On Renormalization Methods in Dynamical Systems, Probability, and Statistical Physics

A Dissertation presented

by

Matthew Dannenberg

to

The Graduate School

in Partial Fulfillment of the

Requirements

for the Degree of

Doctor of Philosophy

in

Mathematics

Stony Brook University

August 2022

Stony Brook University

The Graduate School

Matthew Dannenberg

We, the dissertation committee for the above candidate for the
Doctor of Philosophy degree, hereby recommend
acceptance of this dissertation

Mikhail Lyubich - Dissertation Advisor
Distinguished Professor, Department of Mathematics

Dzmitry Dudko - Chairperson of Defense
Assistant Professor, Department of Mathematics

Marco Martens
Professor, Department of Mathematics

Pavel Bleher
Professor Emeritus, Department of Mathematical Sciences

This dissertation is accepted by the Graduate School

Celia Marshik
Interim Dean of the Graduate School

Abstract of the Dissertation

On Renormalization Methods in Dynamical Systems, Probability, and Statistical Physics

by

Matthew Dannenberg

Doctor of Philosophy

in

Mathematics

Stony Brook University

2022

Renormalization ideas have spread across many mathematical disciplines since their introduction in quantum electrodynamics. While manifestations of renormalization across physics and dynamical systems appear similar, direct connections between them are often mysterious. In this dissertation, we view some common features by analyzing several renormalization operators occurring in probability, statistical physics, and dynamical systems. In probability, we develop a self-contained renormalization proof of local Berry-Esseen type bounds for normalized sums of i.i.d. random variables. In statistical physics, we connect the real-space block-spin renormalization of Dyson hierarchical models explored by Bleher and Sinai to a block renormalization scheme for the corresponding Brydges-Fröhlich-Spencer random walk model representations. These in turn are connected to a block renormalization method for random walks on certain self-similar groups, which is distinguished from the Schur renormalization (Münchausen trick) defined by Bartholdi, Kaimanovich, and Virag. In dynamical systems, we use the renormalization theory for almost commuting pairs of holomorphic maps developed by Gaidashev, Goncharuk, Radu, and Yampolsky to construct Herman rings for highly dissipative two-dimensional holomorphic maps.

Table of Contents

List of Figures

Acknowledgements

I have been extremely fortunate over the past six years to have had the active support and companionship of a great multitude of people, and I am beyond grateful to all of them.

I would like in particular to thank my advisor, Mikhail Lyubich. His patience, concern, and kindness were as valuable to me as his great mathematical insight. I will always be grateful for the effort he took to find interesting projects for me to work on and to connect me to other collaborators. His guidance enabled me to find my way forward.

I would also like to thank my parents, Michael Dannenberg and Carol Dannenberg. They provided me a place of respite, where even in turbulent times I was always able to find shelter. I owe them a debt I can never fully repay.

I am extremely thankful for the friendship of many in the Stony Brook Mathematics Department community. Thank you especially to my housemates, Roberto Albesiano, Nathan Chen, Prithviraj Chowdhury, Deyasini Dasgupta, Jordan Rainone, and Alexandra Victorova. I am glad to have shared my time at graduate school with all of you.

Finally, I would like to especially thank my fiancée, Sarah Lichtman. She has helped me in so many ways; I cannot list them all. I will highlight in particular her companionship on many long walks in the woods, her tolerance for my late nights, her willingness to listen to me talk about abstract mathematical jargon, her availability as a pillar of support in my life, and, of course, her help in editing this thesis. I could not have finished my Ph.D. without her and I will always be grateful for her help.

Chapter 1

What is Renormalization?

Ernest Rutherford is said to have remarked that "All science is either physics or stamp collecting." We regard the mathematics of this thesis as doing some of both: organizing many distinct renormalization operators as one might organize a collection of stamps. Some of these operators will arise from blocking transformations for lattice spin systems, others from return maps of dynamical systems. Nonetheless, we assign the word 'renormalization' to each. It is natural to ask where this term came from, and what common features are present that merit connecting all of these objects together with this single word. In this thesis, we'll explore a few specific cases of renormalization in detail.

The remainder of Chapter 1 is dedicated to a brief history of the idea of renormalization and a basic underlying framework for renormalization which can connect to all of our operators of concern.

Chapter 2 uses a 'block-spin'-type renormalization argument to prove a case of the central limit Theorem, and gives a quantitative rate of convergence.

Chapter 3 concerns block-spin renormalization in the Dyson hierarchical model. We connect this to the behavior of random walks on a self-similar group, where both a block transformation and a return map renormalization are defined and shown to be distinct. The former is then shown to be in some sense equivalent to the renormalization of Dyson models,

demonstrating a case in which common renormalization paradigms in statistical physics and dynamical systems differ.

Chapter 4 is dedicated to the renormalization theories for circle diffeomorphisms and unicritical circle maps. We discuss a unifying perspective of almost-commuting pairs and use this framework to show the survival of Herman rings for highly-dissipative two-dimensional perturbations of one-dimensional analytic maps of $\mathbb{C}$.

1.1 A Brief History of Renormalization

In this section we will pursue a historical explanation as to how the word 'renormalization' came to appear in dynamical systems and statistical physics. We will briefly describe the origins of renormalization in quantum field theory, then will move to discuss how the concept was adapted to statistical physics and dynamical systems.

1.1.1 The Origin of Renormalization Ideas

An excellent overview of the early development of renormalization can be found in Chapters 1 through 3 of [Bro12], upon which this section is based, as well as in [Hua13]. The origin of renormalization itself is somewhat murky. The term originates in a 1936 paper by Robert Serber. Yet, in Chapter 2 of [Bro12], Max Dresden argues that some of the antecedent ideas that would become renormalization date back to the middle of the 19th century. To see a simple example of this sort of reasoning, we turn to fluid dynamics work by Stoke in the 1840s. If one considers a three-dimensional solid object with mass m_0 moving with velocity v through an incompressible three-dimensional fluid with density ρ, boundary conditions and the displacement of fluid material around the object imply that the total kinetic energy is given by

$$T = \frac{1}{2}\left(m_0 + m_*(\rho)\right) v^2,$$

where $m_*(\rho)$ is some 'additional effective mass' depending on the density of the fluid and the geometry of the object which is related to the quantity of fluid being displaced. This gives rise to a hydrodynamic property in which the mechanics of the object will be as if it was more massive than it truly is. In more modern language than was present at the time, we may call $m_0 + m_*(\rho)$ a 'renormalized mass.' Fifty years later, J. J. Thomson took this idea and extended it by analogy using Maxwell's equations to the case of a charged sphere moving through its own electric field. In this case as well, there is an additional 'electromagnetic mass' term which arises. For a uniformly charged sphere of radius r and charge e, this term is given by

$$m_* = \frac{2e^2}{3c^2 r}.$$

This electromagnetic mass is philosophically problematic. While in principle we can remove a solid object from a fluid and measure the true mass, there is no way to remove a charged particle from its own electric field so as to measure the 'true mass' without the electromagnetic mass term being present. As physics moved towards a theory of point particles, this also becomes extremely troubling since $m_* \to \infty$ as $r \to 0$, so that all point particles behave as if they have infinite mass and require an infinite quantity of energy to move.

These divergence problems followed electromagnetism into its quantum formulation as quantum electrodynamics. Hendrik Kramers is generally attributed with the introduction of renormalization to this field in a pair of talks made at the Shelter Island Conference in 1948 and the Solvay Conference in 1949, though his work focused on a non-relativistic case. We call the 'self-energy' the energy of a particle with specified structure of some radius, specified charge, and specified 'true' mass. This quantity generally diverges as the radius tends to 0. One may consider the difference in the self-energy between, say, a free particle and an otherwise identical particle placed in a potential well, for some specified positive radius. In Kramers' framing, both of these individual terms diverge as the radius shrinks but their difference tends to a limiting constant value. This allowed for the calculation of energy differences between different states even while formally the 'total energy' remained divergent.

Renormalization ideas proliferated in the following decade, with work by Tomonoga, Schwinger, Feynman, and Dyson (among many others) developing renormalization into a well-defined perturbative method which could produce asymptotic series approximating quantities of physical interest. This method may be analogized to a principal value integral. Calculating the self-energy of an electron produces an integral which is divergent, but converges if one introduces a cutoff scale $\Lambda > 0$ and considers the integral on the region a distance Λ away from 0. A very clear physical interpretation of this procedure is provided by Gell-Mann and Low in [GL54].

> A test body of "bare charge" q_0 polarizes the vacuum, surrounding itself by a neutral cloud of electrons and positrons; some of these, with a net charge δq, of the same sign as q_0, escape to infinity, leaving a net charge $-\delta q$ in the part of the cloud which is closely bound to the test body (within a distance $\frac{\hbar}{mc}$). If we observe the body from a distance much greater than $\frac{\hbar}{mc}$, we see an effective charge q equal to $(q_0 - \delta q)$, the renormalized charge. However, as we inspect more closely and penetrate through the cloud to the core of the test body, the charge that we see inside approaches the bare charge q_0, concentrated in a point at the center.

Renormalization may then be interpreted as an operation on the parameters of a model (the electron's charge, mass, etc.) which depends on the choice of the cutoff scale Λ. The set of such transformations was first referred to as a 'renormalization group' by Stueckelberg and Petermann in 1953 in [SP53].

1.1.2 The Renormalization Group in Statistical Physics

The connection between renormalization as a tool for eliminating divergences and the more modern perspective of the renormalization group is generally thought to originate in the statistical physics of lattice spin systems. In such systems, one considers an infinite discrete lattice of sites each containing a single particle. Each particle has a 'spin,' an associated

quantity of angular momentum. The interaction of nearby particles with each other gives rise to a complex set of statistical mechanical phenomena. In [Kad66], Leo Kadanoff developed the 'decimation' or 'block-spin' approach to analyzing Ising spin systems, where the spin variable is taken to be ± 1 and particles interact only with those particles directly adjacent to them in the lattice $\mathbb{Z}^d$ with strength proportional to some 'coupling constant' J. In this approach, one divides the lattice into 'blocks' which are L sites across in each direction. On each block, one considers the average spin present in all sites in the block as a 'block-spin'. The space of such blocks is isomorphic to the original lattice, but with a lattice spacing of L times the original lattice spacing.

Adjacent block-spins in this 'block-lattice' interact with one another with a more complicated Hamiltonian than is present for a pure Ising interaction. Yet, to a reasonable degree of approximation, the newly appearing terms may be removed and the block-spins reduced to taking values in ± 1 by altering the coupling constant to some $\hat{J}$. This decimation procedure then amounts purely to a map $J \mapsto \hat{J}$ along with a change in the length scale of the lattice by a factor of L. By building this transformation in such a way so as to preserve to the partition function of a spin model, Kadanoff could derive scaling formulas for the behavior of the thermodynamic functions near a critical point at which a phase transition occurs. The exponents in these scaling laws are often referred to as *critical exponents*.

Kenneth Wilson is most often credited with connecting Kadanoff's decimation transformation with the renormalization phenomenology of quantum field theory in [Wil71a; Wil71b]. Wilson sought to explain universal behavior of critical exponents. Namely, for large classes of qualitatively different spin models, the same power law behavior with the same critical exponents would occur near the critical point. Wilson analyzed the decimation transformation as a dynamical system in the space of all spin models: the renormalization transformation. He found that this dynamical system admits fixed points with a stable manifold of extremely low codimension. Universality classes of spin models could then be identified depending on which fixed point a critical system approached under repeated renormalization.

1.1.3 Renormalization in Dynamical Systems

An excellent survey of the early history of the theory of deterministic chaos and renormalization can be found in [AD02]. The introduction of renormalization methods to dynamical systems is usually attributed to Feigenbaum, as well as to Coullet and Tresser in the late 1970s. They sought to explain the period-doubling phenomena in the quadratic family. Namely, consider the family of maps

$$f_\lambda(x) = \lambda x(1-x),$$

for $\lambda \in (0,4)$. For λ extremely small, f_λ has a stable fixed point. As λ is increased past a sequence of values $\lambda_1, \lambda_2, \ldots$, the system repeatedly undergoes bifurcations and the period of the stable orbit repeatedly doubles. These bifurcation points rapidly converge to a limiting value, in particular with

$$\lim_{k\to\infty} \frac{\lambda_{k+1} - \lambda_k}{\lambda_{k+2} - \lambda_{k+1}} \equiv \delta = 4.669\ldots.$$

Moreover, this all remained true for many qualitatively similar families of unimodal maps of an interval, with the above limit converging to the same universal value. In [Fei78], Feigenbaum proposed a (non-rigorous) explanation for this quantitative universality by means of finding a fixed point of a nonlinear transformation on the class unimodal maps f of the form

$$f \mapsto \alpha f^2(\alpha^{-1}x),$$

for another universal constant α. Much as in Wilson's analysis in statistical mechanics, Feigenbaum could explain universal phenomenon by way of understanding the linearization of a fixed point. This motivated the naming of this nonlinear transformation as a renormalization operator. Rigorous justification followed, with a computer-assisted proof by Lanford in 1982 [Lan17]. Relationships between this phenomena and the thermodynamic formalism were explored in [VSK17].

Rigorous and conceptual proofs without the use of computer assistance have followed more gradually in limited cases. While Feigenbaum's initial renormalization idea seems to hold true for a very large class of unimodal maps, most subsequent proofs have relied upon

methods from holomorphic dynamics. In particular, this usually forces the criticality of f (the exponent α such that with c such that $f'(c) = 0$, $|f(x) - f(c)| \sim |x - c|^\alpha$ sufficiently near c) to be an even integer. Initial connections with holomorphic dynamics were developed by Douady and Hubbard in [DH85], developing the renormalization of quadratic-like germs. In [Sul88], Sullivan provided direct explanation for some renormalization phenomena using holomorphic methods and was able to rigorously prove the existence of a renormalization fixed point, with subsequent work by [McM96] establishing an exponential rate of convergence to this fixed point for infinitely renormalizable maps (a similar result with different methods was produced in by Avila and Lyubich in [AL11]). The Wilsonian picture for renormalization near this fixed point was then completed by Lyubich in [Lyu99], which demonstrated hyperbolicity of the fixed point with a one-dimensional unstable manifold. This picture for one-dimensional maps was extended to the dissipative Henon-like setting for higher-dimensional maps in [CLM05]. Outside of the holomorphic setting, results are more limited, though the existence of renormalization fixed points was shown by Martens in [Mar98].

Renormalization ideas have spread widely in dynamics following this initial period-doubling example. A distinctive renormalization picture for Lorenz maps is analyzed in [MW14]. A renormalization approach to KAM phenomena has been explored in a number of sources, including [KK08]. Renormalization ideas in the study of the dynamics of circle maps, both in the case of circle diffeomorphisms and of critical circle maps, will feature strongly in the discussion in Chapter 4 of this thesis.

1.2 A General Framework for Renormalization Groups

In contrast to the historical explanation of the previous section, this section will focus on a common mathematical structure which undergirds the renormalization operators of interest to this thesis. In general, there are a great variety of renormalization methods across mathematics and physics and we make no claims toward a general theory to unify them

beyond our scope. Our framework bears some similarity to tower constructions in dynamical systems (see Chapter 5 of [McM96] or Chapter II of [DV12]) and is inspired by the general discussion about renormalization groups in Chapter 3.4 of [Les98].

1.2.1 A Change of Scale and then Normalization

A renormalization group will ultimately be some form of group (or semigroup) action upon a space of 'systems' or 'fields' of some type. Intuitively, renormalization will act by taking a system of specified size which is normalized in some way, zooming in/out, and then 'normalizing' the obtained system. For our purposes, the groups in question are simple in nature.

Lemma 1.2.1. *If $(\Omega, \mathscr{F}, \mathbb{P})$ is a probability space, then the set of all real valued random variables τ on Ω which are $\mathbb{P}$-almost surely positive and finite form a group under pointwise multiplication. The identity is pointwise almost surely equal to 1. We'll denote this group by $G_{\mathbb{P}}$. We may also relax the positivity condition to non-negativity, obtaining the semi-group $G_{\mathbb{P},0}$.*

It's worth noting that one could generalize our discussion by allowing random variables taking values in arbitrary groups or algebras. We take our definitions using real-valued random variables simply because it is sufficient to encompass all of the cases of direct interest to this thesis and to distinguish renormalization group actions from generic group actions.

We first need to discuss the space that renormalization will act upon. The specific manner of space can vary considerably depending on context. Let $\mathscr{M}$ be some space of systems under consideration; and $\mathscr{U}$ some larger space in which $\mathscr{M}$ is embedded. Generally $\mathscr{M}$ will be a finite or infinite-dimensional Banach manifold, and $\mathscr{U}$ will be a larger such manifold, or a disjoint union thereof. Informally, $\mathscr{M}$ is the space of systems with some specified fixed 'scale' and $\mathscr{U}$ the space of all systems of all scales (or all smaller/larger scales).

To proceed, it is useful to have a slightly more formal way to discuss the idea of a scale. Here, we'll also introduce a probability measure $\mathbb{P}$ on some auxiliary probability space $(\Omega, \mathscr{F}, \mathbb{P})$. In many cases taking Ω to have cardinality 1 suffices. In fact, the only parts of this measure-theoretic structure which are actually required is the σ-ideal of null sets.

Definition 1.2.1. The *scale space* $\mathscr{S}(\mathscr{M}, \mathscr{U})$ over a triple $(\mathscr{M}, \mathscr{U}, \mathbb{P})$ where $\mathscr{M} \subseteq \mathscr{U}$ and $\mathbb{P}$ is a probability measure is the set

$$\mathscr{S}(\mathscr{M}, \mathscr{U}) = \mathscr{U} \times G_{\mathbb{P},0}.$$

We refer to elements of the form

$$\{(f, \tau) : f \in \mathscr{M}, \tau = 1 \text{ almost surely}\},$$

as *normalized*. A map $T : \mathscr{M} \to \mathscr{S}(\mathscr{M}, \mathscr{U})$ sending $f \mapsto (f, \tau_f)$ is a *scale*.

A scale T encodes information about how much 'zooming' one wishes to do on an element $f \in \mathscr{M}$. For example, in statistical physics a blocking transformation into blocks of size 2 corresponds to the constant scale $T_2 : f \mapsto (f, 2)$. On the other hand, in Feigenbaum's period-doubling renormalization, if the prerenormalizabion of a renormalizable f is defined on an interval I_f, the renormalization operator corresponds to the scale $T : f \mapsto (f, 2\chi_{I_f})$ for χ_E a characteristic function of a measurable set E.

We will in this chapter suppress the presence of $\mathbb{P}$ in notation, for simplicity. Renormalization operators will interact with the structure of this scale space in a natural way. In this interest, it is convenient to 'divide' a renormalization operator into a composition of two operators.

Definition 1.2.2. A *renormalization operator* or *renormalization group transformation* with scale T is a map $\mathscr{R}_T = N \circ R_T$, where for some open $M_T \subseteq \mathscr{M}$, $R_T : M_T \to \mathscr{U}$ is a *zooming transformation* with scale T and $N : \mathscr{U} \to \mathscr{M}$ is a *normalization transformation*. We say that an element of M_T is T-*renormalizable*. Given a renormalization operator or family thereof,

the *renormalization group* $\mathscr{G}$ is the group (if all renormalization operators are invertible) or semigroup of operators on $\mathscr{M}$ generated by the given operators under composition. We'll let $\mathscr{T}$ denote the set of scales present in a given renormalization group.

In addition, we require that

1. N is a projection onto $\mathscr{M}$, that is $N \circ N = N$ and $N|_{\mathscr{M}} = \mathrm{id}_{\mathscr{M}}$,
2. the scale T_1 which sends each $f \in \mathscr{M}$ to a normalized pair is in $\mathscr{T}$ and $R_{T_1} = \mathrm{id}_{\mathscr{M}}$,
3. $\mathscr{T} \subset G_{\mathbb{P}_0}$ is a subgroup (sub-semigroup),
4. if $T_1, T_2 \in \mathscr{T}$ and $f \in M_{T_1}$ and $\mathscr{R}_{T_1} f \in M_{T_2}$, then $f \in M_{T_1 \cdot T_2}$ and

$$\mathscr{R}_{T_1} \circ \mathscr{R}_{T_2} = \mathscr{R}_{T_1 \cdot T_2}.$$

Informally, a zooming transformation takes a system in $\mathscr{M}$ and 'zooms in' (or out) an amount prescribed by T, producing a system in $\mathscr{U}$ which is is now of a larger (smaller) scale. Normalization transformations then prune the largest (smallest) scales off of the element of $\mathcal{U}$ to place the element back inside $\mathscr{M}$.

The crux of the above definition is that the renormalization group is a partially defined multiplicative group (or semigroup) action. There are a great diversity of examples of this type of action. We'll discuss a few below. In general, we are be most interested in cases in which $\mathscr{G}$ is a discrete group (semigroup).

1.2.1.1 Scalar Functions on the Positive Real Axis

One of the simplest cases consists occurs when $\mathscr{U} = C(\mathbb{R}_{>0}, \mathbb{R}_{>0})$ and $\mathscr{M}$ is the subset with $f(1) = 1$.

Proposition 1.2.2. *For all $a > 0$ we may define a 'constant' scale $T_a : f \mapsto (f, a)$. Then let*

$$N(f)(x) = \frac{f(x)}{f(1)},$$

$$R_{T_a}(f)(x) = f(ax).$$

Then $\mathscr{G}_{\mathbb{R}_{>0}} = \{\mathscr{R}_{T_a} : a > 0\}$ *is a renormalization group.*

Inspired by dynamical systems and seeing a group action, it is very natural to ask after fixed points of the renormalization group. We will discuss this at length in a subsequent section, but it's worth considering for this example as well. The following proposition is a simple exercise.

Proposition 1.2.3. *A map* $f \in C(\mathbb{R}_{>0}, \mathbb{R}_{>0})$ *is a fixed point of* $\mathscr{R}_{T_a}$ *for all* $a > 0$ *if and only if* $f(x) = x^p$ *for some* $p \in \mathbb{R}$.

This basic example illustrates already a connection between fixed points of a renormalization group and power laws.

1.2.1.2 In Dynamical Systems

If one wishes to consider dynamical systems on a topological space X, then one natural choice is $\mathscr{M} = C(X, X)$ and

$$\mathscr{U} = \bigsqcup_{Y} C(Y, Y),$$

where the union is taken over all subsets $Y \subset X$ which are homeomorphic to X.

Definition 1.2.3. We say that $f \in \mathscr{M}$ is *renormalizable* if there exists a subset Y such that for almost every $x \in Y$ there exists $0 < \tau_{x,f} < \infty$ such that $f^{\tau_{x,f}}(x) \in Y$ and $f^t(x) \notin Y$ for any $0 < t < \tau_{x,f}$.

The choice of the first return time τ_x depends quite explicitly on f. Nonetheless, we would like to be able to classify maps depending on some qualitative properties of this first return time.

Definition 1.2.4. Two renormalizable maps $f, g \in \mathscr{M}$ with corresponding subsets Y_f, Y_g and first return times $\tau_{x,f}, \tau_{x,g}$ are *combinatorially equivalent at level 1* if there exists a homeomorphism $h : Y_f \to Y_g$ such that

$$\tau_{x,f} = \tau_{h(x),g}.$$

This is an equivalence relation.

We can now define return map renormalization.

Definition 1.2.5. Let $M \subset \mathscr{M}$ be an equivalence class of renormalizable maps which are all combinatorially equivalent at level 1. Consider the scale $T_M(f)(x) = (f, \tau_{x,f}\chi_{Y_f}(x))$ and let

$$p\mathscr{R}_{T_M}(f)(x) \equiv R_{T_M}(f)(x) = f^{\tau_{x,f}}(x),$$

for all $x \in Y_f$ and $f \in M$. In dynamical systems, this zooming transformation is commonly called *pre-renormalization*. Further, for any $g \in \mathscr{U}$ which is a map $g : Y \to Y$, one may take

$$N(g) = h(g) \circ g \circ h(g)^{-1},$$

where $h(g) : Y \to X$ is a homeomorphism. Then we call

$$\mathscr{R}_{T_M} = N \circ p\mathscr{R}_{T_M}$$

a *return map renormalization operator*.

In the study of dynamical systems of the interval, one may take $h(g)$ to depend only on the domain Y and in particular to be affine (see Chapter II of [DV12], for example). In other settings, such as the cylinder renormalization of [Yam03], $h(g)$ depends explicitly on g.

The renormalization operators defined in Chapter 4 are all broadly of this type, though of course much stronger conditions than mere continuity are imposed on the elements of $\mathscr{M}$.

One can also generalize this idea to return maps for stochastic processes, using the corresponding probability measure for the process and taking the scale as the return time to a given set.

1.2.1.3 In Probability

For the central limit theorem renormalization operator discussed in Chapter 2, one can take $\mathscr{U}$ a space of all probability laws on $\mathbb{R}$ with finite mean and variance and $\mathscr{M}$ all such laws with zero mean and unit variance.

Proposition 1.2.4. *If $\nu \in \mathscr{U}$ has mean μ and variance σ, let $\{X_i\}_{i\in\mathbb{N}}$ be a sequence of independent and identically distributed (i.i.d.) random variables on $\mathbb{R}$ distributed with law ν. Then take $N(\nu)$ the law of the random variable*

$$\frac{X_1 - \mu}{\sqrt{\sigma}}.$$

Moreover, for any integer n one may define a scale $T_n(X) = (X, n)$ so that for any $\nu \in \mathscr{M}$ the zooming $R_{T_n}(\nu)$ is the law of the random variable

$$\sum_{i=1}^{n} X_i.$$

Then $\mathscr{G} = \{\mathscr{R}_{T_n}\}_{n\in\mathbb{N}}$ is a renormalization group, as is $\mathscr{G}' = \{\mathscr{R}_{T_{2^n}}\}_{n\in\mathbb{N}}$.

1.2.1.4 In the Statistical Physics of Spin Systems

The statistical physics content is discussed in significantly more depth in Chapter 3. We'll use some terminology from that chapter here, for the sake of brevity here. It suffices for our purposes here to note that we may let $\mathscr{U}$ be the space of all spin models (H_f, ν) on a lattice L (in Chapter 3 we take $L = \bigoplus_{\mathbb{N}} \mathbb{Z}_2$), with isotropic real-valued a priori measures ν and having a formal Hamiltonian of the form

$$H_f(\varphi) = -\sum_{x\neq y\in L} f(d(x,y))\varphi_x\varphi_y,$$

for some decreasing $f : \mathbb{N} \to \mathbb{R}_{>0}$ and with $d : L \times L \to \mathbb{R}_{\geq 0}$ an ultrametric having a self-similar hierarchical (tree-like) structure with

$$\inf_{x\neq y\in L} d(x,y) = C > 0,$$

with the metric only taking values $\{0, 1, s, s^2, s^3, \ldots\}$ for some $s > 1$. In this context φ is a real scalar field on L. Then $\mathscr{M}$ can be taken to be the subset of those spin models for which $f(1) = 1$.

We'll make one additional assumption, that we may choose some sublattice $L' \subsetneq L$ isomorphic to L with finite index n (we take $n = 2$ in Chapter 3) and construct a L'-translation-invariant partition $\{B_j\}_{j\in L'}$ of L into *blocks*, such that there exists some constant

$A > 0$ so that $d(x, y) < A$ if and only if x, y are in the same block. Intuitively, this 'block-spin' approach to renormalization will involve combining all the individual spin variables in a block into a single aggregate spin variable.

Definition 1.2.6. Let $(H_f, \nu) \in \mathscr{U}$. Define $N(H_f, \nu) = (H_{\hat{f}}, \hat{\nu})$ where for $m \in \mathbb{N}$ and $E \subset \mathbb{R}$ a Borel set

$$\hat{f}(m) = \frac{f(m)}{f(1)},$$
$$\hat{\nu}(E) = \nu\left(\sqrt{f(1)}E\right).$$

For $(H_f, \nu) \in \mathscr{M}$, define the scale $T_n(H_f, \nu) = ((H_f, \nu), n)$ and let $R_{T_n}(H_f, \nu) = (H_{\tilde{f}}, \tilde{\nu})$ where for $m \in \mathbb{N}$

$$\tilde{f}(m) = f(sm),$$

and $\tilde{\nu}$ is a new a priori measure for the real scalar field $\tilde{\varphi}$ on L' defined by

$$\tilde{\varphi}_j = \sum_{x \in B_j \subset L} \varphi_x.$$

An explicit equation for $\tilde{\nu}$ can be written as a particular weighted convolution (and is done so in Chapter 3). We refer to $\mathscr{R}_{T_n} = N \circ R_{T_n}$ as a *block-spin renormalization operator*.

1.2.2 Fixed Points and Local Linearization

Given a renormalization group $\mathscr{G}$ acting on a Banach manifold $\mathscr{M}$, it is natural to ask about dynamical behavior. In general, renormalization operators tend to be nonlinear operators on infinite-dimensional spaces, so their behavior can be extremely complicated. The simplest case to analyze occurs when there is a fixed point $f_* \in \mathscr{M}$ of the group action.

As we regard renormalization as a normalized rescaling, f_* is a fixed point if, when viewed at different scales, f_* looks self-similar up to normalization. In other words, f_* is a fixed point of a renormalization group when f_* is *scale-invariant*. This perspective is discussed in detail in Chapter 3.1 of [Les98] and Chapter 9.3 of [Gol18].

For the sake of clarity, we'll suppose now that $\mathscr{G}$ is generated by a single operator $\mathscr{R}$. The cleanest case of this fixed point theory occurs when $\mathscr{R}$ is Frechet differentiable in a neighborhood of f_* and $D\mathscr{R}|_{Tf_*}$ is a compact hyperbolic operator. In such a case, the tangent space $Tf_* = E_s \oplus E_u$ admits a splitting into a stable direction and an unstable direction (in physics these are sometimes called the *irrelevant direction* and the *relevant direction*). Moreover, there exist immersed stable and unstable manifolds $W^s_{\mathscr{R}}(f_*)$ and $W^u_{\mathscr{R}}(f_*)$ (see Chapter 5 of [Shu13]) with $W^s_{\mathscr{R}}(f_*)$ of finite codimension and $W^u_{\mathscr{R}}(f_*)$ of finite dimension.

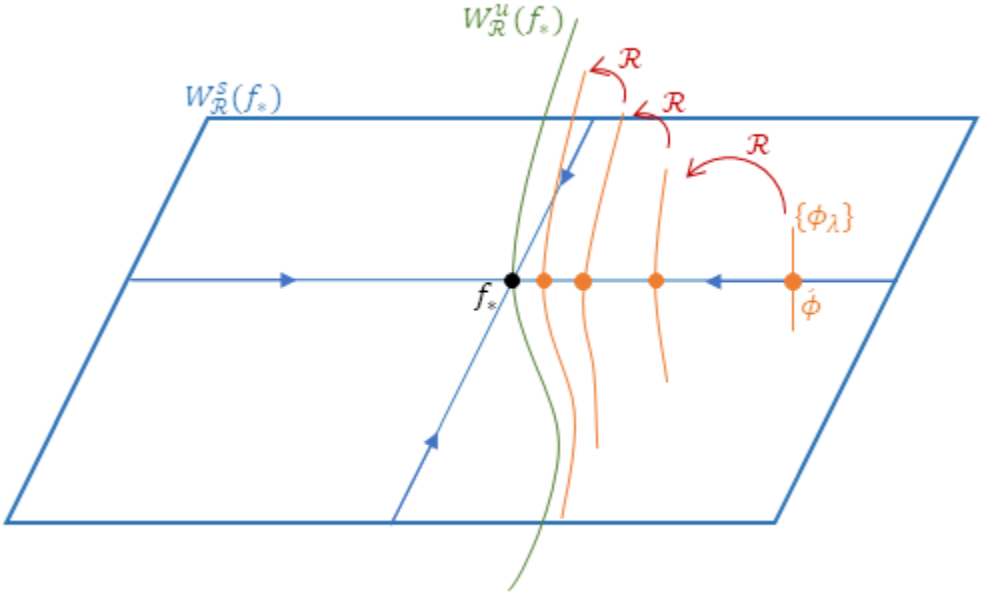

Figure 1.1: The Wilsonian Picture for Renormalization Near a Fixed Point

If $f \in W^s_{\mathscr{R}}(f_*)$, then f is renormalizable infinitely many times and $\mathscr{R}^n(f) \to f_*$ at an exponential rate as $n \to \infty$. In such a sense, if we regard renormalization as a normalized 'zooming in', then on very small scales $\mathscr{R}^n f$ 'looks like' f_*.

On the other hand, if $g \in W^u_{\mathscr{R}}(f_*)$, then there exist infinitely many *anti-renormalizations* of g. That is to say there exists some (maybe not unique) $g_{-1} \in W^u_{\mathscr{R}}(g_*)$ such that $\mathscr{R}g_{-1} = g$. Iterating this to obtain a sequence $\{g_{-n}\}_{n\in\mathbb{N}}$, then $g_{-n} \to f_*$ at an exponential rate as $n \to \infty$.

Since $W^s_{\mathscr{R}}(f_*)$ is of finite codimension, say k, generic k-dimensional families $\{\phi_\lambda\}_{\lambda\in U\subset\mathbb{R}^k}$ in

$\mathscr{M}$ may transversely intersect it at a (locally) unique point $\hat{\phi}$ in the family . Under iteration of renormalization, $\hat{\phi}$ will tend to f_*, dragging nearby points in the family with it until they are pulled away by the unstable behavior of $\mathscr{R}$. Indeed, use of the inclination lemma would suggest that small neighborhoods of $\hat{\phi}$ will converge onto $W^u_{\mathscr{R}}(f_*)$, giving rise to what is often described as 'universal' behavior under rescaling.

Chapter 2

A Renormalization Argument for The Central Limit Theorem

The central limit theorem is a foundational result in probability describing the asymptotic behavior of normalized sums of weakly dependent random variables. In this chapter we pursue a renormalization group approach to proving a quantitative rate of convergence for the central limit theorem in a very fine topology for independent, identically distributed (i.i.d.) random variables satisfying strong moment conditions in a neighborhood of Gaussian distributions. This will yield bounds of Berry-Esseen type. This method need not assume that distributions are even, and does not rely upon the Fourier transform.

2.0.1 Notation

We will take a sequence of independent, identically distributed random variables $\{X_i\}_{i\in\mathbb{N}}$ on $\mathbb{R}$ having mean 0 and unit variance, each distributed according to the law μ, with μ absolutely continuous with respect to Lebesgue measure, having density $p : \mathbb{R} \to \mathbb{R}_{\geq 0}$. We call the space of such probability densities $\mathscr{D}$. We'll write the expectation of some measurable function f with respect to μ as

$$\mathbb{E}_p[f(X)] \equiv \int_{\mathbb{R}} f(x)p(x)dx.$$

We'll write $L^r(\mathbb{R})$ for $1 \leq r \leq \infty$ to denote the L^r space on $\mathbb{R}$ with respect to Lebesgue measure, and $L^r(f)$ for any integrable $f : \mathbb{R} \to \mathbb{R}_{\geq 0}$ to denote the L^r space on $\mathbb{R}$ with respect to the measure $d\nu(x) = f(x)dx$.

