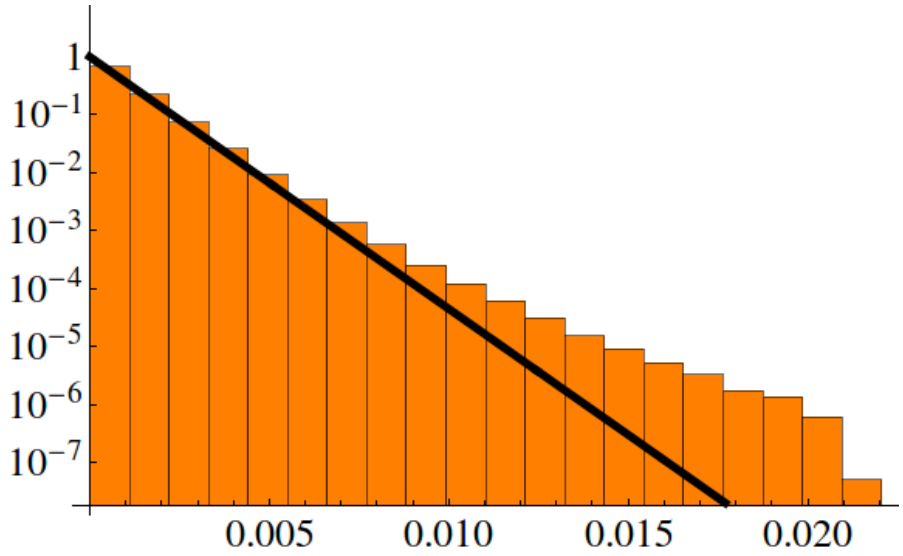

Figure F.3: A histogram of the harmonic energy on a logarithmic scale of the particles along with the straight line corresponding to the Boltzmann distribution. The harmonic energy does not describe the ends of the distribution.

$$H_{\mathrm{BCH}} = \frac{1}{2}\left(x^2\omega^2 + p^2\right) + \frac{K^2\tau x^4 - 2Kx\left(x^2\left(\tau^2\omega^2 - 4\right) - 6p\tau x - 2p^2\tau^2\right) - 2K^2\tau^2 p x^3}{48\tau} \tag{F.3}$$

The 2-d map is a generalization of the sextupole map to higher dimensions. It is given by:

$$x_{n+1} = x_n \cos\omega_1\tau + \frac{px_n}{\omega_1}\sin\omega_1\tau, \tag{F.4}$$

$$px_{n+1} = px_n \cos\omega_1\tau - \omega_1 x_n \sin\omega_1\tau - \frac{K}{2}(x_{n+1}^2 - y_{n+1}^2), \tag{F.5}$$

$$y_{n+1} = y_n \cos\omega_2\tau + \frac{py_n}{\omega_2}\sin\omega_2\tau, \tag{F.6}$$

$$py_{n+1} = py_n \cos\omega_2\tau - \omega_2 y_n \sin\omega_2\tau + Ky_{n+1}x_{n+1} \tag{F.7}$$

The effective Hamiltonian from BCH can be calculated in the same way as given in the main text. The contours are plotted by setting $I_x + I_y = c$ where c is the saddle-point value of $I_x + I_y$. To find the curve that sets the boundary in the space of two invariants, consider setting one of the invariants to a constant and then

finding the value of the other invariant which goes through a saddle point. This can be written using a Lagrange multiplier

$$\frac{dI_x}{dx} - \lambda \frac{dI_y}{dx} = 0, \tag{F.8}$$

$$\frac{dI_x}{dpx} - \lambda \frac{dI_y}{dpx} = 0 \tag{F.9}$$

$$\frac{dI_x}{dy} - \lambda \frac{dI_y}{dy} = 0, \tag{F.10}$$

$$\frac{dI_x}{dpy} - \lambda \frac{dI_y}{dpy} = 0 \tag{F.11}$$

$$I_y = c. \tag{F.12}$$

This gives us 5 equations with 5 unknowns x, px, y, py, λ. These equations were solved numerically and the set of solutions that are closest to the observed numerical boundary are plotted in the main text.