2.1 The Renormalization Operator

The central idea of a renormalization analysis of the central limit theorem is to view the normalized summation of random variables as arising from a (nonlinear) operator on $\mathscr{D}$. Namely, we may define the *renormalized averages*

$$Y_n^1 = 2^{-\frac{n}{2}} \sum_{i=1}^{2^n} X_i,$$
$$Y_n^2 = 2^{-\frac{n}{2}} \sum_{i=2^n+1}^{2^{n+1}} X_i.$$

Notably, this implies the recurrence

$$Y_{n+1}^1 = \frac{1}{\sqrt{2}}(Y_n^1 + Y_n^2).$$

That all X_n are independent implies that Y_n^1 and Y_n^2 are independent. Hence, if we take p_n to be the probability density of Y_n^1 with respect to Lebesgue measure, then $p_0 \equiv p$ and

$$p_{n+1}(x) \equiv \mathscr{R}(p_n)(x) \equiv \sqrt{2} \int_{\mathbb{R}} p_n\left(\frac{x}{\sqrt{2}} + u\right) p_n\left(\frac{x}{\sqrt{2}} - u\right) du. \tag{2.1.1}$$

Definition 2.1.1. The operator $\mathscr{R} : \mathscr{D} \to \mathscr{D}$ is the *renormalization operator*.

As a result, we may understand the asymptotic behavior of dyadic summations of i.i.d. random variables by examining the asymptotic properties of this operator under iteration. Then, one can interpret the central limit theorem as a statement about the basin of attraction of fixed points of $\mathscr{R}$, obtaining a rate of convergence in some neighborhood of such a fixed point in terms of the linearized behavior of $\mathscr{R}$ at that fixed point.

Renormalization ideas have been applied to the central limit theorem for some time, with rigorous arguments dating back at least to the work of Bleher and Sinai [BS73] on Dyson

hierarchical models, which can be applied directly to i.i.d. random variables satisfying very strong conditions. These ideas were connected to field-theoretic renormalization arguments in [Jon75]. Sweeping analyses of the fixed points of several renormalization group actions in probability were carried out using the Fourier transform in [Sin76], pioneering the rigorous study of self-similar processes using renormalization methods. Much recent effort has focused on other probabilistic limit laws, including but not limited to convergence results to stable limit laws in [LS09; LS14], to generalized central limit theorems in [Cal+10; Ami20], and to Wigner's semicircle law in random matrix theory in [Kra17].

2.1.1 Fixed Points of the Renormalization Operator

Here we explore the most basic dynamical properties of the renormalization operator on $\mathscr{D}$: its fixed points.

Lemma 2.1.1. *$\mathscr{R}$ has a unique fixed point in $\mathscr{D}$, given by the normal distribution*

$$g_{0,1}(x) \equiv g(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}.$$

The uniqueness of this fixed point is an immediate consequence of the observation that a fixed point of $\mathscr{R}$ is a stable distribution, since then we must have $\frac{1}{\sqrt{2}}(X_1 + X_2)$ equal in distribution to X_1. The only stable distributions with finite variance are normal distributions, see [Zol86], which are all fixed points of $\mathscr{R}$ if they have mean 0. Fixing the variance yields a unique fixed point in $\mathscr{D}$.

2.1.2 Linearization of the Renormalization Operator

The renormalization operator $\mathscr{R}$ is a quadratic operator, and as such has a fairly unambiguous linearization. Nonetheless, the precise choice of topology on which to analyze $\mathscr{R}$ and show differentiability can be fairly subtle. The desired topology may vary depending on which properties are needed in any given case. Here we'll present one useful topology. We can extend the domain of definition of $\mathscr{R}$ to $L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ without issue.

Lemma 2.1.2. *At any $p \in \mathscr{D}$, $\mathscr{R}$ is Frechet differentiable in $L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ with*

$$D\mathscr{R}|_p(v) = 2\sqrt{2} \int_{\mathbb{R}} v\left(\frac{x}{\sqrt{2}} + u\right) p\left(\frac{x}{\sqrt{2}} - u\right) du,$$

for any $v \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$.

Proof. Observe that

$$\mathscr{R}(p+v) = \mathscr{R}(q) + \mathscr{R}(v) + 2\sqrt{2} \int_{\mathbb{R}} v\left(\frac{x}{\sqrt{2}} + u\right) p\left(\frac{x}{\sqrt{2}} - u\right) du.$$

Moreover,

$$\|\mathscr{R}(v)\|_{L^1(\mathbb{R})} \leq \|v\|^2_{L^1(\mathbb{R})},$$

and

$$\|\mathscr{R}(v)\|_{L^2(\mathbb{R})} \leq \|v\|_{L^2(\mathbb{R})} \|v\|_{L^1(\mathbb{R})},$$

by Young's inequality. The result follows. □

It's worth noting that convergence of probability densities to a Gaussian density in either $L^1(\mathbb{R})$ or $L^2(\mathbb{R})$ can be established by Stein's method, see for example [CGS11; NP12].

2.2 Berry-Esseen Type Bounds for the Central Limit Theorem

In this section, we'll analyze the renormalization operator $\mathscr{R}$ to obtain a local Berry-Esseen type theorem with respect to a fine topology near a Gaussian distribution. Our method mirrors most closely that of [LS14] with key differences in that our proof does not make use of the Fourier transform and holds for distributions which are not even.

2.2.1 The Classical Berry-Esseen Theorem

While the central limit theorem on its own provides conditions for convergence of normalized sums of i.i.d. random variables to a Gaussian distribution, it does not provide any direct

information about the rate of convergence. Such quantitative control is given by the Berry-Esseen theorem, dating back to [Ber41; Ess42] in the Kolmogorov-Smirnov metric. A modern statement is given below.

Theorem 2.2.1 (Theorem 3.6 in [CGS11])**.** *There exists a $C > 0$ such that any i.i.d. sequence $\{X_i\}_{i\in\mathbb{N}}$ of random variables with finite third absolute moment, the normalized sums*

$$S_k = \frac{1}{\sqrt{k}} \sum_{i=1}^{k} X_i,$$

satisfy

$$\sup_{x\in\mathbb{R}} \left| \mathbb{P}(S_k \leq x) - \int_{-\infty}^{x} g(x)dx \right| < \frac{C\mathbb{E}\left[|X_0|^3\right]}{\sqrt{k}}.$$

In general this asymptotic rate of convergence in k may be sharp, see for example [Jir20]. The sup-norm on cumulative distribution functions shown in this theorem is the Kolmogorov-Smirnov metric.

Definition 2.2.1. The *Kolmogorov-Smirnov metric* for two probability densities $p, q : \mathbb{R} \to \mathbb{R}_{\geq 0}$ is given by

$$d_{KS}(p, q) = \sup_{x\in\mathbb{R}} \left| \int_{-\infty}^{x} \left[p(x) - q(x)\right] dx \right|.$$

Finally, it's worth observing that if we have a dyadic sum so that $k = 2^n$, the Berry-Esseen bound states that, adopting the notation of (2.1.1)

$$d_{KS}(p_n, g) \in \mathcal{O}\left(2^{-\frac{n}{2}}\right).$$

2.2.2 Identifying a Relevant Topology

Our analysis of the dynamics of the renormalization operator will focus on a local neighborhood of g. In this neighborhood, we'll define $q : \mathbb{R} \to \mathbb{R}$ so that $p = (1 + q)g$. We will be concerned with the behavior of renormalization when $q \in L^2(g)$. In particular, we may now gain information about the convergence of p to g by observing the manner in which q converges to 0.

This space is a simple case of a Hilbert space over a Gaussian measure. While we will focus on probability distributions on $\mathbb{R}$, much of this discussion can be generalized to other such Hilbert spaces over Gaussian measures, following the basic theory laid out in [Pra06].

The renormalization operator $\mathscr{R} : p \mapsto \tilde{p}$ induces a transformation on q. We'll split that transformation into a 'linear part' T and a 'nonlinear part' S.

Definition 2.2.2. For $q \in L^2(g)$, we define

$$T(q)(x) \equiv \frac{2}{\sqrt{\pi}} \int_{\mathbb{R}} e^{-u^2} q\left(\frac{x}{\sqrt{2}} + u\right) du,$$
$$S(q)(x) \equiv \frac{1}{\sqrt{\pi}} \int_{\mathbb{R}} e^{-u^2} q\left(\frac{x}{\sqrt{2}} + u\right) q\left(\frac{x}{\sqrt{2}} - u\right) du.$$

These operators are defined so that the diagram

$$\begin{array}{ccc} p & \xmapsto{\mathscr{R}} & \tilde{p} \\ \downarrow & & \downarrow \\ q & \xmapsto{T+S} & \tilde{q} \end{array}$$

commutes, with $\tilde{p} = (1 + \tilde{q})g$.

The following lemma notes that our identification of T as the 'linear part' of $\mathscr{R}$ is consistent with the differential found in Lemma 2.1.2.

Lemma 2.2.2. *If $q \in L^2(g)$, then $p \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ and one has $T(q)g = D\mathscr{R}|_g(p)$.*

Proof. Containment of p in $L^1(\mathbb{R})$ follows from the observation that $L^1(g) \supset L^2(g)$ and $\|p\|_{L^1(\mathbb{R})} = \|1 + q\|_{L^1(g)}$. Further, note that

$$\|p\|_{L^2(\mathbb{R})} = \int_{\mathbb{R}} |1 + q(x)|^2 g(x)^2 dx.$$

Since $g(x)^2 \leq \frac{1}{\sqrt{2\pi}} g(x)$, $\|p\|_{L^2(\mathbb{R})} < \infty$ if $q \in L^2(g)$. □

Having $q \in L^2(g)$ is an extremely strong assumption on the behavior of the tails of probability density.

Definition 2.2.3. A probability density p is *sub-Gaussian* if there exists constants $C, b > 0$ such that for any $t \in \mathbb{R}$

$$1 - \int_{-t}^{t} p(x)dx < Ce^{-bt^2}.$$

Equivalently, p is *sub-Gaussian* if $\mathbb{E}_p[e^{aX^2}] < \infty$ for some $a > 0$.

Restricting to $q \in L^2(g)$ only allows us to consider distributions which are sub-Gaussian.

Lemma 2.2.3. *If $q \in L^2(g)$, then p is sub-Gaussian.*

Proof. Note that for $a < \frac{1}{4}$, the function $x \mapsto e^{ax^2}$ is in $L^2(g)$. Thus, by the Cauchy-Schwarz inequality we have

$$\mathbb{E}_p\left[e^{aX^2}\right] = \int_{\mathbb{R}} e^{ax^2}(1 + q(x))g(x)dx < \infty.$$

□

This lemma has an immediate corollary.

Corollary 2.2.4. *If $q \in L^2(g)$, then the moment generating function of p, $E_p[e^{\lambda X}]$ is finite for any $\lambda \in \mathbb{R}$. In particular, all moments of p are finite.*

Finally, it is worth noting that the assumption that $p \in \mathscr{D}$ implies certain behavior of q.

Lemma 2.2.5. *If $p \in \mathscr{D}$ is a probability density, then*

$$\int_{\mathbb{R}} x^i q(x)g(x)dx = 0$$

for $i = 0, 1, 2$. Denote the subspace of $L^2(g)$ of functions satisfying this condition as $\mathcal{H}$.

2.2.3 Hermite Polynomials

In order to proceed with a discussion of the space $L^2(g)$, it is necessary to introduce Hermite polynomials. These play a fundamental role in the behavior of L^p spaces with respect to a Gaussian measure, as shown in [Pra06].

Definition 2.2.4. The *probabilist's Hermite polynomials* $\mathbf{H}_n$ for $n \in \mathbb{N} \cup \{0\}$ are defined by the Rodrigues' formula (see page 785 of [AS64] or section 6.1 of [AAR99])

$$\mathbf{H}_n(x) = (-1)^n e^{\frac{x^2}{2}} \frac{d^n}{dx^n} e^{-\frac{x^2}{2}}.$$

Defined as so, $\mathbf{H}_n$ is monic of degree n. Hermite polynomials have a rich algebraic structure detailed in Chapter 4.2.1 of [Rom05].

Lemma 2.2.6. *The probabilist's Hermite polynomials* $\mathbf{H}_n$ *form an Appell sequence. That is to say that* $\mathbf{H}_n' = n\mathbf{H}_{n-1}$ *for any* $n \in \mathbb{N}$.

We may define a more general family of Hermite polynomials to explain the consequences of this property.

Definition 2.2.5. The Hermite polynomials with variance $\nu > 0$ are defined for $n \in \mathbb{N} \cup \{0\}$ by

$$\mathbf{H}_n^{(\nu)}(x) = \nu^{\frac{n}{2}} \mathbf{H}_n\left(\frac{x}{\sqrt{\nu}}\right).$$

As a consequence of Lemma 2.2.6, we obtain a family of *umbral identities* [Rom05] (the term *umbral* seems to refer to the 'mysterious' or 'shadowy' nature of this manner of identity). In a sense, these identities may allow us to regard Hermite polynomials as 'monomial-like' insofar as they satisfy a kind of binomial formula.

Lemma 2.2.7. *For any* $\nu, \eta > 0$, $n \in \mathbb{N} \cup \{0\}$, *and* $x, y \in \mathbb{R}$, *one has*

$$\mathbf{H}_n^{(\nu+\eta)}(x+y) = \sum_{k=0}^{n} \binom{n}{k} \mathbf{H}_k^{(\nu)}(x) \mathbf{H}_{n-k}^{(\eta)}(y).$$

We will prove the specific case of an umbral identity with $\nu = \eta$ by a different method in the next section.

Being polynomials, Hermite functions are all elements of $L^2(g)$. We are most interested in them in this context, and so will subsequently refer most often to normalized Hermite polynomials with unit variance in $L^2(g)$.

Definition 2.2.6. The (normalized probabilist's) *Hermite polynomials* H_n for $n \in \mathbb{N} \cup \{0\}$ are defined by

$$H_n(x) = \frac{1}{\sqrt{n!}} \mathbf{H}_n(x).$$

The following lemma is a standard result for Hermite polynomials and can be found in [AAR99] or [Pra06].

Lemma 2.2.8. *The Hermite polynomials H_n form a complete orthonormal basis for $L^2(g)$.*

Notably, the space $\mathcal{H}$ then consists of q orthogonal to H_0, H_1, H_2, giving the following corollary.

Corollary 2.2.9. *$\mathcal{H}$ is a closed subspace of $L^2(g)$.*

2.2.4 Sobolev Spaces

Our proof will establish the convergence $q_n \to 0$ in $L^2(g)$. We will, in fact, be able to make a stronger statement, allowing for derivatives of q_n to converge to 0 as well. In this interest, it will be useful to define Sobolev spaces.

Definition 2.2.7. For any $s \geq 0$, the *s-th Sobolev norm* of $f = \sum_{n=0}^{\infty} \alpha_n H_n \in L^2(g)$ is

$$\|f\|_{W^s(g)} = \left(\sum_{n=0}^{\infty} (n+1)^s |\alpha_n|^2 \right)^{\frac{1}{2}}.$$

The subspace of $L^2(g)$ consisting of those f for which $\|f\|_{W^s(g)} < \infty$ equipped with this norm is the *Sobolev space* denoted $W^s(g)$.

Due to the Appell sequence property shown in Lemma 2.2.6, if $s \in \mathbb{N} \cup \{0\}$ this is an equivalent norm with the standard definition of a Sobolev space as in [Pra06]. The following lemma indicates this equivalence of spaces.

Lemma 2.2.10. *If $s \in \mathbb{N} \cup \{0\}$, then $q \in W^s(g)$ if and only if all weak derivatives of q up to order s are defined and are in $L^2(g)$.*

For non-integer s the $W^s(g)$ form natural interpolation spaces between the integer Sobolev spaces. Those q corresponding to probability densities in $\mathscr{D}$ form closed subspaces of each Sobolev space.

Lemma 2.2.11. *For any $s \geq 0$, $\mathcal{H} \cap W^s(g)$ is a closed subspace of $W^s(g)$.*

2.2.5 The Linear Term

Here we'll analyze the behavior of T. We'll also produce a proof of a special case of the umbral identity for normalized Hermite polynomials. To do this, we'll introduce some additional operators.

Definition 2.2.8. Define for all $m \in \mathbb{N} \cup \{0\}$ and any $f \in L^2(g)$

$$T_m(f)(x) = \int_{\mathbb{R}} g(u) H_m(u) f\left(\frac{1}{\sqrt{2}}(x+u)\right) du.$$

Observe that $T = 2T_0$ by a simple change of variables. The following proposition describes the behavior of these operators.

Proposition 2.2.12. *The operator T_0 is a compact, self-adjoint operator from $L^2(g)$ to itself. It has eigenvalues*

$$\lambda_n = 2^{-\frac{n}{2}}.$$

with corresponding orthogonal eigenfunctions H_n. T_m behaves akin to a one-sided shift, satisfying

$$T_m(H_n) = \begin{cases} 2^{-\frac{n}{2}} \sqrt{\binom{n}{m}} H_{n-m} & m \leq n \\ 0 & m > n. \end{cases}$$

Proof. Suppose that f is continuously differentiable and note that

$$\frac{d}{dx} T_m(f)(x) = \frac{1}{\sqrt{2}} T_m(f')(x),$$

or alternatively by making a change of variables and differentiating the other term in the convolution

$$\frac{d}{dx}T_m(f)(x) = \sqrt{m+1}T_{m+1}(f)(x).$$

Now, we note that if $s(x) = xf(x)$, then

$$T_0(s)(x) - \frac{1}{\sqrt{2}}T_1(f)(x) = \frac{x}{\sqrt{2}}T_0(f).$$

As a result, we may note that if $L(f)(x) = \left(x - \frac{d}{dx}\right) f(x)$, then

$$\frac{1}{\sqrt{2}}T_0 \circ L = L \circ T_0.$$

Now, observe that $T_0(1) = 1$, so we inductively conclude that $T_0(H_n) = 2^{-\frac{n}{2}}H_n$. Another induction procedure gives that

$$\begin{aligned} T_m(H_n) &= \frac{1}{(\sqrt{m})!}\frac{d^m}{dx^m}T_0(H_n), \\ &= \frac{2^{-\frac{n}{2}}}{(\sqrt{m})!}\frac{d^m}{dx^m}H_n. \end{aligned}$$

This implies that $T_m(H_n) = 0$ if $m > n$, and otherwise as desired

$$T_m(H_n) = 2^{-\frac{n}{2}}\sqrt{\binom{n}{m}}H_{n-m}.$$

□

With this in hand, we understand the behavior of the linear part of the renormalization transformation.

Corollary 2.2.13. *The operator T is a compact, self-adjoint contraction on $\mathcal{H}$, with a unique fixed point at 0. It has eigenvalues*

$$\lambda_n = 2^{1-\frac{n}{2}}.$$

with corresponding eigenfunctions H_n.

Going further, we may use the found behavior of all of the T_m operators to obtain the following umbral identity.

Lemma 2.2.14. *For $x, y \in \mathbb{R}$ and $n \in \mathbb{N} \cup \{0\}$, one has*

$$H_n(x+y) = 2^{-\frac{n}{2}} \sum_{k=0}^{n} \sqrt{\binom{n}{k}} H_{n-k}(\sqrt{2}x) H_k(\sqrt{2}y)$$

Proof. Since $H_n(x+y)$ is a polynomial in y of degree n, we may expand it as

$$H_n(x+y) = \sum_{k=0}^{n} \alpha_{k,n}(\sqrt{2}x) H_k\left(\sqrt{2}y\right),$$

for some functions $\alpha_{k,n} : \mathbb{R} \to \mathbb{R}$. If we apply T_m to H_n this implies

$$T_m(H_n)(x) = \sum_{k=0}^{n} \alpha_{k,n}(x) \int_{\mathbb{R}} g(u) H_m(u) H_k(u) du.$$

By orthonormality, only the $k = m$ term contributes. Using Proposition 2.2.12 we thus have

$$\alpha_{m,n}(x) = 2^{-\frac{n}{2}} \sqrt{\binom{n}{m}} H_{n-m}(x).$$

□

2.2.6 The Nonlinear Term

Having analyzed the behavior of the linear part, we'd like now to examine the nonlinear term. In order to do so, we'll need to make use of a classical formula for the product of Hermite polynomials, which comes after normalization and rescaling from Theorem 6.8.1 in [AAR99].

Lemma 2.2.15. *If $m \geq n \in \mathbb{N} \cup \{0\}$, then*

$$H_m(x) H_n(x) = \sqrt{\frac{n!}{m!}} \sum_{r=0}^{n} \binom{m}{n-r} \frac{\sqrt{(m-n+2r)!}}{r!} H_{m-n+2r}(x).$$

With this in mind, we may explore the behavior of S. Noting that S is a quadratic operator, we'll denote by $S(f,g)$ the symmetric bilinear polarization of S applied to $f, g : \mathbb{R} \to \mathbb{R}$, so that $S(f,f) = S(f)$.

Lemma 2.2.16. *One has*

$$S(H_n, H_m) = 2^{-\frac{m+n}{2}} \sqrt{\binom{m+n}{n}} H_{m+n}.$$

Proof. We may first make use of our Hermite polynomial sum identity so that

$$S(H_n, H_m)(x) = 2^{-\frac{m+n}{2}} \sum_{i=0}^{n} \sum_{j=0}^{m} \sqrt{\binom{n}{i}\binom{m}{j}} H_{n-i}(x) H_{m-j}(x) \int_{\mathbb{R}} g(u) H_i(u) H_j(-u) du,$$
$$= 2^{-\frac{m+n}{2}} \sum_{i=0}^{\min\{n,m\}} (-1)^i \sqrt{\binom{n}{i}\binom{m}{i}} H_{n-i}(x) H_{m-i}(x).$$

To proceed, we use the identity for products of Hermite polynomials to see

$$S(H_n, H_m)(x) = \sum_{k=m-n}^{m+n} b_{m,n,k} H_k(x),$$

where $b_{m,n,k}$ vanishes if k is not between $|m-n|$ and $m+n$ or if $k - |m-n|$ is not even. Some judicious calculations imply that

$$b_{m,n,k} = \begin{cases} 2^{-\frac{m+n}{2}} \sqrt{\binom{m+n}{n}} & k = m+n \\ 0 & \text{else.} \end{cases}$$

□

The bilinearity of S then gives the following corollary.

Corollary 2.2.17. *Suppose $f_1 = \sum \alpha_n H_n$ and $f_2 = \sum \beta_n H_n$ are in $L^2(g)$. Then*

$$S(f_1, f_2) = \sum_{k=0}^{\infty} \left[2^{-\frac{k}{2}} \sum_{n=0}^{k} \sqrt{\binom{k}{n}} \alpha_{k-n} \beta_n \right] H_k.$$

We'd like to proceed by relating the decay rate of the coefficients of $S(f_1, f_2)$ to those of the coefficients of f_1 and f_2. Most natural is to take $f_1, f_2 \in L^2(g)$.

Proposition 2.2.18. *If $f_1, f_2 \in L^2(g)$, then*

$$\|S(f_1, f_2)\|_{L^2(g)} < \|f_1\|_{L^2(g)} \|f_2\|_{L^2(g)}.$$

Proof. To prove this, it suffices to take f_1, f_2 both Hermite polynomials, with the full result then following from bilinearity of S and completeness of Hermite polynomials in $L^2(g)$. Then we have immediately for any $n \geq k \geq 0$

$$\|S(H_{n-k}, H_k)\|_{L^2(g)} < C,$$

with

$$C = \sup_{n \geq k \geq 0} 2^{-\frac{n}{2}} \sqrt{\binom{n}{k}}.$$

Certainly $C \leq 1$, since no binomial coefficient exceeds 2^n. Taking $n = k = 0$ we conclude that $C = 1$. □

We can in fact take $C < 1$ in the above proof if $f_1, f_2 \in \mathcal{H}$ by restricting the supremum to $n \geq 2$ and $n - k \geq 2$. While an $L^2(g)$ result is sufficient for a basic proof of the central limit theorem, we can strengthen the result to also require convergence of some derivatives of q to 0. To obtain this, we'll instead assume $f_1, f_2 \in W^s(g)$ for some $s > 0$. For this result, we'll need a combinatorial lemma.

Lemma 2.2.19. *For $s_1, s_2 \geq 0$ and $s_3 \leq \frac{1}{4} + s_1 + s_2$ there exists a constant $C_{s_1,s_2,s_3} \geq 1$ such that*

$$\sup_{n \geq k \geq 0} 2^{-\frac{n}{2}} \sqrt{\binom{n}{k}} \frac{(n+1)^{s_3}}{(k+1)^{s_1}(n-k+1)^{s_2}} < C_{s_1,s_2,s_3}.$$

Moreover, if $s_3 \leq \min\{s_1, s_2\}$, then $C_{s_1,s_2,s_3} = 1$.

Proof. Certainly $C_{s_1,s_2,s_3} \geq 1$ since we may take $n = k = 0$. It suffices to argue that each term is bounded whenever n is extremely large. For a given n, the product of the exponential and binomial terms is maximized when $k = \left\lfloor \frac{n}{2} \right\rfloor$. A standard use of Stirling's formula then finds that

$$2^{-\frac{n}{2}} \sqrt{\binom{n}{k}} \sim \left(\frac{\pi}{2} n\right)^{-\frac{1}{4}}.$$

Supposing $k = an$ for $0 \leq a \leq 1$ then the final term behaves asymptotically as

$$a^{s_1}(1-a)^{s_2} n^{s_3 - s_1 - s_2}.$$

In particular, since the first term is asymptotically sharply peaked around $k = \frac{1}{2}n + \mathcal{O}(\sqrt{n} \log n)$, the product of all three terms is bounded provided

$$s_3 - s_1 - s_2 < \frac{1}{4}.$$

If $s_3 \leq \min\{s_1, s_2\}$, we need only note that the final term is maximized when $k = 0$ or $k = n$, where it is bounded by 1. □

As a result, we achieve a result assuming membership of f, g in particular Sobolev spaces.

Corollary 2.2.20. *If $f_1 \in W^{s_1}(g)$ and $f_2 \in W^{s_2}(g)$ for $s_1, s_2 \geq 0$ and if $s_3 < \frac{1}{2} + s_1 + s_2$, then there exists $C_{s_1,s_2,s_3} \geq 1$ such that*

$$\|S(f_1, f_2)\|_{W^{s_3}(g)} < C_{s_1,s_2,s_3} \|f_1\|_{W^{s_1}(g)} \|f_2\|_{W^{s_2}(g)}.$$

If $s_1 = s_2 = s_3$, then $C_{s_1,s_2,s_3} = 1$.

2.2.7 Local Berry-Esseen Bounds

With the analysis of the linear and nonlinear parts of the renormalization operator completed, the argument provides a local Berry-Esseen bound. In all that follows, we take p_0 to be a probability density with mean zero and variance 1, and let $\mathscr{R}(p_n) = p_{n+1}$ for all $n \geq 0$. Further, we write $p_n = (1 + q_n)g$.

Theorem 2.2.21. *Given any integer $k > 2$, there exists an $\epsilon_k > 0$ and an increasing function $A : [0, \epsilon_k) \to \mathbb{R}_{\geq 0}$ with $A'(0) = 1$ so that the following holds. For any probability density p_0 with $q_0 \in \mathcal{H}$ satisfying $\|q_0\|_{L^2(g)} = \epsilon < \epsilon_k$ and where k is the smallest positive integer with $\mathbb{E}_{p_0}[H_k(X)] \neq 0$ one has for all $n \in \mathbb{N} \cup \{0\}$*

$$\left(2^{\frac{k}{2}-1}\right)^n \|q_n\|_{L^2(g)} \leq A(\epsilon),$$

and moreover

$$\lim_{n\to\infty} \left(2^{\frac{k}{2}-1}\right)^n \|q_n\|_{L^2(g)} \to \mathbb{E}_{p_0}[H_k(X)].$$

Proof. We can choose $\epsilon_k = 1 - 2^{1-\frac{k}{2}}$ and we'll denote $\lambda_k = 2^{1-\frac{k}{2}}$. Since $\mathbb{E}_{p_0}[H_j(X)] = 0$ for all $j = 1, \ldots, k-1$, $\langle H_j, q_0 \rangle = 0$ for each such j. This subspace is invariant under both T and S, so the same holds for all q_n. Applying Corollary 2.2.13 and Proposition 2.2.18, we have for any $n \in \mathbb{N} \cup \{0\}$

$$\|q_{n+1}\|_{L^2(g)} \leq f_{\lambda_k}\left(\|q_n\|_{L^2(g)}\right),$$

with $f_\lambda(x) = \lambda x + x^2$ a standard quadratic map. Since $\lambda_k < 1$, the interval $[0, \epsilon_k)$ falls within the attracting basin of the stable fixed point of f_{λ_k} at 0, forcing an exponential rate of convergence to 0. Let $a_n = \lambda_k^{-n} \|q_n\|_{L^2(g)}$, then

$$a_{n+1} \leq a_n \left(1 + \lambda_k^{-1} f_{\lambda_k}^n(\epsilon)\right).$$

Then, we may define

$$A(\epsilon) = \epsilon \prod_{n=1}^{\infty} \left(1 + \lambda_k^{-1} f_{\lambda_k}^n(\epsilon)\right).$$

This infinite product converges since $f_{\lambda_k}^n(\epsilon) \to 0$ exponentially fast for large n. Differentiability follows from noting that as $\epsilon \to 0$, the infinite product is asymptotically equivalent to the q-Pochhammer symbol $(-\lambda; \epsilon)_\infty$.

To see that the limit exists, suppose

$$q_n = \sum_{j=k}^{\infty} \alpha_{n,j} H_j.$$

Necessarily $\alpha_{0,j} = \mathbb{E}_{p_0}[H_j(X)]$. Observe that $S(q_n)$ consists only of Hermite polynomials of degree at least $2k$, so $\alpha_{n,j} = \lambda_j^n \alpha_{0,j}$ for $j = k, \ldots, 2k-1$. Thus, all such $\alpha_{n,j}$ with $j > k$ are exponentially small compared to $\alpha_{n,k}$, so vanish in the limit. For $j \geq 2k$, terms include exponentially small components compared to $\alpha_{n,k}$ added to components bounded by a multiple of $\alpha_{n,k}^2$, so these also vanish in the limit. □

A notable corollary follows from the observation that

$$\|p - g\|_{L^1(\mathbb{R})} = \|q\|_{L^1(g)} \leq \|q\|_{L^2(g)}.$$

Corollary 2.2.22. *Given any integer $k > 2$, there exists an $\epsilon_k > 0$ and an increasing function $A : [0, \epsilon_k) \to \mathbb{R}_{\geq 0}$ with $A'(0) = 1$ so that the following holds. For any probability density p_0 with $q_0 \in \mathcal{H}$ satisfying $\|q_0\|_{L^2(g)} = \epsilon < \epsilon_k$ and where k is the smallest integer with $\mathbb{E}_{p_0}[H_k(X)] \neq 0$ one has for all $n \in \mathbb{N} \cup \{0\}$*

$$\left(2^{\frac{k}{2}-1}\right)^n \|p_n - g\|_{L^1(\mathbb{R})} \leq A(\epsilon).$$

The Kolmogorov-Smirnov distance is trivially bounded by the $L^1(\mathbb{R})$ distance, so we recover the same asymptotic rate of convergence in the Berry-Esseen theorem when $k = 3$. If our distributions are taken to be even, then $k \geq 4$ and we obtain strictly faster convergence than is implied by standard Berry-Esseen bounds.

By making use of Corollary 2.2.20 we may obtain convergence in Sobolev spaces using essentially the same argument.

Theorem 2.2.23. *Let $s > 0$. Given any integer $k > 2$, there exists an $\epsilon_k(s) > 0$ and an increasing function $A_s : [0, \epsilon_k(s)) \to \mathbb{R}_{\geq 0}$ differentiable at 0 so that the following holds. For any probability density p_0 with $q_0 \in \mathcal{H}$ satisfying $\|q_0\|_{L^2(g)} = \epsilon < \epsilon_k(s)$ and where k is the smallest positive integer with $\mathbb{E}_{p_0}[H_k(X)] \neq 0$ one has for all $n \in \mathbb{N} \cup \{0\}$*

$$\left(2^{\frac{k}{2}-1}\right)^n \|q_n\|_{W^s(g)} \leq A_s(\epsilon),$$

and moreover

$$\lim_{n\to\infty} \left(2^{\frac{k}{2}-1}\right)^n \|q_n\|_{W^s(g)} \to (k+1)^{\frac{s}{2}} \mathbb{E}_{p_0}[H_k(X)].$$

Chapter 3

Random Walk Renormalization for the Dyson Hierarchical Model

In this chapter, we explore the Dyson hierarchical model from statistical physics. We'll discuss the basic statistical mechanics and define the block-spin renormalization for such hierarchical models. We then discuss self-similar group theory and define two renormalization operators, termed 'Schur renormalization' and 'block renormalization', which arise naturally from some group-theoretic properties. This self-similar group structure is placed on Dyson hierarchical models, and the corresponding two group theoretic renormalization operators are analyzed and shown to be distinct. Finally, we discuss random walk representations of spin models. A block renormalization operator is introduced in this context as well. The block-spin renormalization operator, group theoretic block renormalization operator, and random walk model block renormalization operator are shown to be, in a certain sense, equivalent for Dyson hierarchical models with Gaussian a priori measures. This equivalence is used to analyze the asymptotic and leading order behavior of correlation functions for such Gaussian Dyson hierarchical models.

The main results of this chapter are the equivalence of block spin renormalization, group theoretic block renormalization, and random walk model block renormalization; as well as

the inequivalence of these three with the group theoretic Schur renormalization. This closes one potential avenue of connection between the renormalization theory of return maps of dynamical systems and the traditional block spin renormalization theory of statistical physics.

3.0.1 Notation

Though this chapter will invoke many different pieces of notation, most will be explained as they appear. The others are listed here.

Given a set $U \subset X$, we denote by U^c the complement of U within X.

We will refer to the Lagrange multiplier β in the definition of a Gibbs state as a 'temperature', instead of as an 'inverse temperature.'

3.1 The Equilibrium Statistical Mechanics of Lattice Spin Systems

In this section, we'll discuss the basic structure of the type of spin systems of interest to our analysis and introduce the related terminology. This presentation of the basic theory of equilibrium statistical mechanics will lead into a discussion of the Dyson Hierarchical Model and the corresponding renormalization theory.

In the sections that follow, we will let $G = (V, E)$ be a graph. Elements of V will be referred to as 'sites' and elements of E as 'edges'. In general, we will be most interested in the case when G is a complete graph on a subset of $\mathbb{Z}$ or $\mathbb{N}$, though the basic theory holds in much greater generality. The size $|V|$ of the graph will be referred to as the 'volume' of G. We will usually need to work with finite-volume graphs, with the infinite-volume cases being understood by taking an appropriate limit over an exhaustion of G by finite subgraphs.

Our early discussion will closely follow those found in standard references. Readers are in particular directed towards Chapters 1 and 2 of [Rue99], Chapters 1 and 4 of [Sin14], Chapter 3 of [Zin00], Chapters 3 and 4 of [Tho15], or Chapter 2 of [Gol18].

3.1.1 Finite-Volume Spin Systems

In this section, we will take $G = (V, E)$ to be a finite graph and $(\Omega, \mathscr{F}, \mathbb{P})$ to be a probability space.

Definition 3.1.1. A *random field* on G taking values in a real vector space X is a random variable $\varphi : \Omega \to X^V$. For any site $v \in V$, let $\pi_v : X^V \to X$ be the projection onto the v-coordinate, and let $\pi_\Lambda : X^V \to X^\Lambda$ for $\Lambda \subset V$ be the projection onto the subspace with coordinates in Λ. We'll say $\varphi_v = \pi_v \circ \varphi$ is the *spin variable*, or *spin*, at v. We'll refer to X^V as the *state space* and to elements of it as *states* or *configurations*.

In general, spin variables may have a complicated correlation structure. In the context of a spin model, this structure is encoded by interactions between sets of spin variables, which represent the energy arising from their mutual interaction with one another.

Definition 3.1.2. An *interaction* on G is a operator-valued set map Ψ which takes in a subset of V and outputs a function on the space of configurations on that subset. We'll often denote $\Psi(\Lambda) \equiv f_\Lambda : X^\Lambda \to \mathbb{R}$. We'll say that an interaction is *pairwise* on G if $f_\Lambda \equiv 0$ unless $|\Lambda| \leq 2$. Moreover, we'll call it *isotropic* if for any $\Lambda \subset V$ and orthogonal transformation T on X one has

$$f_\Lambda\left(T(\cdot), T(\cdot), \ldots, T(\cdot)\right) = f_\Lambda(\cdot, \cdot, \ldots, \cdot).$$

We'll say that an interaction has *finite range* if there exists an $N > 0$ such that if $\operatorname{diam}(\Lambda) > N$ then $f_\Lambda \equiv 0$. Otherwise, we'll say it has *infinite range* or *long range*.

This general theory is most concerned with pairwise interactions that respect the structure of the graph G, so that interactions are only nonzero if $\Lambda \in E$. Interactions enter into the structure of a spin model by way of a Hamiltonian, representing the total energy of the system.

Definition 3.1.3. Given an interaction Ψ on G, the *Hamiltonian* is the function $H : X^V \to \mathbb{R}$ defined by

$$H(\varphi) = \sum_{x \in V} \sum_{\substack{\Lambda \subset V \\ \Lambda \ni x}} f_\Lambda \left(\pi_\Lambda(\varphi) \right). \tag{3.1.1}$$

To fully describe a statistical mechanical system, it is not enough to merely consider the interactions between sites: we must also impose an independent prior distribution on each site which reflects the behavior of the spin variables in the absence of any interaction.

Definition 3.1.4. A collection of *a priori measures* on a state space X^V is an independent collection of probability measures $\{\nu_v\}_{v \in V}$ on X. We'll say it is *homogeneous* if the collection is identically distributed and *isotropic* if X is an inner product space and for all v the measures ν_v and $T_* \nu_v$ are equal in distribution for any orthogonal transformation T on X.

With all of this in mind, we are in a position to define a spin model.

Definition 3.1.5. A *spin model* on G taking values in a real vector space X is a pair $\left(H, \{\nu_v\}_{v \in V}\right)$ where $H : X^V \to \mathbb{R}$ is a Hamiltonian function for a set of pairwise interactions between connected sites in V, and the family $\{\nu_v\}_{v \in V}$ is a set of a priori measures on X.

In the case where the a priori measures are homogeneous, we'll simplify this notation representing a spin model to a pair (H, ν). Traditionally, such spin models are referred to as *scalar valued* if $X = \mathbb{R}$ or as *vector valued* if $X = \mathbb{R}^d$ for some $d \geq 2$. We will tighten this general definition of a spin model to a class of examples as in [FFS92; Ble10]. In summing over the edges in G, we will generally adopt the notation that G is a directed graph.

Definition 3.1.6. A spin model on G taking values in $\mathbb{R}^n$ is ferromagnetic of *pairwise type* if the following conditions hold.

1. The Hamiltonian is of the form

$$H(\varphi) = -\frac{1}{2} \sum_{(v,w) \in E} J_{v,w} \varphi_v \varphi_w,$$

with $J_{v,w} = J_{w,v} \geq 0$. We'll write J as the matrix whose entries are $J_{v,w}$ for some ordering of V.

2. The a priori measures $\{\nu_v\}_{v\in V}$ are isotropic, being either all centered Gaussian measures or satisfying for all $v \in V$ and $A > 0$ the condition that the measure

$$d\nu_{v,A}(\varphi_v) = e^{A\varphi_v^2} d\nu_v(\varphi_v),$$

is a finite measure on $\mathbb{R}$.

3. If

$$d\nu_v(\varphi_v) = g_v(\varphi_v^2) d\varphi_v,$$

for some $g : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$, then we'll say the model has *absolutely continuous a priori measures*.

The structure of a spin model gives rise to particular distributions of random fields which are central to the development of the theory.

Definition 3.1.7. The *Gibbs measure* or *Gibbs state* at temperature $\beta \in \mathbb{R}$ of a spin model $\left(H, \{\nu_v\}_{v\in V}\right)$ on G taking values in X is the measure μ on X^V given by

$$d\mu = \mathcal{Z}^{-1} e^{-\beta H(\varphi)} \prod_{v\in V} d\nu_v(\varphi_v),$$

where the *partition function* $\mathcal{Z}$ is defined as

$$\mathcal{Z} = \int_{X^V} e^{-\beta H(\varphi)} \prod_{v\in V} d\nu_v(\varphi_v).$$

In this finite-volume context, the Gibbs measure is manifestly unique and well-defined for any value of β for which $e^{-\beta H(\cdot)} \in L^1(\prod_{v\in V} \nu_v)$.

Proposition 3.1.1. *If (H, ν) is a homogeneous ferromagnetic spin model of pairwise type on a finite graph G with ν not a Gaussian measure, then the Gibbs measure is well-defined and unique for all $\beta \geq 0$. If ν is a Gaussian measure with variance $\sigma > 0$, then the Gibbs measure is well-defined and unique whenever $\beta \geq 0$ is sufficiently small that $\beta J - \frac{1}{2\sigma}$ is a negative-definite matrix.*

3.1.2 Infinite-Volume Spin Systems

In this section, we'll briefly discuss the infinite-volume case for relevant spin models. As such, we'll now take $G = (V, E)$ to be a countable directed complete graph without self-loops, to simplify notation. In particular, we'll take $G = \mathbb{Z}^n$ or $\mathbb{N}^n$ for some n. We will also begin to identify G with V.

There are two key technical difficulties in defining Gibbs states as in the previous section. First, a Hamiltonian written directly as a summation as in (3.1.1) will often yield a divergent sum, so we may only interpret such an expression as a formal series. Second, while $\prod_{v \in V} \nu_v$ will yield a probability measure on the state space X^V, it may not be the case that Gibbs states are absolutely continuous with respect to it. As such, it is necessary to either define the infinite-volume case by the behavior of finite-dimensional conditional measures or by taking an exhaustion of G by an appropriate sequence of finite subgraphs, obtaining Gibbs states as the 'thermodynamic limit' taken over such a sequence. We will examine both of these approaches in the work that follows.

Defining Hamiltonians and Gibbs states on finite subgraphs $\Lambda \subset G$ is done similarly to the previous section, except that we must account for those interactions between points inside Λ and points outside. Since we are ultimately only concerned with pairwise interactions, we may certainly assume that the set of interactions only contains finite subsets of V.

Definition 3.1.8. Given a finite subgraph $\Lambda \subset G$ and an interaction Ψ on finite subsets of G, a *relative Hamiltonian* is a map $H_\Lambda : X^G \to \mathbb{R}$ given by

$$H_\Lambda(\varphi) = \sum_{x \in \Lambda} \sum_{\substack{\Gamma \subset G \text{ finite} \\ \Gamma \ni x}} f_\Gamma(\pi_\Gamma(\varphi)).$$

For such a Hamiltonian to be well-defined, it is natural to assume that for any site $v \in V$

$$\sum_{\substack{\Gamma \subset G \text{ finite} \\ \Gamma \ni x}} |f_\Gamma(\pi_\Gamma(\varphi))| < \infty.$$

A relative Hamiltonian then gives rise to a relative Gibbs state for a spin model, subject to the added complication that such a Gibbs state must account for the behavior of the random field outside of Λ.

Definition 3.1.9. Given a finite subgraph $\Lambda \subset G$, a *relative Gibbs state* for a spin model $(H, \{\nu_v\}_{v\in V})$ on G taking values in X is a measure μ_Λ^η on X^G given by

$$d\mu_\Lambda^\eta(\varphi) = \mathcal{Z}_\Lambda^{-1} e^{-\beta H_\Lambda(\varphi)} \prod_{v\in\Lambda} d\nu_v(\varphi_v) d\eta\left(\pi_{\Lambda^c}(\varphi)\right),$$

where the *boundary condition* η is some Radon probability measure on X^{Λ^c} and $\mathcal{Z}_\Lambda$ is the *relative partition function*

$$\mathcal{Z}_\Lambda = \int_{X^G} e^{-\beta H_\Lambda(\varphi)} \prod_{v\in\Lambda} d\nu_v(\varphi_v) d\eta\left(\pi_{\Lambda^c}(\varphi)\right).$$

There are a number of ways to proceed with a definition of an infinite-volume Gibbs state. In the approach that follows, we use the set of relative Gibbs states as conditional measures of the Gibbs state of the whole system.

Definition 3.1.10. A *Gibbs state* at temperature β for a spin model $(H, \{\nu_v\}_{v\in V})$ on G taking values in X is a measure μ such that for any finite $\Lambda \subset G$ and μ almost every $\psi \in X^G$ (we write $\pi_{\Lambda^c}(\psi) = \overline{\psi}$), the conditional measure $\mu^{\overline{\psi}}$ defined on the set $\left\{\varphi \in X^G : \pi_{\Lambda^c}(\varphi) = \overline{\psi}\right\}$ satisfies for $\delta^{\overline{\psi}}$ the Dirac measure on X^{Λ^c} centered at $\overline{\psi}$,

$$\mu^{\overline{\psi}} \otimes \delta^{\overline{\psi}} = \mu_\Lambda^{\delta^{\overline{\psi}}}.$$

An alternative, but equivalent, definition takes a 'thermodynamic limit' of relative Gibbs states over an exhaustion of G, defining Gibbs states as any limiting measure thus obtained.

Definition 3.1.11. A *Gibbs state* at temperature β for a spin model $(H, \{\nu_v\}_{v\in V})$ on G taking values in X is a measure μ on X^G which is a weak limit point of a sequence $\{\mu_{\Lambda_n}^{\eta_n}\}_{n\to\infty}$, for some sequence of measures $\{\eta_n\}_{n\in\mathbb{N}}$ with η_n a measure on Λ_n^c for each n, of relative Gibbs states at temperature β for a suitable exhaustion $\Lambda_n \nearrow G$.

In order to guarantee the existence of such a thermodynamic limit, one requires a suitable family of subsets of G.

Definition 3.1.12. For any $\Lambda \subset G$, define

$$\partial\Lambda = \{v \in \Lambda | d(v, \Lambda^c) = 1\}$$

We say that an increasing sequence $\Lambda_n \nearrow G$ in the *sense of Van Hove* if $\bigcup \Lambda_n = G$ and

$$\lim_{n\to\infty} \frac{|\partial\Lambda_n| + |\partial\Lambda_n^c|}{|\Lambda_n|} \to 0.$$

In the case that G happens to be a discrete graph arising from an amenable group, many Fölner sequences provide good examples of sequences which converge in the sense of Van Hove.

In the general circumstances so far described, there are a large variety of different cases in which Gibbs states are known to exist. For Ising models, a classic existence result comes from [GM67]. For more general measures, results date back to [Dob68] (also found as Theorem 1.3 in [Rue99]), under sufficient assumptions of continuity of relative Hamiltonians. The following proposition proceeds from those results, and is sufficient for our purposes.

Proposition 3.1.2. *If (H, ν) is a homogeneous spin model of pairwise type on G taking values in $\mathbb{R}^d$, there exists a Gibbs state on G if for any $v \in G$ one has*

$$\sum_{w \neq v} J_{v,w} < \infty.$$

Markedly different from the finite-volume case, infinite-volume Gibbs measures do not need to be unique. However, the set of Gibbs measures is convex and weakly closed, with extreme points often referred to as *phases* or *ergodic Gibbs states*. We say $\beta_c \in \mathbb{R}$ is a *critical point* or a *phase transition* if the number of phases when $\beta < \beta_c$ differs from the number when $\beta > \beta_c$.

As reasoning directly about infinite-volume Gibbs states can be difficult, it is useful to analyze the properties of these measures indirectly by observing the correlation structure of distinct spin variables.

Definition 3.1.13. Given a Gibbs state μ for a spin model on G taking values in $\mathbb{R}^d$, a *correlation function* is an expectation

$$\langle \varphi_{v_1}^{(\alpha_1)} \varphi_{v_2}^{(\alpha_2)} \cdots \varphi_{v_n}^{(\alpha_n)} \rangle = \int \prod_{j=1}^{n} \varphi_{v_j}^{(\alpha_j)} d\mu(\varphi),$$

where $v_1, \ldots, v_n \in G$, $1 \leq \alpha_1, \ldots, \alpha_n \leq d$, and $\varphi_v^{(\alpha)} \in \mathbb{R}$ denotes the α component of the spin variable φ_v.

3.1.3 Dyson Hierarchical Models

The spin models of interest in this paper are scalar-valued suitable spin models on $\mathbb{N}$ of a particular class, originally constructed by Dyson [Dys69]. Our description of these models will most closely follow [Ble10]. Here and in what follows we'll take $\mathbb{N}$ to contain 0, for simplicity. The 'hierarchical' nature of the spin model will relate to choosing the interaction strength between two points to depend only on their distance in a particular ultrametric $d : \mathbb{N} \times \mathbb{N} \to \mathbb{R}_{\geq 0}$.

Definition 3.1.14. A *dyadic box* in $\mathbb{N}$ is an interval $B_{j,k} = [j \cdot 2^k, \ldots, (j+1) \cdot 2^k - 1] \subset \mathbb{N}$ for $j, k \in \mathbb{N}$. Define $\ell(x, y)$ for $x, y \in \mathbb{N}$ to be the smallest integer k such that there exists a j so that $x, y \in B_{j,k}$. The *hierarchical ultrametric* $d : \mathbb{N} \times \mathbb{N} \to \mathbb{R}_{\geq 0}$ is defined so that $d(x, x) = 0$ for all $x \in \mathbb{N}$, and otherwise

$$d(x, y) = 2^{\ell(x,y)-1}.$$

Being an ultrametric, the hierarchical ultrametric imposes a tree-like metric structure on $\mathbb{N}$, which will be discussed later in a subsequent section.

Given this ultrametric, we are in a position to define the central spin models of interest to this paper.

Definition 3.1.15. A *scalar-valued Dyson Hierarchical Model* is a homogeneous suitable spin model (H, ν) on $\mathbb{N}$ taking values in $\mathbb{R}$ with

$$J_{x,y} = f(d(x, y)),$$

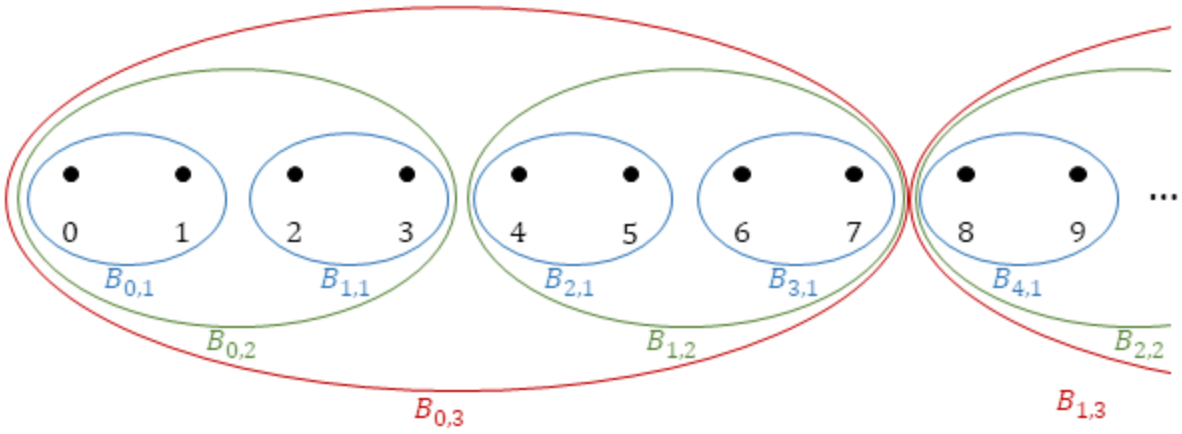

Figure 3.1: The Dyadic Block Structure on $\mathbb{N}$ Respected by the Hierarchical Ultrametric

for some decreasing $f : \mathbb{N} \to \mathbb{R}_{\geq 0}$ which vanishes at ∞. Such a model is said to have *power-law interactions with parameter* a if $f(n) = n^{-a}$.

The original development of these models in [Dys69] sought to understand power law interactions with parameter $1 < a < 2$ and was motivated by the question of existence of phase transitions for Ising models with long range power law interactions on $\mathbb{Z}$. The interactions in such a Dyson model between two points is weaker than the corresponding power law interaction on $\mathbb{Z}$, but is asymptotically within a constant factor as one point is taken to ∞. Dyson's original argument used the Griffiths inequalities (see Theorem 5.4.1 in [Rue99]) to show that phase transitions found in a Dyson hierarchical model implied the existence of phase transitions in a corresponding long range model on $\mathbb{Z}$. Dyson hierarchical models were subsequently found to have interesting properties on their own right, enabling a significant body of research based on renormalization methods.

Dyson hierarchical models may also be obtained by taking a thermodynamic limit of zero-boundary relative Gibbs measures of a particular sequence of subgraphs of $\mathbb{N}$.

Definition 3.1.16. The homogeneous spin model of pairwise type on $V_N = B_{0,N}$ with the

Hamiltonian

$$H_N(\varphi) = -\sum_{x \neq y \in V_N} f(d(x,y))\varphi_x \varphi_y,$$

for $f : \mathbb{N} \to \mathbb{R}_{\geq 0}$ which vanishes at ∞ is said to be a *truncated Dyson hierarchical model on* V_N.

Work in [Dys69; Dys71; BS73; BS75] demonstrates in particular that it suffices to consider this zero boundary case in taking a thermodynamic limit as $N \to \infty$. This holds in the sense that the variance in the mean spin as a function of β serves as an order parameter for a Dyson hierarchical model. This is to say that such a model lacks a phase change if the variance in the mean spin is 0 for all $\beta > 0$, and has a phase change at some $\beta_c > 0$ if the variance in mean spin is 0 for $\beta < \beta_c$ and nonzero for $\beta > \beta_c$.

The early analyses in [Dys69; BS73] serve to demonstrate the existence of a phase change for Dyson models with power law interactions having parameter $1 < a < 2$. For $a > 2$, phase changes do not occur. For $a \leq 1$, the thermodynamic limit is no longer well-defined and one must resort to the analysis of non-extensive systems as in [CT96; Muk09], which falls outside of the scope of this thesis. The marginal case $a = 2$ described in [Aiz+88] yields a distinctive phase change for models on $\mathbb{Z}$. The marginal case for Dyson hierarchical models is demonstrated in [Dys71], with $f(n) = n^{-2} \log \log n$.

One can also define vector-valued Dyson hierarchical models, which have been analyzed in detail in a number of sources, including [BM89; BM07; BM21].

3.1.4 Block-Spin Renormalization for Dyson Models

Renormalization ideas in statistical mechanics are commonly attributed back to the block-spin transformations of Kadanoff [Kad66] and the subsequent work by Wilson [Wil71a; Wil83] describing the renormalization group in terms of scaling transformations. The hierarchical structure of Dyson hierarchical models makes them particularly amenable to such a renormal-

ization description: the model itself is already cleanly divided into a sequence of delineated scales.

In its simplest setting, block-spin renormalization (also called Migdal-Kadanoff renormalization) takes a truncated Dyson model on V_N with a priori measure ν and interaction function $f : \mathbb{N} \to \mathbb{R}_{\geq 0}$ to a truncated Dyson model on V_{N-1} with a new a priori measure $\tilde{\nu}$ and a new interaction function $\tilde{f} : \mathbb{N} \to \mathbb{R}_{\geq 0}$. We'll define the *interaction sequence* $f_n \equiv f(2^n)$ for $n \in \mathbb{N}$ for simplicity. For any $x \in V_N$, we define $\tilde{x} \in V_{N-1}$ so that $x \in B_{\tilde{x},1}$, and moreover for any $\tilde{x} \in V_N$ we define $x_1 = 2\tilde{x}$ and $x_2 = 2\tilde{x} + 1$.

The main operation in block-spin renormalization consists of transforming the random field $\{\varphi_x\}_{x \in V_N}$ into a random field $\{\tilde{\varphi}_{\tilde{x}}\}_{\tilde{x} \in V_{N-1}}$ by

$$\frac{\varphi_{x_1} + \varphi_{x_2}}{2^\kappa} = \tilde{\varphi}_{\tilde{x}},$$

with $\kappa > 0$ a parameter of the block-spin transformation. The action of this renormalization on the interaction sequence and a priori measures are specified by the normalization condition that $f_0 = 1$ and the stipulation that renormalization should leave the partition function unchanged. The following proposition based upon descriptions in [BS73; Ble10; Ble12; BM21] describes the induced transformations.

Proposition 3.1.3. *Subject to the above normalization constraints, we must take*

$$\kappa = -\frac{1}{2}\log_2(f_1),$$

with the induced maps on interaction sequences $f \mapsto \tilde{f}$ with

$$\tilde{f}_n = \frac{f_{n+1}}{f_1},$$

and on absolutely continuous a priori measure densities $\nu \mapsto \tilde{\nu}$ with

$$\tilde{\nu}(E) = \frac{1}{\iint_{\mathbb{R}^2} e^{\beta st} d\nu(s) d\nu(t)} \iint_{\frac{s+t}{2^\kappa} \in E} e^{\beta st} d\nu(s) d\nu(t),$$

for any Borel set E.

Proof. The key observation in this proof is that the Hamiltonian $H_{N,f}$ can be rewritten as a sum over elements in V_{N-1} as

$$\begin{aligned} H_{N,f}(\varphi) &= -\sum_{\tilde{x}\in V_{N-1}} f_0\varphi_{x_1}\varphi_{x_2} - \sum_{\tilde{x}\neq\tilde{y}} 2^{2\kappa} f\left(2d(\tilde{x},\tilde{y})\right)\tilde{\varphi}_{\tilde{x}}\tilde{\varphi}_{\tilde{y}}, \\ &= -\sum_{\tilde{x}\in V_{N-1}} f_0\varphi_{x_1}\varphi_{x_2} + H_{N-1,\tilde{f}}(\tilde{\varphi}). \end{aligned}$$

As $d(\tilde{x},\tilde{y}) = 1$ if and only if $d(x,y) = 2$, to have $\tilde{f}_0 = 1$ we must take

$$\kappa = -\frac{1}{2}\log_2(f_1),$$

giving the desired map on interaction sequences. Writing out the partition function then yields

$$\begin{aligned} \mathcal{Z} &= \int e^{-\beta H_{N-1,\tilde{f}}(\tilde{\varphi})} \prod_{\tilde{x}\in V_{N-1}} e^{\beta\varphi_{x_1}\varphi_{x_2}} d\nu(\varphi_{x_1}) d\nu(\varphi_{x_2}), \\ &= \int e^{-\beta H_{N-1,\tilde{f}}(\tilde{\varphi})} \prod_{\tilde{x}\in V_{N-1}} d\tilde{\nu}(\tilde{\varphi}_{\tilde{x}}). \end{aligned}$$

□

We will eliminate the dependence on β entirely by making a change of variables to a random field $\{\sqrt{\beta}\varphi_x\}_{x\in V_N}$, which only has the effect of rescaling the a priori measures. Formally then, renormalization is a transformation acting jointly on interaction sequences and on a priori measures.

Definition 3.1.17. We define the *full block-spin renormalization transformation* $\mathscr{R} : (f,\nu) \mapsto (\tilde{f},\tilde{\nu})$.

The cases when the interactions are a power law in hierarchical distance is of particular interest.

Corollary 3.1.4. *If* $f(n) = n^{-a}$, *then* $\kappa = \frac{a}{2}$, *and* $\tilde{f} = f$.

This motivates the use of a as a parameter for the renormalization transformation, and allows renormalization of Dyson hierarchical models with power law interactions to be understood purely as a transformation of a priori measures. We will henceforth assume a priori measures are absolutely continuous, with $d\nu(z) = p(z)dz$.

Definition 3.1.18. The *block-spin renormalization transformation* with parameter $a \in \mathbb{R}$ is the operator $\mathscr{R}_a$ on probability densities with faster-than-Gaussian decay given by

$$\mathscr{R}_a(p)(s) = \frac{e^{\frac{s^2}{2^{2-a}}}}{Z} \int e^{-t^2} p\left(\frac{s}{2^{\frac{2-a}{2}}} - t\right) p\left(\frac{s}{2^{\frac{2-a}{2}}} + t\right) dt.$$

The dynamical behavior of this operator yields great insight into the nature of phase changes within Dyson hierarchical models. A survey of such results can be found in [Ble10]. The main results relevant to our analysis are shown below.

Theorem 3.1.5 ([BS73; BS75; Ble77; Sin14])**.** *For $1 < a < 2$, the family of Gaussian distributions $\nu \sim N(0, \sigma)$ for which $\mathscr{R}_a$ is defined is invariant, and $\nu_* \sim N(0, 1 - 2^{1-a})$ is a fixed point of $\mathscr{R}_a$. For $1 < a < \frac{3}{2}$ this fixed point is hyperbolic, with $D|_{\nu_*}$ compact, and with a single unstable direction oriented along the family of Gaussian distributions. For $\frac{3}{2} < a < 2$, this Gaussian fixed point has more than one unstable direction. In particular, for sufficiently small $\epsilon > 0$ and $a = \frac{3}{2} + \epsilon$, new non-Gaussian hyperbolic fixed points of $\mathscr{R}_a$ with codimension 1 stable direction appear.*

3.2 Self-Similar Groups

In this section we will discuss the basic structure of a self-similar group and see how this structure gives rise to a operation on random walks first referred to as the 'Münchausen trick' in [Kai05] or as the 'Schur renormalization' or the 'Bartholdi-Kaimanovich-Virag transformation' in [GN07]. We will then see how Dyson hierarchical models can be fit into this framework.

3.2.1 Basic Terminology of Self-Similar Groups

This discussion will closely follow the work of Grigorchuk and Nekrashevych in Section 2 of [GN07], with similar discussion found in Chapter 1 of [Nek05]. One starting point of the theory of self-similar groups is found in infinite rooted trees, one of the basic self-similar structures in mathematics.

Definition 3.2.1. The *regular degree d rooted tree* T_d is the free monoid $\{1, \ldots, d\}^*$, identified as a graph with edges (w, wk) for any $w \in T_d$ and $k \in \{1, \ldots, d\}$. The *k-th level* of T_d consists of all words of length k. We write $|w|$ to denote the level of $w \in T_d$.

These trees are self-similar in the sense that at any point $w \in T_d$, the subtree wT_d descending from w is isomorphic to the whole tree. Nonetheless, the root of the tree is distinguished as it has a strictly lower degree than every other point in the tree. This gives rise to distinctive behavior in the automorphism group of T_d.

Lemma 3.2.1. *Any automorphism $g \in Aut(T_d)$ preserves distance from the root. In particular, for all $k \in \mathbb{N}$, the k-th level of T_d is invariant under g.*

Due to these two properties, an element $g \in \mathrm{Aut}(T_d)$ may be entirely described by two pieces of data: an element of the symmetric group S_d representing the manner in which g permutes the d subtrees at level 1, and a list of d automorphisms of T_d representing the action of g on each of those subtrees. That is to say,

$$g \cong ((g^{(k)})_{k=1,\ldots,d}, \sigma_g), \tag{3.2.1}$$

with $\sigma_g \in S_d$ and $g^{(k)} \in \mathrm{Aut}(T_d)$ for all k. This equivalence gives $\mathrm{Aut}(T_d)$ the structure of a *permutational wreath product*

$$\mathrm{Aut}(T_d) \cong \mathrm{Aut}(T_d) \wr S_d,$$

and exemplifies the type of self-similarity which self-similar groups will also possess.

Definition 3.2.2. A *self-similar group* is a group $G \subset \mathrm{Aut}(T_d)$ for some $d \geq 2$ such that for any $g \in G$, $g^{(k)} \in G$ for all $k = 1, \ldots, d$.

Written another way, for an element g in a self-similar group G and $w, v \in T_d$, there exists $g|_w \in G$ such that

$$g(wv) = g(w)g|_w(v).$$

Any such group possesses natural families of subgroups. Namely, for any $w \in T_d$, we may consider the stabilizer subgroup $\mathrm{Stab}_G(w)$. Self-similar groups possess a natural group homomorphism $\mathrm{Stab}_G(w) \to G$ defined by

$$g \mapsto g|_w.$$

There are a number of self-similar groups for which these homomorphisms eventually stabilize into a trivial homomorphism as the level of w increases. This can occur if all elements of G act trivially on sufficiently high levels of T_d, for example an action $\mathbb{Z}_2 \hookrightarrow \mathrm{Aut}(T_d)$ which merely permutes two subtrees at level 1. It can also occur if the action of G on all levels has the same period, such as the action $\mathbb{Z}_2 \hookrightarrow \mathrm{Aut}(T_2)$ which swaps every pair of subtrees at all levels. We wish to select groups for which $g \mapsto g|_w$ is surjective for all $w \in T_d$.

Definition 3.2.3. A self-similar group $G \subset \mathrm{Aut}(T_d)$ is *self-replicating* if G acts transitively on the first level of T_d and the map $g \mapsto g|_k$ is onto G for all $k = 1, \ldots, d$.

This allows the same permutational wreath product structure as is possessed by $\mathrm{Aut}(T_d)$ to be inherited by G. If the map $g \mapsto g|_k$ is also injective for some $k \in \{1, \ldots, d\}$, then a self-replicating group is finitely non-co-Hopfian.

Definition 3.2.4. A group G is *finitely non-co-Hopfian* if there exists a finite-index proper subgroup $H \subsetneq G$ such that $H \cong G$.

Some discussions of finitely non-co-Hopfian groups can be found in [Bri+10; Lim21].

3.2.2 Schur Renormalization of Random Walks

Much in the style of renormalization operators defined in dynamical systems via return maps, we can define a renormalization transformation on suitable spaces of random walks

on self-replicating groups. This idea was first developed by Bartholdi and Virag in [BV05] to prove the amenability of the basilica group. Kaimanovich's work in [Kai05] generalized the method to a larger class of self-similar groups, including the Grigorchuk group. Our discussion here will follow [Kai05] and the discussion in Section 7 of [GN07].

Let $\{g_n\}_{n\in\mathbb{N}}$ be a discrete-time random walk on a self-replicating group G starting at the identity $e \in G$ with

$$g_n = h_n h_{n-1} \dots h_1,$$

where all $h_k \sim \mu$ are independent and identically-distributed increments. We'll identify $\mu \in \ell^1(G, \mathbb{R})$ with an element of the ℓ^1-completed group algebra by linearly extending the identification sending Dirac masses $\delta_h \mapsto h$ so that

$$\mu \equiv \sum_{h\in G} \mu_h h.$$

Further, we suppose the support of μ generates G and we impose a normalization condition that $\mu_e = 0$. The law of this random process at time $n \geq 0$ is the n-fold convolution of μ with itself. For $w \in T_d$, the subgroup $\mathrm{Stab}_G(w) \subset G$ has finite index, bounded by $d^{|w|}!$. As such, the stopping time $\tau_w = \min\{n > 0 : g_n \in \mathrm{Stab}_G(w)\}$ is almost surely finite. Thus, the original random walk gives rise to a random walk consisting of returns to $\mathrm{Stab}_G(w)$, which may be mapped by the homomorphism $g \mapsto g|_w$ to a new random walk $\{g_n^w\}_{n\in\mathbb{N}}$ on G with some new law μ^w.

Definition 3.2.5. The *Schur renormalization* at index $k \in \{1, \dots, d\}$ of a probability measure μ having no mass at e is the measure

$$\mathscr{R}_{S,k}(\mu) = \frac{1}{1-\mu^k(e)} \left(\mu^k - \mu^k(e)\delta_e\right).$$

We will refer to μ^k as the *Schur pre-renormalization* at index k of μ and say that the walk is *Schur renormalizable* whenever the Schur pre-renormalization is defined.

This process may be carried out more directly if $w \in \{1, \dots, d\}$ using the equivalence shown in (3.2.1) and identifying this random process as a random walk with internal degrees

of freedom given by $\{1, \ldots, d\}$. Doing so simply entails projecting the process $\{g_n\}_{n\in\mathbb{N}} \subset G$ into a reduced Markov process

$$\left\{\left(g_n^{(w)}, \sigma_{g_n}(w)\right)\right\}_{n\in\mathbb{N}},$$

on $G \times \{1, \ldots, d\}$. One may represent such a walk as an $\ell^1(G, \mathbb{R})$-valued $d \times d$ matrix

$$M = (\mu^{(j,k)})_{j,k\in\{1,\ldots,d\}},$$

with $\mu^{(j,k)}$ a sub-probability measure such that the total mass for each j is

$$\sum_{k\in\{1,\ldots,d\}} \|\mu^{(j,k)}\| = 1.$$

Define the notation $M_{\bar{j}j}, M_{j\bar{j}}, M_{\bar{j}\bar{j}}$ so that a $\bar{j}$ index means the removal of the j-th row or column, and a j index means the selection of only the j-th row or column. For example, with $j = 1$

$$M = \begin{pmatrix} \mu^{(1,1)} & M_{1\bar{1}} \\ M_{\bar{1}1} & M_{\bar{1}\bar{1}} \end{pmatrix}.$$

In particular, $M_{1\bar{1}}$ represents the first row of M with the first column's value omitted, $M_{\bar{1}1}$ represents the first column of M with the first row's value omitted, and $M_{\bar{1}\bar{1}}$ represents what remains of M when both the first column and first row are removed.