The emittance histogram is drawn by sampling the effective energy by starting at the centre and running for a long time with the kicked noisy map. The emittance is often estimated using a harmonic approximation to the energy. We show in Figure F.3 that this approximation does well near the centre but does not accurately describe the ends of the distribution.

BIBLIOGRAPHY

[1] Amnon Aharony and Michael E Fisher. Universality in analytic corrections to scaling for planar ising models. *Physical Review Letters*, 45(9):679, 1980.

[2] Amnon Aharony and Michael E Fisher. Nonlinear scaling fields and corrections to scaling near criticality. *Physical Review B*, 27(7):4394, 1983.

[3] Björn Ahrens and Alexander K. Hartmann. Critical behavior of the random-field Ising model at and beyond the upper critical dimension. *Phys. Rev. B*, 83:014205, Jan 2011.

[4] N Aktekin. The finite-size scaling functions of the four-dimensional Ising model. *Journal of Statistical Physics*, 104(5-6):1397–1406, 2001.

[5] Abu Hamid al Ghazali. *Tahafut al-Falasifa*.

[6] S Amari and H Nagaoka. *Methods of Information Geometry*. Oxford University Press, 2007.

[7] P Anderson. More is different. *Science*, 177, 1972.

[8] V. I. Arnold. *Mathematical methods of classical mechanics*. Springer-Verlag, 1978.

[9] Vladimir I Arnold. Proof of AN Kolmogorovs theorem on the conservation of conditionally periodic motions with a small variation in the Hamiltonian. *Russian Math. Surv*, 18(9), 1963.

[10] Vladimir I Arnold. Instability of dynamical systems with several degrees of freedom(instability of motions of dynamic system with five-dimensional phase space). *Soviet Mathematics*, 5:581–585, 1964.

[11] Vladimir Igorevich Arnol'd. Small denominators and problems of stability of motion in classical and celestial mechanics. *Russian Mathematical Surveys*, 18(6):85–191, 1963.

[12] M Bai, D Jeon, SY Lee, KY Ng, A Riabko, and X Zhao. Stochastic beam dynamics in quasi-isochronous storage rings. *Physical Review E*, 55(3):3493, 1997.

[13] V. Balasubramanian, J. J. Heckman, and Alexander Maloney. Relative entropy and proximity of quantum field theories. *J. High Energ. Phys.*, 104, 2015.

[14] Mustansir Barma and Michael E Fisher. Corrections to scaling and crossover in two-dimensional Ising and scalar-spin systems. *Physical Review Letters*, 53(20):1935, 1984.

[15] Mustansir Barma and Michael E. Fisher. Two-dimensional Ising-like systems: Corrections to scaling in the Klauder and double-Gaussian models. *Phys. Rev. B*, 31:5954–5975, May 1985.

[16] R Batterman. The tyranny of scales. In *The Oxford Handbook of Philosophy of Physics*. Oxford University Press, 2013.

[17] R W Batterman. Emergence, singularities, and symmetry breaking. *Foundations of Physics*, 41, 2011.

[18] A Bazzani, M Giovannozzi, G Servizi, E Todesco, and G Turchetti. Resonant normal forms, interpolating Hamiltonians and stability analysis of area preserving maps. *Physica D: Nonlinear Phenomena*, 64(1-3):66–97, 1993.

[19] A Bazzani, P Mazzanti, G Servizi, and G Turchetti. Normal forms for Hamiltonian maps and nonlinear effects in a particle accelerator. *Il Nuovo Cimento B (1971-1996)*, 102(1):51–80, 1988.

[20] A Bazzani, E Remiddi, and G Turchetti. Variational approach to the integrals of motion for symplectic maps. *Journal of Physics A: Mathematical and General*, 24(2):L53, 1991.