Theorem 3.2.2 ([Kai05; GN07])**.** *The Schur pre-renormalization of μ is given by*

$$\mu^k = \mu^{(k,k)} + M_{k\bar{k}} \left(I - M_{\bar{k}\bar{k}}\right)^{-1} M_{\bar{k}k},$$

and is continuous on $\ell^1_+(G, \mathbb{R})$ and defined whenever $I - M_{\bar{k}\bar{k}}$ is invertible.

The fixed points of Schur renormalization are of particular interest.

Definition 3.2.6. A measure μ is *Schur self-similar* if $\mu = \mathscr{R}_{S,k}(\mu)$ for some $k \in \{1, \ldots, d\}$.

The existing literature has focused primarily on examining cases in which μ has finite support. Work in [BE17] has sought to understand the Poisson-Furstenberg boundary of Schur self-similar walks.

3.2.3 Block Renormalization of Random Walks

In this section we'll discuss a different renormalization procedure which arises from a very simple construction. This construction does not rely on the group G being self-similar, but this will hold in the case of interest. A somewhat different condition on G is instead required.

Definition 3.2.7. A group G is *non-Hopfian* if there exists a nontrivial normal subgroup H such that $G/H \cong G$.

An analysis of random walks on Baumslag-Solitar groups, which satisfy this property, can be found in [CS18]. The following simple proposition indicates how such structure can arise in the context of a self-replicating group.

Proposition 3.2.3. *Suppose G is a self-replicating group and that for some $w \in \{1, \dots, d\}^*$, there exists a normal subgroup H such that*

$$G = H \times Stab_G(w),$$

and the map $g \mapsto g|_w$ from $Stab_G(w)$ to G is injective. Then G is non-Hopfian.

It is not the case that all self-replicating groups are non-Hopfian. In particular, as the Grigorchuk group is finitely generated and was shown to be residually finite in [Gri80], it is Hopfian. We can mimic this argument in more generality.

Proposition 3.2.4. *If G is a finitely generated self-replicating group, then G is Hopfian.*

Proof. If G is self-replicating, then for any non-identity element $g \in G$ there exists a $k \in \mathbb{N}$ such that g acts as a non-trivial permutation on the k-th level of the tree T_d. Then consider the homomorphism $\varphi_k : G \to S_{d^k}$ which sends each element of G to the corresponding permutation on the k-th level of T_d. Certainly $\varphi_k(g) \neq e \in S_{d^k}$, so G is residually finite. As it is also finitely generated, it is Hopfian. □

As such, the block renormalization procedure of this section may not be defined in this manner on all self-replicating groups. A precise condition for when a self-replicating group is non-Hopfian is unknown to us.

We will adopt the notation of the previous section, taking a walk with independent increments all distributed according to a measure $\mu \in \ell^1(G, \mathbb{R})$ where the support of μ generates G. Such a walk induces a walk on cosets of H, which we call *blocks*, and hence on G/H. Through the isomorphism between G/H and G, this gives rise to a new Markov process on G with a new law μ^b.

Definition 3.2.8. The *block renormalization* of μ is the measure

$$\mathscr{R}_b(\mu) = \frac{1}{1-\mu(H)}\left(\mu^b - \mu(H)\delta_e\right).$$

We will refer to μ^b as the *block pre-renormalization* of μ and say that the walk is *block renormalizable* whenever the block pre-renormalization is defined.

Measures which are fixed points under this action are of particular interest.

Definition 3.2.9. A measure μ is *block self-similar* if $\mu = \mathscr{R}_b(\mu)$.

3.2.4 Self-Similar Group Structure in Dyson Models

The ultrametric d defined in Definition 3.1.14 has a natural group theoretic interpretation, by identifying $\mathbb{N}$ with the group $\mathscr{G}_2 = \bigoplus_{\mathbb{N}} \mathbb{Z}_2$, the *hierarchical lattice*. A similar identification appears in [DGW05; BEI07]. This is done by identifying a minimal generating set $\{a_n\}_{n\geq 0} \subset \mathscr{G}_2$ and pushing the group structure forward through a bijection $\psi : \mathscr{G}_2 \to \mathbb{N}$ defined by

$$\psi(a_{n_1}a_{n_2}\cdots a_{n_k}) = \sum_{i=1}^{k} 2^{n_i},$$

for any $n_1, \ldots, n_k \in \mathbb{N}$. We'll write

$$x \star y \equiv \psi\left(\psi^{-1}(x)\psi^{-1}(y)\right), \tag{3.2.2}$$

to indicate this group structure on $\mathbb{N}$. We will henceforth occasionally identify $\mathbb{N}$ and $\mathscr{G}_2$ as convenient for notation.

Proposition 3.2.5. *A dyadic box $B_{0,k} \subset \mathbb{N}$ is the normal subgroup of $\mathbb{N}$ generated by $\{2^0, 2^1, \ldots, 2^{k-1}\}$ for $k \geq 1$. Dyadic boxes $B_{j,k}$ are cosets of $B_{0,k}$. Moreover, for $x > 0$,*

$$\log_2(d(0,x)) = \min\{k : x \in B_{0,k}\} - 1,$$

and $d(x,y) = d(0, x \star y^{-1})$.

The isomorphism $\mathscr{G}_2 \cong \mathscr{G}_2/B_{0,1}$ gives a metric on the group of dyadic boxes $B_{j,1}$ by

$$\tilde{d}(B_{j,1}, B_{k,1}) = d(j,k).$$

If $j \neq k$ then

$$\tilde{d}(B_{j,1}, B_{k,1}) = \frac{1}{2} d(n,m),$$

for $n \in B_{j,1}$ and $m \in B_{k,1}$. This group-theoretic perspective will permit us to regard Dyson hierarchical models as translation-invariant.

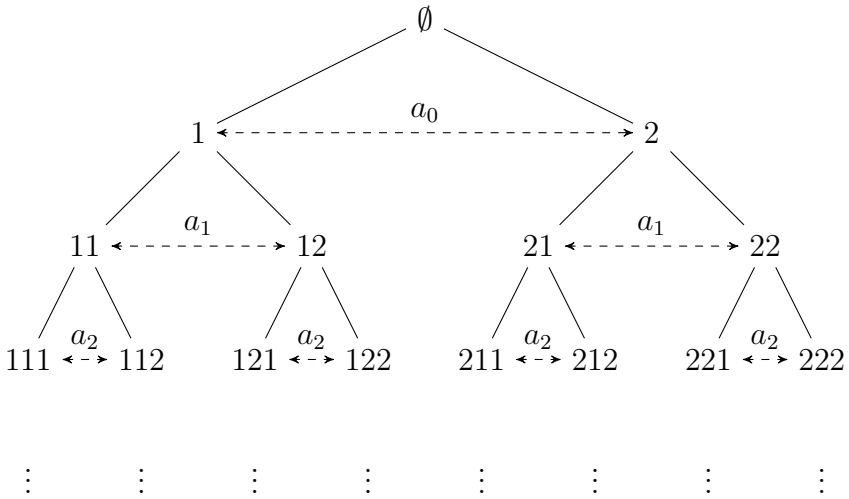

Figure 3.2: The action of $\mathscr{G}_2$ on T_2.

Proposition 3.2.6. *The group $\mathscr{G}_2$ is self-replicating and non-Hopfian.*

Proof. For $x \in T_2$ we may define the action of a sequence $\{a_n\}_{n\in\mathbb{N}}$ of automorphisms (see Figure 3.2.4) by

$$a_0(1x) = 2x \qquad a_{n+1}(1x) = 1a_n(x)$$
$$a_0(2x) = 1x \qquad a_{n+1}(2x) = 2a_n(x).$$

These automorphisms all commute and have order 2, as desired. That $\mathscr{G}_2$ is non-Hopfian comes immediately from taking

$$H = \mathbb{Z}_2 \oplus \bigoplus_{n\geq 1}\{0\}.$$

□

As such, both Schur renormalization and block renormalization operators are defined for probability measures on $\mathscr{G}_2$. In fact, $\mathscr{G}_2$ is also non-co-Hopfian, so that these two operators arise in essence from dual structures on $\mathscr{G}_2$. Random walks on $\mathscr{G}_2$ (and generalizations with $\mathbb{Z}_N$ replacing $\mathbb{Z}_2$) are examined in [DGW05], which introduces the following definition.

Definition 3.2.10. A *hierarchical random walk* is a random walk $\{g_n\}_{n\in\mathbb{N}}$ on $\mathscr{G}_2$ starting at the identity with independent, identically distributed increments having law μ, where for some probability distribution, referred to as the *level distribution*, $\{p_k\}_{k\in\mathbb{N}}$ on $\mathbb{N}$ one has for $2^k \leq n < 2^{k+1}$

$$\mu_{\psi^{-1}(n)} = 2^{-k}p_k,$$

and $\mu_0 = 0$. The space of hierarchical random walks on $\mathscr{G}_2$ will be denoted $\mathscr{H}_2$.

These walks choose a distance of 2^k in $\mathscr{G}_2$ with probability 1, and then choose an element at that distance uniformly at random. We'd like to analyze the Schur and block renormalizations of these highly symmetric walks.

Proposition 3.2.7. *$\mathscr{H}_2$ is invariant under $\mathscr{R}_b$. For a hierarchical random walk $\{g_n\}_{n\in\mathbb{N}}$ on $\mathscr{G}_2$ with law μ having associated level distribution $\{p_k\}_{k\in\mathbb{N}}$, the walk is block renormalizable if*

$p_0 \neq 1$, *and the block renormalization* $\mathscr{R}_b(\mu)$ *has a level distribution satisfying*

$$\tilde{p}_k = \frac{p_{k+1}}{1 - p_0}.$$

Proof. Provided there exists a $k \geq 1$ such that $p_k > 0$, the walk will almost surely leave $B_{0,1}$. The first transition to do so, at time $0 < \tau < \infty$, produces the same probability distribution outside $B_{0,1}$ regardless as to whether the current occupied site is 0 or a_0. As a result, the induced walk satisfies

$$\mathbb{P}\{g_\tau \in B_{j,1}\} = \frac{\mu_{\psi^{-1}(2j)} + \mu_{\psi^{-1}(2j+1)}}{1 - \mu_1}.$$

□

This immediately allows for the classification of fixed points as Markov processes whose transition probabilities decay exponentially in the level k.

Corollary 3.2.8. *A hierarchical random walk* $\{g_n\}_{n\in\mathbb{N}}$ *on* $\mathscr{G}_2$ *with law* μ *having associated level distribution* $\{p_k\}_{k\in\mathbb{N}}$ *is block self-similar if and only if* $0 < p_0 < 1$ *and*

$$p_k = p_0(1 - p_0)^k.$$

Since the hierarchical distance to level k is on the order of 2^k, the one-step transition probability depends on hierarchical distance as a power law.

For Schur renormalization, renormalizing at either index (1 or 2) produces the same operator, so we'll suppress the index and just write $\mathscr{R}_S$. The behavior of this operator is significantly more difficult to analyze than the block renormalization operator. As a result, we restrict ourselves to a proof of Schur renormalizability.

Proposition 3.2.9. *A hierarchical random walk* $\{g_n\}_{n\in\mathbb{N}}$ *on* $\mathscr{G}_2$ *with law* μ *having associated level distribution* $\{p_k\}_{k\in\mathbb{N}}$ *is Schur renormalizable provided* $p_k \to 0$ *sufficiently rapidly, with Schur pre-renormalization given by*

$$(1 + p_0)^2(e - \mu^{(2,2)})^{-1} - (1 + 2p_0)e.$$

Proof. The subtrees $1T_2$ and $2T_2$ are both fixed by $\psi^{-1}(2j)$ for any $j \in \mathbb{N}$, and $\psi^{-1}(2j)$ is sent to $\psi^{-1}(j)$ by the homomorphism $h \mapsto h|_1$. Beginning at $0 \in \mathscr{G}_2$, either the first step is to another even point or there is an unbroken sequence of odd points traversed until returning to an event point. The first case occurs with probability $\frac{1}{2}(1-p_0)$, reaching level k with probability $\frac{p_k}{2}$ for $k \geq 1$ and selecting an even point on that level uniformly at random. This gives

$$\mu^{(1,1)} = \frac{p_1}{2}a_0 + \frac{p_2}{4}(e+a_0)\,a_1 + \frac{p_3}{8}(e+a_0)(e+a_1)\,a_2 + \dots. \tag{3.2.3}$$

In the second case, we may compute directly the measures $\mu^{(1,2)}, \mu^{(2,1)}, \mu^{(2,2)} \in \ell^1(\mathscr{G}_2, \mathbb{R})$ directly. One has $\mu^{(2,2)} = \mu^{(1,1)}$, and

$$\mu^{(1,2)} = \mu^{(2,1)} = p_0 e + \mu^{(1,1)}. \tag{3.2.4}$$

Now, we wish to obtain $(e-\mu^{(2,2)})^{-1}$. In that interest, we rewrite (3.2.3) to obtain

$$e - \mu^{(2,2)} = \left(e - \sum_{k=1}^{\infty} \frac{p_{k+1}}{2^{k+1}} a_k \prod_{i=1}^{k-1}(e+a_i)\right) - \left(\frac{p_1}{2}e + \sum_{k=1}^{\infty} \frac{p_{k+1}}{2^{k+1}} a_k \prod_{i=1}^{k-1}(e+a_i)\right) a_0.$$

We will refer to the image of a point in the group algebra under the homomorphism $a_k \mapsto -a_k$ which fixes all other generators as 'a_k-conjugate'. Multiplying by r_0, the a_0-conjugate of the expression, eliminates a_0 and gives

$$\begin{aligned}
(e-\mu^{(2,2)})r_0 &= \left(e - \sum_{k=1}^{\infty} \frac{p_{k+1}}{2^{k+1}} a_k \prod_{i=1}^{k-1}(e+a_i)\right)^2 - \left(\frac{p_1}{2}e + \sum_{k=1}^{\infty} \frac{p_{k+1}}{2^{k+1}} a_k \prod_{i=1}^{k-1}(e+a_i)\right)^2, \\
&= \left(1 - \frac{p_1^2}{4}\right)e - (2+p_1)\sum_{k=1}^{\infty} \frac{p_{k+1}}{2^{k+1}} a_k \prod_{i=1}^{k-1}(e+a_i), \\
&= \left(\left(1 - \frac{p_1^2}{4}\right)e - \left(1+\frac{p_1}{2}\right)\sum_{k=2}^{\infty} \frac{p_{k+1}}{2^{k}} a_k \prod_{i=2}^{k-1}(e+a_i)\right) \\
&\quad - \left(\left(1+\frac{p_1}{2}\right)\frac{p_2}{2}e + \left(1+\frac{p_1}{2}\right)\sum_{k=2}^{\infty} \frac{p_{k+1}}{2^{k}} a_k \prod_{i=2}^{k-1}(e+a_i)\right) a_1.
\end{aligned}$$

We may then proceed by multiplying by r_1, the a_1-conjugate of this expression, in this way

eliminating a_1. In general, proceeding inductively, if we have

$$(e-\mu^{(2,2)})\prod_{i=0}^{n-1} r_i = \left(A_n e - B_n \sum_{k=n+1}^{\infty} \frac{p_{k+1}}{2^{k-n+1}} a_k \prod_{i=n+1}^{k-1} (e+a_i)\right) - \left(C_n e + B_n \sum_{k=n+1}^{\infty} \frac{p_{k+1}}{2^{k-n+1}} a_k \prod_{i=n+1}^{k-1} (e+a_i)\right) a_n,$$

then multiplying by r_n, the a_n-conjugate, yields the recurrences

$$A_{n+1} = A_n^2 - C_n^2,$$

$$B_{n+1} = B_n(A_n + C_n),$$

$$C_{n+1} = B_{n+1}\frac{p_{n+2}}{2}.$$

Then, if $D_n = \frac{B_n}{A_n}$, we have $D_0 = 1$ and

$$D_{n+1} = \frac{D_n}{1 - D_n \frac{p_{n+1}}{2}}.$$

We may directly solve this recurrence, with

$$D_n = \frac{1}{1 - \frac{1}{2}\sum_{k=1}^{n} p_k}.$$

Thus, the sequence $\{D_n\}_{n\in\mathbb{N}}$ is increasing, with

$$\lim_{n\to\infty} D_n \to D = \frac{2}{1+p_0}.$$

This implies $\frac{C_n}{A_n} \to 0$ at the same asymptotic rate as $p_n \to 0$, and hence the sequence $\{A_n\}_{n\in\mathbb{N}}$ is strictly positive. Moreover, we may conclude that for any $n \in \mathbb{N}$ that $r_n = A_n q_n$ for

$$q_n = e + \frac{2}{2 - \sum_{k=1}^{n} p_k}\left(\frac{p_{n+1}}{2} a_n - (e-a_n)\sum_{k=n+1}^{\infty} \frac{p_{k+1}}{2^{k-n+1}} a_k \prod_{i=n+1}^{k-1} (e+a_i)\right),$$

and that

$$A_n = \prod_{k=1}^{n}\left(1 - \frac{p_{k+1}}{2 - \sum_{j=1}^{k} p_j}\right)^{2^{n-k}}.$$

Let $s_n = \frac{1}{A_n}\prod_{i=0}^{n-1} r_i$; the summability of $\{p_k\}_{k\in\mathbb{N}}$ yields

$$\lim_{n\to\infty} (e-\mu^{(2,2)}) s_n \to e,$$

in $\ell^1(\mathscr{G}_2, \mathbb{R})$. Now, take $n \in \mathbb{N}$ and note that

$$s_{n+1} - s_n = \left(\frac{A_n}{A_{n+1}} r_n - e \right) s_n,$$

and

$$\left\| \frac{A_n}{A_{n+1}} r_n - e \right\| \leq \frac{1}{1 - \frac{D^2 p_{n+1}^2}{4}} \left(\frac{D^2 p_{n+1}^2}{4} + \frac{D}{2} \sum_{k=n+1}^{\infty} p_k \right).$$

For $p_k \to 0$ sufficiently rapidly, we may thus show $\|s_{n+1} - s_n\|$ to be summable, so

$$s_n \to s \equiv (e - \mu^{(2,2)})^{-1} \in \ell^1(\mathscr{G}_2, \mathbb{R}),$$

with

$$s = \prod_{k=0}^{\infty} \left(1 + \frac{p_{k+1}}{2 - \sum_{j=1}^{k+1} p_j} \right) q_k.$$

Then we may use (3.2.4) to directly compute the Schur pre-renormalization

$$\mu^{(1,1)} + \mu^{(1,2)} s \mu^{(2,1)} = (1 + p_0)^2 s - (1 + 2p_0) e.$$

□

In principle, one could obtain the law for $\mathscr{R}_S(\mu)$ based on the preceeding argument, but the task appears computationally intractable, unless a more compact way to express $(e - \mu^{(2,2)})^{-1}$ is determined. However, we are able to observe that Schur renormalization and block renormalization are, in general, different transformations.

Proposition 3.2.10. *Suppose μ is the law of a hierarchical random walk which is both Schur renormalizable and block renormalizable. It need not be the case that $\mathscr{R}_S(\mu) = \mathscr{R}_b(\mu)$. In particular, every nontrivial block self-similar μ is not Schur self-similar.*

Proof. It suffices to show that there need not exist a $u, v \in \mathbb{R}$ such that

$$ue + v\mathscr{R}_b(\mu) = (e - \mu^{(2,2)})^{-1}.$$

Let μ have a level distribution $\{p_k\}_{k\in\mathbb{N}}$. We may then apply (3.2.3), and recall that $\mu^{(2,2)} = \mu^{(1,1)}$ for a hierarchical random walk. Moreover,

$$\mathscr{R}_b(\mu) = \frac{1}{1-p_0}\sum_{k=0}^{\infty}\frac{p_{k+1}}{2^k}a_k\prod_{i=0}^{k-1}(e+a_i).$$

All that remains is to carry out the multiplication to see that

$$\begin{aligned}(ue+v\mathscr{R}_b(\mu))(e-\mu^{(2,2)})^{-1} &= \left(ue+\frac{v}{1-p_0}\sum_{k=0}^{\infty}\frac{p_{k+1}}{2^k}a_k\prod_{i=0}^{k-1}(e+a_i)\right)\left(e-\sum_{\ell=0}^{\infty}\frac{p_\ell}{2^\ell}a_\ell\prod_{j=0}^{\ell-1}(e+a_j)\right),\\ &= ue-\sum_{k=0}^{\infty}\frac{1}{2^k}\left(up_k+\frac{vp_{k+1}}{1-p_0}\right)a_k\prod_{i=0}^{k-1}(e+a_i)\\ &\quad-\frac{v}{1-p_0}\sum_{k=0}^{\infty}\sum_{\ell=0}^{\infty}\frac{p_\ell p_{k+1}}{2^{\max\{k,\ell\}}}a_ka_\ell\prod_{i=0}^{\max\{k,\ell\}-1}(e+a_i).\end{aligned}$$

We call the term on the final line g. Splitting the summation in g into components on which $k=\ell$, on which $k>\ell$, and on which $k<\ell$, we have

$$g = -\frac{v}{1-p_0}\left(\sum_{k=0}^{\infty}\frac{p_kp_{k+1}}{2^k}\prod_{i=0}^{k-1}(e+a_i)+\sum_{k=1}^{\infty}\frac{1}{2^k}\left[p_{k+1}\sum_{\ell=0}^{k-1}p_\ell+p_k\sum_{\ell=1}^{k}p_\ell\right]a_k\prod_{i=0}^{k-1}(e+a_i)\right).$$

Further, note that

$$\sum_{k=0}^{\infty}\frac{p_kp_{k+1}}{2^k}\prod_{i=0}^{k-1}(e+a_i) = \sum_{k=0}^{\infty}\frac{p_kp_{k+1}}{2^k}e+\sum_{k=1}^{\infty}\left[\sum_{j=k+1}^{\infty}\frac{p_jp_{j+1}}{2^j}\right]a_k\prod_{i=0}^{k-1}(e+a_i).$$

Combining these statements results in

$$\begin{aligned}(ue+v\mathscr{R}_b(\mu))(e-\mu^{(2,2)})^{-1} &= \left(u-\frac{v}{1-p_0}\sum_{k=0}^{\infty}\frac{p_kp_{k+1}}{2^k}\right)e-\left(up_0+\frac{vp_1}{1-p_0}\right)a_0\\ &\quad-\sum_{k=1}^{\infty}\left[\frac{1}{2^k}\left(up_k+\frac{vp_{k+1}}{1-p_0}\right)+\frac{v}{1-p_0}\sum_{j=k+1}^{\infty}\frac{p_jp_{j+1}}{2^j}\right]a_k\prod_{i=0}^{k-1}(e+a_i).\end{aligned}$$

In order to have $ue+v\mathscr{R}_b(\mu) = (e-\mu^{(2,2)})^{-1}$, we must thus have u, v a solution to the (infinite) system of equations

$$\begin{aligned}1 &= u-\frac{v}{1-p_0}\sum_{k=0}^{\infty}\frac{p_kp_{k+1}}{2^k},\\ 0 &= up_0+\frac{vp_1}{1-p_0},\\ 0 &= up_k+\frac{v}{1-p_0}\left(p_{k+1}+\sum_{j=k+1}^{\infty}\frac{p_jp_{j+1}}{2^{j-k}}\right) \qquad k\geq 1.\end{aligned}$$

If $p_0 = 0$ but $p_1 \neq 0$, this system has no solution. Alternatively, if $p_k = p_0(1-p_0)^k$ for $0 < p_0 < 1$, then the system can be written as

$$\begin{aligned} 1 &= u - \frac{2vp_0^2}{2-(1-p_0)^2}, \\ 0 &= u + v, \\ 0 &= (u+v)(1-p_0)^k + v\left(\frac{p_0(1-p_0)^{2k+2}}{2-(1-p_0)^2}\right) \qquad\qquad k \geq 1, \end{aligned}$$

which again has no solution. □

3.3 Random Walk Models

In this section, we'll closely follow the work in Chapter 9 of [FFS92] in introducing the structure of a random walk model. These models will, in essence, consist of both probabilistic information about transmission probabilities of a discrete-time random walk which may on its past and an additional layer of combinatorial information used exclusively to construct kernels for the random process. These models will be connected to pairwise type spin models.

3.3.1 Basic Terminology of Random Walk Models

Random walk models are an extremely general combinatorial structure. Here we'll introduce their basic properties.

Definition 3.3.1. A *walk* on a directed graph $G = (V, E)$ is a finite (possibly empty) ordered list of edges in E, called *steps*, such that the endpoint of any given step is the starting point of the subsequent step.

The space of such walks on G will be denoted by Ω_G, with $\Omega_{G,x} \subset \Omega_G$ consisting of those walks which begin at $x \in V$. Further, we will denote by $|\omega|$ the number of steps taken by ω. Moreover, for a walk ω which starts at $x \in V$ and ends at $y \in V$, we will write $\omega : x \to y$ and call the space of such walks $\Omega_{G,x\to y}$. In addition, for any $z \in V$, $n_z(\omega)$ is the number of

times ω touches that point, defined so that

$$\sum_{z \in V} n_z(\omega) = |\omega| + 1.$$

Frequently, we will refer to $\omega(k) \in V$ as the k-th site visited by a walk ω, with $\omega(0)$ its starting point. Finally, we will denote by $\omega_1 \circ \omega_2$ the concatenation: the list of edges formed by first traversing those of ω_1, then those of ω_2. This will only itself be a walk if the endpoint of ω_1 is the starting point of ω_2. If $e \in E$, we'll often identify e with the single-step walk $\{e\}$, for notational convenience.

Definition 3.3.2. A *random walk model* on a finite directed graph G is a collection of non-negative functions, called *weights*, $\{\rho^{[n]} : (\Omega_G)^n \to \mathbb{R}_{\geq 0}\}_{n \in \mathbb{N}}$, with $\rho^{[0]} = 1$.

The assumption that the graph G is finite is made largely for convenience, and may be neglected provided the weights satisfy sufficient summability conditions. These conditions are related to the main objects of concern in these models: the kernels.

Definition 3.3.3. The *kernels* K_{2n} of a random walk model $\{\rho^{[n]}\}_{n \in \mathbb{N}}$ are the functions

$$K_{2n}(x_1 x_2 | \dots | x_{2n-1} x_{2n}) = \sum_{\omega_1 \in \Omega_{G, x_1 \to x_2}} \cdots \sum_{\omega_n \in \Omega_{G, x_{2n-1} \to x_{2n}}} \rho^{[n]}(\omega_1, \dots, \omega_n).$$

We will, in general, assume that random walk models satisfy conditions sufficient to guarantee that these summations are convergent. Our models of interest compare these kernels with the correlation functions of lattice spin models. However, such correlation functions are inherently symmetric in the order of their arguments, whereas there is no a priori reason why these kernels should be. Thus, we must symmetrize the kernels.

Definition 3.3.4. Define a *pair permutation* π as an element π in the symmetric group on $2n$ elements such that

1. $\pi(2k-1) < \pi(2k+1)$ for all $1 \leq k \leq n-1$.

2. $\pi(2k-1) < \pi(2k)$ for all $1 \leq k \leq n$.

Write the set of such permutations as Q_{2n}.

Using pair permutations to symmetrize the kernels results in the central objects of study for random walk models.

Definition 3.3.5. The $2n$-point *Green's Functions* of a random walk model $\{\rho^{[n]}\}_{n\in\mathbb{N}}$ are the functions

$$S_{2n}(x_1,\dots,x_{2n}) = \sum_{\pi\in Q_{2n}} K_{2n}(x_{\pi(1)}x_{\pi(2)}|\dots|x_{\pi(2n-1)}x_{\pi(2n)}).$$

If one wishes to define random walk models for infinite graphs, it is necessary to assume that the summations defining these Green's functions are convergent for all n.

3.3.2 Interpretation of Random Walk Models

As the definition of a random walk model is somewhat opaque, we'll briefly discuss the interpretation of these weights. This is not entirely straightforward, especially for the higher weights $\rho^{[n]}$ for $n \geq 2$. Part of the complexity of interpretation is natural - this structure can encode a diverse family of processes, including but not limited to Markov processes, self-avoiding walks, and whole families of random walk models arising in statistical mechanics including Aizenman's random walk representation for the Ising model [Aiz82]. Broadly, we wish to regard weights applied to walks of fixed lengths and a specified starting point as 'non-normalized probabilities'.

An illustrative example to keep in mind for this discussion is the random walk model $\{\rho_\tau^{[n]}\}$ for $\tau > 0$ corresponding to a simple random walk on $\mathbb{Z}$ with edges connecting integers to their nearest neighbors, where for any n and any walks $\omega_1,\dots,\omega_n \in \Omega_{\mathbb{Z}}$

$$\rho_\tau^{[n]}(\omega_1,\dots,\omega_n) = \tau^{-\sum_{k=1}^n |\omega_k|}.$$

For this model, the corresponding kernels exist when τ is sufficiently small.

For a direct probabilistic interpretation to be coherent, we'll make an assumption about these models.

Definition 3.3.6. A random walk model on $G = (V, E)$ is said to be *unstoppable* if for any $n > 0$ and collection of walks $\omega_1, \ldots, \omega_n$ for which $\rho^{[n]}(\omega_1, \ldots, \omega_n) > 0$ and for any $1 \leq k \leq n$ there exists an edge $e \in E$ such that $\omega_k \circ e$ is a walk and

$$\rho^{[n]}(\omega_1, \ldots, \omega_k \circ e, \ldots, \omega_n) > 0.$$

Self-avoiding walks, for example, are not unstoppable. Much as in the game Snake, it is possible for such a walk to proceed for n steps in such a way that there are no possible $n+1$ steps. The models we will be concerned with are unstoppable. They will, in fact, satisfy the stronger condition that weights are strictly positive for all walks.

If we wish to interpret the first weight $\rho^{[1]}$ as being related to a probability, we could for any $x \in G$ and $t \in \mathbb{N}$ define a time t probability distribution $P_{x,t}$ on $\Omega_{G,x,t} = \{\omega \in \Omega_{G,x} : |\omega| = t\}$ by

$$P_{x,t}(\omega) = \frac{\rho^{[1]}(\omega)}{\sum_{\omega' \in \Omega_{G,x,t}} \rho^{[1]}(\omega')}.$$

It is in this sense that we regard $\rho^{[1]}$ as a non-normalized probability. One can do similarly for $\rho^{[n]}$ by taking n distinct time variables and specifying the n starting points, so that $\rho^{[n]}$ gives information about the correlation structure between multiple walks. If the random walk model represents a situation in which distinct walks are independent of one another, then for all $n \in \mathbb{N}$ one has

$$\rho^{[n]}(\omega_1, \ldots, \omega_n) = \prod_{k=1}^{n} \rho^{[1]}(\omega_k).$$

One may wish to go further and define 'one-step conditional probabilities' $P(e|\omega)$ where ω is a walk ending at the starting point of the edge e, representing the probability that after having traversing the path ω a walk then takes a step along e. A sensible definition would be

$$P(e|\omega) = \frac{\rho^{[1]}(\omega \circ e)}{\sum_{\omega' \in \Omega_{G,e(0),1}} \rho^{[1]}(\omega \circ \omega')}.$$

However, in our generality it need not be the case that this yields a coherent definition of conditional probability.

Definition 3.3.7. We say that the first weight function $\rho^{[1]}$ of a random walk model on $G = (V, E)$ is *dynamical* if for all $\omega \in \Omega_G$ and $e \in E$ starting at the endpoint of ω

$$P(e|\omega) = \frac{P_{x,|\omega|+1}(\omega \circ e)}{P_{x,|\omega|}(\omega)} \in [0, 1].$$

Under this condition, we may truly regard random walk models as being generated 'one step at a time' by a probabilistic procedure. This interpretation amounts to little more than a restatement of the Kolmogorov extension theorem. One could produce similar definitions for higher weight functions as well. However, our models may not necessarily have dynamical weight functions.

In general, the Green's functions are not directly related to their probabilistic equivalents, even if the model admits a direct probabilistic interpretation. But, if for some discrete-time Markov process on G, $\rho^{[1]}(\omega)$ is the (normalized) probability of traversing the walk ω after starting at its initial point, then S_2 coincides with the usual definition of a Green's function, see for example [CY00].

3.3.3 Random Walk Representations of Spin Systems

The work of [BFS82] allows us to translate a pairwise type spin model into a random walk model in what is referred to as the Brydges-Fröhlich-Spencer (BFS) representation of a spin model. In this section we'll explain this method.

A necessary first step to define the BFS representation is to introduce an auxiliary family of measures which arise from the Edwards model (originally found in [Edw65]).

Definition 3.3.8. Define the measures η_n on $\mathbb{R}$ by $\eta_0 = \delta_0$ and for $n > 0$

$$d\eta_n(t) = \frac{t^{n-1}}{\Gamma(n)} \chi_{[0,\infty)}(t) dt.$$

For a graph G and a walk $\omega \in \Omega_G$, let η_ω be the measure on $\mathbb{R}^G$ given by

$$\eta_\omega(\bar{t}) = \prod_{x \in G} d\eta_{n_x(\omega)}(\bar{t}_x).$$

We will also have need of the following notation.

Definition 3.3.9. For $G = (V, E)$ a directed graph, $\omega \in \Omega_G$, and $J_{x,y} \in \mathbb{R}$ defined for any $(x, y) \in E$, we'll write

$$J^\omega = \prod_{(x,y)\in\omega} J_{x,y}.$$

Further, for any $\omega' \in \Omega_G$, we'll write $J^{\omega+\omega'} = J^\omega J^{\omega'}$.

With this notation in mind, the following theorem comes from [BFS82; FFS92].

Theorem 3.3.1 (BFS Random Walk Representation)**.** *For any scalar-valued spin model (H, ν) on G of pairwise type with absolutely continuous a priori measures $d\nu(z) = g(z^2)dz$ and partition function $\mathcal{Z}$, define a map $\mathcal{W} : \mathbb{R}^G \to \mathbb{R}$ by*

$$\mathcal{W}(\bar{t}) = \mathcal{Z}^{-1} \int_{\mathbb{R}^G} e^{\frac{1}{2}(\varphi, J\varphi)} \prod_x g(\varphi_x^2 + 2\bar{t}_x) d\varphi_x.$$

Then the random walk model with

$$\rho^{[n]}_{H,\nu}(\omega_1, \ldots, \omega_n) = J^{\omega_1+\ldots+\omega_n} \int_{(\mathbb{R}^G)^n} \mathcal{W}(\bar{t}_1 + \ldots + \bar{t}_n) \prod_{i=1}^n d\eta_{\omega_i}(\bar{t}_i),$$

satisfies for any $x_1, \ldots, x_n \in G$

$$\langle \varphi_{x_1} \cdots \varphi_{x_n} \rangle = S_{2n}(x_1, \ldots, x_{2n}).$$

3.3.4 Random Walk Models for Gaussian Dyson Models

For Dyson hierarchical models with Gaussian a priori measures, the form of the corresponding random walk model simplifies dramatically. We will denote the weight functions corresponding to a Dyson model with Gaussian a priori measure with variance σ and interaction function f as $\{\rho^{[n]}_{\sigma,f}\}_{n\in\mathbb{N}}$. In the case that $f(n) = n^{-a}$, we may alternatively write $\{\rho^{[n]}_{\sigma,a}\}_{n\in\mathbb{N}}$. We will generally suppress dependence on the graph V_N in notation, with the tacit understanding that results hold for all N and ultimately in the limit as $N \to \infty$.

Proposition 3.3.2. *For a truncated Dyson hierarchical model on V_N with interaction function $f : \mathbb{N} \to \mathbb{R}_{\geq 0}$ and Gaussian a priori measures with variance $\sigma > 0$, one has*

$$\rho^{[n]}_{\sigma,f}(\omega_1, \ldots, \omega_n) = \prod_{i=1}^{n} \rho^{[1]}_{\sigma,f}(\omega_i),$$

for all $\omega_i \in \Omega_{V_N}$ for $i = 1, \ldots, n$. Moreover, for any $\omega \in \Omega_{V_N}$

$$\rho^{[1]}_{\sigma,f}(\omega) = \sigma^{|\omega|+1} \prod_{(x,y)\in\omega} f(d(x,y)). \tag{3.3.1}$$

Proof. Here, one has

$$g(s) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{s}{2\sigma}}.$$

Accordingly, one may simplify

$$\mathcal{W}(\bar{t}) = \left[\prod_{z\in V_N} e^{-\frac{\bar{t}_z}{\sigma}}\right] \mathcal{Z}^{-1} \int e^{\frac{1}{2}(\varphi, J\varphi)} \prod_{z\in V_N} g(\varphi_z^2) d\varphi_z = \prod_{z\in V_N} e^{-\frac{\bar{t}_z}{\sigma}}.$$

As a direct consequence of this, we see that $\mathcal{W}(\bar{t}_1 + \bar{t}_2) = \mathcal{W}(\bar{t}_1) \cdot \mathcal{W}(\bar{t}_2)$, which means that

$$\rho^{[n]}_{\sigma,f}(\omega_1, \ldots, \omega_n) = \prod_{i=1}^{n} \rho^{[1]}_{\sigma,f}(\omega_i).$$

It will thus suffice going forward to consider only $\rho^{[1]}_{\sigma,f}$, which has the form

$$\begin{aligned}
\rho^{[1]}_{\sigma,f}(\omega) &= \prod_{(x,y)\in\omega} f(d(x,y)) \int \prod_{z\in V_N} e^{-\frac{t_z}{\sigma}} d\eta_{n_z}(t_z), \\
&= \prod_{(x,y)\in\omega} f(d(x,y)) \prod_{z\in V_N} \sigma^{n_z}, \\
&= \sigma^{|\omega|+1} \prod_{(x,y)\in\omega} f(d(x,y)).
\end{aligned}$$

□

The Gaussian case of the Dyson hierarchical model is exceptional in having all distinct walks independent. We are also able to interpret this random walk model in a more conventional probabilistic manner.

Proposition 3.3.3. *The weight function $\rho^{[1]}_{\sigma,f}$ is dynamical in the sense of Definition 3.3.7. Moreover, the conditional probability of making a step from $x \to y$ in V_N after a walk $\omega : z \to x$ depends only on $x \star y$ as defined in* (3.2.2)*. This induces a random walk on $\mathscr{G}_2$ with independent increments each having law $\mu_{f,N}$ supported on $B_{0,N}$ which does not depend on σ. Provided f is of sufficient decay that the thermodynamic limit of the Dyson hierarchical model is defined, $\mu_{f,N} \to \mu_f$ in $\ell^1(\mathscr{G}_2, \mathbb{R})$ where μ_f is the law of a hierarchical random walk on $\mathscr{G}_2$ with level distribution $\{p_k\}_{k\in\mathbb{N}}$ satisfying*

$$p_k = \frac{2^k f(2^k)}{\sum_{n=0}^{\infty} 2^n f(2^n)}.$$

Proof. Verifying that the weight function is dynamical follows from a simple computation. As $d(x, y) = d(0, x \star y)$, the independence of conditional probability from all but $x \star y$ follows from (3.3.1). The induced random walk from the model on V_N is hierarchical with law $\mu_{f,N}$ having a level distribution $\{p_{k,N}\}_{k\in\mathbb{N}}$ which satisfies $p_{k,N} = 0$ if $k > N$ and otherwise

$$p_{k,N} \propto 2^k f(2^k).$$

The existence of a thermodynamic limit implies $\sum_{n=0}^{\infty} 2^n f(2^n) < \infty$, which is sufficient to conclude that $\{\mu_{f,N}\}_{N\in\mathbb{N}}$ is Cauchy in $\ell^1(\mathscr{G}_2, \mathbb{R})$. □

3.3.5 Green's Functions of the Gaussian Dyson Model

In the case of a Dyson hierarchical model with power law interactions and Gaussian a priori measures each with variance σ, we can explicitly determine the Green's function $S_2 = K_2 \equiv K^N_\sigma$ (we will use this notation going forward, with the dependence on the power law a being implicit) of the corresponding random walk model on V_N. Though this work was done independently, a very similar method and some of the results of this calculation, particularly Proposition 3.3.5, were subsequently found to have been completed earlier in [AT17]. Their results apply within the slightly narrower context of Markov processes related to Dyson hierarchical models and uses a slightly different definition of the Dyson hierarchical model which is identical in the thermodynamic limit.

In this section, we'll take

$$d\nu(x) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{x^2}{2\sigma}} dx,$$

and $\mathbf{J}_N$ the *interaction matrix* for the truncated Dyson model on V_N with entries

$$J_{x,y} = \begin{cases} 0 & x = y, \\ d(x,y)^{-a} & x \neq y, \end{cases}$$

for $1 < a < 2$. We will suppress the subscripts of $\rho^{[1]}$ and the dependence of K_σ^N on σ, a for ease of notation. The following lemma directly relates the behavior of the truncated Dyson hierarchical model kernel K_σ^N to the behavior of $\mathbf{J}_n$ in a way analogous to the relationship between the Green's function of a Markovian random process and a graph Laplacian.

Lemma 3.3.4. *Take e_x the x-th coordinate vector in $\mathbb{R}^{2^N}$ and $\mathbf{J}_N$ the interaction matrix for the truncated Dyson model on V_N with power law parameter a. For fixed x, the kernel $K_\sigma^N(x,y) \equiv b_y$ is given as the solution, if it exists, to the linear set of equations*

$$-e_x = \left(\mathbf{J}_N - \frac{1}{\sigma}\right) b_y.$$

Proof. We may rewrite the standard summation for the kernel by conditioning on the last step made by the random walk. In doing so, we obtain

$$\begin{aligned} K_\sigma^N(x,y) &= \sum_{\omega: x \to y} \rho_{\sigma,a}^{[1]}(\omega), \\ &= \sigma\delta_{x,y} + \sum_{z \neq y} \frac{\sigma}{d(z,y)^a} \sum_{\omega': x \to z} \rho^{[1]}(\omega'), \\ &= \sigma\delta_{x,y} + \sum_{z \neq y} \frac{\sigma}{d(z,y)^a} K_\sigma^N(x,z), \end{aligned}$$

where the first term is the walk from x to itself with no steps, included only in the case $x = y$. Thus we have a linear system of equations for $b \in \mathbb{R}^{2^N}$. This may be rewritten as

$$-e_x = \left(\mathbf{J}_N - \frac{1}{\sigma}\right) b_y.$$

□

To carry out this computation explicitly, we must analyze the spectrum of $\mathbf{J}_N$ explicitly. Before doing so, we introduce a useful bit of notation.

Definition 3.3.10. Given a dyadic block $B_{j,k}$, we say that the *left* of $B_{j,k}$ is $B_{2j,k-1}$ and the *right* of $B_{j,k}$ is $B_{2j+1,k-1}$.

We are then able to fully characterize the spectrum and eigenvectors of $\mathbf{J}_n$.

Proposition 3.3.5. *$\mathbf{J}_N$ is diagonalizable, with one eigenvector given by the all-ones vector with corresponding eigenvalue*

$$\mu_N = \sum_{i=0}^{N-1} (2^{1-a})^i = \frac{1-(2^{1-a})^N}{1-2^{1-a}}.$$

All other eigenvalues are of the form, for $1 \leq k \leq N$

$$\lambda_k = -(2^{1-a})^{k-1} + \sum_{i=0}^{k-2} (2^{1-a})^i = \frac{1-2\left(1-2^{-a}\right)\left(2^{1-a}\right)^{k-1}}{1-2^{1-a}},$$

where the sum is taken to be empty if $k = 1$. Each such eigenvalue occurs with multiplicity 2^{N-k}, and has eigenvectors v_k^j each supported on a single dyadic block $B_{j,k}$, equal to -1 on the left of the block and 1 on the right of the block. These given eigenvectors form an orthogonal basis for $\mathbb{R}^{2^N}$.

Proof. Here it is most convenient to work by induction on N. In the $N = 1$ case, $\mathbf{J}_1$ is a 2x2 matrix consisting of 0s on the main diagonal and 1s elsewhere. This makes the claim immediate.

Following that case, $\mathbf{J}_{N+1}$ has a convenient form when written as a 2x2 matrix of 2^Nx2^N blocks. Let $\mathbf{L}_N$ denote the 2^Nx2^N matrix with every entry equal to 1. Then

$$\mathbf{J}_{N+1} = \begin{pmatrix} \mathbf{J}_N & (2^{-a})^N \mathbf{L}_N \\ (2^{-a})^N \mathbf{L}_N & \mathbf{J}_N \end{pmatrix}.$$

Thus, if λ is an eigenvalue of $\mathbf{J}_N$ with an eigenvector v such that the sum over the entries of v is 0, then λ is also an eigenvalue of $\mathbf{J}_{N+1}$ with two corresponding independent eigenvectors

$v_1, v_2 \in \mathbb{R}^{2^{N+1}}$ written explicitly as

$$v_1 = \begin{pmatrix} v \\ \vec{0} \end{pmatrix}, \qquad v_2 = \begin{pmatrix} \vec{0} \\ v \end{pmatrix}.$$

Further, let w_1 be a vector equal to -1 for the first 2^N entries and equal to 1 for the final 2^N entries. Since the restriction of w_1 to its first (and last) 2^N entries is an eigenvector of $\mathbf{J}_N$, say with eigenvalue λ, then

$$\mathbf{J}_{N+1} w_1 = \begin{pmatrix} -\lambda\vec{1} + (2^{1-a})^N \vec{1} \\ -(2^{1-a})^N \vec{1} + \lambda\vec{1} \end{pmatrix} = \left(\lambda - (2^{1-a})^N\right) w_1.$$

Thus w_1 is an eigenvector of $\mathbf{J}_{N+1}$ with eigenvalue $\lambda - (2^{1-a})^N$. Finally, let w_2 be the all 1s vector in $\mathbb{R}^{2^{N+1}}$. Again noting that the restriction of w_2 to its first (and last) 2^N entries is an eigenvector of $\mathbf{J}_N$, say with eigenvalue λ, the same computation as previously, mutatis mutandis, shows that w_2 is an eigenvector of $\mathbf{J}_{N+1}$ with eigenvalue $\lambda + (2^{1-a})^N$. This completes the proof. □

With this information on $\mathbf{J}_N$ in hand, we can write an exact solution for the Green's function.

Proposition 3.3.6. *The Green's function $K_\sigma^N(x, y)$ on V_N is given by*

$$K_\sigma^N(x,y) = -\frac{1}{2^N}\left(\mu_N - \frac{1}{\sigma}\right)^{-1} + \frac{1-\delta_{x,y}}{2^{\ell(x,y)}}\left(\lambda_{\ell(x,y)} - \frac{1}{\sigma}\right)^{-1} - \sum_{k=\ell(x,y)+1}^{N} \frac{1}{2^k}\left(\lambda_k - \frac{1}{\sigma}\right)^{-1},$$

where the final sum is taken to be empty if $\ell(x, y) = N$ and the Kronecker delta causes the second term to not be included if $x = y$.

Proof. Noting the symmetry properties of the Dyson model, it suffices to take $x = 0$. Given the previous lemma, let P_N be the matrix whose rows are normalized eigenvectors of $\mathbf{J}_N$, ordered first corresponding to the ordering of eigenvalues

$$\lambda_1, \lambda_2, \ldots, \lambda_N, \mu_N,$$

and then by the natural ordering of dyadic blocks of each given size. Let D_N be the corresponding diagonalization of $\mathbf{J}_N$. Then,

$$b = -P_N^{-1}\left(D_N - \frac{1}{\sigma}\right)^{-1} P_N e_1.$$

First, $P_N e_1$ is the first column of P_N. This yields a vector which is 0 except in the last row, where it takes the value $2^{-\frac{N}{2}}$, and in each row which corresponds to a block of size 2^k containing 0, where it takes the value $-2^{-\frac{k}{2}}$. The diagonal matrix serves only to multiply each of these values by the corresponding inverse of the value $\lambda - \frac{1}{\sigma}$ for that row. Finally, multiplying by $-P_N^{-1}$ takes a linear combination of the corresponding eigenvectors. All this is to say that the result is of the form

$$b = -\frac{1}{2^N}\left(\mu_N - \frac{1}{\sigma}\right)^{-1}\vec{1} + \sum_{k=1}^{N}\frac{1}{2^k}\left(\lambda_k - \frac{1}{\sigma}\right)^{-1} v_k^0,$$

since the 0-th block of every size contains 1. For any fixed $y \in V_N$, we need only determine for which k one has $(v_k^0)_y$ nonzero. If $k < \ell(x,y)$, then 1 and y are in different blocks of size 2^k, so the term is zero. If $k = \ell(x,y)$, then y lies on the right of the block, so $(v_k^0)_y = 1$, unless $k = 0$, as there are no blocks of size 2^0. For k larger, $(v_k^0)_y = -1$. □

Taking $K_\sigma^N(x,y)$ for fixed x, y as a function of σ, we may establish analyticity by providing estimates on the eigenvalues.

Lemma 3.3.7. *The eigenvalues $\mu_N, \{\lambda_k\}$ of the interaction matrix $\mathbf{J}_N$ satisfy*

$$-1 = \lambda_1 < 0 < \lambda_2 < \ldots < \lambda_N < \mu_N < \frac{1}{1-2^{1-a}}.$$

As a result, for any $x, y \in V_N$, $K_\sigma^N(x,y)$ (viewed formally as a function of σ) is analytic for $|\sigma| \leq 1 - 2^{1-a}$.

Proof. Recalling the form of these eigenvalues, we see immediately that

$$0 < \mu_N = \frac{1-(2^{1-a})^N}{1-2^{1-a}} < \frac{1}{1-2^{1-a}},$$

for any $N \geq 1$. For

$$\lambda_k = \frac{1 - 2\left(1 - 2^{-a}\right)\left(2^{1-a}\right)^{k-1}}{1 - 2^{1-a}},$$

the $k = 1$ case can be seen immediately. For larger $k \geq 2$, the numerator is positive and less than 1 because

$$0 \leq 2(1 - 2^{-a})2^{1-a} = \frac{2^a - 1}{(2^{a-1})^2} \leq 1,$$

and the numerator of λ_k is given by this expression multiplied by $(2^{1-a})^{k-2}$, which is itself less than 1. This results in successively larger numerators for each subsequent λ_k. Finally, since

$$2(1 - 2^{-a})(2^{1-a})^{N-1} = \frac{2 - 2^{1-a}}{2^{1-a}}(2^{1-a})^N > (2^{1-a})^N,$$

we conclude that $\lambda_N < \mu_N$. □

With this in mind, we have the ability to take a thermodynamic limit as $N \to \infty$ and obtain the Green's function for the full Dyson hierarchical model.

Corollary 3.3.8. *For each $x, y \in \mathbb{N}$, the Green's function's $K_\sigma^N(x, y)$ on V_N converge uniformly (in σ) on compact subsets of $[0, 1 - 2^{1-a})$ as $N \to \infty$ to the Green's function $K_\sigma(x, y)$ on $\mathbb{N}$. For each $x, y \in \mathbb{N}$, $K_\sigma(x, y)$ is left-continuous at $\sigma = 1 - 2^{1-a}$.*

3.3.5.1 Asymptotics of the Self-Correlation at a Point

There are two regimes of interest to analyze the asymptotic behavior of this Green's function: in the 'high temperature limit' $\sigma \searrow 0$ and near the phase transition as $\sigma \nearrow 1 - 2^{1-a}$. For the high temperature limit we may represent the Green's function at two identical points by a convergent Taylor series expansion.

Proposition 3.3.9. *If $0 < \sigma < 1 - 2^{1-a}$ and $N \geq 2$, then for all $x \in V_N$, we have*

$$K_\sigma^N(x, x) = \sum_{n=0}^{\infty} \left[\frac{1}{2^N}(\mu_N)^n + \sum_{k=1}^{N} \frac{1}{2^k}(\lambda_k)^n \right] \sigma^{n+1}.$$

In particular, $K_\sigma^N(x,x) = \sigma + O(\sigma^3)$. *Moreover, for any fixed* $0 < \sigma' < 1 - 2^{1-a}$, *then there exists positive constants* C_1, C_2 *which depend on* a *and* σ' *but do not depend on* N *such that*

$$C_1 \leq \frac{\partial^3 K_\sigma^N(x,x)}{\partial \sigma^3}(\sigma) \leq C_2,$$

for all $0 < \sigma < \sigma'$.

Proof. Using Proposition 3.3.6 immediately yields

$$\begin{aligned} K_\sigma^N(x,x) &= -\frac{1}{2^N}\left(\mu_N - \frac{1}{\sigma}\right)^{-1} - \sum_{k=1}^{N} \frac{1}{2^k}\left(\lambda_k - \frac{1}{\sigma}\right)^{-1}, \\ &= \sum_{n=0}^{\infty} \left[\frac{1}{2^N}(\mu_N)^n + \sum_{k=1}^{N} \frac{1}{2^k}(\lambda_k)^n\right] \sigma^{n+1}. \end{aligned}$$

A direct computation of the terms of this series up to order 2 yields

$$K_\sigma^N(x,x) = \sigma + O(\sigma^3).$$

Our estimates on eigenvalues imply that for $\lambda = \lambda_k$ or $\lambda = \mu_N$ we have

$$\lambda\sigma - 1 \leq \frac{\sigma - (1 - 2^{1-a})}{1 - 2^{1-a}}.$$

Finally, taking three derivatives and applying the estimates on eigenvalues results in

$$\begin{aligned} \frac{\partial^3 K_\sigma^N(x,x)}{\partial \sigma^3} &= \frac{6\mu_N^2}{2^N(\mu_N\sigma - 1)^4} + \sum_{k=1}^{N} \frac{6\lambda_k^2}{2^k(\lambda_k\sigma - 1)^4}, \\ &\leq \frac{6}{(1-2^{1-a})^2}\left(\frac{1-2^{1-a}}{\sigma - (1-2^{1-a})}\right)^4 \left[\frac{1}{2^N} + \sum_{k=1}^{N} \frac{1}{2^k}\right], \\ &= \frac{6}{(1-2^{1-a})^2}\left(\frac{1-2^{1-a}}{\sigma - (1-2^{1-a})}\right)^4. \end{aligned}$$

This gives the desired upper bound independent of N. To obtain a lower bound, observe that for $N \geq 2$ that,

$$\begin{aligned}
\frac{\partial^3 K_\sigma^N(x,x)}{\partial \sigma^3} &> 6\mu_N^2 \left[\frac{1}{2^N(\mu_N\sigma - 1)^4} + \sum_{k=1}^{N} \frac{1}{2^k(\lambda_k\sigma - 1)^4} \right], \\
&> \frac{6\mu_N^2}{(\lambda_2\sigma - 1)^4} \left[\frac{1}{2^N} + \sum_{k=2}^{N} \frac{1}{2^k} \right] - \frac{3\mu_N^2}{(\lambda_1\sigma - 1)^4}, \\
&= 3\mu_N^2 \left[\frac{1}{(\lambda_2\sigma - 1)^4} - \frac{1}{(\sigma + 1)^4} \right], \\
&> \frac{3}{(1 - 2^{1-a})^2} \left[\frac{1}{((1 - 2^{1-a})^2 - 1)^4} - 1 \right].
\end{aligned}$$

This expression being positive, we may take it as the lower bound, again independent of N. □

Corollary 3.3.10. *The previous proposition also holds in the infinite-volume limit, as* $N \to \infty$.

We can also analyze the Green's function between a point and itself when σ is close to the critical point $\sigma_* = 1 - 2^{1-a}$. To do so, we'll introduce the variable $s(\sigma) = \frac{\sigma_* - \sigma}{\sigma_*}$ and write

$$k_{s(\sigma)}(x,x) \equiv K_\sigma(x,x)$$

for any $x \in \mathbb{N}$.

Lemma 3.3.11. *The Green's function* $k_s(x,x)$ *for* x *in the infinite-volume Dyson model is of the form*

$$k_s(x,x) = \frac{2^{-a}(1 - 2^{1-a})}{1 - 2^{-a}} \sum_{k=1}^{\infty} (2^{a-2})^k \frac{1 - s}{1 + \left(\frac{2^{-a}}{1 - 2^{-a}} (2^{a-1})^k - 1 \right) s},$$

provided $s > 0$.

This representation allows us to investigate the asymptotic behavior of $k_s(x,x)$ as $s \searrow 0$, depending on the value of a.

Proposition 3.3.12. *For any $x \in \mathbb{N}$, then to leading order in s about 0, when $s < 0$, we have for some constants $C_1(a)$ and $C_2(a)$ depending only on a that, if $1 < a < \frac{3}{2}$ then*

$$k_0(x,x) - k_s(x,x) \sim C_1(a)s,$$

and if $\frac{3}{2} < a < 2$ then

$$k_0(x,y) - k_s(x,y) \sim C_2(a) \cdot s^{\frac{1}{a-1}-1},$$

and this expression scales as $-C_{\frac{3}{2}} s \log_2(s)$ for some constant $C_{\frac{3}{2}} > 0$ to leading order if $a = \frac{3}{2}$.

Proof. Direct manipulation shows that

$$k_0(x,x) - k_s(x,x) = \frac{2^{-a}(1-2^{1-a})}{1-2^{-a}} \sum_{k=1}^{\infty} (2^{2a-3})^k \frac{s}{1 + \frac{2^{-a}}{1-2^{-a}} \left[(2^{a-1})^k - 2^a(1-2^{-a})\right] s}.$$

If we only wish to determine the leading order behavior in s, the second term in brackets may be neglected, as its exclusion causes an error less than the above by a factor of s. Thus

$$k_0(x,y) - k_s(x,y) = \frac{2^{-a}(1-2^{1-a})}{1-2^{-a}} \sum_{k=1}^{\infty} (2^{2a-3})^k \frac{s}{1 + \frac{2^{-a}}{1-2^{-a}} (2^{a-1})^k s} + \mathcal{O}(s^2).$$

This sum can be split into two smaller sums, the first consisting of those k sufficiently small that $\left|(2^{a-1})^k s\right| < 1$, and the latter containing all other terms. The value k^* for which this transition occurs is

$$k^* = \left\lfloor \frac{1}{a-1} \log_2 \left(\frac{1}{s} \right) \right\rfloor.$$

The floor can effectively be ignored at the cost of a constant factor error (the larger of 2^{2a-3} and 2^{a-1}). In the sum over small values of k, we have $0 < (2^{a-1})^k s < 1$, so provided s is sufficiently small that $k^* \gg 1$ we have

$$\frac{s}{1 + \frac{2^{-a}}{1-2^{-a}}} \sum_{k=1}^{k^*} (2^{2a-3})^k < \sum_{k=1}^{k^*} (2^{2a-3})^k \frac{s}{1 + \frac{2^{-a}}{1-2^{-a}} (2^{a-1})^k s} < s \sum_{k=1}^{k^*} (2^{2a-3})^k,$$

and provided $a \neq \frac{3}{2}$ we have

$$\sum_{k=1}^{k^*} (2^{2a-3})^k = \frac{2^{2a-3} - 2^{2a-3} \cdot s^{\frac{1}{a-1}-2}}{1 - 2^{2a-3}}.$$

Thus, this sum over the first few terms is on the order of s^1 if $\frac{1}{a-1} - 1 > 1$, hence if $a < \frac{3}{2}$, and on the order of $s^{\frac{1}{a-1}-1}$ if $a > \frac{3}{2}$. It is on the order of $-s\log_2(s)$ if $a = \frac{3}{2}$ as the sum is simply equal to k_*. On the other hand, the second sum over larger values of k may also be bounded as

$$\begin{aligned}\sum_{k=k^*}^{\infty}(2^{2a-3})^k \frac{s}{1+\frac{2^{-a}}{1-2^{-a}}(2^{a-1})^k s} &< \frac{1-2^{-a}}{2^{-a}}\sum_{k=k^*}^{\infty}\frac{(2^{2a-3})^k}{(2^{a-1})^k}, \\ &< \frac{1-2^{-a}}{2^{-a}(1-2^{a-2})}s^{\frac{1}{a-1}-1}.\end{aligned}$$

This term thus contributes at no lower order in s than the previous term. □

We will pursue an investigation of the asymptotic behavior of the Green's function at two different points in a subsequent section after developing some further renormalization theory. Even though in principle the arguments above can be generalized to discuss them directly, this discussion is made significantly simpler through the use of renormalization methods.

3.3.6 Block Renormalization of Random Walk Models

In this section, we'll introduce a renormalization operator designed to mimic the behavior of the block renormalization operator in Definition 3.2.8. We will analyze the behavior of this operator and ultimately show that in the Gaussian measure case, it coincides with the two block renormalization operators already defined in this chapter.

For all of the following, we fix $N > 0$ and take walks to be in Ω_{V_N} unless otherwise specified.

Definition 3.3.11. We say that ω is *block primitive* if ω does not contain steps of length 1.

Lemma 3.3.13. *We may uniquely associate to ω a block primitive random walk $\overline{\omega}$ by removing steps (x,y) of length 1 from ω and replacing the step (y,z) immediately following each with the step (x,z), or if $x = z$ by removing both steps. Notably,*

$$\prod_{(x,y)\in\omega} f(d(x,y)) = \prod_{(x,y)\in\overline{\omega}} f(d(x,y)),$$

for any $f : \mathbb{N} \to \mathbb{R}_{\geq 0}$.

Definition 3.3.12. Define $k_\omega = |\omega| - |\overline{\omega}|$ and the equivalence relation $\omega_1 \sim \omega_2$ if $\overline{\omega}_1 = \overline{\omega}_2$.

Definition 3.3.13. The *Block transformation* of a random walk ω the map $T_b : \Omega_{V_N} \to \Omega_{V_{N-1}}$ with $T_b(\omega) = \widetilde{\omega}$, defined as the unique walk satisfying the conditions that $|\widetilde{\omega}| = |\overline{\omega}|$ and for all i,

$$\widetilde{\omega}(i) = \widetilde{\overline{\omega}(i)}.$$

In particular, $\widetilde{\omega}$ takes a single step each time ω enters a new block.

Of course, this operator is not injective on Ω_{V_N}. However, we may fully characterize the set of random walks mapped to the same result by a block transformation.

Lemma 3.3.14. *Two random walks* ω_1 *and* ω_2 *on* V_N *satisfy* $T_b(\omega_1) = T_b(\omega_2)$ *if and only if* $|\overline{\omega}_1| = |\overline{\omega}_2|$ *and* $d(\overline{\omega}_1(i), \overline{\omega}_2(i)) \leq 1$ *for all* i. *We say that* $\omega_1 \sim_b \omega_2$ *when these conditions are satisfied.*

It is convenient to choose a representative from each $\sim_b$ equivalence class so as to define a right inverse to T_b.

Definition 3.3.14. We say that ω is an *even walk* if $\omega(i)$ is even for all i. Define $S : \Omega_{V_{N-1}} \to \Omega_{V_N}$ by $S : \widetilde{\omega} \to S(\widetilde{\omega})$ an even walk such that $T_b \circ S = \mathrm{id}_{\Omega_{V_{N-1}}}$.

This definition is made unambiguous by the following lemma.

Lemma 3.3.15. *There exists a unique even walk in any* $\sim_b$ *equivalence class.*

Given the block transformation, it is natural to define a block renormalization operator on the space of random walk models by taking the cumulative weight contained in an equivalence class.

Definition 3.3.15. The *block renormalization operator for random walk models* with parameter $a \in \mathbb{R}$ is the map from random walk models on V_N to random walk models on V_{N-1}

defined by $\hat{\mathscr{R}}_{b,a}\left(\{\rho^{[n]}\}_{n\in\mathbb{N}}\right) = \{\widetilde{\rho}^{[n]}\}_{n\in\mathbb{N}}$ where

$$\widetilde{\rho}^{[n]}(\widetilde{\omega}_1,\ldots,\widetilde{\omega}_n) = (2^{-a})^n \sum_{\{\omega_i\}\in\prod_i [S(\widetilde{\omega}_i)]_{\sim_b}} \rho^{[n]}(\omega_1,\ldots,\omega_n).$$

For ease of notation we will also write $\hat{\mathscr{R}}_{b,a}(\rho^{[n]}) = \widetilde{\rho}^{[n]}$ for each individual weight function.

3.3.6.1 Block Renormalization for Gaussian Dyson Hierarchical Random Walk Models

In this section, we discuss the block renormalization of random walk models coming from Gaussian Dyson hierarchical models, defined as in (3.3.1). Following Proposition 3.3.2, it suffices to analyze $\rho^{[1]}_{\sigma,f}$. Changing the variance σ results in a simple change to the weight function.

Lemma 3.3.16. *For any $\sigma,\sigma' > 0$, interaction function f, and walk ω*

$$\rho^{[1]}_{\sigma,f}(\omega) = \left(\frac{\sigma}{\sigma'}\right)^{|\omega|+1} \rho^{[1]}_{\sigma',f}(\omega).$$

The next lemma describes the relationship between the equivalence relations $\sim$ and $\sim_b$ and the Gaussian Dyson hierarchical model weight function.

Lemma 3.3.17. *Every block primitive element in any fixed equivalence class $[\omega]_{\sim_b}$ has the same length, $|\overline{\omega}|$, and the same weight. There are $2^{|\overline{\omega}|+1}$ such block primitive elements in a $\sim_b$ equivalence class.*

The following lemma will allow us to reduce sums over weights in a $\sim$ equivalence class to analyzing the block-primitive case.

Lemma 3.3.18. *For any fixed block primitive random walk $\overline{\omega}$, one has*

$$\sum_{\omega\in[\overline{\omega}]_\sim} \rho^{[1]}_{\sigma,f}(\omega) = \rho^{[1]}_{\sigma,f}(\overline{\omega}) \left(\frac{1}{1-\sigma}\right)^{|\omega|+1}.$$

Proof. Any walk $\omega \in [\overline{\omega}]_\sim$ is identical to $\overline{\omega}$, albeit with the inclusion of k_ω steps of length 1 (each causing the path to oscillate within a single block, but not affecting the larger structure).

Lemma 3.3.13 implies that two walks in this equivalence class with the same value of k_ω have the same weight. We can condition on k_ω to write

$$\sum_{\omega\in[\overline{\omega}]_\sim} \rho^{[1]}_{\sigma,f}(\omega) = \sigma\cdot\sigma^{|\overline{\omega}|}\prod_{(x,y)\in\overline{\omega}} f(d(x,y))\sum_{k=0}^{\infty} \#\left\{\text{walks in } [\overline{\omega}]_\sim \text{ with } k_\omega = k\right\}\cdot\sigma^k.$$

The term outside the summation is manifestly $\rho^{[1]}(\overline{\omega})$. To determine the number of walks in $[\overline{\omega}]$ with $k_\omega = k$, it suffices to note that adding in these k length-1 steps requires placing them between two steps of $\overline{\omega}$ (or at the beginning or end), that they are indistinguishable, and that the order they are placed does not matter. This is the number of ways to write k as a sum of $|\overline{\omega}|+1$ non-negative integers, so

$$\sum_{\omega\in[\overline{\omega}]_\sim} \rho^{[1]}_{\sigma,f}(\omega) = \rho^{[1]}_{\sigma,f}(\overline{\omega})\sum_{k=0}^{\infty}\binom{|\overline{\omega}|+k}{k}\sigma^k.$$

Following this, we note that for any integers n, m, one has $\binom{n+m-1}{m} = (-1)^m\binom{-n}{m}$. Take $n = |\omega|+1$ and $m = k$ and compare with a binomial series to complete the proof. □

We are now in a position to state the main theorem relating block-spin renormalization of Dyson models, block renormalization of random walk models, and block renormalization of random walks on $\mathscr{G}_2$.

Theorem 3.3.19. *Let (H,ν) be a Dyson hierarchical model with $\nu\sim N(0,\sigma)$ with $0<\sigma<1$ and with interaction function f with a well-defined thermodynamic limit. With $a = -\log_2(f(2))$,*

$$\tilde{\sigma}\equiv r_a(\sigma) = 2^{1-a}\frac{\sigma}{1-\sigma}, \tag{3.3.2}$$

and μ_f as in Proposition 3.3.3, the following diagram commutes.

$$\begin{array}{ccc}
(f,\nu) & \xmapsto{\mathscr{R}} & (\tilde{f},\tilde{\nu}) \\
\downarrow & & \downarrow \\
\rho^{[1]}_{\sigma,f} & \xmapsto{\hat{\mathscr{R}}_{b,a}} & \rho^{[1]}_{\tilde{\sigma},\tilde{f}} \\
\downarrow & & \downarrow \\
\mu_f & \xmapsto{\mathscr{R}_b} & \mu_{\tilde{f}}
\end{array}$$

Proof. We can calculate directly using the definition of the full block-spin renormalization operator 3.1.17 that

$$\tilde{\nu} \sim N\left(0, 2f(2)\frac{\sigma}{1-\sigma}\right).$$

Combining Lemmas 3.3.13 and 3.3.18 implies that for any $a > 0$ and $\widetilde{\omega} \in \Omega_{V_{N-1}}$

$$\hat{\mathscr{R}}_{b,a}\left(\rho^{[1]}_{\sigma,f}\right)(\widetilde{\omega}) = 2^{-a}\left(2\frac{\sigma}{1-\sigma}\right)^{|\widetilde{\omega}|+1}\prod_{(x,y)\in S(\widetilde{\omega})} f(d(x,y)).$$

For any $(\tilde{x},\tilde{y}) \in \widetilde{\omega}$, $2d(\tilde{x},\tilde{y}) = d(x,y)$. The renormalized interaction function $\tilde{f}$ is defined so that $\tilde{f}(2^n) = \frac{f(2^{n+1})}{f(2)}$, so we rewrite

$$\hat{\mathscr{R}}_{b,a}\left(\rho^{[1]}_{\sigma,f}\right)(\widetilde{\omega}) = \frac{2^{-a}}{f(2)}\left(2f(2)\frac{\sigma}{1-\sigma}\right)^{|\widetilde{\omega}|+1}\prod_{(\tilde{x},\tilde{y})\in\widetilde{\omega}} \tilde{f}(d(\tilde{x},\tilde{y})).$$

Thus, if we take $2^{-a} = f(2)$ and $\tilde{\sigma} = 2^{1-a}\frac{\sigma}{1-\sigma}$, then

$$\hat{\mathscr{R}}_{b,a}\left(\rho^{[1]}_{\sigma,f}\right) = \rho^{[1]}_{\tilde{\sigma},\tilde{f}}.$$

Finally, taking μ_f with level distribution $\{p_k\}_{k\in\mathbb{N}}$ as in Proposition 3.3.3, we may see directly from Proposition 3.2.7 that $\mathscr{R}_b(\mu_f)$ has level distribution $\{\tilde{p}_k\}_{k\in\mathbb{N}}$ satisfying

$$\tilde{p}_k = \frac{2^{k+1}f(2^{k+1})}{\sum_{n=0}^{\infty} 2^n f(2^n) - 1}.$$

Since $f(1) = 1$, this may be rewritten after some reindexing as

$$\tilde{p}_k = \frac{2^k \tilde{f}(2^k)}{\sum_{n=0}^{\infty} 2^n \tilde{f}(2^n)}.$$

□

The most relevant corollary occurs in the case where μ_f is block self-similar.