[21] Armando Bazzani, Stefano Marmi, and Giorgio Turchetti. Nekhoroshev estimate for isochronous non resonant symplectic maps. *Celestial Mechanics and Dynamical Astronomy*, 47(4):333–359, 1989.

[22] David Bensimon, Mogens H Jensen, and Leo P Kadanoff. Renormalization-group analysis of the global structure of the period-doubling attractor. *Physical Review A*, 33(5):3622, 1986.

[23] Cédric Bény and Tobias J Osborne. Information-geometric approach to the renormalization group. *Physical Review A*, 92(2):022330, 2015.

[24] B Berche, P Butera, and L N Shchur. The two-dimensional 4-state Potts model in a magnetic field. *Journal of Physics A: Mathematical and Theoretical*, 46(9):095001, 2013.

[25] Michael Victor Berry. Regular and irregular motion. In *AIP Conference proceedings*, volume 46, pages 16–120. AIP, 1978.

[26] Martin Berz and Georg Hoffstätter. Exact bounds of the long term stability of weakly nonlinear systems applied to the design of large storage rings. *Interval Computations*, 2:68–89, 1994.

[27] Olivier Bienaymé and Gregor Traven. Approximate integrals of motion. *Astronomy & Astrophysics*, 549:A89, 2013.

[28] Kurt Binder, M Nauenberg, V Privman, and AP Young. Finite-size tests of hyperscaling. *Physical Review B*, 31(3):1498, 1985.

[29] George David Birkhoff. *Dynamical systems*. American mathematical society, 1960.

[30] A. J. Bray and M. A. Moore. Scaling theory of the random-field Ising model. *J. Phys. C*, 18(L927), 1985.

[31] Keith Briggs. Corrections to universal scaling in real maps. *Physics Letters A*, 191(1-2):108–112, 1994.

[32] D. Brody and N. Rivier. Geometrical aspects of statistical mechanics. *Phys. Rev. E.*, 51, 1995.

[33] Kevin S Brown and James P Sethna. Statistical mechanical approaches to models with many poorly known parameters. *Physical Review E*, 68(2):021904, 2003.

[34] AD Bruce and DJ Wallace. Droplet theory of low-dimensional ising models. *Physical Review Letters*, 47(24):1743, 1981.

[35] Marin Bukov, Luca D'Alessio, and Anatoli Polkovnikov. Universal high-frequency behavior of periodically driven systems: from dynamical stabilization to Floquet engineering. *Advances in Physics*, 64(2):139–226, 2015.

[36] J Butterfield. Less is different: Emergence and reduction reconciled. *Foundations of Physics*, 41, 2011.

[37] Duncan S Callaway, John E Hopcroft, Jon M Kleinberg, Mark EJ Newman, and Steven H Strogatz. Are randomly grown graphs really random? *Physical Review E*, 64(4):041902, 2001.

[38] Léonie Canet, Bertrand Delamotte, and Nicolás Wschebor. Fully developed isotropic turbulence: Nonperturbative renormalization group formalism and fixed-point solution. *Physical Review E*, 93(6):063101, 2016.

[39] John Cardy. *Scaling and renormalization in statistical physics*, volume 5. Cambridge university press, 1996.

[40] Sean M Carroll. *Spacetime and geometry. An introduction to general relativity*. 2004.

[41] John R Cary. Lie transform perturbation theory for Hamiltonian systems. *Physics Reports*, 79(2):129–159, 1981.

[42] Michele Caselle, Martin Hasenbusch, Andrea Pelissetto, and Ettore Vicari. Irrelevant operators in the two-dimensional ising model. *Journal of Physics A: Mathematical and General*, 35(23):4861, 2002.

[43] Y Chan, Anthony J Guttmann, BG Nickel, and JHH Perk. The ising susceptibility scaling function. *Journal of Statistical Physics*, 145(3):549–590, 2011.

[44] Patrick Charbonneau, Yi Hu, Archishman Raju, James P Sethna, and Sho Yaida. Morphology of renormalization-group flow for the de almeida-thouless-gardner universality class. *forthcoming*.

[45] Patrick Charbonneau and Sho Yaida. Nontrivial critical fixed point for replica-symmetry-breaking transitions. *Physical review letters*, 118(21):215701, 2017.

[46] Lin Yuan Chen, Nigel Goldenfeld, and Yoshitsugu Oono. Renormalization group theory for global asymptotic analysis. *Physical review letters*, 73(10):1311, 1994.

[47] Alain Chenciner. Poincaré and the three-body problem. In *Henri Poincaré, 1912–2012*, pages 51–149. Springer, 2015.

[48] KG Chetyrkin, SG Gorishny, SA Larin, and FV Tkachov. Five-loop renormalization group calculations in the $g\phi^4$ theory. *Physics Letters B*, 132(4):351–354, 1983.

[49] Boris V Chirikov. A universal instability of many-dimensional oscillator systems. *Physics reports*, 52(5):263–379, 1979.

[50] Noam Chomsky. Language and nature. In *Powers and Prospects*, chapter 2, pages 31–54. Madhyam Books, 1996.

[51] Noam Chomsky. Language and thought: Some reflections on venerable themes. In *Powers and Prospects*, chapter 1, pages 1–30. Madhyam Books, 1996.

[52] Freddy Christiansen, Predrag CvitanoviC, and Hans Henrik Rugh. The spectrum of the period-doubling operator in terms of cycles. *Journal of Physics A: Mathematical and General*, 23(14):L713S, 1990.

[53] Colin Clement, Archishman Raju, and James P Sethna. Normal form of the two-dimensional ising model renormalization group flows. *forthcoming*.

[54] Karin Dahmen and James P Sethna. Hysteresis, avalanches, and disorder-induced critical scaling: A renormalization-group approach. *Physical Review B*, 53(22):14872, 1996.

[55] BBC Damgaard. Non-asymptotic corrections to orbital scaling in period doubling maps. *Physics Letters A*, 132(5):244–248, 1988.

[56] Daniel Dantchev, H. W. Diehl, and Daniel Grüneberg. Excess free energy and Casimir forces in systems with long-range interactions of van der Waals type: General considerations and exact spherical-model results. *Phys. Rev. E*, 73:016131, Jan 2006.

[57] H. W. Diehl and E. Eisenriegler. Walks, polymers, and other tricritical systems in the presence of walls or surfaces. *EPL (Europhysics Letters)*, 4(6):709, 1987.

[58] H. W. Diehl, Daniel Grüneberg, Martin Hasenbusch, Alfred Hucht, Sergei B. Rutkevich, and Felix M. Schmidt. Large-*n* approach to thermodynamic Casimir effects in slabs with free surfaces. *Phys. Rev. E*, 89:062123, Jun 2014.

[59] H. W. Diehl, Daniel Grneberg, Martin Hasenbusch, Alfred Hucht, Sergei B. Rutkevich, and Felix M. Schmidt. Exact thermodynamic Casimir forces for an interacting three-dimensional modelsystem in film geometry with free surfaces. *EPL (Europhysics Letters)*, 100(1):10004, 2012.

[60] Brian P Dolan. A geometrical interpretation of renormalization group flow. *International Journal of Modern Physics A*, 9(08):1261–1286, 1994.

[61] SN Dorogovtsev, JFF Mendes, and AN Samukhin. Anomalous percolation properties of growing networks. *Physical Review E*, 64(6):066110, 2001.

[62] Alex J. Dragt and John M. Finn. Lie series and invariant functions for analytic symplectic maps. *Jour. Math. Phys.*, 17, 1976.

[63] Alex J Dragt and Etienne Forest. Computation of nonlinear behavior of Hamiltonian systems using Lie algebraic methods. *Journal of mathematical physics*, 24(12):2734–2744, 1983.