Corollary 3.3.20. *Let (H,ν) be a Dyson hierarchical model with $\nu \sim N(0,\sigma)$ with $0 < \sigma < 1$ and with interaction function $f(n) = n^{-a}$ for $a > 1$. The following diagram commutes with $\tilde{\sigma} = r_a(\sigma)$ as in (3.3.2).*

$$\begin{array}{ccc} \nu & \xmapsto{\mathscr{R}_a} & \tilde{\nu} \\ \downarrow & & \downarrow \\ \rho^{[1]}_{\sigma,a} & \xmapsto{\hat{\mathscr{R}}_{b,a}} & \rho^{[1]}_{\tilde{\sigma},a} \end{array}$$

3.3.6.2 Green's Function Self-Similarity at the Critical Point

The block-spin renormalization argument allows for simple derivation of self-similar scaling behavior of the Green's function for Gaussian Dyson models with power law interactions $f(n) = n^{-a}$. We will consider $1 < a < 2$ fixed in this section. The key strategy for the following lemma will be to relate the kernels on V_N to those on V_{N+1} by way of the block-spin renormalization summation.

Lemma 3.3.21. *For $\tilde{x}, \tilde{y} \in V_N$ with immediate descendants $x_1, x_2, y_1, y_2 \in V_{N+1}$, respectively, then*

$$K^N_{r_a(\sigma)}(\tilde{x}, \tilde{y}) = 2^{-a}\left[K^{N+1}_\sigma(x_1, y_1) + K^{N+1}_\sigma(x_2, y_1) + K^{N+1}_\sigma(x_1, y_2) + K^{N+1}_\sigma(x_2, y_2)\right].$$

Proof. By definition

$$K^N_{r_a(\sigma)}(\tilde{x}, \tilde{y}) = \sum_{\widetilde{\omega}:\tilde{x}\to\tilde{y}} \rho^{[1]}_{r_a(\sigma),a}(\widetilde{\omega}).$$

The block-spin renormalization argument allows each $\rho^{[1]}_{r_a(\sigma),a}(\widetilde{\omega})$ to be written as a sum over the weights of the entire equivalence class of random walks on V_{N+1} (with Gaussian a priori measure having variance σ) which are sent by a block transformation to $\tilde{\omega}$. Thus

$$K^N_{r_a(\sigma)}(\tilde{x}, \tilde{y}) = \sum_{\tilde{\omega}:\tilde{x}\to\tilde{y}} 2^{-a} \sum_{\omega\in[S(\widetilde{\omega})]_{\sim_b}} \rho^{[1]}_{\sigma,a}(\omega).$$

Taken together, these two sums add the weights of all walks which start in the size 2 block in V_{N+1} corresponding to $\tilde{x}$ and end in the size 2 block in V_{N+1} corresponding to $\tilde{y}$. This allows for the sum to be split into four parts. Then

$$\begin{aligned} K^N_{r_a(\sigma)}(\tilde{x}, \tilde{y}) &= 2^{-a}\left[\sum_{\omega:x_1\to y_1} \rho^{[1]}_{\sigma,a}(\omega) + \sum_{\omega:x_2\to y_1} \rho^{[1]}_{\sigma,a}(\omega) + \sum_{\omega:x_1\to y_2} \rho^{[1]}_{\sigma,a}(\omega) + \sum_{\omega:x_2\to y_2} \rho^{[1]}_{\sigma,a}(\omega)\right], \\ &= 2^{-a}\left[K^{N+1}_\sigma(x_1, y_1) + K^{N+1}_\sigma(x_2, y_1) + K^{N+1}_\sigma(x_1, y_2) + K^{N+1}_\sigma(x_2, y_2)\right]. \end{aligned}$$

□

This lemma has an immediate corollary which relies on the observation that the symmetries of a Dyson model imply that the Green's function of two points depends only on the hierarchical distance between them.

Corollary 3.3.22. *If $\tilde{x} \neq \tilde{y}$, then in the notation of the previous lemma we have*

$$K^N_{r_a(\sigma)}(\tilde{x}, \tilde{y}) = 2^{2-a} K^{N+1}_\sigma(x_1, y_1).$$

On the other hand, if $\tilde{x} = \tilde{y}$, then

$$K^N_{r_a(\sigma)}(\tilde{x}, \tilde{x}) = 2^{1-a} \left[K^{N+1}_\sigma(x_1, x_1) + K^{N+1}_\sigma(x_1, x_2) \right].$$

We can iteratively apply this corollary using the definition of $\ell(x, y)$ for $x, y \in \mathbb{N}$ given in Definition 3.1.14 to obtain a general identity relating the Green's function between two distinct points to the Green's function between a point and itself. To simplify notation here, we'll write

$$K^N_\sigma(j) \equiv K^N_\sigma(x, y)$$

for any $x, y \in \mathbb{N}$ with $\ell(x, y) = j$, and likewise we'll write $K_\sigma(j)$ for the Green's function on $\mathbb{N}$.

Proposition 3.3.23. *The Green's function $K_\sigma(j)$ for $j \geq 1$ satisfies*

$$K_\sigma(j) = (2^{2-a})^{1-j} \left[2^{a-1} K_{r_a^j(\sigma)}(0) - K_{r_a^{j-1}(\sigma)}(0) \right],$$

Proof. Fix $x, y \in \mathbb{N}$. First, take finite $N > 0$ sufficiently large that $x, y \in V_N$. Iteratively applying Corollary 3.3.22 yields

$$K^N_\sigma(x, y) = (2^{2-a})^{1-\ell(x,y)} \left[2^{a-1} K^{N-\ell(x,y)}_{r_a^{\ell(x,y)}(\sigma)}(\tilde{z}, \tilde{z}) - K^{N-\ell(x,y)+1}_{r_a^{\ell(x,y)-1}(\sigma)}(z, z) \right],$$

where $\tilde{z} \in V_{N-\ell(x,y)}$ is the earliest common ancestor of x and y, and $z \in V_{N-\ell(x,y)+1}$ is any immediate descendant of $\tilde{z}$. Then, taking the limit as $N \to \infty$ yields the result. □

This has the clearest implications at fixed points of r_a. The only such nonzero point occurs at the critical point $\sigma_* = 1 - 2^{1-a}$, which gives the following scaling law for the Green's function.

Corollary 3.3.24. *At $\sigma_* = 1 - 2^{1-a}$, we have for any $j \geq 1$*

$$K_{\sigma_*}(j) = \left(2^{2-a}\right)^{1-j} \left(2^{a-1} - 1\right) K_{\sigma_*}(0).$$

In particular, for $x \neq y \in \mathbb{N}$,

$$K_{\sigma_*}(x, y) \propto d(x, y)^{a-2}.$$

We can understand the behavior of the Green's function for $\sigma < \sigma_*$ using the dynamical properties of r_a.

Lemma 3.3.25. *The critical point σ_* is an unstable fixed point of r_a, with*

$$r_a'(\sigma_*) = 2^{a-1}.$$

For all $0 < \sigma < \sigma_*$, $r_a(\sigma) < \sigma$. The only other fixed point of r_a is a stable fixed point at $\sigma = 0$, where

$$r_a'(0) = 2^{1-a}.$$

This allows us to analyze the asymptotic behavior of the Green's function as $\sigma \nearrow \sigma_*$ both for sufficiently small distances and for sufficiently long distances.

Proposition 3.3.26. *Let $s(\sigma) = \frac{\sigma_* - \sigma}{\sigma_*}$ and define*

$$k_{s(\sigma)}(j) = K_\sigma(j).$$

For sufficiently small s (depending on j) and for all $1 \leq i \leq j$, to leading order about $s = 0$, there exists constants $C_1(a), C_{\frac{3}{2}}, C_2(a) > 0$ such that

$$k_s(i) - k_0(i) \sim \begin{cases} C_1(a)(2^{2a-3})^i s & 1 < a < \frac{3}{2} \\ -C_{\frac{3}{2}} \left[1 + i + 2\log_2(s)\right] s & a = \frac{3}{2} \\ C_2(a) s^{\frac{1}{a-1}-1} & \frac{3}{2} < a < 2. \end{cases}$$

For sufficiently large j (depending on s) and for all $i \geq j$,

$$k_s(i) \sim C(a, s)(2^{-a})^i,$$

for some constant C depending only on a and the choice of s.

Proof. Using the preceding lemma, for sufficiently small $\epsilon > 0$, if $s(\sigma) < \epsilon$, then for any $j \in \mathbb{N}$

$$s\left(r_a^{-j}(\sigma)\right) \sim (2^{1-a})^j s(\sigma).$$

We note the continuity of $K_\sigma(j)$ on $[0, \sigma_*]$. Applying Proposition 3.3.23 gives for $s(\sigma) \sim (2^{1-a})^j \epsilon$ that for any $1 \leq i \leq j$

$$k_s(i) \sim (2^{2-a})^{1-i} \left[2^{a-1} k_{(2^{a-1})^i s}(0) - k_{(2^{a-1})^{i-1} s}(0)\right]$$

Using Proposition 3.3.12 and Corollary 3.3.24, we thus have the desired result for all $1 \leq i \leq j$ and some constants $C_1(a), C_{\frac{3}{2}}, C_2(a)$ (redefined from those of Proposition 3.3.12).

On the other hand, by the preceding lemma for sufficiently small $\epsilon_0 > 0$, if $\sigma < \epsilon_0$ then

$$r_a^k(\sigma) = (2^{1-a})^k(\sigma + \sigma^2) + \mathcal{O}(\sigma^3) \tag{3.3.3}$$

for all $k > 0$. As a result, for any $s(\sigma) > 0$ we need only take j sufficiently large that $r_a^j(\sigma) = \epsilon < \epsilon_0$, which is achieved for $j > C(\epsilon) - \frac{1}{a-1} \log_2(s)$ for some $C(\epsilon) > 0$. As a result, for $i \geq j$

$$k_s(i) = (2^{2-a})^{1-i} \left[2^{a-1} K_{r_a^{i-j}(\epsilon)}(0) - K_{r_a^{i-j-1}(\epsilon)}(0)\right].$$

Applying Proposition 3.3.10 and expanding one application of r_a using (3.3.3) reduces this expression to

$$k_s(i) = (2^{2-a})^{1-i} \left(r_a^{i-j-1}(\epsilon)\right)^2 + \mathcal{O}(\epsilon^3).$$

Thus, for some (redefined) constant $C(a, s)$ which incorporates the dependence on j and ϵ

$$k_s(i) \sim C(a, s)(2^{-a})^i.$$

□

The case where $1 < a < \frac{3}{2}$ is of the most interest, as discussed in Theorem 3.1.5, as it is in this case that the renormalization operator at the critical Gaussian distribution has only a single unstable direction. In this case, the previous proposition indicates the existence of two asymptotic behaviors of the Green's function, depending on whether the distance is

taken to be relatively small or relatively large. In either case, one sees power law decay of the Green's function in the hierarchical distance to leading order. This power law behavior indicates self-similarity, and implies the following corollary.

Corollary 3.3.27. *For $1 < a < \frac{3}{2}$, the correlation length of the Dyson hierarchical model near the Gaussian critical point is infinite.*

Chapter 4

Herman Rings for Certain Highly Dissipative Two-Dimensional Complex Analytic Maps

In this chapter we explore recent developments in the theories of analytic circle diffeomorphisms and critical circle maps. We will explore the framework of almost-commuting pairs developed in [GY20a; GY20b] and its extension from a one-dimensional setting to a highly-dissipative two-dimensional setting. Our main result is a schematic proof demonstrating how these ideas can be used to construct golden mean Herman rings up to their boundary curves for two-dimensional perturbations of one-dimensional maps having a golden mean Herman ring.

4.0.1 Notation

- For $U \subset X$ and a map $f : U \to X$ we'll write $f^n = f \circ f \circ \ldots \circ f$ as the n-fold iterate of f, wherever it is well-defined.

- S^1 is the unit circle, which we will sometimes identify as a subset of $\mathbb{C}$, at other times with the real line in $\mathbb{C}/\mathbb{Z}$, and at yet other times with $[0, 1) \equiv \mathbb{R}/\mathbb{Z}$. For $x, y, z \in S^1$,

we'll write $x < y < z$ to mean that these points appear in this order in the given orientation of S^1.

- $\mathrm{Diff}_+(S^1)$ are the orientation-preserving diffeomorphisms of the circle.
- $\mathrm{Crit}_+(S^1)$ are the orientation-preserving unicritical diffeomorphisms of the circle. For $f \in \mathrm{Crit}_+(S^1)$, we will refer to the *critical point* of the map as the point $c \in S^1$ with $f'(c) = 0$ and to the *critical value* $f^{-1}(c)$.
- $\mathbb{A}_r \subset \mathbb{C}/\mathbb{Z}$ denotes the annulus consisting of those points $x + iy$ with $|y| < r$.

4.1 Introduction to the Dynamics of Circle Maps

Circle maps are a classical subject in one-dimensional dynamics. This section will provide a brief overview of the historical development of the field and will introduce the basic combinatorial structures involved.

4.1.1 Some General Theory of Circle Maps

Our basic discussion in this section is based upon Chapter 1 of [DV12], in a which a more thorough history is presented. The most basic maps of the circle are rigid rotations, which play a central role in the theory.

Definition 4.1.1. A *pure rotations* is a map $T_\rho : x \mapsto x + \rho$ for $x, \rho \in S^1$. We'll write $\hat{T}_y : \mathbb{R} \to \mathbb{R}$ for the affine translation map on $\mathbb{R}$.

The dynamics of pure rotations can be described simply using the Archimedean principle. In particular, the orbit of every point under a pure rotation is dense if and only if the orbit of any point under that rotation is dense if and only if ρ is irrational.

In general, other circle homeomorphisms may be more complicated. Nonetheless, if such a homeomorphism can be found to be topologically conjugate to a pure rotation, then some

of its dynamical behavior may immediately be understood. The rotation number provides combinatorial data which suggests which for which ρ such a conjugacy with a pure rotation may be possible.

Definition 4.1.2. Let $f : S^1 \to S^1$ be a homeomorphism with periodic lift $\hat{f} : \mathbb{R} \to \mathbb{R}$. The *rotation number* of f is $\rho_f \in S^1$ and is given by the limit

$$\lim_{n\to\infty} \frac{\hat{f}^n(x) - x}{n} \mod 1,$$

ultimately depending neither on the choice of $x \in \mathbb{R}$ nor on the lift.

The rotation number is invariant under topological conjugacy and the rotation number of T_ρ is ρ. A classic result of Poincaré then relates the dynamics of T_ρ to that of an orientation preserving homeomorphism $f : S^1 \to S^1$.

Theorem 4.1.1 ([Poi85])**.** *If $f : S^1 \to S^1$ is an orientation-preserving homeomorphism with rotation number ρ, then there exists an orientation-preserving continuous map $h : S^1 \to S^1$ such that*

$$h \circ f = T_\rho \circ h.$$

If f has a dense orbit, then h can be chosen to be a homeomorphism.

In what cases and to what extent to regularity of h can be upgraded beyond a homeomorphism is a central question of the theory of circle maps. Answers to this question tend to rely on three factors - the regularity of f, the presence of critical points of f, and the number-theoretic properties of the rotation number ρ. In this chapter, we will be concerned with holomorphic maps f and will distinguish the case when f is a diffeomorphism with that of f with a single critical point. The number-theoretic properties of the rotation number are best described using continued fractions.

Definition 4.1.3. For $\rho \in \mathbb{R}/\mathbb{Z}$, the *continued fraction expansion* of ρ is the natural number sequence $\{a_i\}_{i\geq 0}$ with

$$\rho = \cfrac{1}{a_0 + \cfrac{1}{a_1 + \dots}}.$$

We write $\rho \equiv [a_0, a_1, \dots]$. This expansion is infinite if and only if $\rho \in \mathbb{Q}$. The *n-th partial convergent* to ρ is the reduced rational number

$$\frac{p_n}{q_n} = [a_0, \dots, a_n],$$

and we call q_n the *n-th closest return time*. One has $q_0 = 1$, $q_1 = a_0$, and $q_{n+1} = a_n q_n + q_{n-1}$ for all subsequent n.

Continued fraction expansions are intricately related to the dynamical behavior of the Gauss map.

Definition 4.1.4. The *Gauss map* $G : \mathbb{R}/\mathbb{Z} \to \mathbb{R}/\mathbb{Z}$ is the map given by

$$G(x) = \frac{1}{x} - \left\lfloor \frac{1}{x} \right\rfloor.$$

Lemma 4.1.2. *One has* $G([a_0, a_1, \dots]) = [a_1, \dots]$ *and*

$$a_i = \left\lfloor \frac{1}{G^i(x)} \right\rfloor.$$

The simplest cases which we will primarily concern ourselves with occur when all entries of the continued fraction expansion remain bounded.

Definition 4.1.5. We say an irrational $\rho \in \mathbb{R}/\mathbb{Z}$ is of *bounded type* if for some $n \in \mathbb{N}$ and for all i one has $a_i < n$. Similarly, ρ is *periodic* if $\{a_i\}$ forms a periodic sequence and *eventually periodic* if there exists $N \in \mathbb{N}$ such that $G^N(\rho)$ is periodic. If $a_i = N$ for all i, then ρ is said to be *golden*, and likewise if $a_i = N$ for all sufficiently large i then ρ is *eventually golden*.

It is a classical result that bounded-type rotation numbers, as well as periodic rotation numbers, are dense in $\mathbb{R}/\mathbb{Z}$.

4.1.2 The Dynamics of Circle Diffeomorphisms

Circle diffeomorphisms are, in some sense, a far simpler case than critical circle maps. As a result, global proofs about the behavior of any circle diffeomorphism with appropriate regularity and rotation number are available. This will not hold true in the case of critical circle maps.

One of the earliest results in the study of circle diffeomorphisms comes from Denjoy.

Theorem 4.1.3 ([Den32]). *If $f : S^1 \to S^1$ is a C^2 orientation preserving diffeomorphism with irrational rotation number ρ, then f is topologically conjugate to the pure rotation T_ρ.*

This result cannot in general be strengthened to weaken the regularity assumption to C^1.

The primary issue which prevents one from strengthening the conjugacy from a homeomorphism to a differentiable (or smooth, or analytic) map is the issue of 'small denominators.' Namely, maps with rational rotation numbers will feature periodic points. Maps with rotation numbers which are irrational, but which are 'too close' to a rational number will instead feature points whose orbits pass extremely close to the initial point. Sufficiently many close returns result in non-differentiability of the conjugacy. This connection was made precise by Arnol'd.

Theorem 4.1.4 ([Arn61]). *There exists an $\epsilon(\rho) > 0$ such that if $f : S^1 \to S^1$ is analytic and orientation-preserving with rotation number ρ and ρ satisfies a Diophantine condition*

$$\left|\rho - \frac{p}{q}\right| > \frac{K}{q^{2+\alpha}},$$

for some constants $K, \alpha > 0$ and any $m, n \in \mathbb{N}$, and $\|f - T_\rho\|_{C(S^1)} < \epsilon$, then f is analytically conjugate to T_ρ.

If all of these assumptions hold, but the Diophantine condition failed to hold, Arnol'd was able to demonstrate that the conjugacy could fail to be smooth. Similar methods would lead to the KAM theory in symplectic dynamics. Later work by Herman in [Her79] provided a similar result in the differentiable or smooth settings.

The use of renormalization methods in this problem was pioneered by Yoccoz in [Yoc02]. Here he provided an optimal Brjuno condition on the rotation number for which an analytic conjugacy result could be obtained. He was also able to provide a global theorem on analytic conjugation of circle diffeomorphisms which does not rely on the closeness of f to a pure rotation, but with a slightly stronger 'Herman condition' on the rotation number. Both of these conditions are weaker than the assumption of bounded type.

A distinctive renormalization strategy was developed by Risler in [RF99]. This technique considered analytic maps of a neighborhood of $\mathbb{R}/\mathbb{Z}$ in $\mathbb{C}/\mathbb{Z}$. Such maps are analytic extensions of circle maps when they satisfy a real-symmetry condition. Without this real-symmetry condition, but with an assumption of closeness to a pure rotation, Risler demonstrated the existence of maps which could be analytically conjugated to a pure translation (rotation) within fairly generic families of analytic maps. This approach was connected to that of Yoccoz by Goncharuk and Yampolsky in [GY20b].

4.1.3 The Dynamics of Unicritical Circle Maps

Unicritical maps of S^1 have played a fundamental role in the development of renormalization theory in one-dimensional dynamical systems. These systems possess a second piece of combinatorial datum: the criticality of the critical point. That is, the exponent α such that near the critical point $c \in S^1$ of some unicritical $f : S^1 \to S^1$, one has

$$|f(x) - f(c)| \sim (x - c)^{\alpha}.$$

Holomorphic methods are most suited for cases in which α is an odd integer. In this critical setting, we cannot expect conjugacy by an analytic diffeomorphism with a pure rotation, as such conjugacies would necessarily preserve the critical point. As such, understanding the long-time behavior of critical circle maps is fundamentally more complicated than in the diffeomorphism case.

Renormalization has been a fundamental tool in achieving such understanding. For

bounded type rotation numbers, the commuting pair approach of de Faria in [Far92; De 99] (which will be discussed in detail in the following section) enabled the construction of an attracting renormalization horseshoe, and subsequent work by de Faria and de Melo in [DD00] showed exponential convergence to this horseshoe. A chief difficulty with progressing the theory further to achieve a Wilsonian-like renormalization picture was to show that this horseshoe attractor was a hyperbolic set. This difficulty was eventually addressed by Yampolsky in [Yam03] with the 'cylinder renormalization' method. Extremely recent work by Gaidashev and Yampolsky in [GY20a] has connected this cylinder renormalization with the older commuting pair formalism.

In the simplest case, when the rotation number has a periodic continued fraction expansion, the renormalization operator has a corresponding periodic point. Unlike the diffeomorphism case, this periodic point is not a pure rotation, but still governs the long-time dynamics in a universal manner.

4.2 The Renormalization Theory for Almost-Commuting Pairs

In this section, we'll provide a brief introduction to the framework of commuting pairs, followed by that of almost-commuting pairs. We will see how these are connected to the dynamics of circle maps, both diffeomorphism and unicritical. We will systematically use the same notation to refer both to the diffeomorphism renormalization operator and the critical renormalization operator in an attempt to highlight the similarity of these two cases. That being said, a conclusive formal picture connecting the two as cases of a single globally-defined operator on an appropriate space has not yet been carried out. We will conclude by discussing how these ideas are transferred to the highly-dissipative two-dimensional setting.

4.2.1 The Commuting Pair Formalism

In this section, we willll provide an overview to the commuting pair formalism. The discussion and notation here related to circle diffeomorphisms is based on [Yam21], and that related to critical circle maps is based on [GY20a; Yan20].

Commuting pairs are a classical setting for the renormalization theory of maps of the circle, with holomorphic commuting pairs first being introduced by de Faria in [De 99]. In what follows, we will let f be either an analytic circle diffeomorphism or an analytic unicritical circle map with critical point at 0. We'll refer to 0 as the distinguished point, for convenience. Its orbit we'll call the *distinguished orbit*

$$\mathcal{O}_f = \left\{0, f(0), f^2(0), \ldots\right\}.$$

We'll assume further that the rotation number ρ_f is irrational, and moreover is of bounded type. Then for any $n \geq 0$ one has $f^{q_{2n+1}}(0) < 0 < f^{q_{2n}}(0)$, let the *$n$-th distinguished interval* $\hat{E}_n = E_{2n} \cup E_{2n+1}$ denote the arc joining these two iterates containing the distinguished point, with E_k having endpoints at 0 and $f^{q_k}(0)$. Upon exiting $\hat{E}_n$, the dynamics of f returns only after a long period of time.

Lemma 4.2.1. *We may precisely describe the behavior of the first return to $\hat{E}_n$ under the dynamics of f.*

- *For $x \in E_{2n}$, the first return of f to $\hat{E}_n$ occurs at $f^{q_{2n+1}}(x) \in \hat{E}_n$.*
- *For $x \in E_{2n+1}$, the first return of f to $\hat{E}_n$ occurs at $f^{q_{2n}}(x) \in E_{2n}$.*

This motivates the definition of a pre-renormalization operation.

Definition 4.2.1. The n-th *pre-renormalization* (at 0) of f is $p\mathscr{R}^n f$ which acts as the first return map to $\hat{E}_n$, so that

$$f \mapsto \left(f^{q_{2n+1}}|_{E_{2n}}, f^{q_{2n}}|_{E_{2n+1}}\right).$$

This operation naturally takes a single function on the circle to a pair of interval maps intersecting at 0. We may take $n = 0$ to represent f itself in this form as the pair

$$(f^{a_0}|_{E_0}, f|_{E_1}).$$

If f is a diffeomorphism, we may alternatively take the lift $\tilde{f} : \mathbb{R} \to \mathbb{R}$ such that $0 < \tilde{f}(0) < 1$ and equate f with the pair

$$\left(\tilde{f}|_{[-1,0]}, \hat{T}_{-1}|_{[0,\tilde{f}(0)]}\right).$$

In general, two such maps do not agree at 0, but they do commute. Indeed, for f analytic these two maps may both be extended to a small neighborhood of 0 in $\mathbb{C}/\mathbb{Z}$ where the extensions will commute. The multiple ways to depict f in this form already highlights a key aspect of the conversion between maps of the circle and commuting pairs of interval maps - nonuniqueness. Of course, to proceed to further analysis of commuting pairs we must first carefully define them and discuss a renormalization procedure for them.

Definition 4.2.2. An *analytic commuting pair* (at 0) on intervals $I_\eta, I_\xi \subset \mathbb{R}$ are a pair of analytic orientation preserving interval homeomorphisms $\eta : I_\eta \to I_\eta \cup I_\xi$ and $\xi : I_\xi \to I_\eta$ which satisfy the following conditions.

1. $I_\eta = [0, \xi(0)]$ and $I_\xi = [\eta(0), 0]$. We will commonly require $\xi(0) > 0$ and $\eta(0) < 0$, but this requirement is not strictly necessary (one may also conjugate by the map $x \mapsto -x$ to make it so).

2. Both η and ξ extend to homeomorphisms of neighborhoods in $\mathbb{R}$ of their domains of definition, such that whenever both are defined $\xi \circ \eta = \eta \circ \xi$.

3. If $x \neq 0$, x is a critical point for neither η nor ξ.

Further, we'll say that the commuting pair is *critical* or *cubic critical* if 0 is a critical point for both η and ξ and moreover $\eta''(0) = \xi''(0) = 0$, but $\eta'''(0) \neq 0 \neq \xi'''(0)$. On the other hand, we'll say the commuting pair is *regular* if 0 is not a critical point for either η or ξ.

For convenience, we will frequently suppress the intervals when referring to pairs and call $\zeta = (\eta, \xi)$ a commuting pair. It's worth noting that given a commuting pair we may reverse course to obtain circle maps. To do so, consider the interval $I = \eta(I_\eta)$ and define f_ζ on I by

$$f_\zeta(x) = \begin{cases} \eta \circ \xi(x) & x \in [\eta(0), 0] \\ \eta(x) & x \in [0, \eta \circ \xi(0)]. \end{cases}$$

By gluing $\eta(0)$ and $\eta \circ \xi(0)$, we obtain a topological circle S. Commutativity guarantees regularity of f_ζ at the glued point, giving a well defined projection $\hat{f}_\zeta$. For any sufficiently regular diffeomorphism $\phi : S \to S^1$ we thus obtain a circle map

$$\phi \circ \hat{f}_\zeta \circ \phi^{-1}.$$

Thus, any commuting pair gives rise to a conjugacy class of circle maps.

The combinatorial structure of maps of the circle is well-described by the continued fraction expansion of the rotation number. One would like to develop an equivalent structure for commuting pairs.

Definition 4.2.3. The *height* $\chi(\zeta)$ of a commuting pair $\zeta = (\eta, \xi)$ is the minimal r such that

$$0 \in \eta^r \left([\xi(0), \eta \circ \xi(0)] \right).$$

If no such r exists, then $\chi(\zeta) = \infty$.

A simple compactness argument informs us that I_η contains a fixed point of η if and only if $\chi(\zeta) = \infty$.

Definition 4.2.4. We say that the pair ζ is *renormalizable* if $\chi(\zeta) < \infty$.

The key insight in defining renormalization is in replicating the return map structure for circle maps described earlier by considering returns of a pair $\zeta = (\eta, \xi)$ to $I_\xi \cup [0, \eta^r(\xi(0))]$.

Definition 4.2.5. The *prerenormalization* of a renormalizable commuting pair $\zeta = (\eta, \xi)$ consists of the pair

$$p\mathscr{R}\zeta = \left(\eta^r \circ \xi|_{I_\xi}, \eta|_{[0, \eta^r(\xi(0))]} \right).$$

It's worth noting two features of this prerenormalization operation. First, it necessarily shrinks the domain of definition of the pair. Second, it reverses the orientation of the pair; the location of the intervals of definition of the pair relative to 0 have switched. The latter feature can be addressed either by conjugating the pair by the map $x \mapsto -x$ or simply by considering even iterates of prerenormalization - indeed the prerenormalization defined above for circle maps is equivalent to the second iterate of this commuting pair prerenormalization. As to the former, we compensate by adding a normalization condition to the definition of commuting pairs.

Definition 4.2.6. A commuting pair $\zeta = (\eta, \xi)$ is said to be *normalized* if $|\xi(0)| = 1$.

Given a commuting pair $\zeta = (\eta, \xi)$, we may obtain a normalized commuting pair by conjugating both elements of the pair by the linear transformation $x \mapsto \frac{1}{|\xi(0)|}x$, which we will write as the pair $\tilde{\zeta} = (\tilde{\eta}, \tilde{\xi})$. We may now define a renormalization operator.

Definition 4.2.7. The *renormalization* (at 0) of a pair ζ is the pair

$$\mathscr{R}\zeta = \widetilde{p\mathscr{R}\zeta}.$$

The renormalization of a pair ζ is critical if ζ is, and is regular if ζ is.

We are now capable of defining a rotation number of a renormalizable commuting pair, under the understanding that we will interpret $\frac{1}{\infty}$ as 0.

Definition 4.2.8. The rotation number of a renormalizable commuting pair $\zeta = (\eta, \xi)$ is $\rho_\zeta = [a_0, a_1, \ldots]$ where

$$a_i = \chi(\mathscr{R}^i\zeta).$$

In particular this continued fraction expansion terminates and ρ_ζ is rational if and only if ζ is only finitely renormalizable.

This then ensures that by definition $\rho_{\mathscr{R}\zeta} = G(\rho_\zeta)$. It is worth noting that this rotation number for a commuting pair is consistent with the rotation number defined for circle maps.

Lemma 4.2.2. *If $f : S^1 \to S^1$ is obtained from a commuting pair ζ, then $\rho_f = \rho_\zeta$.*

4.2.2 Dynamical Partitions and Renormalization Iterates

In this section, we base our discussion upon the corresponding portion of [GY20a]. In order to make rigorous use of the fine-scale analysis into which renormalization grants insight, we will need to repeatedly partition the domain of a renormalizable pair $\zeta = (\eta, \xi)$ into finer and finer sets in a way which respects the renormalization dynamics. The basic idea will be to tile $I_\eta \cup I_\xi$ with forward iterates under ζ of the domain of definition of $p\mathscr{R}^n\zeta$.

We'll begin by characterizing which iterates of ζ are appropriate for such a partition.

Definition 4.2.9. Write the multi-index $\overline{s} = (c_1, b_1, c_2, b_2, \ldots, c_m, b_m)$ for $c_i \in \mathbb{N}$ whenever $i > 1$, $b_i \in \mathbb{N}$ whenever $i < m$, and $c_1, b_m \in \mathbb{N} \cup \{0\}$. For $\zeta = (\eta, \xi)$ a commuting pair write

$$\zeta^{\overline{s}} = \xi^{b_m} \circ \eta^{c_m} \circ \ldots \circ \eta^{c_2} \circ \xi^{b_1} \circ \eta^{c_1}.$$

Moreover, we'll take a partial order $\overline{s} \succ \overline{t}$ if $\overline{t} = (c_1, b_1, \ldots, c_k, b_k, d, e)$ with $k < m$ and either $d < c_{k+1}$ and $e = 0$, or $d = c_{k+1}$ and $e < b_{k+1}$.

The notational complexity thus introduced is made somewhat simpler by noting that the domains of definition of η and ξ as well as the commutativity make it so that all b_i may be taken to be either 0 or 1, and all c_i fall between 0 and $\chi(\zeta) + 1$.

Now, supposing that ζ is n-times renormalizable, we may recursively see that $p\mathscr{R}^n\zeta = (\eta_n, \xi_n)$, where η_n and ξ_n are some compositions of η and ξ.

Definition 4.2.10. For an n-times renormalizable $\zeta = (\eta, \xi)$, define multi-indices $\overline{s}_n$ and $\overline{t}_n$ by $p\mathscr{R}^n\zeta = (\zeta^{\overline{s}_n}, \zeta^{\overline{t}_n})$.

This puts us in a position to define the dynamical partition.

Definition 4.2.11. Let $\zeta = (\eta, \xi)$ be n-times renormalizable with $p\mathscr{R}^n\zeta = (\eta_n, \xi_n)$. The *$n$-th dynamical partition* is the collection

$$\mathscr{P}_n = \left\{\zeta^{\overline{r}}(I_{\eta_n}) \text{ for all } \overline{r} \prec \overline{s}_n\right\} \cup \left\{\zeta^{\overline{r}}(I_{\xi_n}) \text{ for all } \overline{r} \prec \overline{t}_n\right\}.$$

Elements of $\mathscr{P}_n$ necessarily have disjoint interiors and taken together cover $I_\eta \cup I_\xi$.

4.2.3 Almost-Commuting Pairs

Despite their friendly appearance, if we wish to analyze the behavior of the nonlinear operator $\mathscr{R}$ acting on the space of normalized (critical or regular) commuting pairs, we are immediately confronted with a technical issue. Commutation of two maps in a neighborhood of 0 imposes an infinitude of conditions on the two maps - namely that all derivatives of the commutator vanish. As a result, the space of commuting pairs defined on sets I, J does not inherit a natural Banach submanifold structure from $C^\omega(I) \times C^\omega(J)$. One way around this is to avoid the use of commuting pairs entirely, turning instead to Yampolsky's cylinder renormalization as in [Yam03]. Instead of this approach, in [GY20a], Gaidashev and Yampolsky sought to enlarge the domain of definition of $\mathscr{R}$ beyond the space of commuting pairs so as to obtain a finite-codimension submanifold. To do this, they adapted an idea originally from Mestel [Mes85] and introduced the larger space of almost-commuting pairs. In proceeding, we'll characterize our spaces of pairs of maps by complex domains, rather than simply subintervals of $\mathbb{R}$.

Definition 4.2.12. Let $\mathcal{A}(Z, W)$ denote the space of pairs $\zeta = (\eta, \xi)$ of analytic maps $\eta : Z \to \mathbb{C}$ and $\xi : W \to \mathbb{C}$, endowed with a Banach space structure by the norm

$$\|\zeta\|_{\mathcal{A}(Z,W)} = \frac{1}{2}\left(\|\eta\|_Z + \|\xi\|_W\right),$$

where $\|\cdot\|_U$ denotes the sup-norm on U.

Going forward we will take it is a tacit assumption that Z and W are topological disks, and that the distinguished point $0 \in Z \cap W$. All spaces and operations we'll henceforth discuss can be defined for a different choice of distinguished point, but then we may simply translate the domains Z and W and the functions η, ξ so as to send the distinguished to 0 and carry on unabated.

Definition 4.2.13. A pair $\zeta = (\eta, \xi) \in \mathcal{A}(Z, W)$ is *normalized* if $\xi(0) = 1$.

Definition 4.2.14. A pair $\zeta = (\eta, \xi) \in \mathcal{A}(Z, W)$ is a *critical pair* if 0 is the only critical point of both η and ξ in Z and W, respectively, and if it is a simple critical point. We'll write $\tilde{\mathcal{C}}(Z, W) \subset \mathcal{A}(Z, W)$ for the space of critical normalized pairs.

Definition 4.2.15. A pair $\zeta = (\eta, \xi) \in \mathcal{A}(Z, W)$ is a *regular pair* if η and ξ have no critical points in Z and W, respectively. We'll write $\tilde{\mathcal{D}}(Z, W) \subset \mathcal{A}(Z, W)$ for the space of regular normalized pairs.

With these spaces in mind, we are in a position to define almost-commutativity. This concept appears in a number of early computer-assisted works, including [Sti94; Bur97], though the precise formulation here is primarily from [GY20a].

Definition 4.2.16. A pair $\zeta = (\eta, \xi)$ is an *almost-commuting pair* to order $r > 0$ if $[\eta, \xi]$ is well-defined in a neighborhood of 0 and the Taylor series for $[\eta, \xi](x)$ at 0 vanishes up to and including the x^r term. We'll denote by $\mathcal{C}(Z, W)$ the space of critical normalized almost-commuting to order 2 pairs and by $\mathcal{D}(Z, W)$ the space of regular normalized almost-commuting to order 2 pairs. We'll write $\mathcal{C}^{\mathscr{R}}(Z, W) \subset \mathcal{C}(Z, W)$ and $\mathcal{D}^{\mathscr{R}}(Z, W) \subset \mathcal{D}(Z, W)$ to represent the subsets of these spaces consisting of renormalizable pairs.

The following proposition demonstrates that weakening commutativity to almost-commutativity does allow our spaces of pairs to inherit topological structure from $\mathcal{A}(Z, W)$.

Proposition 4.2.3 ([GY20a; Yam21])**.** *The spaces $\tilde{\mathcal{C}}(Z, W), \tilde{\mathcal{D}}(Z, W), \mathcal{C}(Z, W)$, and $\mathcal{D}(Z, W)$ are all immersed finite-codimension Banach submanifolds of $\mathcal{A}(Z, W)$.*

One may define heights and renormalization for almost-commuting pairs in exactly the same form as was done earlier for commuting pairs. The addition of renormalizability as a condition does not fundamentally change the Banach submanifold structure.

Proposition 4.2.4. *The spaces $\mathcal{C}^{\mathscr{R}}(Z, W)$ and $\mathcal{D}^{\mathscr{R}}(Z, W)$ are open subsets of $\mathcal{C}(Z, W)$ and $\mathcal{D}(Z, W)$, respectively.*

Now, supposing that ρ has a periodic continued fractions expansion. Two key results in this framework give the existence of a fixed point of an appropriate iterate of renormalization and a corresponding hyperbolicity of the renormalization operator. These theorems differ in details between the critical and regular cases but have fundamental similarities. In the regular case, renormalization fixed points are pairs corresponding to pure rotations.

Theorem 4.2.5 ([Yam21]). *There exists domains Z, W such that the following holds.*

- *There exists an open subset $\mathcal{U} \subset \mathcal{D}(Z, W)$ such that $\mathscr{R}$ is well-defined on $\mathcal{U}$, $\mathscr{R}(\mathcal{U}) \subset \mathcal{D}(Z, W)$ is compact analytic, and $D\mathscr{R}$ is a compact linear operator.*
- *Let ρ_* have a periodic continued fraction expansion and $\zeta_{T_{\rho_*}} \in \mathcal{D}(Z, W)$ be a commuting pair corresponding to the pure rotation T_{ρ_*} on S^1. There exists a minimal $k \in \mathbb{N}$ and a neighborhood $\mathcal{V} \subset \mathcal{U}$ such that $\zeta_{T_{\rho_*}}$ is the unique fixed point of $\mathscr{R}^k$ in $\mathcal{V}$.*
- *$D\mathscr{R}$ is hyperbolic at $\zeta_{T_{\rho_*}}$ with a one-dimensional unstable direction, and has a codimension one analytic strong stable manifold given by*

$$\mathcal{W}^s(\zeta_{T_{\rho_*}}) = \{\zeta \in \mathcal{D}(Z, W) : \rho_\zeta = \rho_*\}.$$

In the critical case, renormalization fixed points do not admit such a convenient representation, but are particular critical commuting pairs.

Theorem 4.2.6 ([GY20a]). *Let ζ_{ρ_*} be a periodic point of $\mathscr{R}$ of period k with rotation number ρ_*. There exists a minimal $p \in \mathbb{N}$ and a pair of domains Z, W such that the following holds.*

- *There exists a neighborhood $\mathcal{U} \subset \mathcal{C}(Z, W)$ on which $\mathscr{R}$ is well-defined such that ζ_{ρ_*} is the unique fixed point of $\mathscr{R}^p$ in $\mathcal{U}$. Moreover, there exists domains $Z' \ni Z$ and $W' \ni W$ such that $\mathscr{R}^p(\zeta_{\rho_*}) \in \mathcal{C}(Z', W')$.*
- *$D\mathscr{R}^p|_{\zeta_{\rho_*}}$ is a compact linear operator. The strong stable manifold of $\mathscr{R}^p$ at ζ_{ρ_*} is given by*

$$\mathcal{W}^s(\zeta_{\rho_*}) = \{\zeta \in \mathcal{C}(Z, W) : \rho_\zeta = \rho_*\}.$$

In principle, one would wish to connect the two preceding theorems together into a global picture in which a separatrix connects $\zeta_{T_{\rho_*}}$ to ζ_{ρ_*}. The precise properties of such a separatrix have not yet been fully explained.

4.2.4 Highly-Dissipative Two-Dimensional Almost-Commuting Pairs

The renormalization theory for almost-commuting pairs discussed previously is a fundamentally one-dimensional theory. Extending results from the one-dimensional case to a higher dimensional setting generally requires sophisticated perturbation arguments. Nonetheless, this renormalization formalism enjoys a key advantage over several other one-dimensional methods: it does not make use of the uniformization theorem. Due to this, in [GY20a] this method was used to extend renormalization results to the analysis of highly dissipative two-dimensional almost-commuting pairs.

To carry out this explanation, we'll discuss the spaces involved and define the manner in which pairs on subsets of $\mathbb{C}$ will be viewed as pairs on subsets of $\mathbb{C}^2$. In what follows, we'll imitate the notation present in [GY20a; Yam21]. The details of the construction can be found in these papers, so we will content ourselves here with defining terminology and the relevant spaces of pairs, as well as stating the main hyperbolicity results.

Definition 4.2.17. Set $\Omega = Z \times Z$ and $\Gamma = W \times W$ for $Z, W \subset \mathbb{C}$. Define $\mathcal{A}_2(\Omega, \Gamma)$ as the space of pairs of analytic maps $\Sigma = (A, B)$ with $A : \Omega \to \mathbb{C}^2$ and $B : \Gamma \to \mathbb{C}^2$. This space is endowed with a Banach space structure with the norm

$$\|(A, B)\|_{\mathcal{A}_2(\Omega,\Gamma)} = \frac{1}{2}\left(\|A\|_\Omega + \|B\|_\Gamma\right),$$

with $\|\cdot\|_U$ denoting the sup-norm on U.

With this in mind, we may embed one-dimensional pairs into spaces of two-dimensional pairs. Let $\pi_1, \pi_2 : \mathbb{C}^2 \to \mathbb{C}$ be projections onto the first and second factors.

Definition 4.2.18. Define the embedding $\iota : \mathcal{A}(Z,W) \to \mathcal{A}_2(\Omega,\Gamma)$ so that for any pair $\zeta = (\eta,\xi) \in \mathcal{A}(Z,W)$ one has

$$\iota : \zeta \mapsto \left(\begin{bmatrix} x \\ y \end{bmatrix} \mapsto \begin{bmatrix} \eta(x) \\ \eta(x) \end{bmatrix}, \begin{bmatrix} x \\ y \end{bmatrix} \mapsto \begin{bmatrix} \xi(x) \\ \xi(x) \end{bmatrix} \right).$$

It's worth noting that ι is an isometry onto its image, and that it has a left inverse. Let $\pi_1, \pi_2 : \mathbb{C}^2 \to \mathbb{C}$ be the coordinate projections.

Definition 4.2.19. Define the operator $\mathscr{L} : \mathcal{A}_2(\Omega,\Gamma) \to \mathcal{A}(Z,W)$ on $\Sigma = (A,B)$ by

$$\mathscr{L}(A,B)(x) = (\pi_1 \circ A(x,0), \pi_1 \circ B(x,0)).$$

This embedding is used to define a renormalization operator $\hat{\mathscr{R}} = \iota \circ \mathscr{R} \circ \iota^{-1}$ on $\iota(\mathcal{A}(Z,W))$. This operator differs slightly from the one described in the one-dimensional section. The combinatorial aspects of this operator are all virtually identical, but the rescaling operation may be slightly non-linear in this context, as the pre-renormalization is carried out at the critical value rather than the critical point, and then transferred to the critical point by the dynamics of the map. A clear discussion may be found in any of [GY20a; GRY21; Yam21].

Now, we'd like to define critical and regular pairs. In both cases, our definitions are perturbative in nature. In that interest, we'll define maps $a, b, h, g : \mathbb{C}^2 \to \mathbb{C}$ by

$$A(x,y) = \begin{bmatrix} a(x,y) \\ h(x,y) \end{bmatrix} = \begin{bmatrix} a_y(x) \\ h_y(x) \end{bmatrix},$$
$$B(x,y) = \begin{bmatrix} b(x,y) \\ g(x,y) \end{bmatrix} = \begin{bmatrix} b_y(x) \\ g_y(x) \end{bmatrix},$$

and will use these directly in our definitions.

The regular case corresponds in a fairly straightforward manner to the one-dimensional version, with an added parameter characterizing the extent of the perturbation into the second dimension.

Definition 4.2.20. A pair $\Sigma = (A, B) \in \mathcal{A}_2(\Omega, \Gamma)$ is $(\mathcal{U}, \delta)$-*regular* for $\mathcal{U} \subset \mathcal{D}(Z, W)$ and $\delta > 0$ if

- $d_{\mathcal{A}_2(\Omega,\Gamma)}(\Sigma, \iota(\mathcal{U})) < \delta$,
- when $y = 0$ both $\partial_x h$ and $\partial_x g$ are nonvanishing,
- we have

$$\sup_{(x,y)\in\Omega} [h(x, y) - h(x, 0)] < \delta,$$

  and likewise for g on Γ.

We'll denote the space of $(\mathcal{U}, \delta)$-regular pairs as $\mathcal{D}_2(\mathcal{U}, \delta)$ (with Ω and Γ implied by the choice of $\mathcal{U}$).

In the critical case there is an extra subtlety. In one-dimension we are able to simply identify the critical point of a pair upon inspection. The corresponding two-dimensional phenomena is a failure of dominated splitting, which is a more difficult property to detect directly. Nonetheless, we imitate the one-dimensional setting by simply analyzing the behavior of the pair on the subspace consisting of points $(x, 0) \in \mathbb{C}^2$.

Definition 4.2.21. A pair $\Sigma = (A, B) \in \mathcal{A}_2(\Omega, \Gamma)$ is $(\mathcal{U}, \epsilon, \delta)$-*critical* for $\mathcal{U} \subset \mathcal{C}(Z, W)$, $\epsilon > 0$, and $\delta > 0$ if

- $d_{\mathcal{A}_2(\Omega,\Gamma)}(\Sigma, \iota(\mathcal{U})) < \delta$,
- when $y = 0$ and $|x| > \epsilon$ both $\partial_x h$ and $\partial_x g$ are nonvanishing,
- we have

$$\sup_{(x,y)\in\Omega} [h(x, y) - h(x, 0)] < \delta,$$

  and likewise for g on Γ.

We'll denote the space of $(\mathcal{U}, \epsilon, \delta)$-critical pairs as $\mathcal{C}_2(\mathcal{U}, \epsilon, \delta)$. Moreover, we'll write

$$\mathcal{C}_2(\mathcal{U}, \delta) = \bigcap_{\epsilon>0} \mathcal{C}_2(\mathcal{U}, \epsilon, \delta).$$

It is notable that neither of the above conditions directly mentions almost-commutativity. Indeed, since these definitions are inherently perturbative, such perturbed pairs may not almost-commute in the same manner as in the one-dimensional setting. It is a key feature of this renormalization procedure that we regard a pair (A, B) as almost-commuting whenever $\mathscr{L}(A, B)$ are almost-commuting.

The main conclusions of this proceess are hyperbolicity results for $\hat{\mathscr{R}}$ for sufficiently highly-dissipative two-dimensional almost-commuting pairs in the critical and regular settings.

Theorem 4.2.7. *[Yam21] There exists domains $Z, W \subset \mathbb{C}$ and $\mathcal{U} \subset \mathcal{D}(Z, W)$ and $\delta > 0$ such that the following hold.*

- *There exists an open subset $\hat{\mathcal{U}} \in \mathcal{D}_2(\mathcal{U}, \delta)$ such that $\hat{\mathscr{R}}$ is a well-defined compact analytic operator on $\hat{\mathcal{U}}$, and $D\hat{\mathscr{R}}$ is a compact linear operator.*

- *For ρ_* having periodic continued fraction expansion, there exists a neighborhood $\hat{\mathcal{V}} \subset \hat{\mathcal{U}}$ and a minimal $k \in \mathbb{N}$ such that $\iota(\zeta_{T_{\rho_*}})$ is the unique fixed point of $\hat{\mathscr{R}}^k$ in $\hat{\mathcal{V}}$.*

- *$D\mathscr{R}$ is hyperbolic at $\iota(\zeta_{T_{\rho_*}})$ with a one-dimensional unstable direction and a codimension-one analytic strong stable manifold.*

Theorem 4.2.8. *[GRY21] Let ρ_* be the golden mean, with corresponding periodic point ζ_{ρ_*} of $\mathscr{R}$ with period k. There exists a minimal even $p \in \mathbb{N}$ and a pair of domains $Z, W \subset \mathbb{C}$ such that the following hold.*

- *The pair $\iota(\zeta_{\rho_*})$ is a fixed point of $\hat{\mathscr{R}}^p$ in $\mathcal{C}_2(\Omega, \Gamma)$.*

- *$D\hat{\mathscr{R}}^p|_{\iota(\zeta_{\rho_*})}$ is a compact linear operator whose spectrum coincides with $D\mathscr{R}^p|_{\zeta_{\rho_*}}$. For $\kappa \neq 0$, Z is an eigenvector of the one-dimensional operator with eigenvalue κ if and only if $\iota(Z)$ is an eigenvector of the two-dimensional operator with eigenvalue κ.*

4.3 On The Existence of and Boundaries of Herman Rings in Two Dimensions

In this section we'll discuss what is meant by a Herman ring for a highly-dissipative map of $\mathbb{C}^2$ and will present a schematic proof outlining how they may be constructed using renormalization theory. This proof will mainly be divided into three parts. The first will discuss sufficient transversality conditions required to obtain a 'triple renormalization intersection,' so that renormalizing a map at three different points results in attraction to particular renormalization fixed points. The second will note a family of one-dimensional maps which satisfies sufficient transversality conditions to satisfy the first part of the proof. The final part of the proof will outline the construction of a Herman ring for a highly-dissipative map of $\mathbb{C}^2$ with a triple renormalization intersection.

For clarity and brevity in this section, we will define the notation

$$\tau_w f(z) = f(z+w) - w.$$

Further, we will refer to a commuting pair of maps (A, B) as a 'pre-renormalization' of a map $F : \mathbb{C}^2 \to \mathbb{C}^2$ if

$$(A, B) = (F^{q_n}, F^{q_{n+1}}),$$

for some $n \in \mathbb{N}$. We will in general assume that pairs commute at 0, with τ_w operating so as to linearly translate z_* to 0 and produce a commuting pair with a distinguished point at 0. We may similarly define the notation $\tau_{(w_1,w_2)}$ for analytic maps of $\mathbb{C}^2$.

Finally, we will denote by $\mathscr{R}F$ a normalized commuting pair obtained from iterates of F, with a distinguished point at 0. Then, the renormalization with a distinguished point z_* of F is $\mathscr{R}\tau_{z_*}F$.

4.3.1 On Herman Rings

Dating back to Fatou, it is a classical fact in one-dimensional holomorphic dynamics that invariant Fatou components of rational maps are either immediate attracting basins, immediate parabolic basins, Siegel disks, or Herman rings (see [Mil11]).

Definition 4.3.1. A *(invariant) Herman ring* $\mathcal{H} \subset \mathbb{C}$ for a map $f : \mathbb{C} \to \mathbb{C}$ is a Fatou component invariant under f on which there is a conformal conjugacy $\phi : \mathcal{H} \to \mathbb{A}_r$ for some $r > 0$ conjugating f to a pure irrational rotation.

We can produce a similar definition for highly-dissipative maps of $\mathbb{C}^2$, now obtaining a submanifold rather than a domain.

Definition 4.3.2. A *(invariant) Herman ring* $\mathcal{H} \subset \mathbb{C}^2$ for a map $F : \mathbb{C}^2 \to \mathbb{C}^2$ is an injectively immersed surface such that $F|_{\mathcal{H}}$ is conformally conjugate to a pure irrational rotation on a one-dimensional annulus.

Such sets are considered for holomorphic endomorphisms of $\mathbb{CP}^2$ in Theorem 7 of [LP14].

4.3.2 Transversality Conditions for a Family

In carrying out our analysis, we will have need of maps which, when renormalized at different points, produce commuting pairs attracted to different renormalization fixed points. We'll state all definitions in terms of holomorphic maps on $\mathbb{C}$, but all definitions in this section may be extended without issue to holomorphic maps $f : U \to \mathbb{C}$ for $U \subset \mathbb{C}$, or highly dissipative two-dimensional holomorphic maps $F : U \to \mathbb{C}^2$ for $U \subset \mathbb{C}^2$.

Definition 4.3.3. A holomorphic map $f : \mathbb{C} \to \mathbb{C}$ has a *regular renormalization intersection* at a point $z \in \mathbb{C}$ for an eventually periodic $\alpha \in \mathbb{R}/\mathbb{Z}$, if there exists an $N, k > 0$ such that $\mathscr{R}^N \tau_z f \in \mathcal{W}^s_{\mathscr{R}^k}(\zeta_{T_\alpha})$.

Definition 4.3.4. A holomorphic map $f : \mathbb{C} \to \mathbb{C}$ has a *critical renormalization intersection* at a critical point z for an eventually periodic $\alpha \in \mathbb{R}/\mathbb{Z}$, if there exists an $N, k > 0$ such that $\mathscr{R}^N \tau_z f \in \mathcal{W}^s_{\mathscr{R}^k}(\zeta_\alpha)$.

Such intersections necessarily occur for rational maps with well-behaved Herman rings, for appropriately chosen points.

Lemma 4.3.1. *Let $f : \mathbb{C} \to \mathbb{C}$ have an invariant Herman ring $\mathcal{H}$ with an eventually periodic rotation number α. Then f has a regular renormalization intersection at z_0 for α for any $z_0 \in \mathcal{H}$.*

Proof. Let $\psi : \mathcal{H} \to \mathbb{A}_r$ be a map conjugating the dynamics of f to a pure rotation on an annulus. Under this conjugation $\overline{\mathcal{O}_f(z_0)}$ is brought to a circle in $\mathbb{A}_r$ on which $\psi \circ f \circ \psi^{-1}$ is a pure rotation. N need only be chosen large enough so that $G^N(\alpha)$ is of periodic type with some period p and so that the linear rescaling transforming prerenormalization to renormalization is sufficiently substantial that $\mathscr{R}^N \tau_{z_0} f \in \mathcal{D}(Z, W)$ with Z, W as in 4.2.5. If α is already periodic, then N can be chosen to depend only on the distance from $\overline{\mathcal{O}_f(z_0)}$ to the boundary of $\mathcal{H}$ and on the behavior of ψ' on the orbit of z_0. □

Toward the converse direction, we note a simple implication of the presence of a regular renormalization intersection.

Lemma 4.3.2. *If $f : \mathbb{C} \to \mathbb{C}$ has a regular renormalization intersection at $z \in \mathbb{C}$ for an eventually periodic $\alpha \in \mathbb{R}/\mathbb{Z}$, z is contained in a rotation domain of f.*

Proof. By using the dynamical partitions of subsequent prerenormalizations of $\tau_z f$ we may construct an analytic embedded circle $\overline{\mathcal{O}_f(z)}$ invariant under the dynamics of f, passing through z. Renormalization stability allows for the construction of an analytic conjugacy between $f|_{\overline{\mathcal{O}_f(z)}}$ and a pure rotation on the unit circle. This conjugacy can then be extended to a biholomorphism of a neighborhood of $\overline{\mathcal{O}_f(z)}$ to an annulus. □

Having established basic conditions for intersections with stable manifolds of renormalization, we would like to consider families of holomorphic maps for which such intersections are robust to perturbations.

Definition 4.3.5. A C^1 family $\{f_\lambda : \mathbb{C} \to \mathbb{C}\}_{\lambda \in U}$ for $U \subset \mathbb{C}$ of holomorphic functions has a *transverse regular (critical) renormalization intersection* at $\lambda_* \in U$, given a holomorphic map of distinguished points $z : U \to \mathbb{C}$ and an eventually periodic $\alpha \in \mathbb{R}/\mathbb{Z}$ if f_{λ_*} has a regular (critical) renormalization intersection at $z(\lambda_*) \in \mathbb{C}$ and moreover, with N, k as defined for this intersection, the family $\mathscr{R}^N \tau_{z(\lambda)} f_\lambda$ intersects this stable manifold transversely at $\lambda = \lambda_*$.

We will be interested in examining families which contain multiple transverse regular and critical intersections for different choices of distinguished point z all coinciding at the same parameter λ_*. Then, sufficiently small perturbations in C^1 are guaranteed to preserve at least finitely many such intersections. That being said, a priori such a perturbed family may have each such intersection at a different parameter value. To get around this potentially thorny issue, we will increase the dimension of the families of maps involved. As such, we'll take the natural generalization of the above definition of a transverse renormalization intersection at a point to higher dimensional C^1-families of holomorphic maps on $\mathbb{C}$, by simply taking $\lambda \in U \subset \mathbb{C}^n$ for some $n \geq 1$.

Lemma 4.3.3. *If a C^1 immersed family $\{f_\lambda : \mathbb{C}^m \to \mathbb{C}^m\}_{\lambda \in U}$ for m either equal to 1 or 2 and $U \subset \mathbb{C}^n$ for $n \geq 1$ of holomorphic functions has a transverse regular (critical) renormalization intersection at λ_* with distinguished point $z : U \to \mathbb{C}^m$ for α the golden mean, then the set of λ for which f_λ has a regular (critical) renormalization intersection at distinguished point $z(\lambda)$ for α contains an analytic codimension 1 embedded submanifold of U containing λ_*.*

Proof. Noting that both the regular and critical renormalization operators of almost-commuting pairs in either the one-dimensional setting or the dissipative two-dimensional setting are analytic with a hyperbolic fixed point having co-dimension 1 stable direction, standard hyperbolic theory (see, for example, Chapter 5 of [Shu13]) guarantees that there exists a

neighborhood E in the local stable manifold such that E is an embedded analytic co-dimension submanifold in the appropriate space of almost-commuting pairs. If some f_λ has a regular (critical) renormalization intersection at $z(\lambda)$, then by possibly enlarging N we may take $\mathscr{R}^N \tau_{z(\lambda)} f_\lambda \in E$. Moreover, for fixed N, $\mathscr{R}^N \tau_{z(\lambda)} f_\lambda$ produces a family of almost-commuting pairs which varies differentiably in λ, giving the desired result. □

Analyticity of these local stable manifolds in the preceding lemma allows for one-parameter families with multiple coinciding transverse renormalization intersections to be perturbed into families with more parameters also having coinciding transverse renormalization intersections.

Proposition 4.3.4. *Let α be the golden mean. Suppose that the family $\{f_\lambda : \mathbb{C} \to \mathbb{C}\}_{\lambda \in U}$ for $U \subset \mathbb{C}$ has transverse regular renormalization intersections at $\lambda_* \in U$ for α with distinct distinguished points $z_1(\lambda), \ldots, z_k(\lambda)$ for some $k \in \mathbb{N}$ and has transverse critical renormalization intersections at λ_* for α with distinct distinguished points $w_1(\lambda), \ldots, w_j(\lambda)$ for some $j \in \mathbb{N}$. Let $\{g_{(\lambda,\nu)} : \mathbb{C} \to \mathbb{C}\}_{\lambda \in U, \nu \in V}$ for $V \subset \mathbb{C}^{k+j}$ containing 0 be an analytic immersed family of holomorphic maps such that $g_{(\lambda,0)} = f_\lambda$ for each $\lambda \in U$, and such that $\{\tau_{(z,w)} g_{\lambda,\nu}\}_{z \in \mathbb{C}, w \in \mathbb{C}, \lambda \in U, \nu \in V}$ is itself an immersed family. Then there exists distinct critical points $w_1^*(\lambda,\nu), \ldots, w_j^*(\lambda,\nu)$ depending holomorphically (for sufficiently small $|\nu|$) on (λ,ν), and with $w_i^*(\lambda,0) = w_i(\lambda)$ for each i. Moreover, there exists some analytic one-dimensional embedded submanifold $M \subset U \times V$ with $M \cap (U \times \{0\}) = \{(\lambda_*, 0)\}$ such that for each $(\lambda_0, \nu_0) \in M$*

1. *$\{g_{(\lambda,\nu)} : \mathbb{C} \to \mathbb{C}\}_{\lambda \in U, \nu \in V}$ has transverse regular renormalization intersections for α at (λ_0, ν_0) with distinguished points $z_i(\lambda)$ for each i,*

2. *$\{g_{(\lambda_0,\nu_0)} : \mathbb{C} \to \mathbb{C}\}_{\lambda \in U, \nu \in V}$ has transverse critical renormalization intersections for α at (λ_0, ν_0) with distinguished point $w_i^*(\lambda,\nu)$ for each i.*

Proof. The family $\{g_{(\lambda,\nu)}\}_{\lambda \in U, \nu \in V}$ is a C^1-small perturbation of $\{f_\lambda\}_{\lambda \in U}$ for all $|\nu|$ sufficiently small. We may choose ϵ sufficiently small that each transverse regular renormalization

intersection survives. Each point $w_i(\lambda)$ must be a critical point of f_λ with $f''_\lambda(w_i(\lambda)) \neq 0$, so we may obtain each $w_i^*(\lambda, \nu)$ as a consequence of the implicit function theorem. By shrinking ϵ if necessary, we may also ensure each transverse critical renormalization intersection survives. Using Lemma 4.3.3, the set of (λ, ν) such that any individual such transverse renormalization intersection survives contains an analytic codimension 1 embedded submanifold of $U \times V$ containing $(\lambda_*, 0)$. Then the intersection of all $k+j$ such submanifolds is at least of dimension 1 in $U \times V$. □

The preceding proposition can also be formulated when the perturbed family consists of highly-dissipative maps of $\mathbb{C}^2$.

Proposition 4.3.5. *Under the assumptions of Proposition 4.3.4, if instead* $\{g_{(\lambda,\nu)} : \mathbb{C}^2 \to \mathbb{C}^2\}_{\lambda \in U, \nu \in V}$ *is an analytic immersed family of two-dimensional maps such that* $g_{(\lambda,0)} = \iota(f_\lambda)$*, the same conclusion holds with the modification that the family* $w_i^*(\lambda, \nu)$ *exists, but need not consist of critical points of* $g_{(\lambda,\nu)}$*.*

Proof. The only modification of the proof of Proposition 4.3.4 which needs to be made is to obtain the family $w_i^*(\lambda, \nu)$ by appealing to Theorem 3.6 of [GY20a], which asserts that in a space of almost-commuting pairs of two-dimensional maps the stable manifold has codimension at most 3, with two dimensions accounting for a loss of criticality. By considering the family $(z, w, \lambda, \nu) \mapsto \mathscr{R}^N \tau_{(z,w)} g_{(\lambda,\nu)}$ for large $N > 0$, we may thus obtain a family $w_i^*(\lambda, \nu)$ depending holomorphically on (λ, ν) with $w_i^*(\lambda, 0) = w_i(\lambda)$. □

4.3.3 Identification of a Cubic Rational Family

While one could make arguments in some generality, we will constrain ourselves to considering one of the simplest examples of a family containing Herman rings.

Definition 4.3.6. Define the family $\{f_{a,b}\}_{a \in \mathbb{C}, b \in \mathbb{C} \setminus \{0\}}$ by

$$f_{a,b}(z) = bz^2 \frac{az+1}{z+a}.$$

If $a \in \mathbb{R}$ and $|b| = 1$, such maps are Blaschke products and manifestly preserve the unit circle. This family has been well-examined by a number of authors, particularly [Buf+05; FH17]. We will find it convenient to label the nontrivial critical points of $f_{a,b}$ in the following lemma.

Lemma 4.3.6. *For $a \neq 0$ sufficiently small, $f_{a,b}$ has three critical points, two of which are nonzero. Their location depends only on a and they are given by*

$$c_{\pm}(a) = \frac{-1 - 3a^2 \pm \sqrt{1 - 10a^2 + 9a^4}}{4a}.$$

As $a \to 0$, $c_+(a) \to 0$ and $c_-(a) \to \infty$.

Classic work by Douady, Ghys, Herman, Shishikura, and Świątek (see Sections 7.2 and 7.3 of [BFB14]) implies that if $f_{a,b}$ has a Herman ring with bounded type rotation number, the boundary components of that Herman ring are quasicircles each containing a single critical point. In this case, if $f_{a,b}$ has a Herman ring, the two critical points contained in the boundary are $c_{\pm}(a)$. From [Buf+05] we have a full characterization of the location of Herman rings within this family. In that interest, a few definitions are needed.

Definition 4.3.7. For a simply connected domain $U \subset \mathbb{C}$ which contains 0, the *conformal radius* of U is the derivative at 0 of a Riemann map $f : \mathbb{D} \to U$ which fixes 0.

Definition 4.3.8. For the annulus $A = \{z \in \mathbb{C} : r_1 < |z| < r_2\}$, the *modulus* is

$$\frac{1}{2\pi} \ln\left(\frac{r_1}{r_2}\right).$$

For a set $U \subset \mathbb{C}$ conformally isomorphic to an annulus A, the *modulus* of U is equal to the modulus of A.

Definition 4.3.9. Let P_α be the quadratic polynomial

$$P_\alpha(z) = e^{2\pi i\alpha} z(1 + z).$$

For α a Brjuno number, 0 falls inside a Siegel disk of P_α with rotation number α. Let r_α denote the conformal radius of that Siegel disk.

In the following theorem, we follow its authors by formally regarding a pure rotation as having a Herman ring with modulus ∞.

Theorem 4.3.7 ([Buf+05]). *For any Brjuno number $\alpha \in \mathbb{R}/\mathbb{Z}$, there exists a holomorphic bijection $\mathscr{F}_\alpha$ from the unit disk to the set of (a, b) such that $f_{a,b}$ has a Herman ring with rotation number α. Moreover, $\mathscr{F}_\alpha(0) = (0, e^{2\pi i\alpha})$ and $\mathscr{F}'_\alpha(0) = (0, r_\alpha)$. Further, the modulus of the Herman ring of $f_{\mathscr{F}_\alpha(\delta)}$ is*

$$\frac{1}{\pi} \log\left(\frac{1}{|\delta|}\right),$$

The relationship between the dynamics of this family of rational maps and the quadratic polynomials comes from Shishikura's surgery procedure relating rational maps with Herman rings to polynomials with Siegel disks initially developed in [Shi87]. In brief, in this procedure one 'cuts' a Herman ring along an invariant circle inside the ring, discarding the interior of the region surrounded by the cut, and 'glues' along this cut a disk with the dynamics of a rigid rotation. The obtained dynamical map $\hat{f}$ on $\hat{\mathbb{C}}$ is quasiregular and preserves a measurable conformal structure with bounded dilatation, so one may apply the measurable Riemann mapping theorem to obtain a quasiconformal homeomorphism $\psi : \hat{\mathbb{C}} \to \hat{C}$ fixing 0 and ∞ which conjugates $\hat{f}$ to a conformal map on $\hat{\mathbb{C}}$. Noting that $f_{a,b}$ fixes ∞ and locally has degree 2 at ∞, our obtained map must be a quadratic polynomial with a Siegel disk. The following proposition summarizes this argument.