[64] William Edward Burghardt Du Bois. The princess steel. *PMLA*, 130(3):819–829, 2015.

[65] William Edward Burghart Du Bois. Sociology hesitant. *boundary 2*, 27(3):37–44, 2000.

[66] Donald A Edwards and Michael J Syphers. *An introduction to the physics of high energy accelerators*. John Wiley & Sons, 2008.

[67] E. Eisenriegler and H. W. Diehl. Surface critical behavior of tricritical systems. *Phys. Rev. B*, 37:5257–5273, Apr 1988.

[68] B. B. Machta et al. Parameter space compression underlies emergent theories and predictive models. *Science*, 342, 2013.

[69] MK Transtrum et al. Perspective: Sloppiness and emergent theories in physics, biology, and beyond. *Journ. Chem. Phys.*, 2015.

[70] Mitchell J Feigenbaum. Quantitative universality for a class of nonlinear transformations. *Journal of statistical physics*, 19(1):25–52, 1978.

[71] Mitchell J Feigenbaum. The universal metric properties of nonlinear transformations. *Journal of Statistical Physics*, 21(6):669–706, 1979.

[72] Mitchell J Feigenbaum, Leo P Kadanoff, and Scott J Shenker. Quasiperiodicity in dissipative systems: a renormalization group analysis. *Physica D: Nonlinear Phenomena*, 5(2-3):370–386, 1982.

[73] Michael E Fisher and Mohit Randeria. Location of renormalization-group fixed points. *Physical review letters*, 56(21):2332, 1986.

[74] Antonio Giorgilli, Ugo Locatelli, and Marco Sansottera. Kolmogorov and Nekhoroshev theory for the problem of three bodies. *Celestial Mechanics and Dynamical Astronomy*, 104(1):159–173, 2009.

[75] Massimo Giovannozzi. Study of the dynamic aperture of the 4d quadratic map using invariant manifolds. Technical report, 1998.

[76] N Goldenfeld. *Lectures on phase transitions and the renormalization group*. Addison-Wesley, 1992.

[77] Nigel Goldenfeld and Leo P Kadanoff. Simple lessons from complexity. *science*, 284(5411):87–89, 1999.

[78] Carl P Goodrich, Simon Dagois-Bohy, Brian P Tighe, Martin van Hecke, Andrea J Liu, and Sidney R Nagel. Jamming in finite systems: Stability, anisotropy, fluctuations, and scaling. *Physical Review E*, 90(2):022138, 2014.

[79] Ryan N Gutenkunst, Joshua J Waterfall, Fergal P Casey, Kevin S Brown, Christopher R Myers, and James P Sethna. Universally sloppy parameter sensitivities in systems biology models. *PLoS computational biology*, 3(10):e189, 2007.

[80] Peter Hänggi, Peter Talkner, and Michal Borkovec. Reaction-rate theory: fifty years after Kramers. *Reviews of modern physics*, 62(2):251, 1990.

[81] Godfrey Harold Hardy. *Divergent series*, volume 334. American Mathematical Soc., 2000.

[82] Lorien Hayden, Archishman Raju, and James P Sethna. Nonlinear scaling theory of the two-dimensional non-equilibrium random-field ising model.

[83] John Z Imbrie. Lower critical dimension of the random-field ising model. *Physical review letters*, 53(18):1747, 1984.

[84] Ernst Ising. Beitrag zur theorie des ferromagnetismus. *Zeitschrift für Physik A Hadrons and Nuclei*, 31(1):253–258, 1925.

[85] G Jona-Lasinio. The renormalization group: A probabilistic view. *Il Nuovo Cimento B (1971-1996)*, 26(1):99–119, 1975.

[86] Giovanni Jona-Lasinio. Renormalization group and probability theory. *Physics Reports*, 352(4):439–458, 2001.

[87] John M Jowett. Introductory statistical mechanics for electron storage rings. In *AIP Conference Proceedings*, volume 153, pages 864–970. AIP, 1987.

[88] Mikko Kaasalainen and James Binney. Construction of invariant tori and integrable Hamiltonians. *Physical review letters*, 73(18):2377, 1994.

[89] Leo P Kadanoff. Notes on migdal's recursion formulas. *Annals of Physics*, 100(1-2):359–394, 1976.

[90] J Kim. Making sense of emergence. *Philosophical Studies*, 95, 1999.

[91] J Kim. Emergence: Core ideas and issues. *Synthese*, 151, 2006.

[92] AN Kolmogorov. On the conservation of conditionally periodic motions under small perturbation of the Hamiltonian. In *Dokl. Akad. Nauk. SSR*, volume 98, pages 2–3, 1954.

[93] Mikhail Kompaniets and Erik Panzer. Renormalization group functions of ϕ^4 theory in the ms-scheme to six loops. *arXiv preprint arXiv:1606.09210*, 2016.

[94] Kyriakos M Kontopoulos. *The logics of social structure*, volume 6. Cambridge University Press, 2006.

[95] J M Kosterlitz and D J Thouless. Ordering, metastability and phase transitions in two-dimensional systems. *Journal of Physics C: Solid State Physics*, 6(7):1181, 1973.

[96] JM Kosterlitz. The critical properties of the two-dimensional xy model. *Journal of Physics C: Solid State Physics*, 7(6):1046, 1974.

[97] JM Kosterlitz. The d-dimensional Coulomb gas and the roughening transition. *Journal of Physics C: Solid State Physics*, 10(19):3753, 1977.

[98] Pik-Yin Lai and KK Mon. Finite-size scaling of the Ising model in four dimensions. *Physical Review B*, 41(13):9257, 1990.

[99] AI Larkin and E Khmel'nitskii. Phase transition in uniaxial ferroelectrics. *WORLD SCIENTIFIC SERIES IN 20TH CENTURY PHYSICS*, 11:43–48, 1995.

[100] Michael Lässig. Geometry of the renormalization group with an application in two dimensions. *Nuclear Physics B*, 334(3):652–668, 1990.

[101] Richard C Lewontin. Its even less in your genes. *The New York Review of Books*, 58(9), 2011.

[102] A. Lichtenberg and M. Lieberman. *Regular and Chaotic motion*. Springer, 1992.

[103] Per Hkan Lundow and Klas Markström. Critical behavior of the Ising model on the four-dimensional cubic lattice. *Physical Review E*, 80(3):031104, 2009.

[104] BA Magradze. An analytic approach to perturbative QCD. *International Journal of Modern Physics A*, 15(17):2715–2733, 2000.

[105] Robert S Maier and Daniel L Stein. Escape problem for irreversible systems. *Physical Review E*, 48(2):931, 1993.

[106] Jian-min Mao and Bambi Hu. Corrections to scaling for period doubling. *Journal of statistical physics*, 46(1-2):111–117, 1987.

[107] G Martinelli and G Parisi. A systematical improvement of the migdal recursion formula. *Nuclear Physics B*, 180(2):201–220, 1981.

[108] J. H. Meinke and A. A. Middleton. Linking physics and algorithms in the random-field Ising model. *eprint arXiv:cond-mat/0502471*, February 2005.

[109] G Hart MK Transtrum and P Qiu. Information topology identifies emergent model classes. arxiv:1409.6203, 2014.

[110] P Møller. Beam-residual gas interactions. Technical report, CERN, 1999.

[111] István Montvay and Peter Weisz. Numerical study of finite volume effects in the 4-dimensional Ising model. *Nuclear Physics B*, 290:327–354, 1987.

[112] J Möser. On invariant curves of area-preserving mappings of an annulus. *Nachr. Akad. Wiss. Göttingen, II*, pages 1–20, 1962.