Proposition 4.3.8 ([Buf+05]). *If $f_{a,b}$ has a Herman ring with rotation number α, then Shishikura's quasiconformal surgery procedure produces the quadratic polynomial P_α.*

It is also worth considering another family, which we may relate to the family of quadratic polynomials even more directly.

Definition 4.3.10. Define the family $\{g_{a,b}\}_{a\in\mathbb{C},b\in\mathbb{C}\setminus\{0\}}$ by

$$g_{a,b}(z) = bz^2\frac{z+1}{z+a}.$$

We may obtain this family from the family $\{f_{a,b}\}$ in one of two ways, provided $a \neq 0$.

Lemma 4.3.9. *When $a \neq 0$, one may note that for $\Lambda_a(z) = az$,*

$$g_{a^2,b} = \Lambda_a \circ f_{a,b} \circ \Lambda_a^{-1}.$$

Moreover, if $q(z) = \frac{1}{z}$, then in addition

$$g_{a^2,b^{-1}} = q^{-1} \circ \Lambda_a^{-1} \circ f_{a,b} \circ \Lambda_a \circ q.$$

For $a > 0$ real, the circle $\partial B(0, \sqrt{a})$ is an invariant circle of $g_{a,b}$. This family has the further convenient property that when $a \to 0$ it degenerates into quadratic polynomials, with

$$g_{0,b}(z) = bz(1 + z),$$

so that if $b = e^{2\pi i\alpha}$ then $g_{0,b} = P_\alpha$. This has a key implication for renormalization at the critical point, originally formulated for cylinder renormalization in [Yam06] and then adapted to almost-commuting pairs in [GRY21].

Lemma 4.3.10. *If α is eventually golden, there exists $N > 0$ such that for b the family $b \mapsto \mathscr{R}^N \tau_{-\frac{1}{2}} g_{0,b}$ is defined on domains Z, W as in Theorem 4.2.6, transversely intersects $\mathcal{W}^s_{\mathscr{R}^2}(\zeta_\alpha)$.*

The following is then a consequence of the Koebe distortion theorem and the robustness of transverse intersections under C^1 perturbations of a family.

Corollary 4.3.11. *Let $c(a)$ denote the critical point of $g_{a,b}$ which varies holomorphically near $a = 0$ and tends to $-\frac{1}{2}$ when $a \to 0$. If α is eventually golden, there exists an $\epsilon > 0$ and an $N > 0$ such that if $|a| < \epsilon$ then the family $b \mapsto \mathscr{R}^N \tau_{c(a)} g_{a,b}$ is defined on domains Z, W as in Theorem 4.2.6, transversely intersects $\mathcal{W}^s_{\mathscr{R}^2}(\zeta_\alpha)$.*

Then, holomorphic conjugacy allows these critical intersections to be ported over to the original family $\{f_{a,b}\}$, albeit possibly by increasing N since the conjugacy becomes degenerate as $a \to 0$.

Corollary 4.3.12. *If α is eventually golden, there exists an $\epsilon > 0$ such that for any $\epsilon_0 < \epsilon$ there exists $N > 0$ depending on ϵ_0 such that if $\epsilon_0 < |a| < 2\epsilon_0$ then both of the families $b \mapsto \mathscr{R}^N \tau_{c_\pm(a)} f_{a,b}$ are defined on domains Z, W as in Theorem 4.2.6 and transversely intersect $\mathcal{W}^s_{\mathscr{R}^2}(\zeta_\alpha)$. These intersections both occur at the same value (a, b) lying in the image of $\mathscr{F}_\alpha$.*

Transverse regular renormalization intersections may also be obtained near $a = 0$.

Lemma 4.3.13. *If α is eventually periodic, there exists an $\epsilon > 0$ and an $N > 0$ such that if $|a| < \epsilon$ then the family $b \mapsto \mathscr{R}^N \tau_1 f_{a,b}$ is defined on domains Z, W as in Theorem 4.2.5 and transversely intersects $\mathcal{W}^s_{\mathscr{R}^2}(\zeta_{T_\alpha})$. This intersection occurs at a value (a, b) lying in the image of $\mathscr{F}_\alpha$.*

Proof. When $a = 0$, $f_{0,b}(z) = bz$ degenerates into a linear map; it is a pure rotation when $|b| = 1$. Such a family translated into commuting pairs around any nonzero distinguished point, say 1, produces the unstable manifold of ζ_{T_α}, which by hyperbolicity of $\mathscr{R}$ transversely intersects the stable manifold of ζ_{T_α}. The result then follows from the observation that

$$f_{a,b}(z) - f_{0,b}(z) = abz\frac{z^2 - 1}{z + a},$$

which is bounded in norm and has bounded derivative provided the orbit of 1 avoids $z = -a$ for a sufficiently long time, which is accomplished by avoiding $B(0, \delta)$ for some $\delta > 0$, and choosing $\epsilon < \delta$ sufficiently small. □

It's worth noting that the choice of distinguished point at 1 in the preceding lemma is not of any consequence, we could just as easily formulate the lemma at any fixed nonzero complex number. That being said, since $f_{a,b}$ preserves the unit circle whenever a is real, one could in principle produce an alternative formulation of this lemma at 1 requiring only that the imaginary part of a is sufficiently small.

4.3.4 Two Dimensional Perturbations with Three Renormalization Intersections

In this section, we'll make a few base choices of notation. Take $f : \mathbb{C} \to \mathbb{C}$ having an invariant golden mean Herman ring $\mathcal{H}$ with quasicircle boundary components each containing a critical point, $z_\pm \in \mathbb{C}$. Let $F : \mathbb{C}^2 \to \mathbb{C}^2$ be an invertible highly-dissipative C^1-small perturbation of $\iota(f)$ in a neighborhood of $\iota(\mathcal{H})$, and where there exist $\mathbf{z}_\pm \in \mathbb{C}^2$ very close to $\iota(z_\pm)$ such that F has a critical renormalization intersection for the golden mean at each of $\mathbf{z}_\pm$. Moreover, suppose that there exists a point $\mathbf{z}_0 \in \mathbb{C}^2$ very close to $\iota(z_0)$ for some $z_0 \in \mathcal{H}$ at which F has a regular renormalization intersection for the golden mean. The main result of this section is that if the $\epsilon > 0$ in each of these 'very close' statements are taken sufficiently small, then F contains an invariant Herman ring in $\mathbb{C}^2$.

The regular renormalization intersection gives a great deal of information about the local behavior of F.

Proposition 4.3.14 ([Yam21])**.** *Suppose $F : \mathbb{C}^2 \to \mathbb{C}^2$ has a regular renormalization intersection for the golden mean at $Z \in \mathbb{C}^2$. Then $\gamma_0 = \overline{\mathcal{O}_F(Z)}$ is an analytic invariant circle on which F is analytically conjugate to a rigid rotation through some angle $\theta \in (0, 1)$ satisfying $G^m(\theta)$ is the golden mean for some $m \in \mathbb{N}$. Moreover, γ_0 is contained in a rotation domain of F.*

Critical renormalization intersections allow us to guarantee the existence of some invariant curves, which we'll call $\gamma_\pm$.

Proposition 4.3.15 ([GRY21])**.** *Suppose $F : \mathbb{C}^2 \to \mathbb{C}^2$ has a critical renormalization intersection for the golden mean at $Z \in \mathbb{C}^2$ and let Ω_n, Γ_n be the domains of definition of the n-th prerenormalization of F at Z. Then there exists an invariant curve $\gamma \subset \mathbb{C}^2$ satisfying all of the following.*

1. *γ is a homeomorphic image of S^1.*

2. *$\gamma \cap \Omega_n \neq \emptyset$ and likewise for Γ_n, for all $n \in \mathbb{N}$.*

3. *For some $\theta \in (0,1)$ satisfying for some $m \in \mathbb{N}$ that $G^m(\theta)$ is the golden mean, there exists a continuous map ψ conjugating the dynamics of $F|_\gamma$ to the dynamics of the rigid rotation through θ on S^1. ψ is not C^1.*

In order to gain some kind of local control over the dynamics of F near such an invariant curve, it is very helpful to have an invariant cone field.

Proposition 4.3.16 ([GRY21])**.** *Suppose U is a domain and we have an analytic map $F : U \times U \to \mathbb{C}^2$ which satisfies*

$$F(z,w) = (g_0(z) + r_0(z,w), g_1(z) + r_1(z,w)),$$

and that $k < |g_0'(z)|, |g_1'(z)| < K$ for all $z \in U$, and F^{-1} is defined on $\Delta = F(U \times U)$. Then there exists $\epsilon, \rho > 0$ such that if we suppose further that r_0, r_1 are uniformly bounded by ϵ on $U \times U$ and if we define a cone field by

$$C^{vert,\rho}_{(z,w)} = \{(u,v) \in T_{(z,w)}(U \times U) : |u| < \rho|v|\},$$

then for any compact subset $\Delta' \Subset \Delta$ and (z,w) such that $F(z,w) \in \Delta'$ one has

$$DF^{-1}|_{F(z,w)}(C^{vert,\rho}_{F(z,w)}) \subset C^{vert,\rho}_{(z,w)}.$$

Moreover, $\|DF^{-1}\| > \mathcal{O}\left(\frac{\kappa}{K\epsilon}\right)$ in $C^{vert,\rho}$.

From here, we'll present a schematic argument, which can be made fully rigorous, for how to construct a Herman ring for F.

To proceed, we first construct a partition of a small neighborhood of $\mathcal{H}$, the Herman ring for f. To do so, take an internal ray within the Herman ring and join it to an external ray outside of $\mathcal{H}$ as is done in the Siegel case in [DLS20]. Pushing the obtained curve forward by f in a manner akin to the construction of dynamical partitions divides a neighborhood of $\mathcal{H}$, giving rise to a partition by 'trapezoidal' sets. This partition can be 'thickened' into a

cover of a neighborhood of $\iota(\mathcal{H})$ by taking a small neighborhood around the image of each trapezoidal set under ι in a manner akin to the thickening procedure used in two-dimensional perturbations in the unimodal setting in [CLM05].

Using Proposition 4.3.14, F has a rotation domain containing an analytic invariant curve γ_0 containing $\mathbf{z}_0$. In some neighborhood of γ_0, this rotation domain is an normally hyperbolic invariant manifold of F. Such manifolds are stable under small perturbations (see Chapter 4 of [HPS70]), so this invariant manifold may be extended extremely close to the boundary of $\iota(\mathcal{H})$. In particular, it may be taken to intersect for some large N the N-th pre-renormalization domain of F around both of $\mathbf{z}_\pm$, U_N.

One may then imitate the proof present in [GRY21] using Proposition 4.3.16 to obtain an invariant cone field on sets compactly containing $U_N \setminus U_{N+k}$ for some $k > 0$. This implies that under iteration, the thickened trapezoidal partition converges in this region in a Hausdorff sense to an analytic invariant manifold - which must agree with the normally hyperbolic invariant manifold. As such, the normally hyperbolic invariant manifold may be extended to reach into the $N + k$-th pre-renormalization domain of F around both of $\mathbf{z}_\pm$.

Iterating this procedure and shrinking the external rays of the trapezoidal cover leads the rotation domain to extend precisely to the invariant curves $\gamma_\pm$ containing $\mathbf{z}_\pm$ whose existence is guaranteed by Proposition 4.3.15. As these invariant circles are not C^1, we may extend no further.

Restricted to the interior of the normally hyperbolic invariant manifold H, F acts as a map of an annulus which extends continuously to the boundary and which is analytically conjugated on γ_0 to a pure rotation. As a result, it is analytically conjugated on the entirety of H, with the conjugacy extending continuously but not differentiably to $\gamma_\pm$. Thus, we have a Herman ring with a rotation number given by a pre-image of the golden mean under the Gauss map.

Bibliography

[AAR99] George E Andrews, Richard Askey, and Ranjan Roy. *Special functions*. Vol. 71. Cambridge university press Cambridge, 1999.

[AD02] David Aubin and Amy Dahan Dalmedico. "Writing the history of dynamical systems and chaos: longue durée and revolution, disciplines and cultures". In: *Historia mathematica* 29.3 (2002), pp. 273–339.

[Aiz+88] M Aizenman et al. "Discontinuity of the Magnetization in One-Dimensional 1 Ix $_y[2IsingandPottsModels$". In: *Journal of Statistical Physics* 50 (1 1988).

[Aiz82] Michael Aizenman. "Geometric analysis of $\varphi 4$ fields and Ising models. Parts I and II". In: *Communications in mathematical Physics* 86.1 (1982), pp. 1–48.

[AL11] Artur Avila and Mikhail Lyubich. "The full renormalization horseshoe for unimodal maps of higher degree: exponential contraction along hybrid classes". In: *Publications mathématiques de l'IHÉS* 114 (2011), pp. 171–223.

[Ami20] Ariel Amir. "An elementary renormalization-group approach to the generalized central limit theorem and extreme value distributions". In: *Journal of Statistical Mechanics: Theory and Experiment* 2020 (1 2020). ISSN: 17425468. DOI: 10.1088/1742-5468/ab5b8c.

[Arn61] Vladimir Igorevich Arnol'd. "Small denominators. I. Mapping the circle onto itself". In: *Izvestiya Rossiiskoi Akademii Nauk. Seriya Matematicheskaya* 25.1 (1961), pp. 21–86.

[AS64] Milton Abramowitz and Irene A Stegun. *Handbook of mathematical functions with formulas, graphs, and mathematical tables*. Vol. 55. US Government printing office, 1964.

[AT17] Elena Agliari and Flavia Tavani. "The exact Laplacian spectrum for the Dyson hierarchical network". In: *Scientific Reports 2017 7:1* 7 (1 Jan. 2017), pp. 1–21. ISSN: 2045-2322. DOI: 10.1038/srep39962. URL: https://www.nature.com/articles/srep39962.

[BE17] Laurent Bartholdi and Anna Erschler. "Poisson–Furstenberg boundary and growth of groups". In: *Probability Theory and Related Fields* 168 (1 2017), pp. 347–372.

[BEI07] David Brydges, Steven N. Evans, and John Z. Imbrie. "Self-Avoiding Walk on a Hierarchical Lattice in Four Dimensions". In: *The Annals of Probability* 20 (1 2007). DOI: 10.1214/aop/1176989919.

[Ber41] Andrew C. Berry. "The Accuracy of the Gaussian Approximation to the Sum of Independent Variates". In: *Transactions of the American Mathematical Society* 49 (1 1941). ISSN: 00029947. DOI: 10.2307/1990053.

[BFB14] Bodil Branner, Núria Fagella, and Xavier Buff. *Quasiconformal surgery in holomorphic dynamics*. 141. Cambridge University Press, 2014.

[BFS82] David Brydges, Jürg Fröhlich, and Thomas Spencer. "The random walk representation of classical spin systems and correlation inequalities". In: *Communications in Mathematical Physics* 83 (1 1982). ISSN: 14320916. DOI: 10.1007/BF01947075.

[Ble10] Pavel Bleher. "CRITICAL PHENOMENA IN THE DYSON HIERARCHICAL MODEL AND RENORMALIZATION GROUP". In: (2010).

[Ble12] Pavel Bleher. "From the seminar on Mathematical Statistical Physics in Moscow State Hierarchical models and renormalisation group critical phenomena in the Dyson hierarchical model and renormalisation group". In: *Eur. Phys. J. H* 37 (2012), pp. 605–618. DOI: 10.1140/epjh/e2012-10053-x.

[Ble77] P. M. Bleher. "Phase transitions of the second kind in asymptotic hierarchical Dyson models". In: *Uspekhi Mat. Nauk* 32 (6(198) 1977), pp. 243–244. URL: http://www.mathnet.ru/php/archive.phtml?wshow=paper%5C&jrnid=rm%5C&paperid=3328%5C&option_lang=eng.

[BM07] P. M. Bleher and P. Major. "Critical Phenomena and Universal Exponents in Statistical Physics. On Dyson's Hierarchical Model". In: *The Annals of Probability* 15 (2 2007). DOI: 10.1214/aop/1176992155.

[BM21] Pavel Bleher and Peter Major. "A Thouless-Like Effect in the Dyson Hierarchical Model with Continuous Symmetry". In: *Pure and Applied Functional Analysis* 6 (1 2021), pp. 101–175. URL: http://yokohamapublishers.jp/online2/oppafa/vol6/p101.html.

[BM89] P. M. Bleher and P. Major. "The large-scale limit of Dyson's hierarchical vector-valued model at low temperatures. the marginal case {Mathematical expression}". In: *Communications in Mathematical Physics* 125 (1 1989). ISSN: 00103616. DOI: 10.1007/BF01217768.

[Bri+10] Martin R. Bridson et al. "Cofinitely Hopfian groups, open mappings and knot complements". In: *Groups, Geometry, and Dynamics* 4 (4 2010). ISSN: 16617207. DOI: 10.4171/GGD/101.

[Bro12] Laurie M Brown. *Renormalization: from Lorentz to Landau (and beyond)*. Springer Science & Business Media, 2012.

[BS73] P. M. Bleher and Ja G. Sinai. "Investigation of the critical point in models of the type of Dyson's hierarchical models". In: *Communications in Mathematical Physics* 33 (1 Mar. 1973), pp. 23–42. ISSN: 00103616. DOI: 10.1007/BF01645604.

[BS75] P. M. Bleher and Ya G. Sinai. "Critical indices for Dyson's asymptotically-hierarchical models". In: *Communications in Mathematical Physics 1975 45:3* 45 (3 Oct. 1975), pp. 247–278. ISSN: 1432-0916. DOI: 10.1007/BF01608331. URL: https://link.springer.com/article/10.1007/BF01608331.

[Buf+05] Xavier Buff et al. "HERMAN RINGS AND ARNOLD DISKS". In: *J. London Math. Soc* (2 2005), pp. 689–716. DOI: 10.1112/S0024610705007015.

[Bur97] Andrew D Burbanks. "Renormalization for Siegel discs". PhD thesis. Loughborough University, 1997.

[BV05] Laurent Bartholdi and Bálint Virág. "Amenability via random walks". In: *Duke Mathematical Journal* 130 (1 2005). ISSN: 00127094. DOI: 10.1215/S0012-7094-05-13012-5.

[Cal+10] Iván Calvo et al. "Generalized Central Limit Theorem and Renormalization Group". In: *Journal of Statistical Physics* 141 (3 2010). ISSN: 00224715. DOI: 10.1007/s10955-010-0065-y.

[CGS11] Louis H Y Chen, Larry Goldstein, and Qi-Man Shao. *Normal approximation by Stein's method*. Vol. 2. Springer, 2011.

[CLM05] A De Carvalho, Mikhail Lyubich, and Marco Martens. "Renormalization in the Hénon family, I: Universality but non-rigidity". In: *Journal of Statistical Physics* 121.5 (2005), pp. 611–669.

[CS18] Johannes Cuno and Ecaterina Sava-Huss. "RANDOM WALKS ON BAUMSLAG-SOLITAR GROUPS". In: *ISRAEL JOURNAL OF MATHEMATICS* 228 (2018), pp. 627–663. DOI: 10.1007/s11856-018-1775-0.

[CT96] Sergio A Cannas and Francisco A Tamarit. "Long-range interactions and nonextensivity in ferromagnetic spin models". In: *Physical review B* 54.18 (1996), R12661.

[CY00] Fan Chung and S. T. Yau. "Discrete Green's Functions". In: *Journal of Combinatorial Theory, Series A* 91 (1-2 July 2000), pp. 191–214. ISSN: 0097-3165. DOI: 10.1006/JCTA.2000.3094.

[DD00] Edson De Faria and Welington De Melo. "Rigidity of critical circle mappings II". In: *Journal of the American Mathematical Society* 13.2 (2000), pp. 343–370.

[De 99] Edson De Faria. "Asymptotic rigidity of scaling ratios for critical circle mappings". In: *Ergodic Theory and Dynamical Systems* 19.4 (1999), pp. 995–1035.

[Den32] Arnaud Denjoy. "Sur les courbes définies par les équations différentielles à la surface du tore". In: *Journal de mathématiques pures et appliquées* 11 (1932), pp. 333–376.

[DGW05] D A Dawson, L G Gorostiza, and A Wakolbinger. "Degrees of Transience and Recurrence and Hierarchical Random Walks". In: *Potential Analysis* 22 (2005), pp. 305–350.

[DH85] Adrien Douady and John Hamal Hubbard. "On the dynamics of polynomial-like mappings". In: *Annales scientifiques de l'École normale supérieure*. Vol. 18. 2. 1985, pp. 287–343.

[DLS20] Dzmitry Dudko, Mikhail Lyubich, and Nikita Selinger. "Pacman renormalization and self-similarity of the Mandelbrot set near Siegel parameters". In: *Journal of the American Mathematical Society* 33.3 (2020), pp. 653–733.

[Dob68] PL Dobruschin. "The description of a random field by means of conditional probabilities and conditions of its regularity". In: *Theory of Probability & Its Applications* 13.2 (1968), pp. 197–224.

[DV12] Welington De Melo and Sebastian Van Strien. *One-dimensional dynamics*. Vol. 25. Springer Science & Business Media, 2012.

[Dys69] Freeman J. Dyson. "Non-existence of spontaneous magnetization in a one-dimensional Ising ferromagnet". In: *Communications in Mathematical Physics 1969 12:3* 12 (3 Sept. 1969), pp. 212–215. ISSN: 1432-0916. DOI: 10.1007/BF01661575. URL: https://link.springer.com/article/10.1007/BF01661575.

[Dys71] Freeman J Dyson. "An Ising ferromagnet with discontinuous long-range order". In: *Communications in Mathematical Physics* 21.4 (1971), pp. 269–283.

[Edw65] Sam F Edwards. "The statistical mechanics of polymers with excluded volume". In: *Proceedings of the Physical Society (1958-1967)* 85.4 (1965), p. 613.

[Ess42] Carl-Gustav Esseen. "On the Liapunov limit error in the theory of probability". In: *Ark. Mat. Astr. Fys.* 28 (1942), pp. 1–19.

[Far92] Edson de Faria. *Proof of universality for critical circle mappings*. City University of New York, 1992.

[Fei78] Mitchell J Feigenbaum. "Quantitative universality for a class of nonlinear transformations". In: *Journal of statistical physics* 19.1 (1978), pp. 25–52.

[FFS92] Roberto Fernández, Jürg Fröhlich, and Alan D. Sokal. *Random Walks, Critical Phenomena, and Triviality in Quantum Field Theory*. Springer Berlin Heidelberg, 1992. DOI: 10.1007/978-3-662-02866-7.

[FH17] Núria Fagella and Christian Henriksen. "The Fine Structure of Herman Rings". In: *Journal of Geometric Analysis* 27 (3 July 2017), pp. 2381–2399. ISSN: 10506926. DOI: 10.1007/S12220-017-9764-9/FIGURES/3. URL: https://link.springer.com/article/10.1007/s12220-017-9764-9.

[GL54] Murray Gell-Mann and Francis E Low. "Quantum electrodynamics at small distances". In: *Physical Review* 95.5 (1954), p. 1300.

[GM67] G. Gallavotti and S. Miracle-Sole. "Statistical mechanics of lattice systems". In: *Communications in Mathematical Physics 1967 5:5* 5 (5 Oct. 1967), pp. 317–323. ISSN: 1432-0916. DOI: 10.1007/BF01646445. URL: https://link.springer.com/article/10.1007/BF01646445.

[GN07] Rostislav Grigorchuk and Volodymyr Nekrashevych. "Self-similar groups, operator algebras and Schur complement". In: *Journal of Modern Dynamics. 2007, Volume 1, Pages 323-370* 1 (3 Apr. 2007), p. 323. ISSN: 1930-532X. DOI: 10.3934/JMD.2007.1.323. URL: https://www.aimsciences.org/article/doi/10.3934/jmd.2007.1.323.

[Gol18] Nigel Goldenfeld. *Lectures on phase transitions and the renormalization group*. CRC Press, Jan. 2018, pp. 1–394. ISBN: 9780429962042. DOI: 10.1201/9780429493492/LECTURES-PHASE-TRANSITIONS-RENORMALIZATION-GROUP-NIGEL-GOLDENFELD. URL: https://www.taylorfrancis.com/books/mono/10.1201/9780429493492/lectures-phase-transitions-renormalization-group-nigel-goldenfeld.

[Gri80] Rostislav Ivanovich Grigorchuk. "Burnside problem on periodic groups". In: *Funktsional'nyi Analiz i ego Prilozheniya* 14.1 (1980), pp. 53–54.

[GRY21] Denis Gaidashev, Remus Radu, and Michael Yampolsky. "Renormalization and siegel disks for complex hénon maps". In: *Journal of the European Mathematical Society* 23 (4 2021), pp. 1053–1073. ISSN: 14359855. DOI: 10.4171/JEMS/1028.

[GY20a] Denis Gaidashev and Michael Yampolsky. "Renormalization of almost commuting pairs". In: *Inventiones Mathematicae* 221 (1 July 2020), pp. 203–236. ISSN: 14321297. DOI: 10.1007/S00222-020-00947-W/FIGURES/3. URL: https://link.springer.com/article/10.1007/s00222-020-00947-w.

[GY20b] Nataliya Goncharuk and Michael Yampolsky. "Analytic linearization of conformal maps of the annulus". In: (Apr. 2020). DOI: 10.48550/arxiv.2004.05126. URL: https://arxiv.org/abs/2004.05126v5.

[Her79] Michael R Herman. "Sur la conjugaison différentiable des difféomorphismes du cercle à des rotations". In: *Publications Mathématiques de l'IHÉS* 49 (1979), pp. 5–233.

[HPS70] Morris W Hirsch, Charles C Pugh, and Michael Shub. "Invariant manifolds". In: *Bulletin of the American Mathematical Society* 76.5 (1970), pp. 1015–1019.

[Hua13] Kerson Huang. "A critical history of renormalization". In: *International Journal of Modern Physics A* 28.29 (2013), p. 1330050.

[Jir20] Moritz Jirak. "A Berry-Esseen bound with (almost) sharp dependence conditions". In: *arXiv preprint arXiv:2008.07332* (2020).

[Jon75] G. Jona-Lasinio. "The renormalization group: A probabilistic view". In: *Il Nuovo Cimento B Series 11* 26 (1 1975). ISSN: 03693554. DOI: 10.1007/BF02755540.

[Kad66] Leo P Kadanoff. "Scaling laws for Ising models near T c". In: *Physics Physique Fizika* 2.6 (1966), p. 263.

[Kai05] Vadim A. Kaimanovich. ""Münchhausen trick" and amenability of self-similar groups". In: vol. 15. 2005. DOI: 10.1142/S0218196705002694.

[KK08] Hans Koch and SAŠA KOCIĆ. "Renormalization of vector fields and Diophantine invariant tori". In: *Ergodic Theory and Dynamical Systems* 28.5 (2008), pp. 1559–1585.

[Kra17] Thomas Krajewski. "A renormalisation group approach to the universality of wigner's semicircle law for random matrices with dependent entries". In: *Advances in High Energy Physics* 2017 (2017).

[Lan17] Oscar E Lanford. "A computer-assisted proof of the Feigenbaum conjectures". In: *Universality in Chaos*. Routledge, 2017, pp. 245–252.

[Les98] Annick Lesne. "Renormalization methods". In: *Critical phenomena, chaos, fractal structures* (1998).

[Lim21] Wouter van Limbeek. "Structure of normally and finitely non-co-Hopfian groups". In: *Groups, Geometry, and Dynamics* 15 (2 2021). ISSN: 16617215. DOI: 10.4171/GGD/603.

[LP14] Mikhail Lyubich and Han Peters. "Classification of invariant Fatou components for dissipative Hénon maps". In: *Geometric and Functional Analysis 2014 24:3* 24 (3 May 2014), pp. 887–915. ISSN: 1420-8970. DOI: 10.1007/S00039-014-0280-9. URL: https://link.springer.com/article/10.1007/s00039-014-0280-9.

[LS09] D. Li and Ya G. Sinai. "Asymptotic behavior of generalized convolutions". In: *Regular and Chaotic Dynamics 2009 14:2* 14 (2 Apr. 2009), pp. 248–262. ISSN: 1468-4845. DOI: 10.1134/S1560354709020051. URL: https://link.springer.com/article/10.1134/S1560354709020051.

[LS14] Dong Li and Yakov G. Sinai. "An Application of the Renormalization Group Method to Stable Limit Laws". In: *Journal of Statistical Physics 2014 157:4* 157 (4 Aug. 2014), pp. 915–930. ISSN: 1572-9613. DOI: 10.1007/S10955-014-1098-4. URL: https://link.springer.com/article/10.1007/s10955-014-1098-4.

[Lyu99] Mikhail Lyubich. "Feigenbaum-Coullet-Tresser universality and Milnor's hairiness conjecture". In: *Annals of Mathematics* (1999), pp. 319–420.

[Mar98] Marco Martens. "The periodic points of renormalization". In: *Annals of mathematics* (1998), pp. 543–584.

[McM96] Curtis T McMullen. *Renormalization and 3-manifolds which fiber over the circle*. 142. Princeton University Press, 1996.

[Mes85] Benjamin David Mestel. "A computer assisted proof of universality for cubic critical maps of the circle with golden mean rotation number". PhD thesis. University of Warwick, 1985.

[Mil11] John Milnor. *Dynamics in One Complex Variable.(AM-160):(AM-160)-*. Vol. 160. Princeton University Press, 2011.

[Muk09] David Mukamel. "Notes on the statistical mechanics of systems with long-range interactions". In: *arXiv preprint arXiv:0905.1457* (2009).

[MW14] Marco Martens and Björn Winckler. "On the hyperbolicity of Lorenz renormalization". In: *Communications in Mathematical Physics* 325.1 (2014), pp. 185–257.

[Nek05] Volodymyr Nekrashevych. *Self-similar groups*. American Mathematical Soc., 2005.

[NP12] Ivan Nourdin and Giovanni Peccati. *Normal approximations with Malliavin calculus: from Stein's method to universality*. Vol. 192. Cambridge University Press, 2012.

[Poi85] Henri Poincaré. "Sur les courbes définies par les équations différentielles". In: *J. Math. Pures Appl.* 4 (1885), pp. 167–244.

[Pra06] Giuseppe Da Prato. *An introduction to infinite-dimensional analysis*. Springer Science & Business Media, 2006.

[RF99] Emmanuel Risler and Société mathématique de France. *Linéarisation des perturbations holomorphes des rotations et applications*. Société mathématique de France, 1999.

[Rom05] Steven Roman. *The umbral calculus*. Springer, 2005.

[Rue99] David Ruelle. *Statistical mechanics: Rigorous results*. World Scientific, 1999.

[Shi87] Mitsuhiro Shishikura. "ON THE QUASICONFORMAL SURGERY OF RATIONAL FUNCTIONS". In: *Ann. sclent. EC. Norm. Sup* (1987), pp. 1–29.

[Shu13] Michael Shub. *Global stability of dynamical systems*. Springer Science & Business Media, 2013.

[Sin14] Ya G Sinai. *Theory of phase transitions: rigorous results*. Vol. 108. Elsevier, 2014.

[Sin76] Yu. G. Sinai. "Self-Similar Probability Distributions". In: *Theory of Probability & Its Applications* 21 (1 1976). ISSN: 0040-585X. DOI: 10.1137/1121005.

[SP53] Ernst Carl Gerlach Stueckelberg von Breidenbach and André Petermann. "La normalisation des constantes dans la theorie des quanta". In: *Helvetica Physica Acta* 26.5 (1953), pp. 499–520.

[Sti94] A. Stirnemann. "Existence of the Siegel disc renormalization fixed point". In: *Nonlinearity* 7 (3 1994), pp. 959–974. ISSN: 09517715. DOI: 10.1088/0951-7715/7/3/011.

[Sul88] Dennis Sullivan. "Bounds, quadratic differentials, and renormalization conjectures". In: (1988).

[Tho15] Colin J. Thompson. *Mathematical Statistical Mechanics*. Princeton University Press, Dec. 2015. DOI: 10.1515/9781400868681.

[VSK17] EB Vul, Ya G Sinai, and Konstantin Mikhailovich Khanin. "Feigenbaum universality and the thermodynamic formalism". In: *Universality in Chaos* (2017), pp. 491–530.

[Wil71a] Kenneth G Wilson. "Renormalization group and critical phenomena. I. Renormalization group and the Kadanoff scaling picture". In: *Physical review B* 4.9 (1971), p. 3174.

[Wil71b] Kenneth G Wilson. "Renormalization group and critical phenomena. II. Phase-space cell analysis of critical behavior". In: *Physical Review B* 4.9 (1971), p. 3184.

[Wil83] Kenneth G Wilson. "The renormalization group and critical phenomena". In: *Reviews of Modern Physics* 55.3 (1983), p. 583.

[Yam03] Michael Yampolsky. "Hyperbolicity of renormalization of critical circle maps". In: *Publications Mathématiques de l'Institut des Hautes Études Scientifiques 2003 96:1* 96 (1 May 2003), pp. 1–41. ISSN: 1618-1913. DOI: 10.1007/S10240-003-0007-1. URL: https://link.springer.com/article/10.1007/s10240-003-0007-1.

[Yam06] Michael Yampolsky. "SIEGEL DISKS AND RENORMALIZATION FIXED POINTS". In: (2006).

[Yam21] Michael Yampolsky. "KAM-RENORMALIZATION AND HERMAN RINGS FOR 2D MAPS". In: (2021).

[Yan20] Jonguk Yang. "Renormalization in the golden-mean semi-Siegel Hénon family: universality and non-rigidity". In: *Ergodic Theory and Dynamical Systems* 40 (4 Apr. 2020), pp. 1108–1152. ISSN: 0143-3857. DOI: 10.1017/ETDS.2018.83. URL: https://www.cambridge.org/core/journals/ergodic-theory-and-dynamical-systems/article/abs/renormalization-in-the-goldenmean-semisiegel-henon-family-universality-and-nonrigidity/E8BD68ECAFE2C5943C79F73645504CE6.

[Yoc02] Jean-Christophe Yoccoz. "Analytic linearization of circle diffeomorphisms". In: (Oct. 2002), pp. 125–173. DOI: 10.1007/978-3-540-47928-4_3. URL: https://link.springer.com/chapter/10.1007/978-3-540-47928-4_3.

[Zin00] Michel. Zinsmeister. *Thermodynamic formalism and holomorphic dynamical systems*. American Mathematical Society, 2000, p. 82. ISBN: 978-0-8218-1948-7.

[Zol86] Vladimir M Zolotarev. *One-dimensional stable distributions*. Vol. 65. American Mathematical Soc., 1986.