[113] Jürgen Moser. *Lectures on Hamiltonian systems*. Number 81. American Mathematical Soc., 1968.

[114] Ryszard Mrugala, James D Nulton, J Christian Schön, and Peter Salamon. Statistical approach to the geometric structure of thermodynamics. *Physical Review A*, 41(6):3156, 1990.

[115] James Murdock. Hypernormal form theory: foundations and algorithms. *Journal of Differential Equations*, 205(2):424–465, 2004.

[116] James Murdock. *Normal forms and unfoldings for local dynamical systems*. Springer Science & Business Media, 2006.

[117] Th. Niemeyer and J. M. J. Van Leeuwen. Wilson theory for 2-dimensional ising spin systems. *Physica*, 71, 1974.

[118] Denjoe O'Connor and CR Stephens. Geometry the renormalization group and gravity. *arXiv preprint hep-th/9304095*, 1993.

[119] WP Orrick, B Nickel, AJ Guttmann, and JHH Perk. The susceptibility of the square lattice ising model: New developments. *Journal of Statistical Physics*, 102(3-4):795–841, 2001.

[120] GIORGIO Parisi and NICOLAS Sourlas. Random magnetic fields, supersymmetry, and negative dimensions. *Physical Review Letters*, 43(11):744, 1979.

[121] Andrea Pelissetto and Ettore Vicari. Renormalization-group flow and asymptotic behaviors at the Berezinskii-Kosterlitz-Thouless transitions. *Physical Review E*, 87(3):032105, 2013.

[122] O. Perkovic, K. Dahmen, and J. P. Sethna. Avalanches, Barkhausen noise, and plain old criticality. *Physical Review Letters*, 75:4528–4531, 1995.

[123] A Piwinski. Intra-beam scattering. In *Frontiers of Particle Beams*, pages 297–309. Springer, 1988.

[124] Henri Poincaré. Les nouvelles méthodes de la mécanique céleste. *Gauthier-Villars, Paris*, I-III, 1899.

[125] Hernando Quevedo. Geometrothermodynamics. *Journal of Mathematical Physics*, 48(1):013506, 2007.

[126] Katherine N Quinn, Francesco De Bernardis, Michael D Niemack, and James P Sethna. Visualizing theory space: Isometric embedding of probabilistic predictions, from the Ising model to the cosmic microwave background. *arXiv preprint arXiv:1709.02000*, 2017.

[127] S. Mahapatra R. Maity and T. Sarkar. Information geometry and the renormalization group. *Phys. Rev. E.*, 2015.

[128] Archishman Raju, Colin B Clement, Lorien X Hayden, Jaron P Kent-Dobias, Danilo B Liarte, D Rocklin, and James P Sethna. Renormalization group and normal form theory. *arXiv preprint arXiv:1706.00137*, 2017.

[129] Chandra K Raju. *The eleven pictures of time: The physics, philosophy, and politics of time beliefs*. Sage, 2003.

[130] CK Raju. *Time: towards a consistent theory*, volume 65. Springer Science & Business Media, 2013.

[131] David Rand, Stellan Ostlund, James Sethna, and Eric D Siggia. Universal transition from quasiperiodicity to chaos in dissipative systems. *Physical Review Letters*, 49(2):132, 1982.

[132] David Rand, Stellan Ostlund, James Sethna, and Eric D Siggia. Universal transition from quasiperiodicity to chaos in dissipative systems. *Physical Review Letters*, 49(2):132, 1982.

[133] C Reick. Universal corrections to parameter scaling in period-doubling systems: Multiple scaling and crossover. *Physical Review A*, 45(2):777, 1992.

[134] Peter Reimann and Peter Talkner. Invariant densities and escape rates for maps with weak gaussian noise. *New Trends in Kramers Reaction Rate Theory*, 11:143, 1995.

[135] H. Risken. *The Fokker-Planck Equation*. Springer-Verlag, 1984.

[136] Dominik Rogula. Large deformations of crystals, homotopy, and defects. In *Trends in applications of pure mathematics to mechanics*, pages 311–331, 1976.

[137] G. Ruppeiner. Thermodynamics: A Riemannian geometric model. *Phys. Rev. A*, 20, 1979.

[138] G. Ruppeiner. Riemannian geometry in thermodynamic fluctuation theory. *Rev. Mod. Phys.*, 1995.

[139] Jesús Salas. Exact finite-size-scaling corrections to the critical two-dimensional Ising model on a torus: Ii. Triangular and hexagonal lattices. *Journal of Physics A: Mathematical and General*, 35(8):1833, 2002.

[140] Jesús Salas and Alan D. Sokal. Logarithmic corrections and finite-size scaling in the two-dimensional 4-state Potts model. *Journal of Statistical Physics*, 88(3):567–615, 1997.

[141] Matthew Sands. *The physics of electron storage rings: an introduction*. Stanford Linear Accelerator Center Stanford, CA 94305, 1970.

[142] R Scharf. The Campbell-Baker-Hausdorff expansion for classical and quantum kicked dynamics. *Journal of Physics A: Mathematical and General*, 21(9):2007, 1988.

[143] Graziano Servizi and Ezio Todesco. Resonant normal forms and applications to beam dynamics. In *Chaotic Dynamics*, pages 379–388. Springer, 1992.

[144] Lev N. Shchur, Bertrand Berche, and Paolo Butera. Numerical revision of the universal amplitude ratios for the two-dimensional 4-state Potts model. *Nuclear Physics B*, 811(3):491 – 518, 2009.

[145] H Sonoda. Solving renormalization group equations with the Lambert W function. *Physical Review D*, 87(8):085023, 2013.

[146] Olaf Stenull and Hans-Karl Janssen. Logarithmic corrections to scaling in critical percolation and random resistor networks. *Phys. Rev. E*, 68:036129, Sep 2003.

[147] Robert H Swendsen. Optimization of real-space renormalization-group transformations. *Physical review letters*, 52(26):2321, 1984.

[148] G Toulouse and M Kléman. Principles of a classification of defects in ordered media. *Journal de Physique Lettres*, 37(6):149–151, 1976.

[149] MK Transtrum, BB Machta, and JP Sethna. Gometry of nonlinear least squares with applications to sloppy models and optimization. *Phys. Rev. E.*, 2011.

[150] FJ Wegner. Some invariance properties of the renormalization group. *Journal of Physics C: Solid State Physics*, 7(12):2098, 1974.

[151] Franz J Wegner. Corrections to scaling laws. *Physical Review B*, 5(11):4529, 1972.

[152] Franz J Wegner and Eberhard K Riedel. Logarithmic corrections to the molecular-field behavior of critical and tricritical systems. *Physical Review B*, 7(1):248, 1973.

[153] S Weinberg. Newtonianism, reductionism and the art of congressional testimony. *Nature*, 330, 1987.

[154] Steven Weinberg. Without god. *New York review of books*, pages 73–76, 2008.

[155] Stephen Wiggins. *Introduction to applied nonlinear dynamical systems and chaos*, volume 2. Springer Science & Business Media, 2003.

[156] Kenneth G Wilson and John Kogut. The renormalization group and the ϵ expansion. *Physics Reports*, 12(2):75–199, 1974.

[157] P Yu. Computation of the simplest normal forms with perturbation parameters based on Lie transform and rescaling. *Journal of computational and applied mathematics*, 144(1):359–373, 2002.

[158] Pei Yu and AYT Leung. The simplest normal form and its application to bifurcation control. *Chaos, Solitons & Fractals*, 33(3):845–863, 2007.

[159] MP Zorzano, H Mais, and L Vazquez. Numerical solution for Fokker-Planck equations in accelerators. *Physica D: Nonlinear Phenomena*, 113(2-4):379–381, 1998